Static and dynamic portfolio construction with economic and statistical shrinkage

Erasmus University Rotterdam, Erasmus School of Economics

Bachelor's thesis Econometrics and Operations Research (Finance)

Peter Bruins

(387806)

Supervisor: S.C. Barendse

Second Assessor: K. Scholtus

Date final version: 02-08-2018

Abstract

In this paper I construct 6 portfolios based on static and dynamic policies. The dynamic portfolios are based on theorem provided by DeMiguel et al. (2015). The paper studies the impact of shrinkage, economic as well as statistic, on multi-period portfolio performance and compares those with a static portfolio which trades towards a robust target portfolio Wang & Taylor (2018). Performance of the portfolios is compared with Sharpe ratios. I find that shrinkage can improve portfolio performance, but it needs to be done carefully. Using the wrong shrinkage target can worsen performance. The portfolio with the highest Sharp ratio is a four fund shrinkage portfolio with a estimated covariance matrix shrunk to a homoscedastic shrinkage target.

1 Introduction

1.1 Backgound

Investors have a large variety of possible investment options, they can choose to invest in countless assets. Since the ground braking work done by Markowitz, the search to maximize profit based on mathematic theorems took a lift. Newly developed econometric techniques have shown that smart asset allocation aiming at improving utility for investors has the potential to outperform classic investment strategies Zhu & Zhou (2009). Zhu and Zhou have shown that there are technical trading rules which are robust and able to outperform model-based optimal trading strategies substantially. Knowledge concerning potential capital gains of new strategies is of great importance to investors around the world, as financial markets are known to be competitive. Over the past few years, many researchers have tried to improve the earlier mentioned fundamental investment strategy; the Markowitz mean variance framework. Markowitz (1952) developed a theorem on how risk-averse investors could maximize their expected return. The theorem only takes into account the mean and variance of a portfolio and using these constructs an "efficient frontier" of portfolios offering a maximum expected return given a certain level of risk.

The problem with the Markowitz (1952) strategy is that theoretical and real-world performance are know to be non corresponding. This is caused by the theory being fundamentally based upon three restrictive assumptions that may not hold in practice. The first assumption is that investors are myopic and maximize single period utility. The second says that financial markets are frictionless (i.e. no transaction costs) and the third says there is no parameter uncertainty (DeMiguel et al. 2015).

Problems like these caused researchers to challenge the static mean-variance portfolio policy in the years after its publication. Examples of papers inspired by the mean-variance framework are Garleanu & Pedersen (2013), Moallemi & Saglam (2017) and Michaud (1989). The papers mentioned above all adopt a mean-variance portfolio policy of some sort to solve issues they believe are most fundamental to the problems within the classical policy framework of Markowitz (1952). Garleanu & Pedersen (2013) try to minimize the effect of errors caused by the mean-variance framework with a portfolio based on shrinkage (explanation in the next paragraph). Moallemi & Saglam (2017) show that an approach to portfolio selection which involves dynamic trading with mean-variance preferences is able to produce beneficial results. Michaud (1989) tries to improve the errors caused by faulty assumptions by constraining the portfolio objective based on fundamental investment considerations. In other words; they try to select reliable data.

Garleanu & Pedersen (2013) proposed a strategy that assumes quadratic transaction costs for trading. They found that a strategy optimizing a linear combination of the existing portfolio and an aim portfolio performs best under these assumptions. The optimal linear combination is based on a shrinkage approach that considers the current Markowitz portfolio and all expected future Markowitz portfolios. Using this dynamic strategy to undo the Markowitz (1952) portfolio of assumption one and two, they found a higher net return than the more naive benchmarks. Kan & Zhou (2007) focused on improving portfolios by relaxing the assumption that there is no estimation error. They developed a strategy where two single-period shrinkage portfolios shrink the sample mean-variance portfolio towards a target portfolio to minimize utility loss from parameter uncertainty. DeMiguel et al. (2015) used the strategies proposed above to conclude that a portfolio strategy based on multi-period shrinkage and quadratic transaction costs substantially outperforms portfolios that do not relax the assumption underlying Markowitz (1952).

1.2 Issue to address

In this paper I will review the portfolio strategy developed by DeMiguel et al. (2015) using two methods. The first will be to look at possible improvement by taking into account the covariance estimator proposed by Ledoit & Wolf (2004). I will enrich the shrinkage of the covariance matrix proposed by them using two different shrinkage targets that always provide a positive definite shrinkage estimator of the covariance matrix as introduced by Schäfer & Strimmer (2005). Although the paper uses the two shrinkage targets for biomedical studies it is shown that they might be beneficial to economic studies as well (Martellini & Ziemann 2010). The Second method will be to assess the critical review on the difference between static and dynamic portfolios performed by Wang & Taylor (2018). Wang and Taylor show that a static portfolio policy which trades towards a robust target can perform on par with the policy developed by DeMiguel et al. (2015), this is in contrast to the findings of Demiguel. Based on the discrepancy between DeMiguel et al. (2015) and Wang & Taylor (2018) as well as possible estimation improvements, the central research question of this paper will be as follows:

"Does the DeMiguel et al. (2015) multi period four-fund portfolio outperform portfolios based on different shrinkage methods and a robust static portfolio? "

1.3 Approach

To answer the question, I will first replicate parts of the research done on dynamic portfolio selection by DeMiguel et al. (2015)(DMN), considering quadratic transaction costs and parameter uncertainty. Specifically I will focus the multi period four-fund portfolio with shrinkage in trading rate. Secondly I will construct the following portfolios; a robust static portfolio as done by Wang & Taylor (2018), several portfolios based on DMN but with covariance shrinkage introduced by Ledoit & Wolf (2004), taking in to account different shrinkage targets (Schäfer & Strimmer 2005). Finally, I will compare the results using three empirical data sets. The comparison will be made using Sharpe ratios net of transaction costs as is common practice in papers of this kind. The goal of my research is to explore the weakness of the normal approach for covariance estimation with different shrinkage-based covariance estimates. To thoroughly address the differences between the estimators I will use a dataset known to be subject to large estimation errors.

One of the empirical data sets will consist of real estate data. The real estate data will consist of U.S. Real Estate Investment Trust (REITs), these are publicly traded assets of companies which buy real estate with the purpose of value accumulation for the owners. Because REITs use real estate as their collateral they tend to have much more stable returns with short periods of strong deviation. REITs are known to give estimation problems, which are caused by more extreme results in comparison to normal stock returns Zhou & Anderson (2010). Zhou and Anderson find that REITs generally have higher extreme risk than those stock markets. Furthermore, they find that the financial crisis (2007) has increased the extreme risk exposure for REITs. In theory this might cause traditional volatility estimators to wrongly estimate risk within the data (Vu & Tuan 2013).

Vu & Tuan (2013) summarize the problem regarding real estate portfolios as follows: The application of mean-variance based portfolio construction to real estate is questionable (Cheng & Liang 2000). Since mean-variance analysis is quite sensitive to estimation errors real estate data can be a huge pitfall to portfolio construction. Using the shrinkage estimator proposed by Ledoit & Wolf (2004) to estimate the covariance matrix, they limit the impact of estimation errors on the optimization process. The result of this policy is a distinctive investment strategy.

The rest of the paper will be organized as follows. Section 2 illustrates the theoretical frameworks of the chosen portfolio strategies. Section 3 provides the data used in the analysis. Section 4 shows the results of the different portfolios. Section 5 concludes.

2 Portfolio strategies

2.1 The classic Markowitz portfolio

To be able to fully compare all the portfolios I will use the Markowitz (1952) portfolio as base comparison. The Markowitz portfolios is constructed as follows:

$$\min_{x} \sigma^2_{p,x} = x'\Sigma x$$
$$s.t \quad \mu_p = x'\mu = \mu_{p,0}, \quad and \quad x'\mathbf{1} = 1$$

where $\mathbf{1}$ is a vector of ones, $\mu_{p,0}$ denotes the target level of return, Σ denotes the covariance matrix and x the portfolio weight vector. To compute the Markowitz porfolio the R package "fportfolio" [1] is used. For extensive proofs and derivations of the Markowitz portfolio consult Markowitz (1952).

2.2 The multi period four-fund portfolio

Before the portfolio selection can be performed, we first need to formally define the investor's objective:

$$\max_{x_i} \quad U\big(\{x_i\}\big) = \sum_{i=0}^{\infty} (1-\rho)^{i+1} \left(x_i'\mu - \frac{\gamma}{2} x_i'\Sigma x_i \right) - (1-\rho)^{i} \left(\frac{\lambda}{2} \Delta x_i' \Sigma \Delta x_i \right) \tag{1}$$

Where $x_i \in \mathbb{R}^N$ for $i \geq 0$ contains the number of share held of each of the N risky assets at time i, ρ is the investor's impatience factor and γ is the absolute risk-aversion parameter. The investors objective can be described as the minimization of the multi-period utility loss. Utility loss is largely caused by two flaws; The parameter uncertainty in the mean-variance framework and transaction costs. This is why the objective tries to account for both with a quadratic penalty on the one hand and an adaptation of portfolio shrinkage on the other hand.

The multi period four-fund portfolio will be constructed the as follows:

$$\hat{x}^{4F} = \big(1-\beta\big)\hat{x}_{i-1}^{4F} + \beta\big(\zeta_1 \hat{x}^M + \zeta_2 \hat{x}^{Min}\big) \tag{2}$$

Here $\hat{x}^M$ denotes the Markowitz portfolio and $\hat{x}^{Min}$ denotes the minimum variance portfolio. The main target of the shrinkage is to reduce the effect of parameter uncertainty within the optimization process. The shrinkage taking place is an economic shrinkage, meaning the shrinkage targets are based upon economic theory. This is different to the statistic based shrinkage later on.

The optimal shrinkage rates (ζ_1 and ζ_1) are given by:

$$\begin{aligned} \zeta_1 &= c^{-1} \frac{\Psi^2}{\Psi^2 + \frac{N}{T}} \\ \zeta_2 &= c^{-1} \frac{\Psi^2}{\Psi^2 + \frac{N}{T}} \frac{\mu'\Sigma\iota}{\iota'\Sigma\iota} \\ where \quad \Psi^2 &= \mu'\Sigma\mu - \frac{(\mu'\Sigma\iota)^2}{\iota'\Sigma\iota} \end{aligned}$$

To ensure an optimal trading rate (β) this will also be shrunk. The optimal trading rate is given by the following:

$$\max_{\beta} \quad V_1(x_{-1} - X^c)'\mu - \frac{1}{2}\big(E[(\widehat{x}^c)'\Sigma\widehat{x}^c]V_2 + x_{-1}'\Sigma x_{-1}V_3 + 2x_{-1}'\Sigma\widehat{x}^c V_4\big) \tag{3}$$

Where x_{-1} is the investors initial position, $\widehat{x}^c = \zeta_1\hat{x}^M + \zeta_2\hat{x}^{Min}$ is the optimal combination between the static mean-variance portfolio and the minimum variance portfolio. The expectation needed to optimize the trading rate β is given by:

$$E[(\widehat{x}^c)'\Sigma\widehat{x}^c] = \frac{c}{\gamma^2} * \big(\zeta_1^2\big(\mu'\Sigma^{-1}\mu + \frac{N}{T}\big) + \zeta_2^2\iota'\Sigma^{-1}\big) + \frac{c}{\gamma^2} * 2\zeta_1\zeta_2\mu'\Sigma^{-1}\iota\big)$$

[1] https://cran.r-project.org/web/packages/fPortfolio/fPortfolio.pdf

Furthermore $V_1, ..., V_4$ are given by:

$$
\begin{aligned}
V_1 &= \frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} \\
V_2 &= \gamma\frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} - 2\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} + \tilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2} \\
V_3 &= \gamma\frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2} + \tilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2} \\
V_4 &= \gamma\left(\frac{(1-\rho)(1-\beta)}{1-(1-\rho)(1-\beta)} - \frac{(1-\rho)(1-\beta)^2}{1-(1-\rho)(1-\beta)^2}\right) - \tilde{\lambda}\frac{(1-\rho)\beta^2}{1-(1-\rho)(1-\beta)^2}
\end{aligned}
$$

.

The optimal trading rate β tries to optimize the trade-off between the excess return of the investor's initial portfolio with the optimal portfolios combination and the variability combined with the trading cost of the four-fund portfolio. In other words, it tries to minimize the expected utility loss.

The optimizations are based on the following estimates for μ and Σ:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t, \quad \text{and} \quad \hat{\Sigma} = \frac{1}{T-N-2}\sum_{t=1}^{T}\left(r_t - \hat{\mu}\right)^2 \tag{4}$$

Where T is the estimation window and N is the total amount of observations.

2.3 The multi period four-fund portfolio with covariance shrinkage

The multi period four-fund portfolio with shrinkage will be calculated with the same investor's objective and the same optimal portfolio goal as before (see equation 1 and equation 2). The differences between the portfolios will arise from the covariance matrix used for the optimization process. Instead of the traditional covariance estimator (equation 3) used by DMN, estimators based on different shrinkage targets will be used. Note that this will not only result in a change in the utility function but also the optimal shrinkage rate (ζ_1 and ζ_2) which is based on the covariance matrix will change. Covariance shrinkage might improve the portfolios because in many practical cases, the number of observations in comparison to the number of assets is to small. Likely this will cause the traditional sample covariance estimator to suffer from large estimation errors. To deal with these problems, the shrinkage targets of the covariance estimators try to find an estimator which is robust in the case of a high dimensional covariance matrix. Shrinkage of the covariance estimator is based on finding an optimum between bias added in the shrinkage process and variance. To understand the trade off between bias and variance one could think of the identity matrix as the shrinkage target. This will result in little variance but a huge bias.

2.3.1 The Ledoit and Wolf shrinkage target

Ledoit & Wolf (2004) propose to shrink the covariance matrix $\Sigma^{L\&W}$ by finding an optimal weight minimizing the quadratic loss of the shrinkage covariance estimator in comparison to the (unknown) true covariance estimator. Hereby the shrinkage covariance estimator is based on a linear combination of the identity matrix and the sample covariance matrix. Mathematically this is shown as follows:

$$\min_{\beta_1,\beta_2} \quad E[\parallel \Sigma^{L\&W} - \Sigma^{true} \parallel_{\mathbf{F}}^2] \tag{5}$$

$$s.t. \quad \Sigma^{L\&W} = \beta_1 I + \beta_2 S \tag{6}$$

Here $\parallel ... \parallel_{\mathbf{F}}$ denotes the Frobinius norm, I is the identity matrix, S is the sample covariance matrix (here we use the unbiased covariance estimator) and Σ^{true} is the true covariance matrix. The solution to the minimization problem is as follows:

$$\Sigma^{L\&W} = \frac{\beta^2}{\delta^2}\mu I + \frac{\alpha^2}{\delta^2} S \tag{7}$$

and

$$E[\parallel \Sigma^{L\&W} - \Sigma^{true} \parallel_{\mathbf{F}}^2] = \frac{\alpha^2\beta^2}{\delta^2} \tag{8}$$

where the following holds for α, β and δ:

$$\mu =< \Sigma^{true}, I >, \quad \alpha^2 = \parallel \Sigma^{L\&W} - \mu I \parallel, \quad \beta^2 = E[\parallel S - \Sigma^{L\&W} \parallel^2]$$
$$and \quad \delta^2 = E[\parallel S - \mu I \parallel^2]$$

With this knowledge L&W prove that in case optimal shrinkage the following always hold, $\alpha^2 + \beta^2 = \delta^2$. This property is very important as it makes sure that the linear combination between S and I is convex.

A clear way to interpret equation (5) is to decompose it, this gives us the following (equation (10) of Ledoit & Wolf (2004)):

$$E[\parallel \Sigma^{L\&W} - \Sigma^{true} \parallel_{\mathbf{F}}^2] = E[\parallel \Sigma^{L\&W} - E[\Sigma^{L\&W}]^2 \parallel_{\mathbf{F}} + \parallel E[\Sigma^{L\&W}] - \Sigma^{true} \parallel_{\mathbf{F}}^2 \tag{9}$$

We can now see that we can interpret the L&W shrinkage as a trade-off between variance and bias, where we seek to minimize the mean squared error as stated above. The mean squared error of the shrinkage target μI consists of only bias and no variance. However, the mean squared error of the sample covariance matrix S is the exact opposite of that, constituting only out of variance and no bias.

Since the true covariance matrix is unknown, we rely on the asymptotic behavior of the estimators to be well behaved. L&W show that if we use the notation above with a subscript n, we can show that that all results asymptoticly hold, that is:

$$S = X_n X_n^t/n \quad \mu_n =< \Sigma_n^{true}, I_n >_n, \quad \alpha_n^2 =\| \Sigma_n^{L\&W} - \mu_n I_n \|_n, \quad \beta_n^2 = E[\| S_n - \Sigma_n^{L\&W} \|_n^2]$$
$$and \quad \delta_n^2 = E[\| S_n - \mu_n I_n \|_n^2]$$

then

$$S^* = \frac{\beta_n^2}{\delta_n^2}\mu_n I_n + \frac{\alpha_n^2}{\delta_n^2} S_n \tag{10}$$

is a consistent estimator for $\Sigma^{L\&W}$, which means it has the same asymptotic properties. Thus, we can use μ, α, β and δ to work around the unknown true covariance matrix. We now only need consistent estimators for μ, α, β and δ (m, a, b and d). These are as follows:

$$m_n =< S_n, I_n >_n$$
$$d_n =\| S_n - m_n I_n \|_n^2$$

then let x_k^n be the k_{nh} column of X_n. Define $\bar{b}_k^n = \frac{1}{t^2}\sum_{k=1}^n \| x_k^n(x_k^n)' - S_n \|_n^2$, now a and b follow:

$$b_n^2 = min(\bar{b}_n^2, d_n^2)$$
$$a_n^2 = d_n^2 - b_n^2$$

For further clarification a different approach on L&W shrinkage is given in the next subsection. For extensive proofs of the results above consult Ledoit & Wolf (2004).

2.3.2 The Schäfer en Strimmer shrinkage targets

Schäfer & Strimmer (2005) provide two more shrinkage targets for covariance estimation which always provide a positive definite shrinkage estimator of the covariance matrix. To understand the shrinkage proposed by Schäfer and Strimmer it is useful to use a different approach to equation (5). They define the shrinkage target as an optimization problem of the covariance matrix, this shows as follows:

$$L(\lambda) =\| \Sigma^{shrinkage} - \Sigma^{true} \|_{\mathbf{F}}^2 \tag{11}$$
$$=\| \lambda T + (1-\lambda)\Sigma^{shrinkage} - \Sigma^{true} \|_{\mathbf{F}}^2 \tag{12}$$
$$= \sum_{i=1}^p \sum_{j=1}^p (\lambda t_{ij} + (1-\lambda)s_{ij} - \sigma ij)^2 \tag{13}$$

Here s_{ii} is the i,i value sample variance matrix S, t_{ij} is the i,j value shrinkage target T and σ_{ij} is the true variance. For exact derivation and explanation see Schäfer & Strimmer (2005) paragraph 2.4. To further clarify how the methodology (equation (11), (12) and (13)) can be used for covariance matrix shrinkage we first show the "simple" L&W shrinkage case. Furthermore defining Ledoit & Wolf (2004) within this methodology, gives us the opportunity to compare the three different shrinkage targets better. If we express the L&W shrinkage in this terminology it will give the following:

$$t_{ij} = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases}$$

$$\widehat{\lambda} = \frac{\sum_{i \neq j} \widehat{var}(s_{ij}) + \sum_i \widehat{var}(s_{ii})}{\sum_{i \neq j} s_{ij}^2 + \sum_i (s_{ii} - 1)^2}$$

Here t_{ij} is in fact the identity matrix and $\widehat{\lambda}$ is the optimal lambda. The two shrinkage targets proposed by Schäfer and Strimmer that I will use, are one homoscedastic shrinkage target and one heteroscedastic target. The two shrinkage targets are as follows

Shrinkage target diagonal with common variance:

$$t_{ij} = \begin{cases} v = \overline{s_{ii}} & if \quad i = j \\ 0 & if \quad i \neq j \end{cases}$$

$$\widehat{\lambda} = \frac{\sum_{i \neq j} \widehat{var}(s_{ij}) + \sum_i \widehat{var}(s_{ii})}{\sum_{i \neq j} s_{ij}^2 + \sum_i (s_{ii} - v)^2}$$

Shrinkage target diagonal with unequal variance:

$$t_{ij} = \begin{cases} v = s_{ij} & if \quad i = j \\ 0 & if \quad i \neq j \end{cases}$$

$$\widehat{\lambda} = \frac{\sum_{i \neq j} \widehat{var}(s_{ij})}{\sum_{i \neq j} s_{ij}^2}$$

Here s_{ii} denotes the i, i value of the sample variance matrix S and $\overline{s_{ii}}$ denotes the average variance of the sample variance matrix S. To estimate the sample covariance matrix as in Schäfer & Strimmer (2005) it is recommended to read appendix A of their paper. Furthermore, note that the difference in the optimal lambda between the first S&S shrinkage target and the L&W shrinkage target is very small. The optimal λ only varies slightly in the denominator. Also we can see that the second S&S shrinkage target is really straightforward. The target given by the first of the two S&S shrinkage targets shrinks the estimated covariance to a diagonal common variance matrix. The second shrinkage target will shrink the estimated covariance matrix to diagonal unequal variance matrix, this implies that it will leave the diagonal entries of the original matrix intact i.e. it does not shrink the variances. S&S specifically note that although it might seem circular to use the heteroscedastic sample variance estimates as shrinkage target because those are the ones we try to improve. In practice this "chicken-egg" issue is completely resolved since $\widehat{\lambda}$ remains the same whether standard or shrinkage estimates are substituted into the equation stated above. To correctly compute the given shrinkage targets I will use the R package "corpcor"[2], This is recommended by Schäfer & Strimmer (2005) in appendix B .

2.4 Static portfolio with a robust trading target

The static portfolios as defined by Wang & Taylor (2018) with robust trading target will be used. This portfolio strategy is developed by Garleanu & Pedersen (2013), but is slightly adjusted with the usage of a time-varying covariance matrix. This causes it to be applicable to single-period investors. The portfolio will maximize the excess price change net of risks and transaction costs in each period, being:

$$\max_{x_t} \quad x_t' E(r_{t+1}) - \frac{\gamma}{2} x_t' \Sigma x_t - \frac{\lambda}{2} \Delta x_t' \Sigma \Delta x_t \tag{14}$$

giving:

$$x_t = \left(1 - \frac{\gamma}{\gamma + \lambda}\right) x_{t-1} + \frac{\gamma}{\gamma + \lambda} \left(\gamma \Sigma\right)^{-1} E_t(r_{t+1}) \tag{15}$$

The covariance matrix will thus vary over time while the expected price change will be fixed. Thus Wang & Taylor (2018) set $E(r_{t+1}) = \mu$ and replace Σ by Σ_t. This will cause the expected price change to limit the excessive turnover. To estimate the time-varying risk (the covariance matrix) the RiskMetrics[3] approach will be used. This risk-forecasting technique will be used because it is convenience, but it also has good results in comparison with the more advanced forecasting techniques. (Wang & Taylor 2018).

3 Data

I will use two equity portfolio data sets. Following DeMiguel et al. (2015) I will use the 48 Industry portfolio (48IndP) and the Fama-French 100 portfolio formed on size and book-to-market (100FF). These are downloaded from Kenneth French his website [4]. I will use daily data from July 2004 until September 2012 and construct price change data by assuming all starting prices are equal to one,and computing price changes from the total return data as in DeMiguel et al. (2015). Because I use the same datasets as DMN I will be able to compare the DMN portfolios with the portfolios reviewed in this paper.

In addition to the DMN datasets I will use a real estate dataset. This will be a dataset consisting out of 5 different U.S. based REITs, of which 3 are equity based and 2 are mortgage based. The REIT data set will be downloaded from the nareit website [5] [6]. The Equity REIT's which

[2]https://cran.r-project.org/web/packages/corpcor/corpcor.pdf

[3]https://cran.r-project.org/web/packages/RiskPortfolios/RiskPortfolios.pdf

[4]http://mba.tuck.dartmouth.edu/pages/fa- culty/ken.french/

[5] https://www.reit.com/data-research/reit-indexes/monthly-property-index-values-returns

[6]This website contains the largest free REIT data set. The organization behind the website considers itself to be the worldwide representative voice for REIT's and publicly traded real estate companies with an interest in U.S. real estate and capital markets.

I will select are the REIT's based on retail, office and industrial real estate and the mortgage REIT's I will select are the REIT's based on home financing and commercial financing. I will use monthly data from December 1999 until June 2018 and use the total return (including dividend and real estate price accumulation). Considering the time frame the data will include the major U.S. housing crisis in 2007. With the usage of monthly data I will diverge from Garleanu & Pedersen (2013) focus on daily data (also followed by DMN), but considering the 3 trillion USD market capitalization[7] I believe that it is an interesting market for investors research. Moreover, DeMiguel et al. (2015) also note that it is preferred to use daily data, but they have evaluated weekly and monthly data and found it to be generally robust to use lower frequency data (bottom of page 1458).

Table 1: Market Capitalization of REIT by sub sector

	Marketcap (USDm)	% of REIT market	Date range
Mortgage Home Financing	46164	4.23	12/1999 – 6/2018
Mortgage Commercial Financing	17381	1.59	12/1999 – 6/2018
Equity Retail	167000	15.30	12/1999 – 6/2018
Equity Industrial	82160	7.53	12/1999 – 6/2018
Equity office	98000	8.99	12/1999 – 6/2018

Note: This table shows market capitalization, weight of the REIT on the total REIT market in % and the date range.

We can see that the market for equity REIT's is far larger then the mortgage market. Furthermore, it is clear that retail REIT's are dominant in the market.

Table 2: Descriptive statistics

	Mean	Std. Dev	Min	Max
Mortgage Home Financing	0.94	5.57	-27.98	18.51
Mortgage Commercial Financing	0.58	8.75	-45.09	37.46
Equity Retail	1.16	7.02	-36.78	43.52
Equity Industrial	1.23	9.47	-56.19	70.48
Equity office	0.94	6.42	-31.80	32.46

Note: This table shows some descriptive statistics for the returns (% per month) of all the REITs.

[7] https://www.reit.com/data-research/data/reits-numbers

Figure 1: Plot of the returns (%) of the equity retail dataset

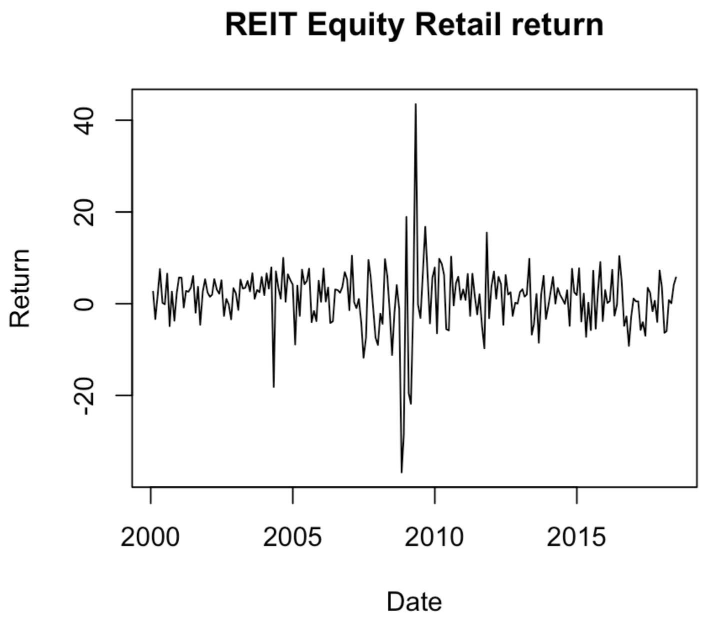
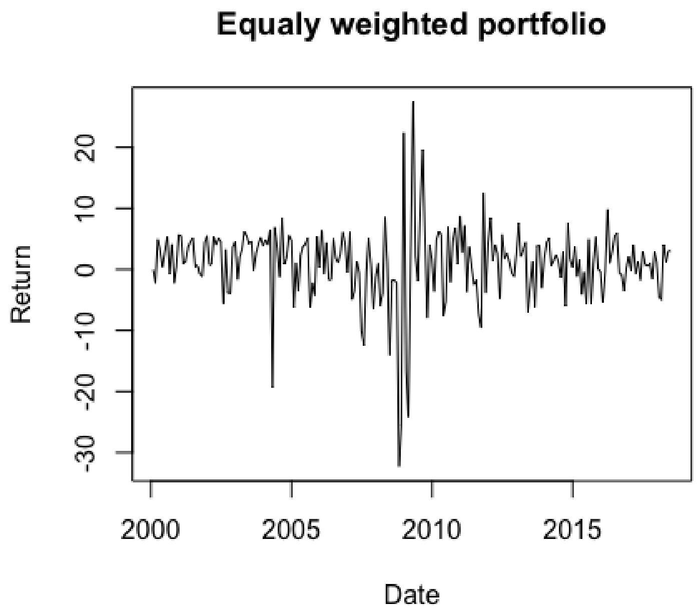


Figure 2: Plot of the returns (%) of the equity retail dataset



The descriptive statics show that although REITs are backed by real estate, they in fact are more volatile. When we take look at the timing of the big negative returns it is eminent that they are caused by the U.S. housing crises; All of the minimum returns take place around 2007 (see figure 1 and 2). Furthermore, when we look at the equally weighted portfolio consisting out of the 5 different REITs we see great resemblance among the figures. This shows that there is strong evidence that the movement in REITs prices is caused by an external force. Strengthening the belief that indeed REITs are largely driven by real estate prices. Concluding, if we look at the bigger picture REITs indeed seems to follow real estate movement and thus can be argued to challenge covariance estimation likewise real estate Vu & Tuan (2013).

4 Results

4.1 Evaluation methodology

To evaluate the performance of the different portfolios, I will use the same methodology as in DeMiguel et al. (2015). This implies that I will do an out-of-sample evaluation of the portfolio gains using the following strategy; First the portfolios will be estimated using a rolling estimation window in which I will use all the information available up to point T. Here I denote the length of the estimation window by T < L, where L is the total number of samples in the data set. For the "rolling-horizon" procedure the length of the estimation window will be constant. To account for the transaction cost, I define the portfolios net gains of trading cost as follows:

$$R_{i+1}^k = (\widehat{x}_i^k)' r_{i+1} - \tilde{\lambda}\Delta(\widehat{x}_i^k)'\Sigma(\widehat{x}_i^k) \tag{16}$$

Where $\widehat{x}_i^k$ denote the estimated portfolio k at time i, r_i is the vector of price changes at the ith out-of-sample period and Σ is the covariance matrix of the assets prices. The Sharp ratios of the portfolios will now be calculated as follows:

$$SR^k = \frac{\overline{R}^k}{\sigma^k} \tag{17}$$

where

$$(\sigma^k)^2 = \frac{1}{L-1}\sum_{i=1}^{L-1}\left(R_{i+1}^k - \overline{R}^k\right)^2$$

$$\overline{R}^k = \frac{1}{L}\sum_{i=1}^{L-1} R_{i+1}^k$$

Where L is the number of out-of-sample periods.

4.2 Comparing the covariance matrices

To clarify the differences between the covariance estimators I will evaluate the differences between the sample covariance estimator as used by DMN and the covariance matrix estimated with the homoscedastic shrinkage target developed by Schäfer & Strimmer (2005). I have chosen these two because they have the biggest difference in Sharp ratios as stated in the next paragraph. To further clarify the differences between the two estimation method, also the corresponding correlation matrices are given.

Table 3: DMN sample covariance

	Home	Commercial	Retail	Industrial	Office
Home	33.85	0.483	0.316	0.270	0.356
Commercial	24.97	78.75	0.653	0.566	0.677
Retail	13.11	41.28	50.70	0.834	0.902
Industrial	15.14	48.29	57.07	92.30	0.827
Office	13.48	39.17	41.83	51.78	42.38

Note: This table shows the sample covariance of the REITs on th left bottom triangle (including the diagonal). The corresponding correlation matrix on the right top triangle (for further clarity the correlations matrix is underlined). Home denotes the mortgage home financing REIT, commercial denotes the mortgage commercial financing REIT, retail denotes the equity retail REIT, industrial denotes the equity industrial REIT and office denotes the equity office REIT

Table 4: Homoscedastic shrinkage sample covariance and correlation

	Home	Commercial	Retail	Industrial	Office
Home	40.62	0.452	0.295	0.253	0.332
Commercial	23.02	63.82	0.610	0.529	0.633
Retail	13.24	53.23	49.33	0.779	0.843
Industrial	13.58	35.60	46.10	70.83	0.774
Office	14.23	33.97	39.76	43.71	45.03

Note: This table shows the covariance matrix based on homoscedastic shrinkage of the REITs on th left bottom triangle (including the diagonal). The corresponding correlation matrix on the right top triangle (for further clarity the correlations matrix is underlined). The labels are the same as in table 3

Comparing the two matrices I find that the difference might be smaller then one expects, especially concerning the correlations. The diagonal of the homoscedastic shrinkage covariance matrix has shrunk to the shrinkage target for about 50%. This is shown by the optimal shrinkage intensity for the variance vector $\lambda_{var} = 0.4688$. The correlations between the REITs in the two covariance estimation techniques are largely intact. $\lambda_{cor} = 0.0651$ shows that only a small part of the correlation matrix has shrunk to the target of 0. We can thus conclude that the largest effect shrinkage seems to have is a reduction between the differences of the individual REIT variances. Thus I conclude that the largest part of possible differences in portfolio estimation is due to differences in the variance. The L&W shrinkage target roughly gives the same result. The heteroscedastic target keeps variances equal by defenition. In the case of L&W there is more of an overall reduction in the variance, whereas the heteroscedastic shrinkage target keeps the variances equal but has stronger shrinkage on the correlation matrix. The optimal shrinkage target for heteroscedastic shrinkage $\lambda_{cor} = 0.113$ is quite a bit higher than the optimum in the homoscedastic case.

4.3 Performance of the portfolios

The tables on the following two pages (table 5 and 6) report the out-of-sample Sharpe ratios of the 6 portfolios considering the 3 different data sets. The two tables consider portfolios based on two kinds of investor's: one investor whose initial portfolio is fully invested in the risk-free asset, and another investor whose initial portfolio is the true Markowitz portfolio x^M. The 'true' Markowitz portfolio x^M is constructed with the use of the entire sample of the three data sets. The return of the risk-free asset is set to be 0.

The following base-case investor is considered in table 3 and 4. The investor has an absolute risk-aversion parameter of $\gamma = 10^{-8}$ this is equal to a relative risk aversion of 1 for an investor who has \$100 million to trade. The impatience factor is set to be $\rho = 1 - \exp(-0.1/260)$, this corresponds to an annual discount of 10%. We estimate the FF100 and the 48IndP portfolio strategy with $T = 500$ observations. The REIT data set is estimated with $T = 111$ (50% of total sample). The portfolio "Four fund shrinkage $S\&S_1$" is the portfolio based on the homoscedastic shrinkage covariance estimation as proposed by Schäfer and Strimmer. The "Four fund shrinkage $S\&S_2$" portfolio corresponds to the portfolio that uses heteroscedastic shrinkage based covariance estimation. Sharpe ratios are discounted by quadratic transaction costs with $\lambda = 3\times10^{-7}$.

Considering the different portfolios and data samples a few things stand out. I will first evaluate the result of the Wang & Taylor (2018) portfolio. The results I find with the portfolio are in contrast with the findings of Wang and Taylor. Where they consider the performance of their portfolio to be on par with the four fund DMN portfolio, I find it to be performing worse. Particularly, using the REIT data set the portfolio struggles to keep up. It even underperforms in comparison with the base portfolio by Markowitz (1952). Since Wang and Taylor used simulated data as well as empirical data to strengthen their results, I find it hard to conclude what caused their portfolio to underperform. Especially the REIT data is difficult to explain. The best possible explanation that I can give, is that the largely stable nature of the REIT data causes the portfolio to misjudge the effects of short periods with extremes. This effect is then stronger because the portfolio policy is static. Another possibility is that transaction costs have a more than average negative effect on the static policy. Because single-period investment strategies are known to trade more frequent, they are penalized more by transaction costs. So it is possible that the more inconsistent REIT dataset caused the portfolio to trade with higher frequency resulting in a high transaction penalty.

However, the optimization process used by Wang & Taylor (2018) does account for quadratic transaction costs.

Table 5: Sharpe ratios discounted with transaction costs

Target portfolios starting with a risk free investment	Data sets		
	100FF	48IndP	REIT
Static policies			
Sample Markowitz	-1.435	-0.627	-0.231
Wang and Taylor	0.203	0.314	-0.434
Multi-period policies			
Four fund shrinkage DMN	0.330	0.515	0.421
Four fund shrinkage $L\&W$	0.329	0.515	0.437
Four fund shrinkage $S\&S_1$	0.347	0.533	0.487
Four fund shrinkage $S\&S_2$	0.322	0.489	0.412

Note:Table 5 reports the annualized out-of-sample Sharpe ratio for the different portfolio strategies that I consider. The settings for λ, ρ and γ are as given above.

The second result I find interesting, is the good performance of the homoscedastic shrinkage portfolio. In both the risk free as well as the true Markowitz case the homoscedastic shrinkage portfolio delivers the best performance. Although with the 100FF and the 48IndP datasets the portfolio still performs worse than the expanding window four fund portfolio as given by DeMiguel et al. (2015) (table 2). The performance in comparison to the normal four fund portfolio is promising. Most notably the homoscedastic shrinkage portfolio outperforms the other portfolio with the REIT dataset. This can be explained by the more equal variances in proportion to the correlation matrix. This can decrease the effect of large outliers causing the portfolio to estimate more robust and thus be better at forecasting risk.

Moreover results show a downside of covariance matrix shrinkage. The performance of the heteroscedastic shrinkage portfolio is worse than the DMN four fund portfolio. The heteroscedastic shrinkage leaves the variances intact and only shrinks the correlations. It seems that this causes the true relation between the 5 REITs to be wrongly presented. The fact that this result is stronger with the REIT data set follows as a result of the asymptotic properties of the shrinkage estimator.

Since the shrinkage estimator will asymptoticly move towards the sample covariance matrix, using larger datasets will decrease the impact shrinkage has. Hence, the REIT data covariance matrix, being only estimated with 111 data points, is undergoing more severe shrinkage than those of 100FF and 48IndP. When the shrinkage target is faulty at interpreting the data, this will than result in a large proportion of the shrunken covariance matrix to exist of the shrinkage target (the optimal λ is larger). This results in poor performance with the REIT data.

Table 6: Sharpe ratios discounted with transaction costs

Target portfolios starting with the Markowitz portfolio	Data sets		
	100FF	48IndP	REIT
Static policies			
Sample Markowitz	-1.465	-0.660	-0.236
Wang and Taylor	0.202	0.315	-0.434
Multi-period policies			
Four fund shrinkage DMN	0.362	0.712	0.526
Four fund shrinkage $L\&W$	0.362	0.710	0.540
Four fund shrinkage $S\&S_1$	0.379	0.750	0.599
Four fund shrinkage $S\&S_2$	0.361	0.691	0.517

Note: Table 6 reports the annualized out-of-sample Sharpe ratio for the different portfolio strategies that I consider. The settings for λ, ρ and γ are as given above.

Furthermore, the L&W shrinkage gives very equal results to the normal DMN portfolio. This is due to the fact that the L&W covariance estimator is nearly equal to the normal sample covariance estimator. With the 100FF and the 48IndP datasets the optimal shrinkage intensity $\lambda = 0.0042$ is very close to 0. In the case of the REIT dataset the optimal shrinkage intensity is a bit higher but still low compared to the S&S shrinkage estimators. The low value for the optimal λ implies that in the 100FF and the 48IndP datasets the L&W estimator finds the trade-off between variance and bias to be close to the optimum.

Having a look at both the results of the portfolios starting with the risk free rate and the portfolios starting with the true Markowitz portfolio. I find that the latter has slightly better results for the multi period portfolios. In the case of the static portfolios performance remains near constant. This falls within our expectations, because starting with the true Markowitz it is optimal to remain within the current portfolio. However, the investor ignores the starting portfolio. Thus the investor benefits from the shrinking trading rate as it causes him to evaluate the starting position given the available information. This is true for all of the multi-period portfolios since they all use the shrunken trading rate (equation (3)).

Overall, all of the multi-period portfolios show great resemblance in performance. This is due to the fact that for all of the portfolios the optimal shrinkage rate are rather close to 0, causing the shrunken covariance matrices to be close to the sample covariance matrix. The homoscedastic shrunken covariance matrix just outperform the others and the static portfolios fail to perform on par with the multi-period portfolios. Moreover, I find that taking into account the impact of both parameter uncertainty and transaction costs is hugely beneficial. The difference is shown by the classic Markowitz portfolio which clearly is the worst performer overall by a quite margin. Demonstrating the importance of relaxing the earlier mentioned assumptions it is based on.

5 Conclusion

I analyze the performance of different static and dynamic portfolio policies on three empirical data sets. I use two of the data sets as used by DeMiguel et al. (2015) and a REIT set. Existing literature shows that the latter can give estimation result. I provide the policies for all of the portfolios and show that there is possible improvement to performance using covariance shrinkage. I find, in contrast to Wang & Taylor (2018), that the dynamic portfolio policies outperform the static policy trading towards a robust target portfolio. My results show, that the homoscedastic shrinkage target for the covariance matrix as presented by Schäfer & Strimmer (2005) results in the best overall performance. Moreover, starting with the true Markowitz portfolio performance is even better. I also conclude that it is important to review different shrinkage approaches to the covariance matrix as results show that shrinkage can also worsen the performance of a portfolio.

All in all the biggest performance gain is caused by taking into account the impact of both parameter uncertainty and transaction costs. All of the multi-period portfolio perform better than the classic Markowitz porfolio. This implies that possible performance gains from economic shrinkage are of larger contribute to overall performance of the portfolios, than those arising from covariance shrinkage. Hence, it remains very important to base portfolio policies on a strong economic foundation as results show performance gains can be quite substantial.

Bibliography

Cheng, P. & Liang, Y. (2000), 'Optimal diversification: Is it really worthwhile?', *Journal of Real Estate Portfolio Management* **6**(1), 7–16.

DeMiguel, V., Martin-Utrera, A. & Nogales, F. J. (2015), 'Parameter uncertainty in multiperiod portfolio optimization with transaction costs', *Journal of Financial and Quantitative Analysis* **50**(6), 1443–1471.

Garleanu, N. & Pedersen, L. H. (2013), 'Dynamic trading with predictable returns and transaction costs', *The Journal of Finance* **68**(6), 2309–2340.

Kan, R. & Zhou, G. (2007), 'Optimal portfolio choice with parameter uncertainty', *The Journal of Financial and Quantitative Analysis* **42**(3), 621–656.

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

Markowitz, H. (1952), 'Portfolio selection', *The Journal of Finance* **7**(1), 77–91.

Martellini, L. & Ziemann, V. (2010), 'Improved estimates of higher-order comoments and implications for portfolio selection', *Review of Financial Studies* **23**, 1467–1502.

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

Moallemi, C. & Saglam, M. (2017), 'Dynamic portfolio choice with linear rebalancing rules', *Journal of Financial and Quantitative Analysis* **52**(3), 1247–1278.

Schäfer, J. & Strimmer, K. (2005), 'A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics', *Statistical applications in genetics and molecular biology* **4**(1).

Vu, E. & Tuan, A. (2013), 'How does the ledoit and wolf shrinkage estimator improve a real estate portfolio?', *Journal of Real Estate Portfolio Management* **19**(1), 89–101.

Wang, J. & Taylor, N. (2018), 'A comparison of static and dynamic portfolio policies', *International Review of Financial Analysis* **55**, 111–127.

Zhou, J. & Anderson, R. (2010), 'Extreme risk measures for international reit markets', *The Journal of Real Estate Finance and Economics* **45**(1), 152–170.

Zhu, Y. & Zhou, G. (2009), 'Technical analysis: An asset allocation perspective on the use of moving averages', *Journal of Financial Economics* **92**(3), 519–544.