ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

Efficient Portfolio Selection in a Large Market

Herbert Hop	Student (id 413423)
Sander Barendse	Supervisor
Bart Keijsers	First Reader

Erasmus University Rotterdam, Erasmus School of Economics

Econometrics and Operations Research

Bachelor Thesis, July 2, 2017

Abstract

The Markowitz estimated mean-variance portfolio is proven to not work well in markets in which the number of assets is relatively high compared to the number of historical observations, which can be attributed to the large estimation error in the sample estimates of the expected return and covariance matrix. Chen & Yuan (2016) introduce a method that reduces the investment universe to the subspace, spanned by the leading eigenvectors of the sample covariance matrix. In this paper, it is shown that the sample mean and covariance matrix can be well used in this so-called subspace mean-variance analysis. These restricted subspace mean-variance portfolios theoretically outperform the usual unrestricted Markowitz mean-variance portfolios, even when more sophisticated, shrinkage estimators are used to estimate the mean and variance. However, empirical analysis does not support these observations. It is shown that in the real market, the subspace method does not necessarily outperform the estimated mean-variance porfolio which is why its practical application is not unchallenged.

1 Introduction

Although the mean-variance portfolio strategy, following from the analysis of Markowitz [7] [8], theoretically guarantee the highest Sharpe ratio, this strategy does not work very well in practice. In fact, those mean-variance portfolios, whose weights are estimated by using the sample first and second moment of the returns, cannot consistently outperform portfolios, based on naive diversification including assigning equal weights to every possible asset. The reason behind this lack of performance can be attributed to the estimation error of the moments of the returns, used to estimate the expected return and the covariance matrix of the returns respectively. Especially when the number of assets, thus the number of parameters to be estimated, grows substantially compared to the number of available historical observations, the estimation error becomes too large, leading to substantial underperformance of the Markowitz mean-variance portfolios.

Chen & Yuan [2] propose a solution to this problem by substantially decreasing the number of parameters to be estimated. By restricting the investment universe to the subspace, spanned by the leading principle components of the sample covariance matrix of the returns, they show that the Sharpe ratio of the estimated subspace mean-variance portfolio approaches that of the global mean-variance portfolio when the number of historical observations and the number of assets increase, because the loss of efficiency due to the restricted investment options is compensated by the gains from smaller estimations errors. Based on both simulation analysis and empirical illustrations they conclude that these proposed subspace mean-variance portfolios perform consistently better than the naive diversification portfolios and the sample Markowitz mean-variance portfolios. Additionally, the subspace mean-variance portfolios can compete very well with alternative, well-established portfolio rules.

The sample mean covariance matrix is a consistent estimator of the true mean and covariance matrix, but suffer from huge estimation errors, especially when the number of assets increases relatively to the number of historical observations. Possible alternatives include shrinkage estimators of the expected return and covariance matrix. Shrinkage estimation attempts to find a balance between the variance and the bias of the estimator. The idea is to shrink the unbiased sample estimator to a shrinkage target, which bears bias, but is numerically more efficient to estimate. Ledoit and Wolf [6] propose a shrinkage estimator which shrinks the sample covariance matrix to a scalar multiple of the identity matrix. Their shrinkage estimator is guaranteed to be invertible in markets in which the number of assets outnumber the historical observations, which is why it is used in this application. Besides, shrinkage estimators for the expected return and covariance matrix, based on the Bayes-Stein estimators, are evaluated, as discussed by Jorion [4].

Simulation analysis shows that the sample moments can be well used as estimators for the mean and covariance matrix when applying the subspace method. With the unrestricted investment universe, more sophisticated estimates of the expected return and covariance matrix result in higher Sharpe ratios, compared to the usual sample mean-variance portfolio. Still, the subspace method outperforms all estimated shrinkage mean-variance portfolios, from which it can be concluded that the estimation error is well captured by restricting the investment universe. Empirical analysis however, shows that the subspace mean-variance portfolios do not necessarily perform better than the estimated mean-variance portfolios, which can possibly be explained by the mass of external effects in real markets.

2 Literature

When the dimension of the covariance matrix is larger than the number of observations available, the traditional sample covariance matrix is known to be not invertible. Even when the number of number of observations exceeds the dimension, the sample covariance matrix is invertible, but suffers from huge estimations errors. In particular, De Miguel et al. [3] show that the Markowitz mean-variance portfolios cannot consistently outperform naive diversified portfolios when the expected return and covariance matrix are estimated using the sample moments.

Several solutions have been proposed in recent literature, all of which have the same goal, namely to impose more structure on the estimator of the expected return and the covariance matrix, which makes it better conditioned and reduces estimation error. One possibility is to assume a K-factor model for the returns, so the expected return and covariance matrix can be estimated from this model. It holds that the lesser the number of factors, the more structure is imposed on the estimators. However, Connor and Korajczyk [1] show that there is no consensus on the number of factors to choose.

Shrinkage estimation has been proposed in recent literature for many applications in portfolio management. This statistical technique is dating back to Stein [10]. Shrinkage estimators aim to shrink the numerically ill-conditioned estimator to a well-conditioned estimator. Ledoit and Wolf [5] provide a shrinkage estimator which shrinks the sample covariance matrix to the covariance matrix obtained from the single-index, one-factor model of Sharpe [9]. Because the shrinkage intensity is determined by minimizing a quadratic loss function, which does not need the inverse of the covariance matrix but is measured by the distance between the shrinkage estimator of the covariance and the true covariance matrix by means of the Frobenius norm, it makes their shrinkage estimation applicable in markets in which the number of assets exceeds the number of observations. An extension is developed by Ledoit and Wolf [6] with an estimator which shrinkage the sample covariance matrix to a scalar multiple of the identity matrix, implying a constant variance matrix with zero covariances. This estimator is also applicable in markets with the dimension of the covariance matrix larger than the number of observations for the same reason. Other shrinkage estimators have been proposed in literature. Well known are the Bayes-Stein estimators. Jorion [4] developed a general Bayes-Stein estimator using utility functions. For a given utility function, the mean-variance optimal portfolio is shrunk towards the minimum variance portfolio. This shrinkage method results in different mean and covariance estimators, which are described in more detail later.

3 Methodology

3.1 Restricting Investment Universe

Let $\mathbf{r} \in \mathbb{R}^N$ be the return of N risky assets in excess of the risk-free rate. Then the Markowitz global mean-variance efficient portfolio is given by solving the equation below:

$$\min_{\mathbf{w} \in \mathbf{R}^N} \frac{\gamma}{2} \mathbf{w}^T \Sigma \mathbf{w} - \mathbf{w}^T E \tag{1}$$

with $E \in \mathbb{R}^N$ and $\Sigma \in \mathbb{R}^{N \times N}$ denoting the mean and variance of $\mathbf{r}$ respectively and with γ being the coefficient of relative risk aversion. The solution of this problem is given by:

$$\mathbf{w}_{\mathrm{mv}} = \frac{1}{\gamma} \Sigma^{-1} E \tag{2}$$

The Sharpe ratio of any portfolio rule with weights $\mathbf{w}$

$$s(\mathbf{w}) = \frac{\mathbf{w}' E}{(\mathbf{w}' \Sigma \mathbf{w})^{\frac{1}{2}}} \tag{3}$$

which is maximal for the theoretical mean-variance portfolio $\mathbf{w}_{mv}$. Because both the population mean E and variance Σ of $\mathbf{r}$ are unknown, those should be estimated from historical data.

Following the Chen & Yuan, the investment universe is limited to a linear subspace of $\mathbb{R}^N$, namely $\mathcal{P}$. The optimization problem now becomes:

$$\min_{\mathbf{w} \in \mathcal{P}} \frac{\gamma}{2} \mathbf{w}^T \Sigma \mathbf{w} - \mathbf{w}^T E \tag{4}$$

When the subspace is chosen such that the first d eigenvectors of the covariance matrix of $\mathbf{r}$ form a basis of $\mathcal{P}$, the solution of the corresponding optimization problem is given as follows:

$$\hat{\mathbf{w}}_d = \frac{1}{\gamma} \sum_{k=1}^{d} \theta_k^{-1} \eta_k \eta_k^T E \tag{5}$$

where θ_k denotes the k-th eigenvalue of Σ and η_k its corresponding eigenvector.

As both the population mean and variance are unknown, all parameters in the above

sum have to be estimated from historical data.
Bai and Ng (2002) propose an information criterion to estimate the number of leading eigenvectors to use, d:

$$\hat{d} = \underset{1 \leq k \leq k_{max}}{\operatorname{argmin}} \left\{ \log \left(\sum_{j>k} \hat{\theta}_k \right) + \frac{k(N+T)}{NT} \log \left(\frac{NT}{N+T} \right) \right\} \tag{6}$$

This information criterion minimizes the variance of the returns, which is explained by the eigenvectors that are omitted, while taking into account a penalty for the number of eigenvectors to include. In the analysis, k_{max} is set equal to 8. Given the estimate of d, the optimal portfolio weights of the subspace mean-variance portfolio can be obtained according to equation (5).

3.2 Improved Moment Estimators

Better estimates of the covariance matrix of the excess returns reduce the necessity of restricting the investment universe. Several models to model mean and variance are analyzed. First of all, the sample covariance matrix is shrunk towards a shrinkage target with slight bias but lesser variance in order to reduce the estimation error. Secondly, the Bayes-Stein estimates of the mean and variance are investigated.

3.2.1 Linear Shrinkage Estimation

The shrinkage estimation of the covariance matrix of the excess returns is based on the idea of Ledoit and Wolf. Because the number of observations is too small, compared to the number of assets, the sample covariance matrix suffers from huge estimation errors. The goal is to find an estimator which is well-conditioned for large-dimensional covariance matrices. The identity matrix as an example of an estimator of the covariance matrix will contain small estimation errors, but will be heavily biased. Therefore, Ledoit and Wolf attempt to find a convex, linear combination between a scalar multiple of the identity matrix and the sample covariance matrix. The optimal weight to allocate to the different covariance matrix estimators is determined according to a quadratic loss function. Mathematically, this can be written down as follows:

$$\begin{aligned} \underset{\rho_1,\rho_2}{\text{minimize}} \quad & E[\|\Sigma^* - \Sigma\|^2] \\ \text{subject to} \quad & \Sigma^* = \rho_1 I + \rho_2 S \end{aligned} \tag{7}$$

with $E[\|\Sigma^* - \Sigma\|^2]$ denoting the expected value of the squared Euclidean distance between the shrinkage estimator of the covariance matrix Σ^* and the true covariance matrix Σ, commonly referred to as mean squared error. S and I denote the sample covariance matrix and identity matrix respectively. To give an interpretation of the shrinkage estimator with respect to the quadratic loss function, it is useful to decompose the mean squared error of the shrinkage estimator in variance and squared bias:

$$E[\|\Sigma^* - \Sigma\|^2] = E[\|\Sigma^* - E[\Sigma^*]\|^2] + \|E[\Sigma^*] - \Sigma\|^2 \tag{8}$$

The shrinkage target will contain no variance and only squared bias. For the sample covariance matrix, the opposite holds: as the estimator is an unbiased estimator of the covariance, the squared bias will be zero, while the variance will be substantially positive due to large estimation errors. The idea is to find the optimal trade-off between variance and squared bias such that the mean squared error of the shrinkage estimator is minimal, which is the solution of the linear optimization problem, given above. Ledoit and Wolf show that:

$$\begin{aligned} \Sigma^* &= \frac{\beta^2}{\delta^2}\mu I + \frac{\alpha^2}{\delta^2} S, \\ E[\|\Sigma^* - \Sigma\|^2] &= \frac{\alpha^2\beta^2}{\delta^2} \end{aligned} \tag{9}$$

with μI denoting the shrinkage target. It can be proven that:

$$\mu = \langle \Sigma, I \rangle, \quad \alpha^2 = \|\Sigma - \mu I\|^2, \quad \beta^2 = E[\|S - \Sigma\|^2], \quad \delta^2 = E[\|S - \mu I\|^2]$$

Besides, the following property holds:

$$\alpha^2 + \beta^2 = \delta^2 \tag{10}$$

to make sure that the linear combination between the sample covariance matrix S and the shrinkage target μI is convex. To obtain the optimal shrinkage intensity, μ, α^2, β^2 and δ^2 need to be known. These are all scalar functions of the true covariance matrix Σ, which mean they are in fact unknown. However, asymptotically there exists consistent estimators for these four unknown scalar functions of Σ. Standard asymptotics, implying that the ratio of the number of assets over the number of observation converges to zero, do not hold. Therefore general asymptotics need to be considered, which means that the number of assets converges at the same speed as the number of observations, which is reasonable in large markets. Under standard asymptotics, the sample covariance matrix is a consistent estimator, which means that the shrinkage intensity vanishes, while under general asymptotics, it converges to a certain constant. Therefore, the point of interest is to consistently estimate the shrinkage estimator under general asymptotics.

Basically, the idea of general asymptotics is as follows. Let $t = 1, 2, \ldots$ be the number of historical observations. Then X_t is an $p_t \times t$ matrix of observations with p_t denoting the number of assets. X_t is assumed to be observed from a system of p_t random variables with mean zero and covariance matrix Σ_t. Recall that p_t cannot converge to infinity while t doesn't. Given t, the sample covariance matrix S_t can be written as $t^{-1}X_tX_t'$. The same procedure can be followed for the four scalar functions:

$$\mu_t = \langle \Sigma_t, I \rangle_t, \quad \alpha_t^2 = \|\Sigma_t - \mu_t I_t\|_t^2, \quad \beta_t^2 = E[\|S_t - \Sigma_t)\|_t^2], \quad \delta_t^2 = E[\|S_t - \mu_t I_t\|_t^2], \tag{11}$$

all of which remain bounded when t grows to infinity. Ledoit and Wolf show that although Σ_t cannot be consistently estimated, the optimal shrinkage target can. Even the shrinkage estimator Σ_t^* can be estimated consistently. First of all, μ_t can be consistently estimated by its so called sample counterpart:

$$m_t = \langle S_t, I \rangle_t \tag{12}$$

The same holds for δ_t^2:

$$d_t = \|S_t - m_t I_t\|_t^2 \tag{13}$$

Let x_k^t be the kth column of X_t. Define $\bar{b}_t^2 = \frac{1}{t^2} \sum_{k=1}^{t} \|x_k^t (x_k^t)' - S_t\|_t^2$.
β_t^2 can be consistently estimated by

$$b_t^2 = \min(\bar{b}_t^2, d_t^2) \tag{14}$$

α_t^2 is then estimated by

$$a_t^2 = d_t^2 - b_t^2 \tag{15}$$

which is in accordance with equation (10). Finally it can be shown that the estimator

$$S_t^* = \frac{b_t^2}{d_t^2} m_t I_t + \frac{a_t^2}{d_t^2} S_t \tag{16}$$

is a consistent estimator of Σ_t^* from which it can be concluded that under general asymptotics, the optimal shrinkage intensity can be consistently estimated. Extensive proof on all results can be found in Ledoit and Wolf [6].

3.2.2 Bayes-Stein Estimation

Below the Bayes-Stein estimators for the mean and variance of $\mathbf{r}$ are described. These can also be found in Chen & Yuan.

$$E_{BS} = (1 - v)\hat{E} + v\hat{E}_g \mathbf{1}$$

and

$$\hat{\Sigma}_{BS} = \left(1 + \frac{1}{T + \lambda}\right) \tilde{\Sigma} + \frac{\lambda}{T(T + 1 + \lambda)} \frac{\mathbf{1}\mathbf{1}'}{\mathbf{1}'\tilde{\Sigma}^{-1}\mathbf{1}}$$

Where

$$\hat{E}_g = \frac{\mathbf{1}'\hat{\Sigma}^{-1}\hat{E}}{\mathbf{1}'\hat{\Sigma}^{-1}\mathbf{1}}, \quad v = \frac{N + 2}{N + 2 + T(\hat{E} - \hat{E}_g\mathbf{1})'\tilde{\Sigma}^{-1}(\hat{E} - \hat{E}_g\mathbf{1})}$$

and

$$\tilde{\Sigma} = \frac{T-1}{T-N-2}\hat{\Sigma}, \quad \lambda = \frac{N+2}{(\hat{E}-\hat{E}_g\mathbf{1})'\tilde{\Sigma}^{-1}(\hat{E}-\hat{E}_g\mathbf{1})}$$

E_{BS} is a shrinkage estimator of the expected return which shrinks the sample mean to the weights return of the minimum variance portfolio multiplied by the sample mean. The shrinkage coefficient is denoted by v. The Bayes-Stein variance estimator $\hat{\Sigma}_{BS}$ is a linear combination of the biased variance estimator $\tilde{\Sigma}$ and the $N \times N$ matrix of ones, scaled by a specific quadratic form of the inverse of $\tilde{\Sigma}^{-1}$. λ is the parameter determining the weights for the different matrices.

4 Results

In this section, the results for the replication of Chen & Yuan and the results on the extension on their research are shown. The first two sections correspond with the replication, the third section corresponds with the extension.

4.1 Simulation Results

Recall that $\mathbf{r}$ is an N dimensional vector of returns in excess of the risk-free rate. For the simulation analysis, the Fama-French three-factor model for the excess returns is assumed:

$$r_{j,t} = \beta_1[r_{M,t}-r_{f,t}]+\beta_2 SMB_t+\beta_3 HML_t+\epsilon_{j,t} \qquad j=1,\ldots,N; \quad t=1,\ldots,T \tag{17}$$

The factors are sampled from a multivariate normal distribution with mean and covariances extracted from July 1963 to August 2007 monthly data on the three factors (available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The factor loadings are sampled from a uniform distribution between 0.9 and 1.2, -0.3 and 1.4 and -0.5 and 0.9 for the factors of the market portfolio, the Small-Minus-Big portfolio and the High-Minus-Low portfolio respectively. The error covariance matrix Σ_ϵ is assumed to be diagonal with elements uniformly sampled between 0.1 and 0.3. Firstly, the effect of the dimension d is shown. Below the results are shown for N = 25 and 100 assets for interval lengths T = 60, 120 and 240 months.

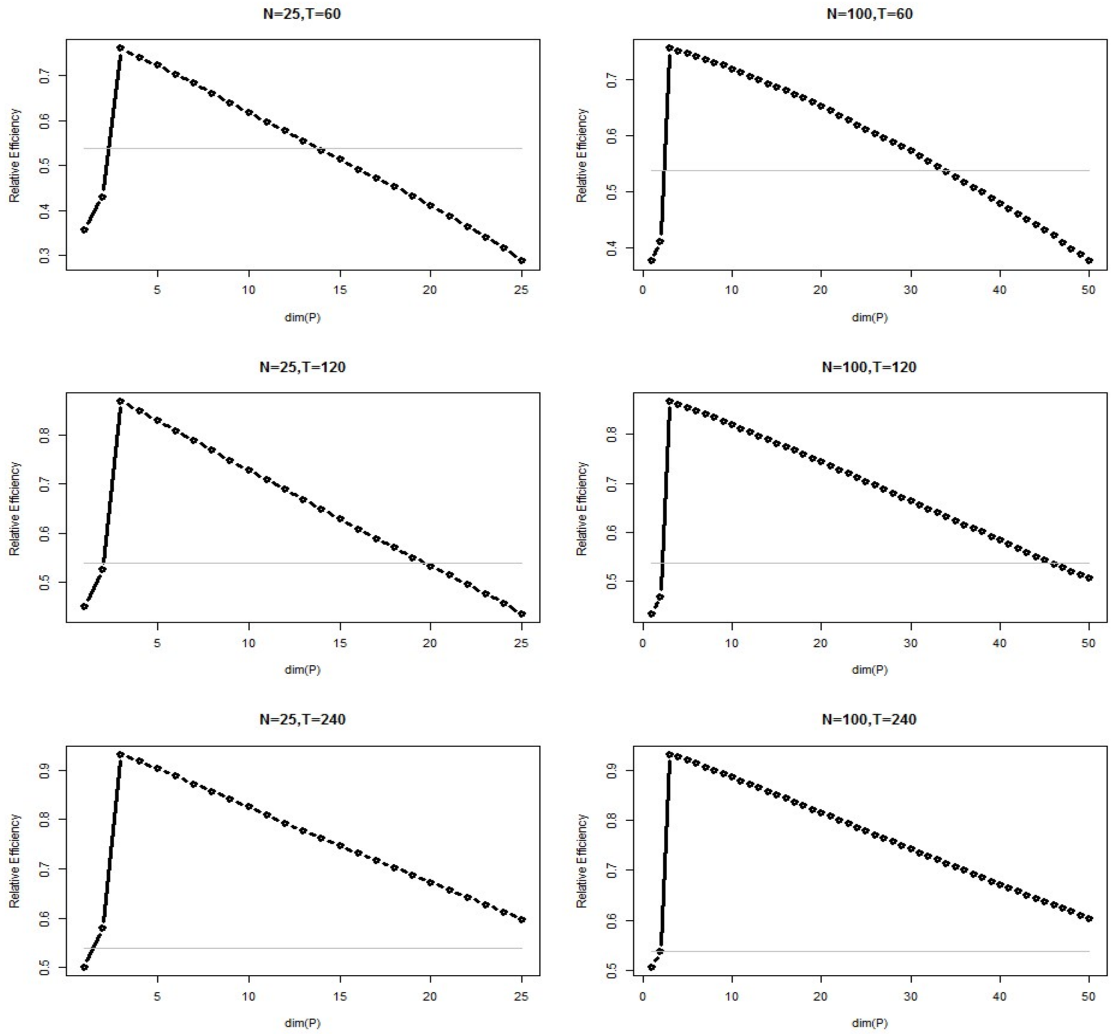

Figure 1: Simulation results from a three-factor model. For each combination of market size N=25 or 100, and estimation window T=60, 120, or 240 months, the relative efficiency, measured by the ratio between the Sharpe ratio relative to that of the true mean–variance portfolio, of $\hat{\mathbf{w}}_d$ is reported here for different choices of d. The results are averaged over 1000 simulated data sets for each plot. The grey horizontal lines correspond to the averaged relative efficiency for the naive diversification.

Clearly, the relative efficiency is maximal for $d = 3$, which is in accordance with the assumption that the returns are generated according with a three-factor model. Furthermore, it can be seen that the the subspace mean-variance portfolio provides significant better Sharpe ratios when choosing d not too small or too large. For $d = 1$ or 2, the first d eigenvectors cannot capture enough information on the covariance matrix and for too large d, the estimation error becomes too large. In the figure 2, the Sharpe ratio is plotted for different number of assets. Four time intervals are considered: $T = 60$, 120, 240 and 360 months. When, the number of assets increases, the estimation errors increases dramatically for the sample mean-variance portfolios. Clearly, the estimated subspace mean-variance portfolio outperforms the sample mean-variance portfolio and the portfolio of naive diversification. Allthough it is not clearly visible in the figure, it appears that the estimated subspace mean-variance portfolio converges slowly to the true mean-variance portfolio as T increases. For increasing N and T, the difference between the true mean-variance portfolio and the estimated subspace mean-variance portfolio declines.

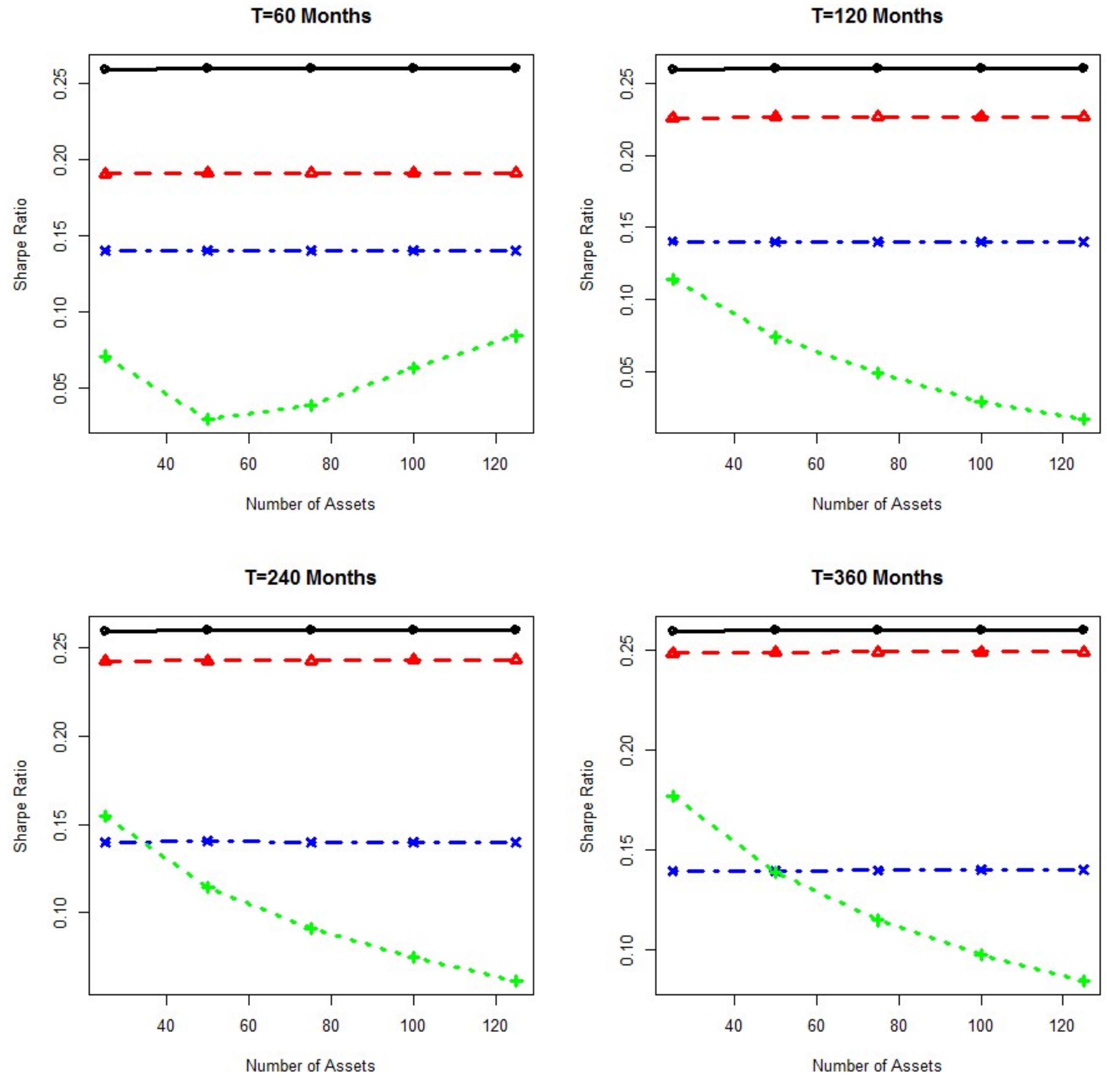

Figure 2: Simulation results based on three-factor model: Data were simulated from a three-factor model. Sharpe ratio of the true mean–variance portfolio (circles), estimated subspace mean–variance portfolio (triangles), sample mean–variance portfolio (pluses) and naive portfolio rule (crosses) are presented for different estimating window size and market size. Note that the true mean–variance portfolio is not feasible and it is added for reference only.

To address the stability of the weights of the subspace mean-variance portfolio, below the weights of the estimated subspace mean-variance portfolio together with the weights of the sample mean-variance portfolio and the true mean-variance portfolio for the case of N=100 assets and T=120 months historical data are depicted.

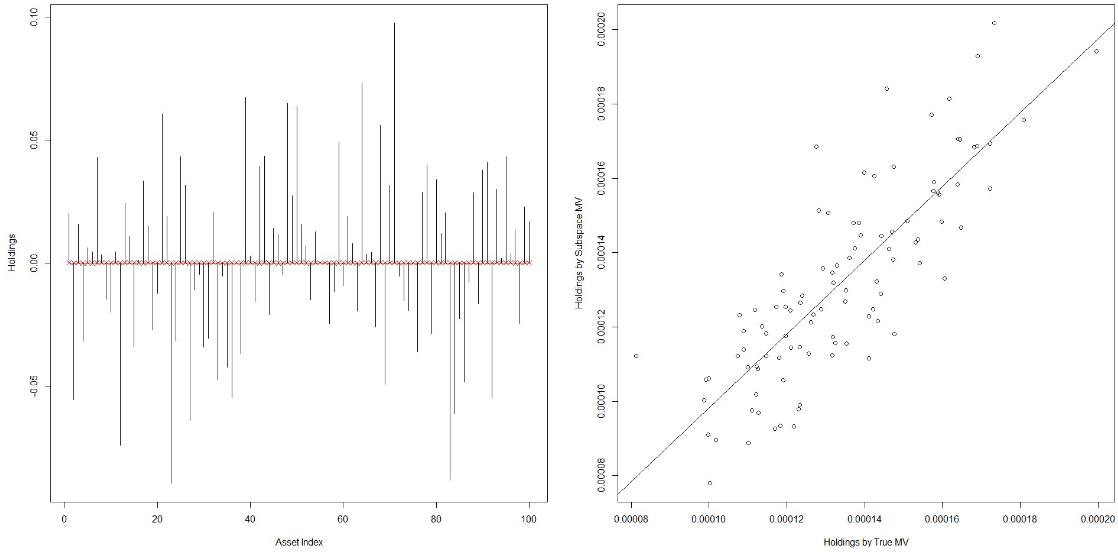

Figure 3: Stability — the left panel compares the holdings of the estimated mean–variance portfolio, represented by the black lines, and subspace mean–variance portfolio, represented by the red crosses. The right panel compares the holdings of the estimated subspace mean–variance portfolio with those of the true global mean–variance portfolio.

Clearly the results in figure 3 show that the portfolio holdings of the subspace mean-variance portfolio are much more stable than those of the sample mean-variance portfolio. Besides, the holdings of the subspace mean-variance portfolio have a nearly one-to-one relation with the weights of the true mean-variance portfolio. To further examine the stability, the monthly rebalancing required by both the subspace mean-variance portfolios and the sample mean-variance portfolios have been evaluated. The rebalancing costs can be calculated by the sum of absolute change in weights over the N assets between two subsequent months, as mathematically shown below.

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

The results of these rebalancing for T=120 months and T=240 months are shown in figure 4. Due to the log scale of the Y-axes of the box plots, the weights of the subspace mean-variance portfolio seem to be less stable than the weights of the sample mean-variance portfolio, which is misleading, because removing the log scale shows very clear the stableness of the subspace mean-variance portfolios, contrary to the sample mean-variance portfolios.

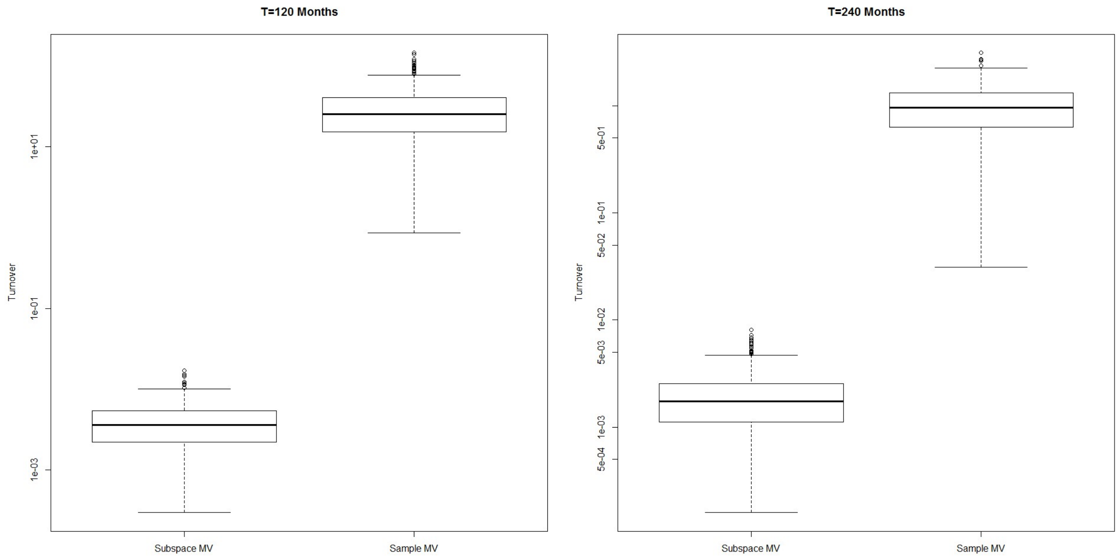

Figure 4: Rebalancing cost — boxplots of the monthly rebalancing cost for the subspace mean–variance portfolio and sample mean–variance portfolio with 100 assets over a period of 50 years. The left panel corresponds to an estimating window of 120 months whereas the right panel 240 months. The Y-axes in both panels are in log scale for better contrast between the two portfolio rules.

4.2 Empirical Illustrations

The setup for the empirical section follows from Chen & Yuan. The used data consists of a 50-year period (from January 1961 to December 2010) of returns for $N = 25$ and 94 portfolios. The data consists of the Fama-French 25 (5×5) and 100 (10×10) portfolios, containing equally weighted returns for the intersections of size and book-to-market portfolios. In case, a portfolio contains missing data, it is erased from the data set. The time frames $T = 60$, 120 and 240 are evaluated. For every N and T, a rolling-window approach is used. For every $t = T, \ldots, 600 - T$, the returns are calculated using observations $\mathbf{r}_{t-T}, \ldots, \mathbf{r}_{t-1}$. The Sharpe ratio is then calculated by dividing the sample mean of the $600 - T$ observations by their sample standard deviation. In figure 5 the results are reported. The Sharpe ratios are plotted for different d's, while the dashed red line corresponds with $\dim(\mathcal{P})$ calculated by means of information criterion (6).

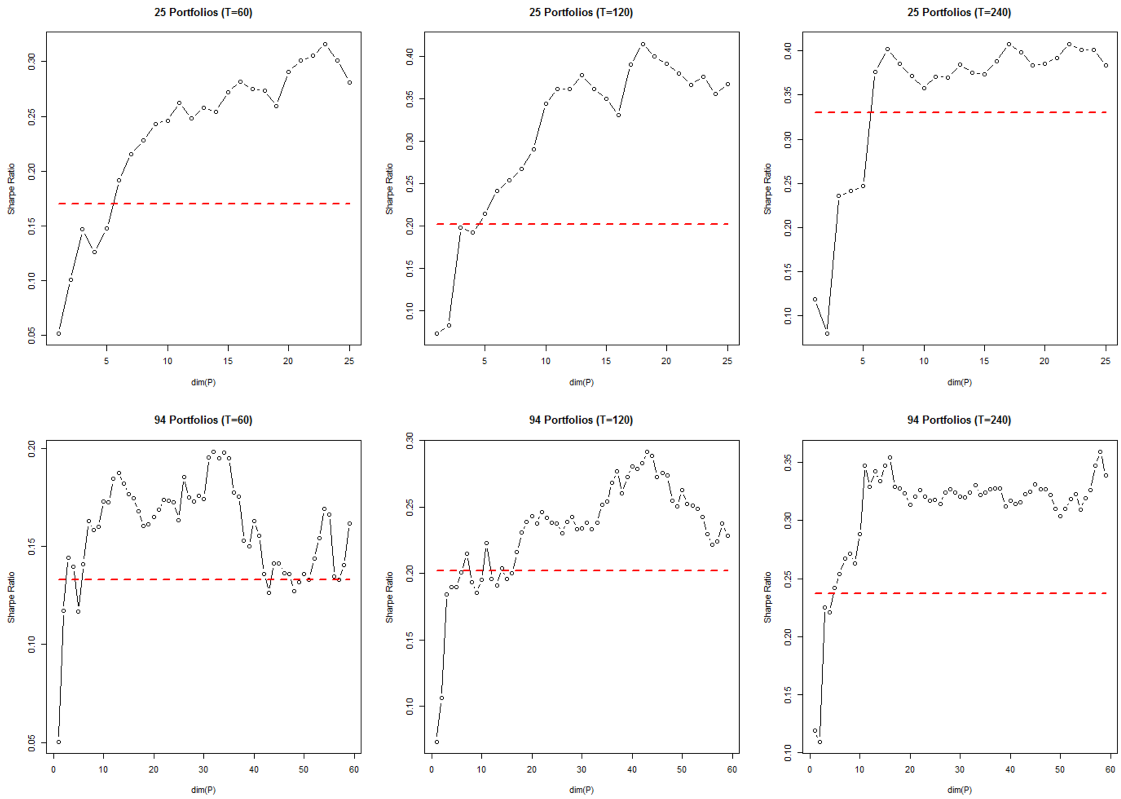

Figure 5: Fama–French portfolio examples — historical performance of $\hat{\mathbf{w}}_d$ for different choices of d over a 50 year period. The dashed horizontal lines represent the Sharpe ratio of the estimated subspace mean–variance portfolio with dim $(\mathcal{P})$ determined using information criterion (6).

Contrary to the simulation analysis, the results are less stable. This can be attributed to external effects which have substantial influence on the returns. Still, it can be seen that for low choices of d, too less variance is explained by the eigenvectors. However, due to external effects, the effects of the estimation error cannot be very clearly extracted from the figure.

Several other portfolio rules are evaluated. These are the three-fund rule of Jorion (1986), the rule of Kan and Zhou (2007), the combination of the sample mean-variance portfolio and the portfolio of naive diversification from Tu and Zhou (2011) and the combination of the portfolio Kan and Zhou and the portfolio of naive diversification, developed by Tu and Zhou (2011). These are abbreviated in respective order by PJ, KZ, S&N and KZ&N. The dimension of $\mathcal{P}$ for the estimated subspace mean-variance portfolio is determined using the discussed information criterion (6).

Table 1: Comparison between the estimated subspace mean–variance portfolio and several other popular alternatives on the Fama–French data sets

Data	T	Sample	Naive	S&N	KZ	KZ&N	PJ	Subspace MV
25 Portfolios	60	0.28	0.13	0.28	0.29	0.28	0.29	0.17
	120	0.37	0.14	0.37	0.38	0.37	0.38	0.20
	240	0.38	0.15	0.39	0.38	0.38	0.38	0.33
100 Portfolios	60	0.16	0.13	NA	NA	NA	NA	0.13
	120	0.21	0.14	0.23	0.21	0.22	0.22	0.20
	240	0.17	0.15	0.17	0.19	0.20	0.20	0.24

The subspace mean-variance portfolio always outperforms the naive diversification portfolio. For $N = 25$, it holds that the subspace mean-variance portfolio performs worse than the other portfolio rules, beside naive diversification. For $N = 100$, it becomes clear that the larger estimation errors result lower Sharpe ratios for the sample mean-variance portfolio. However, for the subspace mean-variance portfolio, the same holds. In fact, for N equals 100, the subspace mean-variance portfolio displays a higher Sharpe ratio only for $T = 240$, which is not expected. The subspace mean-variance portfolio compares favorably well with the other 4 portfolio rules when $N = 100$. In short, based on empirical data, the subspace mean-variance portfolio is not able to consistently outperform the sample mean-variance portfolio.

4.3 Improved Moment Estimators

In this section, the results on the extension of the paper of Chen & Yuan are discussed. First of all, the performance of the subspace mean-variance portfolios making use of the sample moments are compared to the performance of the unrestricted mean-variance portfolios using the shrinkage estimators. Afterwards, the shrinkage estimators are implemented in the subspace method to investigate whether further improvement is possible.

The simulation setup is the same as discussed before: the excess returns are generated according to the three-factor Fama-French model. In figure 6, the Sharpe ratios are portrayed for the subspace mean-variance portfolio and the unrestricted estimated mean-variance portfolios using different shrinkage estimators for the mean and covariance matrix. The figure on the left hand side displays Sharpe ratios for a fixed number of assets $N = 25$ and increasing observation window T. The figure on the right hand side keeps the number of observations T fixed at 360 months, while the number of assets increases.

From the figure on the left hand side, several observations can be made. Clearly, all portfolio rules perform better when T increases, which can be logically explained by the convergence of the sample moments to the true moments of the returns, which reduces estimation error. Furthermore, it can be easily concluded that the subspace mean-variance portfolio rule outperforms all other unrestricted mean-variance portfolios. Thirdly, the shrinkage estimators of Ledoit and Wolf perform substantially better than the estimated mean-variance portfolios using the sample moments, though the difference vanishes for increasing T. This can be explained by the shrinkage intensity converging to zero, because the sample moments approach the true moments when T grows large. Finally, it can be said that the Bayes-Stein estimators perform severely bad, which is remarkable as the Bayes-Stein estimates of the population mean and variance apparently do not shrink to the sample mean and variance.

From the figure on the right hand side, it can generally be concluded that the merits of diversification are way too low compared to the disadvantages due to estimation error, in case the investment universe is not restricted. However, the subspace method shows that increasing N for fixed T does not suffer from large estimation errors and outperforms the other unrestricted mean-variance portfolios. For increasing N, the unrestricted mean-variance portfolios display decreasing Sharpe ratios. The shrinkage estimator of Ledoit and Wolf performs far better than the Bayes-Stein estimators and the sample mean and variance estimates. Finally, it can be seen that when N grows large, the Bayes-Stein estimates and the sample mean and variance estimates perform equally bad.

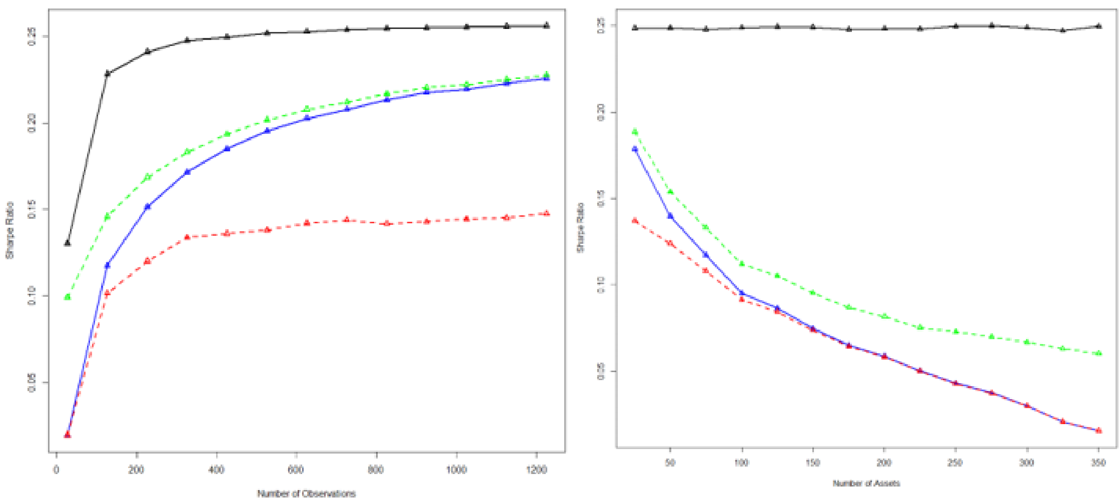

Figure 6: The Sharpe ratios for increasing T (left) and N (right). The black, red, green and blue lines correspond with the subspace mean-variance portfolio, the Bayes-Stein estimators, the linear shrinkage estimator and the sample mean-variance portfolio.

One particular case is examined, namely when N and T converge at the same speed. The results can be seen in figure 7. T is set equal to $N + 1$ to make sure that the Bayes-Stein estimators are defined. Roughly, the same can be said about the performance of the different portfolio rules as earlier done. A few things can be noticed: the unrestricted mean-variance portfolios using the sample moments and the Bayes-Stein estimates perform equally worse, while the shrinkage estimator of Ledoit and Wolf performs much better than those two. When N and T grow at the same speed, the benefits of the subspace method become clear as the Sharpe ratios are increasing, contrary to the unrestricted mean-variance portfolios.

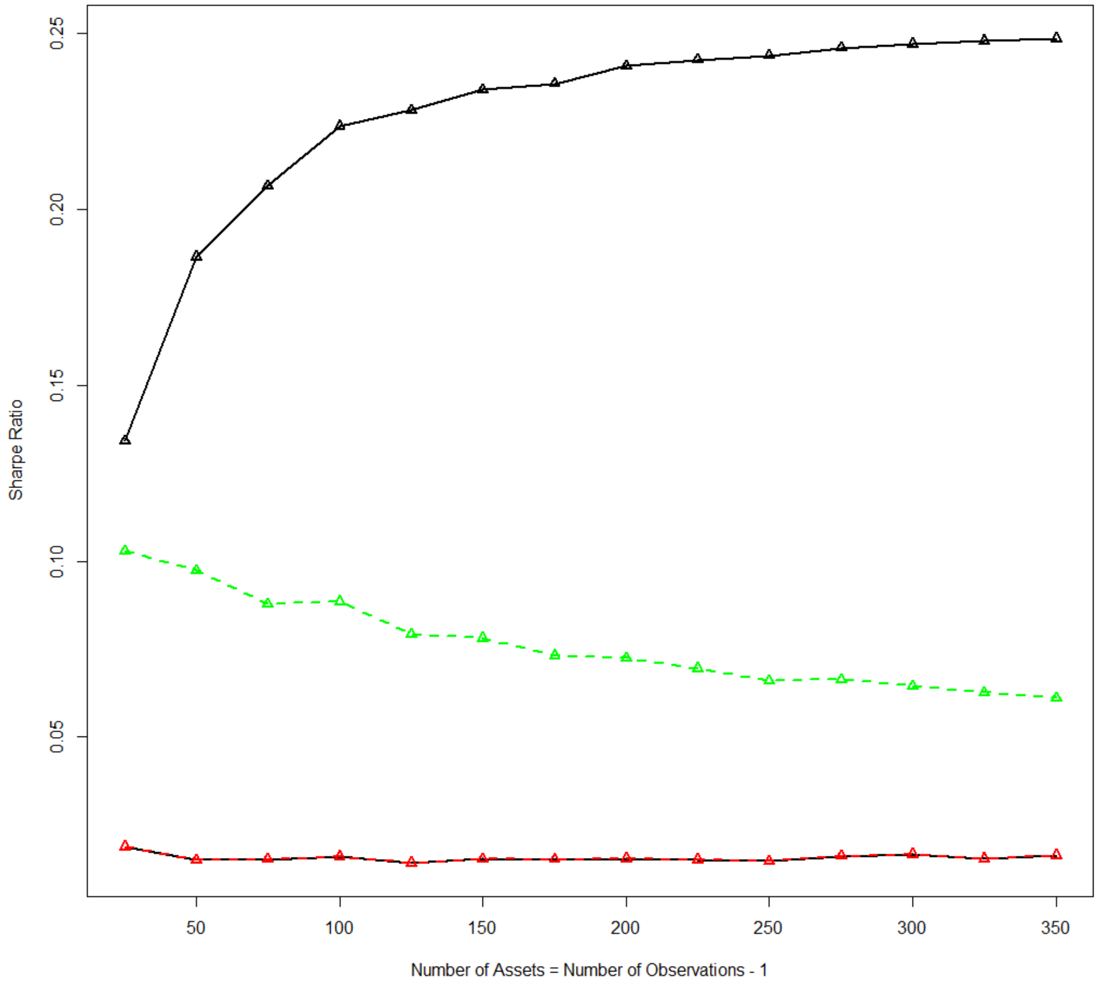

Figure 7: The Sharpe ratios for increasing N and T. The black, red, green and blue lines correspond with the subspace mean-variance portfolio, the Bayes-Stein estimators, the linear shrinkage estimator and the sample mean-variance portfolio.

The different expected return and covariance estimators have been evaluated with the same empirical data procedure as before. Differently than before, the estimates of the mean and covariance matrix have been implemented in the subspace method as well. So, for all three estimators of the mean and covariance matrix, the unrestricted and subspace restricted portfolio Sharpe ratios have been calculated. The results are given in figure 8.

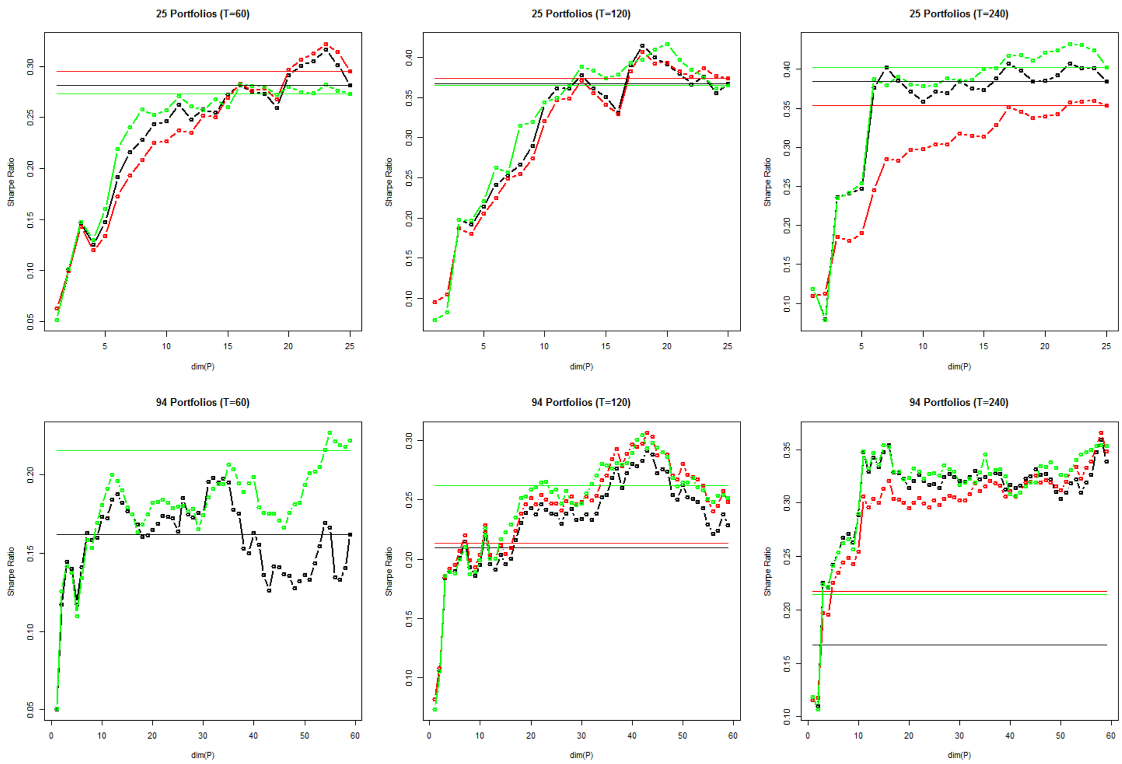

Figure 8: Fama–French portfolio examples — historical performance of $\hat{\mathbf{w}}_d$ for different choices of d over a 50 year period. The dashed black (1), red (2) and green (3) lines correspond with the subspace mean-variance portfolios using the sample moments (1), the Bayes-Stein estimates (2) and the linear shrinkage estimator (3) respectively. The continuous black (1), red (2) and green (3) lines correspond with the unrestricted estimated mean-variance portfolios using the sample moments (1), Bayes-Stein estimates (2) and the linear shrinkage estimator (3) respectively. The left figure on the second row does not include the Bayes-Stein estimates, because $T < N$ in that case.

A few observations can be made, regarding the subspace method. First of all, it can be roughly concluded that the shrinkage estimator outperforms the usual sample covariance matrix, though the difference declines for increasing T. This is in accordance with the definition of the shrinkage estimator which shrinks the covariance estimator to the sample covariance matrix when T grows large. Secondly, it can be seen that the Bayes-Stein estimates perform worse when T grows. This is quite unexpected, as the above mentioned convergence should also hold for the Bayes-Stein estimators. Apparently, this convergence is slower, compared to the shrinkage estimator.

Regarding the unrestricted estimated mean-variance portfolios, the following can be noted. First of all, when the market is small, the differences between the different estimators of the expected return and covariance matrix are minor. There is no clear estimator which provides the best results. For the larger market, it can generally be said that the shrinkage estimator performs much better than the others.

Comparing the subspace mean-variance portfolios with the estimated mean-variance portfolios, the following can be observed. The subspace mean-variance portfolios do not necessarily outperform the estimated mean-variance portfolios. Especially, when T is small, compared to N, the estimated mean-variance portfolios compare favorably well with the subspace mean-variance portfolios, in case the shrinkage estimator of Ledoit and Wolf is applied. For the other estimators, it holds that there is no clear pattern on whether the subspace mean-variance portfolios provide worse or better Sharpe ratios than the estimated mean-variance portfolios. To make a soft remark, it can be said that given a mean and covariance estimator, it seems that the unrestricted estimated mean-variance portfolio performs better than the its restricted subspace counterpart until $\dim(\mathcal{P})$ chosen large enough, although this is not generally true for all estimators.

One final comment to make is that the merits of the subspace-mean variance portfolios, as demonstrated using simulation analysis, do not hold by definition in practice. A reasonable explanation is the vast amount of noise and external effects in real data. This leads to substantial variation in the performance of the different methods with different mean and covariance estimators. Furthermore, the normality of the returns, as assumed in the simulation analysis, is likely to be incorrect. The practical application of the subspace method therefore remains challenged.

5 Conclusion

By restricting the investment universe to the leading eigenvectors of the covariance matrix, the optimal Sharpe ratio can be achieved, when a three-factor model for the excess returns is assumed. This paper has attempted to implement different estimators of the expected returns and the covariance matrix in the restricted and unrestricted investment universe. Simulation analysis has made clear that restricting the investment universe by the subspace, spanned by the leading eigenvectors of the covariance matrix, offers substantial improvement over the estimated mean-variance portfolio with unrestricted investment universe, hereby making use of different mean and covariance estimators. Apparently, restricting the investment universe dramatically captures a lot of the estimation error in the sample mean and covariance matrix. Though the benefits of the subspace method are clear in simulation, empirical analysis does not support this conclusion. In fact, the subspace mean-variance portfolios do not necessarily outperform the estimated mean-variance portfolios. As stated, external effects have major influence on the performance of the different portfolio construction methods with different mean and covariance estimators, which makes it hard to state to determine which moment estimators and investment universe should be implemented in the real stock market.

Further research can be done on examining models that fit particular data well. In this way, the practical merits of the subspace mean-variance portfolios can be made clear. As shown, theoretically the subspace mean-variance portfolio rule outperforms the portfolio rule of naive diversification and the estimated mean-variance portfolio rule, which is why it can be very well implemented, though its practical benefits still need to demonstrated.

References

[1] Connor, G. and Korajczyk, R. A. 1995. *The arbitrage pricing theory and multifactor models of asset returns.* Handbooks in Operations Research and Management Science 9: 87-144.

[2] Chen, J., and Yuan, M. 2016. *Portfolio Selection in a Large Market.* Journal of Financial Econometrics 14: 496-524.

[3] DeMiguel, V., L. Garlappi, and R. Uppal. 2009. *Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?.* Review of Financial Studies 22: 1915–1953.

[4] Jorion, P. 1986. *Bayes-Stein estimation for portfolio analysis.* Journal of Financial and Quantitative Analysis 21: 279–292.

[5] 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: 603–621.

[6] Ledoit, O., and Wolf, M. 2004. *A well-conditioned estimator for large-dimensional covariance matrices.* Journal of Multivariate Analysis 88: 365-411.

[7] Markowitz, H. M. 1952. *Portfolio Selection.* Journal of Finance 7: 77–91.

[8] Markowitz, H. M. 1959. *Portfolio Selection: Efficient Diversification of Investments.* New York: Wiley.

[9] Sharpe, W. F., 1963. *A simplified model for portfolio analysis.* Management Science: 9: 277-293.

[10] Stein, C., 1956. *Inadmissibility of the usual estimator for the mean of a multivariate normal distribution.* Proceedings of the Third Berkeley symposium on mathematical statistics and probability: Vol. 1: No. 399: 197-206.