UNIVERSITETET I AGDER

A Re-Examination of Performance of Optimized Portfolios

Erik Danielsen Nergaard
Andreas Lillehagen Bakke

SUPERVISOR
Valeriy Ivanovich Zakamulin

University of Agder 2017
Faculty of School of Business and Law at UiA

Abstract

DeMiguel, Garlappi, and Uppal (2009) conducted a study demonstrating that mean-variance optimized portfolios do not consistently outperform the naive diversification strategy in out-of-sample tests. This caused a heated debate and several studies claim to defend the value of mean-variance optimization. Kirby and Ostdiek (2012) developed two new methods of mean-variance portfolio optimization and demonstrated that these strategies show superior out-of-sample performance as compared to performance of the 1/N strategy. Several other papers demonstrated that the Global Minimum Variance portfolio outperforms the naive diversification. What all these papers have in common is that they measure the performance using the Sharpe ratio. Zakamulin (2017) argues that to display a convincing demonstration of the value of mean-variance optimization, one needs to show that the superior performance cannot be attributed to some known anomalies. In this thesis, we demonstrate that the strategies of Kirby and Ostdiek and the Global Minimum Variance strategy outperform the naive rule. We use several US datasets with an extended sample period and shorter estimation window. However, after accounting for three known anomalies, there is no longer any evidence of superior performance. Using similar data from the OSE, we also demonstrate that these strategies do not seem to work in Norway, .

Acknowledgements

We would especially like to thank our supervisor Valeriy Zakamulin for guidance and availability throughout the semester. He has given us clear guidelines, quick responses and constructive criticism whenever needed. We are also grateful for the support and the sharing of experience from fellow students, friends, and family.

Contents

1 **Introduction** 1

2 **Literature review** 5

3 **Methodology** 8

3.1 Constructing the portfolios 8

3.1.1 Portfolio mean return and variance 8

3.1.2 Global Minimum Variance 9

3.1.3 Volatility Timing 10

3.1.4 Reward-to-Risk Timing 11

3.1.5 Alternative estimators of conditional expected returns 12

3.2 Performance measures 13

3.2.1 Sharpe ratio 14

3.2.2 Portfolio alpha 14

3.3 Statistical estimations 15

3.3.1 Rolling estimators 15

3.3.2 Estimating conditional betas 16

3.3.3 The Carhart (1997) four-factor model 16

3.4 Statistical outperformance tests 19

3.4.1 Formulating hypotheses 19

3.4.2 Parametric tests 20

3.4.3 Non-parametric tests 21

4 **Data** 23

5 **Empirical results** 27

5.1 Industry datasets 27

5.2 BM datasets 28

5.3 Momentum datasets 29

5.4 Size datasets 30

6 Discussion **32**

7 Summary and Conclusion **34**

References **36**

Appendices **39**

Appendix A Reflection note of Andreas **39**

Appendix B Reflection note of Erik **41**

List of Tables

1 Descriptive statistics 24
2 Results for the Industry dataset 28
3 Results for the BM dataset 29
4 Results for the Momentum dataset 30
5 Results for the Size dataset 31

List of Figures

1 Reward and risk characteristics on the US dataset 25
2 Reward and risk characteristics on the Norwegian dataset 26

1 Introduction

The naive diversification strategy (also called the 1/N strategy) has been an asset allocation strategy for ages. The concept is simple enough: "do not put all your eggs in one basket". The strategy was first mentioned in a Babylonian Talmud around 1500 years ago when Rabbi Issac Bar Aha said: "one should always place his wealth, a third in land, a third in merchandise, and a third at hand". The performance of the naive strategy has in general been questioned since it was introduced due to its simplicity. Nevertheless, without any better strategies, there was no need to change what was already working. In modern times however, there has been written a lot of research literature on how to optimize portfolios using the return distribution parameters.

The concept and theory behind modern portfolio optimization goes back to the early 1950's when Harry Markowitz wrote his doctoral thesis at the University of Chicago. He introduced a model on how to optimize portfolios based on mean and variance using the efficient frontier. The efficient frontier is a tool describing the best possible return given the investor's risk-tolerance (Markowitz, 1952). Today his theory marks a cornerstone of modern portfolio theory, despite practical limitations associated with assumptions required for the model to fulfill its purpose. His theory did not cause an immediate reaction in the academic society, as it was filled with formulas and scribbles. Eventually however, other financial researchers continued building on his work, developing well known models, e.g. the Capital Asset Pricing Model (CAPM). Using modern portfolio theory, financial researchers challenged the naive strategy, and developed more advanced and complex models for portfolio optimization. In later years, modern portfolio theory has been widely taught at universities.

Recently, a number of respectable finance researchers have conducted studies with results questioning the performance of various optimization strategies as compared to the naive strategy. Though there is no question that Markowitz is theoretically correct, these studies argue that mean-variance optimized portfolios does not necessarily outperform the naive strategy with statistically significant margins. As a result, there is an ongoing debate about whether mean-variance optimization adds value and is worth the problems that occur when trying to implement these strategies.

One of the leading research papers on this topic is the study by DeMiguel et al. (2009). They evaluated the out-of-sample performance of some mean-variance optimized portfolios relative to the 1/N strategy. The optimized portfolios are constructed using the sample-based mean-variance strategy, and several extensions proposed in the literature that are designed to mitigate estimation errors. Among the 14 models of optimal asset allocation they examine, they show that typically none of them outperform the 1/N strategy for the seven empirical datasets. This has caused some doubts on whether portfolio optimization adds value, especially since the naive rule is easier to implement. To understand why the optimized portfolios perform poorly, they derived analytically the length of the estimation window needed to estimate the parameters used in the optimization strategies. While these parameters usually are estimated using 60 or 120 months of data, DeMiguel et al. found that for a portfolio of only 25 assets, the estimation window needed is 3000 months, and 6000 months for a portfolio of 50 assets. The extensions designed to mitigate the estimation errors only moderately reduced the needed estimation window. They concluded that there is a need for improvement when it comes to estimating the moments of asset returns. Investors should also use other available information about stock returns, not only statistical information. They also argued for the use of the naive diversification rule as a benchmark when evaluating performance of portfolios.

In defense of mean-variance optimization, other financial researchers claimed to have found evidence of superiority of optimized portfolios. This has resulted in a heated debate in the academic community. Kirby and Ostdiek (2012) conducted a study where they suggested that the study of DeMiguel et al. (2009) focus on portfolios that are exposed to high estimation risk and extreme turnover. To solve this problem they developed two new strategies for mean-variance portfolio optimization. These are distinguished by low turnover and outperform the 1/N strategy even in the presence of large transaction costs. Kritzman, Page, and Turkington (2010) argued that the minimum-variance and mean-variance strategies outperform equally weighted portfolios out-of-sample, and add value when using longer samples for the estimation of expected returns. To improve performance Tu and Zhou (2011) proposed an optimal combination of the naive rule with one of the four strategies: the Markowitz rule, the Jorion (1986) rule, the MacKinlay and Pástor (2000) rule, and the Kan and Zhou (2007) rule. They found that this method not only improves the performance of the four respective strategies, but also outperform the 1/N strategy in most cases.

The commonality of the studies defending the mean-variance optimization is that they used the Sharpe ratio as their performance measure, some without performing statistical tests. Zakamulin (2017) argued that researchers measure performance without further examining whether superior performance can be attributed to established risk factor premiums. Fama and French (1993) identified two such factors based on the Arbitrage Pricing Theory by Ross (1976), and introduced a three-factor model as an extension to the Capital Asset Pricing Model of Treynor (1961, 1962), Sharpe (1964), Lintner (1965), and Mossin (1966) containing the market risk factor. Later, Carhart (1997) included yet another risk factor designed to describe stock returns.

In this study we re-examine the performance of the optimized strategies by Kirby and Ostdiek (2012) by extending the historical period for the US data and testing the strategies using similar Norwegian data. We measure performance by means of the portfolio alpha derived from regression on a multi-factor model in addition to the Sharpe ratio. We also test the Global Minimum Variance portfolio, in similarity of the studies by Kritzman et al. (2010) and Clarke, De Silva, and Thorley (2011), though we do not impose the long-only constraint for this strategy. The purpose of testing the strategies on the Norwegian data is that we want to investigate if the strategies work in Norway. By using the portfolio alpha in addition to the Sharpe ratio, our goal is to account for the possibility that some optimized portfolios show superior performance as a result of profiting from some known market anomalies.

Our results on the US datasets show that all the optimized portfolios outperform the 1/N strategy according to the Sharpe ratio and associated p-values. On the Norwegian datasets however, we see less convincing performance, where none of the optimized portfolios consistently outperform the naive portfolio. When we measure the performance with the portfolio alpha, we see that the advantage of the timing strategies by Kirby and Ostdiek (2012) is reduced and the statistical significance disappears, even for the US data. On this note our results show that the superior performance can be attributed to exploiting known anomalies.

The rest of the thesis is organized as follows: Section 2 reviews the relevant literature to understand the background for our research and how it extends existing research. Section 3 describes the method used for the portfolio strategies, performance measures, and the statistical estimations and tests we use. Section 4 considers the data, data sources, and sample period for our analysis. In this section we also present the descriptive statistics. Section 5 presents and summarizes our empirical results relevant for our discussion. Section 6 discusses the results in compliance with previous literature and emphasizes how our results differ. Section 7 summarizes our thesis with the conclusion and final remarks.

2 Literature review

DeMiguel et al. (2009) issued a research paper questioning the actual performance of optimized portfolios. Their study had a great impact on the academic community and gave rise to the debate on whether mean-variance optimization outperform naive diversification as they were of the first to examine the subject matter. They evaluated the out-of-sample performance of 14 sample-based mean-variance models and compared to the 1/N strategy, using seven different empirical datasets in addition to simulated data. The conclusion of their study was that none of the compared portfolio optimization strategies conveyed results of statistically significant better performance than the naive strategy. Subsequently, many academic researchers have reassessed different mean-variance optimization methods and come up with their own empiric results showing strategies that yield better performances.

Kritzman et al. (2010) claimed that previous research has created an incorrect impression that naive asset allocation outperforms mean-variance optimized portfolios by attributing the sensitivity of optimization to estimation error. Using naive but plausible estimates of expected return, volatility and correlation, their results showed that the optimized portfolios perform better. One of these portfolios is the Global Minimum Variance portfolio. Kritzman et al. (2010) did not believe that the naive strategy is a viable option to optimal diversification. Clarke et al. (2011) examined the composition of the Global Minimum Variance portfolio with focus on the analytic form and the parameters of the individual asset weights. They derived an analytic solution for the long-only constrained Global Minimum Variance portfolio using the simplification associated with a single-factor model for the security variance-covariance matrix.

Kirby and Ostdiek (2012) suggested that DeMiguel et al. (2009) achieved their results because of their research design, focusing on models that are prone to high estimation risk and extreme turnover. Kirby and Ostdiek (2012) found that turnover in the presence of transaction cost removes the advantage of optimized portfolios. In order to address this issue, they developed two simple active portfolio strategies that utilize sample information about the expected returns and variances. These strategies keep the appealing features of the 1/N strategy by prohibiting short-sale, and dismissing optimization and variance-covariance-matrix inversion requirements. To control the turnover and transaction costs, they implemented a tuning parameter that can be interpreted as a measure of timing aggressiveness. They found that both strategies outperform the naive strategy by statistically significant margins, even in the presence of high transaction costs.

Disatnik and Katz (2012) introduced a portfolio strategy that prohibits short positions. This strategy prohibits short positions by investing in a Global Minimum Variance portfolio that is constructed using a block structure to calculate the variance-covariance matrix, and finding the weights analytically. This method avoids generating corner solutions with zero weights in many assets. Their diagonal variance-covariance matrix approach is the same approach as Kirby and Ostdiek (2012) utilized to develop their strategies. Disatnik and Katz (2012) also found that their basic portfolio optimization approach outperforms the 1/N strategy even in the presence of transaction costs. Behr, Guettler, and Miebs (2013) developed a constrained minimum-variance portfolio strategy on a shrinkage-theory based framework and demonstrated that this strategy displays lower out-of-sample variances compared to other mean-variance strategies and consistently returns a Sharpe ratio that is statistically different from that of the 1/N strategy. On the other hand, a study of Haley (2016) presented results that are consistent with those of DeMiguel et al. (2009) and argued additionally that the advantage of the naive strategy extends to individual stock selection and not just portfolios of stocks.

A consistent observation we make is that the performance is more or less measured using the Sharpe ratio. Zakamulin (2017) demonstrated that one can increase the Sharpe ratio by using known anomalies and so it is therefore essential to control whether the better performance can be attributed to mean-variance efficiency or some established risk factor premiums. He also argued that the long-only Global Minimum Variance strategy, Volatility Timing strategy, and Reward-to-Risk strategy exhibit superior performance due to tilting towards the asset with the lowest volatility. Motivated by these issues, we attempt to test the Volatility Timing strategy and the Reward-to-Risk strategy by

Kirby and Ostdiek (2012) for superior performance as compared to the 1/N strategy. In addition, we test the Global Minimum Variance strategy without imposing the long-only constraint. To expand the literature, we extend the sample size of Kirby and Ostdiek using similar datasets from the US. We also use the portfolio alpha as a performance measure in addition to the Sharpe ratio. Further, we consider four Norwegian datasets, sorted on similar criteria as the US datasets, to review the reliability of our results and examine if the strategies work in Norway.

3 Methodology

3.1 Constructing the portfolios

In this section we describe the portfolio construction strategies used in this study. We utilize four strategies where the first strategy is our benchmark strategy, naive diversification, also referred to as the 1/N strategy. This portfolio is constructed by allocating the weights of the assets equally. Throughout this study we assume that there are N risky assets and one risk-free asset. The comparing strategies we use are the Global Minimum Variance portfolio, and two strategies introduced by Kirby and Ostdiek (2012): the Volatility Timing strategy (VT) and the Reward-to-Risk Timing strategy (RRT), both designed to outperform the 1/N strategy.

3.1.1 Portfolio mean return and variance

Mean-variance optimization was introduced by Markowitz (1952). He described how to construct a portfolio with maximized expected returns given a specific level of risk. As a rule, each active dynamic portfolio is rebalanced periodically. Suppose in each period t the investor allocates a portion of his wealth ω_{it} in each asset i. In matrix notation, the mean return μ_{pt} and variance σ_{pt}^2 of the portfolio in period t are then given by:

$$\mu_{pt} = \omega_t' \mu_t \quad and \quad \sigma_{pt}^2 = \omega_t' \Sigma_t \omega_t, \tag{1}$$

where ω_t is the vector of weights, μ_t is the vector of mean returns, and Σ_t is the variance-covariance matrix of the assets in period t. Further, we assume that the asset returns are linearly independent and that the variance-covariance matrix is invertible. The variance covariance matrix is also symmetrical, because $\sigma_{ij} = \sigma_{ji}$, and since the variance is positive, the variance-covariance matrix is positive definite.

3.1.2 Global Minimum Variance

The construction of the Global Minimum Variance portfolio is thoroughly described by Merton (1972), and the literature suggests a two-step approach to estimate the weights of the Global Minimum Variance portfolio in an optimal way. First, estimate the distribution parameters of the stock returns, then minimize the portfolio variance under the assumption that the estimated parameters are true. These steps are repeated periodically. Consider a scenario where the risk-free asset is excluded. After the distribution parameters in section 3.1.1 are estimated, the objective is then to minimize the quadratic program:

$$\underset{\omega_t}{\underline{min}} \quad \frac{1}{2}\omega_t'\Sigma_t\omega_t \qquad \text{subject to} \qquad \omega_t'\mathbf{1} = 1, \tag{2}$$

where $\mathbf{1}$ denotes a $N \times 1$ vector of ones. Using the Lagrangian multipliers is a common way to find the local maxima and minima under equality constraints. Forming the Lagrangian we get the quadratic program:

$$\underset{\omega_t,\gamma}{\underline{min}} \quad L = \frac{1}{2}\omega_t'\Sigma_t\omega_t + \gamma(1 - \omega_t'\mathbf{1}), \tag{3}$$

where γ is the Langrangian multiplier. Solving this program[1] provides us with the vector of weights in period t:

$$\omega_{GMV,t} = \frac{\Sigma_t^{-1}\mathbf{1}}{\mathbf{1}'\Sigma_t^{-1}\mathbf{1}}. \tag{4}$$

To justify the Global Minimum Variance portfolio, consider a scenario where the risk-free asset is included and $\mu_{GMV,t} > r_f$. Then the investment opportunity set is tangent to the efficient frontier of risky asset. The Tangency portfolio weights in period t are given in the literature as:

$$\omega_{tan,t} = \frac{\Sigma^{-1}(\mu - \mathbf{1}r_f)}{\mathbf{1}'\Sigma^{-1}(\mu - \mathbf{1}r_f)}, \tag{5}$$

[1] See the derivation of this solution in Merton (1972)

where r_f is the return on the risk-free asset. If we then assume that the investor has no clue about the mean returns at all, the formula in equation (5) reduces to:

$$\omega_t = \frac{\Sigma_t^{-1}\mathbf{1}}{\mathbf{1}'\Sigma_t^{-1}\mathbf{1}},$$

which is the solution for the weights of the Global Minimum Variance portfolio. This is the closed-form solution for the vector of weights of the Global Minimum Variance portfolio, which means we assume that the market is frictionless and the assets can be bought and sold short without any limitations. We use this solution for comparison in our thesis. In practice however, this strategy is implemented with short-sale restrictions, in which case we can only attain the solution using numerical methods. The subsequent strategies are based on an approach using short-sale restrictions.

3.1.3 Volatility Timing

The VT strategy is an active portfolio strategy in which changes in the estimated variance $\hat{\sigma}^2$ causes rebalancing of weights in the portfolio. It is designed to avoid short sales and keep turnover as low as possible (Kirby & Ostdiek, 2012). Consider the solution for the Global Minimum Variance portfolio in equation 4. To eliminate short-positions entirely, we assume all correlations $\rho_{ij} = 0$ for each period t, so that the estimated variance-covariance matrix $\hat{\Sigma}$ becomes a diagonal matrix. Since the variance is positive, ignoring the correlations results in non-negative weights of assets. Using a diagonal variance-covariance matrix and assuming that the expected returns are equal for all periods ($\mu_t = \mu$), the weights for the Global Minimum Variance portfolio reduces to:

$$\hat{\omega}_{it} = \frac{\left(\frac{1}{\hat{\sigma}_{it}^2}\right)}{\sum_{i=1}^{N}\left(\frac{1}{\hat{\sigma}_{it}^2}\right)} \quad i = 1, 2, \ldots, N, \tag{6}$$

where $\hat{\sigma}_{it}$ is the estimated conditional volatility of the excess return on the ith risky asset.

Kirby and Ostdiek (2012) specified the portfolio weights in terms of conditional return volatility and a tuning parameter that allows some control over portfolio turnover and transaction costs. According to the authors, this tuning parameter η determines how aggressively we make changes in the portfolio weights as a result of changes in the volatility. In reality, increasing this tuning parameter tilts the weights towards the assets with the lowest volatility (Zakamulin, 2017). The weights for the VT(η) portfolio are given by:

$$\hat{\omega}_{it} = \frac{\left(\frac{1}{\hat{\sigma}_{it}^2}\right)^{\eta}}{\sum_{i=1}^{N}\left(\frac{1}{\hat{\sigma}_{it}^2}\right)^{\eta}} \quad i = 1, 2, \ldots, N, \tag{7}$$

where $\eta \geq 0$.

Regarding the tuning parameter. Because the correlations are set to zero, implementing an assumption where $\eta > 1$, should compensate for the lost information. This can be justified by η's effect on the formula above. When η approaches zero we will achieve the weight of the 1/N strategy portfolio and when η approaches infinity the weight of the asset with lowest volatility will approach one (Kirby & Ostdiek, 2012).

3.1.4 Reward-to-Risk Timing

The next strategy introduced by Kirby and Ostdiek (2012) is the Reward-to-Risk (RRT) strategy. Because the VT strategy above ignores information regarding conditional expected returns, one can ask if this information will influence its performance in a way. The RRT strategy is also built on modern portfolio theory and considers this information by adding the conditional expected return μ_{it}. Still considering a situation with the diagonal variance-covariance matrix, the weights of the Tangency portfolio can be expressed as:

$$\hat{\omega}_{it} = \frac{\left(\frac{\hat{\mu}_{it}}{\hat{\sigma}_{it}^2}\right)}{\sum_{i=1}^{N}\left(\frac{\hat{\mu}_{it}}{\hat{\sigma}_{it}^2}\right)} \quad i = 1, 2, \ldots, N, \tag{8}$$

where $\hat{\mu}_{it}$ is the estimated conditional expected excess return in period t for asset i.

Due to the difficulty of estimating expected return as precise as variances, the strategy in equation (8) is likely to involve noteworthy higher estimation risk than the VT strategy. Though setting the off-diagonal elements of $\hat{\Sigma}_t$ to zero reduce this risk, the possibility of extreme weights still remains. This is because negative $\hat{\mu}_{it}$ can cause the denominator of equation (8) to come close to zero. To address this problem, Kirby and Ostdiek (2012) assumes that the investor rejects any assets with $\hat{\mu}_{it} \leq 0$ in period t and express the calculation of weights as:

$$\hat{\omega}_{it} = \frac{\left(\frac{\hat{\mu}_{it}^{+}}{\hat{\sigma}_{it}^{2}}\right)}{\sum_{i=1}^{N}\left(\frac{\hat{\mu}_{it}^{+}}{\hat{\sigma}_{it}^{2}}\right)} \quad i = 1, 2, \ldots, N, \tag{9}$$

where $\hat{\mu}_{it}^{+} = \max(\hat{\mu}_{it}, 0)$ assures non-negative weights for all assets in period t.

Further on, we implement the parameter controlling turnover and construct a final formula for the weights of the RRT($\mu_{it}^{+}\eta$) strategy:

$$\hat{\omega}_{it} = \frac{\left(\frac{\hat{\mu}_{it}^{+}}{\hat{\sigma}_{it}^{2}}\right)^{\eta}}{\sum_{i=1}^{N}\left(\frac{\hat{\mu}_{it}^{+}}{\hat{\sigma}_{it}^{2}}\right)^{\eta}} \quad i = 1, 2, \ldots, N, \tag{10}$$

where $\eta \geq 0$.

3.1.5 Alternative estimators of conditional expected returns

To reduce estimation risk related to expected returns, Kirby and Ostdiek (2012) present another version of the RRT strategy, exploiting the relationship between the first and second moments of excess returns implied by numerous asset pricing models. Assume that a conditional version of the Capital Asset Pricing Model holds. This model implies that the cross-sectional variation in the conditional excess returns is due to cross-sectional variation in the conditional beta coefficients. We can then replace μ_{it}^{+} with the beta coefficient β_{it}^{+}, because the market risk premium μ_m is just a scaling factor that multiplies each of the conditional betas. The weights for the RRT portfolio can then be constructed as:

$$\omega_{it} = \frac{\left(\frac{\beta_{it}^{+}}{\sigma_{it}^{2}}\right)^{\eta}}{\sum_{i=1}^{N}\left(\frac{\beta_{it}^{+}}{\sigma_{it}^{2}}\right)^{\eta}} \quad i = 1, 2, \ldots, N, \tag{11}$$

where β_{it} is the CAPM beta coefficient and $\beta_{it}^{+} = \max(\beta_{it}, 0)$. Replacing μ_{it}^{+} with β_{it}^{+} can lower the sampling variations of the weights. Since $\beta_i = \rho_i \frac{\sigma_i}{\sigma_m}$, the formula in equation (11) can be reduced to:

$$\omega_{it} = \frac{\left(\frac{\rho_i^{+}}{\sigma_i}\right)^{\eta}}{\sum_{i=1}^{N}\left(\frac{\rho_i^{+}}{\sigma_i}\right)^{\eta}} \quad i = 1, 2, \ldots, N, \tag{12}$$

where $\rho_i^{+} = \max(\rho_i, 0)$ and ρ_i reflects the correlation between the excess return of the market and of asset i. Hence, we have replaced μ_i with $\sigma_i \rho_i$. Now, if we assume that the conditional CAPM does not hold, bias is introduced when replacing the estimator $\hat{\mu}_{it}$ with $\hat{\sigma}_{it}\hat{\rho}_{it}$. However, replacing an unbiased estimator characterized by high variance with a biased estimator characterized by low variance may still be beneficial.

Kirby and Ostdiek (2012) argue that this methodology can be extended to multi-factor models. Consider a K-factor model where $\beta_{ij,t}$ denotes the conditinal beta for the ith asset with respect to the jth factor in period t. As a result, the weights for the RRT($\bar{\beta}_t^{+}, \eta$) are calculated as:

$$\hat{\omega}_{it} = \frac{\left(\frac{\bar{\beta}_{it}^{+}}{\sigma_{it}^{2}}\right)^{\eta}}{\sum_{i=1}^{N}\left(\frac{\bar{\beta}_{it}^{+}}{\sigma_{it}^{2}}\right)^{\eta}} \quad i = 1, 2, \ldots, N, \tag{13}$$

where $\bar{\beta}_{it}^{+} = max(\bar{\beta}_{it}, 0$ and $\bar{\beta}_{it} = (1/K)\sum_{j=1}^{K} \beta_{ij,t}$ is the average conditional beta for asset i with respect to the K factors in period t.

The implementation of the beta coefficient is described in section 3.3.2.

3.2 Performance measures

We use two performance measures to evaluate the strategies described previously. The first measure we use is the industry standard, the Sharpe ratio. This measure is also used in many, if not all, of the other studies debating our topic. In addition, we choose to use the portfolio alpha as a performance measure. Zakamulin (2017) showed that using only the Sharpe ratio does not provide information on whether the performance gains can be attributed to exposures to certain risk factors. To accommodate this issue, we estimate the portfolio alphas using a four-factor model described in section 3.3.3.

3.2.1 Sharpe ratio

The Sharpe ratio is a widely used performance measurement within financial analysis. This measurement was developed and defined by Sharpe (1966) as a reward-to-variability ratio. It is used to calculate risk-adjusted return based on volatility. Generally, the greater value of the Sharpe ratio, the more attractive the strategy. The Sharpe ratio reflects the performance of the investment and is a suited measurement for evaluating which strategy performs best.

The Sharpe ratio is known for being a determinative factor when an investor decides on a portfolio to invest in. The formula for the Sharpe ratio can be presented as:

$$\text{Sharpe ratio} = \frac{\mu_p}{\sigma_p}, \tag{14}$$

where $\mu_p = E[r_p - r_f]$ denotes the expected excess return on the portfolio and σ_p denotes the volatility of the excess return.

3.2.2 Portfolio alpha

The portfolio alpha, or Jensen's alpha, is justified as a performance measurement by the CAPM (Jensen, 1968). It solely depends on two factors; expected return on the portfolio, and the beta (systematic risk). The portfolio alpha can be interpreted as the excess return on a portfolio predicted by an asset pricing model relative to the realized portfolio return:

$$\text{portfolio alpha} = \text{realized portfolio return} - \text{predicted portfolio return}.$$

The higher portfolio alpha, the better. It is a popular performance measure, because it is easy to estimate and test for statistical significance with OLS regression. To estimate the alphas one can either use the CAPM or a multi-factor model. Consider a K-factor model:

$$R_p = \alpha_p + \sum_{k=1}^{K} \beta_{p,k} F_k + \varepsilon_p, \tag{15}$$

where $R_p = r_p - r_f$ is the excess return on the portfolio, α_p is the portfolio alpha, $\beta_{p,k}$ is the kth factor loading or systematic risk, F_k is the return of factor k, and ε_p is the disturbance term. We use the Carhart (1997) four-factor model to estimate the portfolio alpha and take advantage of a more complex environment with several risk factors. We do this to get a better reflection of the market and a more precise result of what an investor actually can expect as return. Due to this extension, the systematic risk used to calculate alpha now appeal to all factors that are important for understanding the allocation of the fund.

3.3 Statistical estimations

3.3.1 Rolling estimators

To estimate μ_t and Σ_t for each portfolio's rebalancing date t, we use a fixed-window standard rolling estimation method similar that of Kirby and Ostdiek (2012) and DeMiguel et al. (2009). We use historical data from a window of length L to estimate the parameters for each period t until T, where T is the total number of observations in the out-of-sample period. Common choices of window length for monthly data are $L = 60$ and $L = 120$. Merton (1980) argues that it is necessary with long time series of returns to estimate expected returns. However, to be able to compare the Norwegian data to the US data, we set $L = 60$. This is due to our small sample size of the Norwegian data.[2] The estimators of μ_t and Σ_t follow the expressions:

$$\hat{\mu}_t = \frac{1}{L} \sum_{l=0}^{L-1} r_{t-l} \tag{16}$$

and

$$\hat{\Sigma}_t = \frac{1}{L} \sum_{l=0}^{L-1} (r_{t-l} - \hat{\mu}_t)(r_{t-l} - \hat{\mu}_t)', \tag{17}$$

[2]See section 4.

where $\hat{\mu}_t$ is the estimated conditional mean vector of the excess returns, r_{t-l} is the return of the risky asset, and $\hat{\Sigma}_t$ is the estimated conditional variance-covariance matrix of the excess returns on the risky assets in period t.

3.3.2 Estimating conditional betas

To implement the alternative Reward-to-Risk timing strategy we need a method for calculating the beta risk coefficient. Beta (β) is the undiversifiable risk coefficient associated with the factor return. In a multi-factor model there is one beta associated with each factor. Each of these are systematic risks and dependent on the associated factor. We estimate these betas using a multi-factor model. A generalized formula for a multi-factor model when we have N risky assets can be presented with the following approach:

$$R_{i,t} = \alpha_i + \sum_{k=1}^{K} \beta_{i,k} F_{k,t} + \varepsilon_{i,t} \quad i = 1, 2, \ldots, N, \tag{18}$$

where $R_{i,t}$ is the of excess return for asset i in period t, α_i is the models' pricing error, $\beta_{i,k}$ denotes the beta coefficient for the kth factor associated with the ith risky asset, $F_{k,t}$ is the return of the kth factor at time t and $\varepsilon_{i,t}$ is the time-t disturbance term. We use the Carhart (1997) four-factor model to estimate the betas.

3.3.3 The Carhart (1997) four-factor model

To understand the Carhart (1997) four-factor model we have to go back to the Capital Asset Pricing Model which was introduced by Treynor (1961, 1962), Sharpe (1964), Lintner (1965), and Mossin (1966). Independently they developed the CAPM, by building on the earlier work of Markowitz (1952) as a model for pricing an individual security or portfolio. Later on, Ross (1976) introduced the Arbitrage Pricing Theory, which he proposed as an alternative to the CAPM. The APT formed the foundation for multi-factor models, as it is based on weaker assumptions than the CAPM.

Several "anomalies" have been discovered within the CAPM. This means that there are portfolios of stocks with certain characteristics that have positive and statistically significant alphas in the CAPM. Some of these anomalies are the size anomaly, the value anomaly, and the momentum anomaly, which are related to small cap stocks, stocks with high book-to-market ratio, and selling losers and buying winners respectively. In a rational asset pricing model, higher risk premiums can only be due to higher risk. Therefore one assumes that the anomalies can be explained by specific risk factors. A multi-factor model with two such risk factors was introduced by Fama and French (1993). Their model account for the anomalies related to size and value using the SMB (Small Minus Big) and HML (High Minus Low) factors respectively. Fama and French argue that because these risk factors alone cannot explain the cross-section of average stock returns, the market factor is included in the equation for justification. By including these factors, the three-factor model can be presented as:

$$R_p = \alpha_p + \beta_{p,MKT} F_{MKT} + \beta_{p,SMB} F_{SMB} + \beta_{p,HML} F_{HML} + \varepsilon_p, \tag{19}$$

where R_p is the excess return on the portfolio, α_p represents the portfolio alpha, and F and β denoted MKT, SMB and HML are the factor premiums and risk coefficients for the market, size, and value factors respectively.

All the MKT-, SMB- and HML-factors are each calculated by the use of six value-weigted portfolios (Fama & French, 1992). Fama and French describe the construction of these factor as follows: The portfolios of SMB are constructed by sorting all the non-negative stocks of one index by size (price times shares) at time t and then divide it by its median. The measure of this index is then applied as measure to all indices for which the SMB-factor is constructed, by dividing them into two groups, small (S) and big (B). The SMB-factor is further calculated by subtracting the B (the average return of the three big portfolios) from S (the average return of the three small portfolios). The SMB-factor can be expressed as:

$$\begin{aligned} SMB =& 1/3\big(\text{Small Value} + \text{Low Neutral} + \text{Small Growth}\big) \\ &- 1/3\big(\text{Big Value} + \text{Big Neutral} + \text{Big Growth}\big). \end{aligned} \tag{20}$$

The market factor used is the excess market return $(r_m - r_f)$ calculated by sorting the stocks by size, but this time, include the negative stocks as well. The portfolios of the HML-factor are constructed by sorting the stocks of one index into three groups by their book-to-market value, the lowest 30%, the mid-range 40%, and the highest 30%. Then, similar to the SMB-factor, this measure (or a "breakpoint" (Fama & French, 1993)) is applied to all indices. Further the low 30%(average return of the lowest valued) is subtracted from the high 30% (average return of the highest valued) (Fama & French, 1993).[3] The HML-factor can be expressed as:

$$\begin{aligned} HML = &1/2\big(\text{Small Value} + \text{Big Value}\big) \\ &- 1/2\big(\text{Small Growth} + \text{Big Growth}\big). \end{aligned} \tag{21}$$

By implementing the four-factor model we add yet another factor to this model, the momentum factor (MOM). This factor was first introduced by Jegadeesh and Titman (1993) as a strategy. In their study, Jegadeesh and Titman analyzed the strength of trading strategies with a time horizon of three to twelve months. They include a strategy they refer to as the "$J - month/K - month$" strategy. This strategy selects which stocks to buy according to their returns over the the past J months and holds them for K months.

The "$J - month/K - month$" strategy is constructed by sorting all assets at time t in an ascending order by their returns in the past J months. Then, based on how the assets are sorted and how many securities there are, they are divided into ten equally weighted portfolios, as the weight of each asset equals $\frac{1}{K}$ in the top ten, the second top ten, and so on. The top ten is then defined as the "losers", and the bottom is defined as the "winners". At each t (start of month) the strategy sells the "loser"-portfolio and buys "winner"-portfolio. Jegadeesh and Titman (1993) referred to these as zero-cost portfolios, because the profit is calculated by subtracting the winners from the losers.

Four years later, Carhart (1997) proved that the "anomaly" regarding momentum can be almost completely explained by adding the "$J - month/K - month$"-strategy by

[3]For further reading regarding the calculation and choice of data for the risk-factors SMB and HML, see the article by Fama and French (1993).

Jegadeesh and Titman (1993) as a new risk coefficient. Carhart added the momentum factor to the three-factor model, developing it into the Fama-French-Carhart four-factor model (MOM), an improved multi-factor model for further examination of mutual fund performance. The construction of this momentum factor include the same indices as for the SMB- and HML-factors (Carhart, 1997), but as in the "$J-month/K-month$" strategy by Jegadeesh and Titman (1993) the stocks are now equally weighted. The momentum factor is calculated over a 11 month period lagged one month, and is reconstructed every month (Carhart, 1997). Further Carhart (1997) constructed it by subtracting the lowest average 30% of firms from the highest 30% firms. The formula for the momentum factor can be presented as:

$$\begin{aligned} MOM =& 1/2\big(\text{Small Winners} + \text{Big Winners}\big) \\ &- 1/2\big(\text{Small Losers} + \text{Big Losers}\big). \end{aligned} \tag{22}$$

The four-factor model can then be expressed as:

$$R_p = \alpha_p + \beta_{p,MKT} r_{MKT} + \beta_{p,SMB} r_{SMB} + \beta_{p,HML} r_{HML} \beta_{p,MOM} r_{MOM} + \varepsilon_p, \tag{23}$$

where MOM is the momentum factor. Both the three- and four-factor model is developed to make investments calculations based on a more complex environment than with the CAPM. By taking advantage of the four-factor model, we present a more exact regression based on the deviations of the factors.

3.4 Statistical outperformance tests

3.4.1 Formulating hypotheses

The question that remains after estimating the performance measures is whether the difference between the two measures are significantly different. Given that we want to test the optimized strategies against our benchmark strategy, we formulate our hypotheses as:

$$H_0 : SR_p \leq SR_{1/N} \qquad H_A : SR_p > SR_{1/N}$$

and

$$H_0 : \alpha_p \leq \alpha_{1/N} \qquad H_A : \alpha_p > \alpha_{1/N},$$

where SR_p and $SR_{1/N}$ are the Sharpe ratios of the optimized and 1/N strategies respectively, while α_p and $\alpha_{1/N}$ are the associated portfolio alphas.

There are two ways to test the hypotheses for performance measures, parametric tests and non-parametric tests. A parametric test require a number of assumptions and is a good fit for theoretical hypothesis. A non-parametric test is often a better fit for real-life scenarios due to the difficulty of complying all assumptions for a parametric test. The goal of the tests is to find statistically significant p-values. When we know the p-value, we can establish whether our null-hypothesis can be rejected or not. Common statistical significance levels are 1%, 5 %, and 10 %.

3.4.2 Parametric tests

A parametric test of the hypotheses in section 3.4.1 is based on the assumption that the two excess return series of each strategy follow a normal distribution and are correlated. It is called a parametric test because each random variable is assumed to have the same probability distribution that is parameterized by mean and standard deviation. The hypotheses are tested using a standardized value calculated from the sample data. This test statistic, which has a well known distribution and is simple to calculate, can be used to calculate the p-value. Each performance measure requires a specific test statistic. To implement a parametric test of the Sharpe ratio one can employ the Jobson and Korkie (1981) test, with the correction of Memmel (2003). The test statistic z is then given by:

$$z = \frac{SR_p - SR_{1/N}}{\sqrt{\frac{1}{T}\left[2\left(1-\hat{\rho}\right) + \frac{1}{2}\left(SR_p^2 + SR_{1/N}^2 - 2\hat{\rho}^2 SR_p SR_{1/N}\right)\right]}} \tag{24}$$

where $\hat{\rho}$ is the estimated correlation between the excess returns of the two compared strategies. This test assumes joint normality between the two excess return series and the test statistic is asymptotically distributed as a standard normal when the sample size is large. To implement a parametric test of the portfolio alpha one can use a two-sample t-test. The test statistic y for this test can be expressed as follows:

$$y = \frac{\alpha_p - \alpha_{1/N}}{\sqrt{se_p^2 + 2\hat{\rho} se_p se_{1/N} + se_{1/N}^2}} \tag{25}$$

where se_p and $se_{1/N}$ reflects the standard error of the estimation of alpha from the compared strategies respectively.

3.4.3 Non-parametric tests

Since the parametric tests do not control for time series characteristics in portfolio returns (e.g. autocorrelation, volatility clustering, and absence of normally distributed returns), we employ a block bootstrap approach to compute the p-values. The advantages of using this type of test are that we do not need to make any assumptions, the test provide accurate results even with smaller sample sizes, because it is distribution-free. We can choose the test statistic freely, and the implementation of the test is simple and similar regardless of which statistic we choose. Non-parametric tests, like the bootstrap, use computer-intensive randomization methods to estimate the distribution of the p-values. The bootstrapping method is the most popular non-parametric test that is based on resampling the original data with replacements. If $r_{1/N}$ and r_p represents two original excess returns, this method constructs two pseudo time-series with the same number of observations that retain the historical correlation.

The standard bootstrap was introduced by Efron (1979). This method assumes that the data are serially independent. We cannot use this method because it breaks up the dependency we have in our return data and creates serially independent resamples. To preserve our dependency structure we use blocks instead of individual observations. There are two types of block methods, with overlapping (Künsch, 1989) and non-overlapping blocks (Carlstein, 1986) for uni-variate time-series. Overlapping, also called moving block, is preferred when the sample size is small relative to the block length. Suppose we have a block length l, then the total number of overlapping blocks for a sample of T observations is $T - l + 1$. By construction, the moving block time-series have a non-stationary, or conditional, distribution. We can get a stationary distribution by making the block length random (Politis & Romano, 1994). The length of the blocks are generated from a geometric distribution with probability p. The p is then chosen so that $p = \frac{1}{l}$ where l is the required average block length. The choice of average block length depends on context. A study by Hall, Horowitz, and Jing (1995) express the asymptotic formula for the optimal block length as: $l \sim T^{\frac{1}{h}}$, where $h = 3, 4,$ or 5, depending on what kind of test you are conducting. For one-sided test we use $h = 4$ so that our optimal block length for $M = 10000$ becomes $10000^{\frac{1}{4}} = 10$. The stationary method wraps the data around in a circle so that 1 follows T and so on. The moving block bootstrapping method consists of drawing M resamples of $t^b = \{B_1^b, B_1^b, \ldots, B_m^b\}$ where each block of time indices B_i^b is drawn randomly with replacement from a available blocks $B_1, B_2, \ldots, B_{T-l+1}$. After, the pseudo time-series of the excess returns $r_{1/N}$ and r_p are created by using each resample t_b.

To compute the p-values for the Sharpe ratios, we calculate the difference Δ between the Sharpe ratios for each pseudo time-series and count how many times m the compared strategy does not outperform 1/N strategy. Then we divide this number by the total number of bootstrap resamples M so that $p\text{-}value = \frac{m}{M}$. When the p-value is lower than or equal to a statistical significance level, we can reject the null-hypothesis for that level. We use the same method to calculate the p-values for the portfolio alphas.

4 Data

Our data for the empirical analysis consists of monthly excess returns on broadly based US and Norwegian stock portfolios sorted on industry (Ind), book-to-market ratio (BM), momentum (Mom) and size (Size). They contain value-weighted returns, which indicate that the assets are weighted according to their total market capitalization. The larger assets carry heavier weights while the smaller assets carry lower weights. This means that price changes in the larger assets will have greater effects on the value of the portfolio. The US portfolios contain stocks from NYSE, NASDAQ and Amex, while the Norwegian portfolios contain stocks from the OSE. The US risk-free rates are the one-month Treasury bill rates from Ibbotson Associates and the Norwegian estimate of the risk-free rates are from the OSE data service and Datastream. The data are available from Kenneth R. French's library[4] for the US and Bernt A. Ødegaard's library for Norway.[5] The data libraries are also the source of the factor returns used to estimate the the beta coefficients and alphas for the four-factor model. The risk factors are the excess return on the market ($r_m - r_f$) and the return on three portfolios that are constructed to mimic abnormal excess returns between small and large capitalization stocks (SML), high and low book-to-market equity stocks (HML) and a momentum factor (MOM).

We extend the study of Kirby and Ostdiek (2012) by using a sample period of July 1963 - December 2016, where $T + L = 642$ monthly observations and $L = 60$. Due to lack of available Norwegian data, the sample period for the Norwegian stocks is somewhat shorter. It starts July 1996 and ends December 2016, where $T + L = 246$ monthly observations and $L = 60$. We consider eight datasets in total, four US and four Norwegian, with similar characteristics to be able to compare them. All the datasets consists of ten portfolios. Table 1 shows the descriptive statistics for the out-of-sample period for all data sets used. Figure 1 and 2 describe the cross section of annualized return and standard deviation for the US and Norwegian datasets respectively.

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

[5] http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html

Table 1: Descriptive statistics

	US				Norwegian			
Dataset	**Mean**	**Volatility**	**Min**	**Max**	**Mean**	**Volatility**	**Min**	**Max**
Ind	10.72	18.60	-23.6	20.22	17.43	31.12	-44.72	148.96
BM	12.19	17.05	-25.84	33.58	20.06	26.74	-39.26	38.15
Mom	10.53	19.23	-26.74	45.67	20.15	26.09	-41.17	39.76
Size	12.28	20.10	-30.30	32.95	28.28	20.89	-24.05	43.53

Table 1: This table reports the annualized descriptive statistics for the US and Norwegian out-of-sample data. The period covers July 1968 - December 2016 for the US data and July 2001 - December 2016 for the Norwegian data with value-weighted portfolio returns.

Figure 1: Reward and risk characteristics on the US dataset

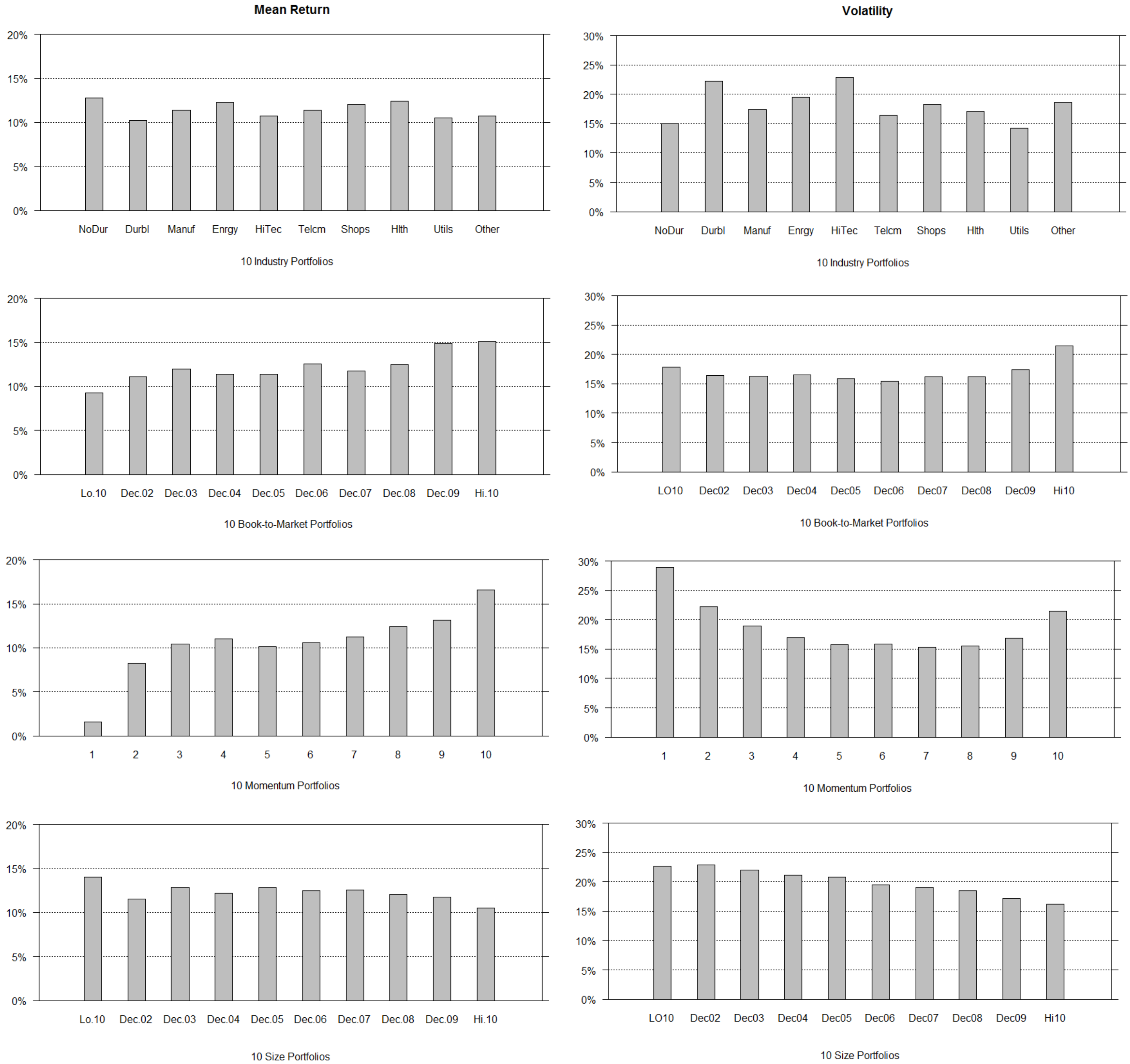

Figure 1: This figure summarizes mean and volatility for each portfolio in the dataset. The graphs on the left shows the cross section of annualized mean returns and the graphs on the right shows the cross section of annualized standard deviations. The reported statistics is associated with the out-of-sample sub-period (observations 61-642).

Figure 2: Reward and risk characteristics on the Norwegian dataset

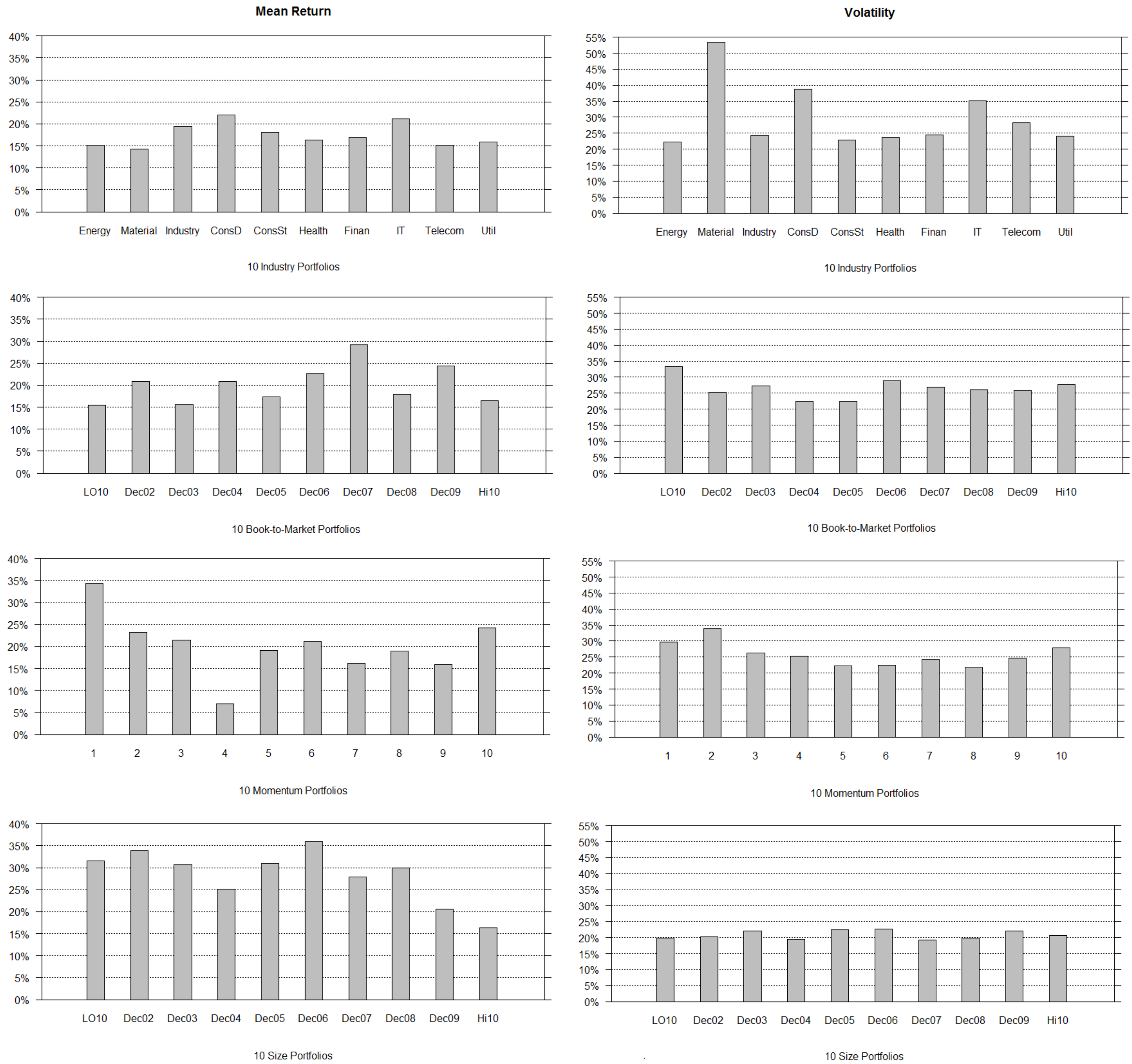

Figure 2: This figure summarizes mean and volatility for each portfolio in the dataset. The graphs on the left shows the cross section of annualized mean returns and the graphs on the right shows the cross section of annualized standard deviations. The reported statistics is associated with the out-of-sample sub-period (observations 61-642).

5 Empirical results

In this section we present the results of our analysis conducted using the methods previously described. We evaluate the out-of-sample performance of each strategy from July 1968 to December 2016 for the US data and from July 2001 to December 2016 for the Norwegian data, with an estimation window length of 60 months. The tables report the estimations for the annualized means ($\hat{\mu}$), annualized volatility ($\hat{\sigma}$), annualized Sharpe ratios ($\hat{SR}$) and the p-values for our hypothesis tests for the Sharpe ratio and the portfolio alpha respectively. Like Kirby and Ostdiek, we set $\eta = 1$ for the baseline analysis delivering VT and RRT strategies similar to basic mean-variance optimization using the diagonal variance-covariance matrix. We set $\eta = 2$ and $\eta = 4$ to mitigate information loss associated with ignoring the estimated return correlations and to be able to compare our results with theirs. We also report the estimated values for the Global Minimum Variance portfolio strategy, to document how the timing strategies perform compared to a basic mean-variance optimization strategy.

5.1 Industry datasets

Our analysis starts with the datasets consisting of portfolios sorted on industry. Table 2 reports the performance of each strategy for the US and the Norwegian datasets. On row one we find for each dataset the performance results of using the 1/N strategy. The means are 6.67% and 14.38%, while the volatility values are 14.84% and 21.20% respectively. This translates into Sharpe ratios of 0.45 and 0.68. The reported alphas are 0.47 and -0.13. In the panels for the optimized strategies the US Sharpe ratios are larger than that of the 1/N strategy for the VT, RRT($\bar{\beta}_t^+, \eta$) and GMV and varies from 0.49 to 0.56. The associated p-values are statistically significant at the 5% level for the VT(1), VT(2), and RRT($\bar{\beta}_t^+, \eta$) strategies. The p-value for the GMV strategy is statistically significant at the 10% level. The Norwegian Sharpe ratios are generally lower except for the RRT($\bar{\beta}_t^+, 1$) strategy and ranges from 0.57 to 0.78. None of the higher Sharpe ratios have p-values that are statistically significant. The alphas for the timing strategies on the US dataset

Table 2: Results for the Industry datasets

	US						Norwegian					
Strategy	$\hat{\mu}$	$\hat{\sigma}$	$\hat{SR}$	**pval**	$\hat{\alpha}$	**pval**	$\hat{\mu}$	$\hat{\sigma}$	$\hat{SR}$	**pval**	$\hat{\alpha}$	**pval**
1/N	6.67	14.84	0.45	-	0.47	-	14.38	21.20	0.68	-	-0.13	-
Panel A												
VT(1)	6.81	13.87	0.49	**0.000**	0.48	0.202	12.87	19.77	0.65	0.613	-0.11	0.854
VT(2)	6.80	13.19	0.52	**0.003**	0.49	0.254	12.15	19.40	0.63	0.684	-0.18	0.857
VT(4)	6.54	12.60	0.52	0.033	0.49	0.360	11.19	19.46	0.57	0.774	-0.27	0.860
Panel B												
RRT($\hat{\mu}_t^+$, 1)	5.81	13.85	0.42	0.560	0.45	0.614	13.38	20.44	0.65	0.607	-0.19	0.886
RRT($\hat{\mu}_t^+$, 2)	5.86	13.85	0.42	0.531	0.46	0.518	12.75	20.33	0.63	0.689	-0.26	0.906
RRT($\hat{\mu}_t^+$, 4)	5.81	14.20	0.41	0.649	0.46	0.505	12.42	20.71	0.60	0.738	-0.28	0.850
RRT($\bar{\beta}_t^+$, 1)	7.03	14.34	0.49	**0.037**	0.46	0.643	14.86	20.91	0.71	0.258	-0.08	0.733
RRT($\bar{\beta}_t^+$, 2)	7.15	14.09	0.51	**0.028**	0.46	0.516	15.74	21.08	0.75	0.206	-0.05	0.598
RRT($\bar{\beta}_t^+$, 4)	7.22	13.89	0.52	**0.050**	0.47	0.454	16.86	21.51	0.78	0.185	-0.01	0.480
Panel C												
GMV	7.14	12.69	0.56	**0.010**	0.64	**0.063**	6.69	20.55	0.33	0.988	-0.53	0.938

Table 2: This table summarize our results for the performance of the 1/N strategy (row one), Volatility Timing strategies (Panel A), Reward-to-Risk Timing strategies (Panel B), and the Global Minimum Variance portfolio strategy (Panel C). We report the following sample statistics for the time series of monthly excess return generated by each strategy: annualized mean ($\hat{\mu}$), annualized volatility ($\hat{\sigma}$), annualized Sharpe ratio ($\hat{SR}$), portfolio alpha ($\hat{\alpha}$) and the associated p-values for our hypotheses from section 3.4.1.

are generally similar to the 1/N portfolio alpha and ranges from 0.45 to 0.49. None of the p-values are statistically significant. For the GMV strategy however, the portfolio alpha is 0.64 with an associated p-value that is statistically significant at the 10% level. For the Norwegian dataset, all the alphas are, like the 1/N alpha, negative. None of the p-values suggests that any of the optimized strategies perform better than the 1/N strategy.

5.2 BM datasets

The next datasets in our analysis are the ones with portfolios sorted on the book-to-market ratio. Table 3 presents the performance results of each strategy for the US and Norwegian datasets. The 1/N strategy on row one reports mean values of 7.41% and 17.01%. The volatility values are 15.75% and 20.91% and the Sharpe ratios are 0.47 and 0.81 respectively. The portfolio alphas are reported as 0.41 and 0.01. The Sharpe ratios for the timing strategies vary from 0.48 to 0.53 for the US dataset with associated p-values statistically significant p-values for the VT and RRT($\bar{\beta}_t^+, \eta$) strategies. The GMV portfolio does not outperform the 1/N strategy on this dataset. For the Norwegian dataset the Sharpe ratios vary from 0.66 to 0.82. Even though some of the values are higher than

Table 3: Results for the BM datasets

Strategy	US						Norwegian					
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{SR}$	pval	$\hat{\alpha}$	pval	$\hat{\mu}$	$\hat{\sigma}$	$\hat{SR}$	pval	$\hat{\alpha}$	pval
1/N	7.41	15.75	0.47	-	0.41	-	17.01	20.91	0.81	-	0.01	-
{Panel A}												
VT(1)	7.42	15.46	0.48	**0.026**	0.42	0.234	16.64	20.42	0.81	0.380	0.05	0.154
VT(2)	7.43	15.27	0.49	**0.028**	0.42	0.292	16.23	20.18	0.80	0.523	0.08	0.167
VT(4)	7.46	15.08	0.49	**0.035**	0.42	0.398	15.90	20.42	0.78	0.664	0.14	0.165
{Panel B}												
RRT($\hat{\mu}_t^+$, 1)	7.30	15.00	0.49	0.255	0.41	0.472	16.58	21.51	0.77	0.783	-0.04	0.703
RRT($\hat{\mu}_t^+$, 2)	7.26	14.99	0.48	0.274	0.40	0.545	16.09	21.81	0.74	0.854	-0.08	0.768
RRT($\hat{\mu}_t^+$, 4)	7.21	15.03	0.48	0.343	0.38	0.625	14.99	22.78	0.66	0.947	-0.15	0.805
RRT($\bar{\beta}_t^+$, 1)	7.92	15.70	0.50	**0.024**	0.40	0.852	17.19	20.90	0.82	0.405	0.04	0.346
RRT($\bar{\beta}_t^+$, 2)	8.18	15.71	0.52	**0.025**	0.39	0.852	17.38	21.07	0.82	0.435	0.09	0.254
RRT($\bar{\beta}_t^+$, 4)	8.43	15.84	0.53	**0.038**	0.38	0.858	16.97	21.49	0.79	0.619	0.12	0.290
{Panel C}												
GMV	6.68	15.89	0.42	0.742	0.39	0.537	17.56	22.10	0.79	0.565	0.61	**0.034**

Table 3: This table summarize our results for the performance of the 1/N strategy (row one), Volatility Timing strategies (Panel A), Reward-to-Risk Timing strategies (Panel B), and the Global Minimum Variance portfolio strategy (Panel C). We report the following sample statistics for the time series of monthly excess return generated by each strategy: annualized mean ($\hat{\mu}$), annualized volatility ($\hat{\sigma}$), annualized Sharpe ratio ($\hat{SR}$), portfolio alpha ($\hat{\alpha}$) and the associated p-values for our hypotheses from section 3.4.1.

that of the 1/N strategy, none of the associated p-values are statistically significant. The alphas for the timing strategies range from 0.38 to 0.42 for the US dataset and from -0.15 to 0.14 for the Norwegian dataset. There are no statiscially significant p-values for the timing strategies. The GMV alphas are 0.39 and 0.61, and the p-value for the alpha on the Norwegian dataset is statistically significant at the 5% level.

5.3 Momentum datasets

We continue our analysis with the datasets formed by portfolios sorted on momentum. Table 4 reports the performance of each strategy for the two momentum datasets. The first row shows us that the means of the 1/N strategy for each dataset are 5.75% and 17.11% respectively. The volatility values are 16.85% and 19.90%, and the Sharpe ratios are 0.34 and 0.86 respectively. The portfolio alphas are 0.38 and 0.11. All of the optimized portfolio strategies outperform the 1/N strategy in terms of Sharpe ratio on the US dataset, with values ranging from 0.35 to 0.55. The associated p-values are statistically significant at a 1% level for VT(η) and RRT($\hat{\mu}_t^+, \eta$), and at the 5% level for the GMV portfolio. For the RRT($\bar{\beta}_t^+, \eta$) strategy, the p-value is statistically significant at a 10% level for $\eta = 1$, while

Table 4: **Results for the Momentum datasets**

	US						Norwegian					
Strategy	$\hat{\mu}$	$\hat{\sigma}$	$\hat{SR}$	**pval**	$\hat{\alpha}$	**pval**	$\hat{\mu}$	$\hat{\sigma}$	$\hat{SR}$	**pval**	$\hat{\alpha}$	**pval**
1/N	5.75	16.85	0.34	-	0.38	-	17.11	19.90	0.86	-	0.11	-
{Panel A}												
VT(1)	6.09	15.96	0.38	**0.000**	0.39	0.389	16.35	19.43	0.84	0.662	0.11	0.491
VT(2)	6.27	15.56	0.40	**0.000**	0.39	0.376	16.08	19.34	0.83	0.655	0.12	0.436
VT(4)	6.37	15.27	0.42	**0.000**	0.39	0.440	16.40	19.87	0.83	0.663	0.17	0.349
{Panel B}												
RRT($\hat{\mu}_t^+, 1$)	6.61	15.37	0.43	**0.009**	0.44	0.213	18.16	19.92	0.91	0.155	0.20	0.200
RRT($\hat{\mu}_t^+, 2$)	7.00	15.52	0.45	**0.008**	0.46	0.148	18.93	20.83	0.91	0.317	0.26	0.176
RRT($\hat{\mu}_t^+, 4$)	7.39	15.73	0.47	**0.008**	0.49	0.102	19.91	22.23	0.90	0.476	0.33	0.178
RRT($\bar{\beta}_t^+, 1$)	5.78	16.11	0.36	**0.098**	0.33	0.977	16.56	19.76	0.84	0.719	0.06	0.820
RRT($\bar{\beta}_t^+, 2$)	5.73	16.00	0.36	0.173	0.30	0.974	16.34	19.92	0.82	0.748	0.05	0.764
RRT($\bar{\beta}_t^+, 4$)	5.60	16.02	0.35	0.296	0.28	0.972	16.23	20.91	0.78	0.860	0.03	0.712
{Panel C}												
GMV	8.43	15.44	0.55	**0.023**	0.56	**0.080**	15.77	22.83	0.69	0.918	0.25	0.336

Table 4: This table summarize our results for the performance of the 1/N strategy (row one), Volatility Timing strategies (Panel A), Reward-to-Risk Timing strategies (Panel B), and the Global Minimum Variance portfolio strategy (Panel C). We report the following sample statistics for the time series of monthly excess return generated by each strategy: annualized mean ($\hat{\mu}$), annualized volatility ($\hat{\sigma}$), annualized Sharpe ratio ($\hat{SR}$), portfolio alpha ($\hat{\alpha}$) and the associated p-values for our hypotheses from section 3.4.1.

when η is set to 2 or 4 the p-value is not statistically significant. The Norwegian Sharpe ratios are only higher than the 1/N Sharpe ratios for the RRT($\hat{\mu}_t^+, \eta$) strategy and none of the associated p-values are statistically significant. When it comes to the US alphas the only ones that are lower than the 1/N alphas are the RRT($\bar{\beta}_t^+, \eta$) strategy alphas. The p-values for the GMV portfolio is statistically significant. The Norwegian alphas are generally higher for the VT and the RRT($\hat{\mu}_t^+, \eta$) strategies, but neither of these associated p-values are statistically significant either. The GMV portfolio has an alpha of 0.25, with a non-significant p-value of 0.336.

5.4 Size datasets

The last datasets we investigate is the datasets consisting of portfolios sorted on size. Table 5 reports the performance results for the US dataset and the Norwegian dataset. In row one the mean, volatility, Sharpe ratio, and alpha are 7.50%, 19.20%, 0.39, and 0.46 for the US dataset, and 25.23%, 15.84%, 1.59, and 0.76 for the Norwegian dataset respectively. Some of the timing strategies on the US dataset slightly outperform the 1/N strategy with Sharpe ratios ranging from 0.39 to 0.43. None of the p-values are

Table 5: Results for the Size datasets

Strategy	$\hat{\mu}$	$\hat{\sigma}$	$\hat{SR}$	pval	$\hat{\alpha}$	pval	$\hat{\mu}$	$\hat{\sigma}$	$\hat{SR}$	pval	$\hat{\alpha}$	pval
	US						**Norwegian**					
1/N	7.50	19.20	0.39	-	0.46	-	25.23	15.84	1.59	-	0.76	-
{Panel A}												
VT(1)	7.38	18.67	0.40	0.101	0.47	0.355	24.24	14.90	1.63	**0.068**	0.78	0.270
VT(2)	7.24	18.18	0.40	0.120	0.47	0.373	23.12	14.06	1.64	0.148	0.84	0.191
VT(4)	6.91	17.45	0.40	0.186	0.46	0.471	21.68	13.98	1.55	0.565	0.91	0.170
{Panel B}												
RRT($\hat{\mu}_t^+, 1$)	7.25	17.74	0.41	0.274	0.57	0.160	25.56	15.15	1.69	**0.000**	0.85	**0.006**
RRT($\hat{\mu}_t^+, 2$)	7.51	17.77	0.42	0.217	0.59	0.136	25.38	14.63	1.74	**0.002**	0.93	**0.008**
RRT($\hat{\mu}_t^+, 4$)	7.75	17.87	0.43	0.201	0.61	0.108	24.22	14.33	1.69	0.131	1.01	**0.032**
RRT($\bar{\beta}_t^+, 1$)	7.55	19.36	0.39	0.760	0.46	0.764	24.89	15.45	1.61	0.121	0.77	0.409
RRT($\bar{\beta}_t^+, 2$)	7.61	19.50	0.39	0.691	0.46	0.723	24.60	15.19	1.62	0.167	0.78	0.382
RRT($\bar{\beta}_t^+, 4$)	7.75	19.68	0.39	0.583	0.45	0.655	23.86	14.88	1.60	0.368	0.79	0.382
{Panel C}												
GMV	9.51	15.88	0.60	**0.011**	0.81	**0.004**	23.89	14.85	1.61	0.452	1.10	**0.064**

Table 5: This table summarize our results for the performance of the 1/N strategy (row one), Volatility Timing strategies (Panel A), Reward-to-Risk Timing strategies (Panel B), and the Global Minimum Variance portfolio strategy (Panel C). We report the following sample statistics for the time series of monthly excess return generated by each strategy: annualized mean ($\hat{\mu}$), annualized volatility ($\hat{\sigma}$), annualized Sharpe ratio ($\hat{SR}$), portfolio alpha ($\hat{\alpha}$) and the associated p-values for our hypotheses from section 3.4.1.

statistically significant. The GMV portfolio properly outperform the 1/N strategy with a Sharpe ratio of 0.60 and a p-value that is statistically significant at the 5% level. For the Norwegian dataset the Sharpe ratios are generally higher and range from 1.55 to 1.74 for the timing strategies. Setting $\eta = 1$ and $\eta = 2$ gives us significant p-values at the 1% level for the RRT($\hat{\mu}_t^+, \eta$) strategy. We get a significant p-value at the 10% level for the VT(1) strategy. The GMV portfolio Sharpe ratio for the Norwegian dataset is 1.61, with a non-significant p-value. When it comes to the alphas for the US dataset, they range from 0.45 to 0.61 for the timing portfolios and 0.81 for the GMV portfolio. Most of them higher are than the 1/N portfolio alpha. The p-values are statistically significant at the 1% level for the GMV portfolio, while none of the other p-values are statistically significant. For the Norwegian dataset all the optimized portfolio alphas are higher than the 1/N portfolio alpha and range from 0.78 to 1.01 for the timing strategies and is 1.10 for the GMV portfolio. The p-values are significant at the 1% level when setting $\eta = 1$ and $\eta = 2$, and at the 5% level setting $\eta = 4$ for the VT strategies. The p-value for the GMV portfolio is significant at the 10% level. The other p-values are not statistically significant.

6 Discussion

Our results show us that in terms of Sharpe ratio the timing strategies perform well on the US data. Especially the VT strategy, which outperforms the naive strategy with statistically significant margins on three out of the four datasets we have tested. The $\mathrm{RRT}(\hat{\mu}_t^+, \eta)$ outperform the naive strategy with statistically significant margins only on the momentum dataset. However, when using the alternative estimator of conditional expected returns the naive strategy is outperformed on three datasets. It is worth noting that on the momentum dataset, $\mathrm{RRT}(\bar{\beta}_t^+, \eta)$ only outperform the naive strategy with statistically significant margins when $\eta = 1$.

The GMV strategy also outperforms the naive strategy on three datasets with statistically significant margins, and was the only strategy to do so on the size dataset. On the Norwegian datasets, there are much fewer Sharpe ratios for the optimized portfolio strategies that are higher than that of the 1/N strategy, and even fewer that have statistically significant p-values. The only time we see the timing strategies outperform the naive strategy with statistically significant margins is when we test them on the size dataset, where $\mathrm{RRT}(\hat{\mu}_t^+, \eta)$ is the best performing strategy.

The observation that attracts our attention is that the results for the portfolio alpha does not correspond with the results for the Sharpe ratio. If we take a look at the US datasets. The optimized portfolio strategies virtually outperforms the naive strategy in terms of the Sharpe ratio. They do not outperform the naive strategy as consistent when using the portfolio alpha as a performance measure. Additionally, none of the timing strategies have statistically significant p-values associated with the portfolio alpha. The GMV portfolio alpha has statistically significant p-values for the industry, momentum, and size datasets. On the other datasets we receive better results for the optimized portfolio strategies, yet statistically significant p-values only for the $\mathrm{RRT}(\hat{\mu}_t^+, \eta)$ portfolio alpha on the size dataset, and for the GMV portfolio alpha on the BM and size datasets.

These results leave us with some uncertainty. First, the different results from using the Sharpe ratio and the portfolio alpha as performance measures. Zakamulin (2017) argued that the superior performance of the optimized portfolios is due to them assigning a lot of weight on the assets with the lowest volatility. He demonstrated that when controlling for the Fama-French High-Minus-Low risk factor, the alphas of the optimized portfolios becomes neither economically nor statistically significant. Blitz (2016) confirmed that the HML factor is a proxy for a distinct low-volatility risk factor. We control for two additional risk factors and see the same results for the timing strategies, while for the GMV portfolio strategy, we observe compliance between the results for Sharpe ratio and portfolio alpha on three of the four US datasets. For the Norwegian book-to-market and size datasets we even see the portfolio alpha report superior performance where the Sharpe ratio does not, at least not with statistically significant p-values. However, since our method of implementing the GMV portfolio strategy does not impose the long-only constraint, the advantage of this strategy may be eroded if we impose transaction costs, as this strategy is characterized by higher turnover than the timing strategies (Kirby & Ostdiek, 2012). Considering all the portfolio alphas that are neither economically nor statistically significant, further doubt emerge to whether portfolio optimization contribute to better performance.

Another issue we encounter is the rather unconvincing performance of the RRT$(\hat{\mu}_t^+, \eta)$ strategy. Kirby and Ostdiek (2012) found that the performance of the RRT$(\hat{\mu}_t^+, \eta)$ improved when there is a good dispersion in the cross-section of average stock returns, due to more information from the conditional means. As we can see in Figure 1, this dispersion is quite small in three of the US datasets we use, while on the momentum dataset, where the means range from 1.56% to 16.59%, the RRT$(\hat{\mu}_t^+, \eta)$ strategy outperforms the naive strategy by statistically significant margins for the Sharpe ratio. However, despite good dispersion in the means of the Norwegian datasets, as we can see in Figure 2, this strategy still disappoints.

7 Summary and Conclusion

There is an ongoing debate about whether portfolio optimization adds value or not. DeMiguel et al. (2009) conducted a highly influential study questioning the value of various optimization strategies compared to the 1/N strategy. In response, many studies claim to defend the optimized portfolios by demonstrating their superior performance. Many of these studies use the Sharpe ratio as their performance measure, which do not account for the possibility that the superior performance can be attributed to established risk factor premiums (Zakamulin, 2017). Motivated by these issues we simulate the strategies of Kirby and Ostdiek (2012) and the Global Minimum Variance portfolio. We compare their performance relative to the performance of the 1/N strategy in the out-of-sample period. Both the Sharpe ratio and portfolio alpha from a multi-factor model were used to examine the performance before and after accounting for several risk factor premiums. In addition, we extend the sample period of Kirby and Ostdiek with more recent data and also add four Norwegian datasets with portfolios sorted on the same criteria to examine whether these strategies work in Norway.

We find that all the optimized portfolio strategies perform well on the US datasets in terms of the Sharpe ratio, which agrees well with the results in the study by Kirby and Ostdiek (2012). However, when we account for the risk factor premiums using the portfolio alpha, we find no evidence of superior performance for the timing strategies. This is in line with the conclusion of Zakamulin (2017). The Global Minimum Variance portfolio show superior performance even in terms of the portfolio alpha, however the advantage of this strategy may be eroded if we impose transaction costs.

The results for the optimized portfolios on the Norwegian datasets are not convincing. Out of the four different datasets we examine, only one of them show promising results for the timing strategies. Though we find a number of optimized portfolios with higher Sharpe ratio than the naively diversified portfolio on this dataset, there are only a few of them with statistically significant p-values. The results for the portfolio alpha are similar, which makes us doubt that these strategies work in Norway.

With these results we argue that there is reason to doubt whether portfolio optimization adds value. The strategies we have tested seem to benefit from some known anomalies. In Norway, the optimized portfolios do not show promising results.

References

Behr, P., Guettler, A., & Miebs, F. (2013). On portfolio optimization: Imposing the right constraints. *Journal of Banking & Finance*, *37*(4), 1232–1242.

Blitz, D. (2016). The value of low volatility. *Journal of Portfolio Management*, *42*(3), 94–100.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*(1), 57-82.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Annals of Statistics*, *14*, 1171-1179.

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

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

Disatnik, D., & Katz, S. (2012). Portfolio optimization using a block structure for the covariance matrix. *Journal of Business Finance & Accounting*, *39*(5-6), 806–843.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, *7*(1), 1–26.

Fama, E. F., & French, K. R. (1992). The economic fundamentals of size and book-to-market equity.
(Unpublished working paper. University of Chicago)

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56.

Haley, M. R. (2016). Shortfall minimization and the naive (1/n) portfolio: an out-of-sample comparison. *Applied Economics Letters*, *23*(13), 926-929.

Hall, P., Horowitz, J. L., & Jing, B.-Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika*, *82*(3), 561-574.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers:

Implications for stock market efficiency. *Journal of Finance*, *48*(1), 65-91.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *Journal of Finance*, *23*(2), 389-416.

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

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

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

Kirby, C., & Ostdiek, B. (2012). It's all in the timing: Simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis*, *47*(2), 437-467.

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

Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, *17*(3), 1217-1241.

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

MacKinlay, C. A., & Pástor, L. (2000). Asset pricing models: Implications for expected returns and portfolio selection. *Review of Financial Studies*, *13*(4), 883–916.

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

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

Merton, R. C. (1972). An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis*, *7*(04), 1851–1872.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, *8*(4), 323 - 361.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the*

Econometric Society, *34*(4), 768–783.

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*(428), 1303-1313.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *13*(3), 341-360.

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

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business*, *39*(1), 119-138.

Treynor, J. L. (1961). Market value, time, and risk.
(Unpublished manuscript. "Rough Draft" dated 8/8/61, 95-209)

Treynor, J. L. (1962). Toward a theory of market value of risky assets.
(Unpublished manuscript. "Rough Draft" dated by Mr. Treynor to the fall of 1962. A final version was published in 1999, in Asset Pricing and Portfolio Performance. Robert A. Korajczyk (editor) London: Risk Books, pp. 15.22)

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

Zakamulin, V. (2017). Superiority of optimized portfolios to naive diversification: Fact or fiction?
(Forthcoming in *Finance Research Letters*)

Appendix A Reflection note of Andreas

Our topic is portfolio management. We re-examine the performance of some optimized portfolios relative to the naively diversified portfolio (allocating equal weights to all assets). There is an ongoing discussion on the value of mean-variance optimization of portfolios. A study in 2009 found that none of the examined optimized portfolios outperformed the naive strategy with statistically significant margins. Later, several financial researchers have claimed to defend mean-variance optimization. However, their methods for measuring performance did not account for the possibility that the superior performance could be attributed to some established risk factors.

We expand the literature by accounting for three well known factors when measuring the performance of some optimization strategies. Additionally, we test the strategies on Norwegian datasets, to see if they work in Norway. Our results show that the optimized portfolios do indeed outperform the naively diversified portfolio before accounting for the risk factors. After accounting for the risk factors however, we find little evidence of superior performance. The results on the Norwegian data show that these strategies do not work in Norway.

Portfolio management is closely related to international trends and globalization. Since the introduction of computers and internet, the financial markets have become more connected and have greater impact on each other. For example, the Norwegian stock exchange is highly correlated with the stock exchanges in the US, in Europe, and in Asia. We also saw how the problems of some economies led to a worldwide financial crisis. Information about the markets are now easily available for the investors anywhere in the world and trades happen instantly. Portfolio managers should take international factors into account when picking stock for their portfolio. The demand from private investors for the opportunity of investing in certain technologies or companies located in other countries is huge. Institutions providing broker services must adapt to comply with these demands, and to do so they need a team of portfolio managers and others with international understanding.

The competition in financial market has grown considerably. Norwegian stock brokers and banks now compete with international institutions and must be innovative to hold on to their customer base and acquire new customers. As a part of this competition, portfolio managers have an increased incentive to provide higher returns than the passive market portfolio. It is therefore essential to find the optimal methods of portfolio selection, so that the return on the portfolio is as high as possible. Many private investors put their money in a passive fund, that follow some market index. The goal for portfolio managers should therefore be to convince customers that they can beat the market, by managing their portfolios using methods that have scientific evidence for superior performance.

Investors rely on portfolio managers to take good care of their invested money. They trust that the portfolio managers will accommodate the statements in the portfolio prospectus and not make exceedingly risky investments. The portfolio managers must therefore invest responsibly. They should also contribute to effective and well-functioning markets by following international standards for responsible ownership. The portfolio managers have the responsibility to perform ethical investments and counteract corruption, child labor, human rights violations etc. In addition, environmental issues have become important for investors, and should be accounted for.

Appendix B Reflection note of Erik

In this thesis, we have tested the performance of some optimized portfolios compared to the naive strategy using out-of-sample testing. We have performed a re-examination of these optimized portfolio strategies by testing for both US and Norwegian data. Most of the time-period for US datasets are tested in similar prior studies, but by adding newer data, we seek to find evidence whether former conclusions still hold. By testing these strategies using Norwegian datasets, we seek to discover if the conclusions found on the US data, will apply to the Norwegian as well. It is argued that prior studies may not find a true evidence due to only using the Sharpe ratio as performance measurement. By implementing the portfolio alpha as a performance measurement alongside the Sharpe ratio.

Our thesis has contributed research of the financial sector by examining more recent data, include data from the Norwegian market, and by adding Alpha to the Sharpe ratio as performance measure when concluding our research. After implementing several risk-factors, we found that for the Norwegian datasets, the optimized strategies do not outperform the naive diversification strategy. Regarding the US datasets, we found that the naïve strategy out-perform three of four optimized strategies. The optimized strategy, Global Minimum Variance out-perform naïve, though we consider this may be because transaction costs are not accounted for. According to these results, we conclude that there is little evidence of the optimized out-performing the naïve strategy.

The professional background-material provided by the University of Agder have been essential for us being able to carry out and complete this thesis. Primarily this thesis is a product of lectures given in econometrics, methodology, and finance theory.

The topic of optimizing portfolio strategies can be related to internationalization because it can be applied to all capital market indices, like e.g the NASDAQ, AMEX, OSE etc. We live in a time where communication rarely is a problem, which have been a depending factor in making portfolio strategies a well-known and well-discussed subject world-wide. Today, financial markets across borders are more than ever, connected and dependent of each other. We have already seen how the financial catastrophe in the US in 2008 affected other financial markets world-wide. Different indices can often be highly correlated, and little suggest that the correlation will decrease in any near future. Due to instant sharing through social media, public stock-information is available to all. Because trading often can be done instantly, portfolio managers should consider more possible factors when investing in stocks and developing portfolios.

Though our study is a re-examination, we consider some of our work to be innovative. Though the topic of comparing the naive diversification strategy to optimized portfolio strategies has been part of a number of financial studies later years, we still consider it innovative due to the implementation and testing using Norwegian data. Also, in contrast to some of the prior studies, we implement alpha as a performance measurement as we believe our conclusions may be more reliable than conclusions depending only on the Sharpe ratios. Later years, more and more people in Norway have started investing in passive funds, rather than collecting their money in a risk-free saving account. By doing so, they "*force*" the portfolio managers to try convincing the costumers that they are able to gain a higher return than the index funds or else they lose their customers. The goal has to be more innovative than others and by maximizing the return at lowest possible risk.

This topic can be related to responsibility due to how portfolio managers should act when investing. By investing and managing other peoples money, portfolio managers can be prohibited from investing in stocks which have any connection to war, drugs or other actions which would seem unethical. Another thing which relate investments to responsibility, is that portfolio managers handling others money should not invest others money in assets that are have an unreasonably high risk.