

NHH



Norwegian School of Economics

Bergen, Fall 2017

Will a factor-based Markowitz implementation beat the market?

Anders Bredenkjær Bjørnson & Fredrik Snarvold Gjerde

Supervisor: Francisco Santos

Master Thesis

MSc Economics and Business Administration, Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Abstract

In this thesis, we look at whether a factor-based implementation of the Markowitz (1952) framework beats the market. Different sets of factors are used in the framework, to see whether this choice affects the results. We find the optimal portfolio by using past data to invest in the future. This is implemented, with annual rebalancing, on US stock data from 1964 to 2016. Portfolios are formed on daily, weekly and monthly data to see whether the frequency of the return measurement gives different results.

Our framework begins by running factor regressions on each stock. These coefficients are used in combination with the expected returns and the covariance matrix of the factors to calculate the expected returns and covariance matrix of the stocks. Performance is evaluated by looking at the mean one/four-year *ex-post* Sharpe- and appraisal ratios, as well as the crash-risk of our portfolios.

Based upon the mean one-year Sharpe ratios, our factor-based portfolios formed on daily data all significantly beat the market, while there is no such significance for the portfolios formed on weekly or monthly data. However, all of our factor-based portfolios have a statistically significant one/four-year appraisal ratio when using the market as our benchmark. We also find that there does not seem to be a higher crash-risk for our portfolios than the market.

Our portfolios, when formed on the same measurement frequency, perform remarkably similar. This shows that our framework is an approximation of the standard Markowitz (1952) theory, rather than a factor-based investment strategy. The only thing not accounted for in our framework is the stock covariance explained by the residuals left in the regressions. We also find that this is a more stable implementation of the Markowitz theory, avoiding non-positive definite covariance matrices.

Preface

This Master thesis is written as part of our Master of Science in Economics and Business administration, with a major in Financial Economics. Our motivation for this thesis came from learning about the concepts we use in the “Investments” course at NHH, taught by Associate Professor Francisco Santos.

We would like to thank Francisco Santos, both for his teachings in the aforementioned course, and especially for his work as our supervisor on this thesis. During our work, he has been an invaluable source of feedback and constructive criticism. When we have encountered problems, he has always been available for guidance, so that we have managed to overcome them. We are certain that this would be a significantly worse paper without his help.

We would also like to acknowledge the contributions of Professor Nils Friewald, whose course “Programming with Applications in Finance” gave us the requisite programming skills to actually do the analyses contained within this thesis. In addition, we would like to thank the R-community at large. The large amount of information made available through this community has helped us overcome problems in our code, and also to find more efficient ways of implementing it.

Contents

1	Introduction	1
2	Creating portfolios	4
2.1	Theory of portfolio optimisation using the Markowitz framework	4
2.2	Calculating portfolio return series	5
2.2.1	Calculating covariance matrix and stock expected returns	6
2.2.2	Calculating stock weightings	8
2.3	Factor models	11
2.3.1	The Capital Asset Pricing Model	11
2.3.2	The Single Index Model	12
2.3.3	Fama-French three-factor model	12
2.3.4	The Carhart four-factor model	13
2.3.5	Fama-French five-factor model	13
2.3.6	Combination of value and momentum	14
2.4	Creating comparison portfolios	15
2.5	Overview of our portfolios	16
3	Evaluating portfolio returns	17
3.1	Performance measures	17
3.1.1	Sharpe- and appraisal ratios	17
3.1.2	Bootstrap for the Sharpe and appraisal ratios	18
3.2	Skewness and kurtosis as measures of crash-risk	19
3.3	Calculating factor loadings	19
4	Data	21
4.1	Stock returns	21
4.1.1	Data management	21
4.1.2	Creating weekly data	24
4.2	Factor returns	24

5	Results and analysis	25
5.1	Summary of portfolio returns	25
5.2	Periodical Sharpe ratio results	27
5.2.1	Daily data	27
5.2.2	Weekly data	28
5.2.3	Monthly data	29
5.2.4	Analysis of the Sharpe-ratio results	31
5.3	Periodical appraisal ratio results	32
5.3.1	Daily data	32
5.3.2	Weekly data	33
5.3.3	Monthly data	34
5.3.4	Analysis of appraisal ratio results	35
5.4	Measuring crash-risk in the portfolios	37
5.4.1	Skewness and kurtosis	38
5.4.2	Crash-risk in the worst times	39
5.4.3	Overall crash-risk	41
5.5	Factor loadings in portfolios	41
5.6	Why are the portfolio performances so similar?	43
5.6.1	Which approximation is better?	44
6	Conclusion	45
7	Bibliography	47

1 Introduction

In this thesis, we will combine a variety of factor models, such as the Fama-French models, with the Markowitz framework. As far as we can tell, factor models have been used more as a replacement to the Markowitz method, rather than in combination with it. We therefore find it interesting to look at how these two concepts would work when combined. The analysis is based on daily, weekly and monthly US stock data from 1964 to 2016. Specifically, we look at whether investment portfolios formed on different factors using a customised Markowitz implementation beats the market. In addition to the market, we will also compare our factor-based portfolios with a $\frac{1}{N}$ portfolio.

The portfolios are based upon established factor models, such as the Fama-French (1993, 2015) models and the Carhart (1997) model, and also some modifications of these. We evaluate portfolio performances by their *ex-post* Sharpe- and appraisal ratios, as well as their crash-risk. Our hypothesis is that these factor-based portfolios should beat the market, as they allow for investing in stocks with positive α 's, and can take advantage of beneficial correlations.

To create our factor-based portfolios, we use past data to invest in the future, and rebalance the portfolios annually. That is to say, for daily and weekly data, we use returns from 1995 to calculate the stock weightings in 1996. For monthly data, because of the low frequency of return observations, we use five years of previous data to calculate these weights.

We run regressions on past returns for each stock, using a set of factors as explanatory variables. For a stock to be included in that year's portfolio, we demand a certain number of return observations in the past data, which we call the observation rate. This is done to improve the regressions, by only including stocks with a sufficient sample size, giving better coefficient estimates.

To calculate the expected returns and covariance matrix of the relevant stocks, we use the regression coefficients in combination with the expected returns and covariance matrix of the factors. The calculation of these inputs is where our framework deviates from the standard Markowitz framework. The inputs are used in the same way as they would be in a standard Markowitz implementation to calculate the stock weightings. This gives us the portfolio with the highest *ex-ante* Sharpe ratio. We have in this implementation included some constraints, the most significant of which, is that we do not allow shorting in our portfolios. Once the weight of each stock has been found, we invest in these stocks for the following year.

The primary performance measure is the one/four-year *ex-post* Sharpe ratio. One Sharpe ratio is calculated for each calendar year for the daily and weekly portfolio returns, while we use four years for the monthly data. To increase the sample size, we use the bootstrap method to produce a sample of 10 000 means of this data.

We find that our factor-based portfolios formed on daily data all perform significantly better than the market on the 5% level, while one of them also beats the $\frac{1}{N}$ portfolio on this significance level. For the weekly and monthly factor-based portfolios there are no significant differences in performance compared to the market and the $\frac{1}{N}$ portfolio.

To further evaluate our portfolios, we calculate a sample of 10 000 one/four-year *ex-post* appraisal ratio means. This was done in the same way as for the Sharpe ratios. We find that all of our factor-based portfolios have significant appraisal ratios by this measure, thereby beating the market on the 5% level.

To assess the most extreme risks in the portfolios we look at their risk of crashing, by analysing the worst returns and years. We also look at their skewness and kurtosis measures, and for the portfolios formed on daily and weekly data, we evaluate these measures in the worst periods. From these analyses, we find that while our factor-based portfolios sometimes do worse than the market, there does not seem to be a larger crash-risk overall.

Based on our analyses, we find that all of our factor-based portfolios have remarkably similar performances, and share many of their best and worst times. To confirm their similarity, and to explain why some of them beat the market, we look at their factor loadings. What we find, confirm their similarities, as the differences in factor loadings are insignificant. There are also significant loadings for other factors than the market, and we find that these factors have a negative correlation with the market. This helps us explain why the factor-based portfolios formed on daily data beat the market, as these beneficial correlations allows the portfolios to increase returns and/or reduce volatility.

Based upon the similarity of our factor-based portfolio performances, we more closely examine our calculations, and find that what is being done is not factor investments per se. Rather, what we end up with is a more stable approximation of the standard Markowitz framework. The only difference in inputs for our framework, is that we do not account for the share of the stock-covariances that would be explained by the residuals of the regressions we run on the stocks. While this means that it is not a perfect approximation of Markowitz, we do avoid the biggest problems, such as the risk of a non-positive definite covariance matrix. We believe the biggest reason for this is the reduction of estimates that are required.

As we never experience a non-positive definite covariance matrix in our analysis, we conclude that improving the covariance matrix is more important than further lowering estimation numbers. Therefore, we would recommend using all available factors when utilising our framework, particularly if it is used as a theoretical basis. As the basis for an investment portfolio however, there are no significant differences between the various combinations of factors.

Based on our findings, we would say that our biggest contribution to existing literature, is the fact that our framework is a more stable way of implementing the Markowitz theory. In

addition, we find that investing based on this framework gives a significantly higher Sharpe ratio than the market portfolio when using daily data. When evaluated by periodical appraisal ratio, this framework also beats the market when using weekly and monthly data. We also find that there does not seem to be a higher crash-risk for portfolios formed on our framework.

This thesis is structured as follows. Section 2 describes in detail how we create our portfolios, and the theoretical basis that we use. Section 3 describes how we evaluate the portfolio returns. Section 4 details what data we use for our thesis, and what we do with this data before we start the analysis. We then present and analyse our results in section 5. We end the thesis with concluding remarks in section 6.

2 Creating portfolios

In this section we present in detail how we create our portfolios, as well as the established theory that we use as a basis for implementing our custom Markowitz framework. We start by introducing the standard Markowitz (1952) theory. The next step is the creation of our factor-based portfolios, where the calculations of the expected returns, covariance matrices and weights of the stocks are described. We then describe the established factor models that we use as a basis for our portfolios, and continue by describing our comparison portfolios. Finally, we present the overview of the portfolios used in this thesis.

2.1 Theory of portfolio optimisation using the Markowitz framework

As we use a customised Markowitz framework in our thesis we find it useful to present the standard framework, to be able to explain where we diverge from it. The Markowitz (1952) theory of optimal portfolio assumes that an investor is risk averse, thus an investor will take on increased risk only if compensated with higher expected return. Further an investor places his funds in the security with the greatest discounted value (Markowitz, 1952). The expected return of a portfolio is shown in equation (1) where R_p is the return on the portfolio, R_i is the return on asset i and w_i is the corresponding weight of the asset.

$$E(R_p) = \sum_{i=1}^{\infty} w_i E(R_i) \quad (1)$$

$$\sigma_p^2 = \sum_{i=1}^{\infty} w_i \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j \rho_{ij} \quad (2)$$

Equation (2) is the variance of the portfolio where σ_i is the sample standard deviation of the periodic returns on an asset and ρ_{ij} is the correlation coefficient between the returns on assets i and j . Based on equation (2) an investor can reduce the portfolio risk (variance of returns) by holding combinations of securities that are not perfectly positively correlated, $\rho_{ij} = 1$ (Markowitz, 1952). Normally this calculation would be done using a matrix containing the individual variances, as well as asset covariances.

Figure 1 shows the basics behind the Markowitz theory with expected return versus standard deviation. Every possible combination of risky assets can be plotted in the model and the set of all possible portfolios defines a region, bounded by the hyperbola (Markowitz, 1952). The upper edge of this region is the efficient frontier and a portfolio lying on this line represents the combination offering the best possible expected return for a given risk level. The optimal portfolio is located where a line, often referred to as the Capital Allocation Line (CAL), from

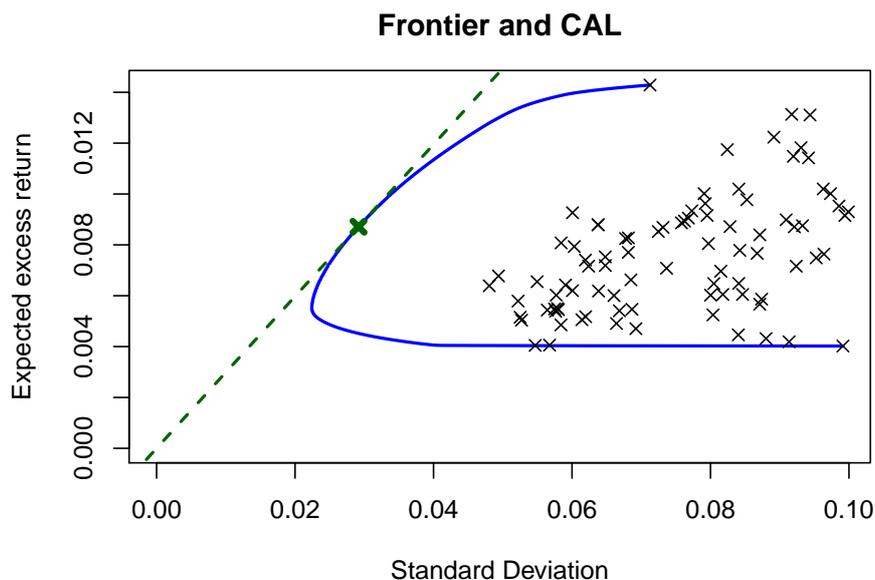


Figure 1: Figure of the Frontier and Capital Allocation Line (CAL). The black x's are plotted individual stocks and the blue line represents the frontier. The dashed green line is the CAL while the green x is the tangent portfolio. The figure is based on the stocks which have return observations every month in our monthly data set from 1964 through 2016. This has been calculated using excess returns, meaning that the CAL goes through the origin. To calculate the frontier and CAL we use the six-factor model, which is described in table 2, and base the optimisation on the entire data set.

the intercept defined by the risk-free rate is tangent to the efficient frontier. This is why the optimal portfolio is sometimes referred to as the tangent portfolio. In figure 1 we have used excess returns, resulting in the intercept being at the origin. This tangent portfolio will be mean-variance-efficient, and have the highest Sharpe ratio of any possible investment. The value of the Sharpe ratio is given by the slope of the CAL, and this ratio is further explained in section 3.1.1.

To calculate the expected returns of stocks, used in equation (1), one would normally use the average return over the period available. Similarly, to calculate the covariance matrix, one would normally use the standard variance and covariance calculations. Our divergence from the Markowitz theory is that we use a different way to calculate these inputs. Exactly how we do this is described in section 2.2.1.

2.2 Calculating portfolio return series

We create our set of factor-based portfolios using a customised Markowitz implementation in conjunction with a variety of factor models, which are described in section 2.3. By using past stock returns in this framework, we find the theoretically optimal portfolio with the highest *ex-ante* Sharpe ratio. The portfolios are rebalanced at the start of every year, and for this reason,

we identify the optimal portfolio for each year in our return period. We have decided to only rebalance our portfolios once a year, as we believe this will prevent transactional costs from significantly affecting our results. For daily and weekly data, we use one year of previous observations to calculate the optimal portfolio, whilst for monthly data we use 5 years of observations. We then invest in this optimal portfolio the following year. The reason for using a greater time period for monthly data is the low observation frequency. We need a significant number of observations to be able to accurately calculate the optimal portfolio.

As we use regressions as part of our portfolio weight calculations, there could still be a problem with a low number of observations. If there is not a sufficient number of observations in the relevant period, it would lead to less accurate estimates of the regression coefficients. To alleviate this, and improve the quality of the regressions, we define a parameter we call *observation rate*. This is defined as the number of available return observations in the period, divided by the theoretical maximum number of observations. That is to say, if the period on which we base the regression has 60 months (5 years), there are theoretically 60 possible observations. However, there might be some months for which specific stocks do not have a return observations, for example if it was only recently listed. By applying a required observation rate, we demand that the stocks have a certain number of return observations in the regression data, which should improve the regression estimates.

Given the low number of observations on which we base our regressions for the weekly (52) and monthly data (60), we have decided to require an observation rate of 75% for the portfolios formed on these frequencies. However, the daily data has a comparatively high number of observations (250), and therefore we demand an observation rate of 50% for these portfolios. By applying these requirements, we are certain that there is a sufficient number of observations to be able to run regressions on the sample, and this is likely to give better coefficient estimates. In addition, we believe it is reasonable for an investor to demand a certain number of observations before investing in an asset. A second requirement is that the stocks has to have an observed return in the last observation period of the regression data. This is done to avoid investing in delisted stocks.

We have complete data for the stocks in our data set and our factors from 1964 to 2016. From this, the portfolios formed on daily and weekly data get realised returns from 1965 to 2016, while the portfolios formed on monthly data get realised returns from 1969 to 2016.

2.2.1 Calculating covariance matrix and stock expected returns

As our portfolios are rebalanced at the start of each year, we need to identify the covariance matrix, and the expected returns of the stocks once a year. This is therefore done 52 times for the portfolios formed on daily and weekly data, and 48 times for those formed on monthly data.

This process is done in an identical fashion each year, with minor differences between the measurement frequencies and the various factor models. Therefore we will only demonstrate one calculation. The principle differences to keep in mind is that the portfolios formed on monthly data use five years of previous returns in the calculation, and that the various factor models have different sets of factors.

The first step in our estimation of the covariance matrix and expected returns is to run a regression for each of the stocks as a function of the relevant factors, using the last year of return data. The regressions take the form of equation (3), where i is the individual stocks and j refers to individual factors. β_{ij} refers to factor j 's coefficient on stock i , while ε_{it} is the residual for stock i at time t . From these regressions, we extract the α 's and the factor-coefficients, as well as the variance of the residual from each regression.

$$r_{it} = \alpha_i + \sum \beta_{ij} r_{jt} + \varepsilon_{it} \quad (3)$$

We calculate the expected return of the different stocks, using the α -values, factor-coefficients and the expected return of the factors. This calculation is shown in equation (4), with n being the number of stocks, and m being the number of factors.

$$\begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_n \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} + \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_m \end{bmatrix} \cdot \begin{bmatrix} \beta_{11} & \beta_{21} & \cdots & \beta_{n1} \\ \beta_{12} & \beta_{22} & \cdots & \beta_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{1m} & \beta_{2m} & \cdots & \beta_{nm} \end{bmatrix} \quad (4)$$

Using the factor covariance matrix, as well as the regression coefficients we calculate the covariance matrix of the stocks. To make the variance of the individual stocks correct in this matrix, we add the residual variance to the primary diagonal. The calculation of this covariance matrix can be seen in equations (5) and (6).

$$\Sigma_s = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nm} \end{bmatrix} \times \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_m^2 \end{bmatrix} \times \begin{bmatrix} \beta_{11} & \beta_{21} & \cdots & \beta_{n1} \\ \beta_{12} & \beta_{22} & \cdots & \beta_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{1m} & \beta_{2m} & \cdots & \beta_{nm} \end{bmatrix} = \mathbf{B}\Sigma_f\mathbf{B}^T \quad (5)$$

$$diag(\Sigma_s) = diag(\Sigma_s) + \sigma(\varepsilon_i) \quad (6)$$

This is where our implementation of the Markowitz framework differs from the standard implementation. Because we base our calculations on estimates from the regression by using

different factors instead of stocks, we create fewer estimates than a standard Markowitz framework. These estimates are then used to calculate the covariance matrix and expected returns of the stocks. An example of the reduction in number of estimates can be seen in table 1. One of the difficulties with the standard Markowitz implementation is the large amount of estimates needed. This high number of estimates can lead to errors in the calculation of the covariance matrix, which potentially produce impossible results (Bodie, Kane & Marcus, 2014). We believe that the reduced number of estimates in our modified model should reduce the problem of getting impossible results considerably because the sensitivity to errors is reduced. Having run a total of 1 520 estimations using our custom Markowitz implementation, none of these had a non-positive definite covariance matrix, while this was a big problem when attempting to run a standard Markowitz implementation.

	Custom Markowitz	Standard Markowitz
μ estimates	6	1000
Variance estimates	1006	1000
Covariance estimates	15	499500
Coefficient estimates	7000	0
Total estimates	8027	501500

Table 1: *Difference in number of estimates between our Markowitz implementation, and the standard implementation. This example is based on a theoretical set of 1000 stocks, using 6 factors, which is the highest number of factors we use. In our implementation, we use the expected returns of the factors, multiplied by their coefficients to calculate the stock expected returns, which reduces number of μ estimates. We have a slightly higher number of variance estimates, as we use both the residual variance, and the variance of the factors in our calculations. However, as the only covariance estimates we need are the covariances of the factors $((6^2 - 6)/2 = 15)$, this is where the largest reduction in estimates is to be found. The standard Markowitz model directly estimates the covariance of each stock, resulting in a high number $((1000^2 - 1000)/2 = 499\,500)$. The largest number of estimations in our model comes from the α 's and factor-coefficients (7 000).*

2.2.2 Calculating stock weightings

The next step in creating our portfolios, is to calculate the weights of the individual stocks. This is done identically for all portfolios at the start of each year, and we will show how it is done in one of these years. To calculate the weights we use the R-package *quadprog* (Turlach & Weingessel, 2013), and solve it as a quadratic programming problem.

To find the weightings that give the highest *ex-ante* Sharpe ratio, we use a solution than minimises variance for a specific return. At first, the solution might describe a portfolio where the sum of the weights do not equal 100%, but this solution will still be on the CAL. Because of our specifications, it finds the point on the CAL with a return of 0.5%, and thus, by dividing the individual weights by the sum of all weights, we end up with the tangent portfolio.

We have decided to make our portfolio somewhat conservative, in that we will not allow shorting. The reasons for this limitation are the difficulties and constraints that exist for shorting (Goetzmann et.al, 2007). We are unable to take these constraints into account in our calculations. There is also no shorting in the market portfolio, which is the primary comparison portfolio in this study. We have also prohibited the weight of a single stock from being above 50%. While the sheer number of stocks make a weight above 50% unlikely, we decide to include this constraint because we want the overall portfolio to be somewhat diversified.

We use Quadratic Programming to implement our custom Markowitz framework, and this is also a way in which one could implement the standard framework. Therefore, what we do is the equivalent of what one would do to find the weightings in a standard Markowitz implementation, given the tools that we have decided to use. Using the *solve.QP* function in R's *quadprog* package (Turlach & Weingessel, 2013), we find the ideal proportional weights of the different stocks in our data set.

Specifically, the quadratic programming function aims to minimise equation (7) with respect to the x -vector, subject to the constraints in equation (8).

$$\text{Min}(-\mathbf{d}^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T D \mathbf{x}) \quad (7)$$

In equation (7) d is a vector, D is a matrix and T denotes the transpose of a vector. In equation (8) A and b are, respectively, a matrix and a vector describing the constraints.

$$A \mathbf{x} \geq \mathbf{b} \quad (8)$$

In our problem, as we are only minimising variance, the \mathbf{d} -vector is empty, reducing equation (7). By replacing x with ω for the weight of each stock we end up with the following minimising problem:

$$\text{Min}(\frac{1}{2} \omega^T \Sigma_s \omega) \quad (9)$$

Σ_s in equation (9) is the covariance matrix of the stocks. The minimising problem in equation (9) is subject to the constraints:

$$A \omega = \mathbf{b} \quad (10)$$

Specifically, what is being minimised is:

$$\begin{bmatrix} \omega_1 & \omega_2 & \cdots & \omega_n \end{bmatrix} \times \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} \times \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{bmatrix} \quad (11)$$

Subject to:

$$\begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_n \\ \hline 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \hline -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{bmatrix} \times \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{bmatrix} \geq \begin{bmatrix} 0.005 \\ 0 \\ 0 \\ \vdots \\ 0 \\ -0.5 \\ -0.5 \\ \vdots \\ -0.5 \end{bmatrix} \quad (12)$$

Only one of the constraints in equation (12) has an equality sign, and that is the constraint above the first line, which says that the portfolio return should equal 0.5%. The constraint between the two lines says that each individual weight has to be equal to or greater than 0, and the constraints below the second line says that no individual weight can be above 50%.

Once the stock weights for the theoretically optimal portfolio have been calculated, these are used to calculate the realised returns for this portfolio. To calculate these returns, the framework runs through each row of the data set, calculating the weighted return of each stock, and summing these to find the portfolio return. The returns of the different stocks will alter their weights in the portfolio. To recalibrate the weights, to account for these stocks returns, the framework multiplies the former weights with $(1 + r_i)$, and dividing by $(1 + r_{portfolio})$. This is displayed in equation (13):

$$\omega_{i,t+1} = \omega_{i,t} \frac{1 + r_{i,t}}{1 + r_{portfolio,t}} \quad (13)$$

2.3 Factor models

We create a set of 10 factor-based portfolios for daily, weekly and monthly realised returns, and the combinations of factors used are largely based upon established factor models. Because of this, we will present the models that we have used in creating our portfolios and explain their theoretical basis. It is worth noting that these models are designed to explain the cross-sectional returns of the stocks, rather than being used as a strict investment strategy. In this thesis however, we want to see how well they perform as a basis for stock-picking and investing in our framework. In addition to using the factor models described throughout this section, we have included some other portfolio designs. This is done to test how different sets of factors, in some cases taken from distinct models, perform in combination with each other.

2.3.1 The Capital Asset Pricing Model

One of the first factor models developed, is the Capital Asset Pricing Model (CAPM), which was published by Sharpe (1964), Treynor (1962), Lintner (1965) and Mossin (1966). Based on the findings of Markowitz (1959) the CAPM is a set of predictions concerning equilibrium expected return on risky assets and gives a theoretical explanation for risk premiums. The model shows a linear relationship between the return of the asset and the market, with β_i being the slope coefficient, formally expressed by equation (14).

$$R_{it} - R_{Ft} = \beta_i(R_{Mt} - R_{Ft}) \quad (14)$$

Sharpe (1964) and Lintner (1965) add two key assumptions to the Markowitz (1959) theory to identify a portfolio that is mean-variance-efficient. The first assumption is complete agreement, which states that investors agree on the joint distribution of asset returns from $t - 1$ to t . This assumption claims that the market is efficient which implies that the market portfolio is the optimal portfolio to invest in. This is why we find it interesting to use the market as the primary comparison portfolio in our thesis. The second assumption is borrowing and lending at a risk-free rate, which is the same for all investors and does not depend on the amount borrowed or lent. However, as Fama & French (2015) argue, the empirical record of the CAPM is poor. This is due to the strong simplifying assumptions which can give theoretical failings and it is difficult to implement valid tests for the model.

Based on the failure of CAPM the Arbitrage Pricing Theory (APT) was developed by Ross (1976). According to this model, asset returns are a linear combination of the returns of multiple systematic risk factors and an asset-specific return. Ross (1976) shows that the idiosyncratic risk can be diversified away by holding portfolios instead of single assets. He argues that the returns

should only incorporate the asset's exposure to factor risk. In contrast to CAPM, APT does not require equilibrium and opens up for more explanatory factors.

2.3.2 The Single Index Model

The Single Index Model (SIM), is a single factor model which uses the market index as a proxy for the common factor (Bodie, Kane & Marcus, 2014). Equation (15) shows the formal definition of the SIM.

$$R_{it} - R_{Ft} = \alpha_i + \beta_i(R_{Mt} - R_{Ft}) \quad (15)$$

In equation (15) α_i is the stock's excess return when the market excess return is zero. By including a portfolio based on the SIM we are able to test whether a simple model, based on only the most common factor and allowing for α 's and beneficial covariances, beats the market.

2.3.3 Fama-French three-factor model

Fama & French's (1993) approach is based in Ross' (1976) APT where they add more factors to the model. Fama & French (1993) argue that higher average returns on firms with small market capitalisation and firms with high book-to-market produce undiversifiable risks in returns that are not captured by the market. In support of this, Fama & French (1993) show that the returns on the stocks of firms with small market capitalisation covary more with one another than with returns on the stocks of firms with high market capitalisation. This also applies for the returns on firms with high book-to-market ratios, commonly known as value firms, when compared to firms with low book-to-market ratios, commonly known as growth firms. This is a continuation of Ball's (1978) findings where valuation ratios help identify variation in expected returns, with higher book-to-market ratios indicating higher required rates. Fama & French (1995) further argue that there are similar size and book-to-market patterns in the covariation of fundamentals like earnings and sales. Based on their findings (1993,1996) they propose a three-factor model for expected returns, shown in equation (16).

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i(SMB_t) + h_i(HML_t) \quad (16)$$

In this equation, small minus big (*SMB*) is the difference between returns on diversified portfolios of stocks with small market capitalisation and stocks with large market capitalisation. High minus low (*HML*) is the difference between the returns on diversified portfolios of stocks with high and low book-to-market ratios.

We have decided to use the Fama-French three-factor model as the basis for one of our portfolios because it is perhaps the best known multi-factor model. It is therefore interesting to see whether this can create a portfolio that is better than other established models.

2.3.4 The Carhart four-factor model

Jegadeesh & Titman (1993) show that an investment strategy defined as buying and selling stocks that have, respectively, high and low returns over the last 3-12 months give a risk-adjusted excess return. They use the US stock market to show this, but the strategy has also been shown to work in 12 European stock markets (Rouwenhorst, 1998). This strategy is called momentum and describes the tendency of a stock price to continue rising if it is going up and continue declining if it is going down.

Based on these findings the Carhart four-factor model was introduced in 1997 and is an extension of the Fama-French three-factor model, including a momentum factor (Carhart, 1997). This model is described by equation (17). The momentum-factor is calculated by subtracting the equal weighted average of the lowest performing firms from the equal weighted average of the highest performing firms, lagged one month (Carhart, 1997).

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i(SMB_t) + h_i(HML_t) + m_i(MOM_t) \quad (17)$$

By including a portfolio formed on the Carhart four-factor model, we will be able to see whether the inclusion of the *MOM*-factor can help improve portfolio performances.

2.3.5 Fama-French five-factor model

Novy-Marx (2013) and Titman, Wei & Xie (2004) claim that the Fama-French three-factor model is an incomplete model for expected returns. They argue that its three factors miss much of the variation in average returns related to profitability and investments. Fama & French (2006) argue that higher expected earnings imply a higher expected return, and shows the effect for the US stock market. Their findings are consistent with the findings of Haugen & Baker (1996) and Cohen, Gompers & Vuolteenago (2002). Novy-Marx (2013) shows that profitability, as measured as the company's gross profits to its assets, has roughly the same power as the book-to-market ratio in predicting the cross section of average returns. Based on this, profitable firms should outperform unprofitable firms.

Titman, Wei & Xie (2004) argue when looking at growth in equity as investments a higher expected growth in investments give a lower expected return and vice versa. An economic intuition behind this is that managers aim to build huge empire to seem more successful and gain

bonuses, instead of focusing on what is actually best for the shareholders, commonly known as over-investment hypothesis (Titman, Wei & Xie, 2004). An alternative explanation is that an investor is willing to increase investments when equity cost of capital is low (Liu, Whited & Zhang, 2009).

Based on the findings of Novy-Marx (2013) and Titman, Wei & Xie (2004), Fama & French (2015) add profitability and investment as factors to the three-factor model, shown in equation (18).

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i(SMB_t) + h_i(HML_t) + r_i(RMW_t) + c_i(CMA_t) \quad (18)$$

Robust minus weak (*RMW*) is the difference between the returns on diversified portfolios of stocks with robust and weak profitability. Conservative minus aggressive (*CMA*) is the difference between the returns on diversified portfolios of the stocks of low and high investment firms. Fama & French (2015) test whether a five-factor model improves the description of average returns compared to a four-factor model that drops *HML* by running regression on US stock data. They find that the five-factor model performs no better than the four-factor model because the *HML*-returns are captured by the factor's exposure to the two new factors, profitability, *RMW*, and investments, *CMA*. Based on their findings Fama & French (2015) conclude that *HML* is a noisy proxy for expected return because the market capitalisation also responds to forecasts of earnings and investment.

The Fama-French five-factor model is an expansion of their three-factor model, and we therefore use it as a basis for one of our portfolios. We also include the four-factor model where the supposedly redundant *HML*-factor is excluded. In addition, we will add the *MOM*-factor to these two models to form two more portfolios, allowing us to further assess the impact of momentum.

2.3.6 Combination of value and momentum

Asness, Moskowitz & Pedersen (2013) argue that market anomalies predominantly focuses on individual US equities and often examines value and momentum separately. In the cases which value and momentum are studied outside of US equities, they are also studied separately, and separate from other markets. Asness, Moskowitz & Pedersen (2013) offer new insights into these two market anomalies by examining their joint returns across several markets and asset classes. They find consistent evidence of value and momentum return premiums across all the markets they study.

Asness, Moskowitz & Pedersen (2013) find that a three-factor model with the market and these two factors explains the data better than including only one of them. Because both strategies

have positive average returns, and yet are negatively correlated with one another, the combination of them obtain higher risk premia.

As a continuation, they show that a combination of the two factors is immune to liquidity risk and generates substantial, abnormal returns. The negative correlation between value and momentum strategies and their high positive expected returns suggests that a combination of the two is more efficient than either strategy in isolation, which is why they construct the *COMBO*-factor shown in equation (19).

$$COMBO_t = 0.5(HML_t) + 0.5(MOM_t) \quad (19)$$

We are interested in the findings of Asness, Moskowitz & Pedersen (2013) in regards to the strength of value- and momentum investing, and add their *COMBO*-factor to our analysis. More specifically, we will test Asness, Moskowitz & Pedersen's (2013) three-factor model with the market, value and momentum, as well as their two-factor model of the market and the *COMBO*-factor. In addition we will also create a portfolio based upon the Fama-French five-factor model, but where we replace *HML* with the *COMBO*-factor. By adding these three portfolios to our analysis we are able to see if the value and momentum strategy add any significant value to the portfolios.

In contrast to the Fama & French (1993, 2015) factors, Asness, Moskowitz & Pedersen (2013), do not use value-weights to calculate the returns for their models, but rather signal-weights. However, we have, for the sake of consistency and availability, decided to use the factors supplied by Kenneth R. French (2017).

2.4 Creating comparison portfolios

To evaluate the performance of the investment strategies, beyond mutual comparisons, they will also be compared with the market portfolio, which will be our primary comparison portfolio. In addition, we wish to create a $\frac{1}{N}$ portfolio to further assess the quality of our framework. This portfolio is based on the same stocks that were used for the factor-based portfolios. That is to say, an observation rate of 75/50% over the preceding year (5 years for monthly data), as well as an observed return on the final measurement point in the preceding year will be required. This $\frac{1}{N}$ portfolio will have an equal weighting in each stock, and will also be rebalanced at the start of every year.

The reason for demanding the same observation rate for the $\frac{1}{N}$ portfolio is that it should make for a more accurate comparison. While the regressions used in the factor-based portfolios *require* a certain observation rate to be able to run, there is also a certain conservatism in the rate we have

set. If an investor demands a certain history in the stocks he invests in, this should also hold for a simple portfolio such as this, and not just the more complicated factor-based portfolios. In addition, this requirement will allow us to check whether the difference in observation rate in the portfolios formed on different measurement frequencies has any significant effects.

Our method of creating portfolios would make a standard Markowitz implementation a natural comparison portfolio. We have decided against this because of certain problems in calculating the covariance matrix. Given the large number of stocks we are investing in, the calculation of this matrix has considerable errors. The fact that there are some return observations that are unavailable is likely to increase these problems, and we get a non-positive definite matrix. This makes it impossible to calculate the weights of the stocks by the methods we use in this thesis.

2.5 Overview of our portfolios

From the factor models in section 2.3, and their underlying theory, we have decided upon 10 different sets of factors on which we will base our portfolios. The different factor bases, as well as our comparison portfolios, are summarised in table 2. Throughout the rest of this thesis, we will refer to the portfolios by the model names given in the table.

Model name	Description	Type
Six-factor	Fama-French five-factor model plus MOM	Factor-based portfolio
FF5F	Fama-French five-factor model	Factor-based portfolio
Mom5F	Fama-French five-factor minus HML, plus MOM	Factor-based portfolio
Combo5F	Fama-French five-factor minus HML, plus COMBO	Factor-based portfolio
FF4F	Fama-French five-factor minus HML	Factor-based portfolio
Carhart	Carhart four-factor model	Factor-based portfolio
FF3F	Fama-French three-factor model	Factor-based portfolio
Value-mom	Three-factor with market, HML and MOM	Factor-based portfolio
Combo	Two-factor with market and COMBO	Factor-based portfolio
SIM	Single Index Model	Factor-based portfolio
$\frac{1}{N}$	Investing equal share in each stock	Comparison portfolio
Market	Market portfolio	Comparison portfolio

Table 2: Name and description of the 10 factor-based portfolios and two comparison portfolios used in this thesis.

3 Evaluating portfolio returns

In this section, we will lay out the basis of how we evaluate the performance of our portfolios. It starts by presenting theory around the Sharpe- and appraisal ratios, and how we calculate them, before we explain the bootstrap method we use to expand our sample of these measures. We will then present skewness and kurtosis, and how these two measures are used to assess the crash-risk of our portfolios, before explaining how and why we measure the factor loadings of our portfolios.

3.1 Performance measures

3.1.1 Sharpe- and appraisal ratios

Sharpe (1966) argues that when optimising the capital allocation, the quality of a portfolio should be measured by the ratio of return to standard deviation. He claims that this best covers the importance of the trade-off between risk and reward. Based upon this, he introduced a measure called the Sharpe ratio. This measurement is defined in equation (20) where $E(r_p)$ is the expected return of a portfolio, r_f is the risk-free rate and σ_p is the standard deviation of the portfolio return. This ratio shows the excess return achieved per unit of risk the investor takes on.

$$SR_p = \left(\frac{E(r_p) - r_f}{\sigma_p} \right) \quad (20)$$

We use the Sharpe ratio to find our optimal portfolio in section 2.2.2. Specifically, we find the stock weightings that give the highest *ex-ante* Sharpe ratio, and we therefore use the *ex-post* Sharpe ratio as our primary performance measure.

Because our realised returns are based on several different frequencies of observation, and because it is customary to do so, we use the annualised Sharpe ratios. The principle of annualising the Sharpe ratio for daily data can be seen in equation (21). The same principle is applied to weekly and monthly returns.

$$SR_{daily} = \frac{\mu_{daily}}{\sigma_{daily}} \rightarrow SR_{annual} = \frac{\mu_{daily} \cdot n}{\sqrt{\sigma_{daily}^2 \cdot n}} = \frac{\mu_{daily} \cdot n}{\sigma_{daily} \cdot \sqrt{n}} = \frac{\mu_{daily}}{\sigma_{daily}} \sqrt{n} \quad (21)$$

We calculate the Sharpe ratio for the entire period, as well as using periodic results for this measure. For daily and weekly data, we calculate one Sharpe ratio for each year of observations, while for monthly data, we calculate the Sharpe ratio for four year periods. This gives us a

sample of 52 Sharpe ratios for the portfolios formed on daily and weekly data, while we get a sample of 12 for the portfolios formed on monthly data.

We will use the Sharpe ratio as our primary performance measure, but add the *ex-post* appraisal ratio to strengthen our analysis. This measure is the ratio of return unexplained by the benchmark portfolio divided by the standard deviation of the residuals, as seen in equation (22) (Kritzman, 1993). In this thesis, the market is used as the benchmark portfolio.

$$AR_p = \frac{\alpha_p}{\sigma(\epsilon_p)} \quad (22)$$

In equation (22) α_p is the return for each portfolio that is not explained by the market and $\sigma(\epsilon_p)$ is the risk unexplained by the market. In other words, the appraisal ratio tells how much value is created above the exposure to the market per unit of unsystematic risk (Warwick, 2003). As with the Sharpe ratios we annualise the appraisal ratios for each measurement frequency by applying the same principle as in equation (21). This measure is also calculated both for the entire return period and for periods of data, in the same way as with the Sharpe ratio. We therefore end up with 52 periodic appraisal ratios for the portfolios formed on daily and weekly data, while we get a sample of 12 for the portfolios formed on monthly data.

3.1.2 Bootstrap for the Sharpe and appraisal ratios

Samples of 52 Sharpe- and appraisal ratios are not an ideal number for making statistical inferences, and the 12 observations for monthly data are definitely insufficient. The bootstrap method can alleviate this problem. Efron & Tibshirani (1993) describe bootstrapping as using the original sample as the “population”, and drawing samples from this population, with replacement. As many samples as necessary can be drawn from the population, and then, by taking the mean of each sample, a distribution for the mean is found. With a sufficient number of means calculated, a confidence interval can be found by using the percentiles of the mean distribution found from the sampling.

The bootstrap method used in this study takes in each set of Sharpe- and appraisal ratios found for one frequency of return observations, and creates mean samples for each series. To minimise bias, the samples are taken for the same observations of Sharpe- or appraisal ratios for each return frequency. That is to say, if we create a sample of 3 Sharpe ratios from the daily data, we would use the same 3 years for each of our portfolios, rather than different years.

In total we produce 10 000 sample means for the one/four-year Sharpe- and appraisal ratios. These samples are then used to create 90% and 95% confidence intervals for the mean periodic value of these measures.

3.2 Skewness and kurtosis as measures of crash-risk

Skewness and kurtosis are measures used to describe the shape of a distribution (Joanes & Gill, 1998). Specifically, skewness is a measure of the asymmetry of the distribution, while kurtosis describes the thickness of the distribution's tails.

Asymmetry in returns could be problematic if it is negative. The reason for this, is that it could mean that a large share of the returns are negative, even if the mean is positive, and this could increase the risk of a portfolio crashing. If this asymmetry on the negative side is coupled with thick tails, it could be particularly problematic. According to Chen, Hong & Stein (2001) nine of the ten biggest one-day changes to the S&P 500 were negative, which suggests a negative skewness in the case of this index. This could also be the case for the market and/or our factor-based portfolios.

Because of the potential risks stemming from these measures, we have decided to use them to assess the crash-risk of our portfolios. To do this we use the R-package *fBasics*, and specifically the functions *skewness*, *kurtosis*, *colSkewness* and *colKurtosis* (Rmetrics Core Team et.al, 2014). These kurtosis functions identify *excess* kurtosis, where the normal distribution has a kurtosis of 0, rather than the absolute value of 3. We therefore add 3 to each kurtosis measurement to get the absolute value.

While different methods of calculating skewness and kurtosis might get different values, Joanes & Gill (1998) find that these deviations are negligible for a large sample size. As we wish to look at these measures both for the entire return series and for the worst times, it is worth noting these possible deviations when looking at the bad times. The bad times are identified by looking at the cumulative returns in calendar years. Because one year has so few return observations for monthly data, we exclude these portfolios in this analysis.

Because we want to look at skewness and kurtosis in combination, we have decided to define what we call the *crash-risk measure*. We define this composite measure as the product of the skewness and the kurtosis of the return distribution. Since this is not an established or tested measure, we need to be careful when using it in comparisons, particularly across different measurement frequencies.

3.3 Calculating factor loadings

We have decided to analyse factor loadings because we believe it may explain some of the performance differences between our factor-based portfolios and the market. In addition, it will help us explain similarities and differences in the returns and performances of the various factor-based portfolios.

To see how the different factors are loaded in the portfolios, we run a series of regressions on the portfolio returns as a function of the factors. The regressions will take the form of equation (23). We have decided to include every factor in these regressions, rather than just the ones used in forming the individual portfolios. The main reason for this is that it makes it easier to compare specific factor loadings between the various portfolios. Secondly, although a factor did not play a part in forming a specific portfolio, it might still be able to contribute to the explanation of its returns, because of interactions between the factors.

$$r_{pt} = \alpha_p + \beta_p(Mkt_t) + s_p(SMB_t) + h_p(HML_t) + r_p(RMW_t) + c_p(CMA_t) + m_p(MOM_t) + e_{pt} \quad (23)$$

4 Data

In this section the data used in this thesis is described. We start by presenting the stock data, and how we move from the raw data to our final data set. The creation of our weekly data set is then described, and the section is finished by presenting the factor data.

4.1 Stock returns

We have chosen to test our model on the US stock market. The stock data is downloaded from The Center for Research in Security Prices (CRSP). Specifically, we download data for all available stocks in the time-period from 30. December 1963 to 31. December 2016 for daily data. For monthly data, we use the period between January 1964 and December 2016. The variables we have chosen to include are described in table 3. Weekly data is not available from CRSP (CRSP Stocks, 2017). We therefore create this from the daily data set, once it has been processed in the manner described in section 4.1.1.

Abbreviation	Name	Description
PERMNO	PERMNO	Permanent company number
Date	Date	Date of the observation
SHRCD	Share Code	Two digit code describing the share traded The first digit describes the type of share traded Combination of first and second digit identify the security type
EXCHCD	Exchange code	Identify the stock exchange, 1 = NYSE, 2 = NYSE MKT 3 = NASDAQ, 31 = NYSE When-issued 31 = NYSE When-issued, 32 = NYSE MKT When-Issued 33 = NASDAQ When-Issued 0 = Unknown, -1 = Suspended, -2 = Halted
PRC	Price	Closing price or Negative Bid/Ask Average on Calendar Date
RET	Return	Return for the stock
DLRET	Delisting Return	Return on delisting of a stock

Table 3: *Abbreviations, names and descriptions of the different variables included in data set from CRSP*

4.1.1 Data management

Most of the data management is conducted in R (R Core Team, 2016), but some minor data clean-up was done in a text editor. Specifically, there were some returns from the CRSP data that were listed with letters, which led R to read the return variable as a factor, and these were removed and set to “not available”.

We start with data from 30. December 1963 up to 31. December 2016 for US stocks (CRPS stocks, 2017). The two observations from 1963 will be used when we calculate weekly returns.

In order to get a data set with only full weeks we need the two observations from 1963 since these observations are Monday and Tuesday in the first week of our data set. However, while using daily returns we remove the two observations from 1963, as 1964 through 2016 are the only full years for which complete data is available both from CRSP stocks (2017) and all of Kenneth French's (2017) factors.

In the data set from CRSP, there are different types of securities included, and from several exchanges (CRSP Stocks, 2017). We decide to limit our data to three stock exchanges, NYSE, NASDAQ and AMEX. This limitation is included because these exchanges are the ones used by Kenneth French when calculating the factors we use (French, 2017). The number of observations that are removed by this filter can be seen in tables 4 and 5.

Furthermore, we will only use the securities that are marked as "Common stocks which need/have not be/been further defined". Analyses of the US stock market are normally restricted to common stocks, as US preferred stocks are closer to bonds in their nature than to stocks (Bruckner et al., 2015) and therefore have different dynamics. The factors provided by Kenneth R. French are based on only common stocks, and therefore we decide to remove the securities not used by French (French, 2017). By only including common stocks, we are able to make the data sets more manageable without affecting the significance of our analysis. The number of observations that are removed by this filter can be seen in tables 4 and 5.

Further, we will also exclude penny stocks from our data sets. Shares of very low value, commonly known as penny stocks, can distort our analysis since slight changes in stock price can give very high/low returns. This is misleading because it does not reflect the rise of a growth company but simply shows minimal price fluctuations of inconsiderable stocks. The US Securities and Exchange Commission (SEC) defines a penny stock as a security that is issued by a very small company and that trades for less than \$5 per share (SEC, 2013).

According to NASDAQ stock market rules (NASDAQ Inc., 2017) penny stocks do not qualify for being listed. This implies that the studies of NASDAQ stock exchange do not contain these stocks. However, we are unable to find equivalent data for the other exchanges, meaning our data sets may contain penny stocks. This is a problem for our analysis due to potential very high/low returns. While NASDAQ Inc (2017) define the limit for penny stocks to be \$4, we have decided to use the definition from the SEC (2013) and remove all observations with a price below \$5. We choose this limit because we use several exchanges, and therefore prefer to use the governmental definition. The number of observations that are removed due to the penny stock filter can be seen in tables 4 and 5.

The delisting return in our data is a stock's return on delisting from the exchange. To include this in the final data sets, we replace normal returns listed as "not available" with the corresponding delisting return. This is done to include all the returns that are realised for our portfolios.

	Observations		Stocks
	<i>Number</i>	<i>Difference</i>	<i>Number</i>
Raw CRSP data	80958729		30998
Stock Exchange	77396016	-3562713	29717
Security Type	63588251	-13807765	23867
Penny Stocks	39493474	-24094777	18987

Table 4: Amount of observations excluded by the various data filters we apply to get from the original to the final daily data set. Also included are the number of unique stocks in the data set, as measured by number of unique PERMNOs. The single steps are described in detail throughout this section.

	Observations		Stocks
	<i>Number</i>	<i>Difference</i>	<i>Number</i>
Raw CRSP data	3909068		30972
Stock Exchange	3710859	-198209	29688
Security Type	3048830	-662029	23847
Penny Stocks	1889300	-1159530	18715

Table 5: Amount of observations excluded by the various data filters we apply to get from the original to the final monthly data set. Also included are the number of unique stocks in the data set, as measured by number of unique PERMNOs. The single steps are described in detail throughout this section.

Once the data has been filtered according to the previous paragraphs, we reshape the data set using the *dcast*-function from R's *data.table*-package (Dowle & Srinivasan, 2017). This changes the format of the data sets, so that the first column is the date, followed by one column for each of the stocks in the data set. This gives us a set of 13 345 observations for 18 987 stocks in the daily data set, which will be reduced by two observations once the 1963 data is removed. For the monthly data set we have 636 observation of 18 715 stocks. Because we have no need for stocks where there is not a single return observations, we calculate the observation rate for each stock, and remove the stocks where it is zero. This reduces the number of stocks by 13 to 18 974 in the daily set. In the monthly data set, this filter reduces the number of stocks by 76, to 18 639.

Because the models we use are based upon excess returns, this is the next step in our data management. For each series of stock returns, we subtract the risk-free rate, based upon the one month US Treasury bill, provided by French (2017).

4.1.2 Creating weekly data

There is no weekly data provided by CRSP, and therefore we have to create this data set ourselves. We start by defining an “auxiliary” date for each observation, which is the date of the Friday of the week in question. To calculate the weekly returns we multiply the daily returns and then subtract 1, as can be seen in equation (24). In R, this is done by using the *aggregate*-function for the observations with the same “auxiliary” dates.

$$R_{week} = (1 + r_{Mon})(1 + r_{Tue})(1 + r_{Wed})(1 + r_{Thu})(1 + r_{Fri}) - 1 \quad (24)$$

The calculation of weekly returns is done in an identical fashion for both the stocks and the factors, and will result in some returns being reported as part of the next year. Specifically, returns seen in late December, if the corresponding Friday is in January, will be reported the following year. One example of this is the two observations from 1963 included in the first week of 1964.

4.2 Factor returns

Our source for the factor returns is the Kenneth R. French Data Library (French, 2017). From French’s library we get the daily and monthly factors from the five-factor model and momentum based on the US stock market. We download the factors for the same period as the stock data. This data set also contains the risk-free rate, which is used to calculate the excess returns. Although we considered constructing the factors ourselves, we decided not to, given the large time period for which they are readily available for the US stock market. Furthermore the purpose of this thesis is not to construct the factors, but to test whether or not our custom Markowitz implementation, using factor models, beat the market.

We download the five-factor data set, as well as the momentum returns. Based upon these, we create Asness, Moskowitz & Pedersen’s (2013) *COMBO*-factor in accordance with equation (19). We download both the daily and monthly factor data, and create the weekly factors in the same way as was done for the stock data. While the weekly factors included in the Fama-French three-factor model is available to download, we decide to create all of them from the daily factors for consistency. The factors downloaded from French (2017) are reported in percentage rather than absolute value, and we therefore divide these by 100.

5 Results and analysis

In our thesis we focus on 10 factor-based portfolios and two comparison portfolios, listed in table 2. It is worth noting that the return results reported here are *realised excess returns*, but we will just use the term *returns* for readability purposes. For the same reason, when discussing *ex-post* Sharpe- and appraisal ratios, we will only use the terms Sharpe- and appraisal ratios.

This section presents the results of our analysis. We start by analysing the portfolio returns, presenting summary statistics for the entire observation period. We then move on to analyses of the periodical Sharpe- and appraisal ratios, assessing how our portfolios perform on the basis of their mean one/four-year statistics by these measures. The next subsection will then try to analyse the crash-risk of our portfolios. This is done by looking both at their worst single observations and worst years of cumulative returns, and by estimating the skewness and kurtosis of the return distributions. Based on the findings from Sharpe- and appraisal ratios, and crash-risk we will look at the factor loadings, to see whether they can help us assess what we see in our performance measures. Finally, we discuss, and try to explain, what we have observed in the portfolio returns.

5.1 Summary of portfolio returns

Table 6 shows returns, standard deviations and annualised Sharpe- and appraisal ratios for the different portfolios after running our model on the entire return series. For daily data all the factor-based portfolios produce similar returns but the six-factor portfolio has the highest Sharpe- and appraisal ratios. The market has both the lowest return and Sharpe ratio.

For weekly data the $\frac{1}{N}$ portfolio gives the highest return but also the highest standard deviation. The six-factor portfolio again produces the highest Sharpe ratio, while the $\frac{1}{N}$ portfolio has the highest appraisal ratio. Meanwhile, the market portfolio produces the lowest Sharpe ratio, with the lowest mean return.

For the monthly data the $\frac{1}{N}$ portfolio produces the highest return and also the highest Sharpe- and appraisal ratios, while the Combo5F portfolio is the best of the factor-based portfolios. Just as in the weekly data the market produces the lowest return and Sharpe ratio, and every other portfolio has a positive appraisal ratio, with the market as a benchmark.

It is hard to conclude anything from table 6 since the returns are not normally distributed. However the table gives an overview of how the different portfolios perform. It is worth noting that the market seems to be the worst portfolio to invest in, especially for the daily returns.

The Sharpe- and appraisal ratios for the factor-based portfolios created on daily returns are unexpectedly large, with no values below 1.4. This is particularly clear when comparing their

Sharpe ratios with that of the market. One potential explanation is that the factor-based portfolios are better at choosing stocks with favourable correlation between the stocks which will reduce the volatility of the returns.

Model	Daily				Weekly				Monthly			
	μ	σ	SR	AR	μ	σ	SR	AR	μ	σ	SR	AR
Six-factor	0.062%	0.672%	1.464	1.517	0.183%	1.81%	0.729	0.682	0.668%	4.58%	0.512	0.355
FF5F	0.062%	0.673%	1.461	1.512	0.179%	1.83%	0.707	0.646	0.671%	4.50%	0.516	0.360
Mom5F	0.062%	0.674%	1.454	1.502	0.182%	1.81%	0.725	0.675	0.663%	4.49%	0.512	0.354
Combo5F	0.062%	0.675%	1.455	1.505	0.183%	1.82%	0.725	0.674	0.689%	4.50%	0.531	0.384
FF4F	0.062%	0.675%	1.452	1.499	0.180%	1.84%	0.707	0.646	0.683%	4.49%	0.527	0.377
Carhart	0.062%	0.669%	1.459	1.510	0.181%	1.82%	0.717	0.659	0.677%	4.48%	0.524	0.372
FF3F	0.062%	0.669%	1.459	1.509	0.180%	1.85%	0.703	0.637	0.671%	4.60%	0.521	0.367
Value-mom	0.060%	0.658%	1.445	1.504	0.188%	1.88%	0.722	0.667	0.669%	4.58%	0.506	0.345
Combo	0.059%	0.656%	1.424	1.477	0.189%	1.88%	0.725	0.671	0.681%	4.54%	0.519	0.365
SIM	0.060%	0.656%	1.440	1.496	0.189%	1.88%	0.723	0.670	0.656%	4.54%	0.500	0.335
$\frac{1}{N}$	0.044%	0.872%	0.806	1.162	0.193%	2.20%	0.633	0.785	0.864%	4.76%	0.629	0.759
Market	0.024%	1.001%	0.375	-	0.115%	2.20%	0.376	-	0.489%	4.56%	0.372	-

Table 6: Average returns, standard deviations and annualised Sharpe- and appraisal ratios calculated for the entire return series of our portfolios for daily, weekly and monthly returns. Appraisal ratio is not calculated for the market portfolio, as it is the benchmark used to calculate this value for the other portfolios. We use return observations from 1965 through 2016 for daily and weekly data, and return observations from 1969 through 2016 for monthly data.

Looking at the correlation between the different factors in our daily data set we find that the market has a negative correlation with all the other factors. This suggests that a positive exposure to the other factors would lead to beneficial correlations when it comes to reducing volatility. This exposure could also contribute to raising the returns of the portfolio, which means that it could improve the Sharpe ratio further. In addition, our framework allows for α 's in the individual stocks, and this could help improve the Sharpe- and appraisal ratios. We will explore this further when analysing the factor loadings of the various portfolios in section 5.5.

The market having the lowest Sharpe-ratio for all measurement frequencies goes against one core assumption of the CAPM theory, namely that CAPM is the ideal portfolio (Sharpe, 1964). Instead, it supports Fama & French's (2015) findings that the CAPM model struggles empirically.

ically. By looking at daily, weekly and monthly data combined for annualised Sharpe ratios the differences are largest for the daily data. One reason for this could be the greater number of return observations, which could lead to better coefficient estimates in the regressions used to calculate the covariances. This could in turn lead to more accurate stock picking for these portfolios, which would give better results.

5.2 Periodical Sharpe ratio results

5.2.1 Daily data

To investigate the Sharpe ratio for each portfolio further, we calculate the one-year Sharpe ratio by looking at one year of daily returns and not the entire sample size as in table 6. Table 7 shows summary statistics of this measure. In this data, the highest maximum ratio is obtained by the six-factor portfolio. The SIM portfolio has the highest mean one-year Sharpe ratio, which is a bit different than what we saw in table 6 where the six-factor portfolio had the highest annualised Sharpe ratio. However, the amount of data does not allow for any solid conclusions from the statistics in table 7.

Model	SR_{min}	SR_{mean}	SR_{max}
Six-factor	-3.105	2.017	6.889
FF5F	-3.078	2.027	6.894
Mom5F	-3.120	2.015	7.424
Combo5F	-3.111	2.029	7.109
FF4F	-3.099	2.037	7.522
Carhart	-3.108	2.014	6.921
FF3F	-3.107	2.027	6.925
Value-mom	-3.154	2.036	6.661
Combo	-3.229	2.037	6.844
SIM	-3.204	2.060	7.101
$\frac{1}{N}$	-2.543	1.049	4.756
Market	-1.824	0.604	3.418

Table 7: Minimum, maximum and mean Sharpe ratios calculated for series of one-year Sharpe ratios, using daily return observations from 1965 through 2016 for our portfolios.

To increase the sample size, we run the bootstrap method on the Sharpe ratio data, and calculate 10 000 one-year Sharpe ratio means, producing 90% and 95% confidence intervals shown in figure 2. What can be concluded from the confidence intervals in this figure is that all the factor-based portfolios are significantly better than the $\frac{1}{N}$ - and market portfolios on a 10% level.

On a 5% significance level all the factor-based portfolios are better than the market and the SIM portfolio is also better than the $\frac{1}{N}$ portfolio. It is also apparent that the portfolios have higher one-year Sharpe ratios than the value they obtained for the entire period, as seen in table 6. This is however also observed for the comparison portfolios. The reason for this might be the fact that returns are autocorrelated, meaning that returns closer in time to each other will have more similar values. This would result in a lower standard deviations, which in turn would increase the Sharpe ratio. The reasons for our factor-based portfolios beating the market, are likely to be the same as those discussed in section 5.1.

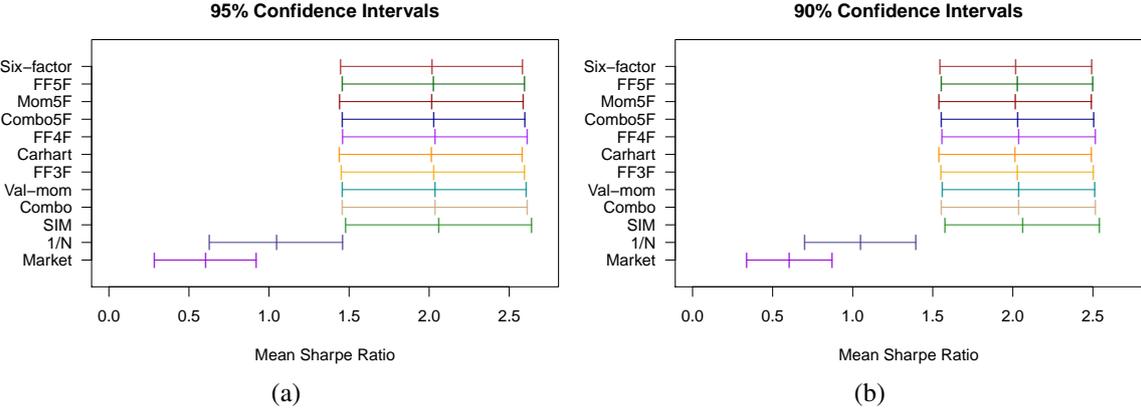


Figure 2: 95%(a) and 90%(b) confidence intervals for the mean one-year Sharpe ratio calculated from daily returns from 1965 through 2016. The confidence intervals are based on the percentiles of the 10 000 Sharpe ratio means from the bootstrap method. The marking in the middle of the confidence intervals show the mean value.

5.2.2 Weekly data

Table 8 shows summary statistics for the time series of annual Sharpe ratios calculated from weekly returns. From this table it is clear that the SIM- and market portfolios have the highest and lowest mean one-year Sharpe ratios, respectively. This matches what we saw in table 6. The SIM portfolio also has the single highest Sharpe ratio observed among these portfolios, while the Combo portfolio has the lowest. The portfolios formed on weekly data also see periodical Sharpe ratios that are higher than that for the entire return period. This is again likely due to the autocorrelation of returns.

We run the bootstrap method to calculate 10000 annual Sharpe ratio means from weekly returns, producing the confidence intervals in figure 3.

As can be seen from figure 3, the SIM portfolio has the highest mean one-year Sharpe ratio, while the market retains the lowest ratio. However, when looking at the confidence intervals, we can conclude that there are no statistically significant differences between the various portfolios. This applies both on the 5% and 10% significance levels.

Model	SR_{min}	SR_{mean}	SR_{max}
Six-factor	-1.950	1.065	3.620
FF5F	-2.034	1.054	3.656
Mom5F	-1.938	1.058	3.946
Combo5F	-2.008	1.060	3.745
FF4F	-2.034	1.050	4.068
Carhart	-1.908	1.053	3.680
FF3F	-2.086	1.052	3.808
Value-mom	-2.537	1.063	3.978
Combo	-2.570	1.068	4.351
SIM	-2.563	1.070	4.428
$\frac{1}{N}$	-1.883	0.839	4.145
Market	-1.558	0.626	3.762

Table 8: Minimum, maximum and mean Sharpe ratios calculated for series of one-year Sharpe ratios, using weekly return observations from 1965 through 2016 for our portfolios.

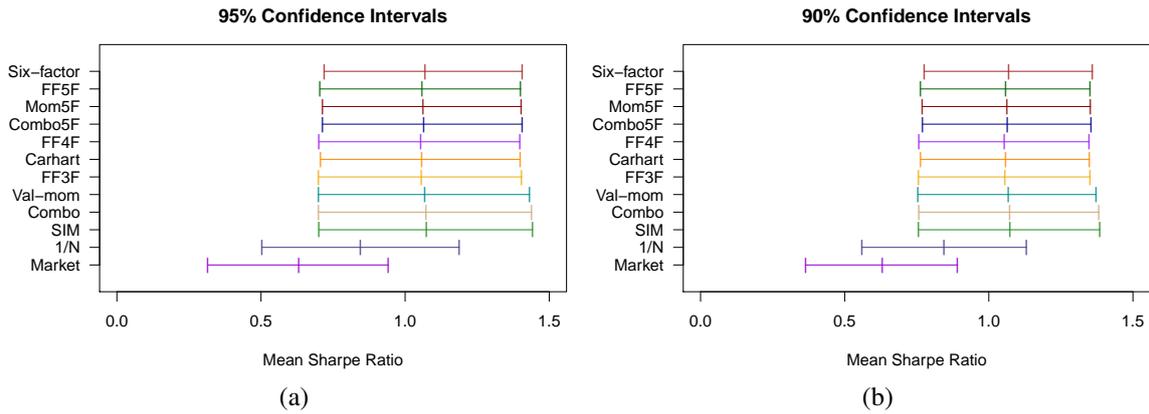


Figure 3: 95%(a) and 90%(b) confidence intervals for the mean one-year Sharpe ratio calculated from weekly returns from 1965 through 2016. The confidence intervals are based on the percentiles of the 10 000 Sharpe ratio means from the bootstrap method. The marking in the middle of the confidence intervals show the mean value.

5.2.3 Monthly data

Table 9 shows a summary of the annualised four-year Sharpe ratios calculated for the monthly returns. For these monthly returns, the $\frac{1}{N}$ portfolio has the highest mean four-year Sharpe ratio, while the market portfolio has the lowest mean ratio. This matches what we see in table 6. The single highest Sharpe ratio is obtained by the FF3F portfolio, while the lowest Sharpe ratio is observed in the SIM portfolio. Unlike in the one-year Sharpe ratios for portfolios formed on daily and weekly data, there seems to be no increase in the periodical values. The reason for this difference is likely to be that the four year periods used to calculate these values are more

diverse in the return observations than what we saw for the other measurement frequencies.

Model	SR_{min}	SR_{mean}	SR_{max}
Six-factor	-0.451	0.666	1.86
FF5F	-0.517	0.677	1.88
Mom5F	-0.468	0.666	1.90
Combo5F	-0.490	0.677	1.89
FF4F	-0.531	0.684	1.88
Carhart	-0.540	0.676	1.90
FF3F	-0.587	0.681	1.94
Value-mom	-0.599	0.658	1.82
Combo	-0.605	0.677	1.83
SIM	-0.667	0.674	1.87
$\frac{1}{N}$	-0.245	0.695	1.48
Market	-0.524	0.429	1.29

Table 9: Minimum, maximum and mean Sharpe ratios calculated for series of four-year Sharpe ratios, using monthly return observations from 1969 through 2016 for our portfolios.

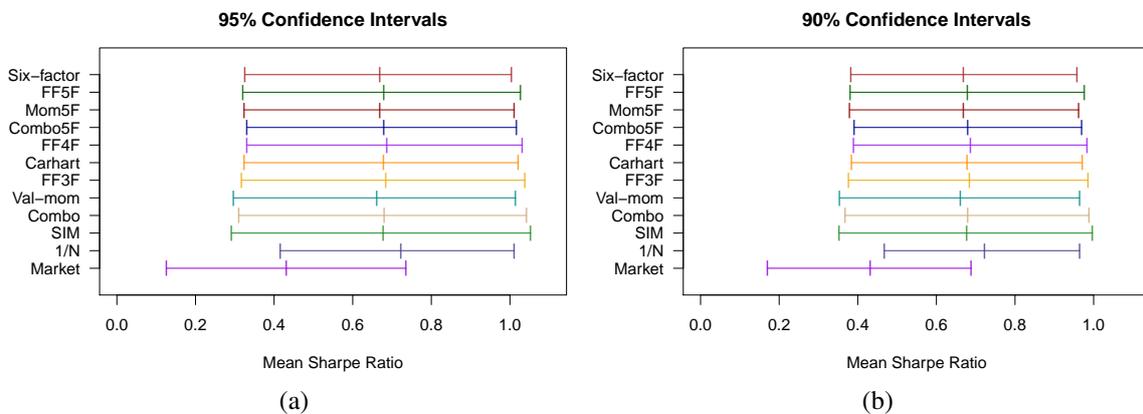


Figure 4: 95%(a) and 90%(b) confidence intervals for the mean four-year Sharpe ratio calculated from monthly returns from 1969 through 2016. The confidence intervals are based on the percentiles of the 10 000 Sharpe ratio means from the bootstrap method. The marking in the middle of the confidence intervals show the mean value.

Using a bootstrap method on the monthly Sharpe ratio data and calculating 10 000 four-year Sharpe ratio means, we obtain the statistics that are summarised in figure 4. From this data, we see that increasing the sample size, does not change which portfolio has the highest or lowest mean Sharpe ratio.

Unlike in the daily data, there are no significant difference in the means of the one-year Sharpe ratios for any portfolios in the monthly data on either the 10% or 5% significance levels. Of the factor-based portfolios, the FF4F portfolio has the highest mean one-year Sharpe ratio, but as

mentioned, the difference is not statistically significant.

5.2.4 Analysis of the Sharpe-ratio results

By looking at the confidence intervals for daily, weekly and monthly returns combined there are some differences in the results. Only the daily data has significant differences between the portfolios. Specifically, the SIM portfolio is significantly better than the market and the $\frac{1}{N}$ portfolio on a 5% level. The rest of the factor-based portfolios are all significantly better than the market on a 5% level, and better than the $\frac{1}{N}$ portfolio on a 10% level. From the monthly and weekly data there are no portfolios that are significantly better than any other. It is also clear from the tables and confidence intervals, that portfolios formed on the same measurement frequency have remarkably similar performances.

One potential reason for the differing results for the different measurement frequencies is the observation rate requirement. We demand a lower observation rate for the daily data, which might contribute to higher returns and/or lower volatility as there are more diversification possibilities. However, there is one argument against the observation rate causing these significant differences. The $\frac{1}{N}$ portfolio is based upon the same observation rate as the factor-based portfolios, which means that the difference between this and the factor-based portfolios can not be explained by the observation rate.

Another potential reason is that even though the observation rate demanded for daily data is lower, the raw number of observations demanded will still be higher. This might produce more accurate regressions, which could lead to better stock-picking, resulting in better returns.

A third factor to consider is that the one-year Sharpe ratios for daily returns is calculated on a higher number of observations than the one-year Sharpe ratios for weekly returns, or the four-year Sharpe ratios for the monthly data. This can drive down the standard deviation in the daily return series, thereby increasing the one-year Sharpe ratios. However, this should also apply to both the market and the $\frac{1}{N}$ portfolio.

The last point of note is that there is likely to be more noise in the daily data, that do not actually represent genuine changes in stock prices. This noise should however result in a higher volatility, but not higher returns, and should also apply to our comparison portfolios.

As we see it, it is therefore likely that the factor-based portfolios formed on daily data is significantly better than the market because they are more efficient in their stock-picking. The models that form the basis for our portfolios look at both the relevant factor returns, and the stock returns, and find the theoretically ideal portfolio. For the daily data this gives significantly better returns than the market.

The CAPM has struggled empirically, and that it should be beaten by factor models that take more information into account is consistent with Fama & French (2015). Although the factor models are only significantly better than the market for daily data, there does seem to be some indication that they are better for the other measurement frequencies as well. However, this cannot be concluded from our data set.

For our further analysis, we will choose a subset of our factor-based portfolios to analyse more deeply. We have done this to reduce the amount of tables and data that would need to be included, without contributing any further value. Given the similarity in performance between the various models when looking at returns and the Sharpe ratios, this should not be problematic. We wish to choose a subset that is as diverse as possible given our original set of portfolios. For this reason, we decide to include both the portfolios formed on the largest and the smallest number of factors, the six-factor and SIM portfolios. Given Asness, Moskowitz & Pedersen's (2013) claim of robustness for the value-momentum strategy, we also include the value-mom portfolios. The Fama-French models are perhaps the best known multi-factor models, and we therefore want to include portfolios formed on one of these, namely the FF5F portfolios. Finally, we have decided to include the Carhart portfolios.

5.3 Periodical appraisal ratio results

To analyse the appraisal ratio further, in a similar way to what we did for the Sharpe ratio, we look at periodical data for this measurement as well. For daily and weekly portfolios, we calculate one-year appraisal ratios, while for the monthly portfolios, we calculate four-year appraisal ratios. For the analysis of the appraisal ratio, we use the market as our benchmark. Therefore, a statistically significant mean one/four-year appraisal ratio is an indication of the portfolio beating the market by this measure.

5.3.1 Daily data

Table 10 shows the summary statistics of the one-year appraisal ratios for the portfolios formed on daily data. There are some slight differences in the highest and lowest appraisal ratio observations, but it is clear that the factor-based portfolios are still similar by this measure.

In the same way as we calculated 10 000 Sharpe ratios, we use the bootstrap method to calculate 10 000 one-year mean appraisal ratios to find if there are any significant differences. The confidence intervals of this data can be seen in figure 5.

The fact that each of the daily factor-based portfolios, as well as the $\frac{1}{N}$ portfolio, has a mean one-year appraisal ratio that is significantly higher than zero means that they all beat the market

Model	AR_{min}	AR_{mean}	AR_{max}
Six-factor	-3.11	2.11	7.01
FF5F	-3.07	2.12	7.02
Carhart	-3.12	2.11	7.06
Value-mom	-3.22	2.16	6.83
SIM	-3.30	2.19	7.30
$\frac{1}{N}$	-3.61	1.29	6.18

Table 10: Minimum, maximum and mean appraisal ratios calculated for series of one-year appraisal ratios, using daily return observations from 1965 through 2016 for our portfolios.

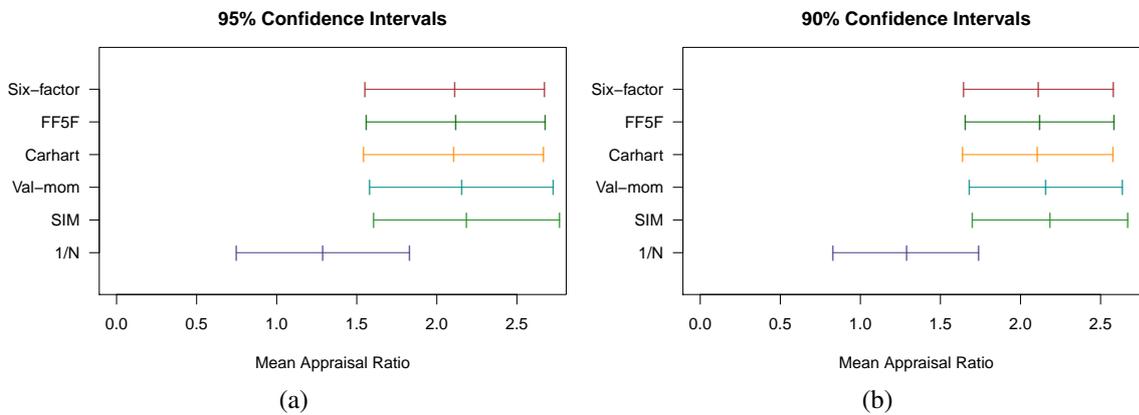


Figure 5: 95%(a) and 90%(b) confidence intervals for the mean one-year appraisal ratio calculated from daily returns from 1965 through 2016. The confidence intervals are based on the percentiles of the 10 000 appraisal ratio means from the bootstrap method. The marking in the middle of the confidence intervals show the mean value.

by this measure on a 5% level. This strengthens the findings from the Sharpe ratio analysis that our factor-based portfolios formed on daily data beat the market. Similarly to what we saw with the Sharpe ratio, we have higher periodical values than the corresponding values for the entire return series seen in table 6. Again, we believe this is likely to be due to more consistency in returns over a shorter period, leading to less volatile residuals.

5.3.2 Weekly data

Table 11 shows summary statistics of the annual appraisal ratios calculated from the weekly returns. The highest mean one-year appraisal ratio is obtained by the six-factor portfolio, while the lowest is obtained by the $\frac{1}{N}$ portfolio. However, the $\frac{1}{N}$ portfolio also has the highest maximum one-year appraisal ratio. Again we see a higher mean periodical value than what we saw for the entire return series in table 6. The autocorrelation of the returns is again likely to be the reason for this.

Model	AR_{min}	AR_{mean}	AR_{max}
Six-factor	-2.67	0.95	3.21
FF5F	-2.52	0.93	3.26
Carhart	-3.00	0.93	3.29
Value-mom	-2.94	0.94	3.72
SIM	-3.04	0.94	4.31
$\frac{1}{N}$	-2.62	0.73	4.45

Table 11: Minimum, maximum and mean appraisal ratios calculated for series of one-year appraisal ratios, using weekly return observations from 1965 through 2016 for our portfolios.

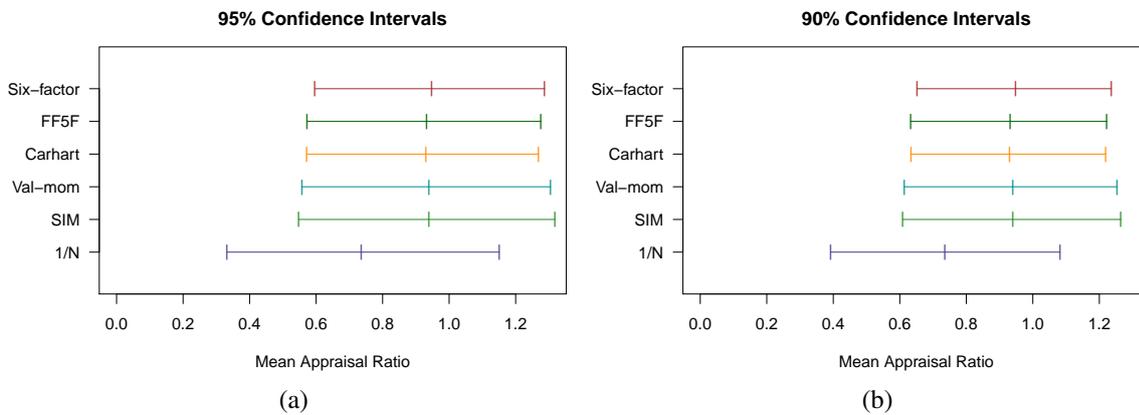


Figure 6: 95%(a) and 90%(b) confidence intervals for the mean one-year appraisal ratio calculated from weekly returns from 1965 through 2016. The confidence intervals are based on the percentiles of the 10 000 appraisal ratio means from the bootstrap method. The marking in the middle of the confidence intervals show the mean value.

We run the bootstrap method to calculate 10 000 annual appraisal ratios from the weekly returns producing the confidence intervals in figure 6. In contrast to the analysis of the Sharpe ratio on weekly data we find that all the selected portfolios are significantly outperforming the market both on a 5% and 10% significance level.

5.3.3 Monthly data

Table 12 shows a summary of the annual appraisal ratios calculated over a period of four years for the monthly returns. The $\frac{1}{N}$ portfolio has the highest mean four-year appraisal ratio. This is consistent with what we found in table 6 where the $\frac{1}{N}$ portfolio also had the highest annualised mean Sharpe ratio. The highest appraisal ratio for our factor-based portfolios is obtained by the FF5F portfolio, while the lowest appraisal ratio is the value-mom- and SIM portfolios.

Unlike what we saw in the periodical Sharpe ratio data, we do get an increase in the four-year appraisal ratio when compared to the equivalent value for the entire return series in table 6.

Model	AR_{min}	AR_{mean}	AR_{max}
Six-factor	-0.67	1.07	2.69
FF5F	-0.72	1.08	2.74
Carhart	-0.65	1.06	2.77
Value-mom	-0.61	0.99	3.10
SIM	-1.02	0.99	2.83
$\frac{1}{N}$	-1.35	1.65	4.45

Table 12: Minimum, maximum and mean appraisal ratios calculated for series of four-year appraisal ratios, using monthly return observations from 1969 through 2016 for our portfolios.

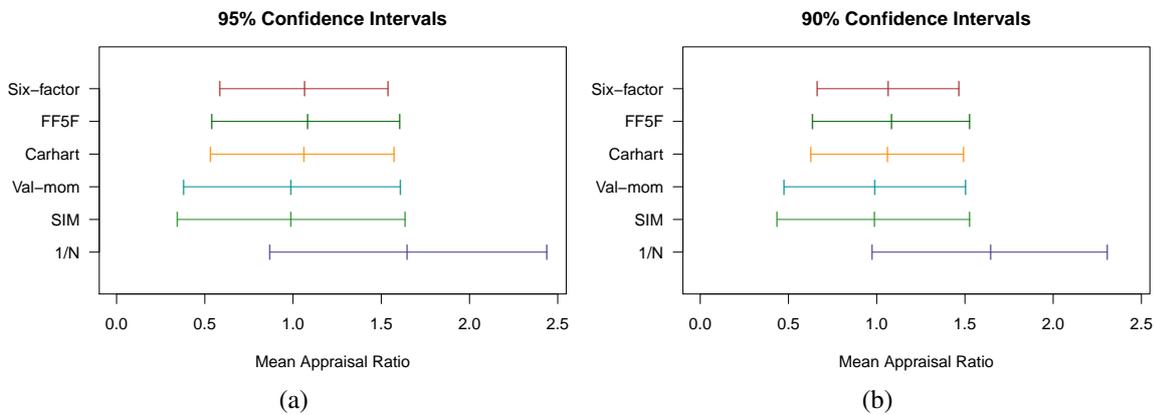


Figure 7: 95%(a) and 90%(b) confidence intervals for the mean four-year appraisal ratio calculated from monthly returns from 1969 through 2016. The confidence intervals are based on the percentiles of the 10 000 appraisal ratio means from the bootstrap method. The marking in the middle of the confidence intervals show the mean value.

This could be due to higher consistency in the residuals for the periods than there are in the full return series, and/or it could be that there are higher α 's in these shorter periods.

Using the bootstrap method on the four-year appraisal ratios and calculating 10 000 annual appraisal ratio means, we obtain the statistics that are summarised in figure 7. Unlike the monthly data for Sharpe ratio we find that these portfolios outperform the market on the 5% and 10% significance levels.

5.3.4 Analysis of appraisal ratio results

All the factor-based portfolios beat the market on a 5% significance level. However, while the $\frac{1}{N}$ portfolio has the worst one-year mean appraisal ratio when formed on daily and weekly data, it has the highest four-year mean when formed on monthly data. As was the case with the Sharpe ratio data, it is clear that the factor-based portfolios formed on the same frequency are remarkably similar.

While all portfolios have a statistically significant mean one/four-year appraisal ratios, there are some differences in the actual value between the measurement frequencies. The reasons for why the portfolios formed on daily data seem to perform better than those formed on other frequencies are likely to be similar to those discussed while analysing the Sharpe ratios. Particularly the greater number of observations, both when picking stocks, and when calculating the appraisal ratio could be contributing to this.

Because of the low number of observations used to form the bootstrapped sample for the monthly portfolios, we wish to apply some caution when interpreting these results in particular. While the bootstrapped data has 10 000 data points, we base these upon only 12 actual observations. This could potentially lead to some bias in the result. While this was true also for the Sharpe ratios, the lack of any significance in those results meant that this was not as big a concern in that analysis. The fact that we see significant appraisal ratios for the portfolios formed on daily and weekly data however, does indicate that our framework beats the market.

We believe that the significant outperformance of the market by the factor-based portfolios is caused by their active stock-picking. By using a model that takes advantage of both α 's and beneficial covariances we get a significantly higher appraisal ratio than the market. This also supports the empirical struggles of the CAPM, consistent with Fama & French (2015).

5.4 Measuring crash-risk in the portfolios

To further assess the quality of our portfolios, we analyse their crash-risk. We first look at the worst days/weeks/months that are observed for each portfolio, as well as the cumulative return in the worst year for each portfolio. By comparing this data, we get a chance to see how bad the most extreme returns are. This can help us assess how likely the portfolios are to collapse. These observations can be seen in table 13.

	Worst single return observation			Worst year by cumulative return		
	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Six-factor	-10.43%	-14.49%	-19.40%	-28.25%	-29.22%	-29.98%
FF5F	-10.56%	-15.01%	-18.95%	-28.62%	-30.09%	-31.40%
Carhart	-10.05%	-15.26%	-19.91%	-27.82%	-28.51%	-30.47%
Value-mom	-9.61%	-15.20%	-22.29%	-27.54%	-28.88%	-31.28%
SIM	-9.82%	-15.13%	-21.58%	-27.72%	-29.33%	-32.17%
$\frac{1}{N}$	-11.30%	-16.31%	-25.21%	-28.19%	-31.81%	-26.49%
Market	-17.44%	-18.00%	-23.24%	-37.62%	-40.19%	-37.78%

Table 13: *The single worst return observations, and the worst years by cumulative returns. Among the single worst observations, the values can not be compared between the measurement frequencies, as they are reported as daily, weekly and monthly returns, respectively. The difference between the worst year for the daily and monthly market portfolios are likely to be caused by rounding errors in the data set. The discrepancy of the worst year of the weekly market portfolio comes from the way we have created the weekly data set as described in section 4.1.2. The returns of one week are all placed on the Friday of said week. This means that some returns that might have actually been at the end of December will be reported as early January instead.*

Looking at crash-risk through these measures, it seems that our comparison portfolios, and especially the market, struggle more than our factor-based portfolios. Among the portfolios formed on daily data, the worst day for any of our factor-based portfolios had a return of -10.56% , while the market, on its worst day, had a return of -17.44% . Even though the equivalent difference among the portfolios formed on weekly and monthly data is smaller, it is still the case that the market performs worse than any of our factor-based portfolios. However, in the case of the portfolios formed on monthly data, it is not the market, but rather the $\frac{1}{N}$ portfolio which experiences the worst month of our portfolios, with a return of -25.21% .

Looking at the worst years by cumulative return, a similar pattern is seen, although there are some noteworthy differences. The market is the worst portfolio for each of the three measurement frequencies. However, by this measure on the portfolios formed on daily data, the $\frac{1}{N}$ portfolio lies in the middle of our factor-based portfolios. For the portfolios formed on monthly data meanwhile, the $\frac{1}{N}$ portfolio has the best return in its worst year.

It is noteworthy that several factor-based portfolios share the same dates or the same years for their worst returns, when they are formed on the same frequency. This is yet another indication of how similar these portfolios are.

5.4.1 Skewness and kurtosis

Two statistical measurements that can be useful when assessing the crash-risk of the portfolios are skewness and kurtosis, which in combination tells us about the shape of the return distribution. To better compare the portfolios by these two measures, we have created a composite measure, which we have called *crash-risk measure*, by multiplying them. A positive skewness is desirable, while a negative skewness is bad, and a high kurtosis strengthens the skewness value. Therefore, a positive skewness along with high kurtosis is good, and a negative skewness with high kurtosis is bad. This means that a negative crash-risk measure is bad, while a positive crash-risk measure is good, and the magnitude says something about how bad/good the measure is.

	Daily			Weekly			Monthly		
	Skewness	Kurtosis	Crash-risk	Skewness	Kurtosis	Crash-risk	Skewness	Kurtosis	Crash-risk
Six-factor	-0.23	31.22	-7.28	-0.83	9.00	-7.50	-0.44	6.60	-2.89
FF5F	-0.34	29.99	-10.25	-0.85	9.38	-8.01	-0.41	6.75	-2.79
Carhart	-0.33	28.69	-9.37	-0.82	9.34	-7.62	-0.40	6.81	-2.74
Value-mom	-0.50	26.43	-13.24	-0.80	9.30	-7.49	-0.52	6.69	-3.45
SIM	-0.62	25.61	-15.90	-0.84	9.20	-7.74	-0.52	7.09	-3.66
$\frac{1}{N}$	-0.46	10.42	-4.83	-0.52	6.45	-3.34	-0.41	5.39	-2.22
Market	-0.50	18.25	-9.20	-0.44	8.14	-3.57	-0.52	4.79	-2.50

Table 14: *The skewness and kurtosis measures of the different portfolios, measured for the entire return series. Thus, the values for the portfolios formed on daily and weekly data are based on returns from 1965-2016, while they are based on returns from 1969-2016 for the portfolios formed on monthly data. We have also included our crash-risk measure, which is the product of the skewness and the kurtosis.*

Looking at the portfolios based on daily data, there are no clear patterns. Some of our factor-based portfolios have a lower skewness than our comparison portfolios, while some have the same or higher. It is however clear that the daily $\frac{1}{N}$ portfolio has the lowest kurtosis, while the factor-based portfolios have higher value for this measure than either comparison portfolio. When we look at the crash-risk measure, the $\frac{1}{N}$ portfolio has the best value. For the factor-based portfolios, meanwhile, the crash-risk measure varies, with the six-factor portfolio being better than the market, and the others worse.

The portfolios formed on weekly data show something different by these measures. Each of the factor-based portfolios has both a worse skewness and a higher kurtosis than either of our comparison portfolios. The market has the best skewness, while the $\frac{1}{N}$ portfolio has the best kurtosis measurement. By our crash-risk measure, the factor-based portfolios formed on weekly data are thus considerably worse.

Lastly, when looking at the skewness of the portfolios formed on the monthly data, there is no clear pattern. All of our portfolios have similar values for this measure, although the best value

is seen by the Carhart portfolio. The comparison portfolios do however have slightly slimmer tails as shown by their kurtosis measures. This leads to the comparison portfolios having slightly better crash-risk measures, although the difference is smaller than what we saw for the weekly portfolios.

Overall, there does seem to be some added crash-risk in our factor-based portfolios when looking at skewness and kurtosis. This is particularly seen in the portfolios formed on weekly data, and to some degree also in the monthly and daily portfolios. However, among the daily portfolios the six-factor portfolio seems to be better than the market. It is however, also worth noting that the comparison portfolios have worse returns at their worst observations. This suggests that, instead of our factor-based portfolios having longer tails when their negative skewness is higher, it is rather that the negative tail is thicker than the positive tail.

5.4.2 Crash-risk in the worst times

When looking at the crash-risk of our portfolios, we believe it is useful to look at the skewness and kurtosis in their worst times. To do this, we have calculated the cumulative return for each portfolio each year, and picked the worst periods. Because of low sample sizes when looking at individual years, we exclude the portfolios formed on monthly data from this analysis. We start by including the four years in which one or more of our portfolios had their worst year, which is 1969, 1973, 1974 and 2008. Since 1973 and 1974 are consecutive years, we decide to combine them into one set, to get a better sample. To get a more diverse sample of bad times, we add two other years. 1990, in which every portfolio had a relatively bad return, and 2002, in which the factor-based portfolios had a positive return, while our comparison portfolios had a negative return. We believe that this sample is relatively diverse, both in terms of being spread out chronologically, and in what the observed returns are.

Year	Skewness					Kurtosis					Crash-risk				
	69	73/74	90	02	08	69	73/74	90	02	08	69	73/74	90	02	08
Six-factor	-0.55	-0.45	-1.50	-0.49	-0.11	6.09	10.23	9.77	5.39	5.63	-3.36	-4.60	-14.61	-2.65	-0.61
FF5F	-0.56	-0.43	-1.52	-0.49	0.03	6.15	10.59	9.91	5.46	5.89	-3.43	-4.59	-15.09	-2.65	0.20
Carhart	-0.57	-0.44	-1.52	-0.53	-0.02	6.09	10.55	9.74	5.42	5.71	-3.46	-4.62	-14.77	-2.90	-0.10
Value-mom	-0.65	-0.40	-1.66	-0.63	-0.02	6.30	10.90	9.58	5.22	5.75	-4.12	-4.38	-15.86	-3.28	-0.12
SIM	-0.62	-0.39	-1.70	-0.63	0.20	6.02	11.03	9.77	5.19	5.28	-3.75	-4.28	-16.59	-3.25	1.23
$\frac{1}{N}$	-0.26	0.20	-0.95	0.20	-0.02	4.15	4.08	6.62	2.96	4.15	-1.11	0.83	-6.33	0.60	-0.10
Market	-0.01	0.36	-0.25	0.48	0.13	3.34	3.82	4.15	3.63	6.56	-0.03	1.38	-1.03	1.72	0.85

Table 15: *The skewness and kurtosis measures of the portfolios formed on daily data, measured for the bad periods we have chosen for further analysis. The periods are identified by cumulative returns in a calendar year. We have also included our crash-risk measure, which is the product of the skewness and the kurtosis.*

In table 15 we present the skewness, kurtosis and our crash-risk measure in these bad periods

for the portfolios formed on daily data. It is clear that the kurtosis measures in all of these periods are lower than what we saw for the entire return series in table 14. This is likely to be caused by these periods being relatively short, which means that there is not as big a spread in the returns as there are in the longer period. This is also supported by the fact that the factor-based portfolios have their highest kurtosis in the two-year period, 1973/74, rather than any of the single years. Another point worth noting, is that there are some portfolios that have a positive skewness in certain periods, especially the comparison portfolios. In 2008, the FF5F- and SIM portfolios have positive skewness values, which leads to the SIM portfolio having the best crash-risk measure of any portfolio in this year.

The year that most clearly sticks out among the portfolios formed on daily data is 1990. Every portfolio has its worst skewness measure, and one of its two worst kurtosis measures in this year. This results in all of the portfolios having their worst crash-risk measurement in 1990. It is however worth noting that the factor-based portfolios have a considerably worse crash-risk measure than the comparison portfolios.

Year	Skewness					Kurtosis					Crash-risk				
	69	73/74	90	02	08	69	73/74	90	02	08	69	73/74	90	02	08
Six-factor	-0.49	-0.17	-1.09	-0.48	-0.36	2.66	5.68	5.79	3.96	4.72	-1.31	-0.95	-6.31	-1.91	-1.68
FF5F	-0.43	-0.16	-1.09	-0.63	-0.41	2.64	5.68	5.76	3.86	4.60	-1.14	-0.93	-6.26	-2.44	-1.86
Carhart	-0.49	-0.19	-1.12	-0.46	-0.30	2.75	5.82	5.93	4.04	4.75	-1.35	-1.09	-6.67	-1.87	-1.42
Value-mom	-0.42	-0.11	-1.06	-0.42	-0.25	2.94	5.53	4.93	4.07	4.45	-1.23	-0.62	-5.24	-1.71	-1.13
SIM	-0.43	-0.12	-1.05	-0.64	-0.30	2.93	5.75	4.91	4.01	4.43	-1.27	-0.72	-5.18	-2.58	-1.31
$\frac{1}{N}$	-0.28	0.67	-0.64	-0.24	-0.18	2.88	3.81	2.87	2.69	4.07	-0.80	2.57	-1.84	-0.65	-0.73
Market	-0.38	0.70	-0.20	0.04	-0.38	2.50	5.46	2.29	3.18	6.19	-0.95	3.82	-0.47	0.12	-2.36

Table 16: *The skewness and kurtosis measures of the portfolios formed on weekly data, measured for the bad periods we have chosen for further analysis. The periods are identified by cumulative returns in a calendar year. We have also added our crash-risk measure, which is the product of the skewness and the kurtosis.*

Table 16 show skewness, kurtosis and our crash-risk measure for the portfolios formed on weekly data. Again we see that the kurtosis is somewhat lower for all portfolios than what we saw for the entire return series in table 14, although the differences are smaller than what we saw for the daily portfolios. Among the weekly portfolios there are fewer instances of positive skewness than what we saw among the daily portfolios, and none of our factor-based portfolios experience any positive skewness.

For our factor-based portfolios, as well as the $\frac{1}{N}$ portfolio, formed on weekly data, 1990 is again the worst year, as measured by skewness, and the composite crash-risk measure. For the market however, 2008 was the worst of these years, and this was also its worst year measured by cumulative returns. It is also clear from the table, that although 1973/74 was a bad period for all of the portfolios, our comparison portfolios were highly resilient to crashing in this period, because of their relatively high positive skewness. This period also has the best crash-risk

measure for all of our factor-based portfolios, although these values remain negative.

5.4.3 Overall crash-risk

It should be noted that it is difficult to compare crash-risk across frequency measurements. This becomes clear when looking at the kurtosis of the market, which is substantially different between the three frequencies, although they ultimately represent the same portfolio. This will naturally extend to our composite crash-risk measure, which is the product of the skewness and kurtosis.

It is difficult to conclude which portfolio has the greatest crash-risk by the measures we have presented. The worst observations and calendar years point towards the market, while skewness and kurtosis indicate that our factor-based portfolios are in greater danger of crashing. Based on these findings, we would conclude that there is no great risk of a complete collapse of our factor-based portfolios.

5.5 Factor loadings in portfolios

In addition to our analysis of the portfolio performance we find it interesting to look at the different factor loadings for our selected sample of factor-based portfolios. We hope that this might explain why our portfolios seem to outperform the market both for Sharpe- and appraisal ratios. We also wish to check the similarity of the various portfolios, given that their performance measures are so similar.

Tables 17 and 18 show the factor loadings for our subset of portfolios formed on daily and weekly data¹. From these tables it is clear that the market, size and momentum factors are significantly loaded for every portfolio formed on both daily and weekly data. These factors are highly significant for every one of our portfolios, regardless of whether they played a part in forming the portfolio. Every one of these loadings are also positive. While all the portfolios are positively exposed to these factors, the portfolios formed on weekly data have higher coefficients than those formed on daily data.

Among the other factors, *CMA* has the most significant exposures. This factor is significantly and negatively loaded in each of the portfolios formed on daily data. The *HML*- and *RMW*-factors meanwhile, are only significantly loaded in the value-mom portfolio formed on daily data, in which they are both negatively loaded.

The significantly loaded factors also seem to have very similar loadings for every portfolio based on the same measurement frequency. This further suggests that there is not a very large

¹We have, for readability purposes, excluded monthly data from this part of our thesis. However, the factor loadings for the portfolios formed on monthly data largely follow the same patterns as those highlighted here.

	Six-factor	FF5F	Carhart	Value-mom	SIM
α	0.0025 (10.367)	0.0025 (10.283)	0.0025 (10.280)	0.002 (10.546)	0.002 (10.283)
Mkt	0.404 (79.996)	0.405 (79.866)	0.403 (79.967)	0.401 (83.721)	0.403 (84.537)
SMB	0.369 (40.484)	0.371 (40.604)	0.369 (40.646)	0.395 (45.705)	0.398 (46.329)
HML	-0.015 (-1.360)	-0.010 (-0.883)	-0.016 (-1.401)	-0.027 (-2.472)	-0.013 (-1.173)
RMW	-0.013 (-1.024)	-0.011 (-0.822)	-0.012 (-0.899)	-0.041 (-3.326)	-0.022 (-1.819)
CMA	-0.051 (-3.238)	-0.053 (-3.377)	-0.043 (-2.761)	-0.065 (-4.385)	-0.063 (-4.265)
MOM	0.121 (17.833)	0.125 (18.389)	0.121 (17.903)	0.128 (19.973)	0.136 (21.350)
Obs.	13090	13090	13090	13090	13090
Adj. R^2	0.3307	0.420	0.420	0.458	0.458

Table 17: Factor exposures for the Six-factor, FF5F, Carhart, Value-mom and SIM portfolios formed on daily data. We run the following regression for the portfolio using daily returns from 1965 through 2016:

$$r_{p,t} = \alpha_p + \beta_p \cdot \text{Mkt}_t + s_p \cdot \text{SMB}_t + h_p \cdot \text{HML}_t + r_p \cdot \text{RMW}_t + c_p \cdot \text{CMA}_t + m_p \cdot \text{MOM}_t + \varepsilon_p$$

where the regression coefficients show the portfolio exposure to the different factors. T -stats are listed in parentheses. Greyed out coefficients and t -stats indicate a factor that was not used in forming the portfolio. Given that there are typically five market days in one week, we have multiplied the daily α 's by five to be able to compare them with α -values of the portfolios formed on weekly data.

difference between the investments of the various portfolios. This helps explain the small, insignificant differences in the Sharpe- and appraisal ratios of these portfolios. There are however some differences between the factor loadings when comparing across measurement frequencies. Specifically, among those factors significantly loaded in all portfolios, they seem to be more heavily loaded in the portfolios formed on weekly data.

The significant factor loadings could also help explain why the portfolios beat the market. When looking at the correlation matrices of the different factor sets², the daily market factor has a negative correlation with all the other factors, while for the weekly factors, all but *SMB* has a negative correlation with the market. This means that our portfolios, being positively loaded in both size and momentum, have taken advantage of beneficial covariances, which are likely to lead to better performance.

As a last point of note, each of the portfolios created on daily and weekly data has a significant α . It is clear that the daily portfolios have a higher α than the portfolios formed on weekly data. Specifically, the daily portfolios get an added weekly return between 0.2% and 0.25%, while the weekly portfolios only get an added return between 0.05% and 0.07%. However, among the

²These matrices are not included in the thesis, but will be made available upon request.

	Six-factor	FF5F	Carhart	Value-mom	SIM
α	0.0005 (3.203)	0.0006 (2.995)	0.0006 (3.058)	0.0007 (3.476)	0.0006 (3.245)
Mkt	0.624 (62.857)	0.631 (62.266)	0.623 (61.143)	0.633 (63.102)	0.638 (63.661)
SMB	0.439 (26.272)	0.438 (25.705)	0.440 (25.669)	0.494 (29.309)	0.504 (29.901)
HML	0.024 (1.125)	0.023 (1.070)	0.021 (0.975)	0.008 (0.367)	0.032 (1.524)
RMW	0.010 (0.434)	0.011 (0.479)	0.009 (0.389)	-0.039 (-1.707)	-0.005 (-0.209)
CMA	-0.003 (-0.099)	-0.012 (-0.379)	-0.008 (-0.254)	-0.037 (-1.206)	-0.048 (-1.586)
MOM	0.161 (14.180)	0.157 (13.574)	0.164 (14.095)	0.176 (15.380)	0.182 (15.890)
Obs	2714	2714	2714	2714	2714
Adj. R^2	0.688	0.684	0.677	0.705	0.707

Table 18: Factor exposures for the Six-factor, FF5F, Carhart, Value-mom and SIM portfolios formed on weekly data. We run the following regression for the portfolio using weekly returns from 1965 through 2016:

$$r_{p,t} = \alpha_p + \beta_p \cdot \text{Mkt}_t + s_p \cdot \text{SMB}_t + h_p \cdot \text{HML}_t + r_p \cdot \text{RMW}_t + c_p \cdot \text{CMA}_t + m_p \cdot \text{MOM}_t + \varepsilon_p$$

where the regression coefficients show the portfolio exposure to the different factors. T-stats are listed in parentheses. Greyed out coefficients and t-stats indicate a factor that was not used in forming the portfolio.

portfolios formed on the same frequency data, there does not seem to be any portfolio that has a significantly better α than the others.

5.6 Why are the portfolio performances so similar?

Throughout our analysis so far, it is clear that the various factor-based portfolios have remarkably similar performances. This is apparent both when looking at their performance measures and their factor loadings. Despite being based upon different numbers and combinations of factors, our portfolios seem to pick very similar stocks. We therefore decide to look deeper at our calculation of both the vector expected returns and the covariance matrix of the stocks.

The first thing we realise, is that the vector of expected returns in our calculations should be identical to that calculated by using the average of the returns. Because all of the average returns should be included in the α 's or the product of the factor coefficients and factor returns, none of the actual returns should be unaccounted for.

When looking at the covariance matrix however, there are some minor differences. The variance of the individual stocks, contained in the main diagonal of the matrix, should be exact, as we include both the variance described by the factors and the residual variance. However, when

we calculate the covariance between the different stocks, some minor discrepancies do appear. Specifically, our covariance matrix does not take into account the share of the covariance that would be explained by the residuals in our regressions.

The reason why our portfolios are so similar to each other is therefore that there are only minor differences in the covariance matrices that we calculate. What we end up with is not a portfolio that invests in a specific set of factors, but rather a different approach to calculating the covariance matrix used in the Markowitz theory. That is to say, our framework creates approximations of the standard Markowitz framework, and it seems to be a better, or at least more stable way of implementing it. The inconsistencies that appear in the standard Markowitz implementation seem to disappear, and despite running 1 520 optimisations, we never experience a covariance matrix that is non-positive definite.

5.6.1 Which approximation is better?

It is impossible to conclude which of our portfolios is best as an investment strategy given the similarities in performance. However, as we seem to have identified a more stable way of implementing the Markowitz framework, we wish to compare them in that regard. Specifically, which set of factors should be considered the best in implementing our framework.

There are two elements to consider in this discussion. These are the number of estimates to be calculated, and the amount of covariance that is unaccounted for by being left in the residuals. The portfolio based on the SIM, using only one factor, will produce the lowest number of estimates, while the six-factor model will produce the greatest number. Returning to table 1 in section 2.2.1, the six-factor model produces 8 027 estimates for 1 000 stocks, while using the SIM would, by the same calculations, produce only 3 002.

On the other hand, the six-factor model, using more explanatory variables, should have smaller residuals, thereby including more relevant information in the covariance matrix. By this measure, the SIM should produce a worse estimate of the covariance matrix, as it only includes one explanatory variable.

We believe that, since all of our models produce stable Markowitz implementations, the calculation of the covariance matrix should be the primary concern. We therefore conclude that all the available factors should be included in the estimation of this matrix, and the six-factor model should be preferred.

6 Conclusion

We create a framework which combines the traditional Markowitz (1952) theory of portfolio optimisation with established factor models. This framework calculates the expected returns and the covariance matrix of the stocks by using factor regression coefficients, and the expected returns and covariance matrix of the factors. It calculates the stock weights in the same way as the standard Markowitz framework, finding the combination with the highest *ex-ante* Sharpe ratio.

We find that, by investing in a portfolio based upon our framework, using daily data allows an investor to beat the market, as measured by its *ex-post* Sharpe ratio. This improvement in performance over the market portfolio is significant on the 5% level. While the Sharpe ratio is not significantly better than the market when the portfolios are formed on weekly or monthly data, there are some indications that our framework is better. The main indication that all the portfolios formed on our framework beat the market comes from the appraisal ratio, for which all of them have a significant mean one/four-year value.

What is most surprising in our results is the magnitude of the Sharpe- and appraisal ratios obtained for our portfolios formed on daily data. The values observed are all considered very high, and might be cause for some concern as to the validity of our results. However, we believe that our methodology is solid, and these ratios are based upon the realised returns of the portfolios, rather than the expected values.

We would however recommend more caution when looking at the periodical Sharpe- and appraisal ratios of the portfolios formed on monthly data. Because of the comparatively low number of return observations, the sample of four-year measures was only 12. Even though we applied the bootstrap method to this data, creating 10 000 sample means of these values, the low number is still some cause for concern. However, the mean of the Sharpe ratio was not significantly better than that of the market, and results of that analysis are, at any rate, inconclusive. However, the mean values of the appraisal ratios were found to be significant, and so, the conclusions based upon this should be handled with care.

While there are strong indications of our framework producing better results than the market, there are no differences between the various portfolios based on it. We found that what our framework truly creates, is not a factor-based investment portfolio as such, but rather a more stable approximation of the Markowitz framework. While the standard framework often experiences problems with a non-positive definite covariance matrix, which could lead to a negative expected variance, this problem did not occur in our implementation. Over the course of creating our factor-based portfolios on this framework, 1 520 covariance matrices were calculated, and all of them were positive definite.

Based upon our results, we find a couple of areas that we believe would be interesting to expand upon in further research:

1. The inclusion of shorting might help show significant improvements over the market for weekly and/or monthly data. While it is difficult to account for the constraints and costs involved in shorting, a significantly better result than ours might be an indication that it is better nonetheless.
2. To further see whether it is possible to get significant results for weekly and monthly data, a longer time period might be used. While the full set of factors is not available further back than mid-1963, some are. Specifically, the factors from the Fama-French three-factor model, plus the momentum factor, is available as far back as 1926. Alternatively, if the necessary accounting data is available, the factors could be created for a greater time period.
3. It could be interesting to find a way to implement both our framework and the standard Markowitz framework, to see how they compare as investment strategies. While our framework is more stable in terms of avoiding problematic covariance matrices, it does not mean that the resulting investment is necessarily better.
4. Checking our framework in a different market could potentially strengthen the conclusions we have reached. This would be particularly interesting if done in conjunction with the previous point of implementing a standard Markowitz framework. We believe our framework would have its greatest advantage over the standard in a big market, so a test on two markets of different size would be interesting.

7 Bibliography

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). *Value and Momentum Everywhere*. *The Journal of Finance*, 68(3), 929-985.

Ball, R. (1978). *Anomalies in relationships between securities yields and yield-surrogates*. *Journal of Financial Economics*, 6(2-3), 103-126.

Bodie, Z, Kane, A, & Marcus, A. J. (2014). *Investments*. Maidenhead: McGraw-Hill Education.

Brückner, R., Lehmann, P., Schmidt, M. H., & Stehle, R. (2015). *Non-U.S. Multi-Factor Data Sets Should be Used with Caution*. Working paper.

Carhart, M. M. (1997). *On Persistence in Mutual Fund Performance*. *The Journal of Finance*, 52(1), 57-82.

Chen, J., Hong, H., & Stein, J. C. (2001). *Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices*. *Journal of Financial Economics*, 61(3), 345-381.

Christy, G. A., & Clendenin, J. C. (1982). *Introduction to investments*. Whitby, Canada: McGraw-Hill Ryerson, Limited.

Cohen, R., Gompers, P., & Vuolteenaho, T. (2002). *Who Underreacts to Cash-Flow News? Evidence from Trading between Individuals and Institutions*. *Journal of Financial Economics*, 66(2), 409-462.

CRSP Stocks (2017). Available: Center For Research in Security Prices. Graduate School of Business. University of Chicago [15.09.2017]. Retrieved from Wharton Research Data Service.

Dowle, M & Srinivasan, A (2017). *data.table: Extension of 'data.frame'*. R package version 1.10.4. <https://CRAN.R-project.org/package=data.table>

Efron, B. & Tibshirani, R. J. (1993). *An introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall

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.

Fama, E. F., & French, K. R. (1995). *Size and Book-to-Market Factors in Earnings and Returns*. *The Journal of Finance*, 50(1), 131-155.

Fama, E. F., & French, K. R. (1996). *Multifactor Explanations of Asset Pricing Anomalies*. *The Journal of Finance*, 51(1), 55-84.

- Fama, E. F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of Financial Economics*, 82(3), 491-518.
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.
- French, K. R. (2017). Kenneth R. French's Data library. Retrieved November 21, 2017, from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
- Goetzmann, W., Zhu, N., & Bris, A. (2003). Efficiency and the Bear: Short Sales and Markets around the World. *The Journal of Finance*, 62(3), 1029-1079.
- Haugen, R. A., & Baker, N. L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41(3), 401-439.
- Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1), 65-91.
- Joanes, D. N., & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), 183-189.
- Kritzman, M. (1993). What Practitioners Need to Know... About Return and Risk. *Financial Analysts Journal*, 49(3), 14-17.
- Lintner, J. (1965). Security Prices, Risk, and Maximal Gains From Diversification. *The Journal of Finance*, 20(4), 587-615.
- Liu, L., Whited, T., & Zhang, L. (2009). Investment-Based Expected Stock Returns. *Journal of Political Economy*, 117(6), 1105-1139.
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77-91.
- Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York, NY: John Wiley & Sons, Inc
- Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), 768-783.
- NASDAQ Inc. (2017). NASDAQ Stock Market. Retrieved 12 7, 2017, from *Equity Rules*: <http://nasdaq.cchwallstreet.com/>
- Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1), 1-28.
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.

Rmetrics Core Team, Diethelm Wuertz, Tobias Setz and Yohan Chalabi (2014). fBasics: Rmetrics - Markets and Basic Statistics. R package version 3011.87.

<https://CRAN.R-project.org/package=fBasics>

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. Journal of Economic Theory, 13(3), 341-360.

Rouwenhorst, K. G. (1998). International Momentum Strategies. The Journal of Finance, 53(1), 267-284.

SEC. (2013, 5 9). United States Securities and Exchange Commission. Retrieved 12 7, 2017, from Penny Stock Rules: <https://www.sec.gov/fast-answers/answerspennyhtm.html>

Sharpe, W. F. (1964). Capital Asset Prices: A Theory Of Market Equilibrium Under Conditions Of Risk. The Journal of Finance, 19(3), 425-442.*

Sharpe, W. F. (1966). Mutual Fund Performance. The Journal of Business, 39(1), 119-138

Titman, S., Wei, K. J., & Xie, F. (2004). Capital Investments and Stock Returns. The Journal of Financial and Quantitative Analysis, 39(4), 677-700.

Treynor, J. L. (1962). Toward a Theory of Market Value of Risky Assets . Available at SSRN: <https://ssrn.com/abstract=628187>

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

Warwick, B. (2003). The Handbook of Risk. Hoboken, NJ: Wiley.