ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis Econometrics and Operational Research

Efficient Portfolio Selection in a Large Market

Supervised by Barendse S.

First reader is Keijsers B.

Abstract

The focus of this paper is on a portfolio rule that approaches the optimal Sharpe ratio in a large market with a realistic amount of historical data and a well-chosen subspace, named "subspace$_{\mathcal{P}}$ mean-variance analysis". This portfolio rule carefully balances the tradeoff between the estimation error and the systematic error. A well-chosen subspace is the key extension on the paper of Chen and Yuan (2016). Also a mathematical comparison is given for the Markowitz (subspace$_{\mathcal{P}}$) mean-variance portfolio with and without the constraint that the sum of the portfolio holdings sums up to one. Another comparison is made between the subspaces $\mathcal{P}_{\Sigma}$ and $\mathcal{P}_{Corr}$, that are created with the eigenvectors of the covariance- or correlation matrix, by using the efficient frontier, the dimension of the subspaces and a comparison over different investment opportunities. To give the last comparison, a mathematical explanation needs to be given that an additional investment option could decrease the Sharpe ratio of the subspace$_{\mathcal{P}}$ mean-variance portfolio.

Diederik Nederhorst (410860)

(1-7-2017)

1 Introduction

Diversification is a risk management technique that mixes a variety of risky assets within a portfolio. The number and properties of these risky assets determine which portfolio rules are optimal. Many rules that exist have been recently examined and compared by Tu and Zhou (2011), and DeMiguel et al. (2009). The Sharpe ratio (Sharpe, 1966), that is a measure for calculating risk-adjusted return, gives an indication which portfolio rules outperform each other. The focus of this paper is on a portfolio rule that approaches the optimal Sharpe ratio in a large market with a realistic amount of historical data and a well-chosen subspace. The name of this portfolio rule is "subspace$_{\mathcal{P}}$ mean-variance analysis". A well-chosen subspace is the key extension on the paper of Chen and Yuan (2016).

The Markowitz mean-variance analysis is used, see Markowitz (1952) and Markowitz (1959), to achieve the optimal Sharpe ratio. The estimates of the first and second central moment conditions are used for the estimated mean-variance analysis. However, to get meaningful moment estimators in a large market, the estimation window should be unrealistic large. For example, just a naïve diversification, which simply assigns an equal weight to each of the assets, already outperforms an estimated mean-variance portfolio in a large market (DeMiguel et al., 2009). The main reason for this outperformance is that, when the market is large, one cannot realistically expect to have enough historical data to do better than naïve diversification.

The solution for the problem mentioned above, the problem of an unrealistic large estimation window to get meaningful moment estimators, is to reduce the amount of estimated parameters in the mean and covariance of the asset returns. The reduction of parameters can be achieved by taking a linear subspace of $\mathbb{R}^N$, where N equals the number of risky assets in the original market. A subspace for investment opportunities could be a set of portfolios, including portfolios corresponding to the leading eigenvectors of the covariance- and/or correlation matrix of asset excess returns. However, due to the selected subspace, an additional investment opportunity does not necessarily mean that the portfolio composition achieves a better Sharpe ratio. It is even possible that an additional investment opportunity results in a decrease of the Sharpe ratio. I will explain this phenomenon using mathematical methods. As is logically the case with other portfolio rules, by introducing an additional investment opportunity, the Sharpe ratio increase or remain equal due to diversification.

The choice of the linear subspace ($\mathcal{P}$) plays an important role in the subspace$_{\mathcal{P}}$ mean-variance portfolio. Instead of choosing only the subspace $\mathcal{P}_{\Sigma}$ as the linear subspace spanned by the first d eigenvectors of the covariance matrix (Σ), which is done by Chen and Yuan (2016), I will also choose $\mathcal{P}_{Corr}$ as the linear subspace spanned by the first d eigenvectors of the correlation matrix ($Corr$). Depending on the relative variances of the individual excess returns, the eigenvectors of Σ and $Corr$ can be quite different. Different eigenvectors for Σ and $Corr$ results of course in different subspaces $\mathcal{P}_{\Sigma}$ and $\mathcal{P}_{Corr}$. These different subspaces results in different subspace$_{\mathcal{P}}$ mean-variance portfolios. It is therefore interesting to look at the performance of $\mathcal{P}_{Corr}$ relative to $\mathcal{P}_{\Sigma}$. This comparison takes place over the dimension of the subspaces, the efficient frontier and, perhaps the most important one, over different investment opportunities. Theoretically, the subspace $\mathcal{P}_{\Sigma}$ should give higher Sharpe ratios than the subspace $\mathcal{P}_{Corr}$ with the subspace$_{\mathcal{P}}$ mean-variance analysis. However, if an investment option meets certain characteristics, my self-made subspace $\mathcal{P}_{Corr}$ outperforms $\mathcal{P}_{\Sigma}$.

Gamma is an indispensable parameter for the comparison with the efficient frontier. The Markowitz subspace$_{\mathcal{P}}$ mean-variance portfolio contains a gamma that did not get any attention in the paper of Chen and Yuan (2016). This gamma is specified as the coefficient of the relative risk aversion. To evaluate the risk aversion, relative to gamma, there will be a simulation over this parameter. For each value of gamma the simulation gives an optimal portfolio composition. All these optimal portfolio compositions together, for each given γ and each subspace $\mathcal{P} \in \{\mathcal{P}_{\Sigma}, \mathcal{P}_{Corr}\}$, should give an indication of the efficient frontier. The two dimensional efficient frontier has the standard deviation of the portfolio return on the x-axes and the expected return on the y-axes. Finally, the highest expected return for a defined level of risk is achieved by one of the two subspaces. So a conclusion can be made about the optimal choice of the subspace $\mathcal{P}$ for the subspace$_{\mathcal{P}}$ mean-variance analysis.

It will be shown that the Sharpe ratio, achieved by the Markowitz subspace$_{\mathcal{P}}$ ($\mathcal{P}_{Corr}$ or $\mathcal{P}_{\Sigma}$) mean-variance analysis mentioned above, approaches the global optimal Sharpe ratio of the entire market that is calculated via the Markowitz mean-variance analysis. This is done on a theoretical way and a practice-oriented way, where real and simulated data sets are used for the practice-oriented way. The asset returns for the simulated data set are driven by systematic risks, which are represented by three research factors. The three factors are simulated from a multivariate normal distribution with the mean and covariance, calibrated from monthly data of July 1963 till August 2007, of the market portfolio, Small Minus Big(SMB) and High Minus Low(HML). The correctness of explaining systematic risks by marketwide factors is confirmed by numerous studies (see, e.g., Connor et al. (2010)).

2 Literature review

Harry Markowitz (Markowitz, 1952) considered the rule that the investor should consider expected return as a desirable thing and variance of return an undesirable thing. However, the desirable part and the undesirable part do not have to be assigned the same weight in the portfolio selection. This is due to the risk tolerance factor. The risk tolerance factor has a one-to-one relationship with the efficient frontier, this is explained in Best and Grauer (1991). This efficient frontier is the set of optimal portfolios that offers the highest expected return for a defined level of risk or the lowest risk for a given level of expected return.

The choice of the linear subspace, in the subspace$_{\mathcal{P}}$ mean-variance analysis, determines how close the portfolio is to the efficient frontier. It determines whether the maximum expected return is achieved for a defined level of risk. In other words, the subspace $\mathcal{P}$ determines if it is possible to achieve a portfolio composition that lies nearby the efficient frontier. Chen and Yuan chose $\mathcal{P}$ to be the linear subspace spanned by the first d eigenvectors of the variance-covariance matrix, this is also suggested by Carrasco and Noumon (2013). The subspace$_{\mathcal{P}}$ mean-variance analysis and the Principal Component Analysis (PCA) use both eigenvectors of the variance-covariance matrix for their analysis. The PCA however, sometimes also derives the eigenvectors of the correlation matrix. Wold et al.

(1987) claimed that it is well known that the principal components are not independent of the scales in which the original returns are measured. So the eigenvectors, and thereby also the subspace $\mathcal{P}$, are likewise dependent on the scales in which the original returns are measured. Wold, Esbensen and Geladi (1987) recommended to derive the principal components from the correlation matrix. The scaling is essential because PCA is a least squares method, which makes variables with large variance have large loadings. To avoid this bias, it is customary to standardize the data matrix so that each column has a variance of one. Standardizing the data matrix and then taking the variance-covariance matrix is of course the same as taking the correlation matrix of the unstandardized data matrix. Taking the eigenvectors of the correlation matrix is done because each variable has the same influence on the PC model. This is why it is very interesting to look at the Markowitz subspace$_{\mathcal{P}}$ mean-variance analysis for the subspace spanned by the first d eigenvectors of the correlation matrix.

3 Data

Kenneth R. French is well known for his research on the three-factor model (Fama and French, 1993). He is considered as an expert on the behavior of security prices and investment strategies.

3.1 Data for simulation

The Fama-French factors are constructed using the six value-weighted portfolios formed on size and book-to-market. The six portfolios are the combinations of two portfolios formed on Market Equity (ME) and three portfolios formed on the ratio of Book Equity to Market Equity (BE/ME). The grouping of the six value-weighted portfolios are

	Median ME	
	Small Value	Big Value
70$^{\text{th}}$ BE/ME percentile	Small Neutral	Big Neutral
30$^{\text{th}}$ BE/ME percentile	Small Growth	Big Growth

With these portfolios, the following three Fama-French factors are constructed:

- Small Minus Big (SMB), is the average return on the three small portfolios minus the average return on the three big portfolios:
  SMB $= \frac{1}{3}$(Small Value + Small Neutral + Small Growth) $-\ \frac{1}{3}$(Big Value + Big Neutral + Big Growth).
- High Minus Low (HML), is the average return on the two value portfolios minus the average return on the two growth portfolios:
  $$\text{HML} = \tfrac{1}{2}(\text{Small Value} + \text{Big Value}) - \tfrac{1}{2}(\text{Small Growth} + \text{Big Growth}).$$
- The excess return on the market:
  $$\text{excess return}_{\text{market}} = r_m - r_f.$$

In Fama and French (1993) there is more explanation about the factor returns. The monthly Fama-French data from July 1963 to August 2007, that is constructed according to the methodology above, can be found on `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

3.2 Data empirical illustration

In the empirical illustration, specific portfolio returns are used. The Fama-French 25 (5×5) and 100 (10×10) portfolios are formed on size and book-to-market. The returns of the 25 portfolios are an equal-weighted combination of one out of five portfolios formed on Market Equity (ME) and one out of five portfolios formed on the ratio of Book Equity to Market Equity (BE/ME). The returns of the 100 portfolios are an equal-weighted combination of one out of ten portfolios formed on Market Equity (ME) and one out of ten portfolios formed on the ratio of Book Equity to Market Equity (BE/ME). The data set that will be used is from January 1961 to December 2010 and can be found on `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

4 Methodology

4.1 Subspace$_{\mathcal{P}}$ Mean-Variance Analysis

The Markowitz mean-variance portfolio is given as the solution to

$$\begin{aligned} &\min_{\mathbf{w}\in\mathbb{R}^N} \left\{ \frac{\gamma}{2}\mathbf{w}^T\Sigma\mathbf{w} - \mathbf{w}^T E \right\} \quad \text{where} \quad \Sigma \in \mathbb{R}^{N\times N},\ E \in \mathbb{R}^N \qquad (1) \\ \text{solution:} \qquad &\mathcal{L}(\mathbf{w}) = \frac{\gamma}{2}\mathbf{w}^T\Sigma\mathbf{w} - \mathbf{w}^T E \\ &\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w}) = \gamma\Sigma\mathbf{w} - E = 0 \ \Leftrightarrow \mathbf{w}_{\text{mv}} = \frac{1}{\gamma}\Sigma^{-1}E. \end{aligned}$$

E and Σ are the mean and variance of $\mathbf{r}$ respectively, where $\mathbf{r} \in \mathbb{R}^N$ is the return of the assets in excess of a risk-free rate and N equals the number of risky assets in the original market. The solution of the minimization in Equation 1, denoted as the mean-variance efficient portfolio ($\mathbf{w}_{\text{mv}}$), achieves the highest possible Sharpe ratio ($s(\mathbf{w}_{\text{mv}})$) for a given γ (Chen and Yuan, 2016):

$$\begin{aligned} s(\mathbf{w}_{\text{mv}}) &= \max_{\mathbf{w}\in\mathbb{R}^N} \{s(\mathbf{w})\} \\ &= \max_{\mathbf{w}\in\mathbb{R}^N} \left\{ \frac{\mathbf{w}^T E}{(\mathbf{w}^T\Sigma\mathbf{w})^{1/2}} \right\} \quad \text{where} \quad \mathbf{w}_{\text{mv}} = \frac{1}{\gamma}\Sigma^{-1}E. \end{aligned} \qquad (2)$$

Obviously, the mean-variance portfolio is something theoretical, because investors do not know E and Σ. So the goal is to find a practical portfolio rule that achieves the best possible approximation of the mean-variance portfolio. Using meaningful estimators of E and Σ could be such a practical portfolio rule. Meaningful moment estimators can be achieved in a small market with a realistic amount of historical data. However, to get meaningful moment estimators in a large market, the estimation window should be unrealistic large. This is caused by the increasing number of parameters in the mean and covariance of the asset returns if the size of the market increases. The consistent estimators of both moment conditions, with $\mathbf{r}_t$ the excess return in month t in the estimation window $[1, T]$, are given by

$$\hat{E} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{r}_t, \quad \text{and} \quad \hat{\Sigma} = \frac{1}{T-1}\sum_{t=1}^{T}(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_t - \bar{\mathbf{r}})^T. \tag{3}$$

Then the estimated mean-variance portfolio, or better known as the sample mean-variance portfolio, can be calculated as follows:

$$\hat{\mathbf{w}}_{\text{mv}} = \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{E}. \tag{4}$$

In case of a large estimation window for a small market, the sample mean-variance portfolio is efficient. If neither the estimation window is large nor the market is small, the estimated moment conditions become useless. Unfortunately, these poor estimations result in a useless portfolio composition where the Sharpe ratio is considerably small.

So useless portfolio compositions are the result of using the sample mean-variance analysis in a large market. The reason for this is that the large number of parameters in the mean and covariance all need to be estimated. A solution to this problem would be to reduce the number of estimated parameters. This reduction is possible if $\mathbf{w}$ is not an element of the whole $\mathbb{R}^N$, but of a carefully selected linear subspace $\mathcal{P}$ with a specific span. The span is defined as $\text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} = \{t_1\boldsymbol{v}_1 + \cdots + t_k\boldsymbol{v}_k : t_1, \ldots, t_k \in \mathbb{R}\}$, with $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ the leading eigenvectors of the covariance- or correlation matrix. The estimated parameters are given by $t_1, \ldots, t_k$, where $k \ll N$. The portfolio composition within a subspace is given as the solution to the Markowitz subspace$_{\mathcal{P}}$ mean-variance analysis:

$$\min_{\mathbf{w}_d^{\mathcal{P}} \in \mathcal{P}} \left\{\frac{\gamma}{2}\mathbf{w}_d^{\mathcal{P}^T}\Sigma\mathbf{w}_d^{\mathcal{P}} - \mathbf{w}_d^{\mathcal{P}^T}E\right\} \quad \text{where} \quad \mathcal{P} \in \{\mathcal{P}_{\Sigma}, \mathcal{P}_{Corr}\}. \tag{5}$$

If E and Σ, in Equation 5, are given by the mean and variance of the return of the assets respectively, instead of the mean and variance of the excess return of the assets, the fraction $\frac{\gamma}{2}$ has the following interpretation: The fraction $\frac{\gamma}{2}$ is the risk tolerance factor, where $\gamma \to \infty$ results in the portfolio with minimal risk and $\gamma \to 0$ results in a portfolio infinitely far on the efficient frontier with both risk and expected return unbounded. The efficient frontier is the curve that shows all efficient portfolios in a risk-return framework. These efficient portfolios are defined as the portfolios that minimize the risk subject for a given expected return, or equivalently to minimize Equation 5 with E and Σ the mean and variance of the return respectively.

To give the solution of the Markowitz subspace$_\mathcal{P}$ mean-variance portfolio, some notation needs to be introduced. $P_\mathcal{P}$, in Equation 6, is a $N \times d$ matrix whose columns are an orthonormal basis of $\mathcal{P}$. Additionally, for $\mathbf{x}_\mathcal{P} \in \mathbb{R}^d$ and $P_\mathcal{P}\mathbf{x}_\mathcal{P} \in \mathcal{P}$, the map $\mathbb{R}^d \to \mathcal{P} : \mathbf{x}_\mathcal{P} \mapsto P_\mathcal{P}\mathbf{x}_\mathcal{P} = \mathbf{w}_d^\mathcal{P}$ is injective and surjective. So the solution of Equation 5 is given by:

$$\begin{aligned}
\mathcal{L}(\mathbf{x}_\mathcal{P}) &= \frac{\gamma}{2} \left(P_\mathcal{P}\mathbf{x}_\mathcal{P}\right)^T \Sigma \left(P_\mathcal{P}\mathbf{x}_\mathcal{P}\right) - \left(P_\mathcal{P}\mathbf{x}_\mathcal{P}\right)^T E \\
\nabla_{\mathbf{x}_\mathcal{P}} \mathcal{L}(\mathbf{x}_\mathcal{P}) &= \gamma P_\mathcal{P}^T \Sigma P_\mathcal{P} \mathbf{x}_\mathcal{P} - P_\mathcal{P}^T E = 0 \ \Leftrightarrow \mathbf{x}_\mathcal{P} = \frac{1}{\gamma} \left(P_\mathcal{P}^T \Sigma P_\mathcal{P}\right)^{-1} P_\mathcal{P}^T E \\
\Rightarrow \ \mathbf{w}_{\mathrm{mv},d}^\mathcal{P} &:= \frac{1}{\gamma} P_\mathcal{P} \left(P_\mathcal{P}^T \Sigma P_\mathcal{P}\right)^{-1} P_\mathcal{P}^T E. \qquad (6)
\end{aligned}$$

However, the portfolio composition of Chen and Yuan (2016) does not use the constraint that the portfolio weights have to sum up to one. This gives very difficult interpretable results. Namely, a solution of the Markowitz (subspace$_\mathcal{P}$) mean-variance analysis gives a portfolio composition where only a fraction of the amount you want to invest is being invested. The self-made Markowitz (subspace$_\mathcal{P}$) mean-variance portfolio with constraint, the constraint that the sum of the portfolio weights sum up to one, is denoted as ${\mathbf{w}_\mathrm{mv}}^*$ (${\mathbf{w}_{\mathrm{mv},d}^\mathcal{P}}^*$) and is given by:

$$\begin{aligned}
{\mathbf{w}_\mathrm{mv}}^* \ &:= \mathbf{w}_\mathrm{mv} + \frac{1}{\gamma} \left(\frac{\gamma - E^T \Sigma^{-1} \iota}{\iota^T \Sigma^{-1} \iota} \right) \Sigma^{-1} \iota \qquad (7) \\
{\mathbf{w}_{\mathrm{mv},d}^\mathcal{P}}^* \ &:= \mathbf{w}_{\mathrm{mv},d}^\mathcal{P} + \left(\frac{1 - \frac{1}{\gamma} \iota^T P_\mathcal{P} \left(P_\mathcal{P}^T \Sigma P_\mathcal{P}\right)^{-1} P_\mathcal{P}^T E}{\iota^T P_\mathcal{P} \left(P_\mathcal{P}^T \Sigma P_\mathcal{P}\right)^{-1} P_\mathcal{P}^T \iota} \right) P_\mathcal{P} \left(P_\mathcal{P}^T \Sigma P_\mathcal{P}\right)^{-1} P_\mathcal{P}^T \iota. \qquad (8)
\end{aligned}$$

I have given the proofs of the Equations 7 and 8 in Appendixes B and C respectively, but there will not be a further discussion in this paper about the the Markowitz (subspace$_\mathcal{P}$) mean-variance analysis subject to the constraint mentioned above.

In the following two equations there will be a more specific solution given for each subspace discussed in this paper, the subspace $\mathcal{P}_\Sigma$ of Chen and Yuan and my own subspace $\mathcal{P}_{Corr}$, of the Markowitz subspace$_\mathcal{P}$ mean-variance portfolio. The solution of the Markowitz subspace$_\Sigma$ mean-variance portfolio is more concisely written down in Equation 9, with a self-made step-by-step procedure that uses orthogonal diagonalization, for the linear subspace $\mathcal{P}_\Sigma$. This subspace is spanned by the first d eigenvectors of Σ. By taking $P_{\mathcal{P}_\Sigma} = [\eta_1, \cdots, \eta_d]$ ($P_\Sigma = [\eta_1, \cdots, \eta_N]$), with η_i and θ_i the eigenvector and eigenvalue of Σ respectively, the solution for subspace $\mathcal{P}_\Sigma$ is given by

$$
\begin{aligned}
\mathbf{w}_{\mathrm{mv},d}^{\mathcal{P}_\Sigma} :=& \frac{1}{\gamma} P_{\mathcal{P}_\Sigma} \left(P_{\mathcal{P}_\Sigma}^T \ \Sigma \ P_{\mathcal{P}_\Sigma}\right)^{-1} P_{\mathcal{P}_\Sigma}^T E \\
=& \ \frac{1}{\gamma} P_{\mathcal{P}_\Sigma} \left(P_{\mathcal{P}_\Sigma}^T \ P_\Sigma \ D_\Sigma \ P_\Sigma^T \ P_{\mathcal{P}_\Sigma}\right)^{-1} P_{\mathcal{P}_\Sigma}^T E \quad \text{with } D_\Sigma = \mathrm{diag}(\theta_1, \cdots, \theta_N) \\
=& \ \frac{1}{\gamma} P_{\mathcal{P}_\Sigma} \left(\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \theta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_N \end{bmatrix} \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \right)^{-1} P_{\mathcal{P}_\Sigma}^T E \\
=& \ \frac{1}{\gamma} P_{\mathcal{P}_\Sigma} \begin{bmatrix} \frac{1}{\theta_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\theta_d} \end{bmatrix} P_{\mathcal{P}_\Sigma}^T E \\
=& \ \frac{1}{\gamma} \begin{bmatrix} \frac{1}{\theta_1}\eta_1 & \cdots & \frac{1}{\theta_d}\eta_d \end{bmatrix} \begin{bmatrix} \eta_1 & \cdots & \eta_d \end{bmatrix}^T E = \ \frac{1}{\gamma} \sum_{k=1}^{d} \frac{1}{\theta_k} \eta_k \eta_k^T E \quad \text{where } \theta_1 \geq \theta_2 \geq \ldots
\end{aligned}
\tag{9}
$$

The solution in Equation 6 is more concisely written down in Equation 10, with a self-made step-by-step procedure that uses orthogonal diagonalization, for my own linear subspace $\mathcal{P}_{Corr}$. This subspace is spanned by the first d eigenvectors of $Corr$. By taking $P_{\mathcal{P}_{Corr}} = [\eta_1, \cdots, \eta_d]$ ($P_{Corr} = [\eta_1, \cdots, \eta_N]$), with η_i and θ_i the eigenvector and eigenvalue of $Corr$ respectively, the solution for subspace $\mathcal{P}_{Corr}$ is given by

$$
\begin{aligned}
\mathbf{w}_{\mathrm{mv},d}^{\mathcal{P}_{Corr}} :=& \frac{1}{\gamma} P_{\mathcal{P}_{Corr}} \left(P_{\mathcal{P}_{Corr}}^T \Sigma P_{\mathcal{P}_{Corr}}\right)^{-1} P_{\mathcal{P}_{Corr}}^T E \\
=& \frac{1}{\gamma} P_{\mathcal{P}_{Corr}} \left(P_{\mathcal{P}_{Corr}}^T D_\sigma \ Corr \ D_\sigma \ P_{\mathcal{P}_{Corr}}\right)^{-1} P_{\mathcal{P}_{Corr}}^T E \quad \text{with } D_\sigma = \mathrm{diag}(\sigma_{11}, \cdots, \sigma_{NN})^{1/2} \\
=& \frac{1}{\gamma} P_{\mathcal{P}_{Corr}} \left(P_{\mathcal{P}_{Corr}}^T D_\sigma \ P_{Corr} \ D_{Corr} \ P_{Corr}^T \ D_\sigma \ P_{\mathcal{P}_{Corr}}\right)^{-1} P_{\mathcal{P}_{Corr}}^T E \text{ with } D_{Corr} = \mathrm{diag}(\theta_1, \cdots, \theta_N).
\end{aligned}
\tag{10}
$$

An element of $I = P_{\mathcal{P}_{Corr}}^T D_\sigma \ P_{Corr} \ D_{Corr} \ P_{Corr}^T \ D_\sigma \ P_{\mathcal{P}_{Corr}}$ is given by the expression

$$
I(i,j) = \sum_{k=1}^{N} \theta_k \left(\left[\eta_i^T e_{\sigma_1} \ \eta_i^T e_{\sigma_2} \cdots \ \eta_i^T e_{\sigma_N}\right] \eta_k \right) \left(\left[\eta_j^T e_{\sigma_1} \ \eta_j^T e_{\sigma_2} \cdots \ \eta_j^T e_{\sigma_N}\right] \eta_k \right) \quad \text{for } i, j = 1, \ldots, d, \tag{11}
$$

where e_{σ_k} is the k^{th} column of $\mathrm{diag}(\sigma_{11}, \cdots, \sigma_{NN})^{1/2}$, $e_{\sigma_k} = \begin{bmatrix} 0 & \cdots & \sqrt{\sigma_{kk}} & \cdots & 0 \end{bmatrix}^T$. The set of eigenvectors (P_{Corr}) is already an orthonormal basis, so the eigenvectors per set are pairwise orthogonal when their eigenvalues are different and the eigenvectors have an Euclidean norm of 1. The proof of Equation 11 is given in Appendix D.

The subspace$_{\mathcal{P}}$ mean-variance portfolio is again a theoretical measurement, because of the unknown Σ and E. To use this portfolio rule in practice, we need to estimate these moment conditions with Equation 3 and use the eigenvectors and eigenvalues of $\hat{\Sigma}$ and $\hat{Corr}$ in Equations 9 and 10. The estimated subspace$_{\mathcal{P}}$ mean-variance portfolios, for $\mathcal{P} \in \{\mathcal{P}_{\Sigma}, \mathcal{P}_{Corr}\}$, are given by

$$\begin{aligned} \hat{\mathbf{w}}^{\mathcal{P}_{\Sigma}}_{\mathrm{mv},d} &= \frac{1}{\gamma}\sum_{k=1}^{d}\frac{1}{\hat{\theta}_k}\hat{\eta}_k\hat{\eta}_k^T\hat{E} \\ \hat{\mathbf{w}}^{\mathcal{P}_{Corr}}_{\mathrm{mv},d} &= \frac{1}{\gamma}\hat{P}_{\mathcal{P}_{Corr}}\hat{I}^{-1}\hat{P}^T_{\mathcal{P}_{Corr}}\hat{E}. \end{aligned} \tag{12}$$

If the number of eigenvectors in the orthonormal basis of the subspace equals the number of risky assets in the market ($\dim(\mathcal{P}) = N$), then the sample mean-variance portfolio is the same as the estimated subspace$_{\mathcal{P}}$ mean-variance portfolio given above. Namely, the linear subspace spanned by N eigenvectors is equal to the space $\mathbb{R}^N$. So in mathematical notation:

$$\hat{\mathbf{w}}_{\mathrm{mv}} = \hat{\mathbf{w}}^{\mathcal{P}}_{\mathrm{mv},d=N}\left(= \hat{\mathbf{w}}^{\mathcal{P}_{\Sigma}}_{\mathrm{mv},d=N} = \hat{\mathbf{w}}^{\mathcal{P}_{Corr}}_{\mathrm{mv},d=N}\right). \tag{13}$$

As indicated in Equation 13, the choice of $d = \dim(\mathcal{P})$ is important, knowing that the estimated subspace$_{\mathcal{P}}$ mean-variance portfolio for $d = N$ is a poor portfolio rule in a large market. There is namely a trade-off between two errors when the subspace dimension vary:

$$s(\mathbf{w}_{\mathrm{mv}}) - s(\hat{\mathbf{w}}^{\mathcal{P}}_{\mathrm{mv},d}) = [s(\mathbf{w}^{\mathcal{P}}_{\mathrm{mv},d}) - s(\hat{\mathbf{w}}^{\mathcal{P}}_{\mathrm{mv},d})] + [s(\mathbf{w}_{\mathrm{mv}}) - s(\mathbf{w}^{\mathcal{P}}_{\mathrm{mv},d})]. \tag{14}$$

The first error is referred to as the estimation error and the second error is referred to as the systematic error. The estimation error measures the loss of optimality due to the estimated moments, also depends on the size of the estimation window, whereas the systematic error is due to less investment options.

4.2 A decreasing Sharpe ratio as a result of an additional investment option

In this section there will be a mathematical explanation about the magnitude of the Sharpe ratio relative to the number of different investment opportunities. Logically, by adding an additional investment opportunity, the Sharpe ratio should increase or remain equal due to diversification. It should not be possible for the Sharpe ratio to fall whenever an additional investment opportunity is introduced. However, it is possible for the subspace$_{\mathcal{P}}$ mean-variance analysis due to the selected subspace. This decreasing Sharpe ratio as a result of an additional investment option, for the subspace$_{\mathcal{P}}$ mean-variance portfolio, will be further explained below. This phenomenon could have consequences for how many assets investors have to incorporate in their investment strategy.

Let $\mathcal{P}^{\mathbb{R}^N} \subseteq \mathbb{R}^N$ and $\mathcal{P}^{\mathbb{R}^{N+1}} \subseteq \mathbb{R}^{N+1}$ be different subspaces of investment compositions. These subspaces are constructed out of the leading eigenvectors of the covariance- or correlation matrix and are used in the subspace$_{\mathcal{P}}$ mean-variance analysis. The only reasonable assumption that will be made for explanatory purposes is that the number of different investment compositions is finite, or in other words $|\mathcal{P}^{\mathbb{R}^N}| < \infty$ and $|\mathcal{P}^{\mathbb{R}^{N+1}}| < \infty$. Reasonable, as an example, because of transaction costs. Because of this assumption, the subspaces $\mathcal{P}^{\mathbb{R}^N} \subseteq \mathbb{R}^N$ and $\mathcal{P}^{\mathbb{R}^{N+1}} \subseteq \mathbb{R}^{N+1}$ can made visible via a set diagram in Figure 1. Obviously, the two subspaces are disjoint, $\mathcal{P}^{\mathbb{R}^N} \cap \mathcal{P}^{\mathbb{R}^{N+1}} = \emptyset$. This is because of the different dimension of the elements of the subspaces $\mathcal{P}^{\mathbb{R}^N} = \{\{\bar{w}^1\}, \cdots, \{\bar{w}^k\}\}$ and $\mathcal{P}^{\mathbb{R}^{N+1}} = \{\{\bar{w}^{1*}\}, \ldots, \{\bar{w}^{m*}\}\}$, where $\bar{w}^i = \begin{bmatrix} w_1^i & \cdots & w_N^i \end{bmatrix}^T$ $\big(|\bar{w}^i| = N\big)$ with $i = 1, \ldots, k$ and $\bar{w}^{j*} = \begin{bmatrix} \tilde{w}_1^j & \cdots & \tilde{w}_{N+1}^j \end{bmatrix}^T$ $\big(|\bar{w}^{j*}| = N+1\big)$ with $j = 1, \ldots, m$.

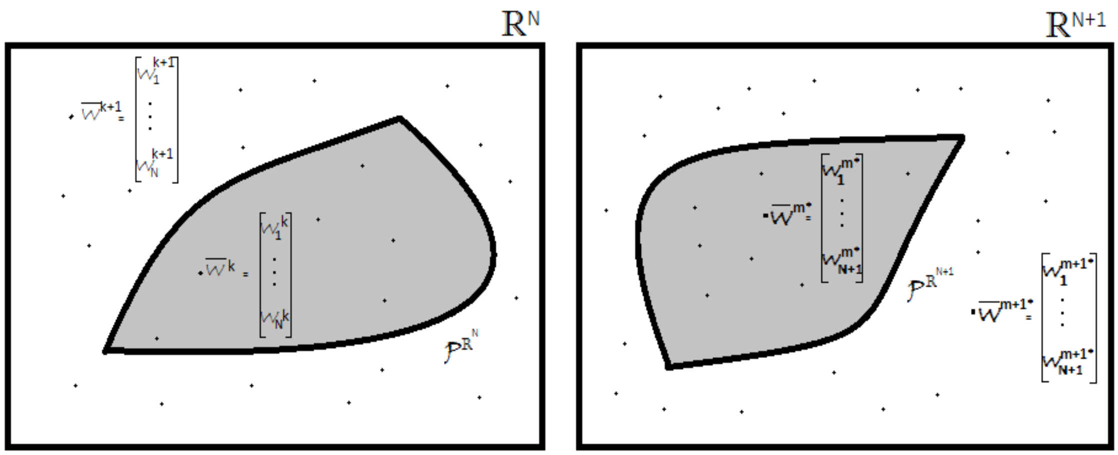

Figure 1: The left panel contains a Venn diagram in $\mathbb{R}^N$ and the right panel contains a Venn diagram in $\mathbb{R}^{N+1}$. All dots represent different investment compositions of dimension N, left panel, or dimension $N+1$, right panel. $\mathcal{P}^{\mathbb{R}^N}$ and $\mathcal{P}^{\mathbb{R}^{N+1}}$ are subspaces made out of the leading eigenvectors of the covariance- or correlation matrix.

Because of the disjoint subspaces, explained above, there can not be any comparison between the sets in the left and right panel of Figure 1 without doing any modifications. However, a comparison between the two subspaces is needed to verify if it is possible that the subspace$_{\mathcal{P}^{\mathbb{R}^{N+1}}}$ mean-variance portfolio has a smaller Sharpe ratio than the subspace$_{\mathcal{P}^{\mathbb{R}^N}}$ mean-variance portfolio. To compare the subspaces, another set diagram needs to be introduced that contains the transformed subspaces $\mathcal{P}_N^{\mathbb{R}}$ and $\mathcal{P}_{N+1}^{\mathbb{R}}$ of the subspaces $\mathcal{P}^{\mathbb{R}^N}$ and $\mathcal{P}^{\mathbb{R}^{N+1}}$ respectively. Let the subspace $\mathcal{P}_N^{\mathbb{R}}$ contain the $k \times N$ possible investment holdings that occur in the subspace $\mathcal{P}^{\mathbb{R}^N}$:

$$\begin{aligned}
\mathcal{P}_N^{\mathbb{R}} &= \{\{w_1^1\}, \ldots, \{w_N^1\}, \{w_1^2\}, \ldots, \{w_N^2\}, \ldots, \{w_1^{k-1}\}, \ldots, \{w_N^{k-1}\}, \{w_1^k\}, \ldots, \{w_N^k\}\} \\
&= \{\{w_1^1\}, \ldots, \{w_1^k\}, \{w_2^1\}, \ldots, \{w_2^k\}, \ldots, \{w_{N-1}^1\}, \ldots, \{w_{N-1}^k\}, \{w_N^1\}, \ldots, \{w_N^k\}\} \\
&= \{w^1, w^2, \ldots, w^{N-1}, w^N\} \quad \text{where } w^q = \{\{w_q^1\}, \ldots, \{w_q^k\}\} \text{ for } q = 1, \ldots, N.
\end{aligned}$$

The subset w^q contains all the different holdings for asset q from the subspace $\mathcal{P}_N^{\mathbb{R}}$, so $w^q \subseteq \mathcal{P}_N^{\mathbb{R}} \subseteq \mathbb{R}$. Finally, let the subspace $\mathcal{P}^{\mathbb{R}}_{N+1}$ contain the $m \times (N+1)$ possible investment holdings that occur in the subspace $\mathcal{P}^{\mathbb{R}^{N+1}}$:

$$\begin{aligned}
\mathcal{P}^{\mathbb{R}}_{N+1} &= \{\{\tilde{w}^1_1\}, \dots, \{\tilde{w}^1_{N+1}\}, \{\tilde{w}^2_1\}, \dots, \{\tilde{w}^2_{N+1}\}, \dots, \{\tilde{w}^{m-1}_1\}, \dots, \{\tilde{w}^{m-1}_{N+1}\}, \{\tilde{w}^m_1\}, \dots, \{\tilde{w}^m_{N+1}\}\} \\
&= \{\{\tilde{w}^1_1\}, \dots, \{\tilde{w}^m_1\}, \{\tilde{w}^1_2\}, \dots, \{\tilde{w}^m_2\}, \dots, \{\tilde{w}^1_N\}, \dots, \{\tilde{w}^m_N\}, \{\tilde{w}^1_{N+1}\}, \dots, \{\tilde{w}^m_{N+1}\}\} \\
&= \{w^{1*}, w^{2*}, \dots, w^{N*}, w^{N+1*}\} \quad \text{where } w^{q`*} = \{\{\tilde{w}^1_{q`}\}, \dots, \{\tilde{w}^m_{q`}\}\} \text{ for } q` = 1, \dots, N, N+1
\end{aligned}$$

The subset $w^{q`*}$ contains all the different holdings for asset $q`$ from the subspace $\mathcal{P}^{\mathbb{R}}_{N+1}$, so $w^{q`*} \subseteq \mathcal{P}^{\mathbb{R}}_{N+1} \subseteq \mathbb{R}$. It holds that both subspaces $\mathcal{P}^{\mathbb{R}}_N$ and $\mathcal{P}^{\mathbb{R}}_{N+1}$ are finite. This can be concluded from the assumption made earlier, that $|\mathcal{P}^{\mathbb{R}^N}| < \infty$ and $|\mathcal{P}^{\mathbb{R}^{N+1}}| < \infty$, so $k \times N < \infty$ and $m \times (N+1) < \infty$. Because both subspaces $\mathcal{P}^{\mathbb{R}}_N$ and $\mathcal{P}^{\mathbb{R}}_{N+1}$ are finite, they can made visible via a set diagram in Figure 1. For illustrative purposes, N is equal to 5 in this figure.

It holds by construction that the cardinality of the finite subset w^q equals k, $|w^q| = k < \infty$ $\left(|w^{q`*}| = m < \infty\right)$ for $\forall q$ ($\forall q`$). Then the first draw out of $\mathcal{P}^{\mathbb{R}}_N$ $\left(\mathcal{P}^{\mathbb{R}}_{N+1}\right)$ is from an arbitrary subset w^q ($w^{q`*}$), so the first draw has $k \times N$ ($m \times (N+1)$) different options and is the $q^{th}$ ($q`^{th}$) element of the vector $\bar{w}^i$ ($\bar{w}^{j*}$). Knowing the first draw from $\mathcal{P}^{\mathbb{R}}_N$ $\left(\mathcal{P}^{\mathbb{R}}_{N+1}\right)$, the remaining $N-1$ $(N+1-1)$ drawings from the remaining subspaces w^q ($w^{q`*}$), to fill the the vector $\bar{w}^i$ ($\bar{w}^{j*}$), are fixed. This is due to the known investment compositions in the left (right) panel of Figure 1.

By adding one extra investment opportunity, the eigenvectors of the covariance- or correlation matrix will change. The span of the leading eigenvectors, that made the subspaces $\mathcal{P}^{\mathbb{R}^N}$ and $\mathcal{P}^{\mathbb{R}^{N+1}}$, will shift or change completely because of the change in the eigenvectors. However, there can still be an overlapping part of the subspaces $\mathcal{P}^{\mathbb{R}}_N$ and $\mathcal{P}^{\mathbb{R}}_{N+1}$, as shown in Figure 2. By the shift of the subspace $\mathcal{P}^{\mathbb{R}}_{N+1}$, new investment opportunities are joining or disappearing. Perhaps better investment opportunities arise in $\mathcal{P}^{\mathbb{R}}_{N+1}$ due to this shift, but more importantly, perhaps the best investment composition in subspace $\mathcal{P}^{\mathbb{R}}_N$ disappears and no better one arises in the subspace $\mathcal{P}^{\mathbb{R}}_{N+1}$! This disappearance means that the best investment opportunity of the set $\mathcal{P}^{\mathbb{R}}_N \cup \mathcal{P}^{\mathbb{R}}_{N+1}$ is in the subset $\mathcal{P}^{\mathbb{R}}_N \setminus \mathcal{P}^{\mathbb{R}}_{N+1}$ of Figure 2. In this case, a drop of the Sharpe ratio is the result after adding one investment opportunity.

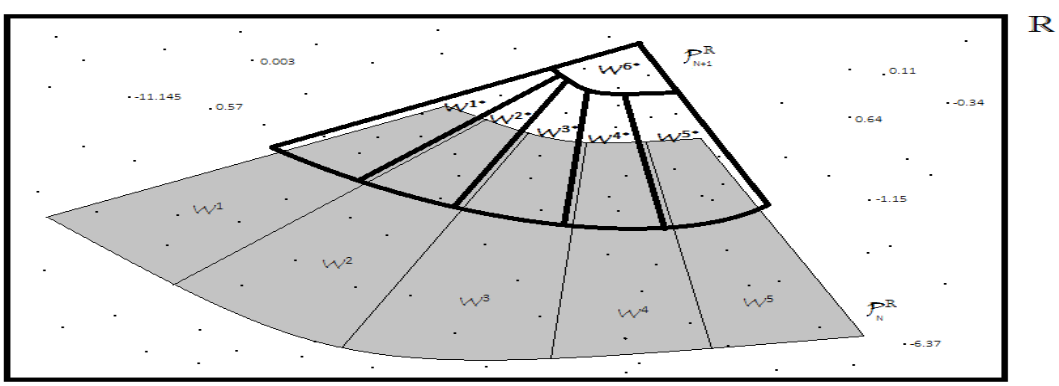

Figure 2: This Venn diagram contains two different subspaces $\mathcal{P}^{\mathbb{R}}_N$ (gray area with thin borders) and $\mathcal{P}^{\mathbb{R}}_{N+1}$ (white area with thick borders) in $\mathbb{R}$. All dots represent a one dimensional holding. For illustrative purposes there is an overlapping part of the subspaces ($\mathcal{P}^{\mathbb{R}}_N \cap \mathcal{P}^{\mathbb{R}}_{N+1}$), so the same holdings are in both subspaces. There are holdings only present in $\mathcal{P}^{\mathbb{R}}_N$ ($\mathcal{P}^{\mathbb{R}}_N \setminus \mathcal{P}^{\mathbb{R}}_{N+1}$) and holdings only present in $\mathcal{P}^{\mathbb{R}}_{N+1}$ ($\mathcal{P}^{\mathbb{R}}_{N+1} \setminus \mathcal{P}^{\mathbb{R}}_N$). The subspace $\boldsymbol{w}^{6*}$ contains the $(N+1)^{\text{th}}$ investment holdings.

4.3 Econometric Properties

In this section, the theoretical explanation will be given that in a large market the Sharpe ratio, achieved by the estimated subspace$_{\mathcal{P}}$ mean-variance analysis, approaches the optimal Sharpe ratio, achieved by the global mean-variance analysis. The assumption is made that the systematic risk, some undiversifiable risk, of the portfolio excess returns ($\boldsymbol{r}_t = [r_{1t}, \ldots, r_{Nt}]^T$) is due to systematic risk factors:

$$r_{jt} = \alpha_j + \beta_{j1} f_{1t} + \cdots + \beta_{jK} f_{Kt} + \epsilon_{jt} \quad j = 1, \ldots, N; \quad t = 1, \ldots, T. \tag{15}$$

Or this factor model in matrix notation:

$$\boldsymbol{r}_t = \boldsymbol{\alpha} + B\boldsymbol{f}_t + \boldsymbol{\epsilon}_t \quad t = 1, \ldots, T, \tag{16}$$

where α_j ($\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_N]^T$) is the Jensen's alpha (Jensen, 1968) that is assumed to be equal to zero, f_{jt} ($\boldsymbol{f}_t = [f_{1t}, \ldots, f_{Kt}]^T$) is the j^{th} common factor, β_{ji} (B is the $N \times K$ matrix) is the factor loading for asset j on factor i and ϵ_{jt} ($\boldsymbol{\epsilon}_t = [\epsilon_{1t}, \ldots, \epsilon_{Nt}]^T$) is the idiosyncratic component or asset-specific factor. The factor realizations $\boldsymbol{f}_t$ are assumed to be stationary with unconditional moments $\mathbb{E}(\boldsymbol{f}_t) = \boldsymbol{\mu}_f > \mathbf{0}$ and $V(\boldsymbol{f}_t) = \mathbb{E}[(\boldsymbol{f}_t - \boldsymbol{\mu}_f)(\boldsymbol{f}_t - \boldsymbol{\mu}_f)^T] = \Sigma_f$. The asset specific terms ϵ_{jt} are assumed to be uncorrelated with each of the common factors f_{jt}. Finally, the assumption that the error terms ϵ_{jt} are serially uncorrelated and contemporaneously uncorrelated across assets, so the variance-covariance matrix of the idiosyncratic risks is equal to the diagonal matrix Σ_ϵ. As a result, the true covariance matrix of the excess returns equals:

$$\begin{aligned}
\Sigma &= \mathbb{E}\left\{(\boldsymbol{r}_t - E(\boldsymbol{r}))(\boldsymbol{r}_t - E(\boldsymbol{r}))^T\right\} \\
&= \mathbb{E}\left\{\left(B(\boldsymbol{f}_t - \boldsymbol{\mu}_f) + \boldsymbol{\epsilon}_t\right)\left(B(\boldsymbol{f}_t - \boldsymbol{\mu}_f) + \boldsymbol{\epsilon}_t\right)^T\right\} \\
&= B\mathbb{E}\left\{(\boldsymbol{f}_t - \boldsymbol{\mu}_f)(\boldsymbol{f}_t - \boldsymbol{\mu}_f)^T\right\}B^T + \mathbb{E}\left\{\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^T\right\} = B\Sigma_f B^T + \Sigma_\epsilon.
\end{aligned} \tag{17}$$

In Appendix A, multiple assumptions about the approximate factor model are made. Because of these previous assumptions, the main result can be introduced in Equation 18, which is proven in Chen and Yuan (2016). This equation shows that the optimal Sharpe ratio achieved by the Markowitz subspace$_{\mathcal{P}}$ mean-variance portfolio approaches asymptotically, in a large market and estimation window, the optimal Sharpe ratio of the entire market.

$$s^2(\hat{\mathbf{w}}^{\mathcal{P}}_{\text{mv},d=K}) = s^2(\mathbf{w}_{\text{mv}}) + O_p(T^{-1/2} + N^{-1/2}). \tag{18}$$

The number of factors for this equation, K, in the Data Generated Process (DGP) is of course unknown. However, Bai and Ng (2002) developed a method to obtain a consistent estimation of the number of factors in the DGP:

$$\hat{K}^{\mathcal{P}} = \underset{1\leq k\leq k_{\max}}{\arg\max} \left\{ \log\left(\sum_{j>k} \hat{\theta}_j^{\mathcal{P}}\right) + \frac{k(N+T)}{NT}\log\left(\frac{NT}{N+T}\right)\right\} \quad \text{with } k_{\max} = 8 \text{ and } \mathcal{P} \in \{\mathcal{P}_{\Sigma}, \mathcal{P}_{Corr}\}. \tag{19}$$

The estimated eigenvalues in Equation 19 are the eigenvalues of $\hat{\Sigma}$ or $\hat{Corr}$. If $\mathbb{P}\{\hat{K}^{\mathcal{P}} = K\} \to 1$ as $N, T \to \infty$, Equation 18 can also be used for $d = \hat{K}^{\mathcal{P}}$.

5 Results

The main result of Section 4 was that the Sharpe ratio, achieved by the subspace$_{\mathcal{P}}$ mean-variance analysis, approaches theoretically the global optimal Sharpe ratio of the entire market, that is calculated via the Markowitz mean-variance portfolio. In this section this will be shown in a practice-oriented way instead of a theoretical way, where simulated and real data sets are used. Because of the great similarities in the simulated results of the two subspaces $\mathcal{P}_{\Sigma}$ and $\mathcal{P}_{Corr}$, only the results for the subspace $\mathcal{P}_{Corr}$ will be given. However, both subspaces will be represented in the empirical illustration.

5.1 Simulation

Stochastic simulations reproduce the behavior of the economy using a mathematical model. The mathematical model in this paper will be a Fama-French three-factor model (Craig MacKinlay and Pástor (2000), DeMiguel et al. (2009) and Tu and Zhou (2011) use the same simulation setup). The economy is represented by a number of risky assets ($N = 25, 100$). The returns of these risky assets are driven by systematic risks that are represented by three simulated research factors. The three factors are simulated from a multivariate normal distribution with the mean and covariance, calibrated from monthly data of July 1963 till August 2007, of the market portfolio, Small Minus Big (SMB) and High Minus Low (HML). The factor loadings needs to be generated from a uniform distribution of a specific interval. These carefully chosen intervals are common in an estimated Fama-French model:

$$\begin{aligned} \text{Loading for the market}: &\quad \beta_{\text{m}} \sim U(0.9, 1.2) \\ \text{Loading for the size}: &\quad \beta_{\text{s}} \sim U(-0.3, 1.4) \\ \text{Loading for the book-to-market}: &\quad \beta_{\text{bm}} \sim U(-0.5, 0.9). \end{aligned}$$

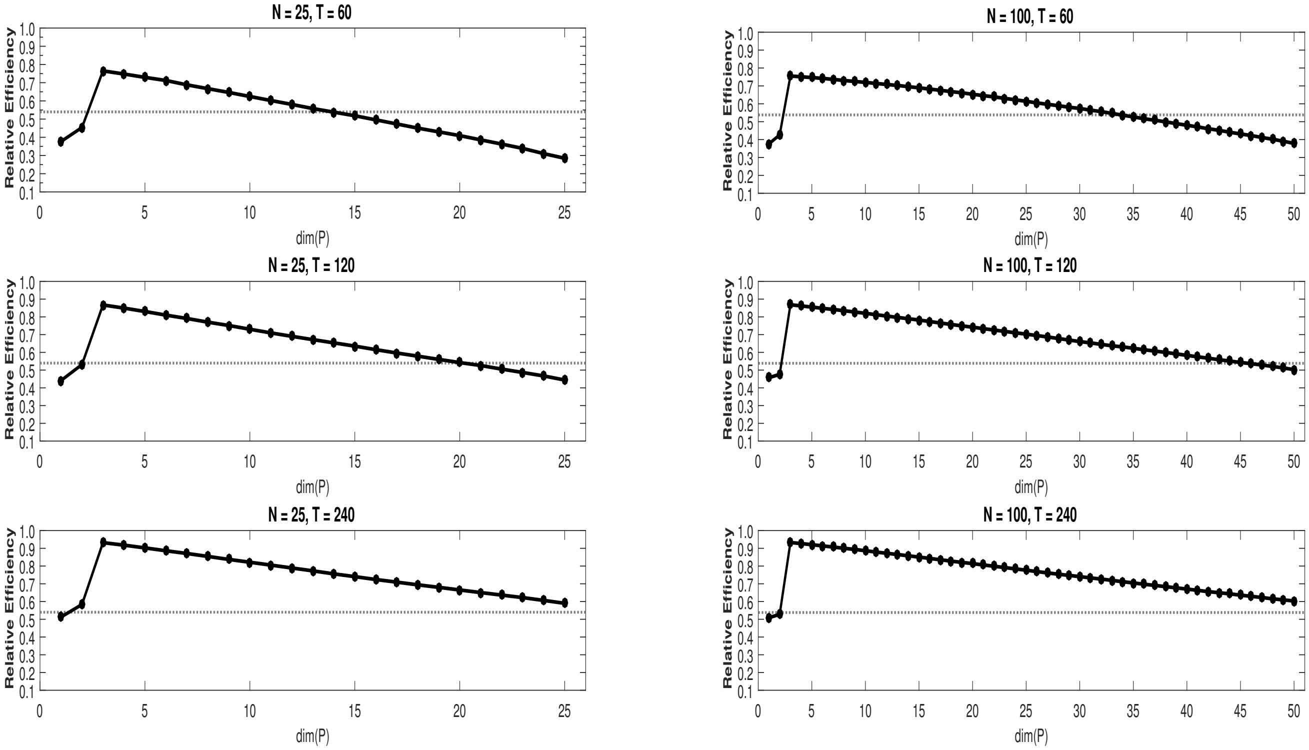

Figure 3: The simulation results are averaged over 1000 data sets that are constructed with a Fama-French three-factor model. The averaging takes place over the relative efficiency, where the relative efficiency equals the ratio $\frac{s(\hat{\mathbf{W}}_{\text{mv},d}^{\mathcal{P}_{Corr}})}{s(\mathbf{W}_{\text{mv}})}$. Each panel shows the relative efficiency of a specific market size ($N = 25, 100$) and estimation window ($T = 60, 120, 240$). All these panels iterate over the dimension of the subspace $d = \dim(\mathcal{P}_{Corr})$. The gray horizontal lines represent the average ratio of Sharpe ratio of the naive diversification over the Sharpe ratio of the true mean-variance portfolio.

First, it is interesting to investigate for which dimension of the subspace $\mathcal{P}_{Corr}$ the Sharpe ratio of the estimated subspace$_{Corr}$ mean-variance portfolio approaches the Sharpe ratio of the true Markowitz mean-variance portfolio. This investigation is done by increasing the dimension d with step sizes of 1 to 25 ($N = 25$) or to 60 ($N = 100$). The main goal of this investigation is to find a dimension that minimizes the summation of the estimation- and systematic error, see Equation 14. The systematic error is simply too large for dimension $d = 1, 2$. So apparently, the selected handful of options from all possible investment options, selected by the subspace $\mathcal{P}_{Corr}$ of dimension one or two, does not contain good enough investment options. For dimension $d > 3$, the estimation error increases rapidly and more intense when the dimension increases. So apparently with a fixed estimation window, the estimation results of the moment conditions (Equation 3) deteriorate more intense when the dimension increases above three. The minimization of Equation 14 occurs for all panels in Figure 3 at dimension $d = 3$. The precise number of factors K in the Data Generated Process is in an empirical example of course unknown, but it is known in a

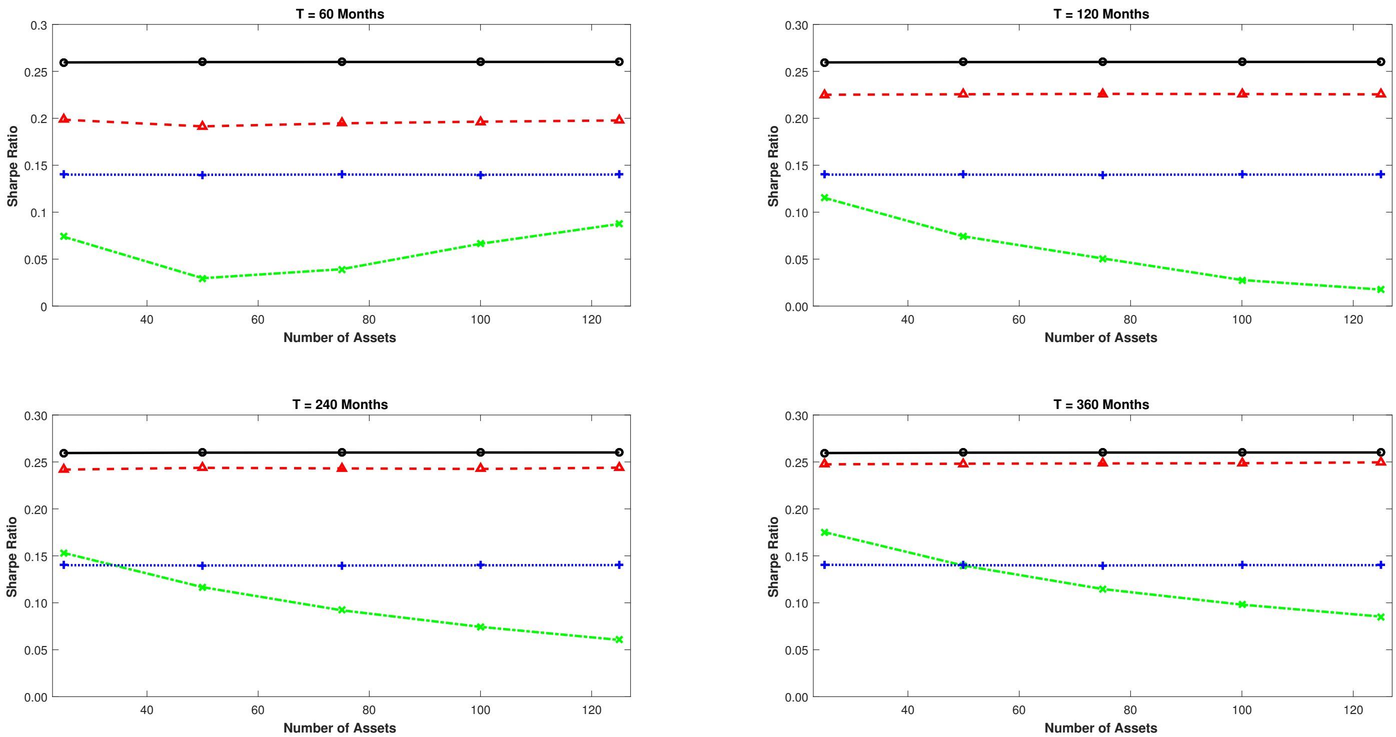

Figure 4: The simulation results are averaged over 1000 data sets that are constructed with a Fama-French three-factor model. The averaging takes place over the Sharpe ratio of the true mean-variance portfolio (circles), the estimated subspace$_{Corr}$ mean-variance portfolio (triangles), the naive diversification (pluses) and the sample mean-variance portfolio (crosses).

simulation example. The returns are generated in this simulation example with three research factors ($K = 3$). Then the practical minimization of Equation 14 at dimension $d = 3$ in Figure 3, can also be theoretically confirmed by Equation 18.

To get an idea of the performance of the estimated subspace$_{Corr}$ mean-variance portfolio in Figure 3, the naive diversification is also reported. There can be seen that the estimated subspace$_{Corr}$ mean-variance portfolio performs better than naive diversification for a wide range of choices for d then only dimension $d = 3$.

Second, it is interesting to compare the motions of different portfolio rules as the number of assets, in which can be invested, increases. The comparison takes place over the Sharpe ratio of the true mean-variance portfolio (normally unobserved), the estimated subspace$_{Corr}$ mean-variance portfolio, the naive diversification and the sample mean-variance portfolio. The number of factors in the DGP, knowing that this is equal to three, is still being estimated using Equation 19. For the estimated subspace$_{Corr}$ mean-variance portfolio is the dimension of the subspace$_{Corr}$ equal to $\hat{K}^{\mathcal{P}_{Corr}}$. It can be seen immediately, from Figure 4, that the estimated subspace$_{Corr}$ mean-variance portfolio outperforms the other portfolio rules for every estimation window and for every number of assets.

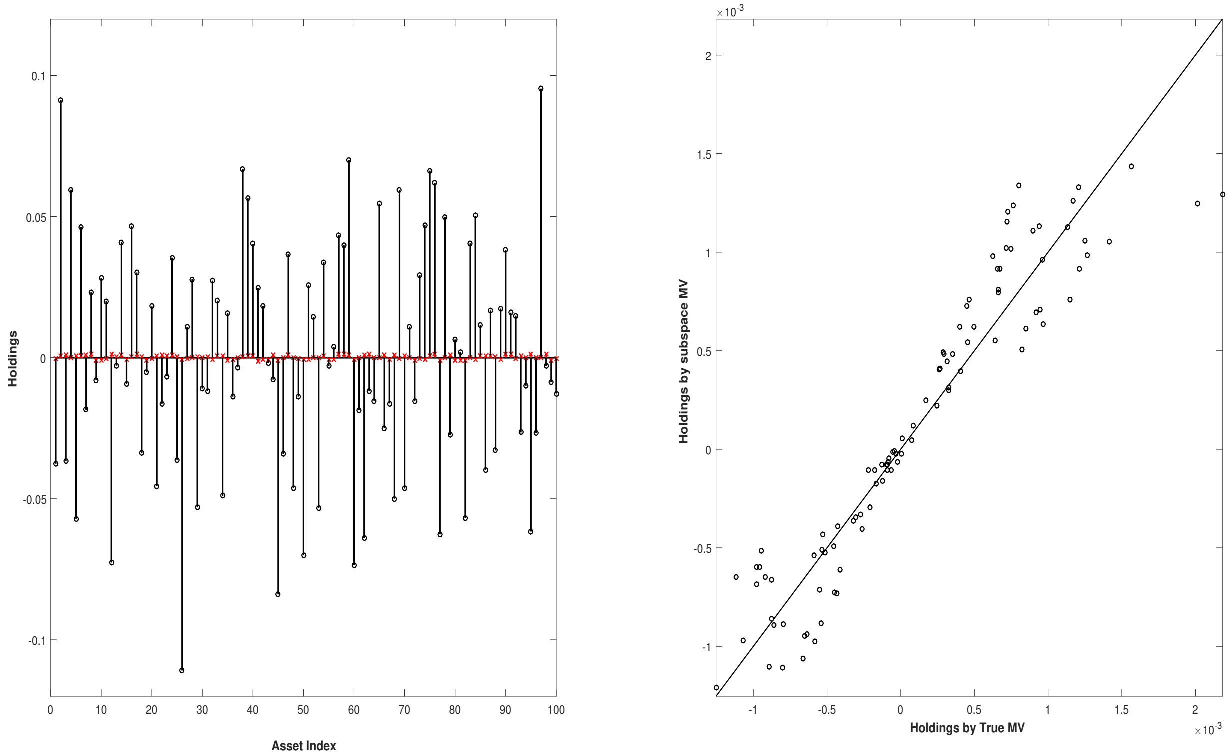

Figure 5: The simulation results are averaged over 1000 data sets that are constructed with a Fama-French three-factor model. Due to a simulation, with an estimation window of $T = 120$ and with $N = 100$ risky assets, is a comparison possible. The left panel compares the holdings of the sample mean-variance portfolio (circles) and the subspace$_{Corr}$ mean-variance portfolio (crosses). The right panel compares the holdings of the mean-variance portfolio (x-axes) and the subspace$_{Corr}$ mean-variance portfolio (y-axes).

Thirdly, a desirable feature of the estimated subspace$_{\mathcal{P}}$ mean-variance portfolio is stability. In the left panel of Figure 5, are the holdings given of the sample- and estimated subspace$_{Corr}$ mean-variance portfolio. The holdings for the subspace$_{Corr}$ mean-variance portfolio have exactly the same scale as the subspace$_{\Sigma}$ mean-variance portfolio, which is again the reason that only one of them is showed in this figure. The conclusion of this panel is, that the scale of short and long investments of the subspace$_{Corr}$ mean-variance portfolio is way smaller than the holding scale of the sample mean-variance portfolio. The left panel compares the holdings of the unobserved true mean-variance portfolio (with the unobserved Σ and E) and the estimated subspace$_{Corr}$ mean-variance portfolio. It can be seen that investments for the true mean-variance portfolio are from a similar scale as investments for the estimated subspace$_{Corr}$ mean-variance portfolio. This can be seen in the right panel of Figure 5, because the scatters are around the 45 degree line.

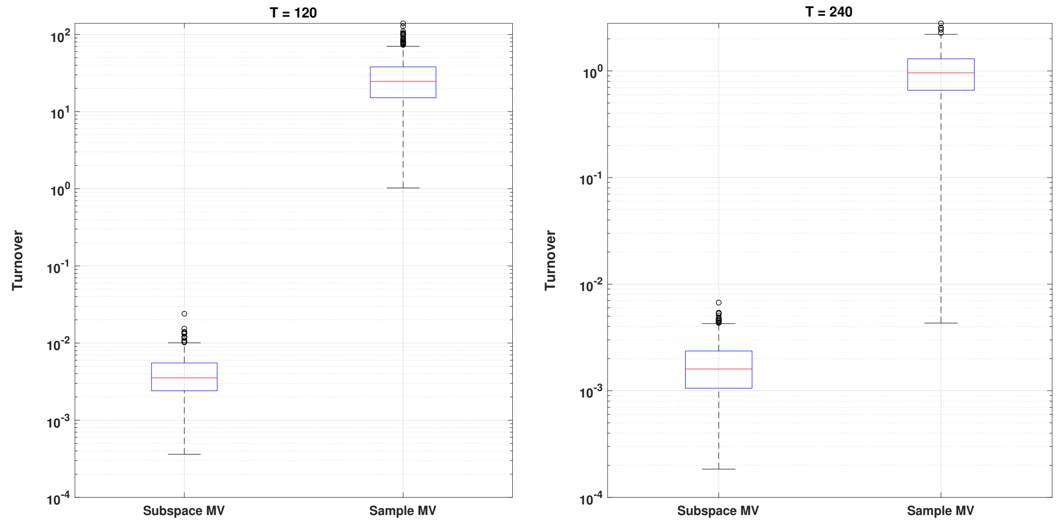

Figure 6: The simulation results of 1 data set that is constructed with a Fama-French three-factor model. A simulation with an estimation window of $T = 120, 240$ and with $N = 100$ risky assets is the rebalancing cost given over 50 years in boxplots for the subspace$_{Corr}$ mean-variance portfolio and the sample mean-variance portfolio.

Fourth and lastly, the rebalancing costs are given over 50 years in a boxplots for the estimated subspace$_{Corr}$ mean-variance portfolio and the sample mean-variance portfolio. The rebalancing costs are of course highly positive correlated with the following turnover variable:

$$\text{Turnover}_t := \sum_{j=1}^{N} |\mathbf{w}_{t+1,j} - \mathbf{w}_{t,j}|. \tag{20}$$

Figure 6 shows that the turnover in 50 years is way larger for the sample mean-variance portfolio than for the estimated subspace$_{Corr}$ mean-variance portfolio. In other words, because of the high correlations mentioned earlier, the rebalancing cost for the sample mean-variance portfolio is over time, on average, more expensive than the rebalancing cost of the estimated subspace$_{Corr}$ mean-variance portfolio.

5.2 Empirical illustration

In this section is represented how the subspace$_{\mathcal{P}}$ mean-variance analysis responds to a non-simulated data set. The data set from January 1961 till December 2010 is used to do the empirical illustration, see Section 3.2 for more explanation. A rolling-window approach is used over this data set to estimate for each window the portfolio composition. More precise, with an estimation window T and N risky assets, the estimated subspace$_{\mathcal{P}}$ mean-variance portfolio will be calculated for each rolling-window $(\boldsymbol{r_{t-T}}, \ldots, \boldsymbol{r_{t-1}})$ and for each subspace. The return of the portfolio will be recorded for each window. Finally, the Sharpe ratio can be calculated with the expectation and standard deviation of the portfolio excess returns.

5.2.1 The relationship between the Sharpe ratio and dim($\mathcal{P}$)

First, it is interesting to look at the relationship between the Sharpe ratio of the estimated subspace$_{\mathcal{P}}$ mean-variance portfolio and the dimension of the subspaces $\mathcal{P}_{\Sigma}$ and $\mathcal{P}_{Corr}$. This is done in Figure 7 by increasing the dimension d with step sizes of 1 for different estimation windows and different numbers of risky assets.

The main goal of using two different numbers of risky assets is to compare the performs of the estimated subspace$_{\mathcal{P}}$ mean-variance analysis in a small- and large market. The Fama-French 100 portfolios, that are formed on size and book-to-market, represent a large market. However, six out of the 100 portfolios contain missing values. The missing values are not randomly distributed across time but are distributed within sub-samples. Because of the intentional separation of a simulation (Section 5.1) and empirical illustration (this section), the missing data is not imputed/simulated but the six portfolios are removed from the data set. This is done to get no simulated values into the empirical illustration. Namely, the goal to compare a small- and a large market is also reached with a comparison of 25 and 94 assets.

On average, for most subspace dimensions with a given T and N, the estimated subspace$_{\Sigma}$ mean-variance portfolio gives better Sharpe ratios then the estimated subspace$_{Corr}$ mean-variance portfolio. So clearly, the handful of investment options captured by the subspace $\mathcal{P}_{\Sigma}$ contains on average better portfolio compositions than the handful of investment options captured by the subspace $\mathcal{P}_{Corr}$.

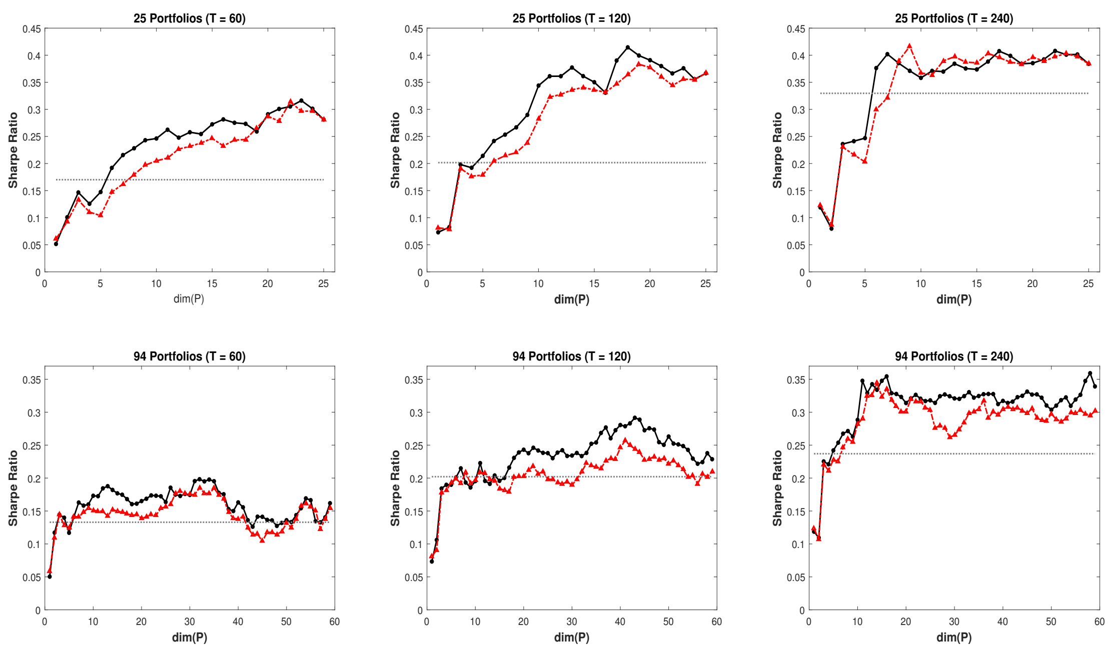

Figure 7: The empirical Sharpe ratios of the estimated subspace$_{\Sigma}$ mean-variance portfolio (circles) and the estimated subspace$_{Corr}$ mean-variance portfolio (triangles). There is an iteration over the dimension of the subspace $\mathcal{P}$ for the estimation windows $T = 60, 20, 240$ and the market sizes $N = 25, 94$. The gray horizontal line represent the Sharpe ratio of the estimated subspace$_{\Sigma}$ mean-variance portfolio, with the dimension estimated using Equation 19.

5.2.2 The efficient frontier

To confirm the statement in Section 5.2.1, that the subspace$_\Sigma$ performs better in the subspace$_\mathcal{P}$ mean-variance analysis than the subspace$_{Corr}$, the efficient frontier is estimated in Figure 8. The efficient portfolios in a risk-return framework are showed in this figure. These efficient portfolios are defined as the portfolios that minimize the risk subject to a given expected return for both subspaces, so to minimize Equation 5 with E and Σ the mean and variance of the return respectively. For multiple values of gamma, the gamma in the risk tolerance factor, the efficient portfolio will be estimated. This gamma increases from 0.7 to 6, using step sizes of 0.02, with a fixed estimation window T and number of risky assets N. The most common tolerance factors (Best and Grauer, 1991) are in the chosen interval of gamma.

Immediately visible from Figure 8, is that the lowest risk for a specific level of expected return is achieved by the estimated subspace$_\Sigma$ mean-variance. So these data confirmed that the subspace$_\Sigma$ gives higher Sharpe ratios in the Markowitz subspace$_\mathcal{P}$ mean-variance analysis than the subspace$_{Corr}$.

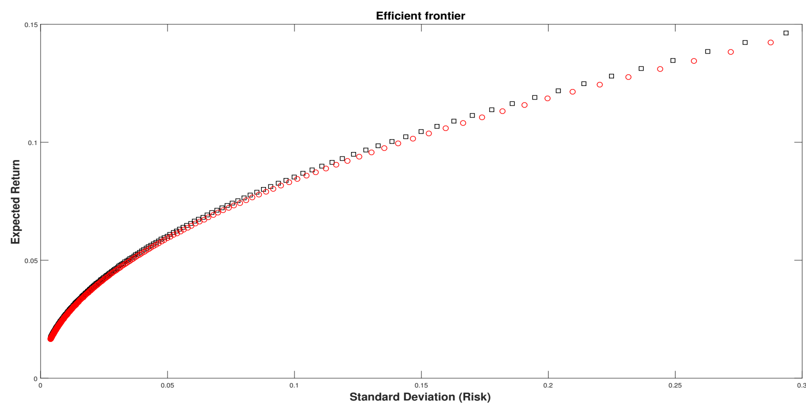

Figure 8: The estimated subspace$_\mathcal{P}$ mean-variance portfolio, for each given $\gamma \in [0.7, 6]$ and each subspace $\mathcal{P}_\Sigma$ (squares) and $\mathcal{P}_{Corr}$ (circles), gives an indication of the efficient frontier.

5.2.3 Outperformance due to a specific investment option

Theoretically, the subspace$_\Sigma$ mean-variance portfolio should give higher Sharpe ratios than the subspace$_{Corr}$ mean-variance portfolio. However, when a specific investment option is added to the investment market, the subspace $\mathcal{P}_{Corr}$ outperforms $\mathcal{P}_\Sigma$ in the subspace$_\mathcal{P}$ mean-variance analysis. This is caused by a simulated investment option with a significantly smaller standard deviation than the other portfolio options. The left panel of Figure 9 shows the motions of the subspace$_\mathcal{P}$ mean-variance portfolios as the number of investment options increases, without the simulated option. The right panel shows the motions of the subspace$_\mathcal{P}$ mean-variance portfolios with a simulated option. This simulated investment option, the 61 added investment option, has a significantly smaller standard deviation than the other investment options. The closer the standard deviation of the investment option approaches zero, the larger the outperformance of $s\left(\hat{\mathbf{w}}_{\text{mv},d}^{\mathcal{P}_{Corr}}\right)$ relative to $s\left(\hat{\mathbf{w}}_{\text{mv},d}^{\mathcal{P}_\Sigma}\right)$!

The occurrence of the overall fall of the Sharpe ratio by adding an additional investment option in Figure 9, is explained in Section 4.2. Because logically, by adding an additional investment opportunity, the Sharpe ratio should increase or remain equal due to diversification. Apparently this is not the case for the subspace$_{\mathcal{P}}$ mean-variance analysis (see Section 4.2).

The large outperformance of the subspace$_{Corr}$ mean-variance portfolio at the 61$^{\text{th}}$ addition, is especially due to the relative differences of the fourth leading eigenvector of the correlation- and covariance matrix. This fourth eigenvector can be roughly interpreted as the return differences between the portfolios with an extreme market equity- and ratio of book equity to market equity score and the portfolios with an average market equity- and ratio of book equity to market equity score. The value at the 61$^{\text{th}}$ position of the fourth leading eigenvector of the correlation matrix differs strongly compared with the other values of the eigenvector. This phenomenon does not occur at the 61$^{\text{th}}$ position of the fourth leading eigenvector of the covariance matrix. The large 61$^{\text{th}}$ value of the fourth eigenvector of the correlation matrix has a large impact on the holding for the 61$^{\text{th}}$ investment option of the subspace$_{Corr}$ mean-variance portfolio. The result of this large value is a large holding for this investment option, that is calculated via Equation 12 with the subspace $\mathcal{P}_{Corr}$, whereas the 61$^{\text{th}}$ holding for the covariance matrix is almost zero. To observe the big differences between the two subspaces, the dimension of the subspace needs to be larger or equal to four.

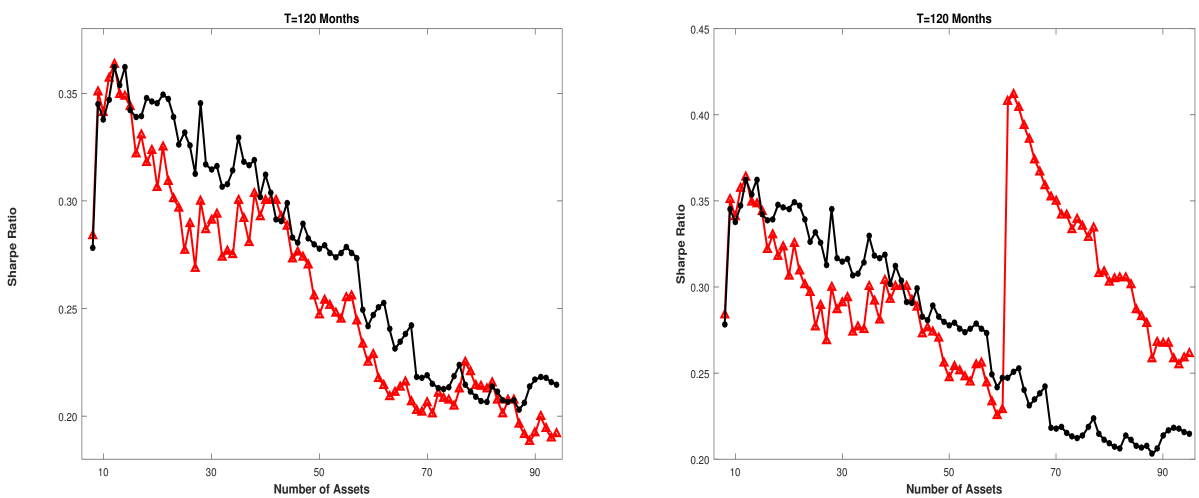

Figure 9: The empirical Sharpe ratios of the estimated subspace$_{\Sigma}$ mean-variance portfolio (circles) and the estimate subspace$_{Corr}$ mean-variance portfolio (triangles). There is an iteration over the number of investment options for a random selected estimation window $T = 120$.

5.2.4 Performance comparison

To demonstrate the performance of the estimated subspace$_{\mathcal{P}}$ mean-variance portfolio, several other portfolio rules are applied to these data. Firstly, more than 40% of all estimated $\hat{K}^{\mathcal{P}_{\Sigma}}$, via Equation 19, are smaller than or equal to four and even more than 50% for $\hat{K}^{\mathcal{P}_{Corr}}$. Because of these low estimated values, not all information needed to make a good composition, will be caught by the first three or four leading eigenvectors. This is why the last two columns are added where the dimension of the two subspaces, for the estimated subspace$_{\mathcal{P}}$ mean-variance portfolios, are equal to $k_{\max}$ ($k_{\max} = 8$). Secondly, other added columns to table 1 are the portfolio rule of Jorion (PJ), the rule of Kan and Zhou (KZ), the combination rule based on the sample mean-variance portfolio and naive diversification of Tu and Zhou (S&N) and the combination rule based on KZ and naive diversification of Tu and Zhou (KZ&N). To compare the performance of the different portfolio rules, the Sharpe ratio is given in Table 1 with different estimation windows T and different numbers of assets N. The number of assets needs to be smaller than the estimation window ($N < T$) for the portfolio rules mentioned above. Clearly, the estimated subspace$_{\mathcal{P}}$ mean-variance analysis outperforms the other portfolio rules in a large market with a big enough estimation window.

Table 1: The comparison of the Sharpe ratio between the estimated subspace$_{\mathcal{P}}$ mean-variance portfolio and several other portfolio rules over a period of 50 years.

Data	T	Naive	PJ	S&N	KZ	KZ&N	$\hat{\mathbf{w}}_{\text{mv}}$	$\hat{\mathbf{w}}^{\mathcal{P}_\Sigma}_{\text{mv},d=\hat{K}^{\mathcal{P}_\Sigma}}$	$\hat{\mathbf{w}}^{\mathcal{P}_{Corr}}_{\text{mv},d=\hat{K}^{\mathcal{P}_{Corr}}}$	$\hat{\mathbf{w}}^{\mathcal{P}_\Sigma}_{\text{mv},d=k_{\max}}$	$\hat{\mathbf{w}}^{\mathcal{P}_{Corr}}_{\text{mv},d=k_{\max}}$
25 Portf.	60	0.13	0.29	0.28	0.29	0.29	0.28	0.17	0.14	0.23	0.18
	120	0.14	0.38	0.37	0.38	0.37	0.37	0.20	0.21	0.27	0.22
	240	0.15	0.38	0.39	0.38	0.38	0.38	0.33	0.28	0.39	0.39
94 Portf.	60	0.13	-	-	-	-	0.16	0.13	0.14	0.16	0.15
	120	0.14	0.22	0.23	0.21	0.22	0.21	0.20	0.18	0.19	0.21
	240	0.15	0.20	0.17	0.19	0.20	0.17	0.24	0.23	0.27	0.26

6 Conclusion

The number and properties of risky assets determine which portfolio rules are optimal to use. Most portfolio rules can be used properly in a small market. However, the performance of these portfolio rules worsens when the size of the market increases due to the large number of estimated parameters. A reduction of the estimated parameters forms a portfolio rule that even can be used in a large market. This portfolio rule is the Markowitz subspace$_{\mathcal{P}}$ mean-variance analysis. The properties of these risky assets determine which subspace is optimal to use for this portfolio rule, a subspace spanned by the leading eigenvectors of the correlation matrix ($\mathcal{P}_{Corr}$) or of the covariance matrix ($\mathcal{P}_\Sigma$). However, a criticism on this rule is that an extra investment opportunity not necessarily results in a better or equivalent Sharpe ratio, this phenomenon occurs due to the selected subspace. Also a reliable method that obtains a consistent estimation of the number of factors in the DGP is needed for this rule, which simultaneously also determines the number of estimated parameters in the subspace$_{\mathcal{P}}$ mean-variance analysis. Thus, even in a large market, the Markowitz subspace$_{\mathcal{P}}$ mean-variance analysis is a reliable portfolio rule if a large, but realistic, amount of historical data is available and a well-chosen subspace is used, which is dependent on the properties of the risky assets.

References

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

[2] Best, M. J. and Grauer, R. R. (1991), On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results, *Review of Financial Studies* 4(2), 315–342.

[3] Chen, J. and Yuan, M. (2016), Efficient portfolio selection in a large market, *Journal of Financial Econometrics* 14(3), 496–524.

[4] Connor, G., Goldberg, L. R. and Korajczyk, R. A. (2010), *Portfolio risk analysis*, Princeton University Press.

[5] Craig MacKinlay, A. and Pástor, L. (2000), Asset pricing models: Implications for expected returns and portfolio selection, *The Review of financial studies* 13(4), 883–916.

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

[7] Fama, E. F. and French, K. R. (1993), Common risk factors in the returns on stocks and bonds, *Journal of financial economics* 33(1), 3–56.

[8] Jensen, M. C. (1968), The performance of mutual funds in the period 1945–1964, *The Journal of finance* 23(2), 389–416.

[9] Jorion, P. (1986), Bayes-stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis* 21(03), 279–292.

[10] Kan, R. and Zhou, G. (2007), Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis* 42(03), 621–656.

[11] Markowitz, H. (1952), Portfolio selection, *The journal of finance* 7(1), 77–91.

[12] Markowitz, H. (1959), *Portfolio Selection, Efficent Diversification of Investments*, J. Wiley.

[13] Sharpe, W. F. (1966), Mutual fund performance, *The Journal of business* 39(1), 119–138.

[14] Tu, J. and Zhou, G. (2011), Markowitz meets talmud: A combination of sophisticated and naive diversification strategies, *Journal of Financial Economics* 99(1), 204–215.

[15] Wold, S., Esbensen, K. and Geladi, P. (1987), Principal component analysis, *Chemometrics and intelligent laboratory systems* 2(1-3), 37–52.

Appendices

A Assumption factor model

The following assumptions are also made in Chen and Yuan (2016), and the proofs are found in the in the appendix of Chen and Yuan (2016).

Assumption A (Factors): The factors have finite fourth moments such that there exist a positive constant $C_1 < \infty$ satisfying

$$\max_{1 \leq k \leq K} \mathbb{E}\left(f_{kt}^4\right) \leq C_1.$$

Assumption B (Factors Loadings): There is a strictly positive matrix Σ_B such that

$$\frac{B^T B}{N} \to \Sigma_B \quad \text{as } N \to \infty.$$

Assumption C (Idiosyncratic Risks): The covariance matrix of $\boldsymbol{\epsilon}_t = \left[\epsilon_{1t}, \ldots, \epsilon_{Nt}\right]^T$, Σ_ϵ has eigenvalues bounded away from both zero and infinity. Moreover, the idiosyncratic risks have finite fourth moments such that there exists a positive constant $C_2 < \infty$ satisfying

$$\max_{1 \leq j \leq N} \mathbb{E}\left[\epsilon_{jt}^4\right] \leq C_2.$$

Assumption D: The factors $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_T$ are weakly dependent in that there exist a positive constant $C_3 < \infty$ satisfying

$$\max_{1 \leq k_1,\ k_2 \leq K} \mathbb{E}\left\{\sum_{t=1}^{T} (f_{k_1 t} f_{k_2 t} - \Sigma^f_{k_1 k_2})^2\right\} \leq C_3 T \quad \text{where } \Sigma^f_{k_1 k_2} \text{ is the } (k_1, k_2)\text{th entry of } \Sigma_f.$$

Assumption E: The idiosyncratic risks $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_T$ are weakly dependent in that there exists a positive constant $C_4 < \infty$ satisfying

$$\max_{1 \leq j_1,\ j_2 \leq K} \mathbb{E}\left\{\sum_{t=1}^{T} (\epsilon_{j_1 t} \epsilon_{j_2 t} - \Sigma^\epsilon_{j_1 j_2})^2\right\} \leq C_4 T \quad \text{where } \Sigma^\epsilon_{j_1 j_2} \text{ is the } (j_1, j_2)\text{th entry of } \Sigma_\epsilon.$$

Assumption F: The factors and idiosyncratic risks are jointly weakly dependent in that there exists a positive constant $C_5 < \infty$ satisfying

$$\max_{1\leq j\leq N,\ 1\leq k\leq K} \mathbb{E}\left\{\sum_{t=1}^{T} \epsilon_{jt}^2 f_{kt}^2\right\} \leq C_5 T.$$

B Proof Equation 7

The portfolio composition with a constraint is given as the solution to the Markowitz mean-variance analysis subject to one constraint:

$$\begin{gathered} \min_{\mathbf{w}^*\in\mathbb{R}^N} \left\{\frac{\gamma}{2}\mathbf{w}^{*T}\Sigma\mathbf{w}^* - \mathbf{w}^{*T}E\right\} \\ \text{s.t. } \mathbf{w}^{*T}\iota = 1 \end{gathered}$$

To incorporate this restriction into one equation, the auxiliary function will be introduced and the respective gradient:

$$\begin{aligned} \mathcal{L}(\mathbf{w}^*, \lambda) &= \frac{\gamma}{2}\mathbf{w}^{*T}\Sigma\mathbf{w}^* - \mathbf{w}^{*T}E - \lambda\left(\mathbf{w}^{*T}\iota - 1\right) \\ \nabla_{\mathbf{w}^*,\lambda}\mathcal{L}(\mathbf{w}^*, \lambda) = 0 &\Rightarrow \begin{cases} \frac{\partial\mathcal{L}}{\partial\mathbf{w}^*} = \ \gamma\Sigma\mathbf{w}^* - E - \lambda\iota = 0 & (1) \\ \frac{\partial\mathcal{L}}{\partial\lambda} = \ -\mathbf{w}^{*T}\iota + 1 = 0 & (2) \end{cases} \end{aligned}$$

So the solution of the minimization is given by:

$$\begin{aligned} (1) : \mathbf{w}^* &= \frac{1}{\gamma}\Sigma^{-1}E + \frac{\lambda}{\gamma}\Sigma^{-1}\iota \\ \text{Substitute (1) into (2)} : \frac{1}{\gamma}E^T\Sigma^{-1}\iota &+ \frac{\lambda}{\gamma}\iota^T\Sigma^{-1}\iota = 1 \\ \Leftrightarrow \ \lambda &= \frac{\gamma - E^T\Sigma^{-1}\iota}{\iota^T\Sigma^{-1}\iota} \\ \text{Substitute } \lambda \text{ into (1)} \quad : \mathbf{w}^* &= \frac{1}{\gamma}\Sigma^{-1}E + \frac{1}{\gamma}\frac{\gamma - E^T\Sigma^{-1}\iota}{\iota^T\Sigma^{-1}\iota}\Sigma^{-1}\iota \\ \Rightarrow \mathbf{w}^*_{\text{mv}} &:= \mathbf{w}_{\text{mv}} + \frac{1}{\gamma}\left(\frac{\gamma - E^T\Sigma^{-1}\iota}{\iota^T\Sigma^{-1}\iota}\right)\Sigma^{-1}\iota \end{aligned}$$

C Proof Equation 8

The portfolio composition, within a subspace and with a constraint, is given as the solution to the Markowitz subspace mean-variance analysis subject to one constraint:

$$\min_{\mathbf{w}_d^{\mathcal{P}*} \in \mathcal{P}} \left\{ \frac{\gamma}{2} \mathbf{w}_d^{\mathcal{P}*T} \Sigma \mathbf{w}_d^{\mathcal{P}*} - \mathbf{w}_d^{\mathcal{P}*T} E \right\} \quad \text{where} \quad \mathcal{P} \in \{\mathcal{P}_\Sigma, \mathcal{P}_{Corr}\}$$
$$\text{s.t. } \mathbf{w}_d^{\mathcal{P}*T} \iota = 1$$

To give the solution of the Markowitz subspace mean-variance portfolio subject to this constraint, some notation needs to be introduced as $P_{\mathcal{P}}$ is a $N \times d$ matrix whose columns are an orthonormal basis of P. Additional, for $\mathbf{x}_{\mathcal{P}} \in \mathbb{R}^d$ and $P_{\mathcal{P}}\mathbf{x}_{\mathcal{P}} \in \mathcal{P}$, the map $\mathbb{R}^d \to \mathcal{P} : \mathbf{x}_{\mathcal{P}} \mapsto P_{\mathcal{P}}\mathbf{x}_{\mathcal{P}} = \mathbf{w}_d^{\mathcal{P}*}$ is injective and surjective. To incorporate this restriction into one equation, the auxiliary function will be introduced and the respective gradient:

$$\mathcal{L}(\mathbf{x}_{\mathcal{P}}, \lambda) = \frac{\gamma}{2} \left(P_{\mathcal{P}}\mathbf{x}_{\mathcal{P}}\right)^T \Sigma \left(P_{\mathcal{P}}\mathbf{x}_{\mathcal{P}}\right) - \left(P_{\mathcal{P}}\mathbf{x}_{\mathcal{P}}\right)^T E - \lambda \left(\left(P_{\mathcal{P}}\mathbf{x}_{\mathcal{P}}\right)^T \iota - 1 \right)$$

$$\nabla_{\mathbf{x}_{\mathcal{P}},\lambda} \mathcal{L}(\mathbf{x}_{\mathcal{P}}, \lambda) = 0 \Rightarrow \begin{cases} \frac{\partial \mathcal{L}}{\partial \mathbf{x}_{\mathcal{P}}} = \gamma P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}} \mathbf{x}_{\mathcal{P}} - P_{\mathcal{P}}^T E - \lambda P_{\mathcal{P}}^T \iota = 0 & (1) \\ \frac{\partial \mathcal{L}}{\partial \lambda} = -\mathbf{x}_{\mathcal{P}}^T P_{\mathcal{P}}^T \iota + 1 = 0 & (2) \end{cases}$$

So the solution of the minimization is given by:

$$(1) : \mathbf{x}_{\mathcal{P}} = \frac{1}{\gamma} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T E + \frac{\lambda}{\gamma} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T \iota$$

$$\text{Substitute (1) into (2)} : \frac{1}{\gamma} E^T P_{\mathcal{P}} \left[\left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} \right]^T P_{\mathcal{P}}^T \iota + \frac{\lambda}{\gamma} \iota^T P_{\mathcal{P}} \left[\left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} \right]^T P_{\mathcal{P}}^T \iota = 1$$

$$\Leftrightarrow \lambda = \frac{\gamma - E^T P_{\mathcal{P}} \left[\left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} \right]^T P_{\mathcal{P}}^T \iota}{\iota^T P_{\mathcal{P}} \left[\left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} \right]^T P_{\mathcal{P}}^T \iota}$$

$$\text{Substitute } \lambda \text{ into (1)} \quad : \mathbf{x}_{\mathcal{P}} = \frac{1}{\gamma} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T E + \frac{1 - \frac{1}{\gamma} E^T P_{\mathcal{P}} \left[\left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} \right]^T P_{\mathcal{P}}^T \iota}{\iota^T P_{\mathcal{P}} \left[\left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} \right]^T P_{\mathcal{P}}^T \iota} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T \iota$$

Thus:

$$\mathbf{w}_{\mathrm{mv},d}^{\mathcal{P}\,*} := P_{\mathcal{P}}\mathbf{x}_{\mathcal{P}} = \frac{1}{\gamma} P_{\mathcal{P}} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T E + \frac{1 - \frac{1}{\gamma} E^T P_{\mathcal{P}} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T \iota}{\iota^T P_{\mathcal{P}} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T \iota} P_{\mathcal{P}} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T \iota$$

$$\Rightarrow \mathbf{w}_{\mathrm{mv},d}^{\mathcal{P}\,*} := \mathbf{w}_{\mathrm{mv},d}^{\mathcal{P}} + \frac{1 - \frac{1}{\gamma} \iota^T P_{\mathcal{P}} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T E}{\iota^T P_{\mathcal{P}} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T \iota} P_{\mathcal{P}} \left(P_{\mathcal{P}}^T \Sigma P_{\mathcal{P}}\right)^{-1} P_{\mathcal{P}}^T \iota$$

D Proof Equation 11

Knowing that $D_\sigma = \begin{bmatrix} e_{\sigma_1} & \cdots & e_{\sigma_N} \end{bmatrix} = \begin{bmatrix} e_{\sigma_1}^T \\ \vdots \\ e_{\sigma_N}^T \end{bmatrix}$, the invertible matrix I can be written as:

$$
\begin{aligned}
I &= P_{\mathcal{P}_{Corr}}^T D_\sigma \ P_{Corr} \ D_{Corr} \ P_{Corr}^T \ D_\sigma \ P_{\mathcal{P}_{Corr}} \\
&= \begin{bmatrix} \eta_1^T \\ \vdots \\ \eta_d^T \end{bmatrix} \begin{bmatrix} e_{\sigma_1} & \cdots & e_{\sigma_N} \end{bmatrix} P_{Corr} D_{Corr} P_{Corr}^T \begin{bmatrix} e_{\sigma_1}^T \\ \vdots \\ e_{\sigma_N}^T \end{bmatrix} \begin{bmatrix} \eta_1 & \cdots & \eta_d \end{bmatrix} \\
&= \begin{bmatrix} \eta_1^T e_{\sigma_1} & \cdots & \eta_1^T e_{\sigma_N} \\ \vdots & & \vdots \\ \eta_d^T e_{\sigma_1} & \cdots & \eta_d^T e_{\sigma_N} \end{bmatrix} P_{Corr} D_{Corr} P_{Corr}^T \begin{bmatrix} e_{\sigma_1}^T \eta_1 & \cdots & e_{\sigma_1}^T \eta_d \\ \vdots & & \vdots \\ e_{\sigma_N}^T \eta_1 & \cdots & e_{\sigma_N}^T \eta_d \end{bmatrix} \\
&= \begin{bmatrix} \eta_1^T e_{\sigma_1} & \cdots & \eta_1^T e_{\sigma_N} \\ \vdots & & \vdots \\ \eta_d^T e_{\sigma_1} & \cdots & \eta_d^T e_{\sigma_N} \end{bmatrix} \begin{bmatrix} \eta_1 & \cdots & \eta_N \end{bmatrix} D_{Corr} \begin{bmatrix} \eta_1^T \\ \vdots \\ \eta_N^T \end{bmatrix} \begin{bmatrix} e_{\sigma_1}^T \eta_1 & \cdots & e_{\sigma_1}^T \eta_d \\ \vdots & & \vdots \\ e_{\sigma_N}^T \eta_1 & \cdots & e_{\sigma_N}^T \eta_d \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} \eta_1^T e_{\sigma_1} & \cdots & \eta_1^T e_{\sigma_N} \end{bmatrix} \eta_1 & \cdots & \begin{bmatrix} \eta_1^T e_{\sigma_1} & \cdots & \eta_1^T e_{\sigma_N} \end{bmatrix} \eta_N \\ \vdots & & \vdots \\ \begin{bmatrix} \eta_d^T e_{\sigma_1} & \cdots & \eta_d^T e_{\sigma_N} \end{bmatrix} \eta_1 & \cdots & \begin{bmatrix} \eta_d^T e_{\sigma_1} & \cdots & \eta_d^T e_{\sigma_N} \end{bmatrix} \eta_N \end{bmatrix} D_{Corr} \begin{bmatrix} \eta_1^T \begin{bmatrix} e_{\sigma_1}^T \eta_1 \\ \vdots \\ e_{\sigma_N}^T \eta_1 \end{bmatrix} & \cdots & \eta_1^T \begin{bmatrix} e_{\sigma_1}^T \eta_d \\ \vdots \\ e_{\sigma_N}^T \eta_d \end{bmatrix} \\ \vdots & & \vdots \\ \eta_N^T \begin{bmatrix} e_{\sigma_1}^T \eta_1 \\ \vdots \\ e_{\sigma_N}^T \eta_1 \end{bmatrix} & \cdots & \eta_N^T \begin{bmatrix} e_{\sigma_1}^T \eta_d \\ \vdots \\ e_{\sigma_N}^T \eta_d \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} \theta_1 \begin{bmatrix} \eta_1^T e_{\sigma_1} & \cdots & \eta_1^T e_{\sigma_N} \end{bmatrix} \eta_1 & \cdots & \theta_N \begin{bmatrix} \eta_1^T e_{\sigma_1} & \cdots & \eta_1^T e_{\sigma_N} \end{bmatrix} \eta_N \\ \vdots & & \vdots \\ \theta_1 \begin{bmatrix} \eta_d^T e_{\sigma_1} & \cdots & \eta_d^T e_{\sigma_N} \end{bmatrix} \eta_1 & \cdots & \theta_N \begin{bmatrix} \eta_d^T e_{\sigma_1} & \cdots & \eta_d^T e_{\sigma_N} \end{bmatrix} \eta_N \end{bmatrix} \begin{bmatrix} \eta_1^T \begin{bmatrix} e_{\sigma_1}^T \eta_1 \\ \vdots \\ e_{\sigma_N}^T \eta_1 \end{bmatrix} & \cdots & \eta_1^T \begin{bmatrix} e_{\sigma_1}^T \eta_d \\ \vdots \\ e_{\sigma_N}^T \eta_d \end{bmatrix} \\ \vdots & & \vdots \\ \eta_N^T \begin{bmatrix} e_{\sigma_1}^T \eta_1 \\ \vdots \\ e_{\sigma_N}^T \eta_1 \end{bmatrix} & \cdots & \eta_N^T \begin{bmatrix} e_{\sigma_1}^T \eta_d \\ \vdots \\ e_{\sigma_N}^T \eta_d \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} I(1,1) & \cdots & I(1,d) \\ \vdots & & \vdots \\ I(d,1) & \cdots & I(d,d) \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\text{with } I(i,j) &= \sum_{k=1}^{N} \theta_k \left(\begin{bmatrix} \eta_i^T e_{\sigma_1} & \eta_i^T e_{\sigma_2} & \cdots & \eta_i^T e_{\sigma_N} \end{bmatrix} \eta_k \right) \left(\eta_k^T \begin{bmatrix} e_{\sigma_1}^T \eta_j \\ e_{\sigma_2}^T \eta_j \\ \vdots \\ e_{\sigma_N}^T \eta_j \end{bmatrix} \right) \\
&= \sum_{k=1}^{N} \theta_k \left(\begin{bmatrix} \eta_i^T e_{\sigma_1} & \eta_i^T e_{\sigma_2} & \cdots & \eta_i^T e_{\sigma_N} \end{bmatrix} \eta_k \right) \left(\eta_k^T \begin{bmatrix} \eta_j^T e_{\sigma_1} \\ \eta_j^T e_{\sigma_2} \\ \vdots \\ \eta_j^T e_{\sigma_N} \end{bmatrix} \right) \\
&= \sum_{k=1}^{N} \theta_k \left(\begin{bmatrix} \eta_i^T e_{\sigma_1} & \eta_i^T e_{\sigma_2} & \cdots & \eta_i^T e_{\sigma_N} \end{bmatrix} \eta_k \right) \left(\begin{bmatrix} \eta_j^T e_{\sigma_1} & \eta_j^T e_{\sigma_2} & \cdots & \eta_j^T e_{\sigma_N} \end{bmatrix} \eta_k \right) \quad \text{for } i,j = 1,\ldots,d
\end{aligned}
$$