UNIVERSITY OF GOTHENBURG

SCHOOL OF BUSINESS, ECONOMICS AND LAW

The benefits of optimized portfolios

An empirical comparison between optimized portfolios and benchmarks

Authors: John Nestenborg & Simon Petersson

Supervisor: Jian Hua Zhang

Master's thesis in Finance, 30 hec

Spring 2022

Graduate School, School of Business, Economics and Law, University of Gothenburg

Acknowledgement

We would like to express our sincere gratitude to our supervisor, Jian Hua Zhang, for his guidance in the writing process of this thesis. His insightful feedback helped us to develop the framework of our thesis and improve the overall quality. Furthermore, we would like to acknowledge the invaluable experience we gained during the fall of 2021 at EDHEC Business School.

Abstract

Uncertainty about the future is an everlasting part of investing. This study aims at testing the historical performance out-of-sample for optimized portfolios and if the performance was superior to benchmarks. 11 different portfolios are compared to two different benchmarks; the naive- and market-capitalized portfolio. The portfolios are optimized using popular weighting schemes aimed at minimizing risk. The tested portfolios comprised 6 different minimum-variance portfolios and 5 equal-risk contributing portfolios. We designed this study to focus on the European market and constituents part of EURO STOXX 50 between 2006 and 2021. Previous studies are contradicting in their findings regarding the performance of optimized portfolios. The methodology is designed to generate an overview of standard performance measures such as annualized return, Sharpe ratios, and maximum drawdowns, and we combined this overview with statistical tests on Jensen's alpha and annualized Sharpe ratios. We found evidence of superiority amongst our optimized portfolios compared to the market-capitalized benchmark, but no evidence of out-performance against the naive portfolio.

JEL Classification: G11

Keywords: Optimized portfolios, Global Minimum Variance, GMV, Equal Risk Contribution, ERC, Naive portfolio, Market-Capitalization portfolio, Comparison between portfolio weighting schemes.

Contents

List of Figures v

List of Tables v

1 Introduction 1

1.1 Purpose . . . 2

1.2 Hypotheses . . . 3

1.3 Structure of paper . . . 4

2 Literature review 5

2.1 Modern portfolio theory . . . 5

2.2 Previous research . . . 7

3 Methodology 11

3.1 Portfolios . . . 11

3.1.1 Naive portfolio . . . 12

3.1.2 Market Capitalization weights . . . 12

3.1.3 Global Minimum Variance . . . 13

3.1.4 Equal-Risk Contribution . . . 13

3.2 Performance measures . . . 15

3.2.1 Maximum drawdown . . . 15

3.2.2 Effective number of holdings . . . 16

3.3 Sharpe ratio . . . 16

3.3.1 Test statistic . . . 17

3.4 Jensen's alpha . . . 18

4 Data 19

5 Results 20

5.1 Statistical tests . . . 22

6 Discussion 25

6.1 Jensen's alpha . . . 27

6.2 Sharpe ratios 28

7 Conclusion **31**

7.1 Future research 32

References **33**

List of Figures

1 The efficient frontier 6
2 In-sample Efficient Frontier 22

List of Tables

1 Portfolios and benchmarks 11
2 Summary of Statistics 20
3 Regression Analysis - Jensen's Alpha 23
4 Sharpe Ratio (SR) - Test 24

1 Introduction

The empirical performance of portfolios out-of-sample is an interesting and important subject. Due to an increased public interest in fund and ETF investments, it has become essential to study and understand the abundant weighting schemes available (Glosten et al. 2021). More capital is allocated to passive alternatives than ever before, and for good reasons - previous studies have indicated that active portfolio management seldom outperforms its benchmark, meaning that it is difficult to defend the high fees associated with such investment (Otuteye & Siddiquee 2020).

Portfolio management literature often aims to find and understand which strategies deliver optimal performance empirically. Generating above-market returns and risk-adjusted returns is a sought after ability. Portfolio literature has generated a vast number of investment strategies aiming at obtaining above-market returns. The history of developed strategies within portfolio literature is often traced back to Modern Portfolio Theory (MPT). MPT is a framework for analyzing and allocating between stocks to maximize overall return for a given level of risk (Markowitz 1952). Markowitz introduced the concept of minimizing risk for a given return expectation by utilizing the correlation between assets. Since then, academic literature has presented various allocation strategies within Markowitz's space of mean and variance. The performance of Mean-Variance Optimized (MVO) portfolios has been heavily researched and debated for years (Baker et al. 2011, DeMiguel et al. 2009). Historical data is the only fundamental measure available in empirical finance and MPT. However, an insufficient amount of research has been conducted concerning relying too heavily on historic properties for out-of-sample allocation. Furthermore, MVO portfolios often bring complexity, and the question regarding its benefits has been contested. Much like previous studies, we will focus heavily on the Global Minimum Variance (GMV) portfolio and its performance out of sample (DeMiguel et al. 2009, Kritzman et al. 2010). We find it imperative to generate a broad spectrum when analyzing optimized portfolios, meaning that we will include various other optimized portfolios. Well-known optimized weighting schemes serve different purposes; for example, the GMV portfolio only minimizes volatility and attracts risk-averse investors. Nevertheless, the fact that the GMV portfolio serves its purpose well in-sample, does not imply minimized risk out-of-sample. The main reason for this is that the method is known to allocate to a low number of effective stocks resulting in lower diversification (Baker et al. 2011). The uncertainty of portfolio investing is one of the main drivers behind our thesis.

1.1 Purpose

The purpose of this paper is to bring clarity to the out-of-sample properties of several of the most popular strategies in portfolio theory literature. There is a broad spectrum of research in this field that we wish to build upon, and the intention is to find results that apply to the European stock market. One of the most popular investment strategies is the naive (1/N) portfolio which allocates equal weights amongst the constituents, and the market-weighted or market-capitalized (MCaP) portfolio, which bases its weights on the size of the compane and often serves as an index. Previous studies have focused on the efficiency of these strategies (DeMiguel et al. 2009, Kritzman et al. 2010, Baker et al. 2011). We aim to clarify ambiguities regarding different passive quantitative strategies and their performance out-of-sample. This thesis will test the efficiency of two benchmarks, the 1/N- and MCaP portfolio. In addition, the study will investigate whether specific strategies can provide statistically superior risk-adjusted returns compared to the benchmarks.

In order to recognize which portfolios serve investors best, only out-of-sample performance is considered. The results of academic literature focused on the out-of-sample performance of portfolios are to no surprise opposing in their findings. Furthermore, several studies suggest that optimized portfolios are superior in risk-adjusted performance, while other studies find no such proof. This thesis extends the studies by DeMiguel et al. (2009) and Kritzman et al. (2010) by examining the out-of-sample risk-adjusted returns of various portfolio models. We will use both the MCaP weights and 1/N weights as benchmarks to broaden our research and compare several key metrics associated with portfolio performance.

The out-of-sample portfolios generate allocations for the coming year based on historical data. In other words, the portfolios will rebalance annually, and the allocations depend on historic properties such as volatility. Gaining insight into which strategies can generate superior risk-adjusted returns out-of-sample may be beneficial for investors looking to diversify and beat the market. The portfolios will be compared in various popular measures but only statistically tested in Jensen's alpha and Sharpe ratio. In addition, since investing can only happen out-of-sample, insights into historical performance measurements can provide investors with broader knowledge about different portfolios. These insights can help mitigate the most significant pitfall of MPT; estimating expected returns. Moreover, investors interested in EURO STOXX 50 Net Total Return (SX5T) will better understand long-term historical differences in performance across strategies.

1.2 Hypotheses

Based on previous research and the assumptions discussed in the next chapter, we have posed the following two hypotheses:

H1. There is a difference in risk-adjusted returns between optimized portfolios and the MCaP portfolio

H2. There is a difference in risk-adjusted returns between optimized portfolios and the Naive (1/N) portfolio

There is a trade-off between rebalancing frequency and associated trading costs. Trading costs will minimize the returns and possibly delete alpha (Novy-Marx & Velikov 2016, Patton & Weller 2020). By not acknowledging the associated trading costs of strategies, results can be misleading. In this paper, we assume no transaction costs. The optimized portfolios, which rebalance annually, will carry transaction costs during each rebalancing period. Novy-Marx & Velikov (2016) found transaction costs to vary heavily depending on strategy and rebalancing frequency. Thus, the portfolio performances in our study will slightly differ from actual market implementation.

This paper also assumes no leavers and joiners in the index. The constituents part of the index by mid-2021 are thus the only index constituents we will focus on throughout our study. Due to not incorporating leavers and joiners, our thesis will only include 48 constituents. By analyzing the historical performance of current constituents, the method of this thesis does not capture the effect of historical leavers and joiners in the index. By disregarding the performance of the actual historical constituents, the performance of the portfolios may be influenced by survivorship bias. However, the impact of this limitation is partly mitigated by the fact that the index returns also reflect the historical performance of the same current constituents.

1.3 Structure of paper

In the next chapter, we will present a literature review of relevant theories and previous research within the scope of this thesis. Chapter 3 contains the methodology of the thesis and presents different methods used to allocate weights for the different portfolios. Chapter 4 describes the data. Chapter 5 presents the overall results for each portfolio, both in terms of the entire period and individual years. Chapter 6 consists of a discussion and analysis. Lastly, a conclusion will be presented.

2 Literature review

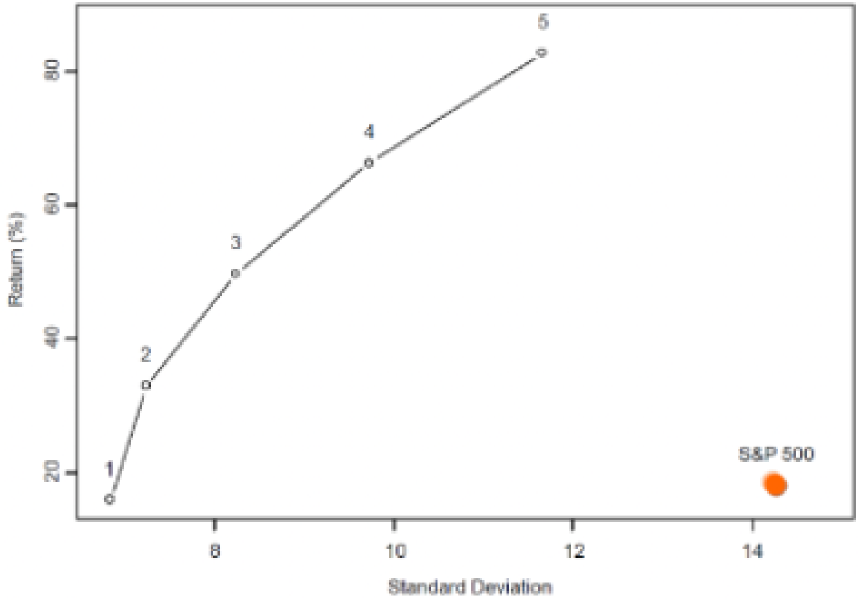
Preceding literature played an essential part in this thesis. Therefore, we will provide a background on the subject and include previous notable research on the portfolios and their methodology, findings, and potential drawbacks.

2.1 Modern portfolio theory

It is well known that Markowitz (1952) pioneered Modern Portfolio Theory (MPT). The results and theories of Markowitz suggested that investors should invest in Mean-Variance Optimized (MVO) portfolios to acquire the best return. The theories of Markowitz have been some of the most important for the field of portfolio theory (Santos 2010). MVO portfolios are quantitative asset allocation investments to create a portfolio for the upcoming period. MPT assumes that investors care solely about expected returns and the risk attributed to these returns, and if the investors' preferences about MVO portfolios are assumed, they should invest such that the return is maximized for a given level of risk.

From the mean-variance framework, the efficient frontier can be created. The efficient frontier is a famous concept consisting of mean-variance efficient portfolio allocations. The efficient frontier is well used in portfolio performance analysis, and its popularity heavily stems from its simplicity. The frontier is made of portfolios that maximize return for a given level of risk (volatility) within the risk-return space that in turn makes the mean-variance space (Markowitz 1952). Expected returns and standard deviations are the generally used axes within the mean-variance analysis. The top hyperbola of these portfolios, shown in figure 1, generate the efficient frontier and are considered efficient portfolios.

Figure 1: The efficient frontier



Schwartz (2000), figure 3, p. 19

The points along the top hyperbola in figure 1 make up the efficient frontier. The portfolios along the efficient frontier are the portfolios that maximize return for the given levels of risk. The return is given on the y-axis in the efficient frontier figure, whereas the volatility, the measure of risk, of portfolios increases along the x-axis. In figure 1, Schwartz (2000) shows the position of the S&P 500 between 1979 and 1998.

The findings of Markowitz (1952) have generated traction, and his theories have been used since their inception. However, the theories have faced criticism throughout the years (Xiong & Idzorek 2011, Statman 2013). Markowitz and the MVO framework have faced criticism because the theories do not analyze the underlying companies. Thus, many vital components will be overlooked if company-specific financial data and the growth of the companies are disregarded. Moreover, some MVO strategies have been criticized for not incorporating investors' true utility. Investors have utility of performance and terminal value, not how the risk is diversified.

2.2 Previous research

It is difficult for investors to predict which allocations will be optimal in future periods as only historical data is available. The future values for the returns, variances and covariances are never known. Optimized portfolios often rely on estimations to allocate for future periods, and the estimates often stem from historical information. Performance discrepancies will naturally exist when crossing the threshold between in-sample and out-of-sample. Hence, the errors will come when investors' go live with investment strategies - such as the GMV. The portfolio may have been optimal in-sample, but the future is always uncertain. A well-used method within finance is to test portfolios and investment strategies on historical data. With financial data, investors can backtest an endless number of investment strategies to find a strategy of choice. In finance, backtesting refers to applying an investment strategy or portfolio to historical data to derive historical statistical- and performance properties. Such properties can be overall performance, risk-adjusted performance, and downside measurements. Backtesting enables investors to create a historical and theoretical path of performance.

According to Monarcha (2019), backtesting bias is a common issue within quantitative finance. Data biases such as survivorship - and look-ahead bias can lead to backtests of portfolio strategies overstating achievable results. Monarcha finds that backtesting biased strategies often brings drops in alpha. Backtesting bias is a common issue in quantitative finance, which is inherent to the natural propension of quants to maximize the risk/return properties of backtested investment strategies before commercialization for obvious marketing purposes (Monarcha 2019). The implications of backtesting biases are part of each investment strategy. The bias transcends to the portfolios studied in our thesis as the optimized portfolios were superior in-sample. We want to elucidate the performances once the portfolios are out-of-sample.

Previous research focused on portfolio theory often focuses on the potential benefits of optimizing. These studies usually revolve around obtaining superior risk-adjusted returns against either a benchmark such as the MCaP- or 1/N portfolio. Since the estimation risk is inherent in expected returns, the fact that the minimum volatility portfolio relies only on risk parameters is considered to be an appealing feature (Kempf & Memmel 2006). In other words, alternatives to return estimating portfolios such as GMV might provide greater accuracy as it is only based on an estimated covariance matrix (Merton 1980). For example, Jagannathan & Ma (2003) studied a Global Minimum Variance Portfolio (GMVP) and found that it can find better risk-adjusted returns than both the 1/N- and MCaP portfolio. Our study builds on this feature as the portfolios tested require no estimation input in future return or risk.

Extensive research has been conducted on low volatility portfolios (DeMiguel et al. 2009, Haugen & Baker 1991, Kritzman et al. 2010, Merton 1980). The risk-adjusted return for low volatility portfolios is presumed to be lower as most portfolios are constructed to minimize risk instead of maximizing return. Haugen & Baker (1991) are viewed as the pioneers of research on minimum-variance portfolios. They studied the 1000 largest U.S stocks over 17 years and concluded that the GMV-portfolio consistently outperformed the MCaP in risk-adjusted returns. They posit that MCaP portfolios and indices are inefficient mainly due to investors' inability to perform short-selling. The view of the value-weighted portfolio being inefficient was further confirmed by Jagannathan & Ma (2003). Schwartz (2000) illustrated the inefficiency of the value-weighted portfolio by comparing it with the efficient frontier shown in figure 1. Schwartz studied the S&P 500 between 1979 and 1998 and found that the GMV outperformed the MCaP in risk-adjusted returns.

As previously mentioned, we will test the portfolios in Jensen's alpha and Sharpe ratios, which are both risk-adjusted return measurements. Jensen's alpha has been heavily used to evaluate funds and fund manager performance. Jensen's alpha builds on the Capital Asset Pricing Model (CAPM), in which alpha represents the average deviation from the CAPM (Jensen 1968). Alpha can take both negative and positive values, and a value above zero represents superiority, also known as abnormal return, to the benchmark; whereas a value below zero means that the benchmark outperforms the tested strategy.

The out-of-sample performance of optimized portfolios is a well-researched field of subject. One of the most renowned studies was done by DeMiguel et al. (2009)

who studied the out-of-sample performance of optimized portfolios in relation to a 1/N portfolio. They studied different portfolios and examined returns over different time periods. After that, the returns were tested against the returns of the 1/N portfolio. The study analyzed 14 different weighting schemes using monthly returns. Looking for statistical significance in Sharpe ratios, they did not find any evidence of superior performance of the optimized portfolios. The allocations used various parameters and sampling windows to allocate for the out-of-sample periods. In terms of data and contrary to this thesis, DeMiguel et al. analyzed out-of-sample performance between 1963 and 2004. The method used by DeMiguel et al. for testing the Sharpe ratios will be incorporated in this thesis in order to be able to extend previous findings and compare outcomes for different datasets.

A similar method used in DeMiguel et al. (2009) was used in Kritzman et al. (2010) who found that optimized portfolios did deliver superior Sharpe ratios compared to a 1/N benchmarks. They examined the returns from minimum-variance and mean-variance portfolios against the 1/N benchmark. The study examined out-of-sample returns in 13 different datasets for 50 000 optimized portfolios. Kritzman et al. incorporated long-only constraints due to the critique of MVO portfolios generating unrealistic allocations. We build on the same critique by having long-only constraints in the studied portfolios. A prominent similarity between DeMiguel et al. (2009) and Kritzman et al. (2010) is that both studies use simulations to generate returns the datasets that the portfolios are compared within. The datasets are generated using past returns with various rolling windows of sampling, while this thesis will build on the methods and theories of these studies, we will use different portfolios to compare with the 1/N and MCaP portfolios.

In addition to GMVPs, we will include different Equal Risk Contribution (ERC) portfolios. This type of portfolio will allocate so that each constituent will contribute with the same amount of risk (Mausser & Romanko 2014). Choueifaty et al. (2013) studied different diversified portfolios, amongst which the ERC, between 1999 and 2010. They found that ERC portfolios deliver significantly higher returns and lower volatility than the MCaP. They posited that the ERC functions as a risk-weighted version of the 1/N portfolio as it had lower risk with similar returns. Furthermore, they found that ERC portfolios obtained higher Sharpe ratios than both the MCaP and 1/N benchmark. We will include different ERC portfolios that do not use volatility as the risk measure. There are disadvantages associated with using volatility as a risk measure. Volatility will not distinguish between positive and negative volatility. Subsequently, according to Alankar et al. (2012), using downside risk measures may generate 20% smaller drawdowns than volatility.

Previous research focused on the performance of optimized portfolios has failed to include adequate analysis of the basics of 1/N portfolios. Kritzman et al. (2010) highlights the appealing features of the 1/N portfolio: Avoiding concentrated positions, selling highs and buying lows on rebalancing dates, never underperforming the worst-performing constituent, and always investing in the best-performing asset. For 1/N portfolios, the efficient number of holdings equals the number of constituents between which it can allocate. In other words, for the 1/N portfolio focused on SX5T, the efficient number of holdings will be 48. For portfolios allocating using other parameters, such as GMVPs, the number of constituents in the portfolio may vary heavily from the efficient number of holdings. A gap in the previous research focusing on the performance of optimized portfolios that this thesis may help fill is to include an analysis of the efficient number of holdings when comparing returns.

The number of constituents to allocate between might also impact the portfolio performance. Duchin & Levy (2009) studied MVO portfolios and found that equally-weighted portfolios performed better than optimized ones when the number of constituents is lower. Moreover, they found that the naively diversified portfolio outperformed MVO portfolios when the number of constituents was between 15 and 25 but lost its superiority when the number of constituents increased to 30.

This study will resemble the studies mentioned above but include more focus on the difference in performance between optimized portfolios and the MCaP benchmark. Even though it is generally accepted to use a MCaP portfolio as a proxy for the market portfolio, it has been heavily criticized for its poor risk-adjusted performance (Sharpe 1991, Markowitz 2005, Arnott et al. 2005).

3 Methodology

This chapter aims to introduce the different allocation methods and investment strategies analyzed to test the hypotheses.

Next, the underlying theories and mathematical equations used to generate the portfolio weights will be presented. The portfolios have been constructed and analyzed using the programming language Python. We chose to analyze using Python as it was one of the most commonly used and practical options for our purposes. The code and methodology were based on solutions constructed by the EDHEC risk institute in Nice, France.

The weighting schemes were generated to our target through collaboration with professors in portfolio management and empirical finance at the EDHEC risk institute. Furthermore, we constructed the code used for the analysis to ascertain that the outputs fit our purposes.

3.1 Portfolios

Within the dataset, we constructed the portfolios displayed in table 1. The portfolios allocated between the constituents part of SX5T. 11 different optimized portfolios were studied to test our hypotheses. Each investment strategy follows different allocation rules and is rebalanced annually.

Table 1: Portfolios and benchmarks

No.	**Portfolio**	**Abbreviation**
1	**Global minimum variance 1yr**	GMV 1yr
2	**Global minimum variance 2yr**	GMV 2yr
3	**Global minimum variance 3yr**	GMV 3yr
4	**1% Global minimum variance 1yr**	1% GMV 1yr
5	**1% Global minimum variance 2yr**	1% GMV 2yr
6	**1% Global minimum variance 3yr**	1% GMV 3yr
7	**Equal risk contribution - Vol 1yr**	ERC vol 1yr
8	**Equal risk contribution - Vol 2yr**	ERC vol 2yr
9	**Equal risk contribution - Vol 3yr**	ERC vol 3yr
10	**Equal risk contribution - MDD**	ERC - MDD
11	**Equal risk contribution - CVaR**	ERC - CVaR
12	**Market-weighted**	MCaP
13	**Naive**	1/N

The various strategies and portfolios and the math behind the different allocations used are presented in the following subchapters. Whereas each portfolio is rebalanced on an annual basis, some portfolios allocated based on different number of historical years. We did not allow any short-selling for our portfolios, implying that the portfolios never held negative weights for any constituent. For example, the GMV 1 - year portfolio allocated based solely on the previous year's volatility, and the GMV 3 - year portfolio allocated based on 3 years' historical volatility. In other words, the portfolios had 3 different windows of estimation: 12 months, 24 months or 36 months. Moreover, certain limitations associated with the different strategies will be highlighted. Of the portfolios in table 1, MCaP and 1/N will be used as our two benchmarks. The remaining 11 portfolios are the optimized portfolios, and their performance was compared and tested against the benchmarks.

3.1.1 Naive portfolio

The naïve portfolio is constructed with equal 1/N weights, where N is the total number of portfolio stocks. Thus, the equal-weighted portfolio has an effective number of stocks equal to the number of stocks in the portfolio. Equation 1 gives the weight allocated to each constituent in the portfolio.

$$w_t^{EW} = 1/N \tag{1}$$

Equation 1 holds for all assets in the portfolio as w_t is the weight of each constituent. The 1/N portfolio does not involve any optimization or estimation. Consequently, the portfolio disregards attributed performance and only attains the number of available constituents as input before assigning the weights. Our dataset contained 48 constituents, and the weight allocated to each constituent was 2.0833%.

3.1.2 Market Capitalization weights

The weighting scheme for market capitalization is constructed by dividing the individual company's free-float value by the total market capitalization of all companies in the index. For this reason, companies with a high stock value hold a greater weight. Subsequently, stocks with lower market capitalization are assigned a smaller weight. The SX5T index follows a market-capitalization weighting scheme implying that this strategy will follow the index.

3.1.3 Global Minimum Variance

GMVPs aim to represent the portfolio with the lowest volatility. The theoretical GMVP is on the left-most point along the efficient frontier, where the variance is lower than all possible risky portfolios. The risk in MPT is volatility, and the GMVP allocates such that the portfolio risk is minimized. However, neglecting returns could be problematic for investors interested in maximizing profit. Equation 2 describes the mathematical solution and optimization problem of the GMVP.

$$w' \sum w \rightarrow \min \text{ s.t. } w'1 = 1 \tag{2}$$

Where w is the vector of weights, 1 is a vector of ones, and $\sum$ is the covariance matrix. The main aim of the GMVP is to minimize the overall variance. Two versions of the GMVP have been included in this paper. First, a GMVP without negative holdings and 100% of the weight is allocated. A GMVP without minimum ownership bounds tends to be heavily concentrated, thus obtaining a low effective number of stocks. In order to help minimize such a problem, a solution is to set bounds of minimum or maximum weights on each constituent. In this paper, a 1% minimum weight was used for 3 GMVPs. The maximum weight factor was not used for either version of the GMV.

3.1.4 Equal-Risk Contribution

ERC strategies allocated are so that each asset in the portfolio will contribute the same risk. Maillard et al. (2010) suggest that the ERC portfolio (in theory) will equal the maximum Sharpe ratio if all assets possess the same risk-reward with identical pairwise correlations. From equation 3 the ERC allocations can be retrieved using volatility as a risk measure. ERC portfolios find weight $w_i j$ so that:

$$w_i \frac{\delta \sigma_p}{\delta w_t} = w_j \frac{\delta \sigma_p}{\delta w_j} \tag{3}$$

Where σ_p is the volatility for the portfolio. 3 different ERC-volatility (ERC-VOL) portfolios were analyzed in this study, where the difference in the portfolios was in the sampling period of the volatility. Similar to the GMV portfolios, the sampling period ranged from 1 to 3 years sampling, indicating that the strategies view the periods of historical risk differently. Thus, the ERC-VOL 1-year strategy sampled the volatility from the constituents from 12 months prior before allocating, whereas the 3-year strategy sampled the volatility 36 months prior.

We also included two ERC portfolios that do not use volatility as risk measure. Instead, they use two more tangible measures of risk, Conditional Value at Risk (CVaR) and Maximum Drawdown (MDD). Using other risk measures such as CVaR or MDD diminishes the drawbacks of using volatility can be solved. CVaR, also called expected shortfall, is the mean of the losses that exceed the value at risk. In this paper, one portfolio will be diversified so that each constituent contributes equal CVaR value to the portfolio. The risk was, in this case, measured as historical CVaR at a 95% level. The portfolio considered the monthly return from each constituent for 3 years before allocating similarly to the ERC-vol strategy. The only difference was the measure of risk. Equation 4 shows the allocation methodology for the ERC-CVaR. The ERC CVaR finds weight w_{ij} so that:

$$w_i \frac{\delta x_p}{\delta w_t} = w_j \frac{\delta x_p}{\delta w_j} \tag{4}$$

Where x_p is the CVaR for each constituent. Another measure of risk used in empirical finance is MDD. The third ERC portfolio used in this paper viewed MDD as its sole risk measure. MDD is the maximum observed loss from a peak to a bottom for an asset or portfolio. By measuring the historical MDD of the constituents in an index, investors can diversify so that ERC criteria are met. For the ERC-MDD, x_p is the MDD for each constituent. Like the ERC-CVaR portfolio, the ERC-MDD considered the monthly returns from 12 months before allocating between the constituents.

3.2 Performance measures

We included 5 different performance measures for our optimized portfolios to provide a deeper analysis. After generating the weights for each portfolio, we obtained the returns for each strategy between 2009 and 2021. Then, we obtained each year's annual return, volatility, and sharpe ratio using the returns from each portfolio. We also generated the maximum drawdown and average efficient number of holdings for each portfolio with the monthly returns and annual weights. We generated the annualized returns for each portfolio over the full sample. We could find the annualized sharpe ratio used to test the hypotheses with the annualized returns and volatilities. Furthermore, with the monthly returns of each strategy, we derived jensen's alphas.

3.2.1 Maximum drawdown

MDD illustrates the maximum observed loss from a previous peak. It is described in percentages and is the minimum of the drawdowns observed over a period. This thesis incorporated the MDD in one of the ERC portfolios and as a performance measure for each portfolio. Even though data from the past does not predict future stock movements, it indicates potential downside risk over an established period. Equation 5 shows the formula for drawdowns and equation 6 shows the MDD.

$$DD_t = min(0, \frac{p_t - p_{max}}{p_{max}}) \tag{5}$$

$$MDD_t = min(DD_t) \tag{6}$$

MDD provides an insight into actual historical losses. The measure is a tangible view of risk as losses over a period solely determine its value. We calculated the MDD for each portfolio to generate a more detailed understanding of the performance of each strategy. We compared the MDD for the portfolios to the benchmark between 2009 and 2021.

3.2.2 Effective number of holdings

In order to understand why portfolios may have generated statistically significantly better returns or Sharpe ratios, we focused on the efficient number of holdings. We extracted the efficient number of holdings for each portfolio in each period. Equation 7 illustrates how to derive the efficient number of holdings for a portfolio.

$$\frac{1}{\sum w_i^2} \tag{7}$$

The efficient number of holdings displays the number of stocks that impact the strategies' return. Even if a portfolio was comprised of 48 stocks, it does not mean that all stocks contribute to the performance equally. A too heavy constituent concentration can become an issue when the portfolio is unevenly distributed, and only a few stocks affect the movements. For example, when using a GMV strategy, the priority for the investor is to finding the lowest volatility. However, while this might hold in-sample, it might not hold out-of-sample due to GMV's tendency to allocate high weights to only a few stocks with low volatility. The issue arises when one or a few of these stocks move drastically, and since those stocks have such high concentration, it will cause undesired volatility.

3.3 Sharpe ratio

The Sharpe ratio helps investors understand the returns compared to its risk (Sharpe 1998). It was first introduced in 1966 by William F. Sharpe and has become a frequently used and cited ratio. The Sharpe ratio measures a portfolio's performance (past or expected) and adjusts it for excess return, as shown in equation 8 (Sharpe 1998). The advantage of this ratio is its simplicity and the fact that it captures the relation between profits made and the additional risk. Nonetheless, there are several limitations, such as assuming normal distribution and potential manipulation from portfolio managers by changing the time interval. Therefore, two measures of the Sharpe ratio were obtained for each portfolio. Firstly, the average annual Sharpe ratio from each year studied was obtained. Subsequently, the annualized Sharpe ratio was derived using annualized return and volatility.

$$\frac{R_p - R_f}{\sigma_p} \tag{8}$$

The risk-free rate is a part of the Sharpe ratio and other portfolio measures. In this thesis, government bond yields 10 year period average was used as a proxy for the risk-free rate. The risk-free rate was composed of multiple 10-year yields in the euro area. The rate was sampled for the periods in which we tested the out-of-sample portfolios for the Sharpe ratios to be as authentic as possible. By matching the yields with the monthly returns, we generated Sharpe ratios for each year that were accurate as the risk-free rate for the same specific period is used. That accuracy helped the accuracy of the average Sharpe ratio measure for the portfolios as it was an average of annual Sharpe ratios. However, for the annualized Sharpe ratio, an average rate was used for the risk-free rate. In contrast to what DeMiguel et al. (2009) used, which was to derive a certainty equivalent return for each different portfolio, we used a longer bond yield which is common practice in portfolio research (Damodaran 2008).

3.3.1 Test statistic

In order to test whether the property measures are statistically discernible or not simply due to chance, we computed the p-values of the alpha and Sharpe ratios. The Sharpe ratios were tested similarly to Jobson & Korkie (1981) and DeMiguel et al. (2009), with a slight difference in the used risk-free rate. The optimized portfolios' returns were tested against the 1/N and MCaP returns to study Jensen's alpha.

We tested the significance of the Sharpe ratios from the optimized portfolios compared to the two benchmarks. In other words, just like Jensen's alpha, each portfolio was tested against the MCaP- and naive portfolios. We computed the z-score for the Sharpe ratio using the formula by Jobson & Korkie (1981) with the extension by Memmel (2003). The full formula for generating the test statistic is described in equation 9.

$$\hat{z} = \frac{\widehat{SR_1} - \widehat{SR_2}}{\sqrt{\frac{1}{n}[2(1-\hat{\rho}^2) + \frac{1}{2}\widehat{SR_1}^2 + \widehat{SR_2}^2 - 2\widehat{SR_1}\widehat{SR_2}\hat{\rho}]}} \tag{9}$$

Where ρ is the correlation between the returns of the optimized portfolio and the benchmark. Pearson's correlation was used for the correlation. $\widehat{SR_1}$ is the annualized Sharpe ratio from the tested portfolio, $\widehat{SR_2}$ is the annualized Sharpe ratio from the benchmark, and n is the sample size. $\hat{z}$ is the test statistic. Equation 9 was used twice for each portfolio to determine statistical differences in the out-of-sample Sharpe ratios, once for the MCaP portfolio and once for the 1/N portfolio. The significance level was set at 5% for the Sharpe ratio test.

3.4 Jensen's alpha

Jensen's alpha helps investors understand if their returns were above or below the one predicted by the CAPM. The measure is the difference between portfolio returns and market returns. The metric is commonly referred to as just alpha. Equation 10 displays the formula we ran the regression on for deriving the alpha.

$$R_{STRATEGY} - RF_t = \alpha_j + \beta_j[R_{BENCHMARK} - RF_t] + u_{jt} \tag{10}$$

Where alpha is defined as $R_{jt} - [RF_t + \beta_j(R_{MKT} - RF_t)]$. We ran a regression for each portfolio against both benchmarks as $R_{BENCHMARK}$ are the returns of the benchmarks. Two tests for each optimized portfolio imply that we had 22 outputs for Jensen's alpha, 11 for MCaP benchmark and 11 for the 1/N benchmark. β_j is the beta of each portfolio. Jensen's alpha was introduced to add another dimension and tested the risk-adjusted returns with another approach than the Sharpe ratio test, which is used in most similar studies. By doing so, we could further test whether our results possess statistical significance. The significance level was set at 5% for Jensen's alpha.

4 Data

Our dataset consists of percentage returns of stocks part of SX5T between 2006 and 2021, where each return represents a monthly gain or loss. We chose to use monthly data as we focused on long-term investments. Annualizing monthly data is one of the most conventional approaches in portfolio research. The weighting schemes for each portfolio will be composed of the individual constituents part of the SX5T index. The data for the constituents was retrieved from Refinitiv Eikon. All data was treated and analyzed using Python.

SX5T is a free-float MCaP index measuring the performance of supersector leaders in Belgium, Finland, France, Germany, Ireland, Italy, the Netherlands and Spain (Stoxx n.d.). The index selects securities annually, in September, with the cut-off review date being the last trading day of August. Two main reasons behind the choice of the index were its recognition and its coverage of major countries in Europe. The geographical advantage of the index helps to reduce biases associated with focusing on one-country markets. The SX5T is a net total return index called EURO STOXX 50 Net Total Return Index. The index includes the net proceeds from constituent dividends, which are reinvested after tax deduction. The net return index was selected because its alternatives, such as gross return indices, disregard taxes associated with dividends.

We also retrieved the long-term interest rates used as the risk-free rate. Government bonds or bonds from central banks are often viewed as risk-free. Rates from 10-year bonds were considered the risk-free rate used in this study to help derive portfolio measures. Therefore, the rate investors can obtain from government bonds can be used as the risk-free rate when analyzing portfolios. We chose to use yields on European government bonds as the SX5T index only contains public companies in Europe. Using a 10-year interest rate as a proxy for the risk-free rate is common practice within finance. When choosing risk-free rates, investors should aim to match the maturity period with the investment horizon (Mukherji 2011). In line with Damodaran (2008), we believe in the use of 10-year rates as the risk-free rate. Moreover, the equity risk premiums and default spreads are easy to derive for the 10-year rates.

5 Results

The portfolios shown in Table 1 were analyzed in Python (Spyder). In order to generate a deeper analysis, different sampling periods were attached to the various portfolios. The different strategies need to be compared over the same period to present reasonable results. The sampling periods varied from 1 to 3 years and all portfolios were compared from the start of 2009.

The various portfolios showed distinguishable features stemming from the different allocation strategies. Comparing the portfolios with each other might fail to bring useful insights due to the different aims of the portfolios. However, since investors have utility over terminal value and return, there is a need for comparison with the overall market and we have decided to focus on strategies with similar purpose. Our study aimed to understand if optimized portfolios can generate better performance compared to our two benchmarks. Table 2 presents the result for the out-of-sample portfolios over the period 2009-2021.

Table 2: Summary of Statistics

Portfolio	**Ann. return**	**Sharpe ratio(μ)**	**Ann. sharpe**	**Efficient holdings(μ)**	**Max DD**
GMV 1yr	10.41%	0.758	0.554	4.08	−30.98 %
GMV 2yr	12.20%	1.030	0.703	4.88	−21.37 %
GMV 3yr	12.05%	1.015	0.686	5.50	−22.36 %
1% GMV 1yr	10.58%	0.811	0.569	8.18	−27.58 %
1% GMV 2yr	13.34%	1.137	0.774	8.67	−20.80 %
1% GMV 3yr	12.52%	1.103	0.700	9.96	−22.17 %
ERC VOL 1yr	11.98%	0.927	0.669	27.85	−23.30 %
ERC VOL 2yr	13.35%	1.037	0.744	35.10	−22.27 %
ERC VOL 3yr	13.20%	1.051	0.725	38.97	−23.46 %
ERC MDD	13.36%	1.063	0.728	39.11	−23.92 %
ERC CVaR	14.19%	1.115	0.758	16.44	−23.47 %
MCaP	7.31%	0.552	0.402	35.04	−25.93 %
1/N	13.56%	0.967	0.667	48.00	−27.11 %

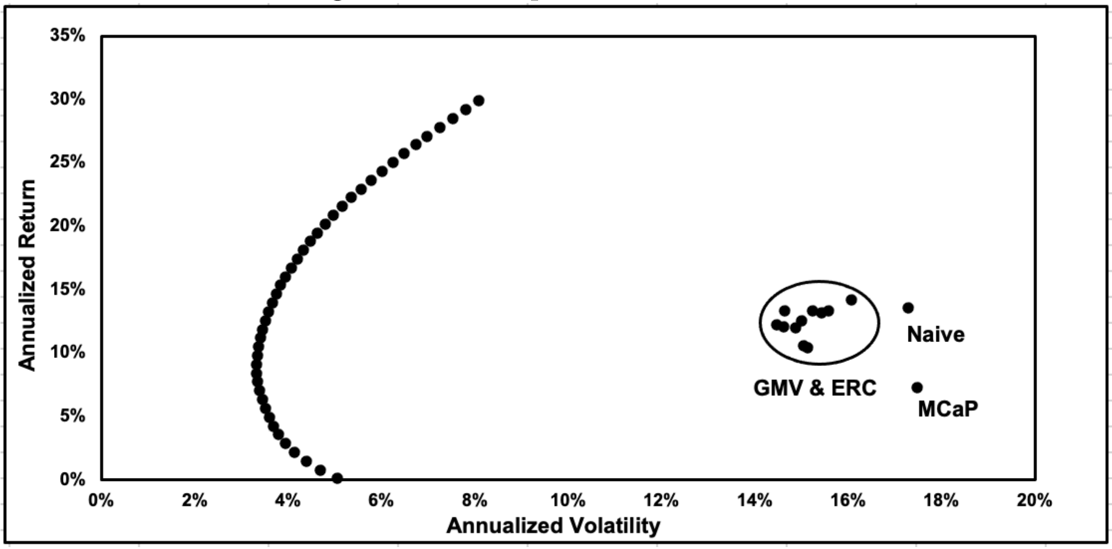
The annualized return for the benchmarks differed heavily as the 1/N portfolio strongly outperformed the MCaP. According to table 2, the out-of-sample 1/N portfolio showed an annualized return of 13.56% between 2009 and 2021. While there is no optimization associated with a 1/N portfolio, it obtained an average Sharpe ratio of 0.967 and an annualized Sharpe ratio of 0.77. As a result, the MCaP benchmark had a lower annualized return than every optimized portfolio.

The three non-constrained GMVPs varied heavily in terms of returns and risk-adjusted performance. Displayed in table 2, GMV - 1 year had the lowest annualized return, average Sharpe ratio, and annualized Sharpe ratio out of all optimized portfolios. One of the highest annualized returns was obtained by the 1% GMV with a 2-year sampling of volatility. All 6 different GMVPs outperformed the MCaP benchmark in annualized return but failed to beat the 1/N benchmark. Each ERC portfolio obtained a higher annualized return than the MCaP benchmark, but only the ERC-CVaR obtained a greater return than the 1/N benchmark. In terms of Sharpe ratios, each portfolio obtained a higher average than the MCaP. Only 3 optimized portfolios obtained a lower average Sharpe ratio than the 1/N benchmark, where the ERC-CVaR had the highest average Sharpe ratio of each portfolio studied. Each portfolio obtained a higher Sharpe ratio on an annualized basis compared with the MCaP benchmark. Furthermore, only 2 portfolios obtained a lower annualized Sharpe ratio than the 1/N portfolio. The GMV 1-year had the lowest annualized Sharpe ratio of 0.554, whereas the ERC-CVaR obtained the highest, 0.758.

As expected, there were large differences in the average efficient number of holdings between the portfolios. The non-constrained GMVPs had the lowest averages, whereas the ERC portfolios obtained the highest. 3 out of 5 ERC portfolios obtained a higher average efficient number of holdings than the MCaP benchmark. The ERC portfolio shares traits with the 1/N portfolio while enabling risk control. The GMVPs showed varying MDD over the whole period, with the largest MDD observed in the non-constrained GMVP with a sample size of one year. Consistent across the GMVPs is that smaller sample sizes imply a larger MDD. The 1% GMV obtained the lowest MDD with 2-year sampling. For the ERC portfolios, there were no large differences in terms of MDD. Interestingly, each ERC portfolio had a lower MDD than both benchmarks as the MCaP and 1/N obtained -25.93% and 27.11%.

Figure 2 displays the in-sample efficient frontier together with our optimized portfolios and benchmarks. The figure is based on annualized returns and volatility over the full period.

Figure 2: In-sample Efficient Frontier



The efficient frontier to the left in figure 2 presents the lowest possible volatility (x-axis) for every given return (y-axis) that was theoretically possible (in-sample) in our data set. The MCaP benchmark performs the worst in both categories and can be seen in the bottom right corner. Shown graphically, the optimized portfolio obtained lower annualized volatility than both benchmarks.

5.1 Statistical tests

When testing the excess returns for the portfolios against SX5T and the 1/N portfolio, the results varied between the portfolios. Table 3 presents the results of the regression analysis performed on the optimized portfolios against the index and 1/N benchmarks. These alphas are displayed monthly and cover the period 2009-2021. The results varied when testing the excess returns for the portfolios against the 1/N portfolio. Equation 10 shows the model used to run the regression.

Table 3: Regression Analysis - Jensen's Alpha

	Market Capitalization				Naive			
Portfolio	**Alpha**	**p-value**	**Beta**	**p-value**	**Alpha**	**p-value**	**Beta**	**p-value**
GMV 1 yr	0.0047	0.0519	0.6400	0.0000	0.0039	0.1514	0.8728	0.0000
GMV 2yr	0.0059	0.0053	0.6596	0.0000	0.0018	0.4736	0.9637	0.0000
GMV 3yr	0.0056	0.0061	0.6794	0.0000	0.0018	0.4544	0.9707	0.0000
1% GMV 1 yr	0.0039	0.0179	0.7632	0.0000	0.0021	0.2191	1.0443	0.0000
1% GMV 2yr	0.0060	0.0000	0.7624	0.0000	-0.0006	0.7185	1.0940	0.0000
1% GMV 3yr	0.0052	0.0002	0.7933	0.0000	0.0002	0.8826	1.0783	0.0000
ERC VOL 1yr	0.0046	0.0000	0.8157	0.0000	0.0001	0.8829	1.1302	0.0000
ERC VOL 2yr	0.0055	0.0000	0.8461	0.0000	-0.0009	0.1907	1.1177	0.0000
ERC VOL 3yr	0.0053	0.0000	0.8612	0.0000	-0.0007	0.2208	1.1090	0.0000
ERC MDD	0.0053	0.0000	0.8705	0.0000	-0.0008	0.2218	1.0973	0.0000
ERC CVAR	0.0060	0.0000	0.8751	0.0000	-0.0007	0.5925	1.0286	0.0000

$$R_{STRATEGY} - RF_t = \alpha_j + \beta_j[R_{BENCHMARK} - RF_t] + u_{jt}$$

As shown in table 3, alpha was positive for each of the portfolios against the MCaP. The right-hand side of the alpha shows the corresponding p-value for the alpha from each optimized portfolio. For each optimized portfolio, except GMV 1 year, the p-values were lower than 0.05. The results of the strategies compared with the 1/N benchmark demonstrate a different outcome; alphas were lower and for some portfolios negative. The corresponding p-values were in turn high, meaning that no statistical conclusions can be made. The significance level for Jensen's alpha was set at 5%. The beta was lower for our GMVPs relative to the other strategies for both the MCaP and 1/N benchmark. The p-values for all betas showed statistical significance, and all strategies displayed a beta of less than 1 for the MCaP, whereas only the non-constrained GMVPs had a beta of less than 1 compared to the 1/N portfolio.

Table 4 displays the Sharpe ratio test statistics for each of the optimized portfolios compared with the benchmarks. The Δ SR is the difference in the annualized Sharpe ratios between a portfolio and the benchmark. If the delta is negative, the benchmark had a higher annualized Sharpe ratio out-of-sample than the portfolio. The test statistic was obtained using equation 9. The associated p-values were derived from the test statistic and are displayed in table 4 below.

Table 4: Sharpe Ratio (SR) - Test

	Market Capitalization		Naive	
Portfolio	**Δ SR**	**p-value**	**Δ SR**	**p-value**
GMV 1yr	0.1524	0.289	-0.1126	0.337
GMV 2yr	0.3011	0.125	0.0361	0.444
GMV 3yr	0.2839	0.129	0.0189	0.470
1% GMV 1yr	0.1668	0.184	-0.0982	0.281
1% GMV 2yr	0.3718	0.025	0.1068	0.257
1% GMV 3yr	0.2977	0.041	0.0327	0.416
ERC VOL 1yr	0.2669	0.017	0.0019	0.493
ERC VOL 2yr	0.3416	0.002	0.0766	0.143
ERC VOL 3yr	0.3229	0.002	0.0579	0.183
ERC MDD	0.3257	0.002	0.0607	0.320
ERC CVAR	0.3555	0.0003	0.0905	0.089

$$\hat{z} = \frac{\widehat{SR}_1 - \widehat{SR}_2}{\sqrt{\frac{1}{n}[2(1-\hat{\rho}^2) + \frac{1}{2}\widehat{SR}_1^{\,2} + \widehat{SR}_2^{\,2} - 2\widehat{SR}_1\widehat{SR}_2\hat{\rho}]}}$$

Compared with the MCaP portfolio, the 1% GMV 2-year and ERC-CVaR portfolios had the highest Δ Sharpe ratio. The lowest Δ SR was obtained by the GMV 1 year portfolio. 7 out of 11 optimized portfolios obtained annualized Sharpe ratios superior to the MCaP benchmark and statistically significant. In other words, 4 out of 11 optimized portfolios failed to obtain Sharpe ratios that were superior to the MCaP benchmark and statistically significant. The 4 portfolios that failed to show superiority against the MCaP were all GMVPs, indicating that all ERC portfolios was statistical significance against the MCaP benchmark. The significance level for the Sharpe ratios was set at 5%.

Compared with the 1/N portfolio, the optimized portfolios were not superior. Only 2 optimized portfolios obtained a negative Δ SR where the GMV 1-year had the lowest Δ SR, -0.1126, and the 1% GMV 2yr had the highest. On the other hand, nine optimized portfolios showed positive Δ SR, but the associated p-values showed no signs of statistical significance.

6 Discussion

Over the full period, we find that our optimized portfolios lie deep within the efficient frontier shown in figure 2 (chapter 5). The figure was generated based on the portfolios' annualized returns and standard deviations. Over multiple years, it is intuitive that out-of-sample portfolios lie deep within the efficient frontier but differences compared with the benchmarks are of interest. Graphically, the 1/N benchmark is superior to the MCaP portfolio, in line with the findings of Schwartz (2000) who also found the MCaP to be inefficient. The 1/N benchmark obtained higher returns with similar risk to the MCaP benchmark. Each of our optimized portfolios is positioned on the two benchmarks' left-hand side, showing that the optimized portfolios obtained lower annualized volatility over the full period. 4 out of 5 ERC portfolios showed similar annualized returns compared with the 1/N benchmark. Nevertheless, they are all positioned on the left-hand side of the 1/N benchmark within the efficient frontier (2). This indicates that the ERCs may act as volatility controlling version of the 1/N portfolio, previously posited by (Choueifaty et al. 2013). Furthermore, as the efficient frontier shows the optimal portfolios for certain levels of risks; in-sample, the GMV portfolios would be closer to the left part of the efficient frontier. The true GMV, without constraints, would be optimal in-sample and thus be the left-most part of the hyperbola. The graphical difference between the actual performance of portfolios and in-sample performance is caused by the inability to predict the future. The difference in performance between in-sample and out-of-sample, which is also highlighted by Monarcha (2019), is visually and numerically clear.

Overall, the annualized returns varied considerably across the portfolios. The portfolio with the lowest annualized return, 10.41%, was the GMV with 1-year sampling. Apart from the MCaP, the GMV 1-year portfolio obtained the lowest average and annualized Sharpe ratios. A reason behind such performance may have been the heavy concentration in a few constituents. In fact, the portfolio obtained the lowest average efficient number of holdings and the most significant maximum drawdown. However, it obtained a higher annualized return and Sharpe ratio than the MCaP benchmark. The MCaP benchmark obtained an annualized Sharpe ratio of 0.402, whereas the 1/N obtained 0.667. The top-performing optimized portfolio was the ERC-CVaR which obtained an annualized return of 14.19%, sufficiently beating both benchmarks. The portfolio had an average Sharpe ratio of 1.115, more than two times the Sharpe ratio for the MCaP benchmark. The ERC-CVaR outperformed each optimized portfolio and benchmark in the annualized Sharpe ratio, obtaining 0.76. Obtaining an annualized Sharpe ratio well above 0.5 over a multiple-year period shows an ability to consistently strong performance. Furthermore, the ERC-CVaR suffered a maximum drawdown of -23.47% with an average efficient number of holdings of 16.44.

Looking at the efficient number of holdings, it is clear that each GMVPs suffered from heavy concentration, considering that the maximum was 48. The efficient number of holdings ranged from 4.08 to 9.96, implying that the portfolios were not diversified in their holdings and allocated large weights to a few constituents. The heavy concentration confirms one of the major limitations with the GMVPs, which is not being as low risk as one might think out of sample. As previously mentioned, GMVP's heavily allocation to a few stocks is inherent from the diversification method. Nonetheless, the efficient number of holdings increased as we incorporated the popular constraint of a minimum holding. As a result, each of the 1% GMV portfolios outperformed the MCaP benchmark in terms of annualized return and annualized Sharpe ratio. However, the out-performance of the optimized portfolio compared to the MCaP benchmark could stem from one main limitation of this study. Apart from leavers and joiners, the MCaP is a buy and hold strategy that will not suffer from transaction costs which our optimized portfolios will. As our optimized portfolios rebalance annually, the returns should be lowered if transaction costs are introduced.

We found that the GMV 1 year portfolios suffered the highest MDD between 2009-2021, however, the MDDs varied heavily amongst the portfolios. The 1% GMV 2 year portfolio obtained the lowest MDD. We see no clarity to either support or contradict the findings of Alankar et al. (2012) who found that volatility-based portfolios suffer more significant drawdowns compared with other risk measures. Although the efficient number of holdings and the MDD indicates higher risk in GMV than often perceived, we did find that the GMV portfolios had the lowest beta out of all optimized strategies, which potentially supports the case for using

GMV portfolios as a low-risk strategy.

6.1 Jensen's alpha

By analyzing the statistical significance of the returns over the entire sample, we found that returns for most of the optimized portfolios are superior to the MCaP benchmark. The returns for all but one GMV portfolio delivered statistically significant out-of-sample returns to the MCaP benchmark. Only the GMV with 1-year sampling failed to show significance in the returns compared to the MCaP benchmark. The portfolio had alpha against the MCaP benchmark, but the p-value was above the threshold and will thus not lead to any conclusion. The GMVPs with volatility sampling of 2- and 3-years were able to show significantly different and better returns over the benchmark. Each GMVP with a 1% delivered significantly better returns than the MCaP out-of-sample.

Our findings regarding the superiority of the GMV to the MCaP benchmark match the findings of Haugen & Baker (1991). They found that the GMV consistently outperforms the MCaP portfolio in both returns and risk, which we confirm as 5 out of 6 GMVPs delivered alpha that was statistically significant to the benchmark. Portfolios such as the GMV have faced criticism for not incorporating investors' true utility as it allocates solely based on volatility. Our findings partly contradict such criticism as our GMVPs evidently outperform the MCaP. Our findings are also supported by Jagannathan & Ma (2003) who found GMVPs to have superior risk-adjusted returns to the MCaP benchmark.

The ERC portfolios all delivered alphas compared to the MCaP benchmark that were statistically significant. Thus, there was statistical proof that each of the five ERC strategies delivered significantly better out-of-sample returns than the MCaP benchmark. The portfolios all build on different risk assumptions and hold different weights, which indicates that the method of diversification is attractive. However, the outperformance is perhaps attributable to the poor performance of the MCaP benchmark.

In line with Jagannathan & Ma (2003) and Haugen & Baker (1991), we find that the MCaP portfolio is inefficient as all but one of the optimized portfolios was able to obtain a Jensen's alpha that was statistically significant. That 10 out of 11 portfolios show significance exemplifies the inefficiency of having a MCaP allocation for indices, supported by the benchmarks location deep inside the efficient frontier. Furthermore, as previously mentioned, the true performance of the optimized portfolios will be lower if transaction costs are measured.

Compared with the 1/N portfolio, we did not find that the optimized portfolios delivering superior Jensen's alphas were statistically significant. These findings are in line with DeMiguel et al. (2009) who compared risk-adjusted returns between optimized portfolios and the 1/N benchmark. They posit that the equally weighted portfolio might be an efficient way of diversification. In terms of Jensen's alpha, we find no proof of the contrary. Furthermore, our findings on Jensen's alpha contradict the results of Kritzman et al. (2010). They suggested that optimized portfolios be superior to the 1/N benchmark in risk-adjusted returns. Our findings regarding the overall inability to outperform the 1/N benchmark partly contradict the findings of Duchin & Levy (2009) as our index had 48 constituents. Duchin & Levy found that the 1/N portfolio obtained no superiority over MVO portfolios when the number of constituents was above 30. Yet, the 1/N benchmark had a higher annualized return than 10 out of 11 optimized portfolios.

6.2 Sharpe ratios

Compared to the MCaP benchmark, the Sharpe ratios were statistically significant for 7 out of 11 optimized portfolios. Hence, 4 portfolios lacked superiority over the MCaP portfolio in annualized Sharpe ratios. Each non-constrained GMVP was unable to outperform the MCaP benchmark according to the Sharpe ratio test. Interestingly, the annualized Sharpe ratios for 2 of 6 tested GMV portfolios significantly outperformed the MCaP benchmark. In other words, 4 of our 6 GMVPs failed to show a Sharpe ratio that was statistically significant. However, combining the superior annualized and average Sharpe ratios with 7 out of 11 portfolios showing statistical significance prove that the MCaP portfolio is inefficient. The overarching outperformance of our optimized portfolios to the MCaP out-of-sample indicates the inefficiency of the MCaP portfolio, matching the findings of Haugen & Baker (1991) and Jagannathan & Ma (2003). As previously stated, this is also displayed graphically in figure 2. Multiple previous studies also share the view of the MCaP being inefficient (Sharpe 1991, Markowitz 2005, Arnott et al. 2005).

Each ERC portfolio showed significantly superior Sharpe ratios compared with the MCaP. Hence, each portfolio aiming to allocate such that each holding contributes the same amount of risk showed superiority in both tested measures of risk-adjusted returns. Whereas volatility is a criticized risk measure, each of the three ERC-vol portfolios showed superior significance over the MCaP. Furthermore, the ERC portfolios obtained higher annualized Sharpe ratios than the MCaP as the Δ SRs ranged between 0.267 and 0.356 against the MCaP benchmark. These findings match the result of Choueifaty et al. (2013) who also found superior Sharpe ratios in the ERC portfolio compared to the MCaP.

Our findings on the performance compared with the MCaP benchmark do not transcend the results of the naive portfolio. None of the portfolios obtained a superior annualized Sharpe ratio that was statistically significant. In other words, each GMV portfolio failed to show superior out-of-sample Sharpe ratios to the 1/N portfolio that were statistically significant. The naive portfolio's annualized out-of-sample Sharpe ratio was 0.667 between 2009-2021. The annualized Sharpe ratio for the GMVPs varied heavily. The GMVPs sampling over 2- and 3-years had a higher annualized Sharpe ratio than the 1/N portfolio. The GMVPs with a lower annualized Sharpe ratio than the 1/N portfolio allocated based on a 1-year sampling of the volatility. The same relationship holds for the average Sharpe ratios. Moreover, as Kritzman et al. (2010) found superiority against the 1/N benchmark for GMVPs, our findings are opposing. 4 of 6 GMVPs only showed slight superiority in terms of annualized and average Sharpe ratios to the 1/N, yet none showed statistical significance. These findings imply matching results to DeMiguel et al. (2009) as none of the GMVPs found a superior annualized Sharpe ratio compared with the 1/N benchmark, which was statistically significant. The annualized and average Sharpe ratios for the 1/N portfolio, 0.967 and 0.667, are signs of a well-performing investment strategy and imply that the portfolio may be a viable alternative, although one must consider the booming stock market over the last 12 years.

Thus, each optimized portfolio allocating based on a 12-month window of the volatility had a lower average Sharpe ratio out-of-sample compared with the 1/N portfolio. The ERC-vol 1 year was the only ERC portfolio with a lower average Sharpe ratio than the 1/N benchmark. However, it still obtained a slightly higher annualized Sharpe ratio. None of the ERC portfolios had a lower annualized Sharpe ratio than the 1/N benchmark. Interestingly, in terms of average and annualized Sharpe ratios, there was also an outperformance by the two remaining ERC-vol portfolios, the ERC-MDD and the ERC-CVaR. Compared with the 1/N, the portfolios obtained Δ SRs of 0.061 and 0.091, respectively. Even though every ERC portfolio showed a positive Δ SR to the 1/N benchmark, the statistical tests showed that none obtained statistical significance. Thus, the indications from the Δ SRs cannot be proven statistically, and we share our findings with DeMiguel et al. (2009). Our findings from the Sharpe ratio test contradict the findings of Choueifaty et al. (2013). None of our optimized portfolios obtained a low enough p-value in the Sharpe ratio tests compared with the 1/N. Conclusively, our optimized portfolios cannot deliver superior Sharpe ratios to the 1/N benchmark that are statistically significant.

7 Conclusion

This paper examined the out-of-sample performance of 11 different optimized portfolios. We included 6 different Global Minimum Variance portfolios and 5 different Equal Risk Contribution portfolios. The out-of-sample results were compared with two benchmarks, the market-weighted- and naive portfolios. The portfolios and benchmarks are rebalanced on an annual basis. In order to compare the out-of-sample performance between the optimized portfolios and benchmarks, we examined certain performance measures of each portfolio. The annualized returns, average Sharpe ratios, annualized Sharpe ratios, average efficient number of holdings, and maximum drawdowns were obtained as overarching portfolio measures. Subsequently, to statistically distinguish the out-of-sample performance, we focused on out-of-sample Jensen's alpha and annualized Sharpe ratios. In other words, we focused on two different risk-adjusted return measures to draw our conclusions. We obtained and tested Jensen's alpha in accordance with Jensen (1968) by examining out-of-sample returns and betas. To the Sharpe ratios, we used the test statistic constructed by the Jobson & Korkie (1981) with the extension from Memmel (2003). Jensen's alpha and the Sharpe ratio test statistics were performed for the optimized portfolios against both benchmarks. Our main purpose was to understand if our optimized portfolios can deliver superior out-of-sample risk-adjusted returns compared to the two benchmarks and, by extension, understand the efficiency of the benchmarks.

Ten of our 11 optimized portfolios exhibit a positive Jensen's alpha that is statistically significant compared to the market-weighted portfolio. The only portfolio that is not statistically different from the benchmark was the Global Minimum Variance with a 1-year sampling of volatility. The findings on the out-of-sample Sharpe ratios indicates a similar relationship of out-performance compared with the market-weighted benchmark. Each portfolio has a higher average and annualized Sharpe ratio to the benchmark, and 7 out of 11 showed statistical significance. Thus, 7 out of 11 portfolios are statistically superior to the market-weighted benchmark in Jensen's alpha and Sharpe ratio, and 3 out of 11 portfolios shows superiority that is statistically significant in Jensen's alpha but not in the Sharpe ratio. Evidently, we can conclude that our optimized portfolios deliver better risk-adjusted returns than the market-weighted portfolio.

The naive portfolio shows strong out-of-sample performance. In terms of Jensen's alpha, none of our portfolios obtains a Jensen's alpha over the naive portfolio that is statistically significant. The same relationship holds for the Sharpe ratio test. None of the optimized portfolios deliver a superior Sharpe ratio that is statistically significant. In other words, no portfolio can statistically outperform the naive benchmark. Thus, we can unequivocally conclude that we find no superior risk-adjusted returns in our optimized portfolios compared with the naive portfolio.

7.1 Future research

Markets are ever-evolving, and our results solely apply to the market conditions of our dataset. Our findings are on the European market and the constituents part of Eurostoxx 50. We implore future research in other markets such as the U.S or Asian markets. Moreover, our data stretched between 2006-2021, with the majority of the years spent in an economic boom. Perhaps future research can include periods of recession to understand the performance of the optimized portfolios tested in this thesis further.

We firmly believe there is a gap to fill regarding research on the portfolios available today. For example, the various risk measurements open up numerous different allocation schemes. Our study indicates that the Equal Risk Contribution portfolios are superior to the other tested portfolios, yet the portfolios are not heavily researched. There are also different ways to measure performance. Today, funds and mutual funds are evaluated by more measures than alpha and Sharpe ratios. For instance, the performance can also be measured in Sortino - and Treynor ratios.

References

Alankar, A., DePalma, M. & Scholes, M. (2012), 'The introduction into tail risk parity', *preprint, AlllianceBernstein research* .

Arnott, R. D., Hsu, J. & Moore, P. (2005), 'Fundamental indexation', *Financial Analysts Journal* **61**(2), 83–99.

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

Choueifaty, Y., Froidure, T. & Reynier, J. (2013), 'Properties of the most diversified portfolio', *Journal of investment strategies* **2**(2), 49–70.

Damodaran, A. (2008), 'What is the riskfree rate? a search for the basic building block', *A Search for the Basic Building Block (December 14, 2008)* .

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

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

Glosten, L., Nallareddy, S. & Zou, Y. (2021), 'Etf activity and informational efficiency of underlying securities', *Management Science* **67**(1), 22–47.

Haugen, R. A. & Baker, N. L. (1991), 'The efficient market inefficiency of capitalization–weighted stock portfolios', *The journal of portfolio management* **17**(3), 35–40.

Jagannathan, R. & Ma, T. (2003), 'Risk reduction in large portfolios: Why imposing the wrong constraints helps', *The Journal of Finance* **58**(4), 1651–1683.

Jensen, M. C. (1968), 'The performance of mutual funds in the period 1945-1964', *The 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* pp. 889–908.

Kempf, A. & Memmel, C. (2006), 'Estimating the global minimum variance portfolio', *Schmalenbach Business Review* **58**(4), 332–348.

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

Maillard, S., Roncalli, T. & Teïletche, J. (2010), 'The properties of equally weighted risk contribution portfolios', *The Journal of Portfolio Management* **36**(4), 60–70.

Markowitz, H. (1952), 'Portfolio selection*', *The Journal of Finance* **7**(1), 77–91. **URL:** *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1952.tb01525.x*

Markowitz, H. M. (2005), 'Market efficiency: A theoretical distinction and so what?', *Financial Analysts Journal* **61**(5), 17–30.

Mausser, H. & Romanko, O. (2014), 'Computing equal risk contribution portfolios', *IBM Journal of Research and Development* **58**(4), 5–1.

Memmel, C. (2003), 'Performance hypothesis testing with the sharpe ratio', *Available at SSRN 412588* .

Merton, R. C. (1980), 'On estimating the expected return on the market: An exploratory investigation', *Journal of financial economics* **8**(4), 323–361.

Monarcha, G. (2019), 'An introduction to alternative risk premia'.

Mukherji, S. (2011), 'The capital asset pricing model's risk-free rate', *The International Journal of Business and Finance Research* **5**(2), 75–83.

Novy-Marx, R. & Velikov, M. (2016), 'A taxonomy of anomalies and their trading costs', *The Review of Financial Studies* **29**(1), 104–147.

Otuteye, E. & Siddiquee, M. (2020), 'Underperformance of actively managed portfolios: some behavioral insights', *Journal of Behavioral Finance* **21**(3), 284–300.

Patton, A. J. & Weller, B. M. (2020), 'What you see is not what you get: The costs of trading market anomalies', *Journal of Financial Economics* **137**(2), 515–549.

Santos, A. A. P. (2010), 'The out-of-sample performance of robust portfolio optimization', *Revista Brasileira de Finanças* **8**(2), 141–166.

Schwartz, T. (2000), 'How to beat the s&p 500 with portfolio optimization', *DePauls University* pp. 1–24.

Sharpe, W. F. (1991), 'The arithmetic of active management', *Financial Analysts Journal* **47**(1), 7–9.

Sharpe, W. F. (1998), 'The sharpe ratio', *Streetwise–the Best of the Journal of Portfolio Management* pp. 169–185.

Statman, M. (2013), 'Invited editorial comment: Is markowitz wrong? investment lessons from the financial crisis'.

Stoxx (n.d.), 'Euro stoxx 50®'.
URL: *https://www.stoxx.com/index-details?symbol=SX5E*

Xiong, J. X. & Idzorek, T. M. (2011), 'The impact of skewness and fat tails on the asset allocation decision', *Financial Analysts Journal* **67**(2), 23–35.