DEGREE PROJECT IN MATHEMATICS,
SECOND CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2019

Maximum Predictability Portfolio Optimization

NAZIM HUSEYNOV

KTH ROYAL INSTITUTE OF TECHNOLOGY
SCHOOL OF ENGINEERING SCIENCES

Maximum Predictability Portfolio Optimization

NAZIM HUSEYNOV

Degree Projects in Financial Mathematics (30 ECTS credits)
Degree Programme in Applied and Computational Mathematics
KTH Royal Institute of Technology year 2019
Supervisor at KTH: Boualem Djehiche
Examiner at KTH: Boualem Djehiche

TRITA-SCI-GRU 2019:032
MAT-E 2019:13

Royal Institute of Technology
School of Engineering Sciences
KTH SCI
SE-100 44 Stockholm, Sweden
URL: www.kth.se/sci

Sammanfattning

Portföljoptimering med maximal prediceringsgrad.

Modern portföljteori har sitt ursprung i Harry Markowitz arbete på 50-talet. Teorin ger investerare kvantitativa verktyg för att sammansätta och utvärdera tillgångsportföljer på ett systematiskt sätt. Huvudsakligen går Markowitz idé ut på att komponera en investeringsportfölj genom att lösa ett kvadratiskt optimeringsproblem.

Det här examensprojektet har utgångspunkt i *Maximally Predictable Portfolio*-ramverket, utvecklat av Lo och MacKinley som ett alternativ till Markowitz problemformulering, i syfte att välja ut investeringsportföljer. En av fördelarna med att använda den förra metoden är att den tar hänsyn till uppskattningsfelen från prognostisering av framtida avkastning. Vår investeringsstrategi är att köpa och behålla dessa portföljer under en tidsperiod och bedöma deras prestanda. Resultaten visar att det mha. MPP-optimering är möjligt att konstruera portföljer med hög avkastning och förklaringsvärde baserat på historisk data. Trots sina många lovande funktioner är framgången med MPP-portföljer kortlivad. Baserat på vår bedömning drar vi slutsatsen att investeringar på aktiemarknaden uteslutande på grundval av optimeringsresultatet inte är en lukrativ strategi.

Nyckelord: Portföljoptimering, linjär optimering, multifaktormodel

Abstract

Harry Markowitz work in the 50's spring-boarded modern portfolio theory. It gives investors quantitative tools to compose and assess asset portfolios in a systematic fashion. The main idea of the *Mean-Variance* framework is that composing an optimal portfolio is equivalent to solving a quadratic optimization problem.

In this project we employ the *Maximally Predictable Portfolio* (MPP) framework proposed by Lo and MacKinlay, as an alternative to Markowitz's approach, in order to construct investment portfolios. One of the benefits of using the former method is that it accounts for forecasting estimation errors. Our investment strategy is to buy and hold these portfolios during a time period and assess their performance. We show that it is indeed possible to construct portfolios with high rate of return and coefficient of determination based on historical data. However, despite their many promising features, the success of MPP portfolios is short lived. Based on our assessment we conclude that investing in the stock market solely on the basis of the optimization results is not a lucrative strategy.

Keywords: Portfolio optimization, linear programming, multi-factor model

Acknowledgements

I would first like to thank my thesis advisor Professor Boualem Djehiche at the Department of Mathematical Statistics at KTH. He made it possible for me to explore this subject and gave invaluable guidance when needed.

I would also like to thank my friend Jonas Ivarsson for all the input and feedback with regards to this paper.

Stockholm, February 2019

Nazim Huseynov

Contents

Chapter 1

Introduction

Harry Markowitz' work in the 50's spring-boarded modern portfolio theory. It gave investors quantitative tools to compose and assess asset portfolios in a systematic fashion. First, Markowitz quantified return and risk of a security, using the statistical measures of its expected return and standard deviation. Second, Markowitz suggested that investors should consider return and risk together, and determine the allocation of funds among investment alternatives on the basis of their return-risk trade-off [1]. Essentially, Markowitz proposition was that composing an optimal portfolio is equivalent to solving a constrained quadratic optimization problem, e.g. risk minimization given a utility constraint (Section 2.4).

Perhaps, the striking simplicity of the model is the reason why it became so popular among researchers [1]. In order to produce the set of optimal weights, it requires both a covariance matrix and a vector of means as input. Since the nature of the inputs is random, they need to be estimated with standard techniques. However, estimation errors in the forecasts have been shown to significantly impact the resulting portfolio weights, which in turn puts reliability of the model in question [1]. Furthermore, the model does not take into account individual investor guidelines such as transaction cost constraints and risk profile. The many criticisms concerning reliability and applicability of the original model have been hindering its broader adoption in the industry. However, they also spurred extensive research in the field resulting in various approaches to alleviate those issues.

One such approach that deals with estimation uncertainty was proposed by Lo and MacKinley [2]. *Maximally Predictable Portfolio* (MPP) framework builds on the idea of *predictability maximization*, that stems from substantial evidence for forecasting ability of various risk-premium models (e.g. Capital Asset Pricing Model, Fama-French three-factor Model) to predict asset returns [2]. Despite considerable variation among assets and over time, predictability, defined in terms of sensitivity to different market risk-premia, is indeed present, and both statistically and economically significant [2].

Two complications arise in the MPP framework. First, forecasting asset returns requires a suitable information structure. *Multi-linear factor modelling* is a well-studied and well-established tool, and of particular interest is the model proposed by Eugene Fama and Kenneth French (Section 2.3). Second, the objective function in the MPP model is a fraction, with quadratic terms in both numerator and denominator. This type of problems are known as *convex-convex fractional programming* problems, solution to these involve mathematical theory beyond what is taught at a university masters program. Luckily, this problem has already been studied in academic research, and a solution method is reviewed in detail in Chapter 2.

The goal of this thesis project is to answer the following questions:

- Can the framework of *Maximally Predictable Portfolio* be used to generate excess returns in the Swedish stock market?
- Are the obtained results valid as an investment strategy?

The scope of the project is limited to answering the above questions in the setting of the Swedish stock market (Large Cap, Mid Cap, Small Cap) during the period January 1, 2012 - December 31, 2018. Only stocks with available historical data will be used in analysis.

Chapter 2

Theory

2.1 Portfolio metrics

Let $[0, T]$ be the historical time period of interest. Let N denote the number of equidistant sub-periods in $[0, T]$. Define $\Delta t = \frac{T}{N}$ as the duration of each sub-period $[t_{k-1}, t_k]$, i.e. $\Delta t = t_k - t_{k-1},\ k = 1, \ldots, N$. Note, that $t_k = k\Delta t$, $t_0 = 0$ and $t_N = T$.

2.1.1 Asset and portfolio return

Let random variable S_{i,t_k} be the price of i-th portfolio asset at time t_k and let C_{i,t_k} be random variable representing the cash flow associated with that asset at time t_k. Then, random variable $R_{i,k}$ of the return of i-th portfolio asset from sub-period k is given by

$$R_{i,k} = \frac{S_{i,t_k} + C_{i,t_k} - S_{i,t_{k-1}}}{S_{i,t_{k-1}}}. \tag{2.1}$$

If no cash flow occurs in the sub-period, the asset's return in sub-period k is

$$R_{i,k} = \frac{S_{i,t_k} - S_{i,t_{k-1}}}{S_{i,t_{k-1}}}. \tag{2.2}$$

The cumulative return of i-th asset the interval $[0, T]$, also known as *performance*, is calculated as

$$R_i = \prod_{k=1}^{N} (1 + R_{i,k}) - 1. \tag{2.3}$$

Given a set of portfolio weights $\left\{w_i\right\}_{i=0}^{p}$, where w_i corresponds to asset i, the portfolio return in sub-period k is

$$R_{pf,k} = \sum_{i=0}^{p} w_i R_{i,k}, \tag{2.4}$$

while the cumulative portfolio return over time interval $[0,T]$ is

$$R_{pf} = \sum_{i=0}^{p} w_i R_i, \tag{2.5}$$

2.1.2 Volatility

Portfolio volatility is a measure of *market price risk*, defined as standard deviation of portfolio returns:

$$\sigma_{pf} = \sqrt{Var(R_{pf})} \tag{2.6}$$

Note, that here R_{pf} denotes the random vector of portfolio returns. For calculation purposes, the unbiased estimator of variance is used:

$$\sigma_{pf} = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(r_{pf,k} - \bar{r}_{pf})^2} \tag{2.7}$$

$$\bar{r}_{pf} = E\Big[R_{pf}\Big] = \frac{1}{N}\sum_{k=1}^{N} r_{pf,k}. \tag{2.8}$$

2.1.3 Beta

The key idea of Modern Portfolio Theory is that *diversification reduces risk*. Specifically, it reduces *unsystematic*, or *idiosyncratic* risk, i.e. risk that is endemic to a particular asset and not portfolio as a whole. Beta is a measure of *systematic* risk, irreducible through diversification, that arises from general exposure to markets. Market beta, that is, beta of a portfolio consisting of all assets traded on the market, is exactly 1. Portfolio beta indicates the sensitivity of an asset portfolio to general market movements. A large portfolio beta value, either positive or negative, indicates strong correlation with market, whereas a beta value close to 0 is characteristic of a market-agnostic portfolio.

Let β_{pf} denote portfolio beta, R_{pf} and R_{bm} denote random variable of portfolio and *benchmark* returns respectively. Typically, a benchmark is some market index, e.g. OMXS30, SIXPRX etc. Then

$$\beta_{pf} = \frac{Cov(R_{pf}, R_{bm})}{Var(R_{bm})} = \frac{1}{N-1}\sum_{k=1}^{N}\frac{(r_{pf,k} - \bar{r}_{pf})(r_{bm,k} - \bar{r}_{bm})}{\sigma_{bm}^2} \tag{2.9}$$

2.1.4 Sharpe ratio

Sharpe ratio, also known as *risk-adjusted return*, is a measure of investment *efficiency*, formally defined as the ratio of expected excess return to total risk:

$$SR = \frac{E\Big[R_{pf} - R_{rf}\Big]}{\sigma_{pf}} \tag{2.10}$$

where R_{rf} is the annual risk-free rate. Sharpe ratio is a *relative* performance metric, which makes it useful in assessing portfolios with different return and risk profiles.

2.1.5 Tracking error

Let $\Delta := R_{pf} - R_{bm}$ be random variable of excess return of portfolio over some benchmark. Tracking error is a *realtive* risk measure, that is defined as volatility of Δ:

$$TE = \sqrt{Var(\Delta)} \tag{2.11}$$

For calculation purposes, the unbiased estimator of variance is used.

$$TE = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(\delta_k - \bar{\delta})^2} \tag{2.12}$$

$$\bar{\delta} = E\Big[\Delta\Big] = \frac{1}{N}\sum_{k=1}^{N}\delta_k \tag{2.13}$$

Here, δ are observations of Δ.

2.1.6 Information ratio

In essence, Information ratio is the *relative* counterpart of Sharpe ratio, which is an *absolute* measure of investment efficiency. Information ratio is used to gauge the ability of fund manager to outperform benchmark. It is defined as ratio of excess portfolio return over some benchmark to Tracking error of same investment:

$$IR = \frac{E\Big[\Delta\Big]}{TE}. \tag{2.14}$$

2.2 Multi-linear regression and Coefficient of determination

Multi-linear regression is a standard tool in statistics used to determine the *linear* relationship between a dependent variable and a set of independent variables. Formally, given a set of n observations $(F_{1,1}, F_{1,2}, \ldots, F_{1,p}, R_1)$, $(F_{2,1}, F_{2,2}, \ldots, F_{2,p}, R_2)$, ..., $(F_{n,1}, F_{n,2}, \ldots, F_{n,p}, R_n)$, where $\left\{F_{k,i}\right\}_{k=1}^{p}$ are the explanatory variables for a given outcome R_i, and assuming, that the true relationship between the dependent and the independent variables is given by

$$R_i = \beta_0 + \beta_1 F_{1,i} + \beta_2 F_{2,i} + \ldots \beta_p F_{p,i} + \varepsilon_i = \boldsymbol{\beta}^T \boldsymbol{F}_i + \varepsilon_i, \tag{2.15}$$

where the ε represents model error such that $E[\varepsilon_i] = 0$ and $Var(\varepsilon_i) = \sigma_{\varepsilon_i}$. The objective is to find the *regression coefficients* $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^T$, such that the squared sum of model errors ε_k

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(R_i - \boldsymbol{\beta}^T \boldsymbol{F}_i\right)^2, \tag{2.16}$$

in matrix notation

$$||\boldsymbol{\varepsilon}||_2^2 = ||\boldsymbol{R} - \boldsymbol{F}\boldsymbol{\beta}||_2^2 \tag{2.17}$$

is minimized:

$$\boldsymbol{\beta}^* = \underset{\beta}{\operatorname{argmin}} ||\boldsymbol{R} - \boldsymbol{F}\boldsymbol{\beta}||_2^2 \tag{2.18}$$

Solution to problem (2.18), known as *Ordinary Least Squares* estimator, is given by

$$\boldsymbol{\beta}^* = (\boldsymbol{F}^T \boldsymbol{F})^{-1} \boldsymbol{F}^T \boldsymbol{R}. \tag{2.19}$$

It is common to assess the quality of the OLS regression by comparing how much the initial variation in the sample can be reduced by regressing onto independent variables. The *coefficient of determination* R^2 is defined as a ratio of model variance to the total variance of the dependent variable. Let $\hat{\boldsymbol{R}}$ denote the predicted values, based on the solution (2.19):

$$\hat{\boldsymbol{R}} = \boldsymbol{F}\boldsymbol{\beta}^* \tag{2.20}$$

Then, coefficient of determination is given by

$$R^2 = \frac{Var(\hat{\boldsymbol{R}})}{Var(\boldsymbol{R})} \tag{2.21}$$

2.3 Fama-French three-factor model

The Fama-French three-factor model (FF3M) was introduced in Fama and French (1993) as an extension of the Capital Asset Pricing Model (CAPM). The difference between CAPM and FF3M is that the former has a theoretical foundation, stemming from Markowitz' original work, while the latter is an ad hoc model that was introduced because it better fits the empirical data.

The main takeaway of CAPM is that an optimal portfolio is a combination of risk-free asset and market portfolio. Risk-free asset is typically represented by government bonds. Market portfolio is combination of all risky assets available in the market, that is typically represented by some suitable market index. In CAPM setting, expected return of an investment portfolio, denoted by R_{pf}, is a linear function of expected excess market return R_m:

$$E[R_{pf}] = \alpha + R_{rf} + \beta \times (E[R_m] - R_{rf}). \tag{2.22}$$

Forecasting portfolio return is a matter of calculating portfolio beta from betas of individual assets. The significance of beta is discussed in Section 2.1.

Validity of the model (2.22) builds on following assumptions [3]:

- All investors have homogeneous expectations, i.e. they expect the same probability distribution of returns.
- All investors want to invest in an optimal portfolio based on Markowitz's mean-variance framework, i.e. for a given expected return, they target the portfolio with the lowest volatility.
- All investors can lend and borrow any amount of money at the risk-free rate.
- All investors have the same one-period horizon.
- All assets are infinitely divisible.
- There are no taxes and transaction costs.
- There is no inflation or any change in interest rates or inflation is fully anticipated.
- Capital markets are *efficient*, i.e. they are in *equilibrium*.

FF3M extends model (2.22) with two additional factors, that represent observed phenomena in the market. Researchers found, that, with regards to returns, (1) small firms outperformed big firms, (2) firms with high book-to-market value outperformed firms with low book-to-market value. Fama and French showed, that these factors were significant predictors of portfolio performance. Denote the difference in rate of return between small and big firms with R_{smb}. Denote the difference in rate of return between firms with high book-to-market value and low book-to-market value with R_{hml}. According to FF3M, expected return of an investment portfolio is given by equation

$$E[R_{pf}] = \alpha + R_{rf} + \beta_1 \times (E[R_m] - R_{rf}) + \beta_2 \times E[R_{smb}] + \beta_3 \times E[R_{hml}]. \quad (2.23)$$

Beta parameters in (2.23) can be viewed as sensitivities to respective factors. However, they are not as significant as the beta in the original model (2.22).

2.4 Mean-Variance optimization

Let R_i be a random variable representing the rate of return of a risky asset S_i, $i = 1, \ldots, n$. Let w_i be the portion of the portfolio capital invested in asset S_i. For simplicity assume no short-selling. Additionally, impose upper bound on individual allocation, α_i. Furthermore, let $\mu_i = E[R_i]$ denote the average rate of return of asset S_i and let $\sigma_{ij} = Cov(R_i, R_j)$ denote the covariance between the i-th and j-th rate of return, where $\sigma_{ii} = \sigma_i^2 = Var(R_i)$ and $\sigma_{ij} = \sigma_{ji}$. Assume, that an investor is risk-averse in the sense that he or she prefers low σ_i:s, and seeks to minimize the risk of an investment portfolio, but at the same time requires some minimal return on investment, ρ. Assume further, that $R_i \sim N(\mu_i, \sigma_i)$. Then the set of optimal portfolio weights $\{w_i\}_{i=1}^n$ is obtained by solving the following optmization problem:

$$\begin{aligned} \underset{w}{\text{minimize}} \quad & \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \\ \text{subject to} \quad & \sum_{i=1}^n w_i = 1 \\ & \sum_{i=1}^n w_i \mu_i \geq \rho \\ & 0 \leq w_i \leq \alpha_i, \quad i = 1, \ldots, n. \end{aligned} \quad (2.24)$$

2.5 Maximally Predictable Portfolio

Consider a collection of n risky assets with returns $\boldsymbol{R}_t := (R_{1,t}, R_{2,t}, \ldots, R_{n,t})^T$. Assume, that $\boldsymbol{R}_t$ is a jointly stationary and ergodic process with finite expectation $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)^T$ and finite autocovariance matrix. Denote by $\boldsymbol{Z}_t$ vector of de-meaned asset returns, i.e. $\boldsymbol{Z}_t = \boldsymbol{R}_t - \boldsymbol{\mu}$. Let $\hat{\boldsymbol{Z}}_t$ denote some forecast of $\boldsymbol{Z}_t$ based on information set available at time $t-1$, and denote it $\boldsymbol{F}_{t-1}$, i.e. ,

$$\hat{\boldsymbol{Z}}_t = E[\boldsymbol{Z}_t|\boldsymbol{F}_{t-1}]. \tag{2.25}$$

We then may express $\boldsymbol{Z_t}$ as

$$\boldsymbol{Z}_t = E[\boldsymbol{Z_t}|\boldsymbol{F}_{t-1}] + \boldsymbol{\varepsilon}_t = \hat{\boldsymbol{Z}}_t + \boldsymbol{\varepsilon}_t, \tag{2.26}$$

Assume, that $\boldsymbol{\varepsilon}_t$ is a conditionally homoskedastic process with zero mean and that the information set $\boldsymbol{F}_t$ is well behaved enough to ensure that $\hat{\boldsymbol{Z}}_t$ is also a stationary and ergodic process. Included in $\boldsymbol{F}_{t-1}$ are ex-ante observable economic variables, e.g. dividend yield, interest rate spreads, or other leading economic indicators [2]. As in previous section, assume that an investor is risk-averse and wishes to diversify portfolio allocation in order to achieve some minimal return. Then, the set of optimal portfolio weights $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)^T$ that maximizes the coefficient of determination of an investment portfolio is the solution of optimization problem

$$\begin{aligned} \underset{w}{\text{maximize}} \quad & R^2 = \frac{Var(\hat{\boldsymbol{Z}}_t^T \boldsymbol{w})}{Var(\boldsymbol{Z}_t^T \boldsymbol{w})} \\ \text{subject to} \quad & \sum_{i=1}^{n} w_i = 1 \\ & \sum_{i=1}^{n} w_i \mu_i \geq \rho \\ & 0 \leq w_i \leq \alpha_i, \quad i = 1, \ldots, n. \end{aligned} \tag{2.27}$$

2.6 Mean absolute deviation

Mean absolute deviation (MAD) of a random variable R is defined as

$$\text{MAD}(R) = E\Big[\Big|R - E[R]\Big|\Big]. \tag{2.28}$$

For a normally distributed random variable, MAD is proportional to standard deviation, as following theorem shows.

Theorem 2.1. *If $R \sim N(\mu, \sigma)$, then*

$$MAD(R) = \sqrt{\frac{2}{\pi}}\sigma. \tag{2.29}$$

Proof. By definition of the expected value of a random variable,

$$
\begin{aligned}
E\Big[\big|R - E[R]\big|\Big] &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} |r-\mu| \exp\Bigg\{-\frac{(r-\mu)^2}{2\sigma}\Bigg\} dr \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{0}^{\infty} (r-\mu) \exp\bigg\{-\frac{(r-\mu)^2}{2\sigma}\bigg\} dr \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{0}^{\infty} t \exp\bigg\{-\frac{t^2}{2\sigma}\bigg\} dt \\
&= \sqrt{\frac{2}{\pi}}\sigma.
\end{aligned} \tag{2.30}
$$

□

This result holds when R is multivariate normal as well [4].

2.7 Optimization problem

First, note that since by definition of standard deviation of a random variable

$$\sigma(R) = \sqrt{Var(R)},$$

maximization of the quotient between variances of two random variables is equivalent to maximization of the quotient of their respective standard deviations, i.e. ,

$$\text{maximize } \frac{Var(\hat{R})}{Var(R)} \quad \text{if and only if} \quad \text{maximize } \frac{\sigma(\hat{R})}{\sigma(R)}$$

The result of Theorem 2.1 implies further, that when $\hat{R}, R$ are normally distributed,

$$\text{maximize } \frac{Var(\hat{R})}{Var(R)} \quad \text{if and only if} \quad \text{maximize } \frac{\text{MAD}(\hat{R})}{\text{MAD}(R)}$$

Finally, maximization of a quotient is equivalent to minimization of its inverse, i.e. ,

$$\text{maximize } \frac{MAD(\hat{R})}{MAD(R)} \quad \text{if and only if} \quad \text{minimize } \frac{\text{MAD}(R)}{\text{MAD}(\hat{R})}$$

Then, model (2.27) can be reformulated as

$$\begin{aligned}
\underset{w}{\text{minimize}} \quad & \frac{E\left[\left|\boldsymbol{R}_t^T\boldsymbol{w} - E[\boldsymbol{R}_t^T\boldsymbol{w}]\right|\right]}{E\left[\left|\hat{\boldsymbol{R}}_t^T\boldsymbol{w} - E[\hat{\boldsymbol{R}}_t^T\boldsymbol{w}]\right|\right]} \\
\text{subject to} \quad & \sum_{i=1}^{n} w_i = 1 \\
& \sum_{i=1}^{n} w_i\mu_i \geq \rho \\
& 0 \leq w_i \leq \alpha_i, \quad i = 1, \ldots, n.
\end{aligned} \tag{2.31}$$

Note that the change of variable from Z_t in (2.27) to R_t in (2.37) was made for notational convenience and consistency with the rest of the chapter.

Model (2.31) can be simplified further. Since $\boldsymbol{R}_t^T\boldsymbol{w} = \sum_{i=1}^{n} w_i R_{i,t}$, it is instructive to look at the individual $w_i R_{i,t}$ terms. Let F_1, F_2 and F_3 represent the three factors in model (2.23) and note, that setting $\beta_0 = \alpha + R_{rf}$ results in a regression equation similar to (2.15):

$$\begin{aligned}
w_i R_{i,t} &= w_i\beta_0 + w_i\beta_{i,1}F_{1,t-1} + w_i\beta_{i,2}F_{2,t-1} + w_i\beta_{i,3}F_{3,t-1} + w_i\varepsilon_{i,t} \\
E[w_i R_{i,t}] &= w_i\beta_0 + w_i\beta_{i,1}E[F_{1,t-1}] + w_i\beta_{i,2}E[F_{2,t-1}] + w_i\beta_{i,3}E[F_{3,t-1}],
\end{aligned} \tag{2.32}$$

which yields

$$\begin{aligned}
w_i R_{i,t} - E[w_i R_{i,t}] &= w_i\beta_{i,1}(F_{1,t-1} - \bar{F}_1) + w_i\beta_{i,2}(F_{2,t-1} - \bar{F}_2) \\
&\quad + w_i\beta_{i,3}(F_{3,t-1} - \bar{F}_3) + w_i\varepsilon_{i,t} \\
&= w_i \sum_{k=1}^{3} \beta_{i,k}(F_{k,t-1} - \bar{F}_k) + w_i\varepsilon_{i,t} \\
\bar{F}_k &= E[F_{k,t-1}].
\end{aligned} \tag{2.33}$$

Thus,

$$\begin{aligned}
\boldsymbol{R}_t^T\boldsymbol{w} - E[\boldsymbol{R}_t^T\boldsymbol{w}] &= \sum_{i=1}^{n} w_i \sum_{k=1}^{3} \beta_{i,k}(F_{k,t-1} - \bar{F}_k) + \sum_{i=1}^{n} w_i\varepsilon_{i,t} \\
&= \sum_{k=1}^{3} \beta_k(\boldsymbol{w})(F_{k,t-1} - \bar{F}_k) + \varepsilon_t(\boldsymbol{w})
\end{aligned} \tag{2.34}$$

where $\beta_k(\boldsymbol{w}) = \sum_{i=1}^{n} w_i\beta_{i,k}$ and $\varepsilon_t(\boldsymbol{w}) = \sum_{i=1}^{n} w_i\varepsilon_{i,t}$.
Because of relation (2.26),

$$\begin{aligned}
\hat{\boldsymbol{R}}_t^T\boldsymbol{w} - E[\hat{\boldsymbol{R}}_t^T\boldsymbol{w}] &= \sum_{k=1}^{3} \beta_k(\boldsymbol{w})(F_{k,t-1} - \bar{F}_k) \\
\beta_k(\boldsymbol{w}) &= \sum_{i=1}^{n} w_i\beta_{i,k}
\end{aligned} \tag{2.35}$$

Assuming equal probability of outcomes $p_t = \frac{1}{T}$, the objective function in (2.31) can be expressed as

$$\frac{E\Big[\Big|\boldsymbol{R}_t^T\boldsymbol{w} - E[\boldsymbol{R}_t^T\boldsymbol{w}]\Big|\Big]}{E\Big[\Big|\hat{\boldsymbol{R}}_t^T\boldsymbol{w} - E[\hat{\boldsymbol{R}}_t^T\boldsymbol{w}]\Big|\Big]} = \frac{\sum_{t=1}^{T}\Big|\sum_{k=1}^{3}\beta_k(\boldsymbol{w})(F_{k,t-1} - \bar{F}_k) + \varepsilon_t(\boldsymbol{w})\Big|}{\sum_{t=1}^{T}\Big|\sum_{k=1}^{3}\beta_k(\boldsymbol{w})(F_{k,t-1} - \bar{F}_k)\Big|} \tag{2.36}$$

The reformulated optimization problem is as follows:

$$\begin{aligned}
\underset{w}{\text{minimize}} \quad & \frac{\sum_{t=1}^{T}\Big|\sum_{k=1}^{3}\beta_k(\boldsymbol{w})(F_{k,t-1} - \bar{F}_k) + \varepsilon_t(\boldsymbol{w})\Big|}{\sum_{t=1}^{T}\Big|\sum_{k=1}^{3}\beta_k(\boldsymbol{w})(F_{k,t-1} - \bar{F}_k)\Big|} \\
\text{subject to} \quad & \sum_{i=1}^{n} w_i = 1 \\
& \sum_{i=1}^{n} w_i\mu_i \geq \rho \\
& 0 \leq w_i \leq \alpha_i, \quad i = 1, \dots, n \\
& \beta_k(\boldsymbol{w}) = \sum_{i=1}^{n} w_i\beta_{i,k} \\
& \varepsilon_t(\boldsymbol{w}) = \sum_{i=1}^{n} w_i\varepsilon_{i,t}
\end{aligned} \tag{2.37}$$

It has been shown, that the optimal solution to 2.37 is equivalent to the optimal solution of following optimization problem [5]:

$$\begin{aligned}
\text{minimize} \quad & \sum_{t=1}^{T}(u_t + v_t) \\
\text{subject to} \quad & \sum_{t=1}^{T}(\xi_t + \eta_t) = 1 \\
& u_t - v_t = \sum_{k=1}^{3}\beta_k(\boldsymbol{y})(F_{k,t-1} - \bar{F}_k) + \varepsilon_t(\boldsymbol{y}) \\
& \xi_t - \eta_t = \sum_{k=1}^{3}\beta_k(\boldsymbol{y})(F_{k,t-1} - \bar{F}_k) \\
& 0 \leq \xi_t \leq a_t z_t \\
& 0 \leq \eta_t \leq b_t(1 - z_t) \\
& u_t \geq 0,\ v_t \geq 0 \\
& z_t \in \{0,1\} \\
& (\boldsymbol{y}, y_0) \in Y \\
& t \in [1, T]
\end{aligned} \tag{2.38}$$

Set Y is the set that satisfies all of the relations below:

$$
\begin{aligned}
y_0 &= \frac{1}{\sum_{t=1}^{T} p_t \left|\sum_{k=1}^{3} \beta_k(\boldsymbol{w})(F_{k,t-1} - \bar{F}_k)\right|} \\
\boldsymbol{y} &= y_0 \boldsymbol{w} \\
\beta_k(\boldsymbol{y}) &= \sum_{i=1}^{n} y_i \beta_{i,k}, \; k = 1, \ldots, K \\
\varepsilon_t(\boldsymbol{y}) &= \sum_{i=1}^{n} y_i \varepsilon_{i,t}, \; t = 1, \ldots, T \\
& \sum_{i=1}^{n} \tilde{r}_i y_i \geq \rho y_0 \\
& \sum_{i=1}^{n} y_i = y_0 \\
& 0 \leq y_i \leq \alpha y_0, \; i = 1, \ldots, n \\
& y_0 \geq 0
\end{aligned}
\tag{2.39}
$$

a_t and b_t are given by

$$
\begin{aligned}
a_t &= \max\left\{ \max\left\{ \sum_{k=1}^{3} \beta_k(\boldsymbol{y})(F_{k,t-1} - \bar{F}_k) \;\middle|\; (\boldsymbol{y}, y_0) \in Y \right\}, 0 \right\} \\
b_t &= -\min\left\{ \min\left\{ \sum_{k=1}^{3} \beta_k(\boldsymbol{y})(F_{k,t-1} - \bar{F}_k) \;\middle|\; (\boldsymbol{y}, y_0) \in Y \right\}, 0 \right\}
\end{aligned}
\tag{2.40}
$$

Solving problem 2.38 yields among other things values $(\boldsymbol{y}^*, y_0^*)$, from which we get optimal portfolio weights $\boldsymbol{w}^* = \frac{\boldsymbol{y}^*}{y_0^*}$.

Chapter 3

Methodology

All parts of the system were built in Python [6]. Following packages (in latest version) were used:

- Selenium - web scraping
- Beautiful soup - parsing and cleaning scraped data
- Pandas datareader - API requests
- Pandas, Numpy, Scikit - essential data analysis packages
- Gurobi - linear and non-linear optimization
- Plotly - plotting graphs

The strategy is to find optimal portfolios based on the data from year 2017, buy these portfolios in the beginning of 2018 and assess their performance at the end of 2018.

3.1 Data acquisition

A list of all currently traded Small, Mid and Large Cap companies was scraped from OMX NASDAQ webpage [7]. Historical monthly adjusted closing price data for the period January 1, 2012 - December 31, 2018 was downloaded and stored for each asset in the above list from Alpha Vantage Free API, which required a private *API key* [8]. Stocks with insufficient histrical data were discarded from the analysis, which brought the total number of assets from 355 down to 255. The Fama-French European market monthly factor data for the period in question was obtained through Kenneth French's website [9].

3.2 Data analysis

Demeaned monthly returns were calculated from adjusted closing price data according to (2.2). The result was then split into a training set, consisting of 60 observations accounting for the first five years of observations (2012-2016), and a test set, consisting of 12 observations accounting for the last year of observations (2017). Regression model (2.23) was fitted on the training set and then used to predict returns. Predictions were then compared to the test set. Regression coefficients , residuals, MSE- and R^2-scores as well as predicted returns were saved and stored locally. The outlined procedure was performed for each stock.

Next, regression coefficients, residuals, factor data and return predictions were used as input to the optimizer specified in accordance with (2.38), with $T = 12$ and $n = 255$. This yielded portfolio weights as output. In total, nine portfolios were constructed using different values of *maximal allocation parameter a* and *minimal annual rate of return ρ* (Table 3.1). Essentially, a regulates the size of the portfolio in terms of total numbers of assets, while ρ represents the "greedyness" of the investor.

Table 3.1: List of portfolios based on choice of parameters a and ρ.

Portfolio	a (%)	ρ (%)
MPP 1	20	5
MPP 2	20	10
MPP 3	20	20
MPP 4	10	5
MPP 5	10	10
MPP 6	10	20
MPP 7	5	5
MPP 8	5	10
MPP 9	5	20

3.3 Portfolio assessment

As benchmark, the *Six Portfolio Return Index* was chosen. SIXPRX reflects the market progress of companies listed on the Stockholm Stock Exchange, subject to the restriction that no company may weigh over ten per cent. Returns received by the shareholders in the form of dividends are reinvested in SIXPRX [10].

Portfolio metrics were calculated using formulas in Section 2.1:

- Return - Equation (2.4)
- Performance - Equation (2.3)

- Volatility - Equations (2.7), (2.8)
- Beta - Equation (2.9)
- Sharpe ratio - Equation (2.10)
- Tracking error - Equations (2.12),(2.13)
- Information ratio - Equation (2.14)

Since period of interest spanned a single year, using regular formulas was equivalent to using the annualized counterparts. Metrics were calculated for each portfolio in Table 3.1.

Chapter 4

Results

MPP portfolios consisted of various combinations of 70 different stocks (Figure A.1). Portfolios were comprised of no more than 30 assets. Stocks were analyzed in terms of coefficient of determination and annual return in 2017. Table 4.1 shows stocks with best/worst predictability in the market (4.1a, 4.1c) versus those selected by optimization algorithm (4.1b, 4.1d) during that period. Table 4.2 presents annual return of stocks in analogous fashion.

Table 4.1: Best and worst stocks in terms of predictability (R^2) in the market (Tables (a), (c)) and in portfolio selection (Tables (b), (d)). Rank column indicates relative position in descending order. Analysis based on financial data from 2017.

Asset	P	rank	R^2	rank
INVE-B.ST	12.904	105	0.533	1
INVE-A.ST	12.150	110	0.504	2
BELE.ST	-15.786	200	0.451	3
INDU-C.ST	22.250	74	0.446	4
SEB-A.ST	6.356	127	0.426	5
LIAB.ST	-5.168	167	0.412	6
DURC-B.ST	54.194	20	0.380	7
LUND-B.ST	11.868	112	0.377	8
STE-R.ST	32.657	49	0.371	9
ENEA.ST	-15.365	198	0.371	9

(a) Ten stocks with highest predictability in the market.

Asset	P	rank	R^2	rank
LIAB.ST	-5.168	167	0.412	6
BMAX.ST	-9.197	180	0.307	32
TELIA.ST	5.065	131	0.298	38
XANO-B.ST	46.986	26	0.282	46
CAST.ST	15.432	96	0.281	47
QLRO.ST	109.249	5	0.270	52
BEGR.ST	19.904	84	0.261	56
BIOT.ST	87.514	9	0.255	59
EOLU-B.ST	27.175	63	0.252	61
REJL-B.ST	-18.347	209	0.250	63

(b) Ten stocks with highest predictability in MPP portfolios.

Asset	P	rank	R^2	rank
HTRO.ST	108.108	6	0.000	255
TRENT.ST	-1.923	158	0.005	254
PROF-B.ST	73.904	12	0.007	252
HPOL-B.ST	0.361	151	0.007	252
RAIL.ST	-18.457	210	0.009	251
SMF.ST	-23.875	218	0.011	250
ORTI-B.ST	-21.119	215	0.014	249
STRAX.ST	2.459	139	0.016	247
CTT.ST	60.468	19	0.016	247
IS.ST	-5.357	169	0.018	246

(c) Ten stocks with lowest predictability in the market.

Asset	P	rank	R^2	rank
HTRO.ST	108.108	6	0.000	255
TRENT.ST	-1.923	158	0.005	254
PROF-B.ST	73.904	12	0.007	252
RAIL.ST	-18.457	210	0.009	251
SMF.ST	-23.875	218	0.011	250
STRAX.ST	2.459	139	0.016	247
CTT.ST	60.468	19	0.016	247
IS.ST	-5.357	169	0.018	246
NAXS.ST	1.721	145	0.022	242
SWEC-A.ST	-16.231	203	0.025	238

(d) Ten stocks with lowest predictability in MPP portfolios.

Table 4.2: Best and worst stocks in terms of performance (cumalative returns) (P) in the market (Tables (a), (c)) and in portfolio selection (Tables (b), (d)). Rank column indicates relative position in descending order. Analysis based on financial data from 2017.

Asset	P	rank	R^2	rank
G5EN.ST	212.836	1	0.049	216
CRAD-B.ST	134.855	2	0.109	163
HMED.ST	114.163	3	0.023	241
NOLA-B.ST	112.587	4	0.081	189
QLRO.ST	109.249	5	0.270	52
HTRO.ST	108.108	6	0.000	255
CORE-A.ST	97.555	7	0.080	191
SCA-A.ST	94.344	8	0.206	94
BIOT.ST	87.514	9	0.255	59
SCA-B.ST	79.812	10	0.210	89

(a) Ten stocks with highest performance in the market.

Asset	P	rank	R^2	rank
G5EN.ST	212.836	1	0.049	216
CRAD-B.ST	134.855	2	0.109	163
NOLA-B.ST	112.587	4	0.081	189
QLRO.ST	109.249	5	0.270	52
HTRO.ST	108.108	6	0.000	255
CORE-A.ST	97.555	7	0.080	191
SCA-A.ST	94.344	8	0.206	94
BIOT.ST	87.514	9	0.255	59
PROF-B.ST	73.904	12	0.007	252
ORTI-A.ST	70.810	14	0.040	223

(b) Ten stocks with highest performance in MPP portfolios.

Asset	P	rank	R^2	rank
ACTI.ST	-83.349	255	0.202	95
VSSAB-B.ST	-75.626	254	0.114	159
FING-B.ST	-73.339	253	0.114	159
INVUO.ST	-70.435	252	0.050	215
OASM.ST	-69.987	251	0.120	156
ICTA.ST	-66.240	250	0.045	221
PREC.ST	-58.750	249	0.094	175
STAR-B.ST	-56.477	248	0.029	234
EPIS-B.ST	-52.612	247	0.105	166
MOB.ST	-51.404	246	0.081	189

(c) Ten stocks with lowest performance in the market.

Asset	P	rank	R^2	rank
VSSAB-B.ST	-75.626	254	0.114	159
FING-B.ST	-73.339	253	0.114	159
PREC.ST	-58.750	249	0.094	175
VRG-B.ST	-33.333	235	0.090	182
SMF.ST	-23.875	218	0.011	250
RAIL.ST	-18.457	210	0.009	251
REJL-B.ST	-18.347	209	0.250	63
SWEC-A.ST	-16.231	203	0.025	238
BILI-A.ST	-16.199	202	0.176	116
TETY.ST	-15.319	196	0.157	131

(d) Ten stocks with lowest performance in MPP portfolios.

Table 4.3 shows, that eight out of nine portfolios had considerable coefficient of determination. MPP9 had the highest R^2-score and MPP2 had lowest, 0.5798 and 0.2683 respectively. Portfolios with high predictability had outperformed market index in 2017. MPP1 had highest annual return and MPP2 had lowest, 0.3706 and 0.0290 respectively. In terms of risk, MPP1 performed worst, with $\sigma = 0.0433$ and $TE = 0.0362$. MPP7 had lowest volatility, 0.02, and MPP9 had lowest Tracking error, 0.0306. MPP9 had highest risk-adjusted returns, $SR = 1.0144$ and $IR = 0.5565$. Worst in this regard was MPP2, with $\sigma = 0.0615$ and $IR = -0.1805$. Overall, MPP portfolios showed low market correlation. In this category, MPP2 did best with $\beta = -0.0081$ and MPP6 worst with $\beta = 0.1003$.

The investment strategy was to buy the optimal portfolios in the beginning of 2018 and assess them at the end of the year in similar fashion as in previous period. With the exception of MPP2, all portfolios underperformed the market (Table 4.4). MPP2 was the only portfolio with positive result in 2018. It yielded 0.0106 in returns, and 0.0623 and 0.2277 in risk-adjusted returns. Worst in this regard was MPP7, with -0.1123, -0.45344 and -0.3378 respectively. MPP8 had lowest risk, $\sigma = 0.0197$ and $TE = 0.0168$. MPP3 had highest risk, $\sigma = 0.0418$ and $TE = 0.0368$. Market correlation increased considerably in all portfolios. Interestingly, MPP3 went from having lowest correlation in 2017 to highest in 2018 with $\beta = 1.2143$. On the other hand, MPP6 reverted from having highest correlation in 2017 to lowest in 2018 with $\beta = 0.6287$. Note, that results for MPP6 and MPP9 were incomplete due to missing data. In 2018 UFLX.ST (Uniflex Sverige AB) was delisted from the stock market due to a merger with POOL-B.ST (Poolia AB).

Table 4.3: Comparison between market (SIXPRX) and optimal portfolios in terms of performance (P), risk (σ, TE), risk-adjusted returns (SR, IR) and correlation (β). Based on financial data from 2017.

	R^2	P	σ	β	SR	TE	IR
SIXPRX	-	0.0948	0.0290	1.0000	0.2517	-	-
MPP1	0.4779	0.3706	0.0433	0.0845	0.6053	0.0362	0.5224
MPP2	0.2683	0.0290	0.0287	-0.0081	0.0615	0.0307	-0.1805
MPP3	0.4814	0.2035	0.0349	0.0743	0.4312	0.0328	0.2363
MPP4	0.4814	0.3407	0.0409	0.0756	0.5930	0.0330	0.5135
MPP5	0.4592	0.2973	0.0402	0.0668	0.5340	0.0336	0.4205
MPP6	0.4935	0.2217	0.0356	0.1003	0.4592	0.0322	0.2803
MPP7	0.5413	0.1960	0.0200	0.0698	0.7262	0.0315	0.2284
MPP8	0.5307	0.2096	0.0238	0.0401	0.6508	0.0330	0.2476
MPP9	0.5798	0.3432	0.0240	0.0990	1.0144	0.0306	0.5565

Table 4.4: Comparison between market (SIXPRX) and optimal portfolios in terms of performance (P), risk (σ, TE), risk-adjusted returns (SR, IR) and correlation (β). Based on financial data from 2018. Portfolios marked with (*) had missing data.

	P	σ	β	SR	TE	IR
SIXPRX	-0.0442	0.0352	1.0000	-0.0904	-	-
MPP 1	-0.1102	0.0315	0.8839	-0.2758	0.0354	-0.1555
MPP 2	0.0106	0.0312	1.2143	0.0623	0.0225	0.2277
MPP 3	-0.0735	0.0418	0.8336	-0.1283	0.0368	-0.0591
MPP 4	-0.1041	0.0318	1.0402	-0.2590	0.0233	-0.2170
MPP 5	-0.0895	0.0307	0.8122	-0.2339	0.0231	-0.1733
MPP 6*	-0.0528	0.0316	0.6287	-0.1286	0.0257	-0.0342
MPP 7	-0.1123	0.0202	1.0006	-0.4534	0.0177	-0.3378
MPP 8	-0.0822	0.0197	0.7246	-0.3416	0.0168	-0.2114
MPP 9*	-0.0831	0.0212	0.8421	-0.3108	0.0225	-0.1515

Chapter 5

Discussion

5.1 Fama-French three-factor model

By examination of Tables 4.1a and 4.1c it is evident, that forecasting ability of model 2.23 is unsatisfactory. For instance, only two stocks, INVE-A.ST and INVE-B.ST (Investor AB), had $R^2 > 0.5$, while most of the assets had coefficient of determination below 0.3. Furthermore, a comparison between values in Tables 4.2a and 4.2c suggests, that the model is better at predicting positive returns than negative returns.

One possible explanation to poor performance is, that factor data used in the analysis was based on much broader European market. Since calculation of factors is based on fiscal data that is not readily available cost free for Swedish firms, a choice was made to use available European data. On a broader scale, recall that model (2.23) is an extension of CAPM that was developed to better suit empirical data. In later years, several researchers from the field of behavioral finance gave possible explanations to phenomena that the Fama-French aims to capture. Namely, that small firms outperform big firms (SMB), and that firms with high book-to-market value outperform those with low corresponding value (applied asset). However, ad-hoc nature of this model raises question of its validity. Perhaps, a three-factor model is an insufficient investment tool. Indeed, recently the authors extended their previous work by addition of two factors, accounting for profitability and investment activity of firms (wikipedia). On the other hand, in the development of the concept of maximally predictable portfolio, Lo and MacKinley used own seven-factor model. The above mentioned models were not subject of the analysis, therefore no remark can be made about their predictive power. But, Occam's razor states that, the simplest is most likely. In this view, the Fama-French model is not incomplete, e.g. missing factors, but rather it is based on wrong assumption of market equilibrium. In other words, poor predictions confirm that financial time-series are non-stationary. From the point of view of a risk-averse investor, this raises question of suit-ability of the model.

5.2 Maximally predictable portfolio-optimization

In this project we maximized coefficient of determination R^2 for portfolio returns constrained by a minimal return requirement. It is therefore reasonable to analyze the results in terms of these quantities. However, it is also important to view the results from a perspective of a real world investor.

Judging by data in Table 4.3, it is safe to say, that MPP9 was the best portfolio out of the nine in Table 3.1. It had the best trade-off between return and risk, although correlation with market was among the highest observed. MPP2 was the worst among portfolios; specifically, it deviated strongly in terms of R^2-score and performance from the rest. Overall, MPP optimization did yield portfolios with considerable predictability and excess return. Perhaps the most surprising feature of MPP portfolios was the low beta values across the board. It seems as if maximizing predictability implicitly minimizes market correlation.

Smaller portfolios had more risk than bigger portfolios, that is, the effect of diversification was observed. Indeed, small portfolios ($a = 20$) consisted of 16 assets, medium portfolios ($a = 10$) consisted of 19 assets, but large portfolios ($a = 5$) consisted of 29 assets. Increase in return parameter ρ in smaller portfolios yielded worse annual performance. On the contrary, in big portfolios being "greedy" led to increased cumulative return. Portfolios with $\rho = 10$, did worse in terms of R^2-score but better in terms of beta than the rest within a portfolio group.

Great performance of MPP9 was accounted by the composition of the portfolio. Table A.13 shows that MPP9 managed to capture some of the top performing stocks in 2017. What really stands out however is the R^2-score of individual assets. The data suggests, that MPP optimization yields portfolios with high predictability *regardless of coefficient of determination of individual assets*. Comparison between Tables 4.1a and 4.1b indicates, that the optimizer does not explicitly prioritize stocks with high predictability, which lends more strength to this conclusion.

5.3 Strategy assessment

From an investors point of view, 2017 was a lucrative year. The net market return was positive, but it was lowered considerably by poor results during summer and late autumn/winter. Nevertheless, it is fair to say that the overall market climate was *positive*. Figure 5.1 suggests, that the exceptional performance of optimal portfolios (except MPP2) was accounted by fairly consistent positive result at the end of each month. Figure 5.2 illustrates this point further. The top performing portfolios (MPP1, MPP4, MPP9) outperformed the market from the start. The rest of portfolios did eventually catch up and outperform market much thanks to not being affected by market downturns.

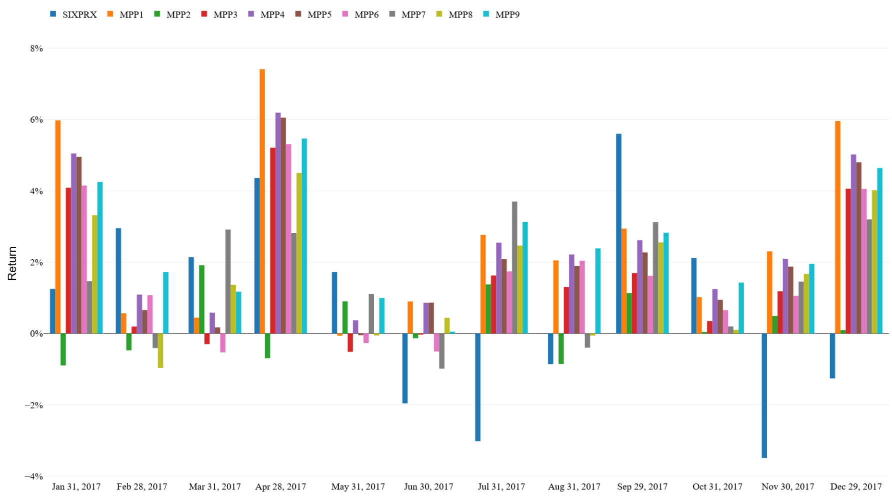

Figure 5.1: Monthly returns of optimal portfolios and index (%) in 2017. Data from Table A.1.

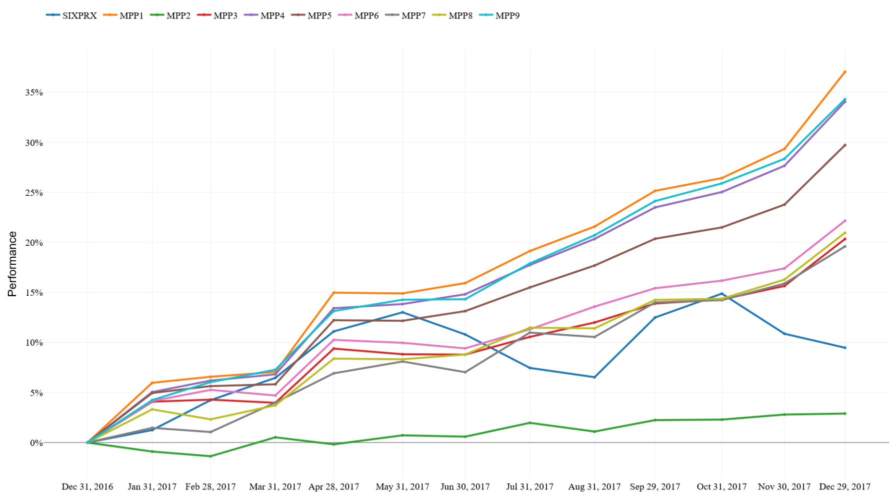

Figure 5.2: Monthly performance of optimal portfolios and index (%) in 2017. Data from Table A.2.

Figure 5.3 indicates, that market sentiment turned *negative* in 2018. Market returns were considerably lower and negative development towards the end of the year lead to a negative annual result. The same can be said about optimal portfolios. Data suggests no clear advantage of MPP portfolios over index, with exception of MPP2 that showed great initial result and ended on plus side. In fact, Figure 5.4 suggests greater correlation with market movements. This confirms findings in Table 4.4, that shows substantial increase in portfolio betas in 2018.

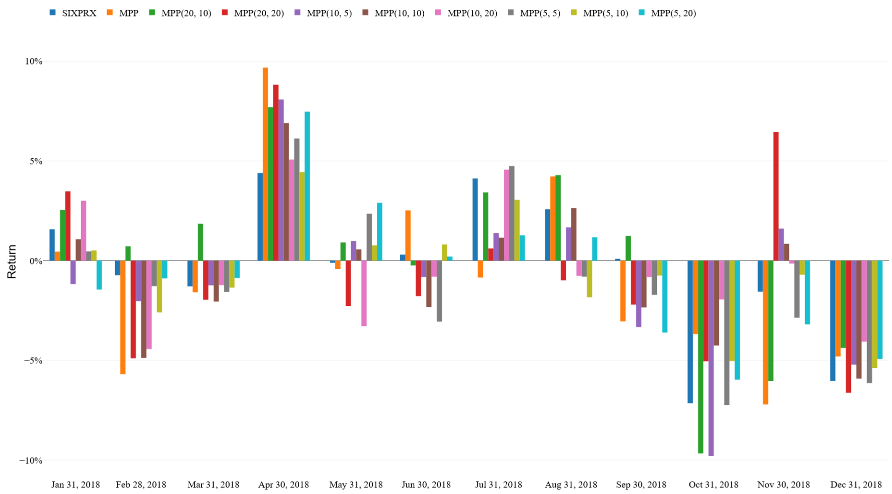

Figure 5.3: Monthly returns of optimal portfolios and index (%) in 2018. Data from Table A.3.

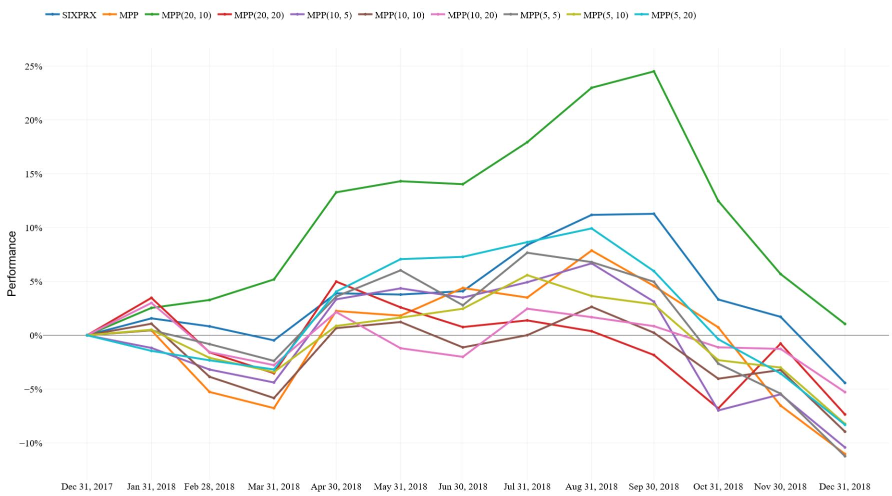

Figure 5.4: Monthly performance of optimal portfolios and index (%) in 2018. Data from Table A.4.

This last point provides further significance to beta-neutrality of MPP portfolios being a key feature.

The inversion of performance was observed even for other MPP portfolios. What caused this effect? As Table A.3 indicates, market development in 2018 was very weak. At the same time portfolio betas increased drastically from 2017 to 2018 (Tables 4.3, 4.4), in other words portfolios became more sensitive to market fluctuations. A possible explanation to poor performance of large portfolios, e.g. MPP9, is that their decline was caused by being exposed at a larger extent to a declining market, despite diversification. The small portfolios, e.g. MPP3, had a limited exposure and were able to avoid the broad market decline.

5.4 Suggestions for further research

Forecasting stock returns is a central part of the method outlined in this paper. A major weakness of the Fama-French model is its assumption of stationarity of financial time series. A more suitable alternative is *autoregressive integrated moving average* (ARIMA) model. It can be used to obtain a stationary time series through appropriate amount of differencing. Presence of seasonal effects in stock market [3] further motivate employment of a seasonal ARIMA model. An added bonus of using these type of model is that it does not require additional data, e.g. factor data, in the analysis, which is can be quite difficult to obtain for a non-institutional investor.

In the original paper Lo and MacKinley do not explicitly assume that the return process follows normal distribution [2]. For instance, Student's t-distribution is superior to normal distribution in terms of modelling asset returns [11]. However, derivation of the problem (2.38) is made under assumption of normality (recall substitution of variance with mean absolute deviation). This was made in order to simplify convex-convex quadratic fractional problem (2.27) to a linear problem. Luckily, the general case has been studied and a algorithmic strategy can be developed on the basis of available research [12].

The outcomes of the project were aligned with results of previous research [2, 5, 12]. Namely, that markets contain substantial predictability that can be used to compose profitable portfolios that beat the market. What is new is that maximum predictability of a portfolio can be understood in terms of market-neutrality. Indeed, the data clearly suggest that low beta is a distinctive feature of MPP portfolios. However, the limited time frame of the project demands further analysis before such claim can be asserted. A suggestion is thus to develop and test this model for several time periods with different duration, preferably in different market conditions.

Lastly, the performance of Gurobi solver was outstanding, the nine optimization problems were solved in under 10s. This suggests that the solver can handle large scale problems in reasonable time. Thus a comparison between using monthly versus using daily data would be interesting in determining the optimal setup for the optimization problem.

Chapter 6

Conclusion

The goal of this project was to develop an algorithmic strategy based on MPP framework. We were curious to see if the success of past research could be replicated in present day Swedish market. In part it did. Solving the optimization problem using historical data from recent years yielded portfolios with exceptional returns and risk profiles. Indeed, these portfolios had relatively high coefficient of determination despite the fact that the assets they were comprised of did not. Another interesting feature that was not discussed in the academic papers is the low beta that was characteristic of MPP portfolios. Thus, the first thesis question was answered affirmatively.

Unfortunately, the promising results did not last particularly long. The Swedish stock market was "bearish" in 2018, most likely due to concerns regarding to the mortgage market and post-election turbulence in later part of the year. In this setting almost all of the benefits of MPP portfolios vanished. With regards to our findings, we concluded that the strategy performed unsatisfactory.

Bibliography

[1] Petter N. Kolm, Reha Tütüncü, and Frank J. Fabozzi. 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2):356 – 371, 2014. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2013.10.060. URL http://www.sciencedirect.com/science/article/pii/S0377221713008898. 60 years following Harry Markowitz's contribution to portfolio theory and operations research.

[2] Andrew W Lo and A. Craig MacKinlay. Maximizing predictability in the stock and bond markets. Working Paper 5027, National Bureau of Economic Research, February 1995. URL http://www.nber.org/papers/w5027.

[3] Marcus Schulmerich, Yves-Michel Leporcher, and Ching-Hwa Eu. *Applied Asset and Risk Management: A Guide to Modern Portfolio Management and Behavior-Driven Markets*. Springer Publishing Company, Incorporated, 1st edition, 2016. ISBN 3662525755, 9783662525753.

[4] Hiroshi Konno and Tomoyuki Koshizuka. Mean-absolute deviation model. *IIE Transactions*, 37(10):893–900, 2005. doi: 10.1080/07408170591007786. URL https://www.tandfonline.com/doi/abs/10.1080/07408170591007786.

[5] Constantin Zopounidis, Michael Doumpos, and Panos M. Pardalos. *Handbook of Financial Engineering*. Springer Publishing Company, Incorporated, 2011. ISBN 1441945733, 9781441945730.

[6] G. van Rossum. Python tutorial. Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, May 1995.

[7] OMX Nasdaq. Historical data. URL http://www.nasdaqomxnordic.com/aktier/historiskakurser.

[8] Alpha Vantage. Documentation. URL https://www.alphavantage.co/documentation/.

[9] Kenneth French. Data library. URL `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[10] Fondbolagens förening. Six portfolio return index. URL `http://www.fondbolagen.se/sv/Statistik--index/Index/Marknadsindex/SIXPRX---aktuell-graf/`.

[11] *Risk and portfolio analysis principles and methods*. Springer series in operations research. Springer Science+Business Media, New York, 2012. ISBN 1-4614-4103-X.

[12] R. Yamamoto and H. Konno. An efficient algorithm for solving convex–convex quadratic fractional programs. *Journal of Optimization Theory and Applications*, 133(2):241–255, 2007. ISSN 0022-3239.

Appendix A

Tables and figures

Table A.1: Monthly returns of optimal portfolios and index (%). Based on financial data from 2017.

Date	SIXPRX	MPP1	MPP2	MPP3	MPP4	MPP5	MPP6	MPP7	MPP8	MPP9
Jan	1.250	5.975	-0.898	4.086	5.048	4.954	4.149	1.468	3.318	4.249
Feb	2.950	0.568	-0.472	0.196	1.091	0.656	1.074	-0.411	-0.967	1.718
Mar	2.140	0.444	1.915	-0.303	0.584	0.173	-0.532	2.914	1.368	1.172
Apr	4.360	7.411	-0.695	5.211	6.192	6.050	5.305	2.811	4.500	5.467
May	1.720	-0.064	0.902	-0.516	0.367	-0.051	-0.266	1.106	-0.057	0.995
Jun	-1.960	0.900	-0.134	-0.031	0.860	0.863	-0.505	-0.987	0.438	0.051
Jul	-3.020	2.765	1.374	1.627	2.546	2.093	1.741	3.698	2.464	3.130
Aug	-0.860	2.048	-0.856	1.301	2.215	1.894	2.040	-0.396	-0.059	2.385
Sep	5.600	2.939	1.135	1.696	2.616	2.274	1.615	3.123	2.552	2.827
Oct	2.120	1.019	0.047	0.349	1.246	0.943	0.655	0.199	0.102	1.428
Nov	-3.490	2.304	0.493	1.185	2.097	1.874	1.058	1.454	1.671	1.951
Dec	-1.260	5.956	0.096	4.058	5.019	4.801	4.053	3.198	4.020	4.638

Table A.2: Monthly performance of optimal portfolios and index (%). Calculated from Table (A.1) and (2.3).

Date	SIXPRX	MPP1	MPP2	MPP3	MPP4	MPP5	MPP6	MPP7	MPP8	MPP9
Jan	1.250	5.975	-0.898	4.086	5.048	4.954	4.149	1.468	3.318	4.249
Feb	4.237	6.577	-1.366	4.290	6.194	5.643	5.267	1.051	2.319	6.040
Mar	6.468	7.050	0.523	3.974	6.814	5.825	4.708	3.996	3.719	7.283
Apr	11.110	14.983	-0.176	9.392	13.428	12.228	10.263	6.919	8.387	13.147
May	13.021	14.909	0.724	8.828	13.844	12.171	9.970	8.101	8.326	14.273
Jun	10.805	15.943	0.589	8.793	14.823	13.139	9.415	7.034	8.800	14.331
Jul	7.459	19.149	1.972	10.563	17.746	15.507	11.319	10.993	11.481	17.909
Aug	6.535	21.589	1.099	12.002	20.354	17.694	13.590	10.553	11.415	20.722
Sep	12.501	25.162	2.247	13.902	23.502	20.370	15.425	14.006	14.259	24.135
Oct	14.886	26.438	2.295	14.299	25.041	21.506	16.181	14.233	14.376	25.908
Nov	10.876	29.351	2.800	15.654	27.663	23.783	17.410	15.894	16.287	28.364
Dec	9.479	37.056	2.898	20.348	34.070	29.725	22.169	19.600	20.962	34.317

Table A.3: Monthly returns of optimal portfolios and index (%). Based on financial data from 2018. Portfolios marked with (*) had missing data.

Date	SIXPRX	MPP1	MPP2	MPP3	MPP4	MPP5	MPP6*	MPP7	MPP8	MPP9*
Jan	1.570	0.451	2.540	3.474	-1.177	1.072	3.001	0.463	0.511	-1.451
Feb	-0.730	-5.691	0.720	-4.896	-2.033	-4.876	-4.433	-1.279	-2.595	-0.889
Mar	-1.290	-1.585	1.847	-1.964	-1.239	-2.053	-1.230	-1.570	-1.352	-0.871
Apr	4.390	9.676	7.691	8.818	8.078	6.894	5.066	6.121	4.438	7.466
May	-0.110	-0.421	0.911	-2.279	0.983	0.564	-3.286	2.350	0.766	2.898
Jun	0.300	2.519	-0.241	-1.780	-0.822	-2.328	-0.807	-3.059	0.811	0.203
Jul	4.120	-0.848	3.419	0.610	1.381	1.148	4.559	4.742	3.046	1.270
Aug	2.580	4.218	4.287	-0.987	1.667	2.634	-0.764	-0.804	-1.835	1.172
Sep	0.090	-3.047	1.236	-2.204	-3.331	-2.346	-0.822	-1.715	-0.748	-3.607
Oct	-7.150	-3.683	-9.665	-5.046	-9.795	-4.257	-1.949	-7.242	-5.035	-5.969
Nov	-1.560	-7.209	-6.033	6.451	1.605	0.843	-0.147	-2.862	-0.706	-3.195
Dec	-6.030	-4.807	-4.381	-6.626	-5.215	-5.916	-4.062	-6.139	-5.384	-4.933

Table A.4: Monthly performance of optimal portfolios and index (%). Calculated from Table (A.3) and (2.3). Portfolios marked with (*) had missing data.

Date	SIXPRX	MPP1	MPP2	MPP3	MPP4	MPP5	MPP6*	MPP7	MPP8	MPP9*
Jan	1.570	0.451	2.540	3.474	-1.177	1.072	3.001	0.463	0.511	-1.451
Feb	0.829	-5.266	3.278	-1.592	-3.185	-3.857	-1.564	-0.822	-2.097	-2.326
Mar	-0.472	-6.767	5.185	-3.525	-4.385	-5.830	-2.775	-2.379	-3.420	-3.177
Apr	3.897	2.254	13.275	4.982	3.339	0.662	2.150	3.596	0.866	4.052
May	3.783	1.823	14.306	2.590	4.355	1.230	-1.206	6.031	1.638	7.067
Jun	4.094	4.388	14.031	0.764	3.497	-1.127	-2.003	2.787	2.462	7.285
Jul	8.383	3.503	17.929	1.379	4.926	0.008	2.464	7.662	5.583	8.648
Aug	11.179	7.869	22.985	0.378	6.675	2.642	1.681	6.796	3.645	9.921
Sep	11.279	4.583	24.505	-1.834	3.122	0.234	0.845	4.965	2.870	5.957
Oct	3.323	0.731	12.472	-6.787	-6.979	-4.033	-1.121	-2.637	-2.310	-0.367
Nov	1.711	-6.531	5.686	-0.774	-5.486	-3.225	-1.266	-5.423	-3.000	-3.551
Dec	-4.422	-11.024	1.056	-7.349	-10.415	-8.950	-5.277	-11.229	-8.223	-8.309

Table A.5: Asset allocation of portfolio MPP1.

Asset	Weight (%)	Performance (%)	rank	R2	rank
VICP-A.ST	7.722	33.927	44	0.040	223
SMF.ST	15.662	-23.875	218	0.011	250
SENS.ST	7.630	12.682	107	0.028	235
SCA-A.ST	10.389	94.344	8	0.206	94
QLRO.ST	2.673	109.249	5	0.270	52
PROF-B.ST	11.198	73.904	12	0.007	252
PREC.ST	5.300	-58.750	249	0.094	175
ORTI-A.ST	0.758	70.810	14	0.040	223
MIDW-A.ST	3.160	-5.181	168	0.089	184
MCAP.ST	0.093	29.683	56	0.088	186
ENDO.ST	9.552	-9.836	182	0.063	205
EAST.ST	2.500	22.472	73	0.154	133
BIOT.ST	13.889	87.514	9	0.255	59
BEGR.ST	4.445	19.904	84	0.261	56
ANOT.ST	5.028	-3.333	160	0.046	219

Table A.6: Asset allocation of portfolio MPP2.

Asset	Weight (%)	Performance (%)	rank	R2	rank
VSSAB-B.ST	0.253	-75.626	254	0.114	159
VRG-B.ST	8.398	-33.333	235	0.090	182
TRENT.ST	12.903	-1.923	158	0.005	254
SWEC-A.ST	10.119	-16.231	203	0.025	238
SMF.ST	8.999	-23.875	218	0.011	250
SENS.ST	4.760	12.682	107	0.028	235
RROS.ST	0.166	-7.911	179	0.051	213
RAIL.ST	0.794	-18.457	210	0.009	251
ORTI-A.ST	1.643	70.810	14	0.040	223
NAXS.ST	13.532	1.721	145	0.022	242
LUPE.ST	20.000	-0.997	154	0.185	109
BONG.ST	9.095	10.465	117	0.087	187
AXFO.ST	2.923	15.337	97	0.197	100
ANOT.ST	2.925	-3.333	160	0.046	219
AGRO.ST	3.492	21.250	78	0.032	230

Table A.7: Asset allocation of portfolio MPP3.

Asset	Weight (%)	Performance (%)	rank	R2	rank
XANO-B.ST	2.625	46.986	26	0.282	46
VICP-A.ST	3.061	33.927	44	0.040	223
TRENT.ST	1.748	-1.923	158	0.005	254
STRAX.ST	20.000	2.459	139	0.016	247
SMF.ST	7.941	-23.875	218	0.011	250
SENS.ST	5.783	12.682	107	0.028	235
PROF-B.ST	6.126	73.904	12	0.007	252
PREC.ST	6.678	-58.750	249	0.094	175
ORTI-A.ST	0.461	70.810	14	0.040	223
MIDW-A.ST	1.900	-5.181	168	0.089	184
FPAR.ST	6.051	40.966	33	0.220	83
ENDO.ST	9.498	-9.836	182	0.063	205
EAST.ST	6.799	22.472	73	0.154	133
BONG.ST	8.068	10.465	117	0.087	187
BIOT.ST	11.043	87.514	9	0.255	59
ANOT.ST	2.220	-3.333	160	0.046	219

Table A.8: Asset allocation of portfolio MPP4.

Asset	Weight (%)	Performance (%)	rank	R2	rank
XANO-B.ST	0.564	46.986	26	0.282	46
VICP-A.ST	5.724	33.927	44	0.040	223
STRAX.ST	10.000	2.459	139	0.016	247
SMF.ST	5.557	-23.875	218	0.011	250
SENS.ST	9.149	12.682	107	0.028	235
SCA-A.ST	10.000	94.344	8	0.206	94
QLRO.ST	0.611	109.249	5	0.270	52
PREC.ST	6.169	-58.750	249	0.094	175
ORTI-A.ST	1.513	70.810	14	0.040	223
NOLA-B.ST	1.204	112.587	4	0.081	189
NAXS.ST	10.000	1.721	145	0.022	242
KDEV.ST	3.904	-4.167	164	0.142	142
IS.ST	4.853	-5.357	169	0.018	246
EOLU-B.ST	2.272	27.175	63	0.252	61
ENDO.ST	2.003	-9.836	182	0.063	205
CTT.ST	9.154	60.468	19	0.016	247
CORE-A.ST	10.000	97.555	7	0.080	191
ANOT.ST	7.177	-3.333	160	0.046	219
AGRO.ST	0.147	21.250	78	0.032	230

Table A.9: Asset allocation of portfolio MPP5.

Asset	Weight (%)	Performance (%)	rank	R2	rank
VICP-A.ST	1.622	33.927	44	0.040	223
SWMA.ST	2.051	14.847	100	0.242	69
STRAX.ST	10.000	2.459	139	0.016	247
SMF.ST	3.188	-23.875	218	0.011	250
SENS.ST	8.484	12.682	107	0.028	235
SCA-A.ST	4.943	94.344	8	0.206	94
PROF-B.ST	8.447	73.904	12	0.007	252
PREC.ST	10.000	-58.750	249	0.094	175
ORTI-A.ST	0.751	70.810	14	0.040	223
NAXS.ST	10.000	1.721	145	0.022	242
MIDW-A.ST	0.860	-5.181	168	0.089	184
HTRO.ST	1.357	108.108	6	0.000	255
ENDO.ST	9.035	-9.836	182	0.063	205
EAST.ST	7.734	22.472	73	0.154	133
CRAD-B.ST	1.767	134.855	2	0.109	163
CORE-A.ST	6.876	97.555	7	0.080	191
BIOT.ST	7.875	87.514	9	0.255	59
ANOT.ST	5.009	-3.333	160	0.046	219

Table A.10: Asset allocation of portfolio MPP6.

Asset	Weight (%)	Performance (%)	rank	R2	rank
VICP-A.ST	3.184	33.927	44	0.040	223
UFLX-B.ST	1.108	19.063	87	0.161	126
SWMA.ST	7.407	14.847	100	0.242	69
STRAX.ST	10.000	2.459	139	0.016	247
SMF.ST	2.218	-23.875	218	0.011	250
SENS.ST	7.423	12.682	107	0.028	235
SCA-A.ST	7.249	94.344	8	0.206	94
SAGA-A.ST	8.649	13.448	103	0.068	199
PROF-B.ST	4.091	73.904	12	0.007	252
PREC.ST	10.000	-58.750	249	0.094	175
ORTI-A.ST	0.566	70.810	14	0.040	223
MIDW-A.ST	1.066	-5.181	168	0.089	184
MCAP.ST	4.347	29.683	56	0.088	186
HTRO.ST	3.894	108.108	6	0.000	255
ENDO.ST	10.000	-9.836	182	0.063	205
EAST.ST	4.303	22.472	73	0.154	133
CRAD-B.ST	1.840	134.855	2	0.109	163
ANOT.ST	2.655	-3.333	160	0.046	219
AGRO.ST	10.000	21.250	78	0.032	230

Table A.11: Asset allocation of portfolio MPP7.

Asset	Weight (%)	Performance (%)	rank	R2	rank
VITR.ST	1.000	65.193	17	0.091	180
TETY.ST	4.042	-15.319	196	0.157	131
TELIA.ST	0.702	5.065	131	0.298	38
SWEC-A.ST	5.000	-16.231	203	0.025	238
STRAX.ST	5.000	2.459	139	0.016	247
SSAB-B.ST	5.000	27.162	64	0.208	91
SSAB-A.ST	5.000	29.844	55	0.208	91
SMF.ST	5.000	-23.875	218	0.011	250
SENS.ST	3.532	12.682	107	0.028	235
SAGA-PREF.ST	4.163	16.703	93	0.049	216
REJL-B.ST	0.109	-18.347	209	0.250	63
POOL-B.ST	1.836	6.613	125	0.120	156
ORTI-A.ST	1.563	70.810	14	0.040	223
NOLA-B.ST	5.000	112.587	4	0.081	189
NAXS.ST	5.000	1.721	145	0.022	242
MEKO.ST	3.652	-9.401	181	0.105	166
LUPE.ST	5.000	-0.997	154	0.185	109
LIAB.ST	3.011	-5.168	167	0.412	6
KLOV-A.ST	5.000	19.282	85	0.095	172
KLED.ST	5.000	9.431	120	0.134	146
HOLM-A.ST	5.000	38.633	35	0.093	177
G5EN.ST	2.193	212.836	1	0.049	216
ENDO.ST	1.153	-9.836	182	0.063	205
CAST.ST	5.000	15.432	96	0.281	47
BMAX.ST	5.000	-9.197	180	0.307	32
BILI-A.ST	2.632	-16.199	202	0.176	116
AXFO.ST	5.000	15.337	97	0.197	100
ANOT.ST	0.410	-3.333	160	0.046	219

Table A.12: Asset allocation of portfolio MPP8.

Asset	Weight (%)	Performance (%)	rank	R2	rank
VIT-B.ST	5.000	16.844	92	0.069	198
TEL2-B.ST	2.678	46.136	28	0.183	110
SWMA.ST	0.982	14.847	100	0.242	69
STRAX.ST	5.000	2.459	139	0.016	247
SSAB-A.ST	5.000	29.844	55	0.208	91
SMF.ST	5.000	-23.875	218	0.011	250
SKIS-B.ST	5.000	6.844	124	0.053	210
SENS.ST	4.981	12.682	107	0.028	235
SCA-A.ST	1.265	94.344	8	0.206	94
SAGA-PREF.ST	5.000	16.703	93	0.049	216
PROF-B.ST	3.455	73.904	12	0.007	252
PREC.ST	0.533	-58.750	249	0.094	175
POOL-B.ST	5.000	6.613	125	0.120	156
ORTI-A.ST	0.630	70.810	14	0.040	223
NOLA-B.ST	3.023	112.587	4	0.081	189
NAXS.ST	5.000	1.721	145	0.022	242
MSON-B.ST	5.000	35.828	40	0.212	85
MIDW-A.ST	4.473	-5.181	168	0.089	184
MCAP.ST	5.000	29.683	56	0.088	186
LATO-B.ST	0.176	28.069	59	0.118	158
KLED.ST	5.000	9.431	120	0.134	146
HIQ.ST	0.100	1.220	148	0.191	107
G5EN.ST	0.909	212.836	1	0.049	216
FING-B.ST	3.350	-73.339	253	0.114	159
ENDO.ST	5.000	-9.836	182	0.063	205
EAST.ST	5.000	22.472	73	0.154	133
CTT.ST	1.859	60.468	19	0.016	247
CORE-PREF.ST	5.000	16.568	94	0.061	206
ANOT.ST	1.586	-3.333	160	0.046	219

Table A.13: Asset allocation of portfolio MPP9.

Asset	Weight (%)	Performance (%)	rank	R2	rank
WIHL.ST	5.000	23.788	71	0.207	93
WALL-B.ST	2.311	13.783	102	0.030	233
VICP-A.ST	5.000	33.927	44	0.040	223
UFLX-B.ST	5.000	19.063	87	0.161	126
SWMA.ST	5.000	14.847	100	0.242	69
STRAX.ST	5.000	2.459	139	0.016	247
SMF.ST	5.000	-23.875	218	0.011	250
SENS.ST	5.000	12.682	107	0.028	235
SCA-A.ST	1.283	94.344	8	0.206	94
SAGA-PREF.ST	5.000	16.703	93	0.049	216
PROF-B.ST	5.000	73.904	12	0.007	252
PREC.ST	0.648	-58.750	249	0.094	175
ORTI-A.ST	0.023	70.810	14	0.040	223
NOLA-B.ST	5.000	112.587	4	0.081	189
NAXS.ST	5.000	1.721	145	0.022	242
MIDW-A.ST	1.542	-5.181	168	0.089	184
LUC.ST	1.075	-13.672	193	0.092	179
KLED.ST	5.000	9.431	120	0.134	146
HOLM-B.ST	2.737	38.115	36	0.099	170
HOLM-A.ST	2.917	38.633	35	0.093	177
G5EN.ST	3.372	212.836	1	0.049	216
FING-B.ST	1.408	-73.339	253	0.114	159
ENDO.ST	1.832	-9.836	182	0.063	205
CTT.ST	5.000	60.468	19	0.016	247
CORE-PREF.ST	4.292	16.568	94	0.061	206
BIOT.ST	5.000	87.514	9	0.255	59
BEIA-B.ST	0.798	21.984	75	0.122	153
AXFO.ST	2.173	15.337	97	0.197	100
ANOT.ST	3.589	-3.333	160	0.046	219

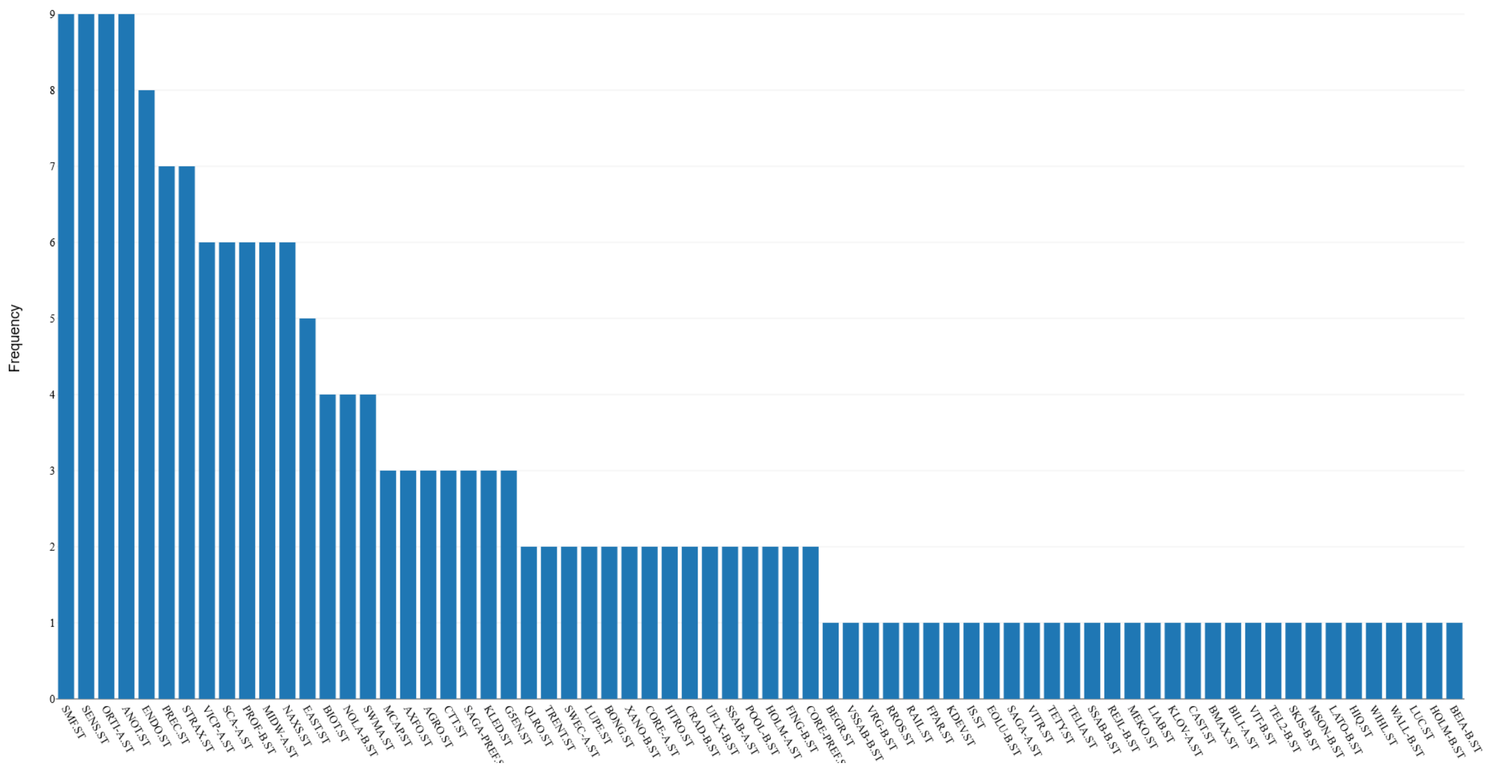

Figure A.1: Assets with frequency in MPP portfolios.

TRITA -SCI-GRU 2019:032