UNIVERSITY OF AGDER

How Good is the Out-of-Sample Performance of Optimized Portfolios?

An empirical comparison of optimal versus naive diversification

Anders Bakke

Supervisor
Valeri Zakamouline

University of Agder, 2014
School of Business and Law

Preface

The submission of this master's thesis concludes my master's degree in Economics and Business Administration with specialization in financial economics at the University of Agder. Throughout the master's programme I have established a deeper understanding of financial theory and how to implement theoretical portfolio models empirically. My interest in this field motivated me to further explore these topics in a master's thesis. This has given me the opportunity of learning some features of the software R, which is a free software environment for statistical computing and graphics. The writing process has been both challenging and time consuming but also interesting and enriching, making this experience worthwhile.

First and foremost, I would like to express my deepest gratitude to my supervisor Professor Valeri Zakamouline for always being available, providing excellent guidance, constructive criticism, and providing me with several of the functions implemented for programming purposes. Without his supervision this thesis would not have been possible. I also wish to thank my fellow classmates for creating a pleasant working atmosphere, providing valuable opinions and input during the writing process, and for making my time at the University of Agder so enjoyable. Last of all, I utter my heartfelt thanks to my closest family for their endless encouragement and support, as well as for putting up with my countless working hours and physical absence.

Anders Bakke

Kristiansand, June 2014

Abstract

Preceding research is inconclusive on the empirical performance of optimized portfolios. In this thesis I evaluate the out-of-sample performance of a minimum-variance portfolio, a mean-variance portfolio, an equally-weighted portfolio, and a value-weighted market portfolio across eight U.S. datasets and seven different out-of-sample time periods. This is done in order to determine if any of the asset-allocation strategies deliver statistically significantly, higher Sharpe ratios compared to the other implemented portfolio models. Although the minimum-variance portfolio is persistent in delivering the highest out-of-sample Sharpe ratio, I find that none of the optimized or equally-weighted portfolios consistently deliver statistically distinguishable Sharpe ratios from each other. The value-weighted market portfolio is found to frequently be statistically suboptimal when compared to the other asset-allocation strategies, suggesting that this strategy should generally be avoided in face of the others. By comparing the in-sample and out-of-sample Sharpe ratio of the mean-variance portfolio, I find that there is estimation error affecting the performance.

Contents

List of Figures

List of Tables

1 Introduction

The empirical performance of asset-allocation strategies sparks both curiosity and conflict among investors and academics alike. Indeed, examining the vast academic literature reveals that much research has been devoted to finding which asset-allocation strategy delivers the best performance when implemented empirically. The empirical findings of these different studies are often opposing, suggesting no definitive answer as to which strategy dominates the others and eventually which one to implement. While several studies produce empirical results supporting the supremacy of optimization, there are also numerous research efforts proposing that a naively constructed portfolio, where each of the portfolio assets are weighted equally, performs just as good or better in terms of Sharpe ratios. Effectively, research related to portfolio optimization is a subject of conflict for many theorists. Naturally, one would expect that optimized portfolios, which depend on expectations about the future and is constructed for this purpose, would perform better than allocation strategies that require no input estimation.

To elucidate this debate, I present two studies with conflicting results that lay the foundation for the development of this thesis. Based on data formed on portfolios of U.S. stocks, DeMiguel, Garlappi, and Uppal (2009) found no evidence to support that optimized portfolios deliver statistically significantly higher Sharpe ratios compared to that of an equally-weighted portfolio (EWP), out of sample. In addition to focusing on the Sharpe ratio of each strategy across each dataset, they also focus on the p-value of the test statistic when determining if the difference in Sharpe ratios between the portfolios are statistically distinguishable. In response to this, Kritzman, Page, and Turkington (2010) presented results where two optimized portfolios, a minimum-variance portfolio (MinVP) and a mean-variance portfolio (MeanVP), delivered higher Sharpe ratios than that of the market portfolio and an EWP in out-of-sample backtesting. They argued for the superiority of optimization by focusing solely on a strategy's Sharpe ratio as an average across the datasets, without any statistical testing.

In addition to the comparison between different optimized portfolios and those constructed in a more naive fashion, there has also been several studies focusing on optimized low-volatility portfolios, such as the MinVP, and their performance compared to a value-weighted market portfolio (VWMP). Recently, these research efforts examining low-volatility investing has immensely increased in popularity. Many papers find that stock portfolios with low volatility or low beta deliver higher risk-adjusted return, than those with higher volatility or higher-beta. These findings are contradictory to traditional financial assumptions, which suggest that taking on more risk should offer a higher expected return. Several findings also suggest that these low-risk asset-allocation strategies perform better than the VWMP, in terms of Sharpe ratio. This surprising empirical finding is often termed as the low-volatility anomaly in the literature.

In this thesis I intend to replicate and extend the existing studies by DeMiguel et al. (2009) and Kritzman et al. (2010) by examining the out-of-sample performance of four different portfolio models. These are the minimum-variance and mean-variance optimized portfolios, in addition to the EWP and the VWMP. The performance of these portfolios will be evaluated based on their out-of-sample Sharpe ratio. In addition, the mean excess return, volatility, and capital accumulation of every implemented strategy will be examined. These portfolio models will be compared based on their out-of-sample performance across eight different datasets. I will replicate certain parts of the findings in Kritzman et al. (2010) and extend on these results by testing the statistical significance of the difference between the out-of-sample Sharpe ratios of the portfolios. Kritzman et al. (2010) and several other studies conclude that certain strategies outperform in terms of Sharpe ratio without testing if the difference is statistically distinguishable from the other Sharpe ratios. While the numerical values of the Sharpe ratios alone can be sufficient for a lot of practitioners, I seek to draw my conclusions from a statistical and scientific standpoint. In addition, I implement the same out-of-sample time period used in DeMiguel et al. (2009) in order to examine if my results are similar to theirs.

As revealed in the literature, the empirical performance of optimized portfolios is clearly ambiguous. Considering the numerous conflicting results, it is plausible that a certain strategy's out-of-sample performance is closely related to a particular dataset or time period. Although a strategy appears to be superior during an isolated time period and dataset, one should be cautious with interpreting this result as though it holds in general. In this regard, I will investigate if the choice of out-of-sample period has any impact on the performance of the implemented strategies. This is done by testing the out-of-sample performance of the portfolios during five additional time periods. The initial out-of-sample time period will be from 1961 to 2012 and from then the time period will decrease by a decade at the time, until the final time period from 1991 to 2012 is reached.

The comparison of the portfolios will enable me to observe if optimized portfolios deliver better empirical performance than portfolios that require no historical information, such as the EWP and the VWMP. In this process I can also investigate the low-volatility anomaly by comparing the performance of the MinVP to portfolios with higher volatility.

The remainder of the thesis is organized as follows. Section 2 will give a review of theory and literature that is relevant for the research problem and the framework of this thesis. In Section 3 I present the data that has been used in the empirical study and from where it has been obtained. The methodology I have followed for portfolio construction and empirical backtesting will be explained in Section 4 of the thesis. Section 5 presents the empirical results that were obtained from this study, while the discussion of these results will be conducted in Section 6. Finally, a conclusion of this empirical study and its findings will be given in Section 7. The R programs in addition to the tables that have been omitted from Section 6 will be presented in the Appendix.

2 Review of relevant theory and literature

2.1 Modern portfolio theory and capital asset pricing theory

As an introduction to the literature, it is beneficial to start with the theoretical framework that modern finance is based upon. According to modern portfolio theory pioneered by Markowitz (1952) an investor should invest in the portfolio that is mean-variance optimal. This statement assumes that the investor cares only about the expected return offered by the portfolio and the risk attributed by holding this portfolio. When the mean-variance assumptions hold, the investor would only be interested in the asset allocations that offer the highest expected return for a given amount of risk. If one considers only risky assets, these mean-variance efficient portfolio allocations make up the efficient frontier. Presented graphically in an expected return-standard deviation space, the efficient frontier is the upper part of the hyperbola of feasible portfolio allocations. The exact allocation chosen by the investor is based on his or her tolerance of risk, or level of risk aversion.

If introducing a risk-free asset that offers a given return at no risk as well as the ability to lend and borrow at the risk-free rate, the efficient frontier shifts from the hyperbola to the optimal capital allocation line (CAL) which is a straight line drawn from the risk-free asset and is a tangent to the hyperbola of risky assets. This tangent point is the optimal combination of risky assets and is known as the tangency portfolio. Every mean-variance rational investor should invest in a portfolio somewhere along this CAL, which is attainable by holding a certain weight of the portfolio in the risk-free asset and a certain weight in the tangency portfolio. The expected return of investing in a portfolio p somewhere along the CAL can be expressed by the following equation:

$$E[r_p] = r_f + \sigma_p \frac{E[r_{tan}] - r_f}{\sigma_{tan}}, \tag{2.1}$$

where r_f is the risk-free rate of return, σ_p is the standard deviation of returns from portfolio p, σ_{tan} is the standard deviation of returns from the tangency portfolio and $E[r_{tan}]$ is the expected return on the tangency portfolio.

Investing on the optimal CAL implies that every investor allocates his or her wealth in a portfolio that offers the best tradeoff between risk and return. The Sharpe ratio is a measure that quantifies this tradeoff. Originally introduced in Sharpe (1966) as the reward-to-variability ratio and later revised in Sharpe (1994) to hold for any benchmark, it has during the times come to be known as simply the Sharpe ratio. Formally, the Sharpe ratio for a portfolio p is given as:

$$SR_p = \frac{E[r_p] - r_f}{\sigma_p}. \tag{2.2}$$

As pointed out in Bodie, Kane, and Marcus (2011) the slope of the CAL equals the increase in the expected excess return of portfolio p per unit of additional standard deviation. When studying Equation (2.2) it is apparent that the Sharpe ratio expresses this same relationship and that the optimal Sharpe ratio is indeed the slope of the optimal CAL. Therefore a portfolio that is allocated along the CAL offers the highest Sharpe ratio.

Where on the CAL the investor will allocate his or her portfolio is determined by the individual investor's utility function. This utility function is ultimately shaped by the investor's risk aversion. Based on the investors utility curves and using leverage, the investor will choose a position somewhere on the CAL that offers the highest Sharpe ratio. Investors who do not want to take on too much risk, will choose an allocation along the CAL closer to the the risk-free asset. Conversely, investors who are very risk tolerant can choose to borrow at the risk-free rate and thus be able to choose an allocation on the CAL that is beyond the tangency portfolio. This implies a mixture of a negative weight in the risk-free asset and a weight larger than 100 percent of the initial capital in the tangency portfolio.

Traditional financial theory states that portfolio returns can be explained by the Capital Asset Pricing Model (CAPM), given that certain assumptions are fulfilled. The development of the CAPM is often attributed to Sharpe (1964), Lintner (1965), and Mossin (1966). CAPM theory states that a portfolio's expected return can be expressed as the sum of the risk-free rate and the product of the expected market premium and the portfolio's beta risk exposure. This beta coefficient can be interpreted as the systematic risk associated with the asset. The CAPM can therefore be formulated as follows:

$$E[r_p] = r_f + \beta_p(E[r_m] - r_f), \tag{2.3}$$

where β_p is the beta coefficient of portfolio p and $E[r_m]$ is the expected return on the market portfolio. According to CAPM theory, every investor will allocate a part of his or her wealth in the tangency portfolio. Because of this, the tangency portfolio must be equal to the market portfolio of risky assets which consists of all existing assets weighted by their market capitalization (Asness, Frazzini, and Pedersen, 2012). This implies that investing in the market portfolio delivers the optimal Sharpe ratio, and thus every investor should allocate a part of their wealth in this portfolio.

Based on this theoretical framework one would expect that, if markets are efficient, investing in the real-world equivalent to the VWMP would be the allocation strategy that has the highest Sharpe ratio. However, the empirical literature suggests that there are many asset-allocation strategies that perform better than the market portfolio. The rest of this section will review several studies that investigate how different asset-allocation strategies perform out-of-sample and also finds that investing in the VWMP does not provide the highest Sharpe ratio.

2.2 Low-volatility anomaly

A popular strand of research is that which centers on low-volatility portfolios. Because the MinVP and its performance relative to the market is often central in these studies, they are interesting for this thesis. Low-volatility strategies are portfolios consisting of less risky assets, with the purpose of lowering the portfolios volatility. Based on traditional assumptions about the risk-reward relationship, such strategies would be expected to deliver lower risk-adjusted returns than their more volatile siblings. This is based on the notion that when taking on more risk one would expect to be compensated by earning higher returns. The same is also expected to hold for another important factor in low-volatility investing, namely low-beta portfolios. This is because, according to the CAPM, portfolios with high beta have higher expected returns than portfolios with low beta. Contrasting this theoretical framework, many studies have uncovered anomalies to this risk-reward relationship. By empirically testing low-beta strategies out-of-sample, they find that such portfolios often deliver equal or higher risk-adjusted returns than high-beta strategies.

The debate concerning the CAPM and its prediction of the risk-reward relationship is not new. Both Black, Jensen, and Scholes (1972) and Haugen and Heins (1975) critiqued the CAPM in their studies, in which they found that low beta stocks deliver higher risk-adjusted returns than high beta stocks. In more recent times, the topic of low-volatility strategies has become more frequently mentioned and several papers reach the same conclusions. Empirical evidence demonstrating that low-volatility stocks perform as well as, or better than, the market, seem to be ever increasing.

Haugen and Baker (1991) performed a study where a MinVP of the 1000 largest U.S. stocks outperformed the VWMP in terms of higher return and lower volatility, further contradicting the notion of higher risk offering higher expected return. Studies by Jagannathan and Ma (2003) and Clarke, de Silva and Thorley (2011) both found similar results in that the MinVP delivered higher returns while having a lower realized volatility, when compared to a value-weighted benchmark. Baker and Haugen (2012) showed in a study that low-risk stocks deliver have higher returns than riskier stocks, for equity markets in 21 developed countries and 12 emerging markets.

The literature suggests several different explanations as for why the MinVP outperforms the market. One approach is to implement factor models in an attempt to explain the returns by their exposure to different sources of risk. Blitz and van Vliet (2007) found that there was still significant alpha present in their low-volatility portfolios after controlling for size, value and momentum effects. Thus they concluded that regression analysis with classical risk factors could not explain the volatility effect in full. They also noted that low-volatility stocks had low betas. Scherer (2011) found that the returns of the MinVP in excess of the returns of the VWMP can be attributed to Fama/French risk factors. The paper also discusses that by constructing a portfolio that loads up on certain risk factor will lead to a statistically significant outperformance over the MinVP.

Others argue from a behavioral standpoint. Baker, Bradley, and Wurgler (2011) found that regardless of whether risk was specified as beta or volatility, both low-beta and low-volatility portfolios outperformed their higher risk counterparts. They argue that the low-risk portfolio prevails through time because institutional investors focus on the benchmark and the information ratio, instead of the benchmark-free Sharpe ratio. This way the mispricing will not be arbitraged away. A similar conclusion is drawn by Brennan, Cheng, and Li (2012) who suggest that the presence of high tracking error in low-volatility stocks makes low-volatility portfolios unattractive for portfolio managers.

Another recently popularized low-volatility asset-allocation worth mentioning is the risk parity strategy. Although there are several approaches to constructing a risk-parity portfolio (RPP), the general objective is to weight each asset in proportion to their risk so that every asset will have an equal risk contribution to the total risk of the portfolio. This way, the portfolio overweights less volatile assets and underweights assets with higher volatility. An advantage of the RPP, similar to that of the MinVP, is that it only requires the covariance matrix in its construction. Asness et al. (2012) compared the historical performance of a VWMP, a 60/40 stock/bond portfolio and a RPP for three different datasets. They found that the RPP delivers a superior Sharpe ratio compared to the other strategies, although it provided lower average returns. The authors discuss the effects of real-life leverage constraints and that even though the RPP is superior in terms of Sharpe ratio, investors may refrain from pursuing a risk-parity strategy due to the lower average returns. However, an investor willing and able to apply leverage can benefit by investing in a portfolio that overweights low-beta assets and underweight high-beta assets and applying leverage to this portfolio.

Frazzini and Pedersen (2014) finds that portfolios of low-beta assets deliver higher alpha and Sharpe ratios, than portfolios of high-beta assets. They also find that the security market line is flatter than the relationship suggested by the CAPM not only for U.S. equities, similarly to the results of Black et al. (1972), but that this also holds for international equity markets. They suggest that the good performance of low-beta portfolios can be exploited by a betting against beta (BAB) factor, which is a portfolio that holds low-beta assets leveraged to a beta of one, and shorts high beta-assets deleveraged to a beta of one. Because of the BAB factor rivals standard asset pricing factors such as the size, value and momentum factors in terms of robustness, statistical significance and robustness, the authors suggests that this is a important factor for cross-sectioning portfolio returns.

Chow, Hsu, Kuo, and Li (2013) provides a comprehensive survey of low-volatility strategies. The paper points out that since the global financial crisis, low volatility portfolios based on U.S assets have outperformed the market by delivering higher returns and Sharpe ratios, with only two-thirds the volatility risk. In their study they also found that low volatility portfolios generally deliver superior returns in the long term across several countries. In terms of the different low volatility strategies, they do not find that one construction method is better than the other from a return perspective.

2.3 Research on the performance of optimized portfolios

Because the minimum-variance and the mean-variance portfolios are part of the Markowitz mean-variance optimization framework, they can be considered as optimized portfolios. The following is a review of studies that focus their research problem on the performance of portfolio optimization compared to more naively implemented strategies that require no preliminary estimation.

By studying 14 different portfolio models on seven datasets of monthly returns, DeMiguel et al. (2009) found that none of the various mean-variance optimized models delivered consistently better performance than an equal-weighting strategy in terms of Sharpe ratio, certainty-equivalent return and turnover. They implemented several portfolio models where some had no constraints, and others with long-only or shrinkage constraints. The performance of these portfolio models were then compared to the performance of the EWP. In terms of Sharpe ratio alone, they find that the EWP delivers higher or statistically indistinguishable Sharpe ratios compared to constrained strategies, which in turn were higher than unconstrained strategies. The time periods used for estimating the needed parameters for the particular allocation strategy consisted of rolling 60- and 120-month windows. The main out-of-sample period implemented in their study was from July 1963 to November 2004. The authors relate the poor performance of the optimized portfolios relative to the equal-weighting strategy to estimation error. In an analytic study of the estimation error they find that based on data from the U.S. stock market data, a portfolio of 25 assets would require an estimation window of more than 3000 months for the sample-based mean-variance strategy to outperform the EWP. If the number of assets is increased to 50, the required estimation window would double.

In an out-of-sample comparison of a long-only MeanVP and an EWP, Duchin and Levy (2009) found that when the portfolios consisted 15, 20 and 25 assets the naive diversification strategy delivered higher average returns. However, for a portfolio consisting of 30 assets the MeanVP had higher average returns than the EWP. They conclude that the naive strategy is better for portfolios of relatively few assets, while the optimized portfolio strategy will outperform the EWP when the number of assets is relatively high. They also found that, in-sample, the MeanVP is superior regardless of the number of assets.

Kritzman et al. (2010) argue against the many findings where the EWP outperforms optimized portfolios. Using what they define as naive but plausible estimates of expected returns, volatilities and correlations, they find that optimized minimum-variance and mean-variance portfolios deliver superior out-of-sample performance, compared to both the VWMP and an EWP. In order to demonstrate that optimization outperforms even in simple forms, the authors only implemented a long-only constraint on the portfolios. The authors' critique as to why the optimized portfolios of other papers deliver poor performance is due to too short sample intervals for mean estimation. Because mean estimates based on trailing 60 or 120-month historical samples might lead to implausible assumptions about future returns, these estimates might cause suboptimal asset allocations. To circumvent this, they obtain their mean estimates by simply computing the mean return of the previous 50 years of historical data, for the relevant dataset. These mean return estimates are held constant for the out-of-sample implementation. The covariance matrix is estimated based on lookback periods of 5, 10 and 20 years, in addition to an expanding-window approach. Eight datasets are considered and the out-of-sample Sharpe ratios for the implemented portfolio strategies are averaged across the datasets. The out-of-sample period used in this paper is from 1978 to 2008. Additionally, they argue that the effect of estimation error is exaggerated. The methodology implemented in this study is not very clear. In addition, the authors proclaim the superiority of optimization without testing for the statistical significance of the difference between Sharpe ratios. This may render the conclusion inaccurate from a statistical standpoint.

2.4 Estimation error

Because the true future values of the means, variances and covariances are not known, the MinVP and MeanVP implemented for out-of-sample testing are dependent on parameter estimation in their construction. These estimates can be obtained using historical information or expectations about the future, and thus the precision of these estimates is uncertain. Due to the findings of many researchers that will be discussed in the sequel, it is reasonable to believe that the performance of the optimized portfolios will be suboptimal when tested out-of-sample. This is often attributed to estimation error, which occurs because of the difference between the estimated parameters and their realized values.

Michaud (1989) reported that unconstrained mean-variance optimization can lead to suboptimal or financially unwise asset allocations. In addition, these portfolio are often significantly outperformed by equal-weighting strategies. He mentions that the term "error maximization" is often used to refer to mean-variance optimization because small estimation errors in the input estimates can lead to large output errors. The paper proposes that introducing several constraints, such as no short-selling and shrinkage estimators can help enhance mean-variance optimization.

Chopra and Ziemba (1993) found that misspecifications of means are more severe than errors in covariances. They suggest that estimates of covariances are the least critical in terms of the errors influence on the optimal portfolio. The relative importance of errors also depends on the investor's risk tolerance. The authors conclude that due to estimated means being more prone to error, investors seeking to minimize the estimation error should consider removing the need for mean return as an input in the portfolio construction. This could be done by considering a minimum-variance allocation as portfolios based on variance and covariance estimates are less affected by errors.

Ledoit and Wolf (2004) suggest that a sample covariance matrix constructed based on historical information, contains errors due to extreme observations and periodical variations. This leads to extreme observations that should be accounted for. The authors propose a new way of estimating the covariance matrix, by shrinking the sample covariance matrix towards the constant-correlation matrix. This is a method that can be implemented in order to reduce the estimation error in the covariance matrix. In the study, they also test the out-of-sample performance of the proposed shrinkage method. They conclude that their suggested shrinkage estimator outperforms the other alternatives. This thesis will implement such a strategy in order to reduce the error in estimation. This will be explained more thoroughly in Section 4. In addition, DeMiguel et al. (2009) finds that portfolios computed with long-only constraints combined with shrinkage usually are more effective in reducing estimation error. In fact, they find that a long-only minimum-variance portfolio with shrinkage delivers better Sharpe ratio than the EWP in five out of seven datasets. Although, only statistically distinguishable in one of the cases.

Opposing these findings, Best and Grauer (1991)Best and Grauer (1991) and Kritzman (2006) argues that the effects of estimation error is not as serious as much of the literature suggests. Although the asset weighting scheme can differ a great deal from the truly optimal weights, the resulting return distribution will not be very different. Kritzman (2006) concludes that the concern for estimation error is exaggerated.

3 Data

The data used in this study is obtained from the Kenneth French online data library,[1] which have been collected and sorted by Fama and French. I will utilize eight different datasets as the investment universe for the MinVP, MeanVP and EWP, each consisting of monthly returns. Six of the datasets covers the time period of January 1927 to December 2012, while the two remaining datasets starts in January 1928 and January 1931, respectively. The relevant datasets along with their purpose, abbreviation and available time periods are summed up in Table 3.1. An additional dataset used is the Fama/French factors which serve multiple purposes. The Fama/French factors dataset contains the risk-free rate of return and the value-weighted market premium. The value-weighted market premiums are utilized when constructing the VWMP and the risk-free rate is subtracted from the three other portfolios in order to obtain the excess returns.

Fama and French constructed the datasets by including all relevant NYSE, AMEX and NASDAQ stocks and sorting them into portfolios based on different properties. This implies that the study will be limited to assets based on U.S. stocks. From this online library I use returns sorted by industry, size, book-to-market, momentum, dividend yield, short-term reversal and long-term reversal to function as the investable universe. Because all the data is collected from the same data source and is restricted to U.S. stocks, they are consistent and comparable. For further details on the construction of each dataset, the reader is encouraged to check out the website in the aforementioned footnote.

The data which serves the purpose of the investable universe are the same as that imple-

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

mented in the study by Kritzman et al. (2010). This was a natural choice as the empirical results of this thesis will be directly comparable to the results of Kritzman et al. (2010).

Table 3.1: Datasets utilized for the empirical study

Dataset	Purpose	Abbreviation	Available time period
10 industries	Investable universe	10 ind	01/1927-12/2012
30 industries	Investable universe	30 ind	01/1927-12/2012
10 size deciles	Investable universe	size	01/1927-12/2012
10 book-to-market deciles	Investable universe	book	01/1927-12/2012
10 momentum deciles	Investable universe	mom	01/1927-12/2012
10 short-term-reversal deciles	Investable universe	short-term	01/1927-12/2012
10 dividend yield deciles	Investable universe	div	01/1928-12/2012
10 long-term reversal deciles	Investable universe	long-term	01/1931-12/2012
Fama/French factors	r_m and r_f	NA	01/1927-12/2012

4 Methodology

This section presents the methodology that has been implemented in order to reach the empirical results. I will show how the MinVP, MeanVP, EWP, and VWMP were constructed for this study, the considerations in relation to moment estimation, how portfolio performance was measured, how the statistical tests are implemented, and which out-of-sample time periods that were considered. Naturally, some of the methods and considerations are heavily influenced by the papers of DeMiguel et al. (2009) and Kritzman et al. (2010). The construction and out-of-sample implementations of the portfolios has been managed by using the free programming language and R, developed by the R Core Team (2013). The specific R code that reproduces my results will be given in the Appendix. In the computation of the portfolio models, transaction costs or taxation related to capital gains will not be considered.

4.1 Description of implemented models and the estimation process

To facilitate the different portfolio strategies implemented in this thesis, it is advantageous to further explore the foundations of which they are built upon, along with the presentation of the notation utilized for the rest of the thesis. Additionally, I introduce the implemented portfolio strategies and how they were constructed along with considerations on how to reduce the possible estimation error. The construction of the different models will be viewed from the perspective of a utility-maximizing investor, who cares only about the ratio between risk and return. This implies that the investor wants to invest in the portfolio strategy that offers the highest Sharpe ratio. According to the CAPM, which was discussed in the literature review, the Sharpe-efficient portfolio is the VWMP. Although other characteristics of the portfolios will be discussed, it will be the Sharpe ratio that ultimately ranks the asset-allocation strategies.

The investor's utility-maximization problem can be stated as a maximization problem where the assets of a portfolio p are weighted in such a way that the investor's utility is maximized:

$$\max_{\mathbf{w}} \; U(r_p) = E[r_p] - \frac{\gamma}{2} Var[r_p], \tag{4.1}$$

where γ is a scalar that represents the investor's risk aversion. As discussed in the literature review, it is the investor's risk aversion that decides where on the CAL the investor will position him- or herself. Graphically, it is where the specific indifference curve produced by a given γ in Equation (4.1) is a tangent to the CAL. Investing in this portfolio is the theoretically optimal allocation for that given investor.

Because the slope of the CAL is constant, all points on the CAL will have the same Sharpe ratio. This implies that the Sharpe ratio will be the same whether the portfolio is invested in only risky assets or if it is partitioned between the risk-free asset and the portfolio of risky assets. Because of this, the Sharpe measure can be implemented as a performance measure for all the portfolios regardless of how the assets are weighted. In the formal presentation of the different portfolio models, which is given later in the thesis, only the MeanVP will be constructed by explicitly considering the investor's level of risk aversion. The MinVP will be the portfolio allocation that offers the highest expected return for the lowest variance for all the risky assets. The inclusion of the VWMP will be the empirical equivalent to the market portfolio from the CAPM theory which is the theoretical tangency portfolio. The EWP will also consist of allocations limited to risky assets. Due to this, the remainder of the formulations and assumptions presented in this section will relate to a scenario where only risky assets are available for investment. The inclusion of a risk-free asset will be discussed in the section concerning the construction of the MeanVP.

Assume that the investor holds a certain amount of wealth and wants to allocate this wealth in a portfolio of N risky assets. The condition is that the investor must be fully invested, so the portfolio weights must sum to unity. This is commonly termed as the budget constraint. The return on the investor's portfolio can now be expressed as:

$$r_p = \sum_{i=1}^{N} w_i r_i, \;\text{ subject to }\; \sum_{i=1}^{N} w_i = 1, \tag{4.2}$$

where r_p is the return on the portfolio, w_i is the weight invested in asset i, and r_i is the return from asset i. Because the future returns are not known beforehand, expected values must be considered:

$$E[r_p] = \sum_{i=1}^{N} w_i E[r_i], \;\text{ subject to }\; \sum_{i=1}^{N} w_i = 1, \tag{4.3}$$

where $E[r_i]$ is the expected value of the return on asset i. The variance of the portfolio return can be expressed as:

$$Var[r_p] = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j Cov[r_i, r_j], \tag{4.4}$$

where $Var[r_p]$ is the variance of the portfolio return and $Cov[r_i, r_j]$ is the covariance of asset i and asset j.

Expected values and variances of the returns can be reformulated so that the expected value of asset i can be expressed as $E[r_i]=\mu_i$ and the variance of asset i can be expressed as $Var[r_i] = \sigma_i^2$. Vector and matrix notation is often utilized in the literature to simplify the mathematical formulation, and this thesis will also rely on such an approach. Then, $\boldsymbol{\mu}$ will denote an $N \times 1$ vector containing the expected returns from all of the the N assets, with $\mathbf{w}$ being an $N \times 1$ vector of the weight invested in each of the N assets. The riskiness of the assets is given by the covariance matrix, $\boldsymbol{\Sigma}$, which is an $N \times N$ matrix with elements consisting of the variance of all the N assets and the pair wise covariance between all the N assets. Now, the vector of expected asset returns, the vector of asset weights and the covariance matrix are given as:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} Var[r_1] & Cov[r_1, r_2] & \cdots & Cov[r_1, r_N] \\ Cov[r_2, r_1] & Var[r_2] & \cdots & Cov[r_2, r_N] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[r_N, r_1] & Cov[r_N, r_2] & \cdots & Var[r_N] \end{bmatrix}.$$

The mean return of a portfolio can now be denoted by:

$$\mu_p = \mathbf{w}'\boldsymbol{\mu}, \tag{4.5}$$

while the variance of the portfolio returns can be expressed as in the following equation:

$$\sigma_p^2 = \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}. \tag{4.6}$$

Because every portfolio model will be weighted differently, the formulas (4.5) and (4.6) holds in general for all portfolios consisting of only risky assets, and will give the unique measures for each respective portfolio. It is assumed that the asset returns are linearly independent and that the covariance matrix is nonsingular.

In this case Equation (4.3) can be expressed as:

$$\mu_p = \mathbf{w}'\boldsymbol{\mu}, \text{ subject to } \mathbf{w}'\mathbf{1} = 1, \tag{4.7}$$

where $\mathbf{w}'\mathbf{1} = 1$ is the budget constraint. In this study I will base the expected return on a sample of the historical mean return. This implies that the notation μ will be used interchangeably to denote both the expected return and the mean return. Table 4.1 displays the

Table 4.1: The different portfolio models implemented for the empirical study

Portfolio model	Abbreviation	Input estimates
Minimum-variance portfolio	MinVP	Covariance matrix
Mean-variance portfolio	MeanVP	Mean returns and covariance matrix
Equally-weighted portfolio	EWP	None needed
Value-weighted market portfolio	VWMP	None needed

four different portfolio models that will be implemented in this study. As shown in the table, the two optimized portfolios are dependent on moment estimation while the EWP and the VWMP do not depend on any preliminary estimation.

4.1.1 Moment Estimation

The MinVP is dependent on the covariance matrix, while the MeanVP needs both the covariance matrix and the mean return when being computed. As one does not know the true realizations of these moments, they must be estimated. In order to obtain certain inputs for constructing the optimized portfolios, a certain period of the dataset must be reserved for estimation. These periods are often referred to as in-sample or lookback periods and can be of different length. While much of the literature reviewed for this study only consider an in-sample period of a fixed length, such as in Blitz and van Vliet (2007) where they implement a rolling-window covariance matrix of 36 months, I will implement several in-sample periods of different lengths and compute the out-of-sample results for each of these cases. This will also reveal if differences in the in-sample length, will affect the out-of-sample results accordingly.

The following is an explanation of the rolling-window approach I have utilized to estimate the both the sample covariance matrix and the sample mean return. In this example, I assume that the allocation strategy starts in 1951, but the same principle holds for every time period tested. Even though the following exemplification utilizes the covariance matrix, the same rolling-window approach also holds for the mean estimation process. The covariance matrix is estimated over a rolling window of T months, where T is set to 60, 120, 240 and an all-data approach. The choice of in-sample length is purposely equal to that of Kritzman et al. (2010). If T=60 the relevant portfolio weights at the time of the investment will be based on the covariance matrix from January 1946 to December 1950. The period for the next covariance-matrix estimate will be rolled one month forward, so that it picks up the month T+1 while discarding the first one. Hence, the new relevant weights are based on the period February 1946 to January 1951. In this manner a new sample covariance matrix will be estimated every month, and the weights of the relevant portfolio will be updated. This means that the portfolio is rebalanced every month.

In the case where T=all data, I will implement an expanding estimation window. In this approach the first covariance matrix estimation will be based on the n months of observations up until the start of the out-of-sample testing. When assuming that the out-of-sample testing starts in 1951 and if the relevant dataset contains historical observations starting from 1927, all 24 years worth of data from January 1927 to December 1950 will constitute the basis for the initial covariance estimate. At time *T+1,* the first month of 1951 will be included but the observation from January 1927 will not be excluded. This way, the estimation window will expand in time along with the portfolio strategy. This implies that the entire history present in the dataset will be considered.

When considering that in theory an investor's expectations reflect the information set that consists of all previous information available up until the current time, and assuming that this information is relevant the all-data expanding approach would reflect this. However, with this expanding-window approach every observation will be deemed as of equal importance towards the future expectations, which might be unrealistic as well as detrimental for the out-of-sample results.

Formulated mathematically, the estimation process is explained in the sequel. First, I assume that the vector of monthly asset returns at time t follows a random walk and is given by:

$$\mathbf{r}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t, \tag{4.8}$$

where $\mathbf{r}$ is the vector of monthly returns, $\boldsymbol{\mu}$ is the vector of mean returns and $\boldsymbol{\epsilon}$ is a vector of random disturbance attributed at time t which is a normally distributed random variable with zero mean and constant variance. Now, the rolling window of estimated means can be specified as:

$$\hat{\boldsymbol{\mu}}_t = \frac{1}{T} \sum_{i=t-T}^{t-1} \mathrm{r}_i, \tag{4.9}$$

while the rolling window of estimated covariance matrices is given as:

$$\hat{\boldsymbol{\Sigma}}_t = \frac{1}{T} \sum_{i=t-T}^{t-1} \boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_i', \tag{4.10}$$

where T is the length of the rolling window.

While many scientific papers implement mean estimates computed by a shorter rolling window of 60 and 120 months, Kritzman et al. (2010) argue that such periods are too short to be economically wise. While implementing rolling covariance matrices of different lengths, they use a constant mean return estimate based on the first 50 years of data when constructing the mean-variance optimized portfolio. In order to achieve results that are directly comparable to that of their study, I will also do this for the out-of-sample period from 1978 to 2008. This implies that the historical mean from the first 50 years of data, or 47 in the case of the "long-term" dataset due to fewer historical observations, will serve as the mean estimate input for the computation of the MeanVP during the Kritzman et al. (2010) replication. The means are assumed to be constant in the entire out-of-sample test in line with Kritzman et al. (2010).

However, for the rest of the out-of-sample periods tested, the means will be estimated in a rolling-window manner.

As mentioned earlier, certain methods can be implemented in order to reduce the effect of the estimation error on the portfolios when working with historical samples. In this thesis, the Ledoit and Wolf (2004) shrinkage of the covariance matrix will be implemented when constructing both the MinVP and the MeanVP. The authors explain that their method will pull extremely high coefficients in the sample covariance matrix downwards and pull extremely low coefficients upwards. This is because extremely high coefficients tend to be estimated with a lot of positive error, while extremely low coefficients tend to be estimated with a lot of negative error. It becomes clear that this is a process that shrinks the extreme values towards the constant-correlation matrix.

The shrinkage estimator proposed in Ledoit and Wolf (2004) can be expressed as:

$$\hat{\Sigma}_{Shrink} = \hat{\delta}^* F + (1 - \hat{\delta}^*)\hat{\Sigma}, \tag{4.11}$$

where $\hat{\Sigma}_{Shrink}$ is the shrunk estimated covariance matrix, $\hat{\delta}^*$ is the estimated optimal shrinkage constant which is a number between 0 and 1, F is the sample constant-correlation matrix and $\hat{\Sigma}$ is the sample covariance matrix. The optimal value of δ^* is the one that minimizes the expected value of the quadratic loss function between the shrinkage estimator and the true covariance matrix. Put more simply, minimizing the distance between the shrinkage estimator and the true covariance matrix. This constant is an estimate which is why it is denoted with a hat in Equation (4.11). The shrinkage estimator assumes that asset returns are independent and identically distributed over time and have finite fourth moments. When implementing the shrinking of the covariance matrix in the statistical software R, I utilize the package constructed by Burns Statistics (2012), which applies the method proposed in Ledoit and Wolf (2004). Because the paper by Kritzman et al. (2010) did not apply shrinkage estimates to the covariance matrix, I will not implement the technique explained above during the out-of-sample replication period from 1978 to 2008.

I will expect that the estimation error in the inputs will be less severe due to several reasons:

- If the amount of investable assets are large compared to the available observations of historical returns, the sample covariance matrix is estimated with a lot of error. This is pointed out in Ledoit and Wolf (2004). Because the datasets I implement in this study only consist of 10-30 different investable assets, I would expect that the error in the estimates are somewhat lower than if this was a problem.

- All the portfolio models will implement a long-only approach. As pointed out by Jagannathan and Ma (2003), restricting short selling does reduce estimation error to some degree.

- Because I implement a shrinkage estimator I would expect that the estimation error will decrease since extreme values from the estimation process are shrunk towards the constant-correlation matrix

4.1.2 Minimum-variance portfolio

The MinVP is the portfolio that offers the lowest variance possible of all risky assets on the efficient frontier. Because the minimum-variance portfolio depends only on the covariance matrix as an input, there is no estimation error contributed from the sample mean. Although estimation error also exists in the sample covariance matrix, Chopra and Ziemba (1993) found that this error is significantly less than that generated when estimating the means. This implies that there should be less estimation risk associated to the MinVP. The MinVP that is constructed in this paper will have certain constraints. This is done in order to make it a more realistic strategy for an investor as well as reduce estimation error. The first constraint will be that the investor's position in the different assets will be long-only. This means that the investor cannot short-sell assets that are expected to give negative returns in the future. Chopra and Ziemba (1993) and Jagannathan and Ma (2003) discuss the benefits of such a constraint. Other reasons for why this constraint is reasonable is that real life investors may not always be permitted to taking short positions in certain assets, not to mention that short selling is a risky endeavor.

When constructing the MinVP, I will not introduce any constraints on how much of the total wealth that can be invested in a single asset. Of course, this can lead to extreme weights in certain assets and a portfolio consisting of very few assets. However, I implement this approach for several reasons. In the papers by DeMiguel et al. (2009) and Kritzman et al. (2010) there are no limitations on the weights. As my methodology is closely related to theirs I will disregard this as well. The second reasoning is that because the datasets used already consists of many stocks sorted into a certain amount of assets. This implies that the MinVP invests in assets that already consist of several single securities. This will also make sure that the portfolio has a certain degree of risk diversification, even though the weights are limited to a few assets of the dataset. Additionally, the primary objective of a MinVP is to achieve minimum variance and not a certain degree of diversification.

If the MinVP allocation were to be rational from a theoretical standpoint, it would have to be the portfolio which offered the highest Sharpe ratio. This would hold if one assumes that the expected return would be equal for every asset. In this scenario, it is obvious that the reasonable thing to do for a mean-variance optimizer is to choose the assets with the lowest standard deviation. The MinVP would then be a point on the CAL that offers the highest Sharpe, and the investor should allocate his wealth somewhere along this CAL.

As the name implies, the MinVP is constructed by finding the weights that return a portfolio with the lowest variance of all feasible portfolios. The weights are found by solving the following minimization problem, expressed in matrix notation:

$$\min_{\mathbf{w}} \quad \sigma_p^2 = \frac{1}{2}\mathbf{w}'\mathbf{\Sigma}\mathbf{w}, \text{ subject to } \mathbf{w}'\mathbf{1} = 1, \tag{4.12}$$

where $\mathbf{1}$ is an $N \times 1$ vector of ones, $\mathbf{w}$ is an $N \times 1$ vector of portfolio weights, $\mathbf{\Sigma}$ is an $N \times N$ covariance matrix and $\frac{1}{2}$ has been added for mathematical convenience. Solving this minimization problem returns the following solution for the vector of weights:

$$\mathbf{w}^{MinVP} = \frac{\mathbf{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}}. \tag{4.13}$$

Because the true covariance matrix is unknown and I have implemented an estimate that is shrunk towards the constant-correlation matrix, this yields the following formulation for the weights:

$$\mathbf{w}^{MinVP} = \frac{\hat{\mathbf{\Sigma}}_{Shrink}^{-1}\mathbf{1}}{\mathbf{1}'\hat{\mathbf{\Sigma}}_{Shrink}^{-1}\mathbf{1}}. \tag{4.14}$$

At each rebalancing point, the $\mathbf{w}^{MinVP}$ is reestimated and thus providing an updated weighting scheme for the MinVP.

Because the MinVP implemented in this thesis has short-selling constraints, this has to be implemented in the minimization problem. There are now several constraints, namely (i) that the weights must sum to one, such that $\sum_{i=1}^{N} w_i^{MinVP} = 1$, and (ii) the weights cannot be negative, such that $\sum_{i=1}^{N} w_i^{MinVP} \geq 0$. I have solved the problem using a quadratic programming solver developed by Turlach and Weingessel (2013). This optimizer will solve quadratic programming problems on the matrix form min $(-d^{-1}b + 1/2b^{-1}Db)$ with the constraints $A^{-1}b >= b_0$. Considering this, the minimization problem for obtaining the weights for the MinVP can now be specified in the following form:

$$\min_{\mathbf{w}} \quad \frac{1}{2}\mathbf{w}'\hat{\mathbf{\Sigma}}_{Shrink}\mathbf{w}, \text{ subject to } \mathbf{A}'\mathbf{w} >= \mathbf{b}, \tag{4.15}$$

where $\mathbf{A}$ is an $(m+l)\times N$ matrix and $\mathbf{b}$ is an $(m+l)\times 1$ vector, both consisting of m equality constraints and l inequality constraints. This is the reason that the notation “>=” is used in Equation (4.15), as some constraints are equalities and other inequalities. By solving this quadratic program I obtain the weights that minimize the variance of the portfolio. Of course, in the case where the MinVP was constructed without the shrinkage estimator the covariance estimate $\hat{\mathbf{\Sigma}}_{Shrink}$ was replaced with the sample covariance estimate $\hat{\mathbf{\Sigma}}$.

4.1.3 Value-weighted market portfolio

Because the true market portfolio is unattainable in real-life financial markets, large value-weighted indices often hold as market portfolio proxies. The proxy for the market portfolio will in this case be constructed by utilizing the market returns provided by Kenneth French, which is computed by taking the return of all the stocks on the NYSE, AMEX and NASDAQ stock exchanges sorted by their market capitalization, as mentioned in Section 3. If one believes the CAPM assumptions to hold, the theory tell us that the optimal strategy for the investor is to hold the market portfolio of risky assets. Based on the investor's preferences, he will be best off when holding a mixture of the risk-free asset and the VWMP of risky assets. Therefore, the VWMP would be expected to be the best performing portfolio according to the CAPM, and the out-of-sample Sharpe ratio of this portfolio should be superior to the other models considered.

By investing in the VWMP one does not pay attention to expected means, variances or covariances, as I would expect the VWMP to already reflect this information. This strategy does not require any portfolio rebalancing, as it should reflect the movement of the market. That implies that this is a buy-and-hold or passive investment strategy. An advantage to this is that there is no need to do consecutive active trading. This is due to the portfolio weights being automatically rebalanced as the price movements of the stocks reflect the current market trend. The only time the investor needs to actively manage the portfolio is when new stocks are to be implemented into the portfolio. This implies that the VWMP has very low turnover, and thus there are very little transaction costs.

To construct the total returns on the VWMP, I have added the risk free rate to the market premium, both measures found in the factor data from the Kenneth French online data library. By its construction the VWMP includes every risky asset available in the investment universe, where each asset is weighted in accordance to their market capitalization. In this study I have chosen an allocation such that all of the investor's wealth is invested in the VWMP and none is allocated to the risk-free asset.

Black et al. (1972) and Haugen and Heins (1975) are among several studies concluding that the CAPM does not hold empirically, and thus not being mean-variance efficient. This can be a reason as for why several studies find that the VWMP does not perform as well out-of-sample.

4.1.4 Equally-weighted portfolio

The equally-weighted portfolio is constructed so that one invests an equal percentage in every available asset. This way, each of the N risky assets are weighted in a manner so that $w_i^{EWP} = \frac{1}{N}$ and the weights sum to unity such that $\sum_{i=1}^{N} w_i^{EWP} = 1$. Obviously, investing in the EWP does not require any moments estimation or optimization, thus there is no estimation error from inputs. Several contributions to the literature suggest that the naively diversified EWP is a better benchmark to assess the performance of other asset-allocation strategies. DeMiguel et al. (2009) concludes that because of the simplicity and low implementation costs the strategy it would serve as a natural benchmark. The empirical evidence suggesting that the EWP performs well out-of-sample is also a convincing addition as to why the strategy should be considered as a benchmark. Therefore, when testing for the statistical distinguishability between Sharpe ratios, I will use the Sharpe ratio delivered by the EWP as the benchmark. As previously mentioned, multiple research efforts suggest that popular asset-allocation strategies do not produce significantly better Sharpe ratios compared to the EWP. By using the EWP as a benchmark I can examine this as well.

The literature mentions several theories on why the EWP performs so well out-of-sample. The following reasonings are mentioned in DeMiguel et al. (2009) and Kritzman et al. (2010):

- Because the estimation of parameters carries with it estimation error, optimization will be suboptimal. The loss in performance due to the naive diversification may be less than of the optimization, and thus EWP could perform better out-of-sample.
- Easy to implement with no prerequired estimation.
- As the strategy weights all assets equally, the portfolio is diversified and there are no concentrated positions.
- The strategy will never perform worse than the worst performing asset, while it will invest in the best performing asset.
- The investable assets provided by the datasets are from 10 to 30 portfolios that already have been presorted based on several individual stocks. This makes them diversified portfolios with lower idiosyncratic volatility than individual stocks. This will probably reduce the loss from an equally-weighted strategy compared to the optimal weighting, and thus the performance from the EWP will be closer to that of the other portfolio models.

4.1.5 Mean-variance portfolio

I will implement the mean-variance model of Markowitz (1952), which optimizes the risk-return tradeoff based on an investor's utility function. Because the construction method is based on the principles of modern portfolio theory, the empirical behavior of this strategy will be interesting and one should expect it to also perform well empirically. The optimal allocation is found given the investors risk aversion. The portfolio weights are optimized so that the portfolio achieves a return that is satisfactory for the investor at the lowest variance feasible. In addition to this, the portfolio needs estimates of mean return and the covariance matrix for construction.

Now assuming that there exists $N+1$ assets, the investor will partition his or her wealth between N risky assets and the additional asset 0 which is the risk-free asset that offers the risk-free rate of return, r_f, so that the portfolio's total expected return is given as:

$$\mu_p = w_0 r_f + \sum_{i=1}^{N} w_i E[r_i], \text{ subject to } \sum_{i=0}^{N} w_i = 1, \tag{4.16}$$

where w_0 is the weight invested in the risk-free asset. Because the amount invested in the risk-free asset can be expressed as $w_0 = 1 - \sum_{i=1}^{N}$, I can rid the equation of the budget constraint so that Equation (4.16) can be rearranged such that:

$$\mu_p = r_f + \sum_{i=1}^{N} w_i E[r_i] - \sum_{i=1}^{N} w_i r_f = r_f + \sum_{i=1}^{N} w_i (E[r_i] - r_f), \tag{4.17}$$

or in matrix notation:

$$\mu_p = \mathbf{w}'(\boldsymbol{\mu} - \mathbf{1}r_f) + r_f. \tag{4.18}$$

I compute the optimal MeanVP by maximizing the investors utility function. The problem can be specified in matrix notation as:

$$\max_{\mathbf{w}} \; U(\mu_p) = (\mathbf{w}'(\boldsymbol{\mu} - \mathbf{1}r_f) + r_f) - \frac{\gamma}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}. \tag{4.19}$$

The weights of the MeanVP are chosen so that the portfolio with the lowest variance for the target mean return, μ_p, computed by Equation (4.19) is obtained. This is expressed by the following minimization problem:

$$\min_{\mathbf{w}} \; \frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}, \text{ subject to } \mathbf{w}'(\boldsymbol{\mu} - \mathbf{1}r_f) + r_f = \mu_p. \tag{4.20}$$

Solving the minimization problem, returns the weights which are given as:

$$\mathbf{w}^{MeanVP} = \frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \mathbf{1}r_f)}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \mathbf{1}r_f)}. \tag{4.21}$$

As one does not know the true mean and covariance matrix, the sample mean and sample covariance matrix will be used as inputs. This implies that there will be a certain degree of estimation error which might affect this portfolio out of sample. Also in this portfolio the sample covariance matrix has been shrunk. The formulation that has been used in order to obtain the implemented MeanVP is given as:

$$\mathbf{w}^{MeanVP} = \frac{\hat{\boldsymbol{\Sigma}}_{Shrink}^{-1}(\boldsymbol{\mu} - \mathbf{1}r_f)}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_{Shrink}^{-1}(\boldsymbol{\mu} - \mathbf{1}r_f)}. \tag{4.22}$$

I must also take the short-selling constraint into consideration when constructing the MeanVP, so that the portfolio weights are long only. Similarly to the construction of the MinVP, I implement a quadratic programming solver and find the weights by specifying the problem as follows:

$$\min_{\mathbf{w}} \; \frac{1}{2}\mathbf{w}'\hat{\boldsymbol{\Sigma}}_{Shrink}\mathbf{w} - \mathbf{d}'\mathbf{w}, \text{ subject to } \mathbf{A}'\mathbf{w} > \mathbf{b}, \tag{4.23}$$

where $\mathbf{A}$ is now an $l \times N$ matrix and $\mathbf{b}$ is now an $l \times 1$ vector consisting of inequality constraints and $\mathbf{d}$ is a vector of $(\boldsymbol{\mu} - r_f)/\gamma$. Also note that I only use the "$>$" sign as there are no equality constraints. Solving this problem returns the vector of optimized weights for the MeanVP. Additionally, in the case where the MeanVP was constructed without the shrinkage estimator the covariance estimate $\hat{\mathbf{\Sigma}}_{Shrink}$ was replaced with the sample covariance estimate $\hat{\mathbf{\Sigma}}$.

4.2 Measuring portfolio performance

This section will cover the computations of the Sharpe ratio, the mean return, the standard deviation and the capital accumulation. All of these measures will be reported in Section 5 and will reveal important characteristics of the different empirical portfolios.

4.2.1 Mean return and standard deviation

The computation of the standard deviations and excess returns are discussed in the sequel. First, the standard deviation and mean return is found from the time series of the respective asset-allocation strategy's excess return. Next, these measures are converted to annual measures by implementing the fact that t=12. The measures are annualized so that they are more easily comparable to other studies. The annualized mean return of a given portfolio strategies implemented in the thesis is computed in the following manner:

$$\bar{\mu}_p = \hat{\mu}_p 12, \tag{4.24}$$

where $\bar{\mu}_p$ is the annualized mean return of the portfolio and $\hat{\mu}_p$ is the realized monthly mean return of the portfolio. The standard deviation of the portfolio's returns is given as the square root of the portfolio variance:

$$\sigma_p = \sqrt{\sigma_p^2}, \tag{4.25}$$

and the annualized standard deviation of a given portfolio strategy is found in the following way:

$$\bar{\sigma}_p = \hat{\sigma}\sqrt{12}, \tag{4.26}$$

where $\bar{\sigma}_p$ is the annualized standard deviation of the portfolio's returns and $\hat{\sigma}$ is the realized monthly standard deviation of the portfolio's returns.

This process is repeated for every relevant dataset and for every specified lookback period. Eventually, the measures are averaged across datasets, to obtain a more general result.

4.2.2 The Sharpe ratio

The main performance measure for judging the performance of the different portfolios will be the out-of-sample Sharpe ratio. While the Sharpe ratio has also been criticized for its flaws, it is still popularly implemented in the literature because of its ability to compare investment with different risk exposures. This thesis will not be concerned with the limitations of the

Sharpe ratio. Because the mean-variance optimizing investor would prefer the portfolio that offers the highest Sharpe ratio, it serves as an important measure for this study.

When computing the monthly out-of-sample Sharpe ratio of the implemented models, the parameters μ and σ are replaced with their realized equivalents $\hat{\mu}$ and $\hat{\sigma}$. Now, the monthly out-of-sample Sharpe ratio for a portfolio p is given as:

$$S\hat{R}_p = \frac{\hat{\mu}_p - r_f}{\hat{\sigma}_p}, \tag{4.27}$$

where $\hat{\mu}$ is the monthly out-of-sample mean return and $\hat{\sigma}$ is the out-of-sample standard deviation of the portfolio's monthly excess returns. The annualized, out-of-sample Sharpe ratio that is computed based on the out-of-sample performance of a portfolio p, can be expressed as:

$$S\bar{R}_p = S\hat{R}_p\sqrt{12}. \tag{4.28}$$

The Sharpe measures reported in the empirical results section are computed in accordance with Equation (4.28).

Many of the papers that compare Sharpe ratios of different asset allocation strategies do not test if these ratios are statistically distinguishable. As a response to this I will implement such a test. Specifically, I will implement the test discussed in Jobson and Korkie (1981) with the correction pointed out in Memmel (2003). This test works in a pair-wise manner, comparing if the Sharpe ratio of a given strategy is statistically different from that of another strategy. Considering the out-of-sample performance of the portfolios, the null hypothesis that $H_0 : S\hat{R}_1 - S\hat{R}_2 = 0$ is tested where the test statistic is given as:

$$\hat{z} = \frac{S\hat{R}_1 - S\hat{R}_2}{\sqrt{\frac{1}{n}[2(1-\hat{\rho}^2) + \frac{1}{2}(S\hat{R}_1^2 + S\hat{R}_2^2 - 2S\hat{R}_1 S\hat{R}_2\hat{\rho}^2)]}}, \tag{4.29}$$

where $S\hat{R}_1$ and $S\hat{R}_2$ are the monthly out-of-sample Sharpe ratios of portfolio strategy 1 and 2 respectively, n is the sample size, ρ is the out-of-sample correlation coefficient, and $\hat{z}$ is the standard normal distributed test statistic.

In line with DeMiguel et al. (2009) and Kritzman et al. (2010), I also compute the in-sample Sharpe ratio of a mean-variance strategy. This measure is found by treating the out-of-sample period as it was in-sample, and thus I find what would be the optimal allocation for a mean-variance strategy during that period. By doing this I will know how close the asset allocation strategies came to the optimal strategy, which would also be a way of assessing the effect of estimation error. The monthly in-sample Sharpe ratio from the mean-variance strategy is computed in the following manner:

$$SR_{IS} = \sqrt{(\hat{\boldsymbol{\mu}}_{IS} - \mathbf{1}r_f)\hat{\boldsymbol{\Sigma}}_{IS}^{-1}(\hat{\boldsymbol{\mu}}_{IS} - \mathbf{1}r_f)}, \tag{4.30}$$

where $\hat{\boldsymbol{\mu}}_{IS}$ is the monthly in-sample mean return, $\hat{\boldsymbol{\Sigma}}_{IS}$ is the in-sample covariance matrix and

r_f is the in-sample risk-free rate. Additionally, this measure is annualized by multiplying with $\sqrt{12}$.

In addition to reporting the individual Sharpe ratios for every strategy, every T and for all the eight datasets, I also report the average Sharpe ratios across the datasets. This can be exemplified as follows: If I compute the MinVP with T=120 for the eight datasets, I end up with eight different Sharpe ratios. The average of these Sharpe ratios is computed so that I have an average Sharpe for that T. The reason for why this is done is because the results will be more general. If a certain strategy performs very well over a given dataset, this might just be a random occurrence. However, if a strategy consistently performs well over several datasets it is more likely that this strategy actually is better. The average measures are also easier to compare across the different strategies and the different T.

4.2.3 Capital accumulation

As an additional performance measure, I will track the capital accumulation by looking at how the portfolio returns will affect the growth of $1 invested at the start of the out-of-sample period. By doing this, I will see which strategy provides the best return, strictly money wise. Because a mean-variance rational investor would only care about the maximum Sharpe ratio, this implementation might seem unintuitive. However, in the real world with restrictive borrowing an investor might be unable to sufficiently lever his position in the optimal portfolio so that it matches his preferred level of expected return. With this argumentation I feel it is interesting to also look at this factor. As I will implement the portfolio strategies using several datasets, also this measurement will also be reported as an average across the datasets.

4.3 Testing periods

Because the focus of the thesis is on replicating and extending previous studies and to determine the consequences of the choice of the out-of-sample period, the empirical testing will be repeated for several time periods. An overview of the implemented periods is presented in Table 4.2. The first time period tested, will be the replication period from January 1978 to December 2008. Following this, will be the out-of-sample time period used in DeMiguel et al. (2009) which is from July 1963 to November 2004. Because DeMiguel et al. (2009) apply different datasets than those I implement in this study; the results cannot be directly comparable. However, I will implement the same out-of-sample period as them in order to see if the empirical results of my study is different than the conclusions reached in their study.

For the rest of the empirical testing I will implement different out-of-sample periods that are not based on any previous studies. The first will be the full sample period from January 1951 to December 2012. Covering 62 years, this is the longest time period allowed by the "div" dataset after setting aside the first 20 years as the in-sample period for estimation purposes. Because the time period is relatively long and includes both market upturns and downturns, the long-term behavior of the implemented portfolios will be apparent. From then,

Table 4.2: Out-of-sample time periods implemented for determining the empirical performance of portfolios

Time period classification	Time period by date
Kritzman et al. (2010) replication period	01/1978-12/2008
DeMiguel et al. (2009) replication period	07/1963-11/2004
Full sample period	01/1951-12/2012
Subsample	01/1961-12/2012
Subsample	01/1971-12/2012
Subsample	01/1981-12/2012
Subsample	01/1991-12/2012

I decrease the out-of-sample period a decade at a time until the last period from January 1991 to December 2012 is reached. This produces four different out-of-sample subperiods starting in 1961, 1971, 1981, and 1991 respectively. By doing this, I will be able to see if the portfolios perform differently when the out-of-sample period has different starting points.

5 Empirical Results

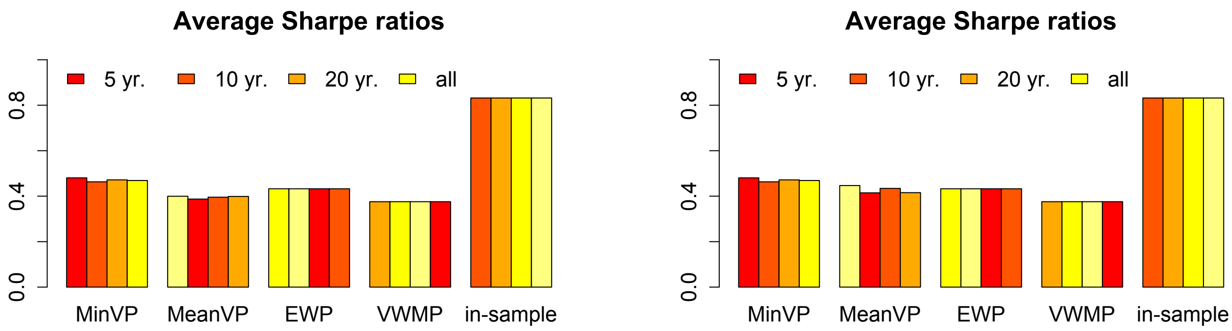
This section will present the results of the empirical study. I will compare the performance of the different asset-allocation strategies in the manner presented in Section 4. The replication of results from previous studies will be introduced first. Results from the procedure of implementing different out-of-sample periods are presented second. The empirical results will be discussed further in Section 6. When presenting the results I will simply use the terms Sharpe ratio, mean returns and standard deviation although I actually refer to the annualized measures.

5.1 Replication of the Kritzman et al. (2010) results

The first empirical backtesting will be the replication of the results found by Kritzman et al. (2010). These out-of-sample results covers the time period from January 1978 to December 2008 and there is no shrinkage of the covariance matrix. In addition the sample mean has been held constant. The results will be presented in the sequel, but the comparison of my findings to those of Kritzman et al. (2010) will be postponed until Section 6.

5.1.1 Sharpe ratios

I will start by considering the Sharpe ratios based on their numerical value and focus on the p-values returned by the statistical test later. Kritzman et al. (2010), along with much of the literature, only report the Sharpe ratios without testing the statistical distinguishability. This can result in erroneous conclusions about a strategy's out-of-sample performance. Figure 5.1 displays graphical comparisons of the out-of-sample Sharpe ratios produced by the different



(a) Constant mean

(b) Rolling mean

Figure 5.1: Graphical comparison of average annualized out-of-sample Sharpe ratios over the period 01/1978-12/2008

strategies when averaged across all eight investable universes. The different colors of the bars reflect the various length of the lookback period T. Of course, the EWP, VWMP, and in-sample MeanVP do not depend on T which is why the bars are of constant height. This idea of graphically comparing the average Sharpe ratios is in line with Kritzman et al. (2010), the results however are not especially similar. The differences in the results will be covered in Section 6. The numerical values of the averages are reported in Table 5.1 where the first row reports the Sharpe ratios when a constant mean was used, while last row reports the Sharpe ratios when a rolling mean was used. Figure 5.1a displays the Sharpe ratios when the MeanVP is constructed with a constant mean. From the graphical comparison it is evident that the optimized MeanVP delivers lower Sharpe ratios than both the MinVP and the EWP, while the VWMP has the lowest of the Sharpe ratios. This ranking holds for every T. Additionally, the in-sample Sharpe ratio is considerably higher than those out-of-sample. This suggests that mean-variance optimization in its theoretical construction delivers better performance than the other strategies, but the estimation error reduces the out-of-sample performance considerably. Another observation from the graphical comparison is that the length of T, when averaged across datasets, seems to have little effect on the empirical Sharpe ratio, except for a few cases.

For the sake of curiosity I also implemented a MeanVP with mean estimates based on a rolling window, as seen in Figure 5.1b, which is the approach I have implemented for the other results. Interestingly enough, this approach delivers a slightly higher Sharpe ratio compared to the MeanVP with a constant mean. Due to Kritzman et al. (2010) suggesting that a rolling mean estimation will result in poor out-of-sample portfolio performance, this result is somewhat surprising. However, this increase in performance does not make a considerable difference to the implications of the results.

When decomposing the analysis of the Sharpe ratios further, it seems as a portfolio's

Table 5.1: Average annualized out-of-sample Sharpe over the period 01/1978-12/2008

MinVP				MeanVP					EWP	VWMP
T=60	T=120	T=240	T=all	T=60	T=120	T=240	T=all	in-sample		
0.4806	0.4632	0.4717	0.4689	0.3996	0.3872	0.3955	0.3988	0.8269	0.4326	0.3757
0.4806	0.4632	0.4717	0.4689	0.4467	0.4147	0.4342	0.4155	0.8269	0.4326	0.3757

Note: The bottom row reports the Sharpe ratios when the MeanVP is constructed with a constant mean as shown in Figure 5.1a, while the row above reports the Sharpe ratios when the MeanVP is constructed with a rolling mean as in Figure 5.1b.

performance is somewhat dependent on the dataset. These results are listed in Table 5.2, stating the Sharpe ratio for every strategy and every dataset in addition to the p-value of the test statistic obtained by Equation (4.29). When determining the significance I consider a 5 percent significance level. The MinVP delivers a higher and statistically distinguishable Sharpe ratio than the benchmark EWP in two cases, when T=120 and T=all data during the "mom" dataset, while the MeanVP delivers a higher and statistically distinguishable Sharpe ratio for every T during the "mom" dataset. This suggests that both of the optimized strategies do very well when the momentum deciles are considered as an investable universe. The MeanVP has very poor performance during the "short-term" dataset, and delivers statistically significant lower Sharpe ratios compared to the EWP in the cases where T=60, 240 and all data. The VWMP has Sharpe ratios that are statistically distinguishable and lower than the EWP in three out of the eight datasets.

5.1.2 Other portfolio measures

I will now consider the empirical standard deviation, mean and capital growth of the implemented portfolio models, as they are reported in Table 5.3. The paper by Kritzman et al. (2010) did not report these measures, so they cannot be directly compared. Yet, I wish to investigate them to get additional information about the portfolios. As observed in the table it becomes clear that the MinVP does indeed deliver the lowest portfolio variance as it has the lowest standard deviation. The VWMP and the EWP end up with being equally volatile, while the MeanVP has the highest standard deviation.

Concentrating on the mean returns, there are some interesting observations. The VWMP actually delivers the lowest of all portfolios with a mean return of 5.89 percent. The MinVP is outperformed by the EWP in all but one case, when T=all data the MinVP has a mean return of 6.97 percent against the 6.75 percent of the EWP. The MeanVP delivers the highest mean return of all the portfolio models, for every length of T.

The last measure reported in the table is the final value of a $1 amount if this was invested in a portfolio at the start of the period. Regarding this capital growth provided by the portfolios, the MeanVP is also superior. Next is the MinVP which delivers the second-most pleasing capital gains, except for the one case where T=120 and the EWP delivers a

Table 5.2: Out-of-sample Sharpe ratios and p-values over the period 01/1978-12/2008 without shrinkage of the covariance matrix

Portfolio strategy		10 ind	30 ind	size	book	mom	short-term	long-term	div
EWP		0.4603	0.4334	0.4203	0.4749	0.3305	0.3751	0.5025	0.4637
MinVP	T=60	0.4985	0.4269	0.3900	0.5188	0.3964	0.4357	0.5797	0.5984
		(0.71)	(0.96)	(0.67)	(0.43)	(0.24)	(0.20)	(0.11)	(0.15)
	T=120	0.5116	0.4503	0.3985	0.4470	0.4381	0.4439	0.5366	0.4799
		(0.64)	(0.89)	(0.77)	(0.53)	(0.03)*	(0.16)	(0.48)	(0.86)
	T=240	0.5439	0.5606	0.3986	0.4397	0.4013	0.4150	0.5260	0.4880
		(0.42)	(0.27)	(0.80)	(0.46)	(0.14)	(0.41)	(0.58)	(0.80)
	T=all data	0.5025	0.5295	0.3590	0.4335	0.5155	0.4478	0.4945	0.4688
		(0.66)	(0.41)	(0.54)	(0.28)	(0.00)*	(0.07)	(0.87)	(0.85)
MeanVP	in-sample	0.7527	1.0392	0.5416	0.7075	1.3353	0.9529	0.6517	0.6734
	T=60	0.4214	0.2930	0.3219	0.4601	0.5613	0.1814	0.4641	0.4934
		(0.66)	(0.12)	(0.20)	(0.83)	(0.02)*	(0.02)*	(0.70)	(0.95)
	T=120	0.3866	0.2868	0.3232	0.4078	0.5888	0.2234	0.4743	0.4071
		(0.39)	(0.08)	(0.22)	(0.32)	(0.01)*	(0.07)	(0.79)	(0.46)
	T=240	0.3943	0.3163	0.3507	0.4179	0.5926	0.1943	0.4702	0.4276
		(0.43)	(0.16)	(0.39)	(0.40)	(0.02)*	(0.03)*	(0.73)	(0.61)
	T=all data	0.4630	0.3213	0.3145	0.4287	0.5690	0.1612	0.4279	0.5045
		(0.97)	(0.27)	(0.14)	(0.51)	(0.04)*	(0.01)*	(0.38)	(0.57)
VWMP		0.3757	0.3757	0.3757	0.3757	0.3757	0.3757	0.3757	0.3757
		(0.03)*	(0.27)	(0.47)	(0.03)*	(0.25)	(0.98)	(0.00)*	(0.13)

Note: The p-value of the difference between the Sharpe ratio of each portfolio strategy from that of the EWP benchmark is reported in the parentheses. Those that are significant on a 5 percent significance level are marked with *.

Table 5.3: Portfolio standard deviations, means and capital gains over the out-of-sample period 01/1978-12/2008

Portfolio strategy		Standard deviation	Excess mean return	Growth of $1
MinVP	T=60	13.87%	6.63%	35.17
	T=120	14.02%	6.48%	32.34
	T=240	14.08%	6.59%	33.47
	T=expanding	14.89%	6.97%	36.33
MeanVP	T=60	17.73%	6.88%	36.75
	T=120	17.81%	6.83%	37.81
	T=240	18.30%	7.14%	41.91
	T=expanding	18.97%	7.30%	41.72
EWP		15.67%	6.75%	32.74
VWMP		15.67%	5.89%	24.58

$0.40 higher gain. Also in this area, the VWMP displeasing by delivering the lowest capital gain of the portfolio strategies.

5.2 Application of the time period used in DeMiguel et al. (2009)

5.2.1 Sharpe ratios

Studying Figure 5.2 and Table 5.4 reveals that some the portfolios behave a little differently than in the previous out-of-sample time period. The VWMP is still the portfolio with the worst performance in terms of Sharpe ratios and the EWP still maintains a Sharpe around 0.43. However, the performance of the MinVP is in general worse than both the EWP and the MeanVP. The MeanVP performs better during this period and is the best performing portfolio of the four. The in-sample MeanVP Sharpe ratio is still considerably higher than the rest suggesting that large parts of the gains from optimization has disappeared due to estimation error.

Moving on to the results presented in Table 5.5 shows that the MinVP delivers a statistically higher Sharpe ratio than the EWP for the “mom” dataset while the EWP delivers a statistically higher Sharpe ratio than the MinVP for the “book” data. In both cases the length of the in-sample period is T=all data. Again, the MeanVP delivers statistically distinguishable and higher Sharpe ratios for all T using the “mom” dataset, although nowhere else.

Table 5.4: Average annualized out-of-sample Sharpe ratios over the period 07/1963-11/2004

MinVP				MeanVP					EWP	VWMP
T=60	T=120	T=240	T=all	T=60	T=120	T=240	T=all	in-sample		
0.4306	0.4202	0.4196	0.4177	0.4716	0.4485	0.4321	0.4369	0.7800	0.4323	0.3648

Table 5.5: Out-of-sample Sharpe ratios and p-values over the period 07/1963-11/2004 with shrinkage of the covariance matrix

Portfolio strategy		10 ind	30 ind	size	book	mom	short-term	long-term	div
EWP		0.4386	0.4316	0.4405	0.4913	0.3349	0.3846	0.4823	0.4547
MinVP	T=60	0.4285	0.4244	0.3504	0.4657	0.3851	0.3720	0.5192	0.4998
		(0.91)	(0.94)	(0.25)	(0.57)	(0.31)	(0.75)	(0.37)	(0.56)
	T=120	0.4025	0.4171	0.3659	0.4404	0.3750	0.3894	0.5164	0.4553
		(0.69)	(0.89)	(0.35)	(0.19)	(0.33)	(0.90)	(0.36)	(0.99)
	T=240	0.4138	0.4337	0.3747	0.4289	0.3540	0.3688	0.5042	0.4789
		(0.78)	(0.98)	(0.42)	(0.10)	(0.62)	(0.69)	(0.54)	(0.75)
	T=all data	0.4318	0.4816	0.3315	0.3815	0.4761	0.3773	0.4282	0.4338
		(0.94)	(0.61)	(0.22)	(0.00)*	(0.01)*	(0.83)	(0.19)	(0.42)
MeanVP	in-sample	0.5901	0.8620	0.6158	0.7079	1.2713	0.8132	0.6932	0.6865
	T=60	0.3824	0.4091	0.4448	0.5231	0.6440	0.4936	0.5570	0.3187
		(0.58)	(0.82)	(0.96)	(0.67)	(0.00)*	(0.14)	(0.33)	(0.12)
	T=120	0.3185	0.3804	0.3808	0.5845	0.6374	0.4306	0.4681	0.3874
		(0.21)	(0.62)	(0.47)	(0.12)	(0.00)*	(0.48)	(0.86)	(0.46)
	T=240	0.2898	0.2918	0.3711	0.4876	0.6339	0.4298	0.4547	0.4983
		(0.12)	(0.16)	(0.23)	(0.95)	(0.00)*	(0.41)	(0.64)	(0.54)
	T=all data	0.3682	0.3526	0.3310	0.5146	0.6573	0.3292	0.4371	0.5055
		(0.31)	(0.28)	(0.01)*	(0.64)	(0.00)*	(0.42)	(0.48)	(0.38)
VWMP		0.3648	0.3648	0.3648	0.3648	0.3648	0.3648	0.3648	0.3648
		(0.01)*	(0.12)	(0.17)	(0.00)*	(0.35)	(0.25)	(0.00)*	(0.04)*

Note: The p-value of the difference between the Sharpe ratio of each portfolio strategy from that of the EWP benchmark is reported in the parentheses. Those that are significant on a 5 percent significance level are marked with *.

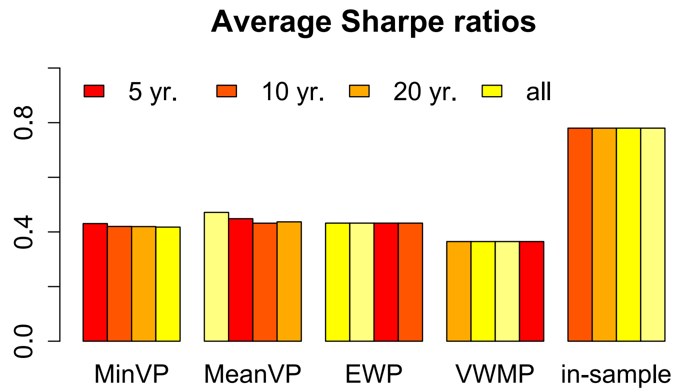


Figure 5.2: Graphical comparison of average annualized out-of-sample Sharpe ratios over the period 07/1963-11/2004

Table 5.6: Portfolio standard deviations, means and capital gains over the out-of-sample period 07/1963-11/2004

Portfolio strategy		Standard deviation	Excess mean return	Growth of $1
MinVP	T=60	13.97%	6.00%	88.94
	T=120	14.02%	5.89%	84.50
	T=240	14.08%	5.90%	84.25
	T=expanding	14.85%	6.19%	89.91
MeanVP	T=60	16.68%	7.89%	203.01
	T=120	17.42%	7.84%	209.11
	T=240	16.98%	7.33%	187.30
	T=expanding	18.21%	7.92%	269.70
EWP		15.80%	6.82%	109.87
VWMP		15.54%	5.67%	66.97

5.2.2 Other portfolio measures

The results from the other portfolios measures that are listed in Table 5.6. It is evident that these measures are very similar to those of the previous out-of-sample period. Not surprisingly the MinVP is the least volatile strategy, followed by the VWMP, EWP and finally the MeanVP. The means of the MinVP are higher than those of the VWMP but again, the EWP and the MeanVP delivers better mean return along with capital accumulation.

5.3 Results from implementation of different out-of-sample periods

In this section, I report the average out-of-sample Sharpe ratios listed in Table 5.7 and shown graphically in Figure 5.3. Additionally, the statistical distinguishability of the Sharpe ratios, as they are summarized in the Appendix, will be briefly commented. Due to the amount of

Sharpe ratio computations, these tables take up quite a bit of space and are therefore moved to the Appendix in order to improve the readability of this section.

5.3.1 Sharpe ratios

When observing the empirical data reported in Table 5.7 and Figure 5.3, there are certain characteristics that are evident. As previously, I will focus on the numerical value of the Sharpe ratios at first, and consider their statistical differentiability later on. The first thing that is noticeable is that the VWMP delivers the lowest Sharpe ratio of all strategies for every out-of-sample time period implemented. These findings suggest that in lieu of the other portfolio strategies, holding the market portfolio is not rational from a risk-adjusted return standpoint. Secondly, the optimized portfolios do seem to outperform the EWP in most cases except for when the out-of-sample periods start in 1981 and 1991, where the MeanVP performs worse than the EWP. From this result one could argue that, based on the numerical value of the Sharpe ratios alone, optimized portfolios do outperform simpler strategies. However, one should also focus on the statistical significance. A third observation is that of the two optimized portfolios, the MinVP consistently outperforms the MeanVP across out-of-sample periods and for the different lengths of the lookback period. The MeanVP that is tested, in-sample as opposed to out-of sample, reveals that if the mean-variance strategy was constructed without estimation error, the MeanVP would do better than all the other portfolio strategies. It also suggests that the estimation error present in the out-of-sample estimation cancels out all the gains from optimization, as the performance of the out-of-sample MeanVP is reduced to that of the EWP. The in-sample Sharpe ratios are nearly double that of the out-of-sample MeanVP.

The individual Sharpe ratios reported in the Appendix show that also during these periods the choice of dataset is predictive for the performance of some of the portfolio strategies. This is evident in the "mom" dataset as the MeanVP in most cases deliver statistically significantly higher Sharpe ratios than the EWP. However, considering the p-values in all the other datasets, there is no strong evidence suggesting that any of the optimized portfolios statistically outperforms the EWP.

5.3.2 Other portfolio measures

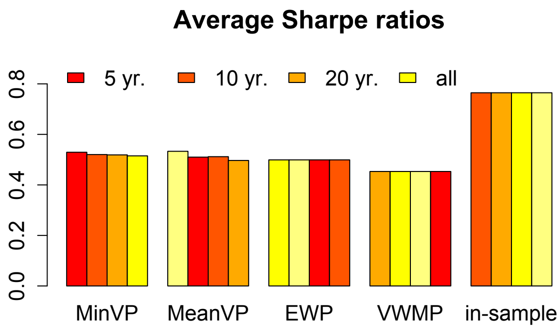
The portfolio standard deviations, means and capital gains from the out-of-sample period from 1951-2012 is reported in Table 5.8. Also this period delivers similar results as the ones before. The MinVP is the least volatile strategy, followed by the VWMP, EWP and finally the MeanVP. The means of the MinVP are higher than those of the VWMP while the EWP and the MeanVP delivers higher mean return and capital accumulation. Because the relationship between the different portfolio's excess means, standard deviations and capital gains are very similar for the remaining out-of-sample periods, I only report the findings from the 1951-2012 time period.

Table 5.7: Average annualized out-of-sample Sharpe ratios with different out-of-sample starting points

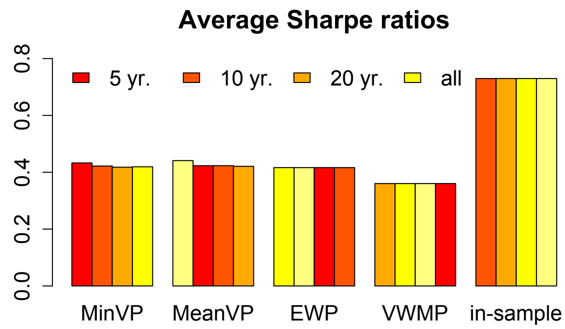
Portfolio strategy		Starting point of out-of-sample period				
		01/1951	01/1961	01/1971	01/1981	01/1991
MinVP	T=60	0.5292	0.4328	0.4503	0.5455	0.5487
	T=120	0.5201	0.4219	0.4421	0.5290	0.5352
	T=240	0.5188	0.4178	0.4494	0.5311	0.5537
	T=all data	0.5150	0.4193	0.4532	0.5239	0.5521
MeanVP	T=60	0.5330	0.4410	0.4478	0.4538	0.4884
	T=120	0.5099	0.4231	0.4170	0.4308	0.5012
	T=240	0.5117	0.4232	0.4271	0.4459	0.5182
	T=all data	0.4967	0.4209	0.4181	0.4423	0.5237
	in-sample	0.7647	0.7298	0.7236	0.8472	0.9046
EWP		0.4991	0.4164	0.4226	0.4736	0.5461
VWMP		0.4531	0.3602	0.3659	0.4170	0.4690

Table 5.8: Portfolio standard deviations, means and capital gains over the out-of-sample period 01/1951-12/2012

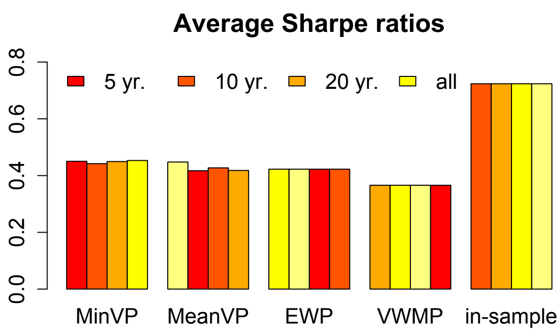
Portfolio strategy		Standard deviation	Excess mean return	Growth of $1
MinVP	T=60	13.51%	7.11%	765.67
	T=120	13.55%	7.01%	711.48
	T=240	13.71%	7.08%	722.16
	T=expanding	14.32%	7.36%	862.60
MeanVP	T=60	15.92%	8.50%	1786.82
	T=120	16.63%	8.49%	2118.77
	T=240	16.09%	8.23%	2221.13
	T=expanding	17.28%	8.55%	3791.59
EWP		15.43%	7.68%	916.71
VWMP		15.02%	6.81%	530.61



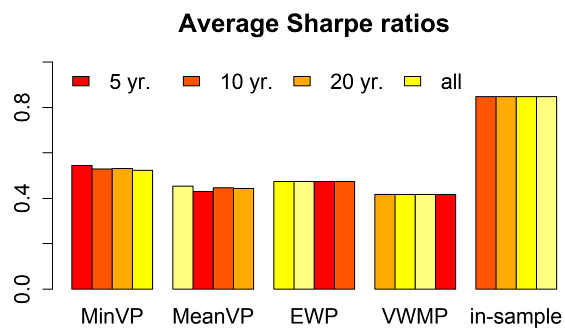
(a) 01/1951-12/2012



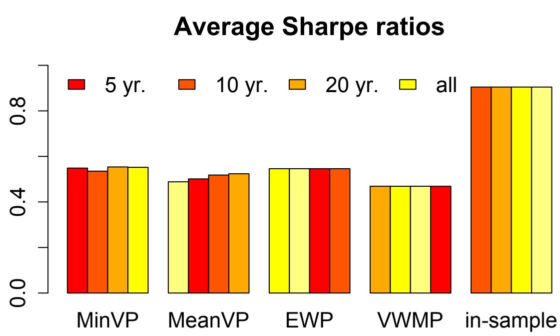
(b) 01/1961-12/2012



(c) 01/1971-12/2012



(d) 01/1981-12/2012



(e) 01/1991-12/2012

Figure 5.3: Graphical comparison of average annualized out-of-sample Sharpe ratios with different out-of-sample starting points

6 Discussion

6.1 Implications of the Kritzman et al. (2010) replication

Assessing the Sharpe ratios produced by the portfolio models in the paper by Kritzman et al. (2010), they find that the VWMP delivers the lowest, the EWP delivers the second lowest, while the optimized MinVP and MeanVP clearly delivers the highest. In their study the portfolios are compared graphically, in the same manner as I have shown in Section 5. However, they do not report the exact values of the Sharpe ratios, so the reader has to approximate them based on the value on the y-axis. Based on this approximation the VWMP has a Sharpe ratio lower than 0.4, the EWP just below 0.5, while the MinVP and MeanVP produces Sharpe ratios somewhere between 0.65 and 0.55. These results are quite different from those that I reproduced. As shown in Section 5, the Sharpe of the VWMP is 0.3757, which is very similar to that of Kritzman et al. (2010), but this is where the similarities end. The MeanVP produces Sharpe ratios that are only a little higher than the VWMP and is beaten by the EWP for every length of T. The MinVP outperforms the other strategies but does not produce Sharpe ratios above 0.5 which is lower than those reported in Kritzman et al. (2010). This suggests that one should be critical to the results from their study.

Although my Sharpe ratios are more modest, they still favor optimization to a degree in that the MinVP delivers the highest Sharpe ratios and outperform the non-optimized portfolios. This is a tempting conclusion, however the statistical significance of the difference between the Sharpe ratios should be considered. Thus, a more important result is that I find that the two optimized portfolio almost never produces statistically distinguishable Sharpe ratios when comparing to that of the EWP. Considering this, there is no evidence of optimization leading to better portfolio performance compared to an equally-weighted approach.

6.2 Implication of the DeMiguel et al. (2009) replication

When implementing the out-of-sample time period from July 1963 to November 2004 it becomes clear that the MinVP performs worse than for the rest of the implemented periods. This is surprising considering the better performance during the other time periods. Similarly to the results obtained by DeMiguel et al. (2009), the EWP is one of the best performing strategies. Considering the statistical significance of the difference between Sharpe ratios, I reach the same conclusions as DeMiguel et al. (2009) in that, except for a few cases, none of the optimized portfolios statistically outperforms the benchmark EWP. While DeMiguel et al. (2009) angles their paper such that it focuses on that optimized portfolios do not statistically outperform the EWP. However, this also holds the other way around, in that an equally-weighted strategy does not outperform optimized portfolios. Even though my data is somewhat different to that which is implemented in their study I still find similar results.

6.3 The general performance of the implemented portfolios

If concluding based on the numerical value of the Sharpe ratio alone, the MinVP comes across as the best strategy. This would be the strategy that delivers the highest risk-adjusted return of all the portfolio models considered in the study. This result is in line with the literature on the performance of low-volatility portfolios, such as Clarke et al. (2011). However, this outperformance does not hold statistically, as there are too few cases where the Sharpe ratio produced by the MinVP actually is statistically distinguishable from the benchmark. The MeanVP comes across as the second best strategy for all periods except during 1978-2008, 1981-2012 and 1991-2012 where it is beaten by the EWP, but also this portfolio fails to produce differences in Sharpe ratios that are statistically significant in general. The EWP statistically outperforms the VWMP in several of the cases. This suggests that holding the market portfolio provides the investor with the lowest risk-adjusted return, and that a investor who wants to maximize the Sharpe ratio should rather choose one of the other portfolios. This is perhaps one of the most apparent results of the out-of-sample comparison.

Because the MinVP in general delivers higher empirical Sharpe ratios than the MeanVP, this suggests that including estimates of mean returns in the optimization process does not add any advantage as opposed to just estimation the covariance matrix. In a situation where no plausible estimates for mean returns are obtainable, it might be best to stick with the optimized minimum-variance strategy. As pointed out by DeMiguel et al. (2009), the difference between the in-sample MeanVP Sharpe ratio and the one produced by the EWP shows the difference between a truly optimal allocation and a naive allocation. It also hints at the severity of the estimation error, due to that fact that the out-of-sample MeanVP is reduced from the optimal in-sample Sharpe ratio to one that equals the EWP which does not require any optimization or stock-picking skill. This suggests that the effect of estimation error is severe enough to impact the out-of-sample performance to a degree where all gains from optimization is lost. The possible gain from optimization, as expressed by difference between the in-sample MeanVP Sharpe ratio and the EWP Sharpe ratio, is completely eroded due to estimation error. This is the same conclusion that DeMiguel et al. (2009) reached.

Another thing to note regarding the in-sample MeanVP is that the Sharpe ratio delivered by this strategy is the highest of all the models. This result helps to justify optimization, because this proves that, at least when the estimates are known, optimization leads to better results. This means that the naive-diversification strategy is not a recommended strategy if somehow one is certain of the true means, variances and covariances. This result is in line with DeMiguel et al. (2009) and Kritzman et al. (2010).

An interesting result is that the MeanVP often performs fairly well out-of-sample, even though some of the literature suggests that due to estimation error it is expected to do poorly. Because constraints on the optimization procedure are present, this may be the reason for why these problems are not evident. This could imply that the portfolio does so well that the negative effect of the estimation error is not severe enough to halt the performance sufficiently.

The comparison between the MinVP and the MeanVP has shown that the less risky MinVP often delivers equal or better risk-adjusted returns as the MeanVP. However, the study shows that capital accumulation is often far greater when implementing the Markowitz strategy. Investors demanding a certain monetary amount at the end of the investment period might be put off by the MinVP. In financial theory one would circumvent this by implementing a levered strategy that would lead to the same amount of capital growth. However, real world investors often face borrowing restrictions and cannot implement such a strategy. Such investors would therefore choose the MeanVP over the MinVP in a decision between the two.

6.4 Implications of choice of dataset and out-of-sample time period

When studying the results it becomes clear both of the optimized strategies often deliver a statistically higher Sharpe ratio than the EWP in some of the cases when considering the dataset formed on momentum deciles. This holds for most of the periods tested and the MeanVP stands out in terms of statistically significance. For example, implementing a MeanVP strategy on the “mom” dataset over the time period 1951 to 2012 will return a portfolio with the highest risk-adjusted return, which is also statistically distinguishable for all T. This gives the impression that an optimized mean-variance strategy strongly outperforms an equally-weighted strategy, but when considering the other datasets as well, it turns out that the MeanVP does not perform nearly as good as in the “mom” data. This demonstrates the importance of analyzing the performance over several datasets, as a strategy can be found to perform well under certain choices.

Analyzing the time period on which the portfolio strategies are tested shows that the Sharpe ratios of the portfolio strategies are higher for some of the time periods while lower for others. However, this holds for most of the portfolios collectively so that when the MinVP delivers a high Sharpe ratio, so do the other portfolios as well. This implies that there is none of the time periods implemented where a single strategy suddenly outperform the other as they are always within a certain relation from each other. The only period where this found to a degree is the out-of-sample period from 1963 to 2004, where the MinVP is not the best performing portfolio in terms of numerical value of the Sharpe ratios. This suggests that the choice of out-of-sample period does not make a significant difference to the empirical performance of the portfolios. Furthermore, there is no single out-of-sample period where a portfolio strategy suddenly and consistently delivers a statistically significantly higher Sharpe ratio when compared to the benchmark.

6.5 Suggested explanations for MinVP performance

Considering the common assumption that taking on more risk should lead to higher expected return, it is surprising that the MinVP delivers higher risk-adjusted return than the other more risky strategies. In terms of the MinVP and MeanVP a possible explanation could be that misspecification of the mean return estimates leads to estimation error severe enough to deteriorate all possible gains from optimization. This is evident when comparing the out-of-sample Sharpe ratios of the MeanVP to the in-sample Sharpe ratios, which is also in line with the findings of DeMiguel et al. (2009). Because the covariance estimates are found to be more accurate (Best and Grauer, 1991; Chopra and Ziemba, 1993) and in addition are shrunk as proposed in Ledoit and Wolf (2004), the out-of-sample MinVP might produce less estimation error so that there is still some gains from optimization remaining.

The literature suggests additional explanations for the good performance of the MinVP. Several studies (Asness et al., 2012; Blitz and van Vliet, 2007; Frazzini and Pedersen, 2013) argue that due to leverage constraints and leverage aversion investors will overweight high-beta securities in their portfolios. This causes the riskier assets to be overpriced and thus offer lower risk-adjusted returns, while less risky assets are underpriced and thus offer higher risk-adjusted return. Indeed, inspecting the capital accumulations of the portfolios reveals that the MinVP generally delivers low capital growth when compared to the MeanVP and the EWP. It becomes evident that investors seeking to maximize capital gain might refrain from investing in the MinVP even though it offers the higher risk-adjusted return.

Papers such as Frazzini and Pedersen (2014) and Scherer (2011) connect the MinVP performance to risk-factor exposure by performing regression analysis and finding that the MinVP loads up on certain risk factors, while Baker and Haugen (2012) argue from a behavioral standpoint suggesting that agency problems make low-volatility stocks unattractive for portfolio managers.

7 Conclusion

Reviewing the academic literature on the empirical performance of various asset allocation strategies, reveals that there is little consensus on whether or not optimization leads to outperformance. Can optimized portfolios deliver better performance than strategies such as equal weighting of assets or investing in the VWMP? To facilitate the opposing views of this area of research I concentrated on two research papers with conflicting conclusions. DeMiguel et al. (2009) suggest that optimized portfolios do not consistently outperform an EWP, and that the out-of-sample Sharpe ratios of the optimized portfolios are not statistically distinguishable from those of the EWP. However, Kritzman et al. (2010) claimed that optimized portfolios definitely perform better. They based this on out-of-sample results where the minimum-variance and mean-variance portfolios deliver considerably higher Sharpe ratios, compared to the EWP and VWMP.

The aim of this thesis has been to compare the out-of-sample performance of optimized portfolios and naively diversified portfolios. This was done by studying the empirical performance of the MinVP, MeanVP, EWP, and VWMP in order to determine if any of the portfolio strategies deliver consistently better performance. By utilizing the same eight datasets and out-of-sample time period implemented in Kritzman et al. (2010), I attempted to replicate their results for further comparison and analysis. In addition to this, I also tested the data on the same out-of-sample time period that was implemented in DeMiguel et al. (2009) to see if I reached similar conclusions. To determine if the choice of out-of-sample time period had any influence on the performance of the portfolios, I performed backtests on five additional out-of-sample time periods.

Assuming a mean-variance optimizing investor, the out-of-sample Sharpe ratio of the portfolios was used to measure the performance. While a strategy might outperform another in terms of the numerical value of the out-of-sample Sharpe ratio, the properties of the data might reveal that the ratios are actually statistically equal to each other. To distinguish the Sharpe ratios from each other, a test examining if the difference between them is statistically significant was applied. Furthermore, the annual mean and standard deviation of a portfolio strategy, as well as the capital accumulation of a $1 initial investment into each of the strategies was examined.

By replicating the results obtained by Kritzman et al. (2010) and following their approach of implementing a constant mean estimate based on the historical data and a covariance matrix estimated on 5, 10 and 20-year rolling windows, as well as an expanding approach, I find dissimilar results in that the out-of-sample Sharpe ratios obtained in this thesis are considerably lower. In addition, I find no consistency in the statistical tests of the difference between the Sharpe ratios of the MinVP, MeanVP, and the EWP. From this I conclude that the out-of-sample performance of the optimized portfolios and the EWP are not statistically distinguishable from each other. When performing the rest of the empirical backtesting I then estimated the means in the same manner as the covariance matrix, in addition to shrinking the covariance matrix estimate towards the constant-correlation matrix. For the remaining periods as well, I find that the difference between the out-of-sample Sharpe ratios of the MinVP, MeanVP, and EWP were not consistently statistically significant. This suggests that the optimized portfolios do not outperform a naively-weighted benchmark, but also that an EWP do not outperform optimized portfolios. This is important considering that some researchers tend to look at the performance of the strategies from a single perspective of either good or bad. The results also suggest that the choice of out-of-sample time period does not affect the performance of the portfolios when compared to each other.

I also find that a MeanVP that is constructed using in-sample estimates obviously outperforms the other strategies. This suggests that in theory optimization does indeed deliver superior Sharpe ratios. However, the apparent estimation error from misspecifications of the inputs is so large that most of the performance is lost. Because the MinVP performs so well, it could be that misspecification of the mean estimates will affect the out-of-sample portfolio performance.

In addition, I find that the VWMP underperforms compared to the other portfolio strategies during all the implemented out-of-sample time periods. The lower Sharpe ratios produced by the VWMP is often found to be statistically distinguishable as well, suggesting that investing in the market portfolio is the least attractive of the four strategies.

When only considering the numerical value of the Sharpe ratios, the optimized portfolios generally deliver higher Sharpe ratios. This suggests that they should be expected to deliver higher Sharpe ratios on average. From a practitioner's point of view, this might lead to optimized portfolios being preferable. The MinVP is found to be the strategy that performs the best out-of-sample. This is particularly interesting considering that it is the least volatile of all the portfolios. Considering the findings of papers examining the low-volatility anomaly, this result might come as no surprise.

If the findings of this thesis should imply anything, it would have to be that more research must be devoted to increasing the accuracy of the parameters used when constructing optimized portfolios. This would further diminish the effects of estimation error.

References

Asness, C., Frazzini, A., Pedersen, L. H., 2012. Leverage aversion and risk parity. *Financial Analysts Journal* 68, pp. 47–59.

Baker, M., Bradley, B., Wurgler, J., 2011. Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal* 67 (1), pp. 40–54.

Baker, N. L., Haugen, R. A., 2012. Low risk stocks outperform within all observable markets of the world (april 27, 2012), available at SSRN: http://ssrn.com/abstract=2055431.

Best, M. J., 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.

Black, F., Jensen, M., Scholes, M., 1972. The capital asset pricing model: Some empirical tests. In: Studies in the Theory of Capital Markets. Praeger Publishers Inc.

Blitz, D., Vliet, P. v., 2007. The volatility effect: Lower risk without lower return. Erim report series research in management, Erasmus Research Institute of Management (ERIM).

Bodie, Z., Kane, A., Marcus, A. J., 2011. Investments and Portfolio Management: Global Edition. McGraw-Hill Education.

Brennan, M. J., Cheng, X., Li, F., 2012. Agency and institutional investment. *European Financial Management* 18 (1), pp. 1–27.

Chopra, V. K., Ziemba, W. T., 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management* 19 (2), pp. 6–11.

Chow, T.-m., Hsu, J. C., Kuo, L.-L., Li, F., 2013. A study of low volatility portfolio construction methods (december 20, 2013), available at SSRN: http://ssrn.com/abstract=2298117 or http://dx.doi.org/10.2139/ssrn.2298117.

Clarke, R. G., de Silva, H., Thorley, S., 2011. Minimum-variance portfolio composition. *The Journal of Portfolio Management* 37 (2), pp. 31–45.

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

Duchin, R., Levy, H., 2009. Markowitz versus the talmudic portfolio diversification strategies. *The Journal of Portfolio Management* 35 (2), pp. 71–74.

Frazzini, A., Pedersen, L. H., 2014. Betting against beta. *Journal of Financial Economics* 111 (1), pp. 1–25.

Haugen, R. A., Baker, N. L., 1991. The efficient market inefficiency of capitalization-weighted stock portfolios. *The Journal of Portfolio Management* 17, pp. 35–40.

Haugen, R. A., Heins, A. J., 1975. Risk and the rate of return on financial assets: Some old wine in new bottles. *The Journal of Financial and Quantitative Analysis* 10 (5), pp. 775–784.

Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance* 58 (4), pp. 1651–1684.

Jobson, J. D., Korkie, B. M., 1981. Performance hypothesis testing with the sharpe and treynor measures. *The Journal of Finance* 36 (4), pp. 889–908.

Kritzman, M., 2006. Are optimizers error maximizers? *The Journal of Portfolio Management* 32 (4), pp. 66–69.

Kritzman, M., Page, S., Turkington, D., 2010. In defense of optimization: The fallacy of 1/ n. *Financial Analysts Journal* 66 (2), pp. 31–39.

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

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics* 47 (1), pp. 13–37.

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

Memmel, C., 2003. Performance hypothesis testing with the sharpe ratio. *Finance Letters* 1, pp. 21–23.

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

Mossin, J., 1966. Equilibrium in a capital asset market. *Econometrica* 34 (4), pp. 768–783.

original by Berwin A. Turlach R port by Andreas Weingessel <Andreas.Weingessel@ci.tuwien.ac.at>, S., 2013. quadprog: Functions to solve Quadratic Programming Problems. R package version 1.5-5.
URL `http://CRAN.R-project.org/package=quadprog`

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL `http://www.R-project.org/`

Scherer, B., 2011. A note on the returns from minimum variance investing. *Journal of Empirical Finance* 18 (4), pp. 652 – 660.

Sharpe, W. F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19 (3), pp. 425–442.

Sharpe, W. F., 1966. Mutual fund performance. *Journal of Business* 39, pp. 119–138.

Sharpe, W. F., 1994. The sharpe ratio. *The Journal of Portfolio Management* 21 (1), pp. 49–58.

Statistics, B., 2012. BurStFin: Burns Statistics Financial. R package version 1.01. URL `http://CRAN.R-project.org/package=BurStFin`

A Presentation of omitted tables

This part of the appendix presents the tables reporting the out-of-sample Sharpe ratio of every portfolio strategy for every dataset, as well as the p-values stating the statistical significance of these measures, from the five different out-of-sample time periods starting in 1951, 1961, 1971, 1981 and 1991 respectively.

Table A.1: Out-of-sample Sharpe ratios and p-values over the period 01/1951-12/2012 with shrinkage of the covariance matrix

Portfolio strategy		10 ind	30 ind	size	book	mom	short-term	long-term	div
EWP		0.5392	0.5032	0.4941	0.5437	0.4155	0.4515	0.5231	0.5222
MinVP	T=60	0.5788	0.5648	0.4584	0.5698	0.4642	0.4665	0.5370	0.5944
		(0.59)	(0.45)	(0.55)	(0.49)	(0.22)	(0.63)	(0.68)	(0.20)
	T=120	0.5642	0.5490	0.4677	0.5469	0.4543	0.4818	0.5522	0.5449
		(0.73)	(0.58)	(0.66)	(0.92)	(0.27)	(0.32)	(0.36)	(0.69)
	T=240	0.5675	0.5605	0.4622	0.5154	0.4761	0.4711	0.5535	0.5438
		(0.70)	(0.48)	(0.61)	(0.37)	(0.07)	(0.52)	(0.31)	(0.69)
	T=all data	0.5523	0.5842	0.4308	0.4611	0.5854	0.4908	0.5069	0.5083
		(0.84)	(0.30)	(0.36)	(0.01)*	(0.00)*	(0.14)	(0.63)	(0.57)
MeanVP	in-sample	0.7350	0.8835	0.5730	0.6552	1.0593	0.8263	0.6689	0.7159
	T=60	0.5093	0.5042	0.4871	0.5738	0.6516	0.5575	0.5227	0.4581
		(0.71)	(0.99)	(0.92)	(0.60)	(0.00)*	(0.06)	(0.99)	(0.33)
	T=120	0.4470	0.4662	0.4272	0.6215	0.6669	0.5102	0.4755	0.4650
		(0.22)	(0.66)	(0.26)	(0.11)	(0.00)*	(0.24)	(0.42)	(0.39)
	T=240	0.4484	0.4243	0.4342	0.5529	0.6805399	0.5005	0.4697	0.5829
		(0.22)	(0.31)	(0.16)	(0.83)	(0.00)*	(0.27)	(0.23)	(0.25)
	T=all data	0.4595	0.5082	0.3968	0.5388	0.6865	0.3861	0.4447	0.5531
		(0.13)	(0.94)	(0.00)*	(0.90)	(0.00)*	(0.23)	(0.12)	(0.47)
VWMP		0.4531	0.4531	0.4531	0.4531	0.4531	0.4531	0.4531	0.4531
		(0.00)*	(0.14)	(0.34)	(0.00)*	(0.14)	(0.91)	(0.00)*	(0.03)*

Note: The p-value of the difference between the Sharpe ratio of each portfolio strategy from that of the EWP benchmark is reported in the parentheses. Those that are significant on a 5 percent significance level are marked with *.

Table A.2: Out-of-sample Sharpe ratios and p-values over the period 01/1961-12/2012 with shrinkage of the covariance matrix

Portfolio strategy		10 ind	30 ind	size	book	mom	short-term	long-term	div
EWP		0.4396	0.4178	0.4255	0.4651	0.3337	0.3645	0.4565	0.4284
MinVP	T=60	0.4660	0.4502	0.3530	0.4641	0.3957	0.3624	0.4812	0.4897
		(0.73)	(0.71)	(0.27)	(0.98)	(0.16)	(0.95)	(0.50)	(0.33)
	T=120	0.4515	0.4320	0.3638	0.4338	0.3823	0.3802	0.4976	0.4339
		(0.88)	(0.87)	(0.35)	(0.38)	(0.20)	(0.64)	(0.22)	(0.93)
	T=240	0.4538	0.4472	0.3662	0.4073	0.3659	0.3718	0.4892	0.4415
		(0.85)	(0.73)	(0.38)	(0.08)	(0.37)	(0.82)	(0.30)	(0.83)
	T=all data	0.4493	0.4942	0.3300	0.3831	0.4816	0.3903	0.4052	0.4207
		(0.89)	(0.36)	(0.20)	(0.02)*	(0.00)*	(0.37)	(0.15)	(0.75)
MeanVP	in-sample	0.6322	0.8651	0.5644	0.6440	1.0194	0.7852	0.6502	0.6778
	T=60	0.3897	0.4036	0.4210	0.4626	0.5751	0.4677	0.4548	0.3538
		(0.57)	(0.87)	(0.95)	(0.97)	(0.00)*	(0.10)	(0.98)	(0.31)
	T=120	0.3366	0.3822	0.3637	0.5112	0.6010	0.4102	0.4231	0.3572
		(0.22)	(0.70)	(0.35)	(0.37)	(0.00)*	(0.41)	(0.61)	(0.34)
	T=240	0.3395	0.3119	0.3670	0.4421	0.5993	0.4119	0.4227	0.4910
		(0.23)	(0.22)	(0.22)	(0.62)	(0.00)*	(0.32)	(0.49)	(0.28)
	T=all data	0.3699	0.3929	0.3334	0.4766	0.6008	0.2967	0.4256	0.4715
		(0.22)	(0.71)	(0.01)*	(0.78)	(0.00)*	(0.25)	(0.58)	(0.37)
VWMP		0.3602	0.3602	0.3602	0.3602	0.3602	0.3602	0.3602	0.3602
		(0.00)*	(0.11)	(0.16)	(0.00)*	(0.36)	(0.79)	(0.00)*	(0.06)

Note: The p-value of the difference between the Sharpe ratio of each portfolio strategy from that of the EWP benchmark is reported in the parentheses. Those that are significant on a 5 percent significance level are marked with *.

Table A.3: Out-of-sample Sharpe ratios and p-values over the period 01/1971-12/2012 with shrinkage of the covariance matrix

Portfolio strategy		10 ind	30 ind	size	book	mom	short-term	long-term	div
EWP		0.4453	0.4220	0.4174	0.4695	0.3373	0.3717	0.4772	0.4406
MinVP	T=60	0.4983	0.4626	0.3670	0.4908	0.4134	0.3737	0.4929	0.5041
		(0.55)	(0.68)	(0.49)	(0.64)	(0.13)	(0.96)	(0.71)	(0.40)
	T=120	0.5019	0.4689	0.3713	0.4467	0.4107	0.3984	0.4974	0.4416
		(0.53)	(0.64)	(0.54)	(0.58)	(0.10)	(0.49)	(0.60)	(0.99)
	T=240	0.5282	0.5101	0.3754	0.4372	0.4006	0.3886	0.4891	0.4663
		(0.33)	(0.37)	(0.58)	(0.39)	(0.12)	(0.66)	(0.74)	(0.72)
	T=all data	0.5234	0.5478	0.3406	0.4089	0.4896	0.4047	0.4568	0.4537
		(0.31)	(0.18)	(0.36)	(0.11)	(0.01)*	(0.31)	(0.61)	(0.60)
MeanVP	in-sample	0.6804	0.9057	0.5265	0.6093	1.0438	0.7372	0.6029	0.6832
	T=60	0.4433	0.4203	0.4606	0.4624	0.5303	0.4687	0.4464	0.3508
		(0.99)	(0.99)	(0.62)	(0.92)	(0.04)*	(0.74)	(0.68)	(0.28)
	T=120	0.3476	0.3615	0.3844	0.5214	0.5566	0.3969	0.4377	0.3300
		(0.33)	(0.57)	(0.67)	(0.37)	(0.01)*	(0.69)	(0.60)	(0.21)
	T=240	0.3579	0.3160	0.3906	0.4710	0.5666	0.3904	0.4179	0.5065
		(0.37)	(0.30)	(0.63)	(0.98)	(0.01)*	(0.73)	(0.28)	(0.33)
	T=all data	0.3845	0.3927	0.3263	0.4750	0.5659	0.2850	0.4042	0.5116
		(0.35)	(0.70)	(0.02)*	(0.90)	(0.02)*	(0.20)	(0.24)	(0.19)
VWMP		0.3659	0.3659	0.3659	0.3659	0.3659	0.3659	0.3659	0.3659
		(0.01)*	(0.18)	(0.32)	(0.01)*	(0.40)	(0.74)	(0.00)*	(0.08)

Note: The p-value of the difference between the Sharpe ratio of each portfolio strategy from that of the EWP benchmark is reported in the parentheses. Those that are significant on a 5 percent significance level are marked with *.

Table A.4: Out-of-sample Sharpe ratios and p-values over the period 01/1981-12/2012 with shrinkage of the covariance matrix

Portfolio strategy		10 ind	30 ind	size	book	mom	short-term	long-term	div
EWP		0.5328	0.4869	0.4420	0.5097	0.3882	0.4046	0.5301	0.4946
MinVP	T=60	0.6395	0.5916	0.4464	0.5548	0.4654	0.4538	0.5724	0.6404
		(0.30)	(0.37)	(0.96)	(0.42)	(0.18)	(0.25)	(0.39)	(0.10)
	T=120	0.6359	0.5840	0.4509	0.4985	0.4696	0.4791	0.5667	0.5473
		(0.34)	(0.42)	(0.91)	(0.82)	(0.13)	(0.09)	(0.42)	(0.56)
	T=240	0.6477	0.6211	0.4568	0.4731	0.4614	0.4692	0.5634	0.5558
		(0.27)	(0.26)	(0.86)	(0.41)	(0.15)	(0.15)	(0.43)	(0.50)
	T=all data	0.6372	0.6504	0.4104	0.4642	0.5285	0.4902	0.5114	0.4988
		(0.23)	(0.13)	(0.74)	(0.27)	(0.02)*	(0.03)*	(0.69)	(0.88)
MeanVP	in-sample	0.8507	1.0972	0.6143	0.6927	1.1481	0.9268	0.6854	0.7626
	T=60	0.4479	0.5417	0.4622	0.4905	0.4169	0.4046	0.4602	0.4065
		(0.39)	(0.64)	(0.79)	(0.79)	(0.75)	(1.00)	(0.27)	(0.31)
	T=120	0.3504	0.3500	0.3917	0.5552	0.4916	0.3976	0.5030	0.4073
		(0.06)	(0.23)	(0.50)	(0.48)	(0.26)	(0.90)	(0.66)	(0.29)
	T=240	0.3886	0.3588	0.3560	0.5232	0.5065	0.3846	0.5031	0.5468
		(0.13)	(0.28)	(0.20)	(0.82)	(0.20)	(0.73)	(0.65)	(0.49)
	T=all data	0.4914	0.4995	0.2835	0.5240	0.5113	0.1948	0.4745	0.5594
		(0.56)	(0.89)	(0.00)*	(0.78)	(0.31)	(0.01)*	(0.38)	(0.28)
VWMP		0.4170	0.4170	0.4170	0.4170	0.4170	0.4170	0.4170	0.4170
		(0.00)*	(0.18)	(0.68)	(0.04)*	(0.49)	(0.53)	(0.00)*	(0.17)

Note: The p-value of the difference between the Sharpe ratio of each portfolio strategy from that of the EWP benchmark is reported in the parentheses. Those that are significant on a 5 percent significance level are marked with *.

Table A.5: Out-of-sample Sharpe ratios and p-values over the period 01/1991-12/2012 with shrinkage of the covariance matrix

Portfolio strategy		10 ind	30 ind	size	book	mom	short-term	long-term	div
EWP		0.5780	0.5427	0.5458	0.5722	0.4598	0.4592	0.6614	0.5495
MinVP	T=60	0.5498	0.4617	0.4917	0.6231	0.5422	0.5344	0.5993	0.5873
		(0.81)	(0.55)	(0.58)	(0.46)	(0.28)	(0.16)	(0.32)	(0.67)
	T=120	0.5529	0.4519	0.4897	0.5555	0.5553	0.5285	0.6659	0.4821
		(0.84)	(0.52)	(0.58)	(0.79)	(0.19)	(0.22)	(0.94)	(0.50)
	T=240	0.5958	0.5222	0.4954	0.5623	0.5458	0.5152	0.6806	0.5125
		(0.89)	(0.89)	(0.64)	(0.87)	(0.21)	(0.32)	(0.74)	(0.73)
	T=all data	0.5345	0.5200	0.4264	0.5634	0.6298	0.5466	0.6240	0.5724
		(0.66)	(0.86)	(0.34)	(0.87)	(0.04)*	(0.09)	(0.56)	(0.52)
MeanVP	in-sample	0.8166	1.2102	0.7363	0.7059	1.1284	0.9931	0.8280	0.8186
	T=60	0.4954	0.4870	0.4935	0.4675	0.5297	0.4985	0.4996	0.4360
		(0.52)	(0.71)	(0.61)	(0.23)	(0.54)	(0.61)	(0.05)*	(0.25)
	T=120	0.4432	0.4412	0.4818	0.5354	0.5959	0.4947	0.6224	0.3950
		(0.24)	(0.49)	(0.52)	(0.64)	(0.25)	(0.55)	(0.59)	(0.10)
	T=240	0.4930	0.3310	0.4887	0.5258	0.6273	0.5083	0.6105	0.5607
		(0.41)	(0.10)	(0.52)	(0.50)	(0.15)	(0.39)	(0.38)	(0.90)
	T=all data	0.5774	0.5376	0.4522	0.5441	0.6061	0.2552	0.6469	0.5698
		(1.00)	(0.97)	(0.01)*	(0.62)	(0.38)	(0.05)*	(0.83)	(0.78)
VWMP		0.4690	0.4690	0.4690	0.4690	0.4690	0.4690	0.4690	0.4690
		(0.04)*	(0.32)	(0.34)	(0.12)	(0.88)	(0.72)	(0.00)*	(0.34)

Note: The p-value of the difference between the Sharpe ratio of each portfolio strategy from that of the EWP benchmark is reported in the parentheses. Those that are significant on a 5 percent significance level are marked with *.

B Implemented R code for achieving the empirical results

This part of the appendix will state the R code that has been used in order to reach the empirical results. The purpose of including this code is to ease the replicability of my results. The package “BurStFin” by Burns Statistics (2012) was applied for computing the Ledoit and Wolf (2004) shrinkage estimator, while the package “quadprog” by Turlach and Weingessel (2013) was used for the quadratic programming problems.

B.1 Program for the rolling-window approach

The following R code shows how the empirical results were obtained when implementing a rolling-window estimation approach. Note that some of the code is changed depending on the relevant datasets, time period, mean estimation approach and whether or not the shrinkage estimator has been applied.

```
rm(list=ls(all=TRUE)) # clear environment

```

```r
# load prerequisites
source("functions_thesis.R")
library("quadprog")
library("BurStFin")

# read data set implemented as the investment universe
ret <- read.table("10_industries.txt", header=TRUE)

# read data on factor returns
fac <- read.table("FF3_factors.txt", header=TRUE)

# construct the time series object
ret.ts <- ts(ret[,2:ncol(ret)]/100, start=c(1927,1), frequency = 12)
fac.ts <- ts(fac[,2:ncol(fac)]/100, start=c(1927,1), frequency = 12)

# define the start and end dates and select the relevant time window
nYears <- 5                         # years set aside for preliminary estimation
year.start <- 1951
start <- c(year.start-nYears,1)     # OOS starting year and month
end <- c(2012,12)                   # OOS ending year and month
ret.ts <- window(ret.ts, start=start, end=end)
fac.ts <- window(fac.ts, start=start, end=end)

# return to data frame
ret <- as.data.frame(ret.ts)
fac <- as.data.frame(fac.ts)

# define lookback period and index start
nobs <- nrow(fac)
lookback.period <- nYears*12
first <- lookback.period + 1

# constructing the market portfolio (VWMP)
fac.period <- fac[first:nobs,]
rf <- fac.period[,"RF"]
```

```r
ret.mkt <- fac.period[,"Mkt.RF"] + rf # return on VWMP

# constructing the equally-weighted portfolio (EWP)
N <- ncol(ret)
w.naive <- rep(1/N, N) # defining the weights
naive.period <- ret[first:nobs,]
ret.naive <- rowSums(t(w.naive)*naive.period) #return on EWP

# constructing the minimum variance (MinVP) and the mean variance portfolio (MeanVP)
n <- length(rf)
ret.minvar <- rep(0,n)
ret.markowitz <- rep(0,n)

for (i in 1:n) {
  # find the indices of the lookback period
  period.end <- first + i - 2
  period.start <- period.end - lookback.period + 1
  if (period.start < 1) period.start <- 1 # index cannot be less than 1

  # estimate the covariance matrix
  covmat <- var.shrink.eqcor(ret[period.start:period.end,]) # using shrinkage estimator
  #covmat <- cov(ret[period.start:period.end,])              # without shrinkage estimator

  # compute the minimum-variance portfolio
  w.minvar <- minvarport(covmat, shorts=FALSE)          # MinVP weights, no shorts
  ret.minvar[i] <- sum(w.minvar*ret[(period.end+1),]) # return on the MinVP

  # compute the mean-variance portfolio
  er <- apply(ret[period.start:period.end,], 2, mean)          # mean return in the period
  w.markowitz <- tanmaxutility(er, covmat, rf[i])              # mean variance weights
  ret.markowitz[i] <- sum(w.markowitz*ret[(period.end+1),]) #
```

```
return on the MeanVP
}

# portfolio capital accumulation
ind.mkt <- cumprod(c(1, 1+ret.mkt))                 # VWMP
ind.minvar <- cumprod(c(1, 1+ret.minvar))           # minimum variance portfolio
ind.naive <- cumprod(c(1, 1+ret.naive))             # 1/N portfolio
ind.markowitz <- cumprod(c(1,1+ret.markowitz))  # mean variance portfolio

# accumulated wealth at the end of the period
finalwealth.minvar <- round(ind.minvar[n+1], digits=2)        # minimum variance portfolio
finalwealth.market <- round(ind.mkt[n+1], digits=2)           # VWMP
finalwealth.naive <- round(ind.naive[n+1], digits=2)          # 1/N portfolio
finalwealth.markowitz <- round(ind.markowitz[n+1], digits=2) # mean variance portfolio

# plot the accumulated wealth
ind <- ts(cbind(ind.mkt, ind.minvar, ind.naive, ind.markowitz), start=c(year.start,1), frequency = 12)
plot( log(ind), plot.type = "single", col = c("red", "blue", "green", "pink"),
      ylab="Log growth of $1", xlab="Year" )
legend(x="topleft", legend=c("Market", "Minimum variance", "Naive", "Markowitz"),
       col = c("red", "blue", "green", "pink"), lty=1)

# computing excess return
mkt <- ret.mkt - rf                  # VWMP
minvar <- ret.minvar - rf            # MinVP
naive <- ret.naive - rf              # EWP
markowitz <- ret.markowitz - rf  # MeanVP

# annualized out-of-sample Sharpe ratio
Sharpe.mkt <- Sharpe(mkt)                  # VWMP
Sharpe.minvar <- Sharpe(minvar)            # MinVP
```

```
Sharpe.naive <- Sharpe(naive)           # EWP
Sharpe.markowitz <- Sharpe(markowitz) # MeanVP

# computing max sharpe when treating out-of-sample period as in-sample
insample.ret <- ret[first:nobs,]                  # returns in in-sample period
insample.er <- apply(insample.ret, 2, mean)   # mean returns in in-sample period
insample.covmat <- cov(insample.ret)           # in-sample covariance matrix

# in-sample Sharpe from mean variance strategy
insample.Sharpe.markowitz <- maxsharpe(insample.er, insample.covmat, mean(rf))

# various annualized measures
an.mean(minvar)     # mean excess return on MinVP
an.std(minvar)      # sample volatility of MinVP
an.mean(mkt)        # mean excess return on VWMP
an.std(mkt)         # sample volatility of VWMP
an.mean(naive)      # mean excess return on EWP
an.std(naive)       # sample volatility of EWP
an.mean(markowitz) # mean excess return on MeanVP
an.std(markowitz)  # sample volatility of MeanVP

# testing statistical significance of Sharpe ratios across portfolios
SharpeTest(minvar, naive)
SharpeTest(markowitz, naive)
SharpeTest(mkt, naive)
```

Listing 1: R code for the rolling-window approach

B.2 Program for the expanding-window approach

The following R code shows how the empirical results were obtained when implementing a expanding-window estimation approach. Note that some of the code is changed depending on the relevant datasets, time period, mean estimation approach and whether or not the shrinkage estimator has been applied.

```r
rm(list=ls(all=TRUE)) # clear environment

# load prerequisites
source("functions_thesis.R")
library("quadprog")
library("BurStFin")

# read data set implemented as the investment universe
ret <- read.table("10_industries.txt", header=TRUE)

# read data on factor returns
fac <- read.table("FF3_factors.txt", header=TRUE)

# construct the time series object
ret.ts <- ts(ret[,2:ncol(ret)]/100, start=c(1927,1), frequency = 12)
fac.ts <- ts(fac[,2:ncol(fac)]/100, start=c(1927,1), frequency = 12)

# define the start and end dates and select the time window
year.start <- 1951
nYears <- year.start-1927           # years set aside for preliminary estimation
start <- c(year.start-nYears,1)     # OOS starting year and month
end <- c(2012,12)                   # OOS ending year and month
ret.ts <- window(ret.ts, start=start, end=end)
fac.ts <- window(fac.ts, start=start, end=end)

# return to data frame
ret <- as.data.frame(ret.ts)
fac <- as.data.frame(fac.ts)

# define lookback period and index start
nobs <- nrow(fac)
lookback.period <- nYears*12
first <- lookback.period + 1

# constructing the market portfolio (VWMP)
fac.period <- fac[first:nobs,]
```

```
rf <- fac.period[,"RF"]
ret.mkt <- fac.period[,"Mkt.RF"] + rf # return on VWMP

# constructing the equally-weighted portfolio (EWP)
N <- ncol(ret)
w.naive <- rep(1/N, N) # defining the weights
naive.period <- ret[first:nobs,]
r.i <- naive.period*w.naive
ret.naive <- rowSums(r.i) # return on EWP

# constructing the minimum variance portfolio (MinVP) and the mean-variance portfolio (MeanVP)
n <- length(rf)
ret.minvar <- rep(0,n)
ret.markowitz <- rep(0,n)

for (i in 1:n) {
  # find the indices of the lookback period
  period.end <- first + i - 2
  period.start <- first - lookback.period
  if (period.start < 1) period.start <- 1 # index cannot be less than 1

  # estimate the covariance matrix
  covmat <- var.shrink.eqcor(ret[period.start:period.end,])   # using shrinkage estimator
  #covmat <- cov(ret[period.start:period.end,])               # without shrinkage estimator

  # compute the weights of the minimum variance portfolio
  w.minvar <- minvarport(covmat, shorts=FALSE)                # MinVP weights, no shorts
  ret.minvar[i] <- sum(w.minvar*ret[(period.end+1),])         # return on the MinVP

  # compute the mean variance portfolio
  er <- apply(ret[period.start:period.end,], 2, mean)         # mean return in the period
  w.markowitz <- tanmaxutility(er, covmat, rf[i])             # mean
```

```
variance weights
  ret.markowitz[i] <- sum(w.markowitz*ret[(period.end+1),])   # return on the MeanVP
}

# portfolio capital accumulation
ind.mkt <- cumprod(c(1, 1+ret.mkt))                 # VWMP
ind.minvar <- cumprod(c(1, 1+ret.minvar))           # minimum variance portfolio
ind.naive <- cumprod(c(1, 1+ret.naive))             # 1/N portfolio
ind.markowitz <- cumprod(c(1,1+ret.markowitz)) # mean variance portfolio

# accumulated wealth at the end of the period
finalwealth.minvar <- round(ind.minvar[n+1], digits=2)          # minimum variance portfolio
finalwealth.market <- round(ind.mkt[n+1], digits=2)             # VWMP
finalwealth.naive <- round(ind.naive[n+1], digits=2)            # 1/N portfolio
finalwealth.markowitz <- round(ind.markowitz[n+1], digits=2) # mean variance portfolio

ind <- ts(cbind(ind.mkt, ind.minvar, ind.naive, ind.markowitz), start=c(year.start,1), frequency = 12)
plot( log(ind), plot.type = "single", col = c("red", "blue", "green", "pink"),
      ylab="Log growth of $1", xlab="Year" )
legend(x="topleft", legend=c("Market", "Minimum variance", "Naive", "Markowitz"),
       col = c("red", "blue", "green", "pink"), lty=1)

# computing excess return
mkt <- ret.mkt - rf               # VWMP
minvar <- ret.minvar - rf         # MinVP
naive <- ret.naive - rf           # EWP
markowitz <- ret.markowitz - rf   # MeanVP

# annualized out-of-sample Sharpe ratio
Sharpe.mkt <- Sharpe(mkt)               # VWMP
```

```
Sharpe.minvar <- Sharpe(minvar)         # MinVP
Sharpe.naive <- Sharpe(naive)           # EWP
Sharpe.markowitz <- Sharpe(markowitz) # MeanVP

# various annualized measures
an.mean(minvar)     # mean excess return on MinVP
an.std(minvar)      # sample volatility of MinVP
an.mean(mkt)        # mean excess return on VWMP
an.std(mkt)         # sample volatility of VWMP
an.mean(naive)      # mean excess return on EWP
an.std(naive)       # sample volatility of EWP
an.mean(markowitz) # mean excess return on MeanVP
an.std(markowitz)  # sample volatility of MeanVP

# testing statistical significance of Sharpe ratios across portfolios
SharpeTest(minvar, naive)
SharpeTest(markowitz, naive)
SharpeTest(mkt, naive)
```

Listing 2: R code for expanding-window approach

B.3 R code for the different functions used

The following R code shows how the various functions implemented in the main programs have been created.

```
# computes the weights for the minimum variance portfolio
minvarport <- function(covmat, shorts=TRUE) {
  covmat <- as.matrix(covmat)
  if (shorts==TRUE) {
    # this part will allow for shorts
    inv.cov <- solve(covmat)           # computes the inverse of the variance covariance matrix
    ones.vec <- rep(1, nrow(covmat)) # creates a vector of ones
    w.minvar <- (inv.cov %*% ones.vec)/as.numeric(ones.vec %*% inv.cov %*% ones.vec) # creates the weights
  }
  else if (shorts==FALSE) {
    # this part will not allow for shorts
    N <- nrow(covmat)
```

```
      zeros <- rep(0, N)           # making d a vector of zeros in order to get the right expression
      Amat <- cbind(t(array(1, dim = c(1,N))), diag(N))
      bvec <- c(1, rep(0, N))      # making the restrictions vector
      result <- solve.QP(Dmat=covmat, dvec=zeros, Amat, bvec, meq=1)
      w.minvar <- result$solution # creates the weights
    }
  return(w.minvar)
}

# computes the mean-variance portfolio weigths
tanmaxutility <- function(er, covmat, r, A=1000) {
  # compute the tangency portfolio by maximizing mean-variance utility
  # in the presence of restrictions on short sale and borrowing
  # Utility function: U(x) = E[x]-A*Var[x]
  #
  # inputs:
  # er               N x 1 vector of expected returns
  # covmat                  N x N covariance matrix of returns
  # r                scalar, risk-free rate return
  # A                scalar, risk-aversion coefficient
  #
  # output is a vector of portfolio weights

  #
  # check for valid inputs
  #
  er <- as.vector(er)
  covmat <- as.matrix(covmat)
  if (!is.numeric(r) || length(r) !=1)
    stop("Risk-free rate is not a scalar")
  if(length(er) != nrow(covmat))
    stop("Mismatch in number of rows")
  if(any(diag(chol(covmat)) <= 0))
    stop("Covariance matrix is not positive definite")

  n = nrow(covmat)
  Dmat <- covmat
```

```
  dvec <- (er-r)/A
  Amat <- cbind(rep(-1,n), diag(1,n))
  bvec <- c(-1, rep(0,n))
  result <- solve.QP(Dmat=Dmat,dvec=dvec,Amat=Amat,bvec=bvec,meq=0)
  w <- round(result$solution, 6)
  if (!all(w == 0)) w <- w/sum(w)

  return(w)
}

# performs the Sharpe significance testing
SharpeTest <- function(ex1, ex2) {
  # test for equality of two Sharpe ratios
  # returns the p-value of the test
  if (length(ex1) != length(ex2))
    stop("Different lengths of two returns!")

  SR1 <- mean(ex1)/sd(ex1)
  SR2 <- mean(ex2)/sd(ex2)
  ro <- cor(ex1,ex2)
  n <- length(ex1)
  z <- (SR2-SR1)/sqrt( (2*(1-ro)+0.5*(SR1^2+SR2^2-2*SR1*SR2*ro^2))/n )
  pval <- 2*pnorm(-abs(z))
  return(pval)
}

# computes the in-sample Sharpe ratio
maxsharpe <- function(er, covmat, r) {
  # compute the maximum Sharpe ratio
  #
  # inputs:
  # er          N x 1 vector of expected returns
  # covmat                N x N covariance matrix of returns
  # r           scalar, risk-free rate return
  #
  # output is the maximum Sharpe ratio

```

```
  #
  # check for valid inputs
  #
  er <- as.vector(er)
  covmat <- as.matrix(covmat)
  if (!is.numeric(r) || length(r) !=1)
    stop("Risk-free rate is not a scalar")
  if(length(er) != nrow(covmat))
    stop("Mismatch in number of rows")
  if(any(diag(chol(covmat)) <= 0))
    stop("Covariance matrix is not positive definite")

  #
  # computations
  #
  ones <- rep(1, nrow(covmat))  # vector of ones
  covmat.inv <- solve(covmat)   # inverse of covariance matrix
  er <- er - r*ones
  s <- sqrt(er %*% covmat.inv %*% er)
  annual.s <- s*sqrt(12)        # NB! input must be monthly returns!
  return(as.numeric(annual.s))
}



# computes the annualized Sharpe ratio from monthly
Sharpe <- function(er) {
  return(mean(er)/sd(er)*sqrt(12))
}

# computes the annualized mean return from monthly
an.mean <- function(er) {
  return(mean(er)*12)
}

# computes the annualized standard deviation from monthly
an.std <- function(er) {
  return (sd(er)*sqrt(12))
```

```
}
```

Listing 3: R code of functions used

B.4 R code for creating the bar charts

The following R code shows how the graphical plots from Section 5 has been created. Note that the numbers has been changed for each out-of-sample time period.

```
### Barplots of average Sharpe ratios ###

# MinVP
SRmvp5y <- 0.4306
SRmvp10y <- 0.4202
SRmvp20y <- 0.4196
SRmvpall <- 0.4177
mvp.vector <- c(SRmvp5y, SRmvp10y, SRmvp20y, SRmvpall)
vector.names <- c("5 yr.", "10 yr.", "20 yr.", "all")
names(mvp.vector) <- vector.names

# MeanVP
SRmv5y <- 0.4716
SRmv10y <- 0.4485
SRmv20y <- 0.4321
SRmvall <-  0.4369
mv.vector <- c(SRmv5y, SRmv10y, SRmv20y, SRmvall)
names(mv.vector) <- vector.names

# EWP
SRewp <- 0.4323

# VWMP
SRvwmp <- 0.3648

# In-sample MeanVP
SRins <- 0.7800

# Creating the plot
group <- cbind(mvp.vector, mv.vector, SRewp, SRvwmp, SRins)
rownames(group) <- vector.names
colnames(group) <- c("MinVP", "MeanVP", "EWP", "VWMP", "in-sample")
```

```
n <- ncol(group)
colors <- heat.colors(n)
barplot(group, beside=TRUE, col=colors, ylim=c(0,1), main="Average Sharpe ratios")
legend("topleft", rownames(group), fill = colors, bty = "n", horiz=TRUE)
```

Listing 4: Creating the bar charts