Master Thesis Econometrics and Management Science

Quantitative Finance

Portfolio Performance With High-Dimensional Inverse Covariance Matrix

Author:
Thomas Mandigers

Supervisor:
dr. S.O.R. Lönn

First reader:
prof.dr. M.van der Wel

Abstract

High-dimensional inverse covariance estimation is a big problem for portfolio managers that want to create portfolios with a high number of assets. Recent research in this area delivered multiple shrinkage techniques that enable for the estimation of these high-dimensional inverse covariance matrices. We analyze the out-of-sample performance of portfolios constructed by high-dimensional inverse covariance matrices, created with a shrinkage technique, and we also evaluate the robustness of the shrinkage technique. Out of the 5 shrinkage technique analyzed the POET, Ledoit and Wolf and Graphical Lasso shrinkage technique prove to be good performers against the equally weighted portfolio. However, only the Ledoit and Wolf, and the Graphical Lasso are a reasonable choice for real world use. They are the only two, out of the investigated techniques, that give comparable inverse covariance estimates for similarly distributed data sets. The performance of all techniques are investigated in combination with different mean-variance portfolios. Generally, the Graphical Lasso is the highest risk adjusted performer and the most robust in creating the inverse covariance matrix.

Contents

1 Introduction

Estimating the inverse covariance in high-dimensional data set often results in a ill fit due to big estimation errors, or is not even possible to construct at all. The classic maximum likelihood estimator of the covariance matrix is ill conditioned or singular in the high-dimensional case. In order to estimate the inverse covariance, in the classical way, the number of observations needs to be equal to or bigger than the amount of variables. This is for most use cases not a big problem. However, this greatly restricts the possible size of the, industry standard, Markowitz 1952 (Markowitz) mean-variance portfolio. When a portfolio manager only trades with stocks in the SP500 stock market index for example, he can trade with 500 different stocks. That means in order to create mean-variance portfolio, at least 500 data points of each stock are needed. That is approximately 2 years of daily data points. However, the parameter structure that correctly describes the stock returns of today, is not necessarily the correct structure for the returns tomorrow. This results in the fact that the stock returns of 2 years ago has negligible informative power on the returns tomorrow, if it has any informative power at all. Therefore, the most recent data is the most influential data on tomorrows stock returns. Consequently, portfolio managers only want to use the most recent data, but less data points inherently means less possible assets for the portfolio.

So the problem at hand is creating an inverse covariance with less daily historical prices than the number of invested assets. This generalizes to estimating the inverse covariance with less number of observations than variables. The literature provides some shrinkage technique which allows for the construction of such inverse covariance. The usefulness of shrinkage in finance applications is addressed as early as 1961 by James and Stein (1961). The idea led to the prominent shrinkage techniques, for the covariance and or inverse covariance, as displayed in: Fan et al. (2013), Ledoit and Wolf (2003), Ledoit and Wolf (2004), and Ledoit et al. (2004). They start with an estimated covariance matrix and shrink this towards a target matrix, this matrix is then suited to be inverted in the classical way. While Banerjee et al. (2008), Friedman et al. (2008), and Meinshausen et al. (2006) iteratively try to optimize the maximum likelihood, of the starting matrix, through a Lasso regression. This starting matrix can be an estimate of the covariance matrix, or, if possible, an estimate of the inverse covariance. The Lasso regression itself was first introduced by Tibshirani (1996). Finally, Kourtis et al. (2012) directly shrink the estimated inverse covariance, with the goal to create mean-variance portfolios. This method therefore does not give an estimate of the inverse covariance as output but directly gives the mean-variance portfolio weights.

Subsequently, most of the shrinkage techniques use their own starting point, regularization method and regularization target. Resulting in unique inverse covariance matrices with more variables than observations. These can in turn be used to create big portfolios, that are only based on the latest returns. This solves the high-dimensional problem for portfolio managers. So the next logical step is to test which shrinkage technique is most suitable to use in portfolio creation. More crudely what technique gives the better performing portfolio.

We analyze the out-of-sample performance of portfolios constructed by high-dimensional inverse covariance matrices, created with a shrinkage technique, and we also evaluate the robustness of the shrinkage technique.

However, there are many different interesting mean-variance portfolios alongside the mean-variance frontier, and each choice naturally gives a different performance outcome for the same inverse covariance. Therefore, in order to give a sound conclusion on the performance of the shrinkage technique multiple portfolios need to be considered. This enables us to compare the performance of the shrinkage technique to each other and across different portfolios to see if there is one particular technique that outperforms across the different portfolio creations. Three different portfolio weighting schemes are used to construct the portfolios. The Global Minimum Variance (GMV) portfolio, the Global Minimum Variance portfolio with no short sale constraint (GMV-C), and a weight restricted Global Minimum Variance portfolio (GMV-W) such that all weights are bigger or equal to $\frac{0.5}{N}$ (N is the number of assets used).

The weighting schemes are chosen following the papers of Jagannathan and Ma (2003), and DeMiguel et al. (2009). The paper of Jagannathan and Ma (2003) proves that setting restrictions on the weights, even if the restrictions are wrong, positively improve the performance of the portfolio. Such a restriction could be no short sale for example. The findings are confirmed in the DeMiguel et al. (2009) paper. In this paper the out of sample performance of multiple different portfolios is investigated with multiple different data sets. They found that the portfolios that incorporated some weight restrictions performed well for every data set. Whereas the portfolios without weight restrictions did well sometimes, but performed quite bad respectively to the other portfolios for other data sets.

Using weight restrictions leads to a portfolio that generally performs well across a multitude of data sets. The benefit of weight restrictions can be explained by the big pitfall of a GMV portfolio. Jobson and Korkie (1980), and Michaud (1989) show that the GMV portfolio has the tendency to maximize the effects of estimation errors in the input data. The weight restrictions can reduce this effect, which generally result in a better performing portfolio. Finally in order to see if the use of shrinkage techniques in portfolios management is a beneficial addition. The portfolio performances are compared to the equally weighted portfolio. As this is a great benchmark to compare to.

We analyze the out-of-sample performance of portfolios constructed by high-dimensional inverse covariance matrices, created with a shrinkage technique, and we also evaluate the robustness of the shrinkage technique. To the best of our knowledge, a direct comparison between the performance of high-dimensional portfolios created with different shrinkage techniques is not done before. In addition, we do not only compare the performance but also the ability of the shrinkage technique to construct a robust estimation of the inverse covariance. With this analysis we can deduce the shrinkage technique that results in the best performing portfolio, with a robust estimation of the inverse covariance matrix. The technique or techniques, that have a robust estimation of the inverse covariance, is a good starting point for portfolio managers that want to create high-dimensional portfolios. With the technique(s) they can be confident that the performance is not substantially influenced by estimation errors, and can from there adjust the portfolio to satisfy the personal preferences. The literature can take the shrinkage technique(s) with a robust estimation of the inverse covariance as baseline for new shrinkage techniques to compare to. In order to see if the new technique is a valid addition for portfolio management.

The portfolio results are studied with different performance measures and the robustness is analyzed with the standard deviation of these performance measures. The data used for this test is the CRSP daily return data of all stocks with a market cap higher than $10,000,000, at the beginning of April 2020. The portfolios are updated at the beginning of each quarter, where the portfolio is constructed with the inverse covariance based on returns of previous quarter. Consequently, the inverse covariance matrices are created out of sample with approximately 63 data points each quarter.

However, directly comparing the performance of the portfolio can give the wrong conclusion. When the shrinkage technique constructs vastly different inverse covariance matrices with only slight deviations in the data set, it is not a robust estimator of the inverse covariance. Meaning that the resulting point estimates of the performance measures are not predominately based on the data. Before the performance measures can be compared, the robustness of the point estimates based on the shrinkage technique need to investigated.

The robustness can be analyzed by estimating multiple point estimates, with similarly distributed data sets and comparing them to see if there are big differences. Many financial papers like Hansen and Lunde (2005), Ledoit and Wolf (2008), Kosowski et al. (2007), Kosowski et al. (2006), Ruiz and Pascual (2002) resort to the idea of Efron (Efron, 1979), to use a bootstrap method to simulate the comparable data sets. They show that with a simple bootstrap procedure a data set can be generated that approaches the true distribution of sample set. This idea is extended by Politis and Romano (1992) to included the time and cross asset influences. They call this the circular bootstrap. So with this bootstrap 100 new unique asset return data sets can be created (from the ± 63 days in a quarter), and these can in turn be used with the shrinkage techniques to create 100 different inverse covariance and sequentially 100 portfolios. When the shrinkage methods are robust they should create really similar portfolios for all 100 data series, as they all follow the same distribution. Looking at the standard deviation between the performance measure, within one quarter, gives an indication on how robust the performance is.

The POET, Ledoit and Wolf (LW) and Graphical Lasso (GL) shrinkage techniques of respectively Fan et al. (2013), Ledoit and Wolf (2004) and Friedman et al. (2008) result in good performing portfolios in comparison to the Equally Weighted (EQ) portfolio. Further, the LW and GL also display a robust estimation of the inverse covariance. Whereas the POET performance noticeably less in that regard. The POET show way bigger standard deviation between the performance measures within the 100 bootstrap sequence of 1 quarter. Consequently, it is inadvisable to use the POET technique for real world use to create mean-variance portfolios with this data set. However, the LW and GL techniques do turn out to be a reasonable to apply for real world use. Where the Graphical Lasso shrinkage technique generally results in the portfolio with the highest risk adjusted returns and has the smallest standard deviation in the performance measures.

The paper continues followingly, in Section 2 the choice of data is explained. In the Methodology, the shrinkage technique of Ledoit and Wolf is explained in Section 3.1, POET in 3.2, Graphical Lasso in 3.3 and the Kourtis, Dotsis and Markellos (KDM) shrinkage technique in Section 3.5. The optimal regularization value is determined in Section 3.4. Then the choice of portfolios and the creation of the portfolios are explained in Section 3.6. In order to assess the performance of the portfolios in combination with the shrinkage technique Section 3.7 covers the performance measures. Then for the robustness check on the shrinkage technique the circular block bootstrap is used to create new data sequences. The circular block bootstrap and the robustness check are further explained in Sections 3.8 and 3.9. The Result Section is separated in multiple parts, first in sections 4.1 and 4.2 some general results are shown for the different portfolios shrinkage techniques and regularization values. In Section 4.3 and 4.4 the portfolio performance and robustness analysis, with the use of the circular bootstrap is shown. Finally Section 4.5 is a summary of all results.

2 Data

Actively traded stocks are needed to get a realistic view of how the portfolios perform in the real world. Therefore all stocks traded at the New York Stock Exchange (NYSE) with a market cap of $10,000,000 or higher at the beginning of April 2020 are included. So we truly have the stocks that are traded the most, in sense of amount and total traded value. The market cap of companies naturally change over time as companies grow, shrink or even go bankrupt. The kind of stocks that are popular also change over time, like the technology stocks for example. They are most prominent in market now, but where almost totally absent two decades ago. Therefore, to find results that are true for the current market, and applicable for the near future, we use all the stocks with a market cap higher or equal to

$10,000,000 in the beginning of April 2020 and not all the stocks that have a market cap higher than $10.000.000 in that certain quarter. The stock prices from January 1st 2010 until December 31st 2019 are used.

In the first quarter of 2010 there are 338 stocks and in the final quarter of 2019 there are 495 stocks. The change in available stocks is due to the creation of new companies that also created a market cap big enough in April 2020 to be selected. These companies are most often technology companies as they have skyrocket over past few years. The full data set is divided in quarters so 4 quarters each year, and the stocks returns of companies are only added to the quarter data set if their returns are available for the full quarter. Finally, we subtract the daily Fama French library [1] risk free rate from the daily CRSP return series without dividend. [2].

Table 1: Average descriptive statistics

	Return	**St.dev**	**Sharpe**	**Kurtosis**	**Skewness**
min	-0.8559	0.1132	-2.0602	3.1703	-11.6168
2.5%	-0.3562	0.1655	-0.7197	4.5654	-1.5437
mean	0.1048	0.4066	0.2764	22.9134	0.2847
97.5%	0.5626	1.1399	0.9929	85.7664	4.4018
max	2.0340	2.9576	1.4436	1.93E+03	41.3762

Note: The tables displays in the return column, the minimal yearly return, the 2.5 percentile yearly return, the average yearly return, the 97.5 percentile yearly return, and the maximal yearly return found in all the asset present in the complete data set from 2010 to 2019. Also the minimal, 2.5% average, 97.5% and maximal yearly values of the standard deviation and Sharpe ratio are displayed of all asset present in the complete data set from 2010 to 2019.

The SP500 is often considered as a good representation of the NYSE, and the average yearly return of the SP500 is between 8% and 12%. The 10.48% shown in Table 1 falls perfectly within that range. Meaning that on average our data set is a good representation of the NYSE. The 97.5% yearly returns also resonates well with the American Stock market. When a company has a extraordinary year the stock return in that year is often between 50% and 200%, Apple for example had a 88% return in 2019, and Amazon displayed a yearly stock return of 118% in 2015. Meaning that 97.5% and maximal value of yearly return in the data set is a good representation of the market. The minimal yearly return does not precisely represent the market, because in the true market bankruptcy is present which would result in a minimal return of -1. This discrepancy is a caused by selecting the stocks with the highest market cap at the beginning of April 2020. Therefore, none of the companies selected went bankrupt between 2010 and 2019. We expect this will not have a big impact on the final results. Because the

[1] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[2] Wharton Research Data Services. "CRSP Daily Stock" wrds.wharton.upenn.edu, accessed 04/1/2020.

disturbance a bankruptcy would cause, should be minimal for our portfolios due to the high number of assets used. In addition, there are still big losses present in the data set, and having asset that show a sudden big loss is far more common than having an asset in your portfolio that goes bankrupt.

3 Methodology

The big problem when the number of assets (N) is higher than the number of observations (T), is that the sample covariance matrix S (the maximum likelihood estimator of the covariance) is not of full rank, and therefore its inverse (S^{-1}) is not defined. Moreover, when $N < T$, the covariance can still be ill conditioned, if too many assets are used (high value for N). This is also referred to as a high-dimensional case. This ill conditioned covariance in turn results in badly or undefined inverse covariance matrices. So $N > T$ or just a high value for N are problematic when the classic maximum likelihood estimator for the covariance is used. The literature presents multiple shrinkage techniques that allow for the estimation of these high-dimensional covariance matrices resulting in well-conditioned and invertible covariance matrices. However, the final goal is not creating high-dimensional inverse covariance matrices, but an asset portfolio. So the question is, which of the techniques gives the better performing portfolio.

We analyze the out-of-sample performance of portfolios constructed by high-dimensional inverse covariance matrices, created with a shrinkage technique, and we also evaluate the robustness of the shrinkage technique. The methodology covers some of the shrinkage techniques suggested by the literature. The first is the Ledoit and Wolf shrinkage method Ledoit and Wolf (2003), Ledoit et al. (2004), Ledoit and Wolf (2004) shown in Section 3.1, the principal orthogonal complement thresholding estimator (POET) introduced by Fan et al. (2013) is explained in 3.2, the Graphical Lasso of Friedman et al. (2008) is presented in Section 3.3, and the inverse covariance shrinkage method of Kourtis et al. (2012) is shown in 3.5. The portfolio creation and performance measures are shown in Sections 3.6 and 3.7. The 100 different data sets for each quarter are created with the circular bootstrap of Politis and Romano (1992) and is explained in 3.8. Then Section 3.9 explains how the circular bootstrap is used to test the robustness of the shrinking techniques. Finally, Section 3.4 shows how the optimal regularization value is found.

3.1 Ledoit and Wolf Shrinkage

The Ledoit and Wolf shrinkage method for the covariance, introduced in Ledoit and Wolf (2003), Ledoit et al. (2004) and Ledoit and Wolf (2004), is a frequently used method as it is efficient in estimating the high-dimensional covariance matrices. They shrink the sample covariance to a stable more central matrix leading to a decrease in magnitude of the extreme values, and thereby reducing the estimation error. The Ledoit and Wolf covariance matrix is estimated by combining the sample covariance matrix S with a target matrix $\hat{F}$:

$$\hat{\Sigma}_{LW} = \hat{\delta}\hat{F} + (1-\hat{\delta})S. \tag{1}$$

As target matrix, Ledoit and Wolf suggest three different targets. In Ledoit and Wolf (2003) the covariance matrix of an one factor model is used, in Ledoit et al. (2004) the identity matrix is suggested and in Ledoit and Wolf (2004) the constant correlation matrix is taken as target. These methods are all fairly similar in the sense that they all shrink the sample covariance matrix to a target matrix in a linear fashion. They differ from each other by target matrix and the linear equation how the target and sample matrix are combined. However, because the general idea is quite similar, we only use the constant correlation one. In order to keep the sheer amount of results within reason. The constant correlation matrix $\hat{F}$ is defined as:

$$\begin{aligned} r_{ij} &= \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}}, \\ \bar{r} &= \frac{2}{(N-1)N}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} r_{ij}, \\ f_{ij} &= \bar{r}\sqrt{s_{ii}s_{jj}}. \end{aligned} \tag{2}$$

In which s_{ij} is the sample covariance, and r_{ij} and $\bar{r}$ are respectively the sample correlation and its average. The final missing part of equation 1 is the $\hat{\delta}$. Ledoit and Wolf show that $\hat{\delta}$ asymptotically acts like a constant over T, and that $\hat{\delta}$ can be constructed in the following way:

$$\begin{aligned} \hat{\kappa} &= \frac{\hat{\pi}-\hat{\psi}}{\hat{\gamma}}, \\ \hat{\delta} &= \max\left\{0, \min\left\{\frac{\hat{\kappa}}{T}, 1\right\}\right\}. \end{aligned} \tag{3}$$

This leaves three new parameters to be constructed. The parameters are defined as follows: $\hat{\pi}$ is the sum of the asymptotic variance of the sample covariance scaled by $\sqrt{T}$. Then $\hat{\psi}$ is the sum of the asymptotic covariance of shrinkage target and the sample covariance. Finally, $\hat{\gamma}$ is the measure of misspecification in the shrinkage target. Mathematically the parameters can be defined as:

$$
\begin{aligned}
\hat{\pi} &= \sum_{i=1}^{N}\sum_{j=1}^{N}\hat{\pi}_{ij} \quad \text{with} \quad \hat{\pi}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\left\{(y_{it}-\bar{y}_i)(y_{jt}-\bar{y}_j)-s_{ij}\right\}^2,\\
\hat{\psi} &= \sum_{i}^{N}\hat{\pi}_{ii} + \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}\frac{\bar{r}}{2}\left(\sqrt{\frac{s_{jj}}{s_{ii}}}\hat{v}_{ii,ij}+\sqrt{\frac{s_{ii}}{s_{jj}}}\hat{v}_{jj,ij}\right) \quad \text{with,}\\
\hat{v}_{jj,ij} &= \frac{1}{T}\sum_{t=1}^{T}\left\{(y_{jt}-\bar{y}_j)^2-s_{jj}\right\}\left\{(y_{it}-\bar{y}_i)(y_{jt}-\bar{y}_j)-s_{ij}\right\},\\
\hat{\gamma} &= \sum_{i=1}^{N}\sum_{j=1}^{N}(f_{ij}-s_{ij})^2,\\
\hat{\kappa} &= \frac{\hat{\pi}-\hat{\psi}}{\hat{\gamma}},\\
\hat{\delta} &= \max\left\{0,\min\left\{\frac{\hat{\kappa}}{T},1\right\}\right\}.
\end{aligned} \tag{4}
$$

The full derivation of $\hat{\delta}$ and the elements involved are shown in Ledoit and Wolf (2003), Ledoit et al. (2004) and Ledoit and Wolf (2004)[3]. The parameter $\hat{v}$ represents the asymptotic covariance between the diagonal sample covariance and the sample covariance. The inverse covariance with the Ledoit and Wolf shrinkage estimator is then given as the the inverse of equation 1: $\hat{\Theta}_{LW} = \hat{\Sigma}_{LW}^{-1}$

3.2 POET

The principal orthogonal complement thresholding estimator (POET) introduced by Fan et al. (2013) use the ranking of the eigenvalues to shrink the sample covariance matrix. The eigenvalues of the sample covariance matrix are ranked from largest to smallest. Then the POET estimator keeps the covariance matrix based on the first K principal components unchanged and applies a thresholding procedure to the covariance matrix based from the remaining $N-K$ principal components. To clarify, below is the mathematical representation, wherein $\hat{S}_K$ is the covariance matrix based on the first K principal components and $\hat{R}_K$ the covariance based on the principal components of $K+1$ to N,

$$
\begin{aligned}
\hat{S}_K &= \sum_{i=1}^{\hat{K}}\hat{\lambda}_i\hat{\xi}_i\hat{\xi}_i' + \hat{R}_{\hat{K}},\\
\hat{R}_{\hat{K}} &= \sum_{i=\hat{K}+1}^{N}\hat{\lambda}_i\hat{\xi}_i\hat{\xi}_i' = (\hat{r}_{ij})_{N\times N}.
\end{aligned} \tag{5}
$$

In the equation $\hat{\lambda}$ and $\hat{\xi}$ are respectively the eigenvalues and eigenvectors of the sample covariance, $\hat{K}$ is the number of the first principal components. The POET applies the shrinkage to $\hat{R}_{\hat{K}}$ and then combines the shrunken matrix $\hat{R}_K$ with $\hat{S}_K$ to form the final covariance matrix. This matrix can in turn be

[3] Ledoit and Wolf code available at: https://www.econ.uzh.ch/en/people/faculty/wolf/publications.html

inverted to construct the inverse covariance matrix. The shrinkage to $\hat{R}_K$ is represented followingly:

$$
\begin{aligned}
\hat{R}^{\zeta}_{\hat{K}} = (\hat{r}^{\zeta}_{ij})_{N\times N}, \qquad \hat{r}^{\zeta}_{ij} &= \begin{cases} \hat{r}_{ii}, & i=j, \\ \hat{r}_{ij} I(|\hat{r}_{ij}| \geq \tau_{ij}), & i \neq j, \end{cases} \\
\hat{\Sigma}_{POET} &= \sum_{i=1}^{K} \hat{\lambda}_i \hat{\xi}_i \hat{\xi}'_i + \hat{R}^{\zeta}_{\hat{K}}, \\
\hat{\Theta}_{POET} &= \hat{\Sigma}^{-1}_{POET},
\end{aligned} \tag{6}
$$

where I is an indicator function, τ_{ij} is the threshold specific to each element in $\hat{R}_K$, and is specified as follows:

$$
\tau_{ij} = \tau\sqrt{(\hat{r}_{ii}\hat{r}_{jj})}, \quad \text{for } \tau > 0. \tag{7}
$$

In this paper τ is equal to $\frac{1}{N}$, following the adaptive thresholding for sparse covariance estimation setup of Cai and Liu (2011). Now only the value of $\hat{K}$ is still left to be determined. This is done with the paper of Bai and Ng (2002)[4] determining the number of factors, they estimate $\hat{K}$ by:

$$
\hat{K} = \text{argmin}_{0\leq K_1 \leq M=N} \log\left(\frac{1}{NT}||Y - T^{-1}Y\xi_{k1}\xi_{k1}||^2_F\right) + K_1 g(T,N). \tag{8}
$$

Where M is the predetermined upper bound of $\hat{K}$, ξ_{k1} is a $N \times K_1$ matrix where the columns are $\sqrt{N}$ times the eigenvectors. That correspond to the largest K_1 eigenvalues of the full $N \times N$ eigenvalue matrix $Y'Y$. Where Y represents the matrix of input variables (for this paper that is the matrix of asset returns of one quarter). Finally $g(T,N)$ is the penalty function

$$
g(T,N) = \frac{N+T}{NT}\log\left(\frac{NT}{N+T}\right). \tag{9}
$$

The eigenvalue and eigenvectors, used in equation 5, are constructed with the sample covariance matrix. However, the maximum likelihood estimator of the covariance matrix in a high-dimensional case can lead to big estimation errors. In turn that can lead to even bigger estimation errors in the eigenvalues and vectors. Therefore an additional POET technique is run, where the eigenvectors and eigenvalues are constructed with the sample correlation matrix. When there is a big difference between the two inverse covariance matrices it indicates that the estimation errors does have a significant impact on the performance of the POET technique. The POET with these values based on the sample correlation matrix is called the POET-C and is constructed exactly the same as the POET technique.

3.3 Graphical Lasso

The graphical lasso as described in Friedman et al. (2008)[5], is a special lasso regresison. The standard lasso regression introduced by Tibshirani (1996) is often used to reduce the estimation error

[4] Bai and Ng Determining the Number of Factors in Approximate Factor Models code of Christophe Hurlin and Valerie Mignon available at: http://www.runmycode.org/companion/view/69

[5] Graphical lasso code of Hossein Karshenas available at: http://statweb.stanford.edu/ tibs/glasso/

in a model with a high number of variables. This is accomplished by setting the weights (β), of the variables with insignificant influence, to zero. This is done by penalizing the likelihood for every non-zero β. Meaning that the increase in likelihood, for giving a weight to β, needs to be higher than the penalty. When this is not the case the maximum likelihood is found with a zero weight for that particular β. Thereby reducing the number of variables used. The β is estimated in the following way:

$$\hat{\beta}^i = \text{argmin}_\beta \left\{ ||y_i - X\beta||^2 + \rho||\beta||_1 \right\}. \tag{10}$$

This is the classic lasso regression in which y and X are respectively the response variable and explanatory variables, $||\beta||_1$ is the L_1 norm: the sum of the absolute values of the elements. The ρ is the regularization parameter. This parameter that determines the penalty height for each β that is non zero. Therefore, a higher value of ρ causes more values of β to be equal to zero and vice versa. This Lasso shrinkage method can be extended to not only shrink the number of parameters in a model, but can also be used on the inverse covariance matrix. Meinshausen et al. (2006) show with their *neighbourhood selection* approach that with use of the lasso method, a more sparse inverse covariance can be formed that is non-singular and well conditioned in a high dimensional case and or when $N > T$. However, this method is only seen as an approximation to the real problem. They consistently estimate a set of zero and non-zero values for the inverse covariance matrix, but do not maximize the L_1 penalized log-likelihood. Whereas, Banerjee et al. (2008) have created a different algorithm that does maximize the L_1 penalized log-likelihood. They start with the maximisation problem of the penalized log-likelihood:

$$L = \max \left\{ \log \det \hat{\Theta} - \text{tr}(S\hat{\Theta}) - \rho||\hat{\Theta}||_1. \right\}. \tag{11}$$

In the formula, L is defined as the log-likelihood, $\hat{\Theta}$ is the estimate of Σ^{-1}, S is the sample covariance matrix and tr(.) denotes the trace. Banerjee et al. (2008) show that Equation 11 is convex, and thereby go over to the estimation of Σ in the following way. They set W as the estimate of Σ. Then show that the problem can be solved by optimising over each row and corresponding column of W. This is done with a block coordinated descent, by defining W and S as follows:

$$W = \begin{pmatrix} W_{11} & \omega_{12} \\ \omega_{12}^T & \omega_{22} \end{pmatrix}, \quad S = \begin{pmatrix} S_{11} & s_{12} \\ s_{12}^T & s_{22} \end{pmatrix}. \tag{12}$$

Then ω_{12} can subsequently be computed as:

$$\omega_{12} = \text{argmin}_y \left\{ y^T W_{11}^{-1} y : ||y - s_{12}||_\infty \leq \rho \right\}. \tag{13}$$

Banerjee et al. (2008) go on and solve this problem with a interior-point procedure, shifting the rows and columns to ensure the final column is always the target column. Equation 13 is solved for each column and W is updated after each calculation of ω_{12}. This process is repeated until convergence. When the initial matrix W is positive definite matrix, it continues to be so after each and every update. Using this convexity Banerjee et al. (2008) also show that solving 13 is the same as solving:

$$\operatorname{argmin}_{\Phi}\left\{\frac{1}{2}||W_{11}^{1/2}\Phi-\phi||^2+\rho||\Phi||_1\right\}. \tag{14}$$

With ϕ defined as $\phi = W_{11}^{1/2}s_{12}$. So when Φ solves 14, then $\omega = W_{11}\Phi$ solves expression 13. This minimization problem mirrors the standard lasso regression in equation 10. This is the basis used by Friedman et al. (2008) to build their graphical lasso. Friedman et al. (2008) start by stating that when:

$$W=\hat{\Sigma} \quad \text{and} \quad \Theta=\Sigma^{-1} \quad \text{it follows that} \quad W\Theta=I. \tag{15}$$

This can in turn be split up in the following way:

$$\begin{pmatrix} W_{11} & \omega_{12} \\ \omega_{12}^T & \omega_{22} \end{pmatrix}\begin{pmatrix} \hat{\Theta}_{11} & \hat{\theta}_{12} \\ \hat{\theta}_{12}^T & \hat{\theta}_{22} \end{pmatrix}=\begin{pmatrix} I & 0 \\ 0^T & I \end{pmatrix}. \tag{16}$$

With Equations 15 and 16, the maximization problem in Equation 11 can mathematically be transformed to:

$$\begin{gathered} W-S-\rho\cdot\Gamma=0 \\ \begin{pmatrix} W_{11} & \omega_{12} \\ \omega_{12}^T & \omega_{22} \end{pmatrix}-\begin{pmatrix} S_{11} & s_{12} \\ s_{12}^T & s_{22} \end{pmatrix}-\rho\cdot\begin{pmatrix} \Gamma_{11} & \gamma_{12} \\ \gamma_{12}^T & \gamma_{22} \end{pmatrix}=0. \end{gathered} \tag{17}$$

Wherein, $\Gamma_{ij}=\text{sign}(\Theta_{ij})$ if $\Theta_{ij}\neq 0$, and when $\Theta_{ij}=0$, then $\Gamma_{ij}\in[-1,1]$. This leads to the observation that the upper right part of equation 17 and equation 14 can respectively be written as

$$\omega_{12}-s_{12}-\rho\cdot\gamma_{12}=0, \tag{18}$$

$$W_{11}\Phi-s_{12}+\rho\cdot\nu=0, \tag{19}$$

with $\nu\in\text{sign}(\Phi)$. With these equation the graphical lasso algorithm can be run. First start with $W = S+\rho I$, and a given value ρ, which in this case is the highest absolute value in the initial matrix W. [6] Note, that the Graphical Lasso also allows for use of a regularization matrix instead of a single value. The advantage of using a matrix is that every variable has its own regularization value, and when optimized this can lead to even higher maximum likelihood values. Then with the chosen regularization value solve the lasso problem of Equation 14 for each $j=1,\ldots,N$ with the inputs W_{11} and s_{12}. This results in a $N-1\times 1$, $\hat{\Phi}$ vector. Use the estimated Φ vector to replace the corresponding column and row in W with $\omega_{12}=W_{11}\hat{\Phi}$. The inverse covariance $\hat{\Theta}_{GL}$ is in turn easily computed given

[6] This is a generic choice of ρ in order to have a more restricted inverse covariance as a baseline.

equation 16. The partitioning enables us to calculate $\hat{\theta}_{12}$ and $\hat{\theta}_{22}$ for each iteration of the graphical lasso as follows:

$$\begin{aligned}
\hat{\theta}_{12} &= -W_{11}^{-1}\omega_{12}\hat{\theta}_{22}, \\
\hat{\theta}_{22} &= 1/\left(\omega_{22} - \omega_{12}^{T} W_{11}^{-1}\omega_{12}\right).
\end{aligned} \tag{20}$$

The full estimated inverse covariance matrix with use of the graphical lasso regression is defined as $\hat{\Theta}_{GL}$.

3.4 Optimal Regularization Value Graphical Lasso

The performance of the portfolio based on Graphical Lasso inverse covariance, is greatly influenced by the choice of its regularization value ρ. When the value of ρ becomes large, the shrinkage is so strong the inverse covariance becomes a diagonal matrix. When the regularization value is extremely small, there is very little shrinkage, which in turn can lead to bigger estimation errors. Subsequently, the optimal regularization value needs to be found in order to fully utilize the capabilities of the Graphical Lasso. In the Section 3.3 the highest absolute value of the covariance is suggested the regularization value. This regularization value results in a complete shrinkage, meaning the covariance matrix is a diagonal matrix. This is most likely not the optimal regularization value.

Therefore, 50 equally spaced regularization values are tested between zero and the highest absolute value of the sample covariance matrix as the largest value (the 50th value). Further, the Graphical Lasso technique also allows for the use of a regularization matrix instead of a single value. Therefore, in addition to the 50 single regularization values also 10 different regularization matrices are examined. The 10 matrices are constructed as the absolute value of the sample covariance matrix divided by 10,9,...,2 and 1. With the set regularization values the inverse covariance matrix is estimated and subsequently the different portfolios. This is repeated 100 times with the use of the circular bootstrap, to create the data sets, for each quarter.

The values set in the single regularization value and regularization matrix are not constant across the different quarter. The regularization value is determined with the sample covariance matrix of the current quarter. Meaning that single regularization value in quarter t at index 1,10 and 50 are always defined as:

$$\rho_{1,t} = \frac{\max(|S_t|)}{50}, \quad \rho_{10,t} = \frac{\max(|S_t|)}{41}, \quad \text{and} \quad \rho_{50,t} = \max(|S_t|). \tag{21}$$

This is also true for the regularization matrix, wherein the full absolute sample covariance matrix is divided by 10 to 1 for the different values of ρ, instead of the maximum value of the absolute covariance matrix.

Finally, to see if there is some correlation between the regularization values that result in the better performing portfolio, we make a combination regularization value. In this combination value we start of in the first quarter with the 50th regularization value $\rho_{50,1}$ (this corresponds to the full shrinkage case). Then the regularization index that in quarter 1 resulted in the highest average Sharpe ratio, over 100 bootstrap data set, is used as the regularization index of quarter 2, and the index that resulted in the highest average Sharpe in quarter 2 is used as index for quarter 3 and so on. For example in quarter 1, $\rho_{30,1}$ resulted in the highest average Sharpe ratio then in quarter 2, $\rho_{30,2}$ is used as the regularization value. Let say in quarter 2, $\rho_{33,2}$ gave the highest average Sharpe ratio then in quarter 3, $\rho_{33,3}$ is used as regularization value, and so on to the final quarter.

The average performance of this combination regularization portfolio can indicate if there is any correlation in the choice of regularization value and Sharpe ratio, across subsequently quarters. When there is a correlation the combination regularization value can result in a substantially better performing portfolio compared to portfolios based on single regularization value. This set-up is also copied for the matrix case and is called the combination regularization matrix. For both the single regularization value and the matrix value only the GMV portfolio is created. Then the regularization value (single value and matrix) that gives the best performing portfolio and the combination case (single and matrix) are used in combination with the GMV-C and GMV-W portfolios. [7] In order to see if the extra restriction can further improve the performance

3.5 Inverse Covariance Shrinkage

The method introduced by Kourtis et al. (2012) (KDM) builds upon the shrinkage regression of Ledoit and Wolf (2003). However, this technique does not shrink the covariance but directly targets the inverse covariance. A logical approach when the final goal is the creation of mean-variance portfolios. Because it can ease the complications the inversion brings with it, when trying to optimize a portfolio with respect to portfolio performances. Therefore instead of equation 1, Kourtis et al. (2012) propose,

$$\hat{\Theta}_{KDM} = c_1 S^{-1} + c_2 \hat{\Omega}, \tag{22}$$

as the regression for the inverse covariance. Where $\hat{\Omega}$ is an $N \times N$ symmetric positive definite target matrix. However, an estimation of the inverse covariance matrix is needed for equation 22. This is not possible with the classic maximum likelihood estimator when $N > T$. Therefore the Moore-Penrose inverse is used

[7] The GMV-C and GMV-W portfolios are not constructed with all the different regularization values as this would results in a extremely long running times due to the optimization time needed to create one of the portfolios. Let alone 62 portfolios for all simulation in each quarter.

Consequently, just as in the POET technique, this can result in big estimation errors in the eigenvalues and eigenvectors. Therefore the inverse covariance is constructed with the eigenvalues and eigenvectors based on the sample covariance and sample correlation matrix. The respective models are indicated as KDM and KDM-C and all other steps are identical for both models. The target matrix, in equation 22 can also be a combination of two different matrices, for example the identity matrix and inverse covariance based on the 1-factor model. Using two target matrices, equation 22 changes to:

$$\hat{\Theta}_{KDM} = c_1 S^{-1} + c_2 I + c_3 \hat{F}. \tag{23}$$

The target matrices I and $\hat{F}$ are respectively the identity and target matrix[8]. This shrinkage technique is optimised for portfolio creation in the sense computation time. The inverse covariance itself is therefore not an output of this regression, because the weights of the portfolio are created by a linear combination of the weights constructed with the sample inverse covariance and the target inverse covariance matrix. This eliminates the need to directly construct the covariance matrix based on the sample and target covariance matrix. The final weights correspond to the weights of a Global Minimum Variance (GMV) portfolio[9]. The computation of the weights corresponding to the inverse covariance in Equation 22 and 23 are shown in the Appendix Section A.1. In this paper two target matrices are used. The two used are the suggested matrices in Kourtis et al. (2012) namely: a combination of the identity matrix and the inverse covariance of the 1 Factor model, where the market return is the factor.

3.6 Portfolio Creation

There are many different weighting schemes to create a portfolio, but the Global Minimum Variance (GMV) has proven over the years to create a reliable and reasonable well performing portfolio. Further, Jagannathan and Ma (2003) shows that the inclusion of no-short sale constraint can reduce the optimal portfolio risk even when this constraint is wrong. This is emphazised by the results of DeMiguel et al. (2009), they show that the GMV with short sale constrained (GMV-C) and the GMV with generalized weight constraints (GMV-W) where the overall best, out of sample, performing portfolios with different number of assets and observations. They tested 15 different portfolios, and the GMV-C and GMV-W where the most stable across the different data sets with different number of assets. They where not the best performing when looking at the Sharpe ratio for each data set separately, but they where one of the better performing portfolios in each tested data set. Therefore,

[8] We use the identity matrix and the 1 Factor model inverse covariance as target matrix, the identity matrix however can also be swapped for a different target matrix and thereby using 2 different target matrices

[9] Different weighting schemes can be used. These need to be modified in the same way as w_{KDM} in order for this regression to work.

because they generally give high Sharpe ratios regardless of the data set, the GMV-C and GMV-W weighting schemes are also used to create portfolios.

Finally, to see if the inclusion of a shrinkage technique in portfolio management has any real world usage. We compare the created portfolios to the equally weighted portfolio. As this portfolio is really simplistic, but has proven really hard to beat as shown by DeMiguel et al. (2009). When the equally weighted portfolio dominates the performance of the mean-variance portfolios based on the inverse covariance from the shrinkage techniques, it is not beneficial for portfolio managers to us these methods in a real world application.

In order to truly test the real world performance, all portfolios are made out-of-sample. Meaning, that at the start of each new quarter, the weights of the portfolio, for that quarter, are based on the inverse covariance created with the asset returns of the previous quarter. This can cause a miss match in available assets. As stated before the number of assets is not constant throughout the test period. Therefore, it can be the case that the inverse covariance of the previous quarter consist of fewer asset than in the current quarter are available. This then leads to the exclusion of the assets created in the current quarter. Subsequently, the new assets will be included in next quarters portfolio as the asset information is then available as previous quarter data. The GMV portfolio is constructed like:

$$w_{ij} = \frac{\hat{\Theta}_i \iota}{\iota' \hat{\Theta}_i \iota}, \quad j = GMV, \tag{24}$$

where i represents the different shrinkage techniques, $i \in \{CBB, LW, POET, GL, KDM\}$, and j is for notational consistency, to differentiate between the different weight schemes later on. Then the GMV-C and GMV-W are represented in the following way:

$$\begin{aligned}
\min \quad & \sigma_p^2 = w_{ij} \hat{\Sigma}_i w_{ij}, \quad j = \begin{cases} \text{GMV-C}, & \text{if } x = 0, \\ \text{GMV-W}, & \text{if } x = \frac{0.5}{N}, \end{cases} \\
\text{st.} \quad & \iota' w_{ij} = 1, \quad j = \begin{cases} \text{GMV-C}, & \text{if } x = 0, \\ \text{GMV-W}, & \text{if } x = \frac{0.5}{N}, \end{cases} \\
& w \geq x.
\end{aligned} \tag{25}$$

In which σ_p^2 and $\hat{\Sigma}_i$ are respectively the portfolio variance and the constructed covariance by technique i. When $x = 0$ the optimization creates the GMV-C weights (j=GMV-C), if $x = 0.5 \times \frac{1}{N}$ it constructs the GMV-W weights (j=GMV-W). The value $x = 0.5 \times \frac{1}{N}$ forces every asset to have a weight higher or equal to $0.5 \times \frac{1}{N}$, and $x = 0$ forms the no short sale constraint. The three weighting schemes will all be used in combination with every shrinkage technique covered before. When the GMV-C or GMV-W portfolio is used in combination with one of the shrinkage techniques it can be seen as a double shrinkage. First a shrinkage on the inverse covariance matrix estimate and then a shrinkage on the constructed weights. This is a beneficial attribute when the shrinkage on the inverse covariance is not successful enough in reducing the big estimation errors that come with high-dimensional data sets.

3.7 Portfolio Performance

The industry standard performance measures are: standard deviation, Sharpe ratio, turnover and certainty equivalent(CEQ). All the measures are out of sample, as the portfolios are created only with historical returns. The Sharpe ratio (Sharpe, 1966) is a risk adjusted return measure and is calculated like:

$$SR_{ij,t} = \frac{\hat{\mu}_{ij,t} - RF_t}{\hat{\sigma}_{ij,t}}. \tag{26}$$

The parameters $\hat{\mu}$ and $\hat{\sigma}$ are the estimate average return and standard deviation, RF_t is the risk free rate in quarter t, $SR_{ij,t}$ is the Sharpe ratio in quarter t based on portfolio ij, where i is the shrinkage technique $\left(i \in \{CBB, LW, POET, GL, KDM\}\right)$ and j is the weighting scheme $\left(j \in \{GMV, GMV-C, GMV-W\}\right)$. Because daily asset returns are used, the resulting Sharpe ratio is also daily. To get the yearly Sharpe ratio multiply by the square root of 252 (there are approximately 252 tradings days in one year). The CEQ is the lowest risk-free return an investors is willing to accept instead of the risky return of the portfolio, and it is constructed by:

$$CEQ_{ij,t} = \hat{\mu}_{ij,t} - \frac{\gamma}{2}\hat{\sigma}^2_{ij,t}, \tag{27}$$

the resulting $CEQ_{ij,t}$ stands for the certainty equivalent measure of a portfolio with shrinkage technique i and weighting scheme j in quarter t. The risk aversion level of the investor is displayed by γ. The CEQ is calculated with γ equal to 1,5 and 10. The turnover is used to indicate how much trading is needed every time the portfolio is updated. Turnover is defined as the average sum of the absolute difference between the weights before and after updating of all assets:

$$\text{Turnover}_{ij} = \frac{1}{Q-M}\sum_{t=1}^{Q-M}\sum_{l=1}^{N}\left(|w_{ij,l,t+1} - w_{ij,l,t+}|\right). \tag{28}$$

Where Q is equal to the total number of quarters, M is the update rate in quarters, in this case each quarter so $M = 1$, $w_{ij,l,t+1}$ is the portfolio weight of asset l when shrinkage technique i, and weighting scheme j are used at quarter $t+1$, and $w_{ij,l,t+}$ is the portfolio weight right before updating. In the case of a equally weighted portfolio $w_{ij,l,t} = w_{ij,l,t+1}$, but due to price changes in the asset is $w_{ij,l,t+}$ $\neq \text{w}_{ij,l,t+1}$. Therefore, the portfolio needs to be update to become an equally weighted portfolio again, and the average of absolute sum of this difference is the Turnover. So when the portfolios are updated quarterly, $w_{ij,l,t+}$ is then given as:

$$w_{ij,l,t+} = w_{ij,l,t} \prod_{t^*=1}^{Q^*} (1 + R_{l,t^*}), \tag{29}$$

R represents the excess return of stock l on day t^* in quarter t, and Q^* is number of days in quarter t. As mentioned before the number of assets used in the portfolio can change with each quarter. This of course influences the turnover massively, not only by the increase (decrease) in weight to that assets (from 0 to something) but also by the reduction(increase) in weight it causes to the other assets. However, this bias is included in all portfolios, so no portfolio will suddenly have a better turnover than other portfolios due to this bias. Consequently, the turnover of these portfolios can still be compared to each other. The turnover calculation itself also stays the same as described. The only addition is that the weight of the new assets in the portfolio is set to zero at $t+$ and then the turnover can be constructed.

3.8 Circular Block Bootstrap

All the techniques base their inverse covariance on the given data set. However, in order to use the most relevant data only 3 months of stock returns are used to create the inverse covariance. This itself is of course not a problem, because the methods are designed to work with less data points than variables. However, small data sets usually result in higher estimation errors, which in turn results in volatile portfolio performance. In order to test what technique gives the better performing portfolio, we can not only look at the performance of the portfolios, but must also look at the robustness of the shrinkage technique. Does the shrinkage technique give similar inverse covariance for identically distributed data sets. Therefore, instead of creating one portfolio for each methods for every quarter, 100 portfolios are created for each quarter (for each method).

In order to create 100 different portfolios, 100 different data sets are needed for every quarter. Due to the fact asset returns are used, we need a model that can simulate time series data that is cross correlated. The classic multivariate volatility models or multivariate GARCH models are not designed to use hundreds of variables, as this will lead to big estimation errors. There are some more sophisticated approaches that can deal with very large numbers of variables, like the model proposed by Chib et al. (2006). They show that Bayesian estimations can be used to create a multivariate stochastic volatility model with more than 500 variables. However, this method needs to run an extensive Monte Carlo simulation to converge before the model can be used. This would greatly increase the running time, because for every quarter a new model needs to be created. Additionally, our goal is not to find an optimal multivariate volatility model for this data set, but to test the stability of the shrinkage technique and their performance when used in portfolio management. Therefore a quick method is needed that can simulate a new data set that incorporates the fact that time series data is used and that the assets are correlated to each other.

This leads to the circular block bootstrap technique presented by Politis and Romano (1992). The method is based on the paper of Efron (1979) who showed that the bootstrap method can satisfactory estimate the probability density distribution F when given a random sample sequence of observations from that same distribution. This means that by bootstrapping a sample set of observations a new data sequence can be obtained that recreates the distribution, and therefore in essence simulate a new data set with the same distribution. The bootstrap itself is a simple procedure, it randomly chooses a data point from the sample set with replacement until it has sampled the amount of observations specified beforehand. This idea can be applied to our problem with some modifications.

First all the assets in the quarterly data set needs to be bootstrapped at the same time, otherwise the general market influence and the cross influences are lost. So instead of univariate drawing data points, a index number is drawn and all values on that index number in the original data is added to the bootstrap sequence. Secondly the data is time dependent, meaning that return of yesterday have an influence on the return today. This characteristic can be included by bootstrapping blocks of data instead of single data point. Using blocks ensures that influence of time is still present. A block bootstrap can draw blocks of any size b from the data set. However, by using blocks the bootstraps no longer chooses the observations completely random. Because some data points now have more chance of getting picked than others. When $b = 5$ for example the first and last observation are only in 1 possible block while observation 6 to $T-5$ are in 5 possible blocks.

Politis and Romano (1992)[10] circumvent this problem by modifying the method of Efron (1979). The initial set up is the same as of a standard block bootstrap setting the block size to b consecutive observations. However, instead of not being able to choose the final few indices because the block size is larger than the remaining data points, it uses a modulo operation. By using the modulo operation they ensure all observation have the same possibility to be drawn. So when T = 63 and index 61 is drawn with block size 5, the block that is added consist of the values on the index (61,62,63,1,2). This added restriction to the normal block bootstraps creates the circular block bootstrap. They show in the paper that this method results in an unbiased bootstrap distribution. The circular block bootstrap is used in many financial applications as shown in Hansen and Lunde (2005), Ledoit and Wolf (2008), Kosowski et al. (2006), Kosowski et al. (2007) and Ruiz and Pascual (2002) for example.

[10] Cicular bootstrap code Ledoit and Wolf available at: https://www.econ.uzh.ch/en/people/faculty/wolf/publications.html

3.9 Robustness Shrinkage Techniques

The circular block bootstrap is performed for each quarter (approximately 63 days), with a block size $b = 5$ [11], on all stocks simultaneously and keeps drawing until a new quarter of data is sampled. So, when a quarter has 63 days of observation, 65 new observations are drawn. The newly sampled data set is then used in combination with all shrinkage techniques, as shown in Sections 3.1,3.2, 3.3, and 3.5, to create the inverse covariance and subsequently the GMV, GMV-C and GMV-W portfolios. This is repeated 100 times to create, for each technique, 300 unique mean-variance portfolios. The simulation is done from 2010 to 2019.

Consequently, the use of the circular bootstrap results in multiple data sequences generated from the same data set. This enables us to test the robustness of the shrinkage techniques in creating high-dimensional covariance matrices an or inverse covariance matrices. Because, the bootstrap gives an unique data set that resemble the distribution of the original data set. Therefore, it is expected that the shrinkage technique gives comparable inverse covariance matrices, and thereby comparable portfolios, for every circular bootstrap sequence. However, when the performance of the portfolios are vastly different (within a singular quarter) it indicates that the shrinkage technique does not give comparable inverse covariance matrices, and is therefore not reliable to use for making mean-variance portfolios. The robustness of the shrinkage technique can easily be shown by looking at standard deviation of the performance measures within on quarter. Because large standard deviation in the performance measures for the same mean-variance portfolio can only be caused by significant differences in the inverse covariance matrix.

3.10 Direct Implementation Shrinkage Techniques

Directly constructing the point estimates of the performance measure for each quarter can give a good indication how well the different shrinkage technique perform. All quarters of asset returns are directly used in combination with the different shrinkage technique to create the three mean-variance portfolios. By comparing the average performance measures of the portfolios we can see if there are any shrinkage technique that result in a subpar portfolio. When there is a shrinkage technique that creates a noticeably lesser performing portfolio than the equally weighted portfolio, this shrinkage technique can be disregard for the remaining of the paper. Even if the disregarded shrinkage technique turns out to be extremely robust in estimating the inverse covariance and subsequently the mean-variance portfolio, it does not matter for a portfolio manager. The simplistic equally weighted portfolio already gives better performance, and does not suffer from any robustness

[11] b=5 is a generic choice to shorten the running time. The block size calibrate calculates the optimal block size and is also included in the Ledoit and Wolf Circular bootstrap code

issues. Consequently, making the use of that shrinkage technique irrelevant in combination with this data set.

3.11 Exclusion Financial Factors

The inclusion of Financial Factor information in the estimation of the covariance matrix and subsequently the inverse covariance matrix, generally result in better performing portfolios. The Financial Factor information reduce the estimation errors in the classic maximum likelihood estimator of the covariance matrix. However, we have decided not to include the Financial Factor information in this paper.

The reason not to include them is because in the first place this paper tries to identify if the shrinkage techniques give reasonable portfolio results with high-dimensional data sets. Meaning, do the performance measures give realistic values, but more importantly can the shrinkage technique robustly estimate the inverse covariance with similarly distributed data sets. The goal is not, to find the absolute best performing portfolio in a high-dimensional set up. In addition, the inclusion of the Financial Factor information would bring a number of arbitrary choices with it that could impact the robustness investigation.

Where is the Financial Factor information included, in the sample covariance matrix, in the target covariance matrix, into both, or is the shrinkage technique changed to incorporate the information directly. Every different choice could impact the robustness of the shrinkage technique, and then it is not clear if the shrinkage technique estimates a robust inverse covariance due to its own set-up or due to the Financial Factors. Furthermore, which Factors do you use and how many Factors. Therefore, the inclusion of Financial Factor information is interesting, and will most likely result in better performing portfolios and perhaps even a robuster estimate of the covariance matrix, but it is something for further research outside of this paper.

4 Results

The Result Section covers in 4.1 the performance results when the quarterly asset return data is directly used with the different shrinkage techniques. In Section 4.2 multiple regularization values and regularization matrices for the Graphical Lasso are analyzed on their resulting portfolio performances. In Section 4.3 the portfolio performances with the use of the 100 bootstrap data sequences in each quarter are displayed. The portfolios in Section 4.3 are examined on their robustness across the 100 bootstrap sequences in Section 4.4. Finally, Section 4.5 gives a summary of the main results.

4.1 Results Direct Implementation Shrinkage Techniques

Before running all shrinkage techniques in combination with the circular bootstrap, we first run them directly on the quarter return series. The average quarterly performance is shown in Table 2, the standard deviation, Sharpe ratio and CEQ are scaled to yearly performance measures. Looking at the GMV portfolio performance it is clear that only the Graphical Lasso can come close to the performance of the equally weighted portfolio. With a Sharpe ratio and standard deviation of 0.9292 and 0.0978 for the Graphical Lasso and 0.9308 and 0.1009 for the equally weighted portfolio. All the other portfolios perform subpar in one or multiple performance measures in comparison with the equally weighted portfolio.

The Kourtis Dotsis Markellos (KDM) and POET show extreme results in turnover and negative Sharpe ratio values. The KDM-C performs better than the KDM when looking at standard deviation and Sharpe ratio. However the turnover, which already is extremely big, becomes even bigger. So changing the way the eigenvalues and vectors are constructed still does not make the KDM technique reasonable for real world use. Because an average quarterly turnover of 116.000 means you need to shift 116.000 times your entire invested capital every time you update the portfolio. This is of course absurd and not realistic.

The POET-C does improve some of the measurements to be reasonable values compared to the POET, but these values are still extremely bad and the turnover keeps being unrealistically large. These large turnovers in Global Minimum Variance portfolios are however not unique for this paper. In the paper of DeMiguel et al. (2009), turnover point estimates of 13.10E+03 are found when testing the performance of a GMV portfolio. In Appendix Section A.2 Table 9 is shown. This is the same table as Table 2, but with yearly input data instead of quarterly. With a yearly data set the standard deviation and Sharpe ratio do improve to more reasonable values for KDM, KDM-C and POET, but the turnover is still extremely large. Meaning that the techniques do work better when larger sample sets are used, but still suffers from unreasonable big values in the turnover. The final shrinkage technique used with the GMV portfolio is the Ledoit and Wolf method. This method gives reasonable results for all performance measures. However it is dominated by the performance of the equally weighted and Graphical Lasso portfolios, and also shows a substantial value for the turnover. The turnover value is not as unreasonably high as the KDM and POET techniques, but having to shift 2.7809 times the invested wealth every quarter is still quite high.

The bad performance of the POET and KDM techniques can partly be explained by the aggregation of estimation errors. They both use the eigenvalues and eigenvectors to estimate an initial inverse covariance matrix. However, these eigenvalues and eigenvectors are based on the sample covariance or correlation matrix, and due to the high number of variables it is likely that these matrices suffer from high estimation errors. Then basing eigenvalue and vectors on these matrices can result in values with even bigger estimation errors. In addition to the higher estimation errors, the big pitfall of a GMV portfolio is that is has a tendency to maximize the effects of estimation errors in the input data, this affect is shown by Jobson and Korkie (1980) and Michaud (1989). Therefore combining the inverse covariance matrix, that suffers from higher estimation errors, with a GMV portfolio most likely results in the extreme performance measure shown in Table 2.

Table 2: Performance measure results GMV, without circular bootstrap

		EQ	KDM	KDM-C	POET	POET-C	LW	G-Lasso
GMV	**St.dev**	0.1009	7.9115	0.4323	-4.69E+04	1.7342	0.0676	0.0978
	Sharpe	0.9308	-0.1264	0.9187	-0.4757	-0.5704	0.6459	0.9292
	CEQ $\gamma = 1$	0.0772	-4.74E+287	0.1720	-1.02E+49	-1.13E+46	0.0406	0.0745
	CEQ $\gamma = 5$	0.0122	$-\infty$	-0.7224	-2.28E+151	-3.10E+138	0.0284	0.0111
	CEQ $\gamma = 10$	-0.0677	$-\infty$	-3.1466	-1.53E+213	-2.44E+197	0.0134	-0.0669
	Turnover	0.0598	1.10E+03	1.16E+05	8.05E+02	4.32E+03	2.7809	0.1261

Note: The table shows the average performance measure for the GMV portfolio in combination with each shrinkage technique, and the Equally Weighted portfolio, for the time period 2010-2019. The average Expected Return (ER), Standard deviation (St.dev), Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures. The KDM-C and POET-C are the same shrinkage technique as KDM and POET, but the eigenvalues and vectors are estimated with the sample correlation matrix instead of the sample covariance. The values with ∞ are too big to display or to calculate.

Consequently, the KDM, KDM-C, POET, and POET-C techniques are subpar choices to use in combination with a GMV portfolio for this data set. The LW technique does give reasonable performance measure, but they are subpar to the equally weighted and the GL. The LW is therefore not the best choice to create a GMV portfolio with the given data set.

The GMV-C and GMV-W portfolios in Table 3 show a general performance improvement for all shrinkage techniques, except for the Graphical Lasso. The KDM and KDM-C based portfolio however still show extremely large turnovers in the range of 100,000. This is a peculiar result, because the weight restriction should restrict the weights to be between 0 and 1 (or 0.5/N and 1 for the GMV-W). This means that the maximum difference in weight between two quarters for one assets is 1.[12] Therefore the average turnover must be smaller than 495 (maximum amount of assets is 495). Hence, it is easy to conclude that optimization function could not restrict some of the assets weights to adhere to the set weight restrictions. This can lead to extreme negative and or positive weights in one quar-

[12]It could technically be case that when there is a massive price increase in an assets, that the position in that asset grows so much that is bigger than the initial investment. When this is the case the maximum difference in weight between two quarters can be higher than 1. However, for simplicity 1 is taken as the bound as this bound will never be crossed for all assets at once or a high number of assets at once.

ter and a normal weight that adheres to the restriction in the following quarter, creating a extreme turnover between the two quarters.

The optimization can fail because it is simply impossible to find weights that follow all the restrictions, but the more likely scenario is that the optimization function reached one or multiple of its termination criteria and gave its final iteration as output. In order to get results that do adhere to the given restriction the termination criteria in the optimization function need to be widened. This will naturally lead to a longer computation time for the portfolios, and that is not desirable for portfolio managers. Therefore, the KDM and KDM-C are not analysed any further.

Comparing the POET, POET-C and LW results of Tables 2 and 3 we see a clear improvement in Table 3. All portfolios now show reasonable performance measures. So it is clear that the inclusion of weight restrictions is beneficial. The 3 techniques even come close to the the performance of the equally weighted portfolio. The POET and POET-C now show positive Sharpe ratios for both the GMV-C and GMV-W portfolio wherein the POET is the clear better performer of the two in all the performance measures. Therefore, we further only cover the POET results and no longer look at the POET-C.

The LW based portfolio also increase massively, it now has a higher Sharpe ratio than the EQ portfolio and lower standard deviations. Diving deeper in the performance measure we see quite a noteworthy difference in standard deviation between the GMV-C and GMV-W portfolio but only a slight difference in Sharpe ratio. This means that the expected return of the GMV-W must be much higher than the GMV-C portfolio in order to offset this difference in standard deviation.

Then looking at the two POET portfolios (GMV-C and GMV-W in Table 3) they show fairly the same performance with a few small differences. Most noteworthy are the slight increase in standard deviation and decrease in turnover for the GMV-W portfolio. This is most likely caused by the difference in lower weight bound. In the GMV-W every assets has a weight bigger than zero. Whereas in the GMV-C portfolio some assets can have a weight of zero. This can be case because the variance of this asset is really high. This big variance is then excluded from the GMV-C portfolio, but has to be included with the GMV-W as every assets has a weight of $\frac{0.5}{N}$.[13] Additionally, this also directly explains the smaller turnover. Because, the weights have a smaller window to move in which results in a more stable weight across the quarters.

Finally, for the GL it is clear that the inclusion of weight restriction did not impact the performance in any significant way. There are also no big differences between the GMV-C and GMV-W performance in combination with the POET technique, but the GMV-W does greatly improve the expected return for LW, with comparable results for the other performance measures. Therefore, we

[13]The diversification effect is really minor in this case as the asset portfolio contains over 300 assets at all time. So any additional variance caused by adding an extra asset is not offset by the diversification effect.

continue by only analyzing the GMV-W portfolio as this give slightly more promising, but still really similar performance as the GMV-C portfolio.

Table 3: Performance measure results GMV-C/W, without circular bootstrap

		EQ	KDM	KDM-C	POET	POET-C	LW	G-Lasso
GMV-C	**St.dev**	0.1009	0.3695	0.3886	1.2912	6.0091	0.0668	0.0991
	Sharpe	0.9308	0.9865	1.0046	0.7748	0.1664	1.0430	0.9282
	CEQ $\gamma = 1$	0.0772	0.1704	0.1779	0.0779	-0.1330	0.0587	0.0754
	CEQ $\gamma = 5$	0.0122	-0.5784	-0.6486	-0.0040	-0.4678	0.0159	0.0114
	CEQ $\gamma = 10$	-0.0677	-2.3930	-2.7702	-0.1082	-1.0281	-0.0364	-0.0675
	Turnover	0.0598	1.23E+05	1.37E+05	0.6711	1.5405	0.3558	0.1233
GMV-W	**St.dev**	0.1009	0.3694	0.3898	1.3965	1.9708	0.0819	0.1004
	Sharpe	0.9308	0.9860	0.9187	0.7163	0.5075	0.9963	0.9296
	CEQ $\gamma = 1$	0.0772	0.1712	0.1525	0.0611	-0.0081	0.0683	0.0765
	CEQ $\gamma = 5$	0.0122	-0.5721	-0.6734	-0.0237	-0.1199	0.0166	0.0120
	CEQ $\gamma = 10$	-0.0677	-2.3652	-2.7941	-0.1353	-0.2772	-0.0465	-0.0676
	Turnover	0.0598	1.24E+05	1.34E+05	0.4879	0.7955	0.2249	0.1226

Note: The table shows the average performance measure for the GMV-C and GMV-W portfolios in combination with each shrinkage technique for the time period 2010-2019. The Average Expected Return (ER), Variance, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures all other measures are quarterly. The KDM-C and POET-C are the same shrinkage technique as KDM and POET, but the eigenvalues and vectors are estimated with the sample correlation matrix instead of the sample covariance.

4.2 Results Regularization Values

The regularization value determines the shrinkage of the covariance matrix as explained in Section 3.4. Finding the optimal regularization value for the Graphical Lasso can therefore be extremely useful. In order to find the optimal value multiple regularization values and matrices are used as stated in Section 3.4. In total 62 different regularization values are used, 51 single values that apply to all assets and 11 regularization matrices. The matrix regularization results are shown in the Appendix Section A.3.

Table 4 shows the GMV portfolio performance measures for the different single regularization values. In the Table 4 $\rho_{1,t}$ is the smallest regularization value, $\rho_{50,t}$ is the largest value corresponding to the complete shrinkage case, and combination is the changing regularization value. As explained in 3.4 it takes the regularization value that resulted in the highest average Sharpe ratio in the previous quarter. For example if in quarter 1 $\rho_{23,1}$ gave the portfolio with the highest average Sharpe ratio then in quarter 2 the combination regularization value would take $\rho_{23,2}$ as its regularization value. Subsequently, when in quarter 2 $\rho_{26,2}$ gives the highest average Sharpe ratio, then in quarter 3 the combination regularization value would use $\rho_{26,3}$ as its regularization value.

There are some clear trends in Table 4, the standard deviation is increasing with higher regularization values, while turnover is clearly decreasing with more restrictive regularization values (higher values for ρ). The CEQ values all follow different trends, for $\gamma = 1$ the CEQ is increasing with higher regularization values, and for $\gamma = 10$ the opposite is true. The CEQ at $\gamma = 5$ follows a complete different path it starts of decreasing and from $\rho_{3,t}$ onward start increasing again. It mimics the trend shown by the Sharpe ratio where the highest value can be found at $\rho_{3,t}$. So the lowest CEQ value coincide with the highest Sharpe ratio.

However, looking at the combination regularization value we see the highest average Sharpe ratio of the whole table, but also the highest CEQ at $\gamma = 5$. This is opposite to the result found in the set regularization values where a higher Sharpe ratio leads to a lower CEQ value at $\gamma = 5$. This can not directly be explained by just way higher values of expected return and standard deviation, because the standard deviation is lower than for most of the other single regularization values. This indicates that the standard devation in the combination case is more volatile with occasionally way higher standard errors. This can cause a substantial increase in the CEQ value. Therefore both $\rho_{3,t}$ and the combination regularization value are used in the further research. As they have the highest average Sharpe ratio, but are still significantly different from each other.

Finally, given the performance of the combination regularization value, there most likely is a correlation between the regularization value that gives the highest average Sharpe ratio in the current quarter with the regularization value that gives the highest average Sharpe ratio in the next quarter.

Table 4: Performance measure of different single regularization values in GMV portfolio

Regularization index	ER	St.dev	Sharpe	CEQ $\gamma = 1$	CEQ $\gamma = 5$	CEQ $\gamma = 10$	Turnover
$\rho_{1,t}$	0.0849	0.1210	0.9363	0.0426	0.0050	-0.0400	0.5357
$\rho_{2,t}$	0.0983	0.1313	0.9572	0.0486	0.0042	-0.0486	0.3713
$\rho_{3,t}$	0.1062	0.1371	**0.9583**	0.0521	0.0036	-0.0539	0.2968
$\rho_{4,t}$	0.1123	0.1409	0.9526	0.0552	0.0039	-0.0569	0.2520
$\rho_{5,t}$	0.1172	0.1435	0.9464	0.0581	0.0046	-0.0585	0.2223
$\rho_{6,t}$	0.1214	0.1455	0.9414	0.0606	0.0055	-0.0594	0.2016
$\vdots$							
$\rho_{15,t}$	0.1388	0.1523	0.9297	0.0709	0.0101	-0.0610	0.1417
$\vdots$							
$\rho_{50,t}$	0.1478	0.1561	0.9295	0.0757	0.0117	-0.0630	0.1240
Combination	0.1386	0.1420	**0.9935**	0.0771	0.0237	-0.0393	0.2802

Note: The table shows the average performance measure for the GMV portfolio in combination with a few of the tested regularization indices/values for the time period 2010-2019. The Regularization indices are ranked from lowest to highest with the 50th regularization index resulting in full shrinkage (diagonal covariance matrix). The Standard deviation, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures all other measures are quarterly.

4.3 Results Portfolio Performances

Following the results in Sections 4.1, 4.2, only the POET, Ledoit and Wolf, and the Graphical Lasso are covered in combination with the Circular Bootstrap, to create GMV and GMV-W portfolios. The GMV-C portfolio is left out, because the performance is so similar to the GMV-W portfolio, as seen in Table 3.

Table 5 shows the average GMV portfolio performances measures with 100 circular bootstraps sequences. As expected, having seen Table 2, the POET technique, in combination with a GMV portoflio, results in unrealistic performance measures, and is consequently unreasonable to use for a real world application with these assets. On the other hand, the two Graphical Lasso based portfolio do perform well. They both have a slightly higher risk adjusted return than the equally weighted portfolio. The GL with the combination regularization value only displays a slightly smaller standard deviation, while the GL with the single regularization value show a substantial lower standard deviation than the equally weighted portfolio.

This result is also captured in the certainty equivalent values. When $\gamma = 1$ we see almost identical CEQ values between EQ and the combination ρ GL, and a lower value for the single ρ. However, when the risk aversion increases we see that the slight better performance of the GL has more of a impact. There are some noticeable difference in CEQ values for the risk aversion levels 5 and 10. The difference between the EQ and combination GL gets bigger for $\gamma = 5$ and $\gamma = 10$ giving the GL portfolio noticeably better CEQ values than the EQ portfolio. [14] However, with single regularization value GL we still see a better CEQ value for the equally weighted portfolio at $\gamma = 5$ and only a slightly better value for the GL when $\gamma = 10$. This can most likely be attributed to the much lower expected return of this GL, and that this difference is so big it can not be offset by the lower standard deviation. To summarize, it can generally be said that risk averse people prefer to invest in the combination ρ GL. However, when the trading cost are incorporate this could change as the turnover of the EQ portfolio is quite a bit lower than all portfolio displayed in Table 5.

Finally the LW technique does display realistic results that could conceivably be recreated in the real world. However, the performance is subpar to the EQ portfolio and the two GL portfolios in terms of Sharpe ratio and turnover, but is equal or better in terms of standard deviation. The strange result is that the CEQ of the LW is better for $\gamma = 5$ and $\gamma = 10$ than the other techniques, but is worse for $\gamma = 1$. This emphasises the fact that lower standard deviation only really have impact on the CEQ at higher values of risk aversion.

[14] Better value is here defined as the highest value. The value indicates the height of risk free return the investor wants to receive to be indifferent between the investment or the risk free return. Meaning higher CEQ indicates a better investment opportunity. Negative CEQ means the amount the investor want to pay in order to be indifferent between investing or not.

Table 5: Performance measure results GMV, with circular bootstrap

		EQ	POET	LW	G-Lasso ρ	G-Lasso ρ Combination
GMV	**St.dev**	0.1010	6.10E+06	0.0671	0.0674	0.0914
	Sharpe	0.9308	0.0381	0.7506	0.9583	0.9935
	CEQ $\gamma = 1$	0.0772	$-\infty$	0.0468	0.0521	0.0771
	CEQ $\gamma = 5$	0.0122	$-\infty$	0.0326	0.0036	0.0237
	CEQ $\gamma = 10$	-0.0677	$-\infty$	0.0151	-0.0570	-0.0409
	Turnover	0.0598	9.56E+03	2.3460	0.3115	0.2757

Note: The table shows the average performance measure for the GMV portfolio in combination with 4 different shrinkage technique and the Equally Weighted portfolio for the time period 2010-2019. G-Lasso ρ and combination are the Graphical Lasso results with a single regularization value. The average Standard deviation, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures. The Turnover is the quarterly Turnover.

Comparing Table 5 with Table 6 it directly jumps out that the POET and, the Ledoit and Wolf technique perform much better in combination with the GMV-W portfolio. The POET technique now shows reasonable and even good performances for all measures. The Ledoit and Wolf method does show a higher standard deviation, but also a way higher Sharpe ratio. It is even beating the EQ portfolio with a higher Sharpe ratio and a lower standard deviation. The turnover is also substantially lower compared to the GMV portfolio in Table 5. Making the LW shrinkage technique in combination with the GMV-W portfolio a worthy option to consider for real world use. Whereas the POET does perform considerably better than in Table 5, but still shows a higher standard deviation resulting in an overall lower Sharpe ratio. Furthermore, the turnover is also considerably larger than the other techniques. Suggesting that the POET is still not a sensible alternative to the EQ portfolio or the other portfolios.

Looking at the Graphical Lasso performances we see a higher standard deviation for both, which in turn results a in lower Sharpe ratio for the combination regularization and an equivalent value for the single regularization value compared to Table 5. Both the Sharpe ratios are now even lower than the LW based portfolio. However, they have a lower turnover than LW. Next to that, they still have a higher Sharpe ratio than the EQ portfolio and have lower turnovers than their own counterpart in the GMV portfolios. In addition, the CEQ values are a bit higher for $\gamma = 1$, but the values are lower for $\gamma = 5$ and $\gamma = 10$. This behaviour coincide with the higher standard deviation found in the GMV-W that in turn are less important at a risk level of $\gamma = 1$, but do influence the CEQ values more when the risk level increase.

Finally, comparing all performances in Table 6, the LW technique has the best risk adjusted returns, but for low levels of risk aversion the combination ρ GL is preferred over the LW technique. Looking at higher risk aversion levels the LW technique is the preferred choice. However, the difference in performance between all the techniques is pretty small. The EQ portfolio does differentiate it self by having substantially lower turnover values. Meaning, that when the trading cost are included the higher Sharpe ratio of the other techniques could be offset by the extra cost of the higher turnovers.

Table 6: Performance measure results GMV-W, with circular bootstrap

		EQ	**POET**	**LW**	**G-Lasso ρ**	**G-Lasso ρ Combination**
GMV-W	**St.dev**	0.1010	0.1295	0.0831	0.0893	0.0993
	Sharpe	0.9308	0.8127	0.9944	0.9541	0.9543
	CEQ $\gamma = 1$	0.0772	0.0806	0.0692	0.0700	0.0789
	CEQ $\gamma = 5$	0.0122	-0.0236	0.0168	0.0112	0.0176
	CEQ $\gamma = 10$	-0.0677	-0.1609	-0.0471	-0.0613	-0.0573
	Turnover	0.0598	0.4251	0.2376	0.1840	0.1616

Note: The table shows the average performance measure for the GMV-W portfolio in combination with 6 different shrinkage technique and the Equally Weighted portfolio for the time period 2010-2019. G-Lasso ρ and combination are the Graphical Lasso results with a single regularization value. The G-Lasso matrix and combination, are the Graphical Lasso results with a regularization matrix. The average Expected Return (ER), Variance, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures.

In Appendix Section A.4 the full Table of 5, and 6 are shown including the matrix regularization results of the Graphical Lasso. The matrix regularization results are similar to the single regularization results discussed before. With the combination matrix showing slightly higher Sharpe ratios due to smaller standard deviation. However, with the significantly longer running time to create the combination regularization matrix, the small increase might not be enough for portfolio managers to justify the the longer running time .

4.4 Results Robustness

In Section 4.3 we have seen that the LW and GL techniques perform quite similar with the performance of the POET technique lacking a bit behind. Leaving the question which technique is more robust. In Table 7 the minimal, average and maximum standard deviation between the performance measures are shown. The standard deviation is taken from the 100 simulation within one quarter. So the minimal standard deviation in the Sharpe ratio can be formulated as:

$$\min\left\{\text{Std}\left(SR_{ij}\right)\right\} = \min\left\{\sigma_{SR_{ij,t}}\right\}. \tag{30}$$

The standard deviation in the Sharpe ratio between the 100 simulation is defined as σ_{SR}, the indicators i, and j respectively represent the shrinkage technique, and weighting scheme, and t indicates to which quarter the Sharpe ratio belongs. This set up is the same for the standard deviation of the standard deviation (St.dev σ), the standard deviation of the Turnover and the standard deviations of the Weights shown in Table 7. In the final row of Table 7 displays the biggest percentage difference in weight, given to the same asset over the 100 bootstrap sequences, in one quarter. Mathematically this is defined followingly:

$$\Delta W_{ij,t} = \frac{1}{N}\sum_{l=1}^{N}\frac{\max w_{ij,l,t} - \min w_{ij,l,t}}{\min w_{ij,l,t}}. \tag{31}$$

In the formula $\Delta\ W_{ij,t}$ is the biggest percentage difference in weights for quarter t, with shrinkage method i and weighting scheme j. The parameter max $w_{ij,l,t}$ is the maximum weight given to asset l in quarter t for shrinkage method i and weighting scheme j. So $\Delta\ W_{ij,t}$ can also be defined as the average, of biggest percentage weight difference over all assets in quarter t. There are 40 quarters so $W_{ij,t}$ contains 40 values. From $W_{ij,t}$ the minimal, average and max value are shown in Table 7.

In Figure 1 the standard deviation of the Sharpe ratios between the 100 bootstrap sequences of each of the 40 quarters are displayed. The Sharpe ratios come from the Graphical Lasso with single regularization value in combination with a GMV portfolio (G-Lasso ρ in table 7). In the figure the average standard deviation is shown by the red vertical line. From the figure it is clear that the standard deviations in the Sharpe ratio a generally small, but do display some bigger values in the upper tail. The interesting values from the distribution in Figure 1 are the minimal standard deviation (0.0275), the average (0.1264) and the maximal standard deviation (0.4032) in Sharpe ratio. Therefore, instead of plotting distributions for all the performance measures in combination with every shrinkage technique and portfolio, the minimal, average and maximal standard deviation of all performance measures are displayed in Tables 7 and 8. Table 7 shows the standard deviations in the GMV portfolios and Table 8 in the GMV-W portfolios.

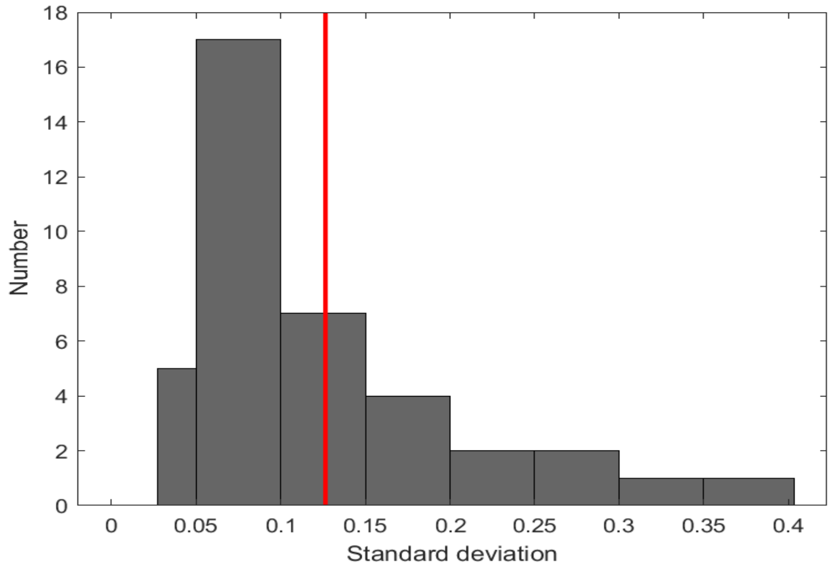

Figure 1: Standard deviations of the Sharpe ratios in the 40 quarters, G-Lasso ρ GMV

Note: The histogram displays the standard deviation in the Sharpe ratio over the 100 bootstrap sequences in one quarter, for all 40 quarters. The Sharpe ratios are from the Graphical Lasso with a single regularization value in combination with a GMV portfolio for the period 2010 to 2019. The red vertical line indicates the average standard deviation of the 40 quarters. The average is 0.1264.

Looking at the standard deviation in σ (σ is the standard deviation of the portfolio displayed as St.dev(σ) in the table) in Table 7, it is clear that the POET has way bigger differences in his σ then the other three techniques. The lowest σ is 1.3349, that is about 20 to 50 times bigger than the maximal σ of the other three. Meaning that the σ of the portfolio is tricky to predict as small changes in the data set result in really big difference in σ. This was to be expected knowing the previous results of the POET in combination with a GMV portfolio.

The standard deviation in σ for the LW technique are fairly low for the minimal and average values, but the maximal value can be a bit daunting. The maximal standard deviation in σ is 0.0697, while the average σ in the GMV portfolio is 0.0671 as shown in Table 5. In the worst situation there can be standard deviation in the calculated σ of more than the average value of σ. The Graphical Lasso performs better is that regard,but the maximal values are still substantial at about 31% and 42% of their respective average σ in Table 5. However, the mean of standard deviations in σ are approximately 6.7% and 3.7% of the average values shown in Table 5, which is a pretty small difference between the 100 simulation.

The trend of reasonable or good values for the min and mean, but slightly high values for max standard deviations, continues throughout the table for all performance measures. Comparing the LW technique to the GL, it is clear that the GL techniques performs better looking at the standard deviations of the Sharpe ratio, weight, and ΔW_{ij}. The strange thing however, is that the LW technique shows a substantially lower standard deviation in the turnover. This is a strange observation because ΔW_{ij} shows there are some really big weight differences within one quarter for the LW technique, and the ΔW_{ij} of both the Graphical Lasso techniques are lower than the LW technique. One explanation could be that the differences between the turnover is indeed bigger for the LW technique. However, because the turnover itself is also much larger than the GL technique, as seen in Table 6, that the standard deviation turns out lower for the LW technique. Because the baseline measured on is so much higher.

Finally the standard deviation results found for σ are also generally true for all the other performance measures. Therefore, as expected, the standard deviation in the performance measure indicate that the POET and LW struggle to perform consistently over the 100 bootstrap sequence and the Graphical Lasso performs quite stable in combination with a GMV portfolio. However, this is the first time we see that the Graphical Lasso is not impeccable in combination with the GMV portfolio, as it does show quite high values for the maximal standard deviations in the performance measures especially the percentage weight difference ΔW_{ij}.

Table 7: Standard deviation within the one quarter for the GMV portoflios

GMV		**POET**	**LW**	**G-Lasso ρ**	**G-Lasso ρ Combination**
	Min	1.3349	2.50E-03	9.06E-04	1.07E-04
St.dev (σ)	**Mean**	1.07E+02	0.0085	0.0060	0.0037
	Max	2.91E+03	0.0697	0.0279	0.0413
	Min	1.6298	0.2214	0.0275	0.0036
Sharpe	**Mean**	1.9790	0.6057	0.1264	0.0951
	Max	2.6393	1.2837	0.4032	0.4311
	Min	1.09E+02	0.0024	0.0060	0.0003
Turnover	**Mean**	9.02E+03	0.0205	0.0539	0.0355
	Max	1.91E+05	0.0692	0.1757	0.3262
	Min	0.6960	1.56E-03	1.38E-04	1.08E-05
Weight	**Mean**	21.7695	3.62E-03	4.04E-04	2.85E-04
	Max	4.62E+02	5.62E-03	1.24E-03	1.90E-03
Biggest percentage	**Min**	0.991183	0.6435	0.2095	0.1388
difference in weights	**Mean**	3.74E+09	23.2912	5.6360	2.1056
in 1 quarter $\Delta\mathbf{W}_{ij}$	**Max**	9.52E+11	6.51E+03	1.14E+03	1.91E+02

Note: The table shows standard deviation of the performance measure within one quarter for the GMV portfolio. The min displays the lowest standard deviation between the 100 bootstrap sequences found within one quarter, across the 40 quarters. The mean displays the average standard deviation between the 100 bootstrap sequences found across the 40 quarters. Finally max displays the maximum standard deviation between the 100 bootstrap sequences within one quarter found across the 40 quarter. G-Lasso ρ is the single set regularization value found to perform the best overall. G-Lasso ρ combination, is the regularization value that consist of the regularization values that where optimal for previous quarter. Finally the biggest percentage difference in weights displays the average percentage difference between the largest and smallest weight of each asset over the 100 bootstrap data sequences. The data set used is from 2010 to 2019.

In table 8 the standard deviation in the performance measures of the GMV-W portfolios are shown. Immediately we see a big improvement for the POET technique compared to the GMV portfolio values in Table 7. The most promising difference between Table 7, and 8 is that the values for ΔW_{ij} are substantially lower. This directly indicates that the POET standard deviations in the performance measures will be lower. This is in fact also the case, but more importantly it is true for the maximal standard deviation values. The maximal values take a hard drop compared to Table 7. The maximal standard deviation in the performance are now slightly bigger than the average standard deviation values displayed in Table 7. This means that using GMV-W weighting scheme results in way robuster estimation of the portfolio weights for the POET.

The min, and mean values of the LW and GL technqieus where already quite good in combination with a GMV portfolio in Table 7. Combining them with a GMV-W portfolio the standard deviations become even lower as seen in Table 8. More importantly the GMV-W portfolio resolves the big maximum standard deviations that are present in the GMV portfolio. The max values in Table 8 are substantially lower than in Table 7. Indicating that only using the shrinkage technique of the LW or the GL is not enough to get rid of big values in the tails. Using the double shrinkage, by also shrinking the weights, is enough to get rid of the big extreme values.

The robustness of the POET is nowhere near the stability of the other techniques, even though it shows a big improvement with the GMV-W portfolio. The POET still displays worse standard deviations in the performance measures than the LW does with the GMV portfolio in Table 7. So consequently, it is easy to conclude that the POET still struggles to give a robust and stable portfolio.

Comparing the GMV-W standard deviation of the LW and Graphical Lasso techniques in Table 8 to each other. We see that both the Graphical Lasso techniques show lower values for all min, mean and max standard deviations compared to the LW technique. Where there is not much difference between the two Graphical Lasso techniques themselves, but the combination regularization values does give slightly lower standard deviations.

Concluding, the LW and Graphical Lasso can create inverse covariance matrices that result in acceptable or even good average standard errors in the performance measures, for similarly distributed data sets. However, they do suffer from some extreme values in the tails. By incorporating a double shrinkage, like weight restrictions, this issue is mostly resolved. Meaning LW and GL do benefit from the inclusion of the second shrinkage. With this double shrinkage both the LW and GL are robust and stable in creating the portfolio weights. Where, the GL techniques are slightly better than the LW, and the combination regularization value tops them all.

In Appendix Section A.5 the Graphical Lasso with a regulation matrix is included in the table together with the standard deviation of the certainty equivalent. The standard deviation in the CEQ performance measure follows the general results of the other performance measures across the shrinkage techniques, and both the Graphical Lasso with the regularization matrices follow the results of their single regularization value counterpart, but with slightly worse results.

Table 8: Standard deviation within the one quarter for the GMV-W portoflios

GMV-W		POET	LW	G-Lasso ρ	G-Lasso ρ Combination
	Min	0.0154	5.94E-04	1.38E-04	5.24E-07
St.dev	**Mean**	0.0702	0.0023	0.0025	0.0015
	Max	0.2788	0.0084	0.0076	0.0092
	Min	0.1597	8.71E-03	5.89E-03	4.81E-05
Sharpe	**Mean**	0.6719	0.0797	0.0422	0.0244
	Max	1.4394	0.2535	0.1900	0.1229
	Min	0.1097	2.78E-03	8.22E-04	5.82E-06
Turnover	**Mean**	0.1565	0.0208	0.0185	0.0104
	Max	0.2367	0.0692	0.0697	0.0616
	Min	0.0012	8.54E-05	1.14E-05	8.48E-08
Weight	**Mean**	0.0017	2.57E-04	1.44E-04	8.35E-05
	Max	0.0023	5.45E-04	5.27E-04	5.38E-04
Biggest percentage	**Min**	1.2073	0.1055	0.0771	0.0531
difference in weights	**Mean**	8.9188	0.6451	0.3525	0.1978
in 1 quarter	**Max**	3.31E+02	6.9469	4.2913	2.1426

Note: The table shows standard deviation of the performance measure within one quarter for the GMV-W portfolio. The min displays the lowest standard deviation between the 100 bootstrap sequences found within one quarter, across the 40 quarters. The mean displays the average standard deviation between the 100 bootstrap sequences found across the 40 quarters. Finally max displays the maximum standard deviation between the 100 bootstrap sequences within one quarter found across the 40 quarter. G-Lasso ρ is the single set regularization value found to perform the best overall. G-Lasso ρ combination, is the regularization value that consist of the regularization values that where optimal for previous quarter and is therefore not consistently the same regularization value across quarters. Finally the biggest percentage difference in weights displays the average percentage difference between the largest and smallest weight of each asset over the 100 bootstrap data sequences. The data set used is from 2010 to 2019.

4.5 Results Summary

Summarizing the results of Section 4.3, and 4.4. We have seen that the Graphical Lasso technique is the only analyzed technique that could be a reasonable choice if the investor wants to create a Global Minimum Variance portfolio. As only the GL technique displays realistic performance measure that can compare to the performance of the equally weighted portfolio. This is true for all regularization values considered. However, looking at the robustness results we see that in the tails, the standard deviations in the performance measure can become quite large. This is something to keep in mind when creating GMV portfolios with the use of the Graphical Lasso.

Further, looking at the weight restricted Global Minimum Variance (GMV-W) portfolios we see that the performance of the POET technique increases massively and has reasonable performance measures, but is still substantially worse than the equally weighted portfolio or the other techniques.

The Ledoit and Wolf technique results in reasonable or even good performance measures in combination with the GMV-W portfolio. It even displays the highest Sharpe ratio and the lowest standard deviations of all the analyzed techniques, including the equally weighted portfolio. The performance measures of the GL technique does not change much in combination with a GMV-W portfolio compared to the GMV portfolio. Except for the turnover, the turnover sees a substantial drop in value when compared to the GMV portfolio turnovers. This drop in turnover is present for all techniques and is a direct result of the weight restrictions.

The double shrinkage of the GMV-W portfolio, by adding the weight restriction, also resolves the high maximal values of the standard deviations in the performance measures found across the 40 quarters. Where for the GMV portfolio we state that these large standard deviations in the tails is something to keep in mind when creating the portfolio, is this not an issue with the GMV-W portfolio for neither the LW nor the Graphical Lasso technique.

5 Conclusion

In the paper we have analysed the out-of-sample performance of portfolios constructed by high-dimensional inverse covariance matrices, created with a shrinkage technique, and we also evaluated the robustness of the shrinkage technique. The portfolios are constructed with the daily returns of approximately 500 stock from the New York Stock Exchange. One quarter of daily returns is used in combination with five different shrinkage techniques to create five high-dimensional inverse covariance matrices. The matrices are in turn used to construct multiple portfolios. The portfolio results are analyzed with different performance measures. In order to analyze the robustness of the inverse covariance matrix estimation, a bootstrap method is used to create multiple matrices based on the same data set. All the newly created matrices are then used to construct portfolios resulting in 100 portfolio in each quarter for the same shrinkage technique. The robustness in then analyzed with use of the standard errors between the performance measures of these 100 portfolios.

The results do not indicate that the KDM technique will create robust inverse covariance matrices, that in turn creates a portfolio that can compete with the equally weighted portfolio. The Graphical Lasso shrinkage technique is the only investigated technique that performs well enough, in combination with a Global Minimum Variance portfolio, to beat or compare to the equally weighted portfolio. However, in the extremes their are some large instabilities in the estimation of the inverse covariance, but on average it is quite robust.

With the GMV-W portfolio the POET, Ledoit and Wolf and Graphical Lasso shrinkage techniques all show better performance measures than the equally weighted portfolio. However, the POET is lacking behind a bit in performance compared to the LW and GL techniques. When looking at the robustness of the techniques it becomes clear that the POET is still not advisable to use. The Ledoit and Wolf and Graphical Lasso techniques show really low standard deviations in there performance measure across the 100 bootstrap sequences. Wherein there is a slight advantage for the Graphical Lasso. Looking at the estimated performance measures there is not much difference between the LW and GL in combination with the GMV-W portfolio. In the end the Graphical Lasso has in our opinion the edge over all the other analyzed shrinkage techniques. As it has the best performance and is at the same time also the most robust. However, the turnover of all the shrinkage techniques are substantially higher than the equally weighted portfolio. Meaning that when the trading cost are incorporated the equally weighted portfolio can still come on top to beat them all.

For further research one can try different shrinkage targets for the Ledoit and Wolf technique and they could try to optimize the regularization values of the Graphical Lasso even further. The paper itself can be extend by a replication of the Jobson and Korkie (1980) paper. In order to see if using a shrinkage technique improves the ability to recreate the mean-variance frontier and if this coincide with our findings of the better performing shrinkage technique. In order to find even better performing portfolios one could also try to implement Financial Factor information. Further research could also focus on using a different technique to the construct the eigenvalue and vectors beside the sample covariance and correlation matrix. Because, of the high-dimensional matrices it is reasonable to assume that the estimations errors of the covariance or correlation matrix are quite big. This then gets aggravated by estimating the eigenvalues and vectors based on these matrices with big estimation errors. This is most likely one of the reasons why the POET performance so bad in the GMV portfolio and why the KDM performs bad overall in this data set.

References

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Banerjee, O., Ghaoui, L. E., and d'Aspremont, A. (2008). Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data. *Journal of Machine learning research*, 9(Mar):485–516.

Cai, T. and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association*, 106(494):672–684.

Chib, S., Nardari, F., and Shephard, N. (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 134(2):341–371.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, 22(5):1915–1953.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer.

Fan, J., Liao, Y., and Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(4):603–680.

Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3):432–441.

Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a garch (1, 1)? *Journal of applied econometrics*, 20(7):873–889.

Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4):1651–1683.

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

Kosowski, R., Naik, N. Y., and Teo, M. (2007). Do hedge funds deliver alpha? a bayesian and bootstrap analysis. *Journal of Financial Economics*, 84(1):229–264.

Kosowski, R., Timmermann, A., Wermers, R., and White, H. (2006). Can mutual fund "stars" really pick stocks? new evidence from a bootstrap analysis. *The Journal of finance*, 61(6):2551–2595.

Kourtis, A., Dotsis, G., and Markellos, R. N. (2012). Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix. *Journal of Banking & Finance*, 36(9):2522–2531.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5):603–621.

Ledoit, O. and Wolf, M. (2004). Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110–119.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

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

Markowitz, H. M. Portfolio selection. *The Journal of Finance*.

Meinshausen, N., Bühlmann, P., et al. (2006). High-dimensional graphs and variable selection with the lasso. *The annals of statistics*, 34(3):1436–1462.

Michaud, R. O. (1989). The markowitz optimization enigma: Is 'optimized'optimal? *Financial Analysts Journal*, 45(1):31–42.

Politis, D. N. and Romano, J. P. (1992). A circular block-resampling procedure for stationary data. *Exploring the limits of bootstrap*, 2635270.

Ruiz, E. and Pascual, L. (2002). Bootstrapping financial time series. *Journal of Economic Surveys*, 16(3):271–300.

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business*, 39(1):119–138.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

A Appendix

A.1 KDM Weight Construction

The weights construction for the inverse covariance shrinkage technique of Kourtis Dotsis and Markellos is shown below for both the single target matrix and the double target matrix.

$$\hat{\Theta}_{KDM} = c_1 S^{-1} + c_2 \hat{\Omega}, \tag{32}$$

$$\hat{\Theta}_{KDM} = c_1 S^{-1} + c_2 I + c_3 \hat{F}. \tag{33}$$

The weights of Equation 32 and 33 are computed in the following way:

$$w_{KDM} = \frac{c_1 S^{-1}\iota + c_2 \hat{\Omega}\iota}{c_1 \iota' S^{-1}\iota + c_2 \iota' \hat{\Omega}\iota}, \quad \text{or} \quad w_{KDM} = \frac{c_1 S^{-1}\iota + c_2 I\iota + c_3 \hat{F}\iota}{c_1 \iota' S^{-1}\iota + c_3 \iota' I\iota + c_3 \iota' \hat{F}\iota}, \tag{34}$$

in which ι is a vector $N \times 1$ of ones and w_{KDM} the resulting weights. This can be written as:

$$w_{KDM} = \hat{d}_0 w_{GMV} + (1 - \hat{d}_0) w_{\hat{\Omega}}, \quad \text{or} \quad w_{KDM} = \hat{d}_1 w_{GMV} + \hat{d}_2 w_I + (1 - \hat{d}_1 - \hat{d}_2) w_F. \tag{35}$$

Where w_{GMV} are the global minimum variance weights based on S^{-1}. Subsequently, w_Ω, w_I and w_F are respectively the GMV weights based on Ω, I and F. The great advantage of using equation 35 is the fact that it enables us to estimate one variable less than in equation 34. Therefore, only d_0, or d_1 and d_2 need to be estimated in the following way:

$$\begin{aligned} \hat{d}_0 &= \frac{\text{var}(R_t^\Omega) - \text{cov}(R_t^{GMV}, R_t^\Omega)}{\text{var}(R_t^{GMV}) - 2\text{cov}(R_t^{GMV}, R_t^\Omega) + \text{var}(R_t^\Omega)}, \\ \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix} &= \begin{pmatrix} v_{11} + v_{33} - 2v_{13} & v_{12} - v_{13} - v_{23} + v_{33} \\ v_{12} - v_{13} - v_{23} + v_{33} & v_{22} + v_{33} - 2v_{23} \end{pmatrix}^{-1} \begin{pmatrix} v_{33} - v_{13} \\ v_{33} - v_{23} \end{pmatrix}. \end{aligned} \tag{36}$$

Where R_t^i is given as the sample return in quarter t, times the weights based on GMV or the target matrix: $R_t^i = R_t w_i$. Further v_{ij} are the respective components in the following covariance matrix:

$$\text{V} = \begin{pmatrix} \text{var}(R_t^{GMV}) & \text{cov}(R_t^{GMV}, R_t^I) & \text{cov}(R_t^{GMV}, R_t^{\hat{F}}) \\ \text{cov}(R_t^{GMV}, R_t^I) & \text{var}(R_t^I) & \text{cov}(R_t^I, R_t^{\hat{F}}) \\ \text{cov}(R_t^{GMV}, R_t^{\hat{F}}) & \text{cov}(R_t^I, R_t^{\hat{F}}) & \text{var}(R_t^{\hat{F}}) \end{pmatrix}. \tag{37}$$

A.2 Results Yearly Data Set

The tables below show the performance measures when a full year of asset returns is used in combination with the the different shrinkage techniques for the GMV, GMV-C and GMV-W portfolio. It is clear that the KDM and POET technique benefit most of the bigger input data sequence, as they displays more reasonable performance measures than when only one quarter of asset returns is used.

Table 9: Performance measure results GMV, without circular bootstrap and with yearly input data

	EQ	KDM	KDM-C	POET	POET-C	LW	G-Lasso
	St.dev	7.4509	0.7185	0.0572	-0.2226	0.0973	0.1297
	Sharpe	0.4892	0.1011	0.6650	-0.3812	0.3910	0.6542
GMV	**CEQ** $\gamma = 1$	-3.26E+02	-0.6009	-5.50E+03	-1.24E+16	0.0311	0.0697
	CEQ $\gamma = 5$	-2.53E+14	-1.28E+01	-1.67E+21	-1.57E+66	0.0034	0.0100
	CEQ $\gamma = 10$	-4.45E+27	-1.96E+02	-5.73E+39	-3.54E+107	-0.0313	-0.0639
	Turnover	5.50E+03	1.89E+03	4.88E+04	7.50E+03	86.2346	3.8887

Note: The table shows the average performance measure for the GMV portfolio in combination with each shrinkage technique, and the Equally Weighted portfolio, for the time period 2010-2019. Wherein the complete year is used as data set. The average Expected Return (ER), Variance, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures. The KDM-C and POET-C are the same shrinkage technique as KDM and POET, but the eigenvalues and vectors are estimated with the sample correlation matrix instead of the sample covariance.

Table 10: Performance measure results GMV-C, without circular bootstrap and with yearly input data

	EQ	KDM	KDM-C	POET	POET-C	LW	G-Lasso
	St.dev	0.3149	0.3423	0.2002	0.1428	0.1034	0.1304
	Sharpe	0.2244	0.2234	0.5655	0.2189	0.8430	0.6551
GMV-C	**CEQ** $\gamma = 1$	-0.6054	-0.5878	0.0837	0.0036	0.0777	0.0702
	CEQ $\gamma = 5$	-12.8528	-12.4013	-0.0294	-0.1113	0.0392	0.0098
	CEQ $\gamma = 10$	-1.99E+02	-1.87E+02	-0.1801	-0.2737	-0.0071	-0.0647
	Turnover	1.80E+03	1.89E+03	13.1735	23.7896	6.2754	3.9019

Note: The table shows the average performance measure for the GMV-C portfolio in combination with each shrinkage technique, and the Equally Weighted portfolio, for the time period 2010-2019. Wherein the complete year is used as data set. The average Expected Return (ER), Variance, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures. The KDM-C and POET-C are the same shrinkage technique as KDM and POET, but the eigenvalues and vectors are estimated with the sample correlation matrix instead of the sample covariance.

Table 11: Performance measure results GMV-W, without circular bootstrap and with yearly input data

	EQ	KDM	KDM-C	POET	POET-C	LW	G-Lasso
	St.dev	0.3181	0.3623	0.1316	0.1252	0.1158	0.1312
	Sharpe	0.2297	0.1802	0.6755	0.6276	0.7418	0.6563
GMV-W	**CEQ** $\gamma = 1$	-0.6030	-0.6195	0.0701	0.0604	0.0739	0.0707
	CEQ $\gamma = 5$	-12.8724	-13.1827	-0.0029	-0.0106	0.0257	0.0099
	CEQ $\gamma = 10$	-2.00E+02	-2.07E+02	-0.0956	-0.1020	-0.0325	-0.0653
	Turnover	1.80E+03	1.79E+03	8.2585	13.8560	4.8863	3.9253

Note: The table shows the average performance measure for the GMV-W portfolio in combination with each shrinkage technique, and the Equally Weighted portfolio, for the time period 2010-2019. Wherein the complete year is used as data set. The average Expected Return (ER), Variance, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures. The KDM-C and POET-C are the same shrinkage technique as KDM and POET, but the eigenvalues and vectors are estimated with the sample correlation matrix instead of the sample covariance.

A.3 Results Matrix Regularization Values

In Table 12 the Graphical Lasso with a regularization matrix performance measures are shown. The first interesting fact shown in the table is that the Sharpe ratio of with the matrix regularization values are significantly higher than the Sharpe ratio shown in Table 4. Further the expected returns starts off by decreasing with the stricter regularization values, and only start to increase again after $\rho_{5,t}$. However the most intriguing observation is the huge difference between $\rho_{9,t}$, and $\rho_{10,t}$, standard deviation takes a big jump up. While the Sharpe and turnover decreases significantly. This indicates that the full shrinkage case results in a substantially different portfolio than when the covariance matrix is almost a diagonal matrix. Looking at the performance measures and most predominately at the Sharpe ratio there is most likely still some room for small improvements. As the highest average Sharpe is found at the $\rho_{9,t}$ and the increase in Sharpe ratio even seems to grow in size when the regularization values approaches the full shrinkage value. So creating some small steps between $\rho_{9,t}$ and the full shrinkage at $\rho_{10,t}$ could results in slightly higher Sharpe ratios. Finally, the combination regularization matrix did not come on top this time when looking at the average Sharpe ratio, as was the case in the combination of the single regularization values. The combination matrix still performed pretty well compared to the other regularization matrices. It has almost identical turnovers as $\rho_{9,t}$, but does have a higher standard deviation which resulted in a small decrease in Sharpe ratio. Looking at the other values in table it is safe to say that the combination path mostly consist of the regularization value of the complete shrinkage cases($\rho_{10,t}$), and $\rho_{9,t}$. As all the performance measures are really similar to each other we further only look at $\rho_{9,t}$ and the combination index. Because these are the two with the highest average Sharpe ratio and are still different looking at the standard deviation.

Table 12: Performance measure of different regularization matrix values in GMV portfolio

Regularization index	St.dev	Sharpe	CEQ $\gamma = 1$	CEQ $\gamma = 5$	CEQ $\gamma = 10$	Turnover
$\rho_{1,t}$	0.0753	1.0753	0.0536	0.0385	0.0200	2.0818
$\rho_{2,t}$	0.0752	1.0777	0.0531	0.0381	0.0197	1.9770
$\rho_{3,t}$	0.0751	1.0778	0.0526	0.0376	0.0193	1.8754
$\rho_{4,t}$	0.0752	1.0805	0.0522	0.0373	0.0189	1.7720
$\rho_{5,t}$	0.0755	1.0870	0.0521	0.0370	0.0185	1.6624
$\rho_{6,t}$	0.0761	1.0980	0.0523	0.0369	0.0181	1.5410
$\rho_{7,t}$	0.0772	1.1146	0.0528	0.0369	0.0174	1.3993
$\rho_{8,t}$	0.0795	1.1376	0.0539	0.0369	0.0161	1.2199
$\rho_{9,t}$	0.0854	**1.1633**	0.0566	0.0367	0.0124	0.9585
$\rho_{10,t}$	0.1179	1.0757	0.0725	0.0343	-0.0115	0.4097
Combination	0.1010	**1.1414**	0.0742	0.0458	0.0113	0.9543

Note: The table shows the average performance measure for the GMV portfolio in combination with the tested regularization matrices for the time period 2010-2019. The Regularization indices are ranked from lowest to highest with the 10th regularization index resulting in full shrinkage (diagonal covariance matrix). The Average Expected Return (ER), Variance, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures all other measures are quarterly.

A.4 All Portfolio Performance Results

The following three tables show the performance measures of the different shrinkage technique in combination with the GMV, and GMV-W portfolios. The table also include the Graphical Lasso with a matrix regularization value unlike the tables in the main text.

Table 13: Performance measure results GMV, with circular bootstrap

		EQ	POET	LW	G-Lasso ρ	G-Lasso ρ Combination	G-Lasso Matrix	G-Lasso Matrix Combination
GMV	**St.dev**	0.0102	6.10E+06	0.0.671	0.0674	0.0914	0.0529	0.0713
	Sharpe	0.9308	0.0381	0.7506	0.9583	0.9935	1.1633	1.1414
	CEQ $\gamma = 1$	0.0772	$-\infty$	0.0468	0.0521	0.0771	0.0566	0.0742
	CEQ $\gamma = 5$	0.0122	$-\infty$	0.0326	0.0036	0.0237	0.0367	0.0458
	CEQ $\gamma = 10$	-0.0677	$-\infty$	0.0151	-0.0570	-0.0409	0.0124	0.0113
	Turnover	0.0598	9.56E+03	2.3460	0.3115	0.2757	0.9594	0.9381

Note: The table shows the average performance measure for the GMV portfolio in combination with 6 different shrinkage technique and the Equally Weighted portfolio for the time period 2010-2019. G-Lasso ρ and combination are the Graphical Lasso results with a single regularization value. The G-Lasso matrix and combination, are the Graphical Lasso results with a regularization matrix. The average Expected Return (ER), Variance, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures.

Table 14: Performance measure results GMV-W, with circular bootstrap

		EQ	POET	LW	G-Lasso ρ	G-Lasso ρ Combination	G-Lasso Matrix	G-Lasso Matrix Combination
GMV-W	**St.dev**	0.0102	0.1295	0.0831	0.0893	0.0993	0.0940	0.0923
	Sharpe	0.9308	0.8127	0.9944	0.9541	0.9543	0.9432	1.0167
	CEQ $\gamma = 1$	0.0772	0.0806	0.0692	0.0700	0.0789	0.0729	0.0794
	CEQ $\gamma = 5$	0.0122	-0.0236	0.0168	0.0112	0.0176	0.0121	0.0236
	CEQ $\gamma = 10$	-0.0677	-0.1609	-0.0471	-0.0613	-0.0573	-0.0627	-0.0441
	Turnover	0.0598	0.4251	0.2376	0.1840	0.1616	0.1362	0.2217

Note: The table shows the average performance measure for the GMV-W portfolio in combination with 6 different shrinkage technique and the Equally Weighted portfolio for the time period 2010-2019. G-Lasso ρ and combination are the Graphical Lasso results with a single regularization value. The G-Lasso matrix and combination, are the Graphical Lasso results with a regularization matrix. The average Expected Return (ER), Variance, Sharpe ratio, and Certainty Equivalent are scaled to yearly performance measures.

A.5 All Robustness Results

The following 2 tables show the standard deviations in all performance measures for all tested shrinkage techniques in combination with the GMV and GMV-W portfolio.

Table 15: Standard deviation within the one quarter for the GMV portoflios

GMV		**POET**	**LW**	**G-Lasso ρ**	**G-Lasso** Combination	**G-Lasso ρMatrix**	**G-Lasso Matrix Combination**
	Min	1.3349	2.50E-03	9.06E-04	1.07E-04	1.29E-03	7.16E-04
St.dev	**Mean**	1.07E+02	0.0085	0.0060	0.0037	0.0052	0.0045
	Max	2.91E+03	0.0697	0.0279	0.0413	0.0293	0.0586
	Min	1.6298	0.2214	0.0275	0.0036	0.1009	0.0298
Sharpe	**Mean**	1.9790	0.6057	0.1264	0.0951	0.2958	0.2315
	Max	2.6393	1.2837	0.4032	0.4311	0.5967	0.7707
	Min	8.5432	0.0203	0.0036	0.0003	0.0119	0.0050
CEQ $\gamma = 1$	**Mean**	∞	0.0502	0.0292	0.0207	0.0274	0.0224
	Max	∞	0.1090	0.1133	0.3108	0.0730	0.0611
	Min	1.56E+77	0.0198	0.0033	0.0003	0.0120	0.0052
CEQ $\gamma = 5$	**Mean**	∞	0.0494	0.0323	0.0264	0.0281	0.0227
	Max	∞	0.1073	0.2037	0.5077	0.0773	0.0796
	Min	3.24E+121	0.0193	0.0030	0.0003	0.0123	0.0054
CEQ $\gamma = 10$	**Mean**	∞	0.0517	0.0393	0.0367	0.0298	0.0248
	Max	∞	0.1054	0.3826	0.8681	0.1059	0.0970
	Min	1.09E+02	0.0024	0.0060	0.0003	0.0335	0.0165
Turnover	**Mean**	9.02E+03	0.0205	0.0539	0.0355	0.0573	0.0501
	Max	1.91E+05	0.0692	0.1757	0.3262	0.0944	0.1951
	Min	0.6960	1.56E-03	1.38E-04	1.08E-05	8.32E-04	4.43E-04
Weight	**Mean**	21.7695	3.62E-03	4.04E-04	2.85E-04	1.36E-03	1.38E-03
	Max	4.62E+02	5.62E-03	1.24E-03	1.90E-03	1.83E-03	4.12E-03
Biggest percentage	**Min**	0.9912	0.6435	0.2095	0.1388	0.5945	0.6167
difference in weights	**Mean**	3.74E+09	23.2912	5.6360	2.1056	23.6510	7.9984
	Max	9.52E+11	6.51E+03	1.14E+03	1.91E+02	3.72E+03	1.21E+03

Note: The table shows standard deviation of the performance measure within one quarter for the GMV portfolio. The min displays the lowest standard deviation between the 100 bootstrap sequences found within one quarter, across the 40 quarters. The mean displays the average standard deviation between the 100 bootstrap sequences found across the 40 quarters. Finally max displays the maximum standard deviation between the 100 bootstrap sequences within one quarter found across the 40 quarter. G-Lasso ρ is the single set regularization value found to perform the best overall. G-Lasso ρ combination, is the regularization value that consist of the regularization values that where optimal for previous quarter and is therefore not consistently the same regularization value across quarters. Finally the biggest percentage difference in weights displays the average percentage difference between the largest and smallest weight of each asset over the 100 bootstrap data sequences. The data set used is from 2010 to 2019.

Table 16: Standard deviation within the one quarter for the GMV-W portoflios

GMV-W		**POET**	**LW**	**G-Lasso ρ**	**G-Lasso** Combination	**G-Lasso** ρ**Matrix**	**G-Lasso Matrix Combination**
	Min	0.0154	5.94E-04	1.38E-04	5.24E-07	1.68E-04	6.04E-05
St.dev	**Mean**	0.0702	0.0023	0.0025	0.0015	0.0009	0.0019
	Max	0.2788	0.0084	0.0076	0.0092	0.0046	0.0095
	Min	0.1597	8.71E-03	5.89E-03	4.81E-05	3.69E-03	4.62E-03
Sharpe	**Mean**	0.6719	0.0797	0.0422	0.0244	0.0180	0.0562
	Max	1.4394	0.2535	0.1900	0.1229	0.0460	0.2765
	Min	0.0909	1.67E-03	9.41E-04	5.74E-06	8.50E-04	6.42E-04
CEQ $\gamma = 1$	**Mean**	0.5076	0.0144	0.0107	0.0074	0.0040	0.0126
	Max	5.4075	0.0357	0.0413	0.0820	0.0113	0.0479
	Min	0.0913	1.85E-03	8.65E-04	5.46E-06	7.81E-04	6.73E-04
CEQ $\gamma = 5$	**Mean**	4.5385	0.0155	0.0114	0.0097	0.0041	0.0135
	Max	1.59E+02	0.0639	0.0769	0.1510	0.0145	0.0704
	Min	0.0900	2.30E-03	8.81E-04	5.13E-06	7.01E-04	6.81E-04
CEQ $\gamma = 10$	**Mean**	7.20E+04	0.0180	0.0140	0.0143	0.0049	0.0161
	Max	2.81E+06	0.1202	0.1500	0.2913	0.0289	0.1315
	Min	0.1097	2.78E-03	8.22E-04	5.82E-06	5.16E-04	5.94E-04
Turnover	**Mean**	0.1565	0.0208	0.0185	0.0104	0.0047	0.0127
	Max	0.2367	0.0692	0.0697	0.0616	0.0354	0.0431
	Min	0.0012	8.54E-05	1.14E-05	8.48E-08	2.26E-05	1.43E-05
Weight	**Mean**	0.0017	2.57E-04	1.44E-04	8.35E-05	6.39E-05	1.90E-04
	Max	0.0023	5.45E-04	5.27E-04	5.38E-04	2.13E-04	4.99E-04
Biggest percentage	**Min**	1.2073	0.1055	0.0771	0.0531	0.0356	0.0659
difference in weights	**Mean**	8.9188	0.6451	0.3525	0.1978	0.1354	0.4606
	Max	3.31E+02	6.9469	4.2913	2.1426	0.8615	6.0821

Note: The table shows standard deviation of the performance measure within one quarter for the GMV-W portfolio. The min displays the lowest standard deviation between the 100 bootstrap sequences found within one quarter, across the 40 quarters. The mean displays the average standard deviation between the 100 bootstrap sequences found across the 40 quarters. Finally max displays the maximum standard deviation between the 100 bootstrap sequences within one quarter found across the 40 quarter. G-Lasso ρ is the single set regularization value found to perform the best overall. G-Lasso ρ combination, is the regularization value that consist of the regularization values that where optimal for previous quarter and is therefore not consistently the same regularization value across quarters. Finally the biggest percentage difference in weights displays the average percentage difference between the largest and smallest weight of each asset over the 100 bootstrap data sequences. The data set used is from 2010 to 2019.