On the Performance of Mean-Risk Portfolio Optimization Strategies

Andreas Blanck
On the Performance of Mean-Risk Portfolio Optimization Strategies

Andreas Blanck

Bachelor of Science in Economics

Umeå University
June 2018
On the Performance of Mean-Risk Portfolio Optimization Strategies
Bachelor of Science in Economics, Umeå University
Author: Andreas Blanck, anbl0063@student.umu.se
Supervisor: Carl Lönnbark, Department of Economics, Umeå School of Business and Economics
Abstract

Finding a portfolio strategy that entails optimal performance and risk diversification may be a complicated task for investors. In this thesis, we explore and evaluate the performances of several non-trivial portfolio optimization strategies, based on various measures of risk, to identify the optimal choice. Furthermore, as in contemporary papers, we will also assess whether or not trivial strategies are outperformed by those that rely on rigorous theoretical frameworks.

These portfolio strategies were backtested on historic U.S. stock market data ranging from 2000 to 2015, to evaluate how they had performed in reality at the time.

In accordance with similar studies, the performance of a trivial equally weighted strategy was not significantly distinguishable from the non-trivial counterparts analyzed in this paper. However, it required 42% more reallocations, possibly implying higher operational costs. By contrast, a non-trivial Expected Shortfall strategy performed better in general, particularly considering its Sharpe ratio. Based on this and mainly its appealing theoretical properties, it is deemed the best strategy to pursue in portfolio management.

**Keywords:** Portfolio Optimization, Variance, Expected Shortfall, Mean-Absolute Deviation, Lower Partial Moment, Sharpe Ratio, Backtesting
Contents

1 Introduction 1
   1.1 Background ...................................................... 1
   1.2 Approach & Aim .................................................. 3

2 General Strategy 5
   2.1 Basic Definitions ............................................... 5
   2.2 Problem Formulation ............................................. 6
   2.3 Convex Optimization .............................................. 8

3 Portfolio Strategies 10
   3.1 Variance .......................................................... 10
       3.1.1 Standard Method .......................................... 11
       3.1.2 Benchmark Hedge ......................................... 12
   3.2 Expected Shortfall ............................................... 13
   3.3 Mean-Absolute Deviation ....................................... 16
   3.4 Lower Partial Moment .......................................... 17
   3.5 Naïve Allocation ................................................ 20

4 Method 21
   4.1 Strategy Implementation ........................................ 21
       4.1.1 Variance .................................................... 22
       4.1.2 Expected Shortfall ........................................ 23
       4.1.3 Mean-Absolute Deviation ................................ 24
       4.1.4 Lower Partial Moments .................................... 24
       4.1.5 Optimization Methods ..................................... 25
   4.2 Portfolio Rebalancing .......................................... 25
<table>
<thead>
<tr>
<th>CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3 Performance Evaluation .................................. 27</td>
</tr>
<tr>
<td>4.3.1 Standard Hypothesis Test .......................... 27</td>
</tr>
<tr>
<td>4.3.2 Bootstrap Hypothesis Test .......................... 29</td>
</tr>
<tr>
<td>4.3.3 Portfolio Turnover .................................. 31</td>
</tr>
<tr>
<td>4.4 Data ..................................................... 32</td>
</tr>
<tr>
<td>4.5 Parameter Setting ...................................... 33</td>
</tr>
<tr>
<td>5 Results ....................................................... 35</td>
</tr>
<tr>
<td>5.1 Performance Plots .................................... 35</td>
</tr>
<tr>
<td>5.2 Performance Statistics ............................... 42</td>
</tr>
<tr>
<td>6 Conclusion .................................................. 45</td>
</tr>
<tr>
<td>References .................................................... 49</td>
</tr>
<tr>
<td>A Supplementary Results .................................. 52</td>
</tr>
<tr>
<td>B Portfolio Weights ...................................... 55</td>
</tr>
</tbody>
</table>
List of Figures

Figure 2.1 Efficient frontier ........................................... 7
Figure 2.2 Convex function ............................................. 9
Figure 3.1 Loss distribution ............................................. 15
Figure 4.1 Evaluation window ......................................... 26
Figure 5.1 SM (EWMA) performance ................................. 36
Figure 5.2 SM (Bootstrap) performance ......................... 37
Figure 5.3 BH (S&P 500) performance ............................... 38
Figure 5.4 ES performance ........................................... 39
Figure 5.5 MAD performance ......................................... 40
Figure 5.6 LPM performance ........................................... 41
Figure B.1 Naive and ES portfolio weights ......................... 56
Figure B.2 SM portfolio weights ..................................... 57
Figure B.3 BH and MAD portfolio weights ....................... 58
Figure B.4 LPM portfolio weights ................................... 59
List of Tables

Table 4.1  Portfolio constituents ........................................ 32
Table 5.1  Performance with monthly target $\mu_0 = 1.01$ ........ 42
Table 5.2  Test statistics based on target $\mu_0 = 1.01$ .......... 43
Table A.1  Performance with monthly target $\mu_0 = 1.015$ ....... 52
Table A.2  Test statistics based on target $\mu_0 = 1.015$ .......... 53
Table A.3  Portfolio asset returns ....................................... 54
Chapter 1

Introduction

1.1 Background

A delicate problem for investors is to find an optimal portfolio strategy that yields profitable returns at the lowest possible risk. That is, a pursued strategy should preferably withstand adverse market movements and be able to generate returns in excess of risk-free rates. Since risk-free rates are associated mainly with limited or even negligible returns, considerable investments need to be in risky assets.

The question that arises as a consequence of this is; how do we measure risk and diversify our capital prudently across different risky assets in a portfolio such that the probability of loss is reduced? After the introduction of modern portfolio theory by Markowitz (1952), this question has been less cumbersome since numerous strategies have emerged for the purpose of resolving this particular issue. These are specifically designed to facilitate the problem of measuring a portfolio’s individual assets’ risks in order to accomplish optimal diversification, such that we face the minimum level of aggregate risk given an imposed target expected rate of return (Shreve, 2004).\footnote{There are also strategies without a specific target, e.g. the Sharpe and minimum risk portfolios, these will however not be in consideration here.}

In this thesis, we will shed light on four types of mean-risk optimization strategies and analyze their performances compared to each other.\footnote{Expected target return is for simplicity estimated by a mean in this thesis, as to...}
thermore, can it really be motivated to put effort into implementing such strategies if they are unable to outperform trivial strategies? For instance, DeMiguel et al. (2009) backtest several variance optimization strategies on historical market data and find that most of these are statistically indistinguishable from a naïve method adopting equal asset allocations. In addition, Frahm et al. (2012) conduct a similar study where although more robust statistical tests with respect to portfolio performance are considered. Their results are chiefly the same as that of DeMiguel et al. (2009) and they conclude their paper by pointing out that none of the analyzed optimization strategies can be advocated or considered better than a trivial equally weighted strategy on statistical grounds. Altogether, these facts thus dispute the application of any strategy based on rigorous theoretical frameworks.

However, the shortcoming of these studies as well as the original work by Markowitz (1952) is the application of variance to measure risk. This may only be convenient when we can justify normally distributed financial returns, a property that is nearly always contradicted by what we observe in reality, supported by stylized facts (Danielsson, 2011; Cont, 2001). This implies that variance can yield flawed risk estimations and may in turn be inadequate for portfolio optimization.

The reason for this is that a deficient risk measure in optimization can give rise to asset allocations that deviate from what may be truly optimal in order to attain the best possible diversification. Consequently, an important question is how strategies adopting alternative risk measures perform over time and if these are statistically superior to those portfolios managed by trivial strategies. This is why we in this thesis will concentrate on mean-risk optimization strategies rather than those limited to mean-variance, for the purpose of decreasing the knowledge gaps that exist in this particular subject.

explain the designation mean-risk optimization.
1.2 Approach & Aim

In essence, we will concentrate on portfolio strategies applying different risk measures that although use the same imposed target portfolio rate of return. In this way, their ability to handle risk will be apparent.

To evaluate their performances, a rolling-window backtest procedure applied on historic market data, within a specified investment time horizon, was implemented in MATLAB® R2017b. By doing this, we will be able to see how the analyzed portfolio strategies had performed in reality if they were used at the time.

Additionally, it is worth to note three limitations that were made to ease the analysis, where we will omit new capital injections and taxes. The third critical restriction regards short selling, which will not be in consideration here. Short selling could of course have been included since there are elegant analytical allocation formulas in modern portfolio theory that do not constrain this. Yet, because it is relatively uncommon, probably due to the fact it is costly as well as banned or restricted occasionally (Hull, 2009; Danthine and Donaldson, 2014), it is better left out completely, leaving this study entirely dependent on numerical solutions. Another fact is that these formulas are limited to variance, a fact that is not appealing regarding the purpose of the thesis.

With this in mind, the aim is to see which of the non-trivial strategies that performs best and if these really outperform trivial methods during an investment horizon of 15 years, including the 2008 financial crisis acting as a stress test. During this horizon, every portfolio managed by a particular strategy will be rebalanced monthly to update the optimal allocations in the twelve large-cap U.S. stocks chosen to be the portfolios’ constituent risky assets. Moreover, the trivial strategies included are a naïve method embracing equal portfolio weights as in DeMiguel et al. (2009) and Frahm et al. (2012), together with a passive investment in the S&P 500 index that here serves as a proxy for an index fund.

With the background and approach to the thesis outlined, we will now continue to look at how the mean-risk optimization problem can be formu-
lated generally and elaborate on a necessary property risk measures must possess to be applicable in this context. In this way will better understand the idea behind each strategy to be presented in Chapter 3 that then follows.
Chapter 2

General Strategy

Before we jump to the general strategy that is consistent with each non-trivial portfolio strategy to be presented, we start with some definitions, notations and important constraints that will imbue this thesis.

2.1 Basic Definitions

Inspired by the outlines in Hult et al. (2012), we first define the necessary portfolio budget constraint at time $t$ defined as

$$V_t = \sum_{i=1}^{n} h_{i,t} S_{i,t},$$

(2.1)

where $h_{i,t}$ denotes the integer quantity of stock $i \in \{1, 2, ..., n\}$ with price $S_{i,t} \in \mathbb{R}_+$ at time $t$. $V_t \in \mathbb{R}_+$ is our initial endowment and portfolio value, thus implying full budget investments.\(^1\) Further, $h_{i,t} \geq 0$ since we do not allow short selling within the framework of this study. This constraint is accompanied by the one capturing the portfolio value at the end of the investment period, $t + 1$, i.e.

$$V_{t+1} = \sum_{i=1}^{n} h_{i,t} S_{i,t+1}.$$  

(2.2)

\(^1\)There are extensions of modern portfolio theory that also involve a risk-free asset to invest in, implying that we would have an inequality instead. Here though, we will only focus on how to optimally allocate our capital across the $n$ stocks exclusively.
Equations (2.1) and (2.2) do provide a good intuition of how the composition of a portfolio actually looks like, henceforth however, we will not employ them but rather work with monetary weights and returns, why we replace Equation (2.1) with

\[ V_t = \sum_{i=1}^{n} w_{i,t} \tag{2.3} \]

and Equation (2.2) is similarly rewritten to

\[ V_{t+1} = \sum_{i=1}^{n} w_{i,t} R_{i,t}, \]

where the weights are defined as \( w_{i,t} = h_i S_{i,t} \) and \( R_{i,t} = \frac{S_{i,t+1}}{S_{i,t}} \) is the gross return. With these definitions that are easier to work with, we are now ready to look at the general risk minimization problem formulation.

### 2.2 Problem Formulation

The application of the mean-risk strategies in this thesis can be seen as a convex minimization problem with the same set of constraints that must be fulfilled. So if we denote a particular risk measure in question by \( \rho \), the general problem formulation can be defined as follows

\[
\begin{align*}
\text{minimize} & \quad \rho(\bar{\mu}_t) \\
\text{subject to} & \quad \bar{\mu}_t^T \bar{\mu}_t = \mu_0 V_t \\
& \quad \bar{\mu}_t^T \bar{1} = V_t \\
& \quad w_{i,t} \geq 0, \quad \forall i \in \{1, 2, ..., n\}. 
\end{align*}
\tag{2.4}
\]

Here \( \bar{1} \) is a \( n \times 1 \) unit vector with ones and as we can see, we have included the budget constraint in Equation (2.3) supplemented by the imposed portfolio target return, \( \mu_0 \), at which rate we want our portfolio to grow.\(^2\)

To this problem, we assume there exists a solution in form of an optimal

\(^2\)The constraints in the formulation are by no means exclusive. For instance, if we remove the portfolio target return requirement \( \mu_0 \), we will end up with a minimum risk (MR) problem. This is the strategy with the lowest possible risk given the choice of \( \rho \), but then also with the lowest expected rate of return of all optimal allocations.
Figure 2.1: The figure depicts the efficient frontier as the black curve on the upper part of the feasible set, i.e. the hyperbolic region containing all portfolios that can be formed given an amount of assets. Along this curve, we find all optimal portfolios, $P_{\mu_0}$, consistent with a specific portfolio target return, $\mu_0$.

allocation vector, $\vec{w}^*$, that determines how we invest our capital in the $n$ available assets in such a way that the risk $\rho(\vec{w}^*)$ it entails is the smallest possible given our target rate of return $\mu_0$. This can be seen as choosing a portfolio in the feasible set, i.e. the set of all portfolios which can be formed by the $n$ assets, that yields the best return in relation to risk and thus implies optimal diversification.

To be concrete, we see in Figure 2.1 the feasible set highlighted in gray and embodied by a hyperbola. Still, only few allocations imply optimal solutions as noted above and these trace out an efficient frontier. This frontier begins in the portfolio with minimum risk, $P_{MR}$ and continues along the upper part of the hyperbola marked by the bold line (Merton, 1972).

However, in order to obtain an optimal solution in a tractable manner, a risk measure needs to be convex, as must the set of possible allocation vectors (Hult et al., 2012). If so, measured risk in a portfolio will reduce...
in conjunction with greater diversification and not the opposite, which by contrast is not guaranteed when there is a lack convexity (Föllmer and Schied, 2002; McNeil et al., 2015).

2.3 Convex Optimization

To get a better understanding of what convexity of a risk measure implies, we will briefly analyze this property in the outlines of Hult et al. (2012) and Föllmer and Schied (2002).

As noted above, we need to have a set containing all possible allocations and a risk measure that are both convex in order to guarantee globally optimal solutions, leading us to the following definition.

\textbf{Definition 2.1. Convex Risk Measure} If we denote the convex set of allocation vectors as \( \Omega \subset \mathbb{R}^n \) and then let two arbitrary allocation vectors \( \vec{w}_A, \vec{w}_B \in \Omega \) such that \( \phi \vec{w}_A + (1 - \phi) \vec{w}_B \in \Omega \) with \( \phi \in [0, 1] \), the convex risk measure with the mapping property \( \rho : \Omega \rightarrow \mathbb{R} \) satisfies

\[
\rho(\phi \vec{w}_A + (1 - \phi) \vec{w}_B) \leq \phi \rho(\vec{w}_A) + (1 - \phi) \rho(\vec{w}_B).
\]

Hence, the risk in the mixed portfolio including allocations \( \vec{w}_A \) and \( \vec{w}_B \) is less or equal than that of two individual portfolio allocations’ weighted average, thus \( \rho \) promotes diversification.

A simple example of this in two dimensions, just to understand the intuition, would be if we only have two stocks to invest in, where \( w_1 \) and \( w_2 \) are the amounts we allocate in the two stocks respectively as usual. The convex property of a risk measure then entails the characteristics seen in Figure 2.2.

In this figure, we clearly see that the aggregated risk of holding the stocks separately is higher than the risk in the portfolio containing both stocks in relative proportions given by \( \phi \). Hence, the problem formulated in (2.4) will have a solution in form of an allocation vector \( \vec{w}^*_t \) that is globally optimal, given that \( \exists \Omega \in \mathbb{R}^n \) and \( \rho \) is convex. So, we see that convexity is a crucial property in order to face tractable minimization problems with reliable solutions.
The idea that a proper risk measure ought to be convex was introduced by Föllmer and Schied (2002) in order to augment four properties presented by Artzner et al. (1999) a sound risk measure should satisfy in order to be designated as what they call coherent. Although, in this thesis we will only concentrate on risk measures that are convex and not necessarily coherent, which is a stronger criteria.

In conclusion, with the gained knowledge of the basic problem formulation consistent with all risk measures to be presented and the intuition behind the important property of convexity they possess, we are now ready to look at the first portfolio strategy. This strategy involves variance, precisely as in the very first work in modern portfolio theory by Markowitz (1952).
Chapter 3

Portfolio Strategies

Through this chapter, we will explore several strategies adopting four different methods of measuring the risk, \( \rho \), in a portfolio with the aim of finding an optimal allocation vector \( \vec{w}_t^* \). These strategies originate from various sources of literature and a remark is that the variable notations and interpretations in the original works have been slightly altered to be consistent with those used in this thesis.

We will start with a strategy that resembles the original theoretical framework introduced by Markowitz (1952) to minimize the risk in a portfolio by using variance, but where will follow the outlines of Hult et al. (2012) for simplicity.

3.1 Variance

The first strategy presented will hereinafter will be denoted as the Standard Method (SM) to separate it from the second variance minimization strategy of this section, namely the one including a benchmark as we will soon see.\(^1\)

\(^1\)An important remark is that, even though we are working with variance, we implicitly work with volatility as the formal risk measure since this is equivalent with standard deviation.
3.1.1 Standard Method

Recall that we seek an optimal allocation vector \( \vec{w}^*_t = [w^*_1,t, w^*_2,t, ..., w^*_n,t]^T \) that minimizes the risk in our portfolio investment. When using variance, we can state this as trying to minimize of the future random variable \( V_{t+1} \), that is, our portfolio value in the next period that depends on its constituent assets’ return random variables \( \vec{R}_t = [R_{1,t}, R_{2,t}, ..., R_{n,t}]^T \) in the interval \([t, t+1]\).

Hence we need to minimize

\[
\text{Var}[V_{t+1}] = \text{Var}[\vec{w}_t^T \vec{R}_t] = E \left[ (\vec{w}_t^T (\vec{R}_t - \vec{\mu}_t))^2 \right] = E \left[ (\vec{w}_t^T (\vec{R}_t - \vec{\mu}_t)) (\vec{w}_t^T (\vec{R}_t - \vec{\mu}_t))^T \right] = \vec{w}_t^T E \left[ (\vec{R}_t - \vec{\mu}_t) (\vec{R}_t - \vec{\mu}_t)^T \right] \vec{w}_t
\]

where we have the expected asset return vector \( \vec{\mu}_t = E[\vec{R}_t] \) and applied the definition of variance, which is stated as follows.

**Definition 3.1. Variance** Variance of asset returns is the defined as

\[
\text{Var}(\vec{R}_t) = E[(\vec{R}_t - \vec{\mu}_t)(\vec{R}_t - \vec{\mu}_t)^T] = \text{Var}(\vec{R}_t)
\]

and is thus the expected value of the squared mean dispersions.

The definition of variance thus implies the \( n \times n \)-sized return covariance matrix

\[
\Sigma = \begin{pmatrix}
\sigma_1^2 & \cdots & \sigma_{1,n} \\
\vdots & \ddots & \vdots \\
\sigma_{n,1} & \cdots & \sigma_n^2
\end{pmatrix}
\]

By assumption, \( \Sigma \) is positive definite and \( \sigma_{ij} \) symbolizes the covariance between the returns of assets \( i \) and \( j \).\(^2\) Further, the square roots of the main diagonal represent the assets’ respective volatilities, \( \sigma_i \), which emphasizes that we are indirectly working with volatility as the formal risk measure.

\(^2\)Note that the time subscript has been suppressed to ease notations. \( \Sigma \) should still be understood as the covariance at time \( t \), i.e. \( \Sigma_t \).
With the objective function of the minimization problem now at our hand, we set up the apparent quadric minimization problem in the following way

\[
\begin{align*}
\text{minimize} & \quad \vec{w}_t^T \Sigma \vec{w}_t \\
\text{subject to} & \quad \vec{w}_t^T \vec{\mu}_t = \mu_0 V_t \\
& \quad \vec{w}_t^T \vec{1}_t = V_t, \\
& \quad w_{i,t} \geq 0, \quad \forall i \in \{1, 2, \ldots, n\}.
\end{align*}
\]

This problem is analytically solvable by using Lagrangian multipliers, but since we have ruled out short selling, this is unfortunately as far we can simplify this problem. This is naturally also the case for our next strategy which base its allocation upon on a benchmark in order to hedge the portfolio.

### 3.1.2 Benchmark Hedge

If we instead want our portfolio to deviate as little as possible from the movement of a benchmark, e.g. a stock index, we can set up a similar but somewhat different optimization problem in contrast to what we saw above. The idea of this Benchmark Hedge (BH) strategy is that we want the variance of the difference between the benchmark random value, \(L_t\), and our portfolio value to be as little as possible, while we at the same time require a specified portfolio return, \(\mu_0\). So we want to minimize

\[
\text{Var}[V_{t+1} - L_t] = \text{Var}[\vec{w}_t^T \vec{R}_t - L_t]
\]

\[
= \text{Var}[\vec{w}_t^T \vec{R}_t] + \text{Var}[L_t] - 2 \text{Cov}[\vec{w}_t^T \vec{R}_t, L_t]
\]

where

\[
\text{Cov}[\vec{w}_t^T \vec{R}_t, L_t] = \vec{w}_t^T \mathbb{E} \left[ \left( \vec{R}_t - \vec{\mu}_t \right) \left( L_t - \mathbb{E}[L_t] \right) \right] = \vec{w}_t^T \Sigma_{L,\vec{R}_t}
\]
and we have benchmark return covariance vector defined as
\[
\Sigma_{L,R} = \mathbb{E}\left[ (\bar{R} - \bar{\mu})(L - \mathbb{E}[L]) \right] = \begin{pmatrix} \sigma_{R_1,L} \\ \sigma_{R_2,L} \\ \vdots \\ \sigma_{R_n,L} \end{pmatrix}.
\]

Thus we can set up the new optimization problem when using a benchmark as
\[
\begin{align*}
\text{minimize} & \quad \bar{w}_t^T \Sigma \bar{w}_t - 2 \bar{w}_t^T \Sigma_{L,R,t} + \sigma_{L,t}^2 \\
\text{subject to} & \quad \bar{w}_t^T \bar{\mu}_t = \mu_0 V_t \\
& \quad \bar{w}_t^T \bar{1} = V_t \\
& \quad w_{i,t} \geq 0, \quad \forall i \in \{1, 2, \ldots, n\}.
\end{align*}
\]
Here \(\sigma_{L,t}^2\) is the variance of the benchmark random variable.

The solution to this problem can be regarded as twofold, on the one hand the composition of \(\bar{w}^*\) intends to meet our required target portfolio return \(\mu_0\), but on the other hand strives to mimic the benchmark to reduce the variance between \(V_{t+1}\) and \(L\). Yet this is for obvious reasons not clear without an analytical expression for \(\bar{w}^*\) that would clarify this fact. Permitting short selling however, the problem can be solved analytically by using Lagrangian multipliers as previously and would prove the assertion above.

Now we have seen two ways of optimizing a portfolio using variance, which although may entail improper weights as a direct consequence of possibly biased risk estimates. Therefore, we will now carry on to see the first alternative to variance.

### 3.2 Expected Shortfall

We will begin with the proven coherent risk measure Expected Shortfall as our first alternative to minimize portfolio risk and this approach was first introduced by Rockafellar and Uryasev (2000). Before we acquaint ourselves in greater detail with this risk measure, it is necessary to first define the related Value at Risk (VaR), which is the risk measure mostly used in finan-
cial regulations (Daníelsson, 2011) and is defined by McNeil et al. (2015) as follows.

**Definition 3.2. Value at Risk** The \( \alpha \) confidence level VaR is the quantile of a portfolio’s loss distribution, \( f \), where the probability of a loss, \( \ell \), being equal or in excess of this real-valued number is \( 1 - \alpha \), that is

\[
\text{VaR}_\alpha := \inf \{ \ell \in \mathbb{R} : F(\ell) \geq \alpha \},
\]

where \( F(\ell) \) is the loss distribution’s cumulative distribution function.

Now following Rockafellar and Uryasev (2000), the \( \alpha \) confidence level VaR can be defined by the cumulative distribution function of losses, namely

\[
F(\text{VaR}_\alpha) = \int_{\ell \leq \text{VaR}_\alpha} f(\tilde{y}) \, d\tilde{y} = \alpha.
\]

Here \( f(\tilde{y}) \) is the assumed continuous joint probability density function of the net rate of return vector \( \tilde{y} = \tilde{R} - \tilde{I} \), where \( \tilde{y} \in \mathbb{R}^n \). Note that we integrate up to the returns that yield a loss smaller than or equal to VaR\(_\alpha\), that is, we assume the loss, \( \ell(\vec{w}, \tilde{y}) \), is a function of the fixed allocation \( \vec{w} \) and return \( \tilde{y} \) vectors.\(^3\)

We are now equipped with the necessary tools to define Expected Shortfall (ES), this is done in line with Rockafellar and Uryasev (2000) as well as Daníelsson (2011).

**Definition 3.3. Expected Shortfall** The \( \alpha \) confidence level Expected Shortfall is defined as

\[
\text{ES}_\alpha(\vec{w}) := \mathbb{E} \left[ \ell(\vec{w}, \tilde{y}) \mid \ell(\vec{w}, \tilde{y}) \geq \text{VaR}_\alpha \right].
\]

and is thus the expected portfolio loss beyond the VaR\(_\alpha\) quantile of the loss distribution, governed by \( f(\tilde{y}) \).

Before we continue to the ES optimization problem, it can be helpful to see a demonstration of how the two risk measures relates to each other and

\(^3\)Assuming \( \vec{w} \) is held constant, what really governs losses are the return outcomes, described by their joint distribution \( f(\tilde{y}) \). Furthermore, worth to point out is that \( \text{VaR}_\alpha(\vec{w}) \) is also a function of \( \vec{w} \) even though it is not stated explicitly.
a hypothetical loss distribution that in turn depends on the distribution of returns, this is depicted in Figure 3.1.\footnote{\text{Losses are here seen as positive numbers, whereas profits are negative. VaR is often related to portfolio risk capital requirements for the purpose of acting as a cushions against adverse market movements and a positive number is thus more intuitive.}}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig3.1.png}
\caption{VaR$_\alpha$ is the quantile that separates losses that occurs with $\alpha\%$ and those with $(1-\alpha)\%$ probability. The shaded area is to highlight that the losses from which ES$_\alpha$ is estimated. $E[\ell]$ is the expected loss which we here assume is negative (our expected portfolio profit).}
\end{figure}

To arrive to the proper problem formulation, we begin by writing ES in integral form as follows

$$\text{ES}_\alpha(\vec{w}) = (1-\alpha)^{-1} \int_{\ell \geq \text{VaR}_\alpha} \ell(\vec{w}, \vec{y}) f(\vec{y}) d\vec{y}. $$

If we then rewrite this expression by instead considering the expected deviations from VaR$_\alpha$ of larger losses than this quantity, we end up with the objective function in the optimization problem as given in Rockafellar and Uryasev (2000), that is

$$\text{ES}_\alpha(\vec{w}) = \text{VaR}_\alpha + (1-\alpha)^{-1} \int_{\ell \in \mathbb{R}^n} \max \left( \ell(\vec{w}, \vec{y}) - \text{VaR}_\alpha, 0 \right) f(\vec{y}) d\vec{y}. \quad (3.2)$$

As we can see, this expression thus connects the two elaborated risk measures.

Equation (3.2) is the objective function in the problem, but in order to apply it we must approximate the integral, this is simply done by a mean
value, why we end up with

$$\begin{align*}
\text{minimize} & \quad \text{VaR}_{\alpha,t} + \frac{1}{N(1-\alpha)} \sum_{j=1}^{N} \max \left( \ell_j(\vec{w}_t, \vec{y}_t) - \text{VaR}_{\alpha,t}, 0 \right) \\
\text{subject to} & \quad \vec{w}_t^T \vec{\mu}_t = \mu_0 V_t \\
& \quad \vec{w}_t^T \vec{1} = V_t \\
& \quad w_{i,t} \geq 0, \quad \forall t \in \{1, 2, ..., n\},
\end{align*}$$

where we take a loss to be $\ell_j(\vec{w}_t, \vec{y}_t) = \vec{w}_t^T (\vec{1} - \vec{R}_{j,t})$ given a return vector scenario forecast at time $t$; $\vec{R}_{j,t} = [R^*_1,t,j, R^*_2,t,j, ..., R^*_n,t,j]^T$ with $j \in \{1, 2, ..., N\}$.

To summarize, we now have the necessary optimization problem which we can apply in order to solve for the allocation vector $\vec{w}^*$. Moreover, an important remark is that while ES is suitable for portfolio optimization, VaR is not, even though it is widely applied in the financial industry in many other contexts. This is due to that it is not a convex function, why it may entail solutions that are not globally optimal nor reliable (Rockafellar and Uryasev, 2000).

### 3.3 Mean-Absolute Deviation

Next, we consider the Mean-Absolute Deviation (MAD) risk measure presented by Konno (1988) to streamline the model of Markowitz (1952) and remedy some of its deficiencies. In particular, the mean-variance framework needs a covariance matrix determined in advance before the actual optimization, which may be associated with heavy calculations. This is on the contrary not necessary in MAD since it entails a linear problem as in ES above, meaning that it should be solvable with less computational effort. Another advantage argued by Konno (1990) is that a greater number of zero-valued asset weights ($\vec{w}^*$ has more elements that are zero) is expected in contrast to the mean-variance approach. This will hence reduce possible transaction costs incurred from purchasing assets to be included in and rebalancing the

---

5This is of course only an advantage if the number of assets to invest in is large. The benefits of diversification come from the fact that we spread our wealth across several and not few risky assets.
portfolio (Konno and Yamazaki, 1991).

With its close connection to variance, MAD is not surprisingly defined in similarity to the definition seen in Equation (3.1).

**Definition 3.4. MAD** Mean-Absolute Deviation is the expected value of mean dispersions, or formally

\[
\text{MAD}(\vec{\omega}) := E\left[|\vec{\omega}^T(\tilde{R} - \vec{\mu})|\right]
\]  

(3.3)

and is thus an analog to the Manhattan norm.

Furthermore, Konno and Yamazaki (1991) show that, if \(\tilde{R}\) is normally distributed, variance and the MAD differ only by a factor of \(\sqrt{2}\) and will thus yield the same allocation vector \(\vec{\omega}^*\) in a minimization problem. Yet this is as we know not the statistical properties of returns we observe in reality and different results from the two risk measures are therefore expected.

To conclude this section, we state the optimization problem using MAD and where we simply approximate the expected value in Equation (3.3) with a mean as follows

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{N} \sum_{j=1}^{N} |\vec{\omega}_t^T(\tilde{R}_{j,t} - \tilde{\mu}_t)| \\
\text{subject to} & \quad \vec{\omega}_t^T \tilde{\mu}_t = \mu_0 V_t \\
& \quad \vec{\omega}_t^T \vec{1} = V_t \\
& \quad w_{i,t} \geq 0, \quad \forall i \in \{1, 2, ..., n\},
\end{align*}
\]

using \(j \in \{1, 2, ..., N\}\).

### 3.4 Lower Partial Moment

We now move on to the last risk measure to be considered in this thesis and it is the Lower Partial Moment (LPM) which has the same basic approach as ES. That is, both these measures only pay attention to the tail of (positive) losses rather than the entire distribution of losses, including those that are negative (i.e., profits). Contrary, the formulations of MAD and variance
above are unable to distinguish whether extreme variations in asset returns are outcomes of adverse or desirable market movements, why both events will be deemed equally risky. This is of course a contradiction to what we intuitively define a risk (Harlow, 1991).\(^6\)

Before we elaborate on the reason why the formulation of LPM is more appropriate, at least theoretically, than variance and MAD in terms of measuring risk, it is worth noting that there is a broad range of literature capturing this framework, plenty of which with focus on its mathematical properties. The definition of LPM with connection to portfolio optimization is although well described in Harlow (1991) as well as Harlow and Rao (1989), where it is stated similarly as

\[
LPM_\gamma(\tau, \vec{w}) = \int_{[\tau, y_T]} \left( \tau - \vec{w}^T \hat{R} \right)^\gamma f(\vec{y}) \, d\vec{y},
\]

(3.4)

where we integrate up to a threshold return \( y_T \) and recall the definition of gross returns \( \hat{R} = (\vec{y} + \vec{1}) \). This threshold return defines a benchmark value in reference to our initial portfolio value \( V_t \), i.e. \( \tau = (1 + y_T) V_t \), below which we deem outcomes are losses. To clarify this, we rewrite the expression in Equation (3.4) by integrating over all space \( \mathbb{R}^n \), that is

\[
\int_{\mathbb{R}^n} \max \left( \tau - \vec{w}^T \hat{R}, 0 \right)^\gamma f(\vec{y}) \, d\vec{y} = E \left[ \max \left( \tau - \vec{w}^T \hat{R}, 0 \right)^\gamma \right],
\]

enabling us to define LPM more neatly in the following way.

**Definition 3.5.** LPM Lower Partial Moment risk measure with moment \( \gamma \) is defined as

\[
LPM_\gamma(\tau, \vec{w}) := E \left[ \max \left( \tau - \vec{w}^T \hat{R}, 0 \right)^\gamma \right],
\]

(3.5)

and is thus dispersions from the benchmark value \( \tau \) in the distribution’s tail of positive losses.

We therefore see that we end up with an expectation value in similarity to the definitions of MAD and variance, only that we merely take the tail of losses into consideration and have the ability to determine \( \gamma \), which re-

\(^6\)A risk in a portfolio should naturally be the associated with a loss its of value, not an increase.
lates to an investor’s perception of risk and thus his or hers utility function. Specifically, we are able to apply primarily two measures, the target shortfall and semivariance depending on the choice of moment, $\gamma = 1$ or $\gamma = 2$.\footnote{Semivariance was actually considered by Markowitz prior to the original portfolio theory in place of variance because of its more appropriate properties. However, since variance is easier to handle mathematically, it was the measure of choice in his original work (Mao, 1970).}

Moreover, instead of focusing on mean dispersions, we consider asymmetric deviations from a benchmark $\tau$ whose value we set optionally and best suit our preferences of risk (Harlow, 1991). For example, if we define losses to be where the portfolio value in the next period fulfills $V_{t+1} < V_t$, we set $y_r = 0$.

With this short introduction of LPM, we now define the relevant optimization problem, where we again approximate the expectation value in Equation (3.5) with a mean value in the following way

$$\text{minimize} \quad \frac{1}{N} \sum_{j=1}^{N} \max \left( \tau - \bar{\mathbf{w}}_t^T \bar{\mathbf{R}}_{j,t}, 0 \right)^\gamma$$

subject to

$$\bar{\mathbf{w}}_t^T \bar{\mathbf{\mu}}_t = \mu_0 V_t$$

$$\bar{\mathbf{w}}_t^T \bar{\mathbf{I}} = V_t$$

$$w_{i,t} \geq 0, \quad \forall i \in \{1, 2, ..., n\},$$

with $j \in \{1, 2, ..., N\}$. Analyzing the above minimization problem, we see that if we choose $\gamma = 1$ (target shortfall), we end up with a linear problem, whereas $\gamma = 2$ (semivariance) yields a quadric formulation.

In summary, LPM have some superior properties over variance and MAD since it allows us to determine what is a loss and above all, only focuses on losses, making it a more sound measure of risk. However, whether or not this along with all other risk measures elaborated in this chapter are superior to a trivial naïve strategy is an important question to answer. We therefore continue by introducing the last and indubitably most simple dynamic investment strategy of this thesis.
3.5 Naïve Allocation

Naïve allocation should here be interpreted as an *equally weighted* portfolio strategy as in DeMiguel et al. (2009) as well as Frahm et al. (2012) and is thus not reliant on any theoretical framework whatsoever. Therefore, we do not need to set up an optimization problem, nor do we have to worry about any types of estimation bias that may flaw the results.

To allocate our capital in this manner, we simply set all weights equally as

\[ w_{i,t} = \frac{1}{n} V_t, \quad \forall i \in \{1, 2, \ldots, n\}, \]

so that every time we reallocate, we spread our wealth uniformly among all assets we opt to constitute our portfolio.

To conclude, we have now explored the theoretical backgrounds of all strategies that relate to dynamic management in this study and we will therefore move on to next chapter describing the pursued methodology.
Chapter 4

Method

In this chapter, the approach in which the study was carried out is presented in detail. We will firstly see how all models were implemented, the way the portfolio was rebalanced monthly, how performance was assessed and lastly, which data as well as parameters that were applied.

4.1 Strategy Implementation

Before we look at the details about the portfolio rebalancing, we better start with how the strategies were implemented to understand the concept of a rolling-window procedure soon to be presented. Further, an important remark is that two approaches to apply returns in the risk optimization were used, namely

1. Forecast risk using monthly historic return data, $\tilde{R}$, from a window $W_H$ by a multivariate Exponentially Weighted Moving Average (EWMA) scheme. However, it is only applicable to the variance models.

2. Forecast risk based on return scenarios one time step ahead. Specifically, historical returns bootstrapping was used to draw 20 daily returns with replacement from $W_H$ to form a monthly counterpart, $\tilde{R}$, repeated $N$ times.
4.1.1 Variance

For both strategies using variance, the main problem is to find an estimation of the covariance matrices before the actual optimization. Here, the covariance matrices were estimated by a EWMA scheme using historic data since it is an easy method to implement and still performs relatively well compared to parametric methods such as Generalized Autoregressive Conditional Heteroskedasticity (GARCH), that although are more complex to implement. A multivariate version of what is presented in Danielsson (2011) was used, that is

\[
\hat{\Sigma}_t = \frac{1 - \lambda}{\lambda(1 - \lambda^W)} \sum_{j=1}^{W_H} \lambda^j (\hat{\bar{R}}_{t-j} - \hat{\bar{\mu}}_t)^\tau (\hat{\bar{R}}_{t-j} - \hat{\bar{\mu}}_t),
\]

where \( \lambda \) is a decay factor that determines the impact of historic returns on the future, thus capturing possible volatility clustering effects. Expected returns \( \hat{\bar{\mu}}_t \) are simply estimated by the mean

\[
\hat{\bar{\mu}}_t = \frac{1}{W_H} \sum_{j=1}^{W_H} \bar{R}_{t-j},
\]

thus clarifying the title of this thesis. To estimate \( \Sigma_{L,\bar{R},t} \), we can in fact use the same scheme only by extending \( \bar{R}_t \) to become a \((n + 1) \times 1\) vector with the last element holding the return of the benchmark variable, \( R_{L,t} \). We can then employ EWMA including this lengthened return vector in place of \( \bar{R}_t \) and extract \( \hat{\Sigma}_{L,\bar{R},t} \). The demonstration of this is omitted to save space, but it is easily shown by using basic matrix algebra.

When bootstrapped return scenarios, \( \bar{R} \), were used, the analysis was limited to the SM strategy to save time. Here though, EWMA could not be used since it weighs historical returns with respect to time and applying it to future scenarios would not make sense. Consequently, a simple moving average scheme as in Frahm et al. (2012) was implemented, i.e.

\[
\hat{\Sigma}_t = \frac{1}{N} \sum_{j=1}^{N} (\hat{\bar{R}}_{j,t} - \hat{\bar{\mu}}_t)^\tau (\hat{\bar{R}}_{j,t} - \hat{\bar{\mu}}_t),
\]

where \( \hat{\bar{\mu}}_t \) is calculated as in Equation (4.1), only that the returns are now
Method Strategy Implementation

from simulated scenarios and $W_H$ is replaced by $N$.\(^1\) This estimation of $\tilde{\mu}_t$ is consistent with all forthcoming strategies that are also based on bootstrapped return scenarios.

### 4.1.2 Expected Shortfall

The optimization formulation of ES, and all other alternatives to variance for that matter, we saw in Chapter 3 looks cumbersome and not that easy to implement at first glance. However, Rockafellar and Uryasev (2000) provide a clever suggestion of how to solve this issue. They introduce auxiliary variables in order to form a convex programming approach in similarity to

$$
\begin{align*}
\text{minimize} & \quad \text{VaR}_{\alpha,t} + \frac{1}{N(1-\alpha)} \sum_{j=1}^{N} u_{j,t} \\
\text{subject to} & \quad u_{j,t} - \vec{w}^T(\vec{1} - \tilde{R}_{j,t}) + \text{VaR}_{\alpha,t} \geq 0, \quad \forall j \in \{1, 2, ..., N\} \\
& \quad u_{j,t} \geq 0, \quad \forall j \in \{1, 2, ..., N\} \\
& \quad w_{i,t} \geq 0, \quad \forall i \in \{1, 2, ..., n\} \\
& \quad \vec{w}^T \tilde{\mu}_t = \mu_0 V_t \\
& \quad \vec{w}^T \vec{1} = V_t.
\end{align*}
$$

Here, we see that the auxiliary variable $u_{j,t}$ replaces the former expression in the sum. To invoke the max function, $u_{j,t}$ most also be greater or equal to this expression as well as always be greater than zero. So with the two new supplemented constraints, we can assure that this formulation is equivalent to the previous, only that this is easier to implement.

\(^1\)This scheme could also have been applied when using historic data. The problem with this model is that it does not weigh more recent returns heavier than older and we face risk of something called ghost effects. That is, since all returns in the window $W_H$ thus become equally probable, an updated window where older returns are discarded might dramatically alter risk forecasts abruptly, in contrast to what actually reflect real market conditions. This was the reason for choosing EWMA over this simpler estimator.
4.1.3 Mean-Absolute Deviation

In the same way as ES, we can rewrite the MAD problem in similarity to what is done in Konno and Yamazaki (1991), that is

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{N} \sum_{j=1}^{N} u_{j,t} \\
\text{subject to} & \quad u_{j,t} + \bar{w}^T(\tilde{R}_{j,t} - \tilde{\mu}) \geq 0, \quad \forall j \in \{1, 2, ..., N\} \\
& \quad u_{j,t} - \bar{w}^T(\tilde{R}_{j,t} - \tilde{\mu}) \geq 0, \quad \forall j \in \{1, 2, ..., N\} \\
& \quad u_{j,t} \geq 0, \quad \forall j \in \{1, 2, ..., N\} \\
& \quad w_{i,t} \geq 0, \quad \forall i \in \{1, 2, ..., n\} \\
& \quad \bar{w}^T \tilde{\mu} = \mu_0 V_t \\
& \quad \bar{w}^T \tilde{1} = V_t.
\end{align*}
\]

The principle in this setup is the same where the auxiliary variable replaces the argument in the sum. Here though, we have three new constraints, where the two first assures the absolute sign in the original expression.

4.1.4 Lower Partial Moments

In the exact same way, we can replace the expression in the sum in the LPM strategy by the auxiliary variable as

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{N} \sum_{j=1}^{N} u_{j,t}^{\gamma} \\
\text{subject to} & \quad u_{j,t} + \bar{w}^T(\tilde{R}_{j,t} - \tilde{\mu}) \geq 0, \quad \forall j \in \{1, 2, ..., N\} \\
& \quad u_{j,t} - \bar{w}^T(\tilde{R}_{j,t} - \tilde{\mu}) \geq 0, \quad \forall j \in \{1, 2, ..., N\} \\
& \quad u_{j,t} \geq 0, \quad \forall j \in \{1, 2, ..., N\} \\
& \quad w_{i,t} \geq 0, \quad \forall i \in \{1, 2, ..., n\} \\
& \quad \bar{w}^T \tilde{\mu} = \mu_0 V_t \\
& \quad \bar{w}^T \tilde{1} = V_t.
\end{align*}
\]

As we can see, the moment $\gamma$ was not included in the auxiliary variable for convenience. Further, depending on if we set $\gamma = 1$ or $\gamma = 2$, the problem is either linear or quadric.
### 4.1.5 Optimization Methods

In the above sections, we discussed how the different risk optimization problems were reduced to become more tractable as well as describing ways of estimating the expected return vectors and covariances. To actually optimize, the MATLAB® functions `linprog` and `quadprog` were used in conjunction with the above problem formulations. Implementing any optimization algorithm alone would be too time-consuming, why built-in functions were relied upon.

For the interested however, the `linprog` function uses a dual-simplex algorithm to solve the linear optimization problem, whereas `quadprog` applies an interior-point-convex method for the quadric counterparts.²

### 4.2 Portfolio Rebalancing

The pursued approach was to backtest all portfolio strategies on real data and analyze their performance in relation to the passive and naïve (trivial) strategies during an investment time horizon we will call the evaluation window, or $W_E$. For this purpose however, a sample of historic asset prices corresponding to a time span exceeding $W_E$ was needed to enable portfolio rebalancing at the first $W_H$ months in $W_E$. The reason for this is that a set of prices prior to the first date in $W_E$ is needed for bootstrapping return scenarios and EWMA calibrations.

To be concrete, a monthly rolling-window strategy was employed where time $t = 0$ marks the starting point (start of the first period in $W_E$) and an initial allocation vector $\vec{w}_0^*$ for each strategy is obtained with the help of `linprog` or `quadprog` by applying historic stock returns prior to $t = 0$ from a window called $W_H^0$. Directly after, the returns $\vec{R}_0$ in between $t \in [0, 1]$ are used to calculate the portfolio value at $t = 1$, $V_1 = \vec{w}_0^{*T} \vec{R}_0$. Then we role one step forward to a new window $W_H^1$ for bootstrapping and EWMA calibrations such that a new optimal allocation $\vec{w}_1^*$ and portfolio value $V_2 = \vec{w}_1^{*T} \vec{R}_1$ can be determined. This is then repeated in the same fashion at every time step.

²For additional information about these functions, the reader is referred to MathWorks® homepage.
until we reach our final period and have $T$ portfolio values for each portfolio strategy. The exact approach is summarized in the following Algorithm.

**Algorithm Portfolio Rebalancing**

1: for $t = 0$ to $T - 1$ by 1 do 
2:   • Bootstrap from or estimate covariances by using $W^t_H$
3:   • Find $\vec{w}^*_t$ using a strategy
4:   • Update $V_{t+1}$ using $\vec{w}^*_t$ and $\bar{R}_t$

To get a feeling of how the rolling-window procedure is performed, we see in Figure 4.1 window $W^0_H$ from which historic stock returns are used at time $t = 0$. Subsequently, the window is rolled one step (one month) forward, discarding the oldest stock returns and at time $t = 1$, $W^1_H$ that contains more up-to-date returns is used instead.

![Figure 4.1](image)

**Figure 4.1:** We use the initial sample window $W^0_H$ in order to find our first allocations $\vec{w}^*_0$ and first portfolio value $V_1$. We then move the window one step ahead while at the same time discard the oldest returns in $W^0_H$, yielding $W^1_H$ and the process repeats.

Regarding bootstrapping returns based on historic data, this has been mentioned frequently up to now, why it is appropriate to see how these where sampled. We do that with the help of the algorithm that follows, inspired by Hult et al. (2012).

So at each time step, we create a sample of $N$ vector return scenarios based on recent historic returns. Furthermore, worth to point out is that all draws consist of a vector with asset returns from the same date, implying that market implied correlations in between assets are always maintained.
Algorithm  

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td>for $j = 1$ to $N$ by 1 do</td>
</tr>
<tr>
<td>2:</td>
<td>• Draw 20 daily return vectors $\tilde{R}_D$ from $W_H$ with replacement.</td>
</tr>
<tr>
<td>3:</td>
<td>• Create monthly return scenario vector $\tilde{R}_j = \tilde{R}_D^1 \tilde{R}_D^2 \ldots \tilde{R}_D^{20}$</td>
</tr>
<tr>
<td>4:</td>
<td>where the product should be read as componentwise.</td>
</tr>
</tbody>
</table>

4.3 Performance Evaluation

Once a series of $T$ portfolio values for each strategy are obtained, some tests are needed to properly judge which performed best and if trivial strategies statistically differ from optimized methods as in similar studies by DeMiguel et al. (2009) and Frahm et al. (2012).

To do this, we will primarily use the Sharpe ratio defined as

$$\Gamma_i = \frac{\mu_{P,i} - r}{\sigma_{P,i}},$$

where $\mu_{P,i} = \mathbb{E}[R_{P,t,i}]$ and $R_{P,t,i}$ is portfolio strategy $i$’s estimated return in the evaluation window, $r$ is the risk-free interest rate and $\sigma_{P,i}$ is the standard deviation of portfolio returns. So it is simply a measure of the risk reward in relation to the riskiness of a portfolio and this will be used in the forthcoming hypothesis tests.

4.3.1 Standard Hypothesis Test

Regarding hypothesis tests, we will as in DeMiguel et al. (2009) use a version of a hypothesis test presented by Jobson and Korkie (1981) to evaluate portfolio performance using Sharpe ratios. The foundation of this test unfortunately relies on the assumption of normally distributed portfolio returns that are also independent and identically distributed (i.i.d.) (Jobson and Korkie, 1981; DeMiguel et al., 2009). This is of course a drawback and may bias the results since financial returns are as discussed seldom exhibiting such characteristics, rather, they usually portray heavy distributional tails and autocorrelation (Cont, 2001; Danielsson, 2011).

Nevertheless, this method is very easy to implement compared to a more
appropriate statistical test used in Frahm et al. (2012) that is partially based on an approach by Ledoit and Wolf (2008). Instead of an assumption of normality and \( i.i.d. \), they bootstrap observed portfolio returns from an evaluation window and form a more realistic distribution in order to make statistical inference.

However, the more robust methods used in both these papers are too complicated to be implemented within the time frame of this thesis and are therefore omitted. A simpler version of this test described in Ledoit and Wolf (2008) will although be considered as a complement and this is presented in the subsequent section.

If we return to this method, the idea is to test whether the Sharpe ratios of two portfolio strategies \( i \) and \( j \) differ. If we thus set \( \Delta_{ij} = \Gamma_i - \Gamma_j \), we can formulate the proper hypothesis as

\[
H_0 : \Delta_{ij} = 0 \\
H_1 : \neg H_0.
\]

The estimated test statistic \( \hat{z}_{ij} \) is then stated in DeMiguel et al. (2009) as

\[
\hat{z}_{ij} = \frac{\hat{\sigma}_{P,j} \hat{\mu}_{P,i} - \hat{\sigma}_{P,i} \hat{\mu}_{P,j}}{\sqrt{\hat{\theta}}} \xrightarrow{d} \mathcal{N}(0, 1),
\]

that is, it converges in distribution to a standard normal distribution and where

\[
\hat{\theta} = \frac{1}{T} \left( \hat{\sigma}_{P,i}^2 \hat{\sigma}_{P,j}^2 - \hat{\sigma}_{P,i} \hat{\sigma}_{P,j} \hat{\sigma}_{P,ij} + \frac{1}{2} \hat{\mu}_{P,i}^2 \hat{\sigma}_{P,j}^2 + \frac{1}{2} \hat{\mu}_{P,j}^2 \hat{\sigma}_{P,i}^2 - \frac{\hat{\mu}_{P,i} \hat{\mu}_{P,j} \hat{\sigma}_{P,ij}}{\hat{\sigma}_{P,i} \hat{\sigma}_{P,j}} \right).
\]

Here \( \hat{\sigma}_{P,ij} \) is the estimated covariance between the two portfolios’ returns.\(^3\)

To calculate all these parameters, ordinary sample mean, variance and covariances are used and since it is a simple two-tailed hypothesis test, the \( p \)-value is calculated by \( p = 2 \Phi(|\hat{z}_{ij}|) \), where \( \Phi \) is the standard normal cumulative distribution function.

To conclude, it is worth pointing out that Memmel (2003) discovered

\(^3\)To clarify, all variables assigned a hat are estimations of their corresponding random variables, this fact holds throughout the entire thesis.
that \( \hat{\vartheta} \) in the test statistic has a typographical error and hence corrected it, as to explain why the original test statistic in Jobson and Korkie (1981) was not used.

### 4.3.2 Bootstrap Hypothesis Test

The second statistical test of this thesis is a bootstrap version presented in Ledoit and Wolf (2008) that uses the same hypothesis we saw in the previous section. The drawback of this method is yet another assumption of i.i.d. returns. For this reason, the authors naturally enter a caveat about applying it and instead advocates a more robust time series bootstrap alternative. But as mentioned earlier, it is much more complicated to implement.

Despite the i.i.d. assumption, it should be superior to the test described above since it relies on an empirical rather than a normal distribution. This enables distributions with heavy-tails that assign non-negligible probabilities to extreme outcomes, which is more realistic.

In this method, the idea is redefine the Sharpe difference to a function \( f(\vec{\nu}) = \Delta_{ij} \), where \( f \) is set as

\[
f(\vec{\nu}) = \frac{\mu_{P,i}}{\sqrt{\xi_{P,i} - \mu^2_{P,i}}} - \frac{\mu_{P,j}}{\sqrt{\xi_{P,j} - \mu^2_{P,j}}},
\]

with \( \vec{\nu} = (\mu_{P,i}, \mu_{P,j}, \xi_{P,i}, \xi_{P,j})^\top \) and \( \xi_{P,i} = E[R_{P,t,i}] \). Furthermore, Ledoit and Wolf (2008) state the assumption

\[
\sqrt{T}(\vec{\nu} - \hat{\vec{\nu}}) \xrightarrow{d} \mathcal{N}(\vec{0}, \Psi),
\]

so that the difference converges in distribution to a normal distribution with zero mean and a \( 4 \times 4 \)-sized positive semidefinite as well as symmetric covariance matrix \( \Psi \) that is unknown. By contrast, Ledoit and Wolf (2008) show a similar approach to the standard test seen above. There they use the assumptions in Jobson and Korkie (1981), implying that \( \Psi \) has a specific known composition and this is in fact the root of this method’s shortcomings.

We are however not interest in the formulation in (4.2) with \( \vec{\nu} \) but rather that of the (modified) Sharpe ratio differences \( \Delta_{ij} = f(\vec{\nu}) \). Ledoit and Wolf
Method Performance Evaluation

(2008) therefore apply the Delta method on (4.2).\footnote{The Delta method is a theorem that relates the distributions of a function \( f(\theta) \) and its argument, the random variable \( \theta \) (Casella and Berger, 2001).} The result of this yields the expression

\[
\sqrt{T} (\hat{\Delta}_{ij} - \Delta_{ij}) \xrightarrow{d} \mathcal{N}(0, \nabla^T f(\hat{\nu}) \Psi \nabla f(\hat{\nu})),
\]  

(4.3)

where we have the transposed gradient of \( f(\hat{\nu}) \) defined as

\[
\nabla^T f(\hat{\nu}) = \begin{pmatrix}
\frac{\xi_{P,i}}{(\xi_{P,i} - \mu_{P,i})^{1.5}}, \\
\frac{\xi_{P,j}}{(\xi_{P,j} - \mu_{P,j})^{1.5}}, \\
-\frac{1}{2} \frac{\mu_{P,i}}{(\xi_{P,i} - \mu_{P,i})^{1.5}}, \\
\frac{1}{2} \frac{\mu_{P,j}}{(\xi_{P,j} - \mu_{P,j})^{1.5}}
\end{pmatrix}.
\]

From (4.3) we can easily derive the expression for the estimation \( \hat{\Delta}_{ij} \)'s standard error

\[
\text{se}(\hat{\Delta}_{ij}) = \sqrt{\frac{\nabla^T f(\hat{\nu}) \Psi \nabla f(\hat{\nu})}{T}}.
\]  

(4.4)

Before we continue to find the test’s \( p \)-value, it is important to note that it is the way we estimate \( \Psi \) that separates this simpler version from the more robust ones. In this simplification though, we use the matrix \((R_{P,t,i}, R_{P,t,j}, R_{P,t,i}^2, R_{P,t,j}^2)^\tau, t = \{1, 2, ..., T\}\) and calculate its sample covariance which we set to \( \hat{\Psi} \). This should be an unbiased estimator when the returns are \textit{i.i.d} (Ledoit and Wolf, 2008).

What we need now is to create a distribution for \( \Delta_{ij} \) to which we can relate our observed standard error \( \text{se}(\hat{\Delta}_{ij}) \) and it is here bootstrapping comes into the picture. Specifically, we most create a total of \( N_B \) scenario standard errors \( \text{se}(\hat{\Delta}^{*k}_{ij}), k \in \{1, 2, ..., N_B\} \) and the procedure is summarized in the below algorithm.

\begin{algorithm}
\textbf{Algorithm} \hspace{0.5cm} Portfolio Return Bootstrapping
\begin{enumerate}
\item \textbf{for} \( k = 1 \) to \( N_B \) \textbf{by} 1 \textbf{do}
\item \textbf{for} \( t = 1 \) to \( T \) \textbf{by} 1 \textbf{do}
\item \hspace{0.5cm} \bullet \hspace{0.5cm} Draw a pair \( (R_{P,t,i}, R_{P,t,j})^\tau \) with replacement.
\item \hspace{0.5cm} \bullet \hspace{0.5cm} Estimate \( \hat{\nu} \) and \( \hat{\Psi} \) using samples \( (R_{P,t,i}^*, R_{P,t,j}^*)^\tau \)
\item \hspace{0.5cm} \bullet \hspace{0.5cm} Estimate \( \text{se}(\hat{\Delta}^{*k}_{ij}) \) as in Equation (4.4) but with \( \hat{\nu}^* \) and \( \hat{\Psi}^* \).
\end{enumerate}
\end{algorithm}

Once the bootstrap procedure is done, we can use a studentized test
statistic based on the observed portfolio returns, that is

\[ d = \frac{|\hat{\Delta}_{ij}|}{se(\hat{\Delta}_{ij})}. \]

We then define the centered statistic serving as the reference distribution and it is formed by the bootstrapped portfolio scenario returns as follows

\[ d^{*,k} = \frac{|\hat{\Delta}^{*,k}_{ij} - \hat{\Delta}_{ij}|}{se(\hat{\Delta}^{*,k}_{ij})}. \]

Finally, we can now state the \( p \)-value for the hypothesis seen in the previous section in the following way

\[ p = \frac{\sum_{k=1}^{N_B} 1\{d^{*,k} \geq d\} + 1}{N_B + 1}, \]

where \( 1\{\cdot\} \) is the indicator function.

With all details of this slightly more robust test presented, we continue to the last performance test to be utilized in the study.

### 4.3.3 Portfolio Turnover

This final test is the so-called portfolio turnover measure. The turnover, \( \Lambda \), for a strategy is a way to estimate how much on average, at every time step, of the portfolio value that is reallocated. This can thus give us an indication of how costly a particular strategy may be. That is, more reallocations mean an increased number of transactions and these may be costly.

The turnover is defined in DeMiguel et al. (2009) as

\[ \Lambda = \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{n} (|x_{i,t+1} - x_{i,t+1}|). \]

Here \( x_{i,t} = \frac{w_{i,t}}{V_t} \) denotes the fractional weight in asset \( i \) instead of its monetary counterpart \( w_{i,t} \) at time \( t \). Time \( t+1 \) should be understood as time \( t + 1 \), but where we still maintain the optimal asset allocations as of time \( t \). Since each asset price may have changed till the new period, these weights are although
likely to differ from that of time $t$.

### 4.4 Data

With every strategy and performance test described, we will now have a look at the data that was chosen for the study. In contrast to similar studies by DeMiguel et al. (2009) and Frahm et al. (2012), only one portfolio comprised of merely twelve stocks was used and not, for example, stock indices. One might argue that this approach may only be valid for this particular portfolio. To at least account for this fact, the choices of portfolio constituents were thus somewhat of greater importance. Hence, the adopted criteria was that, even though only U.S. stocks are included, the portfolio ought to be prudently diversified and include companies with large stock market capitalizations.\(^5\)

Thus will larger market effects be captured and thereby mimic a stock index, at least to some extent.

In Table 4.1, we see the included stocks, tickers, which exchange they are listed on and their respective sectors.\(^6\) The stock data consists of daily

<table>
<thead>
<tr>
<th>Stock</th>
<th>Ticker</th>
<th>Exchange</th>
<th>Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple Inc.</td>
<td>AAPL</td>
<td>Nasdaq</td>
<td>IT</td>
</tr>
<tr>
<td>Intel Corporation</td>
<td>INTC</td>
<td>Nasdaq</td>
<td>IT</td>
</tr>
<tr>
<td>Amazon.com Inc.</td>
<td>AMZN</td>
<td>Nasdaq</td>
<td>Consumer Discretionary</td>
</tr>
<tr>
<td>AT &amp; T Inc.</td>
<td>T</td>
<td>Nasdaq</td>
<td>Telecom Services</td>
</tr>
<tr>
<td>Berkshire Hathaway Inc.</td>
<td>BRK-B</td>
<td>NYSE</td>
<td>Financials</td>
</tr>
<tr>
<td>JPMorgan Chase &amp; Co.</td>
<td>JPM</td>
<td>NYSE</td>
<td>Financials</td>
</tr>
<tr>
<td>Johnson &amp; Johnson</td>
<td>JNJ</td>
<td>NYSE</td>
<td>Health Care</td>
</tr>
<tr>
<td>Walmart Inc.</td>
<td>WMT</td>
<td>NYSE</td>
<td>Consumer Staples</td>
</tr>
<tr>
<td>Exxon Mobil Corporation</td>
<td>XOM</td>
<td>NYSE</td>
<td>Energy</td>
</tr>
<tr>
<td>Boeing</td>
<td>BA</td>
<td>NYSE</td>
<td>Industrials</td>
</tr>
<tr>
<td>DowDuPont Inc.</td>
<td>DWDP</td>
<td>NYSE</td>
<td>Materials</td>
</tr>
<tr>
<td>The Home Depot Inc.</td>
<td>HD</td>
<td>NYSE</td>
<td>Consumer Discretionary</td>
</tr>
</tbody>
</table>

\(^5\)Another criteria was of course that stock data had to be available in the given investment horizon that was chosen.

\(^6\)Stock tickers are included to reduce plot labels seen in Appendix B.
adjusted closing prices in the range 1998-01-08 to 2014-12-31, all of which were retrieved from Yahoo! Finance. The evaluation window although starts at 2000-01-03 such that it lasts 15 years, or equivalently $T = 188$ months in total. The stock prices before this date, or exactly 500 trading days with corresponding stock prices, are needed for the bootstrapping and EWMA procedures as remarked. The length of $W_H = 500$ days was chosen in order capture a fair amount of varying market movements other than just those close to a rebalancing point and thus broadening potential outcomes of simulated returns. Equally important however, it could not be too large either such that obsolete returns that are not resembling actual market conditions would affect estimations.\footnote{Similar choices of lengths for historical simulations are made in Daníelsson (2011).}

Moreover, we need a risk-free interest rate in order to calculate the Sharpe Ratios during the evaluation window. For this purpose, the U.S. 1-Month Treasury Constant Maturity Rate (T-Bill) was primarily used since it is virtually risk-free (Hull, 2009) and its rates were downloaded from FRED\textsuperscript{E} Economic Data. What was meant by primarily was that data only existed from 2001-07-31 and onwards. To accommodate for the lack of data in between this date and 2000-01-03, the Swedish 1-Month T-Bill rates were used in this interval instead, acquired from the central bank of Sweden, Riksbanken.\footnote{Using two different risk-free rates should not matter substantially. It is after all the performance of the methods that are important and all strategies are treated under the same conditions, i.e. subtracted by the same risk-free rates.}

\subsection*{4.5 Parameter Setting}

Before we evaluate the results, we briefly recap mentioned parameter values that were used and present the values of those that were not.

The SM and the BH (with S&P 500 index hedging) strategies applying the EWMA scheme used a rolling-window with $W_H = 25$ historic monthly returns (corresponding to 500 daily returns) for calibration. For monthly returns, the optimal choice of the decay factor is $\lambda = 0.97$ as stated in RiskMetrics\textsuperscript{TM} (1996) and was set accordingly.

The bootstrap scenarios based on historic data adopted in the remaining
non-trivial strategies also used this rolling-window, from which 20 random
draws of daily returns were made to form a monthly counterpart. This was
repeated \( N = 2500 \) times at every time step, deemed sufficient to guarantee
reliable results.\(^9\) While we are on the subject of strategies, we will note
that Expected Shortfall applies the confidence level \( \alpha = 0.975 \), matching
the requirement set for this measure applied in market risk as stated by
Basel Committee on Banking Supervision (2017). Further, the two LPM
strategies both used the benchmark \( \tau = V_t \), meaning that every portfolio
value \( V_{t+1} < V_t \) are considered a loss.

An initial portfolio value for all strategies was also needed and it was
set to \( V_0 = 1 \) to ease subsequent analyses. Then all non-trivial strategies
were set to have a monthly gross target portfolio rate of return of \( \mu_0 = 1.01 \)
corresponding to a yearly net return of about 12.7\%, a realistic and a fair aim
in order to avoid too risky exposures one might incur otherwise. In Appendix
A, results are also provided for the more ambitious target \( \mu_0 = 1.015 \), equal
to a yearly net rate of approximately 19.6\%.

The last thing to comment before we analyze the primary results of
the study is the bootstrapping hypothesis test. In this test we are in need
of a number of bootstrap samples and Ledoit and Wolf (2008) argue that
\( N_B = 5000 \) is sufficient, hence this was the number of choice. So, with this
along with all other parameters defined, we will now move on to see what
results that were obtained.

\(^9\)A larger number would entail much more computational efforts and thus longer computational times.
Chapter 5

Results

With the approach and all theory presented, we will now evaluate the results of the study. Before we do so, there are a few important remarks. First, every non-trivial strategy are plotted versus the two trivial strategies separately. This is done in order to get less messy plots and is also the reason why we only look at portfolio values every sixth months, still enough to see each strategy’s key features. Second, the strategy performances are summarized in tables and these are naturally based on all \( T = 188 \) monthly portfolio values. Third and the last point, we will see a second table with \( p \)-values resulting from comparing all non-trivial with the two trivial strategies’ Sharpe ratios.

5.1 Performance Plots

In this section, we will see plots depicting the portfolio value every sixth month in the evaluation window resulting from applying a particular investment strategy. In this way, we will get a visual feeling of how the various strategies perform in reference to their common initial portfolio value \( V_0 = 1 \).

However, the plot of the LPM\(_1\) strategy is omitted. Since this strategy’s and LPM\(_2\)’s resulting curves nearly overlapped each other, there was no reason to include it. Yet no important information is lost by doing so since its performance is anyway summarized in Table 5.1 and 5.2.

In the figures, we clearly see that the passive investment in the S&P 500
index is always outperformed by all other methods and the resulting portfolio value is just above the initial value of $V_0 = 1$ at the last month. Moreover, we can also see that all non-trivial strategies, except MAD, seemingly perform better than the naïve method, however not substantially. Two other things to point out is that even though the naïve method give rise to fairly good portfolio values and usually follow the trends of the non-trivial strategies, it is not able to match the sudden growth boom these display around 2001. The second point is that the effect of the financial crisis that started in 2008 clearly affected all strategies. Although there are rather heavy losses at this time, all strategies but MAD recover quite rapidly from this.

These comments summarizes the most significant features we observe in the plots. Hence, we will now carry on to analyze the performances measured by using the previously described tools.
Figure 5.2: Performance of the SM (bootstrap) in comparison to the trivial strategies. Note, portfolio values are only plotted every sixth month.
Figure 5.3: Performance of BH with S&P 500 hedge (EWMA) in comparison to the trivial strategies. Note, portfolio values are only plotted every sixth month.
Figure 5.4: Performance of the ES in comparison to the trivial strategies. Note, portfolio values are only plotted every sixth month.
Figure 5.5: Performance of the MAD in comparison to the trivial strategies. Note, portfolio values are only plotted every sixth month.
Figure 5.6: Performance of the LPM$_2$ in comparison to the trivial strategies. 
Note, portfolio values are only plotted every sixth month.
5.2 Performance Statistics

So, if we begin by analyzing the portfolio turnovers summarized in Table 5.1, we see that most of the non-trivial strategies have similar values lying in the range of 3 – 4%. This means that this amount of the portfolio value is reallocated each month on average. Another thing to point out is that the naïve strategy requires on average 42% more reallocations than the other non-trivial methods (MAD excluded).

<table>
<thead>
<tr>
<th>Portfolio Strategy</th>
<th>Value $V_T$</th>
<th>Net Return $\mu_{Net}$</th>
<th>Sharpe Ratio $\Gamma$</th>
<th>Turnover $\Lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naïve</td>
<td>4.86</td>
<td>0.0097</td>
<td>0.172</td>
<td>0.0511</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1.38</td>
<td>0.0028</td>
<td>0.036</td>
<td>-</td>
</tr>
<tr>
<td>SM (EWMA)</td>
<td>5.04</td>
<td>0.0096</td>
<td>0.191</td>
<td>0.0330</td>
</tr>
<tr>
<td>SM (Bootstrap)</td>
<td>5.34</td>
<td>0.0098</td>
<td>0.209</td>
<td>0.0376</td>
</tr>
<tr>
<td>BH (S&amp;P 500)</td>
<td>4.78</td>
<td>0.0094</td>
<td>0.181</td>
<td>0.0308</td>
</tr>
<tr>
<td>ES97.5%</td>
<td>6.29</td>
<td>0.0107</td>
<td>0.225</td>
<td>0.0371</td>
</tr>
<tr>
<td>MAD</td>
<td>2.35</td>
<td>0.0061</td>
<td>0.090</td>
<td>0.0203</td>
</tr>
<tr>
<td>LPM1</td>
<td>5.42</td>
<td>0.0099</td>
<td>0.211</td>
<td>0.0376</td>
</tr>
<tr>
<td>LPM2</td>
<td>5.48</td>
<td>0.0100</td>
<td>0.209</td>
<td>0.0392</td>
</tr>
</tbody>
</table>

*Note: All values in the table are for the last month in the evaluation window, $W_E$, consisting of $T = 188$ months, corresponding to the date 2014-12-12.*

Therefore, if transactions are expensive, we can expect that the naïve strategy will be subject to higher operational costs. MAD would on the other hand be the least expensive alternative, but a plot of the allocation weight evolution seen in Appendix B reveals that this can be explained by the fact that most of the weights are set to zero. That is, we only allocate in few stocks.

Furthermore, we see that the ES97.5% strategy got the highest portfolio value and net return over the 15 years. Another interesting thing to point out is that this along with all other non-trivial strategies, except MAD, actually fulfilled the portfolio target rate of $\mu_0 = 1.01$. Peculiarly, the naïve method did result in a similar rate of return.

The best way to judge performance is however by looking at the Sharpe ratios since these can tell us how much return we gain in relation to the risk.
we take. Here, ES\textsubscript{97.5\%} performs best once again, but the other methods are not much worse since all lie close to $\Gamma = 0.2$. We also see that the naïve strategy performs nearly the same, whereas the passive S&P 500 investment and MAD method perform poorly, yielding low values.

Regarding Sharpe ratios, we will now look at results from the hypothesis tests in order assess whether or not non-trivial strategies really are superior to their trivial counterparts on statistical grounds, these are summarized in Table 5.2.

Table 5.2: Test statistics based on target $\mu_0 = 1.01$. Sharpe ratios of trivial strategies are $\Gamma_{\text{Naïve}} = 0.172$ and $\Gamma_{\text{S&P 500}} = 0.0360$. $p$-values are ordered as pairs; Jobson and Korkie (1981) version at the top and the Ledoit and Wolf (2008) bootstrap version beneath.

<table>
<thead>
<tr>
<th>Portfolio Strategy</th>
<th>Sharpe Ratio $\bar{\Gamma}$</th>
<th>Naïve p-values</th>
<th>S&amp;P 500 p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM (EWMA)</td>
<td>0.191</td>
<td>0.719</td>
<td>0.0067</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.743</td>
<td>0.0088</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.402</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.436</td>
<td>0</td>
</tr>
<tr>
<td>SM (Bootstrap)</td>
<td>0.209</td>
<td>0.874</td>
<td>0.0156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.885</td>
<td>0.0186</td>
</tr>
<tr>
<td>BH (S&amp;P 500)</td>
<td>0.181</td>
<td>0.254</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.287</td>
<td>0</td>
</tr>
<tr>
<td>ES\textsubscript{97.5%}</td>
<td>0.225</td>
<td>0.162</td>
<td>0.322</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.190</td>
<td>0.347</td>
</tr>
<tr>
<td>MAD</td>
<td>0.090</td>
<td>0.377</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.412</td>
<td>0</td>
</tr>
<tr>
<td>LPM\textsubscript{1}</td>
<td>0.211</td>
<td>0.373</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.409</td>
<td>0</td>
</tr>
<tr>
<td>LPM\textsubscript{2}</td>
<td>0.209</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The test is based on the null hypothesis, $H_0 = \Gamma_i - \Gamma_j = 0$, that is, the Sharpe ratios of two strategies do not differ. We test all non-trivial strategies versus the two trivial (Naïve and S&P 500). Furthermore, $p$-values less than $p = 0.001$ are set to zero.

In the above table, we see the $p$-values resulting from testing a trivial and a non-trivial strategy against each other one at a time. For each comparison, we see two $p$-values, where the top and bottom correspond to the methods of Jobson and Korkie (1981) and Ledoit and Wolf (2008) respectively.

What we can see is that none of the tests involving the naïve strategy
are significant at any standard significance level (e.g. at the 1%, 5% or 10% confidence level) regardlessly of which of two tests we apply. The S&P 500 index investment is however statistically distinguishable from all strategies but MAD. This is in accordance with the plots and the poor Sharpe ratios we saw earlier. At last, the EWMA SM strategy is merely significant at the 1% confidence level and BH is significant at the 5% confidence level.
Chapter 6

Conclusion

We have now seen the main results of this study and these indicate that a trivial naïve strategy can in fact perform almost as good as a non-trivial method. For instance, we see in the plots above that the portfolio value follows a similar path over time as all other non-trivial strategies except MAD, only that the portfolio value is somewhat lower, as is its Sharpe ratio. Compared to similar studies, the Sharpe ratio of the naïve method in this study is lower than the optimization strategies in all but one cases, where instead Frahm et al. (2012) find mixed results and DeMiguel et al. (2009) obtain mostly superior performance of the naïve method regarding this quantity. On the other hand, they use other portfolio constituents, types of strategies and do not consider specific portfolio target returns.

However, we cannot reject the null hypotheses that all these Sharpe ratios are indistinguishable from each other and this is in accordance with the findings of both DeMiguel et al. (2009) and Frahm et al. (2012). The result of this can hardly be explained by that the optimization strategies perform poorly, the monthly target requirement is in fact fulfilled in most cases. Rather, the naïve strategy performs remarkably good, probably because it is independent of any estimation errors as well as always promoting a careful diversification that happens to entail a relatively high rate of return.

Moreover, we also see that all non-trivial strategies along with the naïve method outperform the S&P 500 investment substantially, thus giving us an indication of that a passive investment might not be optimal. Worth to bear
in mind though is that in the early days of 2000, the dot-com bubble prevailed and from which this index recovered slowly from. Logically, the financial crisis around 2008 did not help it to get back on track. Thus, another choice of evaluation window might have revealed other findings. However, a remark is that this is only one possible choice of passive investment and there are certainly index funds (of other compositions than the S&P 500 index) that would have performed significantly better than this, yet the question is open whether some of these really have the capability of reaching or even exceeding the portfolio value levels associated with the best performing strategies seen here.

Returning to the discussion about target returns, in Appendix A and the results for the more aggressive target $\mu_0 = 1.015$, this return is by contrast not matched by the optimization strategies, so a return target fulfillment is not always guaranteed.\footnote{There is also a performance limitation of a portfolio, we can of course not outperform the best performing constituting asset. More aggressive targets will likely also lead to a larger allocation in this asset and thereby higher risk exposures.} Hence, other choices of targets, investment horizons (evaluation windows), and risky assets etc. than those analyzed in this study might all exacerbate the optimization, possibly yielding poor risk forecast and in the end, flawed allocation vectors $\mathbf{\hat{w}}_t^\star$.

To reduce such errors and possibly even enhance the main results seen here, we will now elaborate on some suggestions for improvements of and further developments to this study. For example, one can improve covariance estimations from which errors possibly stem, as discussed in DeMiguel et al. (2009). In this case, there are theoretically better alternatives to apply than EWMA, such as OGARCH, DCC and BEKK models, that although require more implementation and computational efforts, particularly BEKK (Danielsson, 2011).\footnote{Abbreviations for Orthogonal Generalized Autoregressive Conditional Heteroskedasticity (OGARCH), Dynamic Conditional Correlation (DCC), Baba-Engle-Kraft-Kroner (BEKK).} Another fact is that the choice of estimator of the expected asset returns may be the greatest source of error (Chopra and Ziemba, 1993), something both DeMiguel et al. (2009) and Frahm et al. (2012) take into account. Without going into details, alternative estimators are discussed in these papers and briefly by Lindberg (2009), where he particularly...
Conclusion

mentions the one by Black and Litterman (1992) involving a mixture of subjective predictions along with the Capital Asset Pricing Model (CAPM) and the model by Ross (1976) regarding the Arbitrage Pricing Theory (APT). Here though, mean values are applied for convenience and may thus have induced seriously biased estimations, which of course is worth to point out.

Furthermore, there may have been an insufficient number of bootstrapped return scenarios, \( N \). However, a more likely source of error could be the historical bootstrapping procedure that assigns the same probability of being drawn to every return, in combination with the choice of the rolling window \( W_H \)’s size. Specifically, the return scenarios may not reflect current market conditions at the particular time \( t \) in the evaluation window and may therefore not be realistic (Daníelsson, 2011). Applying more sophisticated scenario generators such as e.g. Vector Autoregression (VAR), filtered historical simulation bootstrapping or Monte Carlo simulations could most likely assure refined results regardless of target returns or selected portfolio assets.

Another important remark is that we only considered one portfolio and the results we see here might only be valid within this framework. A study including several portfolios of various compositions and sizes, in similarity to DeMiguel et al. (2009), will indubitably entail results from which general conclusions can be made with greater credibility.

The last room for improvement lies in the subject of statistical tests. The tests here are not justified in general as discussed in Ledoit and Wolf (2008). The obtained \( p \)-values might thus not be completely trustworthy to make inference from. A more sound way would be to conduct both joint and single hypothesis tests based on time series bootstrapping in accordance with the approach in Frahm et al. (2012) to reduce the likelihood of rejection errors.

Despite the fact that the approach of this study could have been enhanced based on the ideas above, the time frame was limited to do so and worth to remember is that the methods used here are still motivated since they are to some extent applied in practice. For example historical simulation (Pérignon and Smith, 2010) and the hypothesis test in the paper by DeMiguel et al. (2009). The point is, the results here are still a contribution to the field and an opening for further developments. Additionally, given the applied tools
here, the strategies are tested on historic data and thus show us how they had performed in reality at the time and compared to each other, which is the most important fact.

So what can we then say about the portfolio optimization strategies, which performs best? Well, within the framework of this study, we cannot say that any of the non-trivial strategies are superior since they all perform nearly the same (MAD excluded). Furthermore, they can neither be said to perform better than the naïve strategy statistically speaking, in accordance to the referred papers. Yet, some strategies may imply somewhat better diversification such as ES, which we see in the allocation plots in Appendix B where there are no extreme single stock exposures at any time. Contrary, MAD implies more zero valued weights, precisely as mentioned in Chapter 3 and does hence not promote diversification substantially. This fact makes it highly risky and inappropriate for any use in practice, at least when the number of assets to invest in is small, like in this analysis.\textsuperscript{3} However, worth to point out is that this many zero-valued allocations was not expected before the strategy was implemented.

In conclusion, one can discuss whether any of the non-trivial strategies that actually perform good are worth to implement when they are indistinguishable from the naïve method. However, this also depends on the trading costs (turnover) the strategies entail and the naïve method is on average about 42\% more expensive, so if it these costs are not negligible, alternative strategies are perhaps worth to pursue. In particular, the Expected Shortfall strategy performs best in general regarding turnover, portfolio value and the more important Sharpe ratio when using both target returns analyzed (without regard to statistics). The fact of this and that Expected Shortfall possess sound theoretical properties, which likely is the reason for it now has become the risk measure banks must adopt for market risk (Basel Committee on Banking Supervision, 2017), probably makes it the most robust strategy to apply in portfolio management.

\textsuperscript{3}The real advantage of MAD, if at all, may only be apparent when we consider portfolios including several hundred or thousands of assets.
References


Appendix A

Supplementary Results

Here we will see the result obtained by using the more aggressive monthly target return of $\mu_0 = 1.015$. In Table A.1, the performances of all strategies are summarized.

**Table A.1:** *Strategy performance with monthly target $\mu_0 = 1.015$.*

<table>
<thead>
<tr>
<th>Portfolio Strategy</th>
<th>$V_T$</th>
<th>Net Return $\mu$</th>
<th>Sharpe Ratio $\Gamma$</th>
<th>Turnover $\Lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naïve</td>
<td>4.86</td>
<td>0.0097</td>
<td>0.172</td>
<td>0.0511</td>
</tr>
<tr>
<td>S&amp;O 500</td>
<td>1.38</td>
<td>0.0028</td>
<td>0.036</td>
<td>-</td>
</tr>
<tr>
<td>SM (EWMA)</td>
<td>6.79</td>
<td>0.0115</td>
<td>0.205</td>
<td>0.0345</td>
</tr>
<tr>
<td>SM (Bootstrap)</td>
<td>5.17</td>
<td>0.0098</td>
<td>0.191</td>
<td>0.0396</td>
</tr>
<tr>
<td>BH (S&amp;O 500)</td>
<td>6.71</td>
<td>0.0115</td>
<td>0.200</td>
<td>0.0329</td>
</tr>
<tr>
<td>ES$_{97.5%}$</td>
<td>6.63</td>
<td>0.0112</td>
<td>0.215</td>
<td>0.0391</td>
</tr>
<tr>
<td>MAD</td>
<td>1.75</td>
<td>0.0046</td>
<td>0.061</td>
<td>0.0205</td>
</tr>
<tr>
<td>LPM$_1$</td>
<td>5.33</td>
<td>0.0100</td>
<td>0.195</td>
<td>0.0396</td>
</tr>
<tr>
<td>LPM$_2$</td>
<td>4.94</td>
<td>0.0096</td>
<td>0.180</td>
<td>0.0412</td>
</tr>
</tbody>
</table>

*Note: All values in the table are for the last month in the evaluation window, $W_E$, consisting of $T = 188$ months, corresponding to the date 2014-12-12.*

Looking at the portfolio turnovers, there are no major differences since these are again of the same order of magnitude as previously when using $\mu_0 = 1.01$. Moreover, in contrast to the main results, the two variance strategies using EWMA got the highest portfolio values, although with ES in close proximity. The net returns on the other hand, are not at the level of the target return this time. At last, looking at the Sharpe ratios, these are again of the same order of magnitude. Additionally, even though the ES
Supplementary Results

strategy did not get the highest portfolio return and value, we can see it got
the highest Sharpe ratio.

The results of the statistical tests now follow in Table A.2.

**Table A.2:** Test statistics based on target $\mu_0 = 1.015$. Sharpe ratios of
trivial strategies are $\Gamma_{\text{Naive}} = 0.172$ and $\Gamma_{\text{S&P 500}} = 0.0360$. $p$-values are
ordered as pairs; Jobson and Korkie (1981) version at the top and the Ledoit

<table>
<thead>
<tr>
<th>Portfolio Strategy</th>
<th>Sharpe Ratio</th>
<th>Naïve p-values</th>
<th>S&amp;P 500 p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM (EWMA)</td>
<td>0.205</td>
<td>0.542</td>
<td>0.0036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.536</td>
<td>0.0052</td>
</tr>
<tr>
<td>SM (Bootstrap)</td>
<td>0.191</td>
<td>0.685</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.710</td>
<td>0.0024</td>
</tr>
<tr>
<td>BH (S&amp;P 500)</td>
<td>0.200</td>
<td>0.626</td>
<td>0.0072</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.627</td>
<td>0.0062</td>
</tr>
<tr>
<td>ES$_{97.5%}$</td>
<td>0.215</td>
<td>0.356</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.383</td>
<td>0</td>
</tr>
<tr>
<td>MAD</td>
<td>0.061</td>
<td>0.061</td>
<td>0.643</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.069</td>
<td>0.646</td>
</tr>
<tr>
<td>LPM$_1$</td>
<td>0.195</td>
<td>0.629</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.658</td>
<td>0.0020</td>
</tr>
<tr>
<td>LPM$_2$</td>
<td>0.180</td>
<td>0.844</td>
<td>0.0028</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.852</td>
<td>0.0038</td>
</tr>
</tbody>
</table>

*Note: The test is based on the null hypothesis, $H_0 = \Gamma_i - \Gamma_j = 0$, that is, the Sharpe ratios of two strategies do not differ. We test all non-trivial strategies versus the two trivial (Naïve and S&P 500). Furthermore, $p$-values less than $p = 0.001$ are set to zero.*

Here we see that all non-trivial strategies but MAD are clearly not sign-
ificantly distinguishable from the naïve method. The test with MAD is
however significant at the 10% confidence level and is also near that of the
5% confidence level.

The tests versus the passive S&P 500 strategy yield higher $p$-values in
general this time. Yet, the tests are still significant at the 1% confidence
level, while MAD is once again indistinguishable from this trivial strategy.
In Table A.3 below, we see the monthly net returns of all stocks in the portfolio and the S&P 500 index calculated by using all 188 months in the evaluation window.

<table>
<thead>
<tr>
<th></th>
<th>AAPL</th>
<th>INTC</th>
<th>AMZN</th>
<th>T</th>
<th>BRK-B</th>
<th>JPM</th>
<th>JNJ</th>
<th>WMT</th>
<th>XOM</th>
<th>BA</th>
<th>DWDP</th>
<th>HD</th>
<th>S&amp;P 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net Return $\mu_{Net}$</td>
<td>0.0281</td>
<td>0.0059</td>
<td>0.0159</td>
<td>0.0038</td>
<td>0.0092</td>
<td>0.0081</td>
<td>0.0077</td>
<td>0.0040</td>
<td>0.0077</td>
<td>0.00110</td>
<td>0.0079</td>
<td>0.0071</td>
<td>0.0028</td>
</tr>
</tbody>
</table>

Worth to remark is that only two stocks have a rate of return exceeding the portfolio targets used this study. A very high target rate of return may thus imply larger risk exposures since we need to invest more in the stock that in fact can fulfill this. However, the values are just means of monthly stock returns over a long time and this fact does not necessarily imply large exposures during the entire investment horizon.
Appendix B

Portfolio Weights

In the following, we see plots of the assets’ respective fractional weights arising from pursuing a particular strategy. In essence, we will see how the stock allocations change over time due to rebalancing as a consequence of new optimal solutions $\vec{w}_t^*, \ t \in \{1, 2, ..., 188\}$. As previously though, weights are only shown in steps of six months to get less messy plots. Finally, we will also see the naïve strategy and even though the shape of this plot is trivial, it serves as a good reference.
Figure B.1: A representation of the asset allocation evolution through time. Here though, we only see the weights every sixth months.
Figure B.2: A representation of the asset allocation evolution through time. Here though, we only see the weights every sixth months.
Figure B.3: A representation of the asset allocation evolution through time. Here though, we only see the weights every sixth months.
Figure B.4: A representation of the asset allocation evolution through time. Here though, we only see the weights every sixth months.