Department of Mathematics

Masters Thesis

Smart Beta - index weighting

Financial Mathematics

Author:

Oscar Blomkvist oscarbl@kth.se

Supervisors:

Filip Lindskog lindskog@kth.se
Salla Franzén salla.franzen@seb.se

7th June 2015

Abstract

This study is a thesis ending a 120 credit masters program in Mathematics with specialization Financial Mathematics and Mathematical Statistics at the Royal Institute of Technology (KTH).

The subject of Smart beta is defined and studied in an index fund context. The portfolio weighting schemes tested are: equally weighting, maximum Sharpe ratio, maximum diversification, and fundamental weighting using P/E-ratios. The outcome of the strategies is measured in performance (accumulated return), risk, and cost of trading, along with measures of the proportions of different assets in the portfolio.

The thesis goes through the steps of collecting, ordering, and "cleaning" the data used in the process. A brief explanation of historical simulation used in estimation of stochastic variables such as expected return and covariance matrices is included, as well as analysis on the data's distribution.

The process of optimization and how rules for being UCITS compliant forms optimization programs with constraints is described.

The results indicate that all, but the most diversified, portfolios tested outperform the market cap weighted portfolio. In all cases, the trading volumes and the market impact is increased, in comparison with the cap weighted portfolio. The Sharpe ratio maximizer yields a high level of return, while keeping the risk low. The fundamentally weighted portfolio performs best, but with higher risk. A combination of the two finds the portfolio with highest return and lowest risk.

Key words: *Smart beta, portfolio optimization, Sharpe ratio, equal weights, diversification, fundamental analysis, P/E-ratio, performance, risk, trading cost, market impact*

Sammanfattning

Denna studie är ett examensarbete som avslutar ett 120 poängs mastersprogram i Matematik med inriktning mot Finansiell Matematik och Matematisk Statistik på Kungliga Tekniska Högskolan (KTH).

Ämnet Smart beta studeras i kontexten av en indexfond, där de olika testade principerna för viktning i portföljerna är: likaviktad, maximerad Sharpe-kvot, maximerad diversifiering, och fundamental viktning användandes av P/E-tal. Utfallet i testerna utvärderas i ackumulerad avkastning, portföljrisk, kostnad att handla i portföljen, och ett antal mått på fördelningen av tillgångarna.

Studien går stegvis igenom processen för att samla in, ordna, och "tvätta" data. En kort förklaring av historisk simulering, metoden för att estimera stokastiska variabler såsom kovariansmatriser, är inkluderad, såväl som en analys av distributionen av data.

Processen för att optimera portföljerna och hur regler för att vara en UCITS-fond kan omformas till optimeringsvillkor beskrivs.

Resultaten indikerar att alla utom den mest diversifierade portföljen har högre ackumulerad avkastning än den marknadsviktade portföljen under testperioden. I alla testade fall ökar handelsvolymen liksom marknadspåverkan när en annan strategi än marknadsviktad används. Portföljen med maximerad Sharpe-kvot ger en hög avkastning med bibehållen låg risk. Den fundamentalt viktade portföljen ger bäst avkastning, men med en litet förhöjd risk. Kombinationen av de båda metoderna ger den portföljen med högst ackumulerad avkastning och samtidigt lägst risk under testperioden.

Nyckelord: *Smart beta, portföljoptimering, Sharpe-kvot, likaviktad, diversifiering, fundamental analys, P/E-tal, avkastning, risk, handelskostnad, marknadspåverkan*

Contents

Nomenclature

t_k	The kth point in time
U	The universe of investable instruments
X_i	An instrument in the universe U
μ_{i,t_k}	The expected return of instrument X_i at time t_k. Variations may occur, with and without time index. $*$ denotes estimate.
Σ_{t_k}	Covariance matrix at time t_k. Variations may occur, with and without time index.
σ_{i,t_k}	Volatility of instrument X_i at time t_k. Variations may occur, with and without time index.
I_U	The indicator matrix of when the instruments (rows) in time (columns) are allowed in the portfolio
P	Adjusted price matrix, with entries as below
R	Return matrix, with entries as below
r_{i,t_k}	The return of instrument X_i at time t_k
MC_{i,t_k}	Market capitalization of instrument X_i at time t_k
EDV_{i,t_k}	Expected Daily Volume of instrument X_i at time t_k
V_{t_k}	Value of the total portfolio at time t_k
VaR_{p,t_k}	Value at Risk at level p of the total portfolio at time t_k
ES_{p,t_k}	Expected Shortfall at level p of the total portfolio at time t_k
I_N	$N * N$ unit matrix

1 Introduction

1.1 Background

During discussions between myself and the head of the Index & Solutions team at Skandinaviska Enskilda Banken AB (SEB), namely Salla Franzén, smart beta came up as a new portfolio strategy for which they needed help with evaluation. A task well suited as a thesis ending the master program in Financial Mathematics and Mathematical Statistics at KTH.

An index portfolio is one that is designed to reflect a certain market's performance and commonly consists of a large number of instruments on that market. Investing in an index fund is said to provide a large market exposure, while keeping the operating expenses low. Indexes are a good tool for determining the wellness of a market along with its direction and trend. For example the Russell 3000 consists of stocks of the US' 3000 largest companies, representing 98% of the investable US equity market and therefore a good representation of the market as a whole. Another example is the Dow Jones Industrial Average, commonly quoted as "the Dow" or "the market" (the US) in media and financial reporting, consisting of 30 significant companies traded on the New York stock exchange. An index consisting of fewer instruments in comparison, but still with a world wide reputation of being a great market indicator.

The common practice of the business today is to weight the instruments in the index relative to the market capitalization of the respective instrument. This is useful for a number of reasons; for example as one company's stock changes in price so does the market capitalization, meaning that the weighting follows along, yielding less need for rebalancing. The market capitalization weighted index has served as a benchmark for the market and an active manager beating this could thus be considered good.

One might wonder why an index weighting strategy that well reflects the market behavior should be complemented or replaced by another. A reason being discussed much, which is also simple to justify with logical reasoning, is that an overpriced asset is weighted higher relative to an underpriced asset. The portfolio representing the index will suffer a harder impact when prices even out, than if the opposite weighting would have been held. In other words mis-pricings in the market has a large negative impact, the opposite of what a skilled active manager would have considered.

An attempt at trying a different weighting scheme was done by Wells Fargo[13] in 1970 by equally weighting the NYSE index, being the first index fund. This, however, proved to be very costly and time consuming; considering the brokerage commissions were set by the NYSE to a level of 10-15 times greater than those of today. They retreated to a buy-and-hold strategy using market capitalization weights, just like the rest of the market.

1.2 What is smart beta?

The phrase "seeking alpha" is often heard or read in the context of portfolio management. It is not necessarily the extra risk that draws the attention, but rather the higher likelihood of a greater return. Smart beta is not inferring that other strategies are "dumb" beta, but perhaps simpler in comparison. Defining what smart beta actually consists of is not an easy task, many funds claim to be managed via smart beta, without actually saying why. Research Affiliates (RAFI)[2] defines it according to Definition 1.1, which is the definition that this thesis will consider true.

Definition 1.1. *A category of valuation-indifferent strategies that consciously and deliberately break the link between the price of an asset and its weight in the portfolio, seeking to earn excess returns over the cap-weighted benchmark by no longer weighting assets proportional to their popularity, while retaining most of the positive attributes of passive indexing. Any strategy that is not valuation-indifferent, that does not break the link between the weight in the portfolio and price (or market cap), is not smart beta.*

The idea is hence to weight assets according to other measures than the size or price of the instruments. One might ask: how should the portfolio be weighted? The easiest and most straight forward strategy is to look at the **equally weighted**, $1/n$, scheme tried by Wells Fargo. This will be the first tested. Secondly, related to Markowitz's theory and the trade off problem is **maximizing the Sharpe ratio**. As described in Markowitz's "Portfolio selection" [19] a **maximization of the diversification** could lead to higher performance and will hence be tested. Theses strategies satisfy the requirements in the smart beta definition, Definition 1.1, by no longer weighting assets proportional to their popularity, while retaining the positive attributes of passive indexing. The final test will be on a fundamentally weighted scheme involving the **P/E-ratios** of the constituents in the index. This strategy does not fall precisely under the definition, as it is dependent on the assets' prices, but is worth to include as it *does break* the link to market capitalization. The exact definitions of the strategies and their mathematical implications will be stated in Sections 2.12 - 2.15.

In various articles regarding smart beta and advertisements for smart beta funds, there is a fifth strategy mentioned, called thematic weighting. This is based on analysis of whether an instrument is mis-priced, and the weight is set accordingly. However, no article or fund that I have come across during this study wishes to share their thematic scheme, for natural reasons. Such a scheme would also be considered active management and is not in the traditional line of index management. I have chosen to leave this out and solely rely on that mis-pricing will be detected by a fundamental strategy.

1.3 Purpose

The purpose of this thesis is to evaluate a handful of strategies for index weighting falling under the definition of smart beta and determine whether it is a good idea to initiate a smart beta strategy for SEB or not. The subject is fairly new in the financial industry and has yet to be implemented by the larger institutions, who seek a greater understanding and evaluation of new practices rather than being first to try a, perhaps, failing strategy.

The evaluation will be in the form of measuring return, risk, outperformance, and the liquidity of the portfolio against that of the market capitalization weighted index. Theses are four key measures which play an important role in the decision making process of a trader.

An institution the size of SEB and others similar to it have a great impact on the market when trading volumes of significant size. The final purpose is to determine the market impact that the tested weighting strategies make.

As described and referenced in Section 1.6 there are a number of studies already done; not as a complete comparison but as individual studies, all showing great results. The problem is that most studies and articles that have been considered and reviewed for this thesis show great results for a specific market and conditions of the study. The question is: how will the strategies perform when they are all tested on the same set of data and given the same possibilities? And not only how they perform, but also what implications the strategies yield for an investor.

1.4 Thesis outline

Section 1 covers the background of the problem and what work has already been done. Section 2 describes the notation of the variables of the portfolio before briefly explaining the theory behind the assumptions made for the strategies in a mathematical sense. The strategies are then explained and motivated. Included are explanations of the optimization's aspects, and risk- and cost measures as well as how the parameter estimation is carried out. Section 3 covers the experimental setup, including data processing and the rules associated with the portfolios. Section 4 displays the main results and references to the complete results. Section 5 interprets the results and Section 6 concludes them. Appendix I contains the majority of the result plots. Appendix II contains proofs of theorems stated without reference.

1.5 Remarks

The intended reader of the thesis is a fellow student at the master program. One is assumed to be familiar with concepts such as: return, volatility, value at risk, expected shortfall, and so on. However, a reader with interest and mathematical knowledge should be able to follow the argumentation and understand the concepts of the thesis.

The index business in general and the business carried out at SEB is referred to as "the business" or "the practice of the business" in the continuation of the report.

The strategies discussed are, after proper introduction, denoted as: CW cap weighed, EQ equally weighed, SR maximum Sharpe ratio weighed, SRc maximum Sharpe ratio weighed with rebalancing constraint, DR most diversified, DRc most diversified with rebalancing constraint, PE fundamentally weighted, and PEc fundamentally weighed with rebalancing constraint. Portfolios with weighting according to a strategy are equally referred to.

Disclaimer: I, the author, will not hold any responsibility to any investment decisions based on the statements of this thesis.

1.6 Review

This section is divided into subsections of articles, not strategies, since the reviewed articles overlap in some areas.

Sharpe and Equal (I)

The idea for the different weighting schemes arouse from an article, Smart Beta 2.0, by EDHEC Risk Institute [1]. They discuss a handful of different schemes, but the ones seemingly most interesting for the thesis and for SEB are the equally weighted, the maximal Sharpe ratio, and the maximal diversification schemes. They show proof of great performance for the equally weighted and the maximal Sharpe ratio weighted strategies, as in Table 1.

	Max Sharpe	Equal
Annual return	7.79%	8.09%
Excess returns over cap	1.72%	2.02%
Annual volatility	20.49%	22.71%
Sharpe ratio	0.30	0.28

Table 1: Results produced by EDHEC Risk Institute [1] for the maximal Sharpe ratio and equally weighted schemes. The period is 21 June 2002 to 31 December 2012 with the Scientific Beta USA Cap weighted index as allowed universe.

An issue with the results might be their real lite implementations. The optimization is not carried out via an algorithm, but rather using the analytical solution for the maximization problem of the Sharpe ratio:

$$\max \quad \frac{\bar{\mu}'\bar{w}}{\bar{w}'\Sigma\bar{w}}$$
$$\Rightarrow \quad \bar{w} = \frac{\Sigma^{-1}\bar{\mu}}{\bar{1}'\Sigma^{-1}\bar{\mu}}$$

as well as in the equally weighted case, by using $\bar{w} = \bar{1}/N$, with N being the number of instruments in the universe, $\bar{w}$ the calculated weights, $\bar{\mu}$ the expected returns, and Σ the covariance matrix. This without any consideration of trading cost or regulations, along with a daily rebalancing. Although great results are shown, a real life implementation needs to be tested.

Sharpe and Equal (II)

Previous studies on the equally weighted and, what they call, a mean-variance weighting have been carried out by De Miguel, Garlappi, and Uppal for their paper Optimal Versus Naive Diversification [5]. The mean-variance strategy is an optimization of the trade off problem:

$$\max \quad \bar{w}'\bar{\mu} - \frac{c}{2}\bar{w}'\Sigma\bar{w}$$

which, later proven, is equal to maximizing the Sharpe ratio. The results they find are as in Table 2.

This shows a somewhat similar story as the results from EDHEC Risk Institute,

	Max Sharpe	Equal
Sharpe ratio	0.36	0.17

Table 2: Results produced by De Miguel, Garlappi, and Uppal showing the Sharpe ratios of the maximal Sharpe ratio weighting and equal weighting schemes. The data are mean values from 6 different universes.

but with a larger difference between the two strategies. More results are available, such as turnover and certainty equivalent return, these have intentionally been left out of this thesis due to their poor quality. However, the results do show proof of greater stability in the equally weighted strategy than in the others tested.

The study uses a universe of 10 industry portfolios and 25 size- and book-to-market-sorted portfolios. The strategies are therefore to select between these preset portfolios and weight them. When the resolution of the universe is increased (more and smaller preset portfolios) the performance per risk of the equally weighted strategy does not evolve at the same rate as for the others, including the Sharpe ratio maximizer. The trading costs increase in comparison to that of the equally weighted, but as performance increases, the equally weighted portfolio is outperformed even with trading costs taken under consideration. With the universe of this thesis being much larger than that used in De Miguel, Garlappi, and Uppal's article, one can expect the equally weighted scheme to be outperformed.

For the equally weighted strategy a reasonable question is: How large does N have to be? The answer is not a simple one and, naturally, different for all universes of instruments. The previous study found that the equal weight strategy outperformed others when $M < 3000$ if $N = 25$ and $M < 6000$ if $N = 50$, with M being number of months that parameter estimations are based on for the benched strategies [5]. This would imply that the "1/N"-strategy will outperform in all cases of this study, a fact that must be tested. This does not have any reasonable explanation other than having a poor method of parameter estimation will be reflected in the results.

Diversification and Equal

In 2008 Journal of Portfolio Management published the article Toward Maximum Diversification by Choueifaty and Coignard [4]. This covers the equally weighted strategy and the maximum diversification strategy and compares these to a cap weighted index. The results for Eurozone equities are as displayed in Figure 1 and Table 3 below.

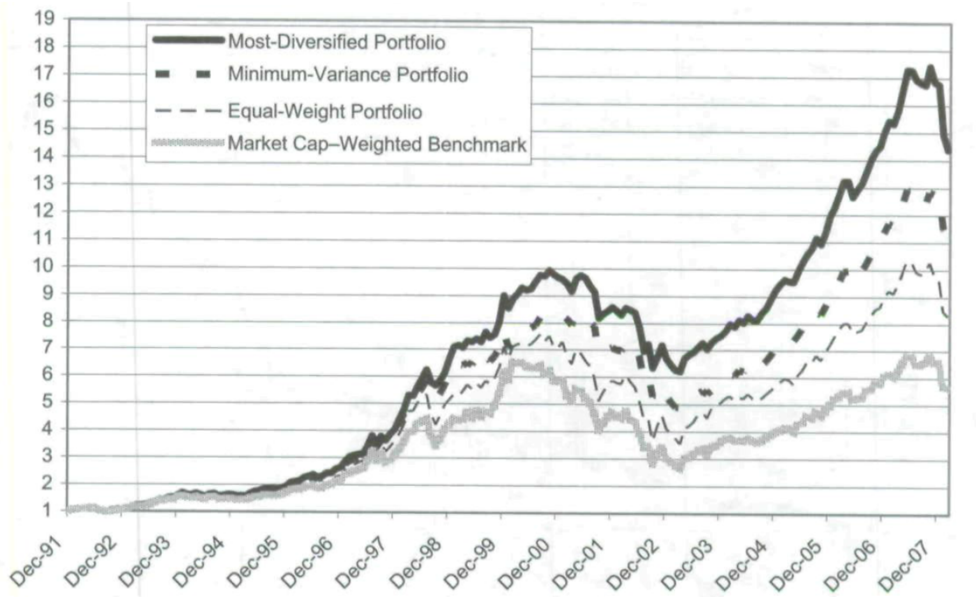

Figure 1: A visualization of the complete series of results presented in Table 3. Graph from Choueifaty and Coignard [4].

Full period 92-08	Max diversification	Equal	Cap
Annual return	17.9%	14.0%	11.3%
Annual excess return	13.3%	9.4%	6.6%
Annual volatility	13.9%	18.1%	17.9%
Sharpe ratio	0.96	0.52	0.37
92-00			
Annual return	28.7%	24.2%	21.6%
Annual excess return	22.9%	18.4%	15.8%
Annual volatility	13.8%	16.4%	16.8%
Sharpe ratio	1.66	1.12	0.94
01-08			
Annual return	5.7%	2.3%	-0.5%
Annual excess return	2.6%	-0.8%	-3.7 %
Annual volatility	13.4%	19.6%	18.9%
Sharpe ratio	0.21	-0.06	-0.20

Table 3: Results produced by houeifaty and Coignard. The cap benchmark is the Dow Jones EuroStoxx Large Cap Total Retum Index, which is also the universe for the other two strategies.

The results are all stating that the most diversified portfolio is the best, showing greater returns while having less volatility and thereby highest Sharpe ratio in all three scenarios. It seems that the most diversified portfolio has a higher volatility in bearish times than the others while having more stable growth in bullish times. The equally weighted portfolio picks out a larger portion of small assets than the cap weighted, due to its nature. From the results, it seems that the smaller assets have a higher volatility, reflected in the larger movements of the equally weighted portfolio in comparison. It is expected that the results of this thesis should reflect those of Choueifaty and Coignard as the universes used are similar.

Fundamental

Investing in undervalued instruments is the sole purpose of a Value Investor and there are many ways of finding those instruments. A much discussed measure is the P/E-ratio, price per share over earnings per share, or a multiple stating how much the market is willing to pay per unit of earnings. Pettersen showed in his thesis [21] that there are proof of a behavior of Swedish stock with low P/E-ratio outperforming the market. A portfolio was selected and rebalanced annually using the 25 stocks with lowest P/E-ratio among the 50% most capitalized. The results are as in Figure 2. Nothing is said about their weighting.

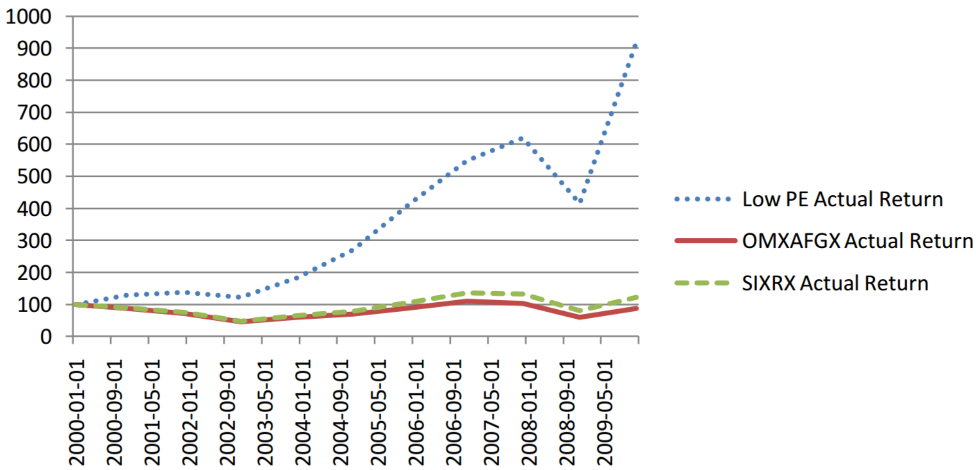

Figure 2: A visualization of the accumulated performance of the P/E-ratio scheme suggested by Pettersen, along with the performance of OMX Affarsvarldens Generalindex and SIX Return Index. Graph from Pettersen [21].

The results show proof of great outperformance of the P/E-ratio strategy proposed. An issue being that much thought has gone into the selection of stocks, the question of the results being sought out, by selecting historical winners in beforehand, arises. This needs to be investigated via testing.

As the half with highest market capitalization of the universe of assets is used, these results might differ from the results in this thesis, using a bigger universe. Looking at the findings of Choueifaty and Coignard [4] and the outperformance of smaller companies, one might expect even better results from the fundamentally weighted strategy of this thesis than that of Pettersen's.

2 Mathematics of the models

This section describes the variables in the construction of the portfolio, the theory of which the weighting schemes are based, the mathematics regarding the models, and the optimization algorithm along with its mathematically formulated constraints. Followed with a description of the calculation of trading volumes as well as the risk measures used and ended with a brief explanation of stochastic parameter estimation.

2.1 Variables of the portfolio

This section covers the main variables of the portfolio and their implications.

Let U be defined as the universe of risky assets $X_i \in (X_1, X_2, \ldots, X_N)$, with their corresponding portfolio weights as the column vector:

$$w = (w_1, w_2, \ldots, w_N)', \quad s.t. \sum_{\forall i} w_i = 1$$

where $'$ denotes the transpose. The instruments have the expected returns $\bar{\mu} = (\mu_1, \mu_2, \ldots, \mu_N)'$ where a return of 1.2 corresponds to a 20% value increase. The covariance matrix is the symmetric matrix:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \cdots & \rho_{1,N}\sigma_1\sigma_N \\ \rho_{1,2}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2,N}\sigma_2\sigma_N \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N,1}\sigma_N\sigma_1 & \rho_{N,2}\sigma_N\sigma_2 & \cdots & \sigma_N^2 \end{pmatrix}$$

where instruments' volatilities are the square root of its diagonal elements, $\bar{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)' = (Diag\,(\Sigma))^{\frac{1}{2}}$ and $\rho_{i,j}$ is the coefficient of correlation between instrument X_i and X_j, taking values -1 to 1.

For the application and model simulation a time index is added with $t_k \in (t_1, t_2, \ldots, t_K)$, of K trading days. The corresponding matrix of weights is hence $\bar{\bar{w}} = (\bar{w}_{t_1}, \bar{w}_{t_2}, \ldots, \bar{w}_{t_K})$, where $\bar{w}_{t_k} = (w_{1,t_k}, w_{2,t_k}, \ldots, w_{N,t_k})'$. Similarly for the returns $\bar{\bar{\mu}}_{t_k} = (\bar{\mu}_{t_1}, \bar{\mu}_{t_2}, \ldots, \bar{\mu}_{t_K})$ and the covariance matrix Σ_{t_k}, with the volatilities $\bar{\sigma}_{t_k} = (\sigma_{1,t_k}, \sigma_{2,t_k}, \ldots, \sigma_{N,t_k})' = (Diag\,(\Sigma_{t_k}))^{\frac{1}{2}}$.

Along with the basic measures, data is associated with the instruments. An $N * K$ indicator matrix I_U is constructed, showing when in time an instrument is allowed in the portfolio, as:

$$I_U = \begin{pmatrix} u_{1,t_1} & u_{1,t_2} & \cdots & u_{1,t_K} \\ u_{2,t_1} & u_{2,t_2} & \cdots & u_{2,t_K} \\ \vdots & \vdots & \ddots & \vdots \\ u_{N,t_1} & u_{N,t_2} & \cdots & u_{N,t_K} \end{pmatrix}$$

where u_{i,t_k} is one when instrument X_i is allowed in the portfolio at time t_k and zero otherwise. The exact construction of this matrix is discussed in Section 3.2.

The matrix of prices adjusted for dividends and splits is denoted P with entries p_{i,t'_k}, for $t'_k \in (t_0, t_1, t_2, \ldots, t_K)$, corresponding to the price of one share of instrument X_i at time t'_k. From this the matrix of returns R with entries r_{i,t_k} may be calculated as:

$$r_{i,t_k} = \frac{p_{i,t'_k}}{p_{i,t'_{k-1}}}$$

As the resulting return matrix R has one less column than the matrix of prices P, an extra historical time point t_0 of prices is required to match the size of R with the other variables.

The market capitalizations of the instruments are found in the matrix MC with entries MC_{i,t_k} of the price, p_{i,t_k} of instrument X_i at time t_k multiplied with the corresponding total number of shares, or simply the total value of each company. When indexed as MC_{t_k} the column of all instruments at time t_k is referred to.

The volume of trades on the market in monetary units for a specific instrument in time, or more commonly denoted the expected daily volume is the matrix EDV, with entries EDV_{i,t_k} for instrument X_i at time t_k. Calculated from the five day average daily volumes ADV with entries ADV_{i,t'_k}, for $t'_k \in (t_0, t_1, t_2, \ldots, t_K)$, as $EDV_{i,t_k} = ADV_{i,t'_{k-1}}$. The extra time point for the same reason as above.

The risk free rate of return will simply be denoted r_0 or for a specified rate for a period in time T, r_T. For examle the three month rate is denoted $r_{3/12}$. This none or single index notation of the interest rate is not to be mixed up with the double index notation of the entries r_{i,t_k} of the return matrix R.

The total value of the portfolio at time t_k is denoted V_{t_k} and is calculated using the initial value of the portfolio V_0 and the weighted returns, as:

$$V_{t_k} = V_0 \prod_{t \leq t_k} \bar{r}'_t \bar{w}_t$$

2.2 Distributions

In the estimation of the distribution of which to base the calculations of the risk measures, one would prefer to analyze every distribution separately at each time point, this is united with manual inspection after fitting of different distribution types, and would lead to an unreasonable amount of computations needed.

Looking at the distribution of all assets' returns for all times in an unsorted environment, one may get a hint of the individual distributions. As seen in Figure 3, a t-location scale (Students t) distribution captures the form of the histogram of returns, as well as falling neatly in a straight line passing through the origin in the QQ-plot. The relatively few points deviating from the straight line in the QQ-plot's tails are indicating that the sample has slightly heavier tails than the fitted distribution. This could affect the calculations of Expected Shortfall, but with as few as a handful of points out of nearly 70000, these will not be adjusted for.

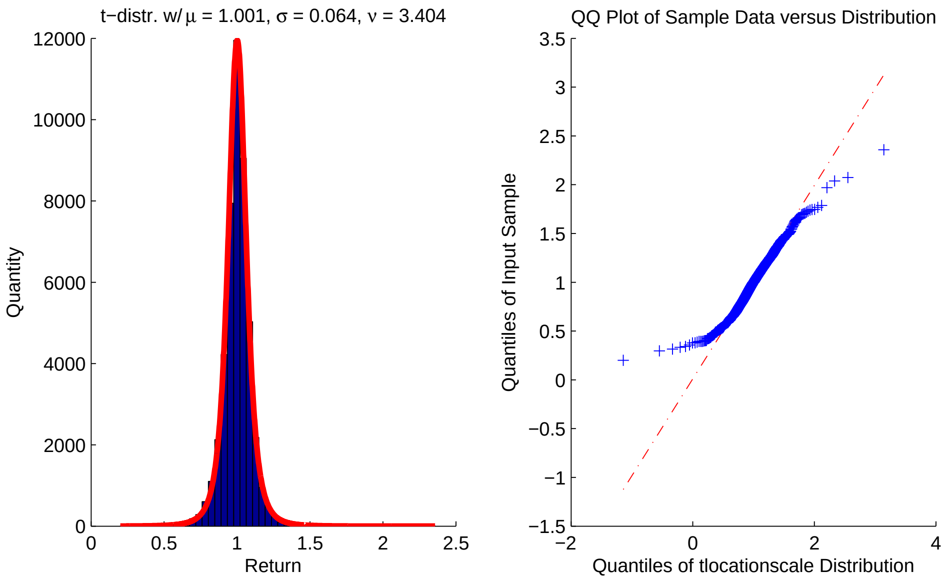

Figure 3: *Left* Histogram of unsorted monthly returns for all instruments, with fitted t-location scale distribution. *Right* QQ-plot of said sample's quantiles and the fitted distribution's quantiles.

Looking at a sample of size 70000 of unsorted monthly returns simulated from daily returns, as described in Section 2.3, seen in Figure 4, a similar behavior as for the real monthly returns is observed. The estimated variables are the same, apart from the variance, which is slightly higher, also observed in the QQ-plot. Using the simulated returns as sample for estimation of the distribution when calculating risk, might hence compensate for the actual distribution's under-weighted tails. In all, use of the historical simulations is justified.

2.3 Parameter estimation using historical simulation

The return data is assumed independent and identically distributed on the interval used for estimation and the size of this interval is one year. Justification of this assumption is discussed in Section 3.1.

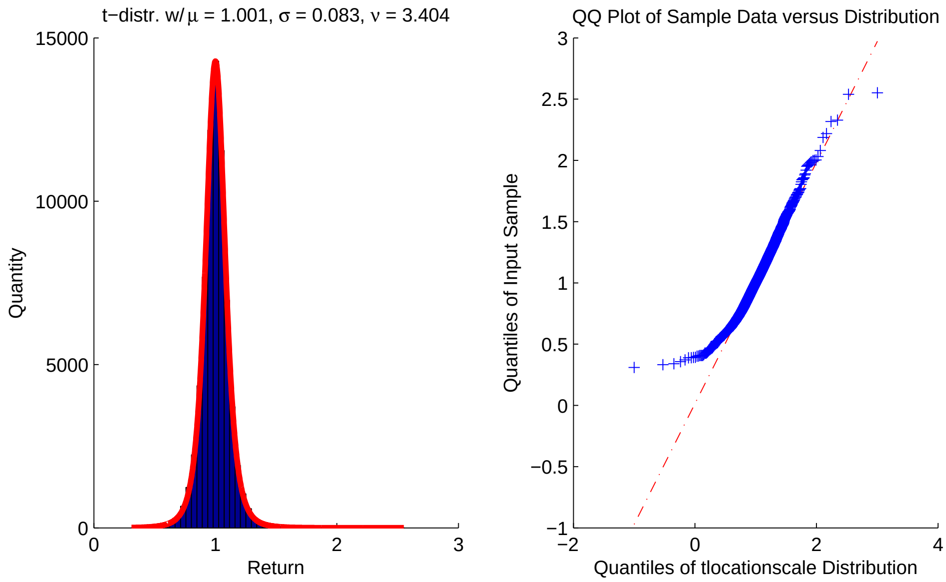

Figure 4: *Left* Histogram of unsorted monthly returns simulated from daily returns for all instruments, with fitted t-location scale distribution. *Right* QQ-plot of said sample's quantiles and the fitted distribution's quantiles.

The typical number of instruments used in the portfolio at each time point is roughly 400 instruments, even by using all historical daily return data from the 252 trading days of the selected interval, the covariance matrix, Σ will not have full rank. Since if n is the number of variables and t is the number of observations,

$$\text{rank}(\Sigma) \leq \min(n, t-1)$$

but Σ has size $n * n$. This means that more observations are needed in order to get a positive semi-definite covariance matrix. This does not imply that a larger interval of historical return data is needed, since this would put restrictions on what instruments are allowed in the portfolio, but rather more observations within that interval.

There are plenty of ways to do this; the easiest would be to scale daily return data with the square root difference in time to monthly data, but 252 is still smaller than 400. A more robust and elegant method is historical simulation, requiring the assumption of a possible dependence in space, but independence in time of the return data.

Let S_{i,t_k} represent the set of the 252 historical daily returns preceding time t_k of instrument i. From this set, 21 (=252/12) constituents are taken at random with replacement and multiplied to form the simulated monthly return S^1_{i,t_k}. This is done a number of times, preferably a number larger than the number of instruments, in this case 5000 times, to form the simulated set of monthly returns $S^*_{i,t_k} = \left(S^1_{i,t_k}, S^2_{i,t_k}, \ldots, S^{5000}_{i,t_k}\right)$. These sets are those from which the covarinace matrices at each rebalancing time are estimated. Note that returns are picked as complete columns to keep the dependence in space.

The estimation of the covariance matrix is calculated as:

$$\Sigma_{t_k} = \frac{1}{5000-1}\sum_{j=1}^{5000}\left(\bar{S}_{t_k} - \bar{\mu}_{t_k}\right)\left(\bar{S}_{t_k} - \bar{\mu}_{t_k}\right)', \quad \bar{S}_{t_k} = \begin{pmatrix} S^j_{1,t_k} \\ S^j_{2,t_k} \\ \vdots \\ S^j_{N,t_k} \end{pmatrix}, \quad \bar{\mu}_{t_k} = \begin{pmatrix} \mu_{1,t_k} \\ \mu_{2,t_k} \\ \vdots \\ \mu_{N,t_k} \end{pmatrix}$$

The assumption of the returns being independent in time implies that reordering the daily returns and therefrom creating new monthly returns will not affect the estimations in a negative way, but rather improve them.

From a set of 252 data points there are $\binom{252}{21} \approx 10^{30}$ ways of choosing 21 elements. A selection of 5000 will mean a low probability of choosing the same elements for two or more calculations of simulated returns, thus a low probability of getting linearly dependent columns in the, therefrom, calculated sample covariance matrix.

Estimating the expected return μ_{i,t_k} for instrument i at time t_k is done differently. Although the assumption of no dependence in time for the detrended returns, one may assume that there is a trend or momentum in the sampled series. Using the simulated returns, S^*_{i,t_k}, the possibility of finding such trends would be lost. A more suitable estimation method is to use the twelve sampled monthly returns and weight them, in this case linearly, with the vector $\bar{v} = (1, 2, \ldots, 12)/78$. This picks up and weights resent movements while adjusting for any great happenings in the past. The calculations are hence as:

$$\mu_{i,t_k} = \left(r_{i,t_{k-11}}, r_{i,t_{k-10}}, \ldots, r_{i,t_k}\right)\bar{v}'$$

The notation Σ_{t_k} and $\bar{\mu}_{t_k}$ are the parameters estimated at time t_k and are those assumed to be the expected parameters at time t_{k+1}.

Remark One might suggest that a risk factor model should be used in a setting like this, there are however several institutions using complete covariance matrices, one being Ossiam [20].

2.4 Risk

Measuring risk is not a trivial task: the volatility of the portfolio gives a hint to the level of risk, but it does not tell the whole story. One way to study risk is to look at the distribution of the possible outcomes of the portfolio to determine the behavior in the tails.

The easy approach is to look at the ordered historical outcomes of the portfolio. E.g. a set of 100 historical outcomes are ordered in ascending order, the fifth constituent would represent the outcome where a worse outcome is at most 5% likely.

A more robust method is to fit a distribution to the possible outcomes. At each point in time examined; from the set of historical returns for the individual instruments S^* explained in Section 2.3, the weighted mean is calculated. The weights are those chosen by the strategy at this point in time, creating a set of historical outcomes of this specific portfolio, S_V^*.
To these outcomes, S_V^*, a distribution function f_V is fitted using a maximum likelihood method [17] and the parameters, μ, σ and ν are estimated in the t-location scale distribution. Motivation for using this distribution is found in Section 2.2.

The first measure, the **"Value at Risk", VaR**, at level $p \in (0,1)$ of a portfolio with value V_{t_k} at time t_k is: [10]

$$\mathrm{VaR}_p\left(V_{t_k}\right) = \min\{m : \mathrm{P}\left(m * r_0 + V_{t_k} <\right) \leq p\}$$

Or, in words, the amount needed to be invested at time t_{k-1} in the risk free asset with return r_0 to ensure that the probability of a strictly negative portfolio at time t_k is at most p. In practice this is calculated from the fitted distribution at time t_k as:

$$\mathrm{VaR}_{p,t_k} = V_{t_{k-1}} - F_V^{-1}(1-p)V_{t_{k-1}}/r_0$$

where $F_V^{-1}(1-p)$ is the ordinary inverse of the $(1-p)$ quantile of the return distribution, f_V. Or simplified; the return where there is a probability p of getting a worse return.

The second measure, **"Expected Shortfall", ES**, at level $p \in (0,1)$ of a portfolio with value V_{t_k} at time t_k is: [10]

$$\mathrm{ES}_p\left(V_{t_k}\right) = \frac{1}{p}\int_0^p \mathrm{VaR}_u\left(V_{t_k}\right) du$$

Essentially the same as considering VaR_p at all levels below and including p and calculating their average. This is in practice calculated as:

$$\mathrm{ES}_{p,t_k} = V_{t_{k-1}} - FF_V^{-1}(1-p)V_{t_{k-1}}/r_0$$

where $FF_V^{-1}(1-p) = \frac{1}{p}\int_0^p F_V^{-1}(1-p)$.

Including VaR is simply because it is the measure mostly used in the business, the choice of also including ES is because of its possibility to detect any risk "hidden in the tail". E.g. a value of ES much higher than that of VaR implies that the distribution of returns for the portfolio has a heavy left tail, meaning that lowering p would dramatically increase VaR.

The measures are lastly divided with the portfolio value $V_{t_{k-1}}$ to get a percentage value comparable in time.

2.5 Trading costs (without market impact)

The cap weighted index has a significant advantage regarding rebalancing: it is *only* necessary when there are in- or outflows from the fund, or when an asset enters or leaves the index, unlike other strategies which need, more or less, to be rebalanced "continuously" in order to stay within the strategy. Doing so is costly and keeping the absolute rebalancing required low is preferred. Calculating the needed addition c_{i,t_k} to each instrument i at time t_k may be done in several ways, although Theorem 2.1 will be used in this study. One could argue that the rebalancing volume is just a matter of calculating the difference between weights at times t_k and t_{k-1} and multiplying by the portfolio size. This is, however, only true when all instruments have performed equally. All other cases cause the weights to change from time to time due to difference in performance, so called drift. Hence the more advanced method described.

The trading cost may then be calculated as a fixed percentage of the volume, in this study 2 bps.

Theorem 2.1. *The volume added at time t_k to instrument i denoted c_{i,t_k} is calculated as:*

$$c_{i,t_k} = V_{t_{k-1}} \left(w_{i,t} R_{t_k} - w_{i,t_{k-1}} r_{i,t_k} \right)$$

where $V_{t_{k-1}}$ is the value of the total portfolio at time t_{k-1}, R_{t_k} the return of the total portfolio from time t_{k-1} to t_k, and r_{i,t_k} the return of instrument i from time t_{k-1} to t_k.

2.6 Liquidity and Market Impact

An issue with a growing portfolio is the risk of losing liquidity. A small private investor with a portfolio of a couple thousand Euro will never have the problem of his portfolio being illiquid when average daily volumes are in the size of several million Euro per instrument. However, a portfolio in the magnitude of several billion Euro and around 1000 instruments (roughly this study's proportions), will have positions in the size of the daily volumes. The time for trading increases in order to keep a low market impact and hence the liquidity decreases.

There are several measures of market impact, non of which can be completely trusted, since they are all based on assumptions that cannot be tested (you cannot interact in the market and *not* interact in the market at the same time). Two measures have been considered; the time to sell the complete portfolio and a measure derived by JP Morgan, which we will call Market Impact I and II respectively.

The time to sell the complete holding of asset X_i at time t_k is defined as:

$$L_{i,t_k} = \frac{V_{t_k} * w_{i,t_k}}{EDV_{i,t_k} * 0.1}$$

where EDV_{i,t_k} is the expected daily volume traded of asset X_i at time t_k. The factor is chosen to 10% market participation, with motivation to it being a guideline of the business. The total time to sell the portfolio at time t_k is calculated as:

$$TL_{t_k} = \max\left(L_{1,t_k}, L_{2,t_k}, \ldots, L_{N,t_k}\right)$$

Referred to as Market impact I. Assuming that price impact caused by participation in the market is uncorrelated for all instruments.

The measure derived by JP Morgan, Market Impact II, consists of two terms; the permanent impact and the temporary impact caused by market participation. It is stated as: [8]

$$JPM = \underbrace{\frac{5}{100}I}_{\text{perm.}} + \underbrace{1.4\frac{95}{100}\frac{volume}{EDV}I}_{\text{temp.}}, \quad I = 0.187\sqrt{\frac{volume}{EDV}}\sigma^2$$

Using the asset and time dependent notation of this thesis and setting the volume as the trading volumes c_{i,t_k} defined in Section 2.5 we get the equation:

$$JPM_{i,t_k} = \frac{5}{100}I_{i,t_k} + 1.4\frac{95}{100}\frac{|c_{i,t_k}|}{EDV_{i,t_k}}I_{i,t_k}, \quad I_{i,t_k} = 0.187\sqrt{\frac{|c_{i,t_k}|}{EDV_{i,t_k}}}\sigma^2_{i,t_k}$$

The two parts (permanent and temporary) are later displayed separately, but are generally summed over all assets X_i to get the total market impact at time t_k as:

$$JPMT_{t_k} = \sum_{i=1}^{N}\frac{5}{100}I_{i,t_k} + \sum_{i=1}^{N}1.4\frac{95}{100}\frac{|c_{i,t_k}|}{EDV_{i,t_k}}I_{i,t_k}, \quad I_{i,t_k} = 0.187\sqrt{\frac{|c_{i,t_k}|}{EDV_{i,t_k}}}\sigma^2_{i,t_k}$$

Referred to as Market impact II. While the first measure tells a somewhat understandable story the second is not an absolute measure that may be applied to the return of the portfolio, but rather a measure that may be compared between the strategies.

2.7 Rules and regulations

There are a number of rules set up, for the Undertaking for Collective Investment in Transferable Securities (UCITS), to the extent that they can work as constraints in an optimization, Definition 2.1. Following these will not only ease the optimization, but also ensure that the results from the study may be used by an actual investor without modification.

Definition 2.1. *The UCITS rules contain, but are not restricted to:[23]*

1. *UCITS can invest in an absolute minimum of 16 assets: 4 holdings of up to 10% each plus 12 holdings of up to 5% each.*
2. *A UCITS fund may invest no more than 5% of its value in approved securities or money market instruments issued by any one body. This limit can be increased to*

10% provided that the total value of any holdings between 5% and 10% does not exceed 40% of the fund.

3. *No more than 20% of the fund as deposits with any one bank*
4. *No more than 20% of the fund invested in any one other fund*
5. *Up to 35% of the fund in any one bond issue provided the rest of the fund is invested in other types of assets; or a minimum of six issues if the fund is over 35% invested in Government bonds.*
6. *No more than 10% exposed in derivatives with another bank as counterpart*
7. *Hold no more than 20% of the voting shares of a company*
8. *Hold no more than 10% of the bonds issued by a company*
9. *Hold no more than 20% of the value of another fund*

As this study will only cover stock as instruments, rules 4, 5, 6, 8, and 9 will not apply. Rule 1 is only a result from taking the limit of as few holdings as possible of rule 2. For the special case of this study the remaining rules will apply, Definition 2.2.

Definition 2.2. *The constraints of the studied portfolios will follow the rules:*

1. *Portfolios may contain up to 5% of its value in approved securities or money market instruments issued by any one body. This limit can be increased to 10% provided that the total value of any holdings between 5% and 10% does not exceed 40% of the portfolio.*
2. *Portfolios may contain no more than 20% of the voting shares of a company*

How the rules are modeled as mathematical optimization constraints may be further read about in Section 2.8.

2.8 Optimization constraints, including UCITS rules

The chosen Matlab optimizer `fmincon`() takes inequality constraints $A * \bar{w} \leq b$, equality constraints $A_{eq} * \bar{w} = b_{eq}$, upper and lower bounds, and non-linear constraints $f_{nl}(\bar{w})$ as inputs. Leading to natural separation of linear and non-linear constraints.

Even though all constraints, including the linear, may be modeled in the function file $f_{nl}(\bar{w})$ as general constraints; doing so has proven to slow down the optimizer as well as reduce its possibilities of finding an optima.

Linear constraints

The first condition, evolving from the common business practice, is that no short selling is allowed and that the portfolio is completely invested according to the index, meaning no fraction in a risk free asset, along with a financing condition. In our notation:

$$w_{i,t_k} \geq 0 \quad \forall i, t_k \tag{2.1}$$

$$\sum_{\forall i} w_{i,t_k} = 1 \quad \forall t_k \tag{2.2}$$

The **UCITS** rules and regulations for the portfolios, Section 2.7, may be mathematically represented as optimization constraints with the linear parts in Equations (2.3) and (2.4). Firstly the constraint of no weight being greater than 10% of the total portfolio and secondly that no more than 20% of a company may be owned.

$$w_{i,t_k} \leq 0.1 \quad \forall i \in U, \quad \forall t_k \tag{2.3}$$

$$V_{t_k} * w_{i,t_k} \leq MC_{i,t_k} * 0.2 \quad \forall i \in U, \quad \forall t_k \tag{2.4}$$

Equations (2.1) and (2.3) serve as the lower and upper bounds respectively. Equation (2.4) is modeled as an inequality constraint with $A_{t_k} = I_N$ the identity matrix, and $\bar{b}_{t_k} = \frac{0.2*\bar{MC}_{t_k}}{V_{t_k}}$, with V_{t_k} being the total portfolio value and MC_{i,t_k} the market capitalization of instrument X_i at time t_k. Equation (2.2) is modeled as an equality constraint, with $A_{eq} = \bar{1}$ the row vector of ones and $b_{eq} = 1$.

Non-linear constraints

The often referred to **5-10-40 rule** of the UCITS rules makes for a typically non-linear constraint. It says that the weights between 5% and 10% cannot, when summed up, constitute more than 40% of the total portfolio value. In our notation:

$$\sum_{\forall i} (w_{i,t_k} | w_{i,t_k} > 0.05) \leq 0.4 \quad \forall t_k \tag{2.5}$$

$$\Rightarrow \sum_{\forall i} w_{i,t_k} * I\left(w_{i,t_k} > 0.05\right) \leq 0.4 \quad \forall t_k \tag{2.6}$$

where, in this case, I represents the indicator function taking value 1 when the argument is greater than or equal to zero and value 0 otherwise. Modeling this equation yields a non-smooth step function and as it is explained on Mathworks' web site regarding the requirements on the non-linear functions put into `fmincon()` [18]:

> *[The function] is not smooth, which is a general requirement for constraint functions [...] Nevertheless, the method often works.*

Interpreted as: a non-smooth function works sometimes. As no assumptions of this being one of those cases will be made, a more precise method is needed.

A simple solution of using the Heaviside function, which derivative may be represented by Dirac's delta, proved to be useless.

Kienitz and Wetterau [11] states that `fmincon()` is applicable when both the objective and the constraint functions are twice differentiable. Finding a good approximation for the step function which is at least twice differentiable would hence solve the problem. The obvious choice would be a hyperbolic tangent, $\tanh(w)$, shifted in location and amplitude so that it has its "upward jump" from 0 to 1 at 0.05. Resulting in:

$$I\left(w_{i,t_k} > 0.05\right) \approx \frac{1}{2} + \frac{1}{2}\tanh\left(h\left(w_{i,t_k} - 0.05\right)\right) \tag{2.7}$$

Where h is a constant managing the form of the step; the higher h the better the approximation. Figure 5 shows the approximation for different h. Using this approxim-

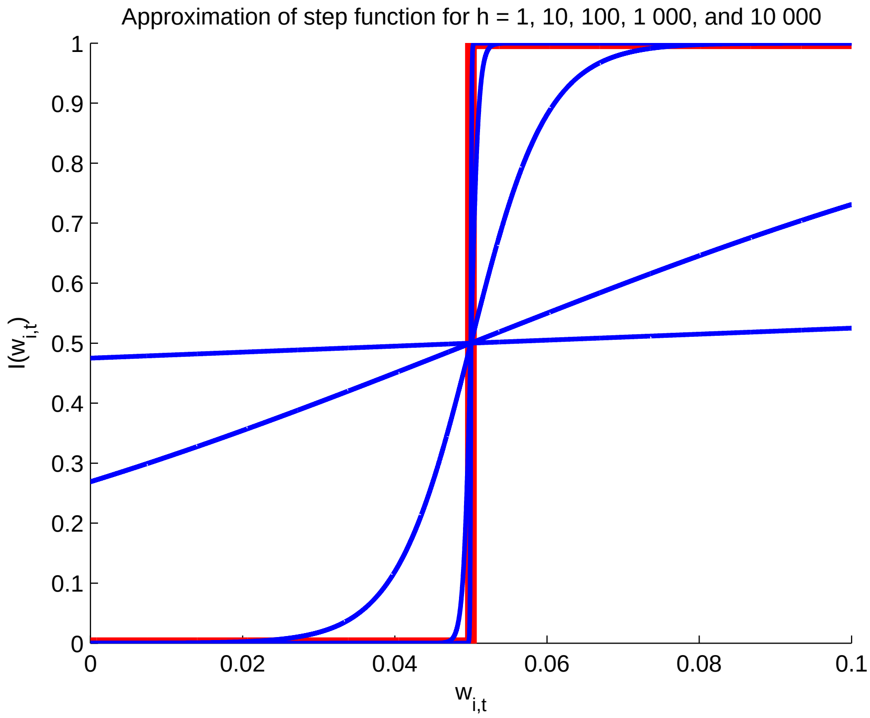

Figure 5: Plot of the Heaviside function around 0.05 along with the approximation Equation (2.7) for increasing h. For $h = 1000$ it is still possible to spot the difference between the two functions, but for $h = 10000$ there is little to tell them apart.

ation as constraint (2.8) with $h = 5000$ gave good results in terms of no time points being infeasible (all optimizations could be solved) as well as coming close to the 5-10-40 rule's upper limit in several cases, meaning no apparent restriction caused by the approximation. In financial terms; one does not wish to rebalance to a level on the boundary, but rather just below it, in order to have a certain margin for asset movements. Thus the approximation is satisfactory.

$$\sum_{\forall i} w_{i,t_k} * \left(\frac{1}{2} + \frac{1}{2}\tanh\left(h\left(w_{i,t_k} - 0.05\right)\right)\right) \leq 0.4 \quad \forall k \tag{2.8}$$

A **limitation of the rebalancing** in terms of restricting the value of the parameters c_{i,t_k} from Theorem 2.1 is wanted in some of the strategies, implemented as Equation (2.9). Say for example the strategy of maximizing the Sharpe ratio; it might give (has proven to give) large variations to what instruments should be chosen in each time step, hence causing large transfer volumes at rebalancing. In some cases a complete change of the portfolio. Restricting this will not only make for a decrease in trading costs, but also force a certain stability to the portfolio.

$$\sum_{\forall i} c_{i,t_k} \leq C \quad \forall t_k \tag{2.9}$$

Where C is the chosen maximum level of rebalancing.

2.9 Optimization algorithm

The problem to be solved will consist of around 400 variables at each point in time and should hence be considered large. However, mistaking a large problem for a *large-scale* problem is easily done. A *large-scale* problem is one that partly is large, but also sparse, meaning that the algorithm may use sparse algebra when solving for optimality. This problem should be considered *medium-scale* , meaning that dense algebra should be used by the solver. While `fmincon`() can handle both variants not all algorithms can. The algorithm recommended by Mathworks is the interior-point algorithm, this can handle both *large-* and *medium-scale* problems and the constraints needed.

According to Mathworks the interior-point approach handles most problems given to it, as long as the functions are smooth.

The interior-point traverses the interior of the feasible region to find an optima, unlike the simplex method which searches along the boundary set by the constraints; the edge of the feasible region, visualized in Figure 6

As the maximal diversification and maximal Sharpe ratio strategies use estimated parameters in their objective functions, there is a possibility of getting the results `NaN` or `Inf`, due to poor estimations. The interior-point approach can recover from such results and continue the search, making it well suited. [15]

The interior-point algorithm, also referred to as Barrier methods, works via `fmincon`() as to solve the problem: [16]

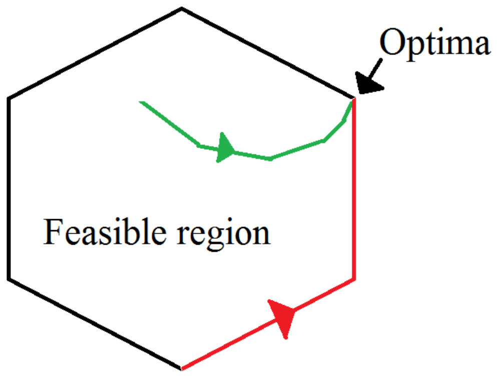

Figure 6: Visualization of a path of the interior-point method (green) and the simplex method (red).

$$\begin{aligned} \min_{x} \quad & f(x) \\ s.t. \quad & g_1(x) = 0 \\ & g_2(x) \leq 0 \end{aligned}$$

This is approximated so that for each $\gamma > 0$, with a slack variable $s_j > 0$ (greater than zero to ensure the natural logarithm is bounded) for every constraint, the problem becomes:

$$\begin{aligned} \min_{x,s} \quad f_\lambda(x,s) = \min_{x,s} \quad & f(x) - \gamma \sum_j \ln(s_j) \\ s.t. \quad & g_1(x) = 0 \\ & g_2(x) + s = 0 \end{aligned} \tag{2.10}$$

as $\gamma \to 0$ the minimum of the approximate $f_\lambda(x,s)$ should converge to the minimum of $f(x)$. The logarithmic term is what is referred to as the barrier function.

The now equality constrained minimization problem is easier to solve than the original problem, this is done via taking either i direct step (Newton step) or a conjugate gradient step, the first is the initial try. If the step does not lower the value of the merit function $f_\gamma(x,s) + \nu||\left(g_1(x), s + g_2(x)\right)||$ or the result is `NaN` or `Inf`, a new step is tried. The step is calculated by solving the system:

$$\begin{bmatrix} H & 0 & J'_{g_1} & J'_{g_2} \\ 0 & S\Lambda & 0-S \\ J_{g_1} & 0 & I & 0 \\ J_{g_2} & -S & 0 & I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta s \\ -\Delta y \\ -\Delta\lambda \end{bmatrix} = \begin{bmatrix} \nabla f - J'_{g_1} y - J'_{g_2}\lambda \\ S\lambda - \gamma\bar{1} \\ g_2 + 1 \end{bmatrix}$$

where J denotes the respective Lagrangian, $S = Diag(s)$, λ is the Lagrange multiplier associated with the constraints g_2, $\Lambda = Diag(\lambda)$, y is the Lagrangian multiplier associated with the constraints g_1 and H is calculated according to:

$$H = \nabla^2 f(x) + \sum_j \lambda_j \nabla^2 g_{1,j}(x) + \sum_j y_j \nabla^2 g_{2,j}(x)$$

If a direct step is, for some reason, not possible, the algorithm takes a conjugate gradient step. The algorithm adjusts both s and x while keeping the slack variables s positive, with the objective to minimize a quadratic approximation of the system 2.10 while fulfilling the linearized constraints. The Lagrange multipliers are obtained by, in a least squares sense, approximately solving:

$$\nabla_x L = \nabla_x f(x) + \sum_j \lambda_j \nabla g_{1,j}(x) + \sum_j y_j \nabla g_{2,j}(x)$$

The step $(\Delta x, \Delta s)$ is taken in order to solve:

$$\begin{aligned}
\min_{\Delta x, \Delta s} \quad & \nabla f' \Delta x + \frac{1}{2}\Delta x' \nabla^2_{xx} L \Delta x + \gamma \bar{1}' S^{-1} \Delta s + \frac{1}{2} \Delta s' S^{-1} \Lambda \Delta s \\
s.t. \quad & g_1(x) + J_{g_1} \Delta x = 0 \\
& g_2(x) + J_{g_2} \Delta x + \Delta s = 0
\end{aligned}$$

The algorithm tries to minimize the norm of the, yet again, transformed constraints.

The algorithm stops when one of the limits is reached, for example the length of the current step is smaller than its tolerance, the result of the objective function is within its tolerance, the number of iterations has reached its maximum, and so forth.

2.10 Markowitz's theory

Initially investors picked out assets which they considered good in terms of return and risk and created a portfolio from there. Harry Markowitz's theory dating back to 1952 is what we refer to as Modern Portfolio Theory today. He introduced not only the concepts of looking at assets' expected return and risk, but also their interrelations in risk and movement, measured as correlation.

In Modern Portfolio Theory each asset X is represented by a normal distribution with mean μ and volatility σ (standard deviation). The key, though, is the correlation between any two assets i and j, the correlation coefficient:

$$\rho_{i,j} = \frac{Cov(X_i, X_j)}{\sigma_i \sigma_j} = \frac{\mathrm{E}\left[(X_i - \mu_i)(X_j - \mu_j)\right]}{\sigma_i \sigma_j}$$

The case of a universe with two assets X_i and X_j where:

$$\mu_i \geq \mu_j, \quad \sigma_i \leq \sigma_j$$

X_i would be considered the better asset and finding a combination of the two having a higher return per unit risk than any of the individual assets, might not be possible and is solely dependent on the value of $\rho_{i,j}$.

A portfolio consisting of two assets i and j where:

$$\mu_i \geq \mu_j, \quad \sigma_i \geq \sigma_j$$

will according to Markowitz perform better when combined, giving the portfolio a higher return per unit of risk than any of the individual assets. Finding this portfolio is however a more delicate task that involves the correlation between the assets, as well as finding the efficient frontier.

Efficient frontier

Consider a portfolio with combinations of an arbitrary number of assets with returns μ, volatilities σ and correlation coefficients ρ. For each combination the resulting portfolio will have a pair of portfolio return and portfolio risk (μ_p, σ_p), and can be referred to as possible investments. Compare with Figure 7.

If for each possible level of portfolio return, the level of portfolio risk is minimized then the resulting pairs of $(\mu_p, \sigma_p) = (\mu_p, \min[\sigma_p|\mu_p])$ will create a line, namely the efficient frontier. Green in Figure 7.

The risk free investment has zero risk and a return r_0 at some level, 7% in Figure 7. If a tangent line is drawn through the portfolio consisting of just the risk free investment and the upper tangent point of the efficient frontier, this is called the capital market line. The line between the two points is the efficient frontier of a portfolio including the risky assets and the risk free assets. The tangent point is the market portfolio, the one considered the preferred portfolio. The theory states that one should not hold any other portfolio than a combination of the risk free asset and the market portfolio.

In a universe where the risk free rate exists, but is not allowed in the portfolio the market portfolio should be held. Finding this portfolio is perhaps not easy, but may be done through maximization of the trade off problem:

$$\max \ \mu_p - c\sigma_p^2$$

where a sought level of risk c of the investor determines the optimal portfolio. A strategy often defined in basic courses in portfolio theory. According to Theorem 2.2 this is equal to maximizing the Sharpe ratio of the portfolio, defined as:

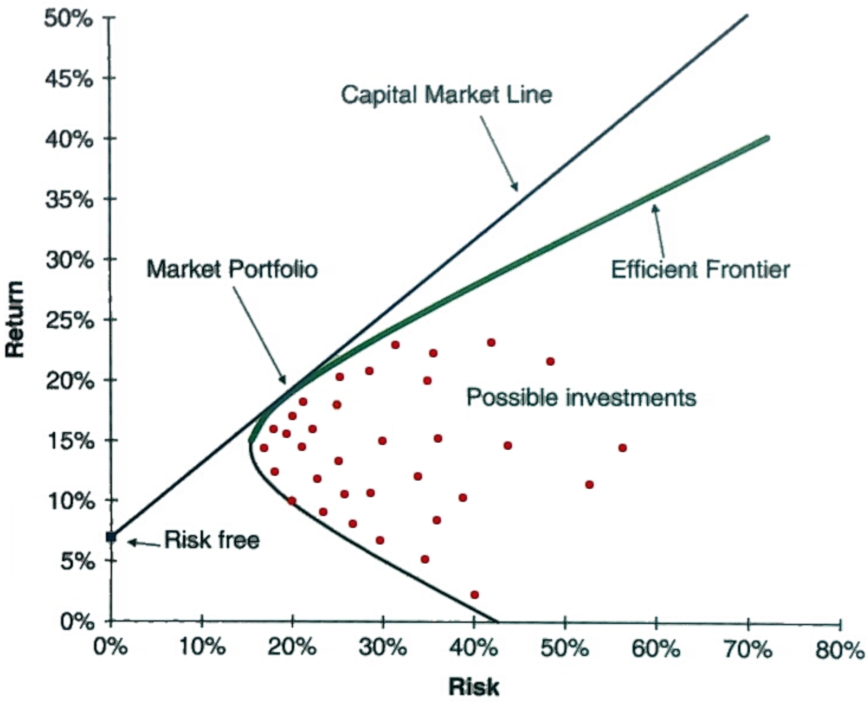

Figure 7: The efficient frontier for a universe of assets. Original graph from [26].

$$SR_p = \frac{\mu_p - r_0}{\sigma_p} \tag{2.11}$$

The advantage with maximizing the Sharpe ratio, being that the trade-off parameter c is determined via the optimal portfolio in Equation (2.4) rather than arbitrarily.

Kopman and Liu [12] stated the following theorem:

Theorem 2.2. *Assume one wishes to solve the program:*

$$\begin{aligned} &\max \frac{A(x)}{\sqrt{B(x)}} \\ &s.t.\ g_i(x) \leq 0, \quad \forall i \end{aligned} \tag{2.12}$$

where $A(x)$, $B(x)$ *and* $\frac{A(x)}{\sqrt{B(x)}}$ *are all convex functions, then there exists a* c *for which the optimal* x^* *in program* (2.12) *is the same as the optimal in* (2.13)*:*

$$\begin{aligned} &\max A(x) - cB(x) \\ &s.t.\ g_i(x) \leq 0, \quad \forall i \end{aligned} \tag{2.13}$$

Not involving the mathematics of optimization, one can easily determine this by looking at Figure 7. The slope of the capital market line is calculated as $\frac{\mu_p - r_0}{\sigma_p - 0} = \frac{\mu_p - r_0}{\sigma_p} = SR_p$ which is the definition of the Sharpe ratio precisely. Hence maximizing the Sharpe ratio is equal to maximizing the slope of the capital market line, i.e. finding the tangent point, being the market portfolio.

These results arise the interest of investigating the strategy of maximizing the portfolio's Sharpe ratio.

Diversification by Booth and Fama

Dating back to the bible (Ecclesiastes 11:2 NLT) a division of your assets among different instruments is preferred, since you do not know what risks lie ahead. A more modern understanding is given by Markowitz [19] in the 50's explaining diversification as one almost certain way of lowering risk without lowering the expected return. Markowitz states that:

> *The hypothesis (or maxim) that the investor does (or should) maximize discounted return must be rejected. If we ignore market imperfections the foregoing rule never implies that there is a diversified portfolio which is preferable to all non-diversified portfolios. Diversification is both observed and sensible; a rule of behavior which does not imply the superiority of diversification must be rejected both as a hypothesis and as a maxim.*

In the essence, although taken out of context, it is clear a diversified portfolio should be preferred over a non-diversified at all times.

The main idea with diversification is to remove reli on chance; the outcome of the investment should not be solely dependent on choosing the best performing instruments, but rather overall performance of the market. Making it quite important when trying to reflect the market via an index.

Booth and Fama [3] showed that with a constant percentage invested in each asset, the portfolio compound return is greater than the weighted average of the compound returns on the assets in the portfolio. Meaning that an asset's compound return is smaller than its contribution to the portfolio compound return. They state that this difference is an incremental return due to diversification. This may be shown, in short as follows.

An asset X_i's continuously compounded return is $\ln[1 + R_i]$, where R_i is the simple return. Taylor series expansion around the mean return of the asset $E[R_i]$ can be used to express the expected value of the compound return $E[\ln[1 + R_i]]$.

$$E[\ln[1+R_i]] = \ln[1+E[R_i]] - \frac{M_2}{2(1+E[R_i])^2} + \frac{M_3}{3(1+E[R_i])^3} - \frac{M_4}{4(1+E[R_i])^4} + \dots \tag{2.14}$$

where $M_k = E\left[(R_i - E[R_i])^k\right]$ is the kth moment of the asset return around its mean. Well estimated by:

$$E[\ln[1+R_i]] = \ln[1+E[R_i]] - \frac{\sigma_i^2}{2(1+E[R_i])^2} \tag{2.15}$$

with σ_i^2 being the variance of the simple returns of X_i. Similarly as in Equation (2.14) the expected value of the compound return of portfolio P may be expressed as:

$$E[\ln[1+R_P]] = \ln[1+E[R_P]] - \frac{\sigma_P^2}{2(1+E[R_P])^2} \tag{2.16}$$

Essentially, Equations (2.15) and (2.16) say that the higher the variance the lower the expected value of the compounded return gets. Next the relation between the risk of the individual asset and the portfolio in the relative term $\beta_{i,P}$ is stated as: (read Section 2.11 for an explanation of the β relationship)

$$Cov(R_i, R_P) = \sigma_P^2 \beta_{i,P}$$

Where it is true that $\sigma_P^2 = \sum_{\forall i \in P} w_i \beta_{i,P} \sigma_P^2$ with w_i being the weight of the ith asset. Another estimate of the ith asset's contribution to the portfolio's compounded return is hence:

$$E[\ln[1+R_i]] = \ln[1+E[R_i]] - \frac{\sigma_P^2 \beta_{i,P}}{2(1+E[R_i])^2} \tag{2.17}$$

The difference between Equation (2.15) and (2.17) is in the second term, more precisely the difference between σ_i^2 and $\sigma_i^2 \beta_{i,P}$. Asset X_i's contribution to the variance of the portfolio return, $\sigma_P^2 \beta_{i,P}$, is less than the variance of X_i's return, σ_i^2, hence asset X_i's return contribution is greater than its compound return. This enhanced contribution to the portfolio's compound return is a result by lowering the risk via diversification. Or in other words, one may find instruments with higher volatility than their contribution to the portfolio volatility.

Markowitz's theory and the results by Booth and Fama suggest something in the line of Figure 8, where an increase in diversification could choose a more efficient pair of μ and σ in a return/risk perspective. The above argumentation give reason to believe that a scheme maximizing the diversification would be suitable in an index context.

2.11 CAPM-theory

The Capital Asset Pricing Model derived in the 1960's, but credited to William Sharpe, is a simplification and continuation of the work done by Markowitz. The model determines the theoretically appropriate expected return of an asset, μ_i, given that of the market, μ_M, the risk free rate, r_0, and the β_i-value compared with the market, as:

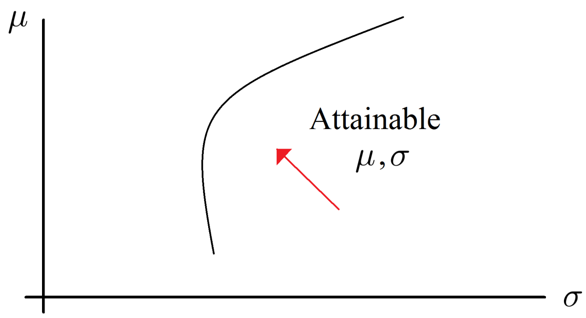

Figure 8: Visual representation of possible return/risk pairs. The red curve represents a possible shift in μ, σ pairs obtained by diversification.

$$\mu_i = r_0 + \beta_i \left(\mu_M - r_0\right) \tag{2.18}$$

Meaning that a more leveraged expected excess return over the risk free rate, should yield a higher valuation of the asset X_i, as it should yield a higher return. Investigating this measure further one may derive the Sharpe ratio-rho relationship [27] of an asset compared to the market, as:

$$\begin{aligned}
\mu_i - r_0 =& \beta_i \left(\mu_M - r_0\right) \\
\{\beta_i = \frac{Cov\left(X_i, X_M\right)}{\sigma_M^2} =& \frac{\rho_{i,M}\sigma_i\sigma_M}{\sigma_M^2} = \rho_{i,M}\frac{\sigma_i}{\sigma_M}\} \\
\mu_i - r_0 =& \rho_{i,M}\frac{\sigma_i}{\sigma_M}\left(\mu_M - r_0\right) \\
\frac{\mu_i - r_0}{\sigma_i} =& \rho_{i,M}\frac{\mu_M - r_0}{\sigma_M} \\
SR_i =& \rho_{i,M} SR_M
\end{aligned}$$

similar for a complete portfolio instead of a single asset. This does not say much about how to choose a well performing portfolio, but may give some enlightenment to when one expects the the equally weighted portfolio to perform well.

Finding instruments with high β is analogous to finding those with high σ, that at the same time follows the market's movements. An equally weighting scheme adds weight to small cap assets while removing from large cap assets, in comparison to a cap weighted index. In doing so, one expects the smaller assets to follow the market more precisely, perhaps due to large dependency on the bigger companies in the market. As well as suggesting that the smaller assets have a higher volatility and therefore will respond more to market movements.

Just looking at the correlation coefficient, if the scheme actually increases the overall correlation with the market for each asset in the portfolio, one can expect a higher return per risk, or Sharpe ratio.

Looking more closely at Equation (2.18), an increase in β suggests an increase in return. Increasing β is, as proposed, related to increasing the constituents volatility,

σ, which is precisely the enumerator of the maximization of diversification scheme, Definition 2.5, further justifying that scheme.

2.12 Equal weighting

Referred to as the $1/N$-method, or in our case EQ, where all instruments have the same weight in the portfolio.

An apparent advantage of an equally weighted scheme is that the portfolio will, in some sense, be automatically diversified, with significant holdings in the whole universe. Another would be that the issue regarding the market cap weighting scheme of overweighting overpriced and underweighting underpriced instruments will no longer be systematic, but rather random.

A significant disadvantage, only reduced in the market cap weighting scheme, is the need of substantial rebalancing due to fluctuations in prices. Along with this comes the strongly related issue that rebalancing forces the investor to sell instruments that have performed well during the last period and buy those having performed poorly. A universe containing instruments of vastly different market capitalizations reduces the opportunity for large portfolios, since the smaller companies set an absolute upper limit. [25]

An absolute condition set by the UCITS-rules, Definition 2.1, is that $N \geq 20$ in order for $w_i \leq 0.05 \; \forall i$, this should not be a problem as the universe considered in this thesis has around 400 possible assets to include at all times.

As the portfolio grows, eventually an upper limit to when it is possible to weight all assets equally while honoring the UCITS rules will be reached. To overcome this issue, an optimization will be carried out, as:

$$\min \quad (\bar{w}_{t_k} - w_{eq,t_k})' \, (\bar{w}_{t_k} - w_{eq,t_k})$$

where $w_{eq,t_k} = \frac{1}{N_{t_k}}$, the quota of one over the number of instruments allowed in the portfolio at time t_k. Ensuring feasibility at all points in time.

Retreating to an equally weighted scheme would imply belief in a hypothesis stating that **the outperformance of an instrument is random and independent of the market size. An equal spread within the universe will take use of this randomness. The unweighted average over an arbitrary, but sufficiently long, period of time of returns within the universe is positive.**

2.13 Weighting for maximal Sharpe ratio

The Sharpe Ratio in its simplest form is the excess return over risk, Definition 2.3.

Definition 2.3. *The Sharpe ratio is defined as:[22]*

$$SR_i = \frac{E\left[X_i - r_0\right]}{\sqrt{Var\left(X_i\right)}} = \frac{\mu_i - r_0}{\sigma_i}$$

with μ_i being the return of the risky asset, r_0 the risk free return, and σ_i the volatility of the risky asset.

Formulating the Sharpe ratio for a complete portfolio P is rather straight forward, beginning with the expected return:

$$\mu_p - r_0 = \bar{\mu}'\bar{w} - r_0 \tag{2.19}$$

where μ_i is the expected return for asset X_i. The volatility is a bit trickier; but is derived as:

$$\begin{aligned}
\sigma_P^2 =& Var\left(P\right) = Var\left(\sum_{\forall i \in U} w_i X_i\right) = \\
=& \sum_{\forall i \in U} w_i^2 Var\left(X_i\right) + 2 \sum_{1 \leq i} \sum_{i<j} w_i w_j Cov\left(X_i, X_j\right) = \\
=& \sum_{\forall i,j \in U} w_i w_j Cov\left(X_i, X_j\right) = \\
=& \sum_{\forall i,j \in U} w_i \Sigma_{i,j} w_j = \bar{w}'\Sigma\bar{w} \\
\Rightarrow \quad & \sigma_P = \sqrt{\bar{w}'\Sigma\bar{w}}
\end{aligned} \tag{2.20}$$

where Σ is the covariance matrix. Equation (2.19) and (2.20) hence lead to the definition of the Sharpe ratio for a portfolio, Definition 2.4.

Definition 2.4. *The Sharpe ratio for a portfolio P is defined as:*

$$SR_P = \frac{\bar{\mu}'\bar{w} - r_0}{\sqrt{\bar{w}'\Sigma\bar{w}}} \tag{2.21}$$

The optimization problem to be solved is hence:

$$\max \quad \frac{\bar{\mu}_{t_k}'\bar{w}_{t_k} - r_0}{\sqrt{\bar{w}_{t_k}'\Sigma_{t_k}\bar{w}_{t_k}}} \tag{2.22}$$

Theorem 2.3. *Let G be defined the feasible region, and define $\bar{w}(\lambda) = \arg\max_{w \in G} \bar{\mu}'\bar{w} - \lambda\bar{w}'\Sigma\bar{w}$. Then:*

$$SR(\lambda) = \frac{\bar{\mu}'\bar{w}(\lambda)}{\sqrt{\bar{w}'(\lambda)\Sigma\bar{w}(\lambda)}}$$

is unimodal; increasing to the left and decreasing to the right of the optimal point.

Equation (2.22) is, according to the proof of Theorem 2.2 and the resulting Theorem 2.3, convex and hence possible to use as an objective function in an optimization problem along with arbitrary constraints. Meaning that it is possible to find a feasible maximum in a surrounding, making it suitable in portfolio optimization.

The importance of comparing the Sharpe ratio lies perhaps not in comparing absolute terms, even though many consider a Sharpe ratio above 0.4 as good for an active investor, it says little when taken out of the perspectives of general market performance and risk. Consider instead the following example; portfolio A has an expected return of 7% and volatility 20%, while portfolio B has the same expected return, its volatility is 25%. The linear proportionality of risk and return is thus broken, and there is no mathematical incentive for investing in B, having the lower Sharpe ratio of the two. Sharpe ratios of the different strategies will be compared. One might ask what the point wold be, since this strategy clearly aims at maximizing it. The issue is that the optimization is highly dependent of the estimation of the return and covariance matrix in the next point in time. Poor estimations could lead to a strategy such as the equally weighting having a higher actual Sharpe ratio when measured on the present point in time.

The theory and the background arise the hypothesis that **the maximization of the Sharpe ratio will yield a more constant return at a steady level of risk, compared to other models. In a handful of cases, poor estimation of the stochastic parameters will yield a lower actual Sharpe ratio than the other strategies.**

2.14 Weighting for maximal diversification

Diversification could be done in a number of ways, an easy, but not very sophisticated method could be to use the weights from the market cap weighted index $\bar{w}_{mc}$ and create new weights according to the scheme:[1]

$$\bar{w} = \frac{\bar{w}_{mc}^{p*}}{\bar{1}'\bar{w}_{mc}^{p*}} \quad 0 \leq p \leq 1 \tag{2.23}$$

With $p*$ being the element-wise power of the vector. Although Equation (2.23) would perhaps yield a more diversified portfolio in comparison to the market cap weighted, it would be difficult to analytically justify it in terms of the theory presented in Section 2.10, as well as make UCITS compliant. A more suited method would be one independent of the market capitalization (also needed to fall under Definition 1.1), beginning with defining the level of diversification according to Definition 2.5.

Definition 2.5. *The diversification ratio $D(P)$ for a portfolio P is defined by Choueifaty and Coignard [4] as:*

$$D(P) = \frac{\bar{w}'\bar{\sigma}}{\sqrt{\bar{w}'\Sigma\bar{w}}} \tag{2.24}$$

where $\bar{w}$ are the portfolio weights, $\bar{\sigma}$ the assets' volatilities and Σ their covariance matrix. $D(P)$ is thus the weighted average of volatilities over the portfolio volatility. Expressed with time dependence in Equation 2.26.

Lets investigate this ratio further. To simplify the calculations, introduce a universe of synthetic assets $U_S = \{Y_1, Y_2, \ldots, Y_N\}$ such that:

$$Y_i = \frac{X_i}{\sigma_i} + \left(1 - \frac{1}{\sigma_i}\right) B$$

Where B is a risk free asset (bond), thus the volatilities for the synthetic assets will become $\bar{\sigma}_{Si} = \bar{1}$. Since the weights $\bar{w}$ in Equation (2.24) add up to one then $\bar{w}\bar{\sigma}_S = \bar{w}\bar{1} = 1$ and thus Equation (2.24) is simplified to:

$$D(S) = \frac{\bar{w}'\bar{\sigma}_S}{\sqrt{\bar{w}'\Sigma_S\bar{w}}} = \frac{1}{\sqrt{\bar{w}'\Sigma_S\bar{w}}} \tag{2.25}$$

Since correlation does not change with leverage, Σ_S is equal to the correlation matrix Σ of the original assets in U. Maximizing the diversification ratio of S, Equation (2.25), is thus the same as minimizing $\bar{w}'\Sigma_S\bar{w} = \bar{w}'\Sigma\bar{w}$. Which is precisely the same as minimizing the portfolio volatility and thus reducing the risk of the portfolio. This somewhat analytically proves part of the hypothesis. One might only say "somewhat...part" since this is a special case of assets with equal volatilities.

A further argument that risk is linearly related to return in a fashion as $\bar{w}\bar{\mu} \propto \bar{w}\bar{\sigma}$ would in some sense explain the argument in the enumerator in Equation (2.24). Also related to the argumentation in Section 2.11 where an increase of the individual assets risk implies a higher expected return. Maximizing the diversification ratio would hence mean maximizing the return per portfolio risk, similar to maximizing the Sharpe ratio and approaching the market portfolio.

The optimization problem to be solved is hence:

$$\max \quad \frac{\bar{\sigma}'_{t_k}\bar{w}_{t_k}}{\sqrt{\bar{w}'_{t_k}\Sigma_{t_k}\bar{w}_{t_k}}} \tag{2.26}$$

Where the Diversification ratio at each point in time is maximized.

The results by Booth and Fama, Section 2.10, and the CAPM-theory, Section 2.11, arise the hypothesis; **increasing the diversification will reduce the risk at the same time as the expected return will increase in comparison to the market cap weighted index.**

2.15 Fundamental weighting

Fundamentally weighted portfolios, or in other words; portfolios dependent on fundamental values in the companies' (assets') books. A method many active investors - particularly value investors - use to find undervalued assets. Doing so is by many not considered something that a machine can perform, but something that is rather a form of art. This along with the unwillingness of successful investors to share their secrets on value investing, gives a difficult task at hand.

A measure much discussed is the P/E-ratio, the price per share over earnings per share. Some state that this measure is useless and others that it is of great value. It is thus an interesting measure to test.

P/E-ratios typically range from 5 to 30, but outliers are of course possible.

It is obvious that an asset with a high P/E-ratio is more risky; since there is evidently an anticipated growth causing the high ratio, not fulfilling this will have a large impact on the price. As long as there is an anticipated growth the ratio will stay high, but as soon as the growth slows down the growth in price will slow down as well and the ratio stabilize around a normal value. Having assets with high P/E-ratio in the portfolio requires the investor to know when to pull out and sell, this might be hard to learn a machine in a simple model, though.

An asset having a too low P/E-ratio is connected with a poor future outlook or an asset that has matured. Having an index as universe all assets will be considered "good" and no consideration of having a too low ratio will be taken. [7]

The assumption made is that a lower P/E-ratio in relation to the rest of the universe is a sign of an undervalued asset, and similarly a high P/E-ratio is a sign of over valuation. The weights should hence be relative to the inverted P/E-ratio as $\bar{w} \propto \left(\frac{E}{P}\right)^{p*}$. The constant $p*$ is chosen to 2, to create a reasonable spread between the highest and lowest weighted assets, this is highly dependent on the ratios for a particular universe. As in the equally weighting scheme, the weights must be calculated through solving an optimization problem to ensure honoring the UCITS rules, namely:

$$\min \quad \left(\bar{w}_{t_k} - \bar{w}_{t_k}^{pe}\right)' \left(\bar{w}_{t_k} - \bar{w}_{t_k}^{pe}\right)$$

where

$$w_{i,t_k}^{pe} = \frac{1}{(PE_{i,t_k})^2} * \sum_{i=1}^{N_{t_k}} (PE_{i,t_k})^2$$

is the squared E/P-ratio adjusted so that $|\bar{w}_{t_k}^{pe}| = 1$, where PE_{i,t_k} is the P/E-ratio of instrument X_i at time t_k.

The theory yields the hypothesis that **the fundamentally weighted portfolio using P/E-ratios will pick under valued assets and thus generally perform better than the other strategies.**

3 Simulation of performance

3.1 Setup

Markets and asset classes

When reading about previous studies carried out in the field of smart beta, one discovers a consensus that a too concentrated market is in general not well suited.[24] Looking at the Swedish market and perhaps the OMX Nordic 40 would be an interesting case, although an index consisting of 40 companies of a, in the context, small market size would not be a practical application. A combination of a larger developed market and an emerging market (EM) seems like the preferred option, but due to lack of data in EM's this part will be ruled out as some of the strategies require historical data.

The easiest choice would be to look at a well established index of large companies and easy access to data, such as the Standard and Poor's 500. The downside is that many studies has taken this easy path and any new discoveries would not be probable. Along with this is the fact that SEB's focus is not on the US, but rather on Europe.

The choice of examining the MSCI Europe Developed index is for a couple of reasons; the index is quoted in Euro rather than the domestic currencies, sufficient amounts of historical data is available in an easy fashion, and the market is wide enough for a smart beta strategy to perform. This is also a preferred market for SEB and their customers.

One could make the remark that each strategy should be applied to its preferred market and universe of assets. This would be the case in a real life application, although before this could be done one must establish which market that is. Choosing a large market with many different assets and asset sizes yields for possibilities for the different strategies to find the instruments which it prefers. At the same time as the strategies have the same conditions to work on, therefore easing the comparison of performance and risk, say.

Price (and other data)

The practice of the business is to trade on closing prices. Closing prices are according to Nasdaq defined as: [9]

> *Price of the last transaction of a particular stock completed during a day's trading session on an exchange.*

When trading "after hours" this is the most up to date valuation of an asset, even though the actual price might change before the market opens again the next trading day.

This study chooses to look at closing prices and therefrom calculate daily returns and so forth. This is for three particular reasons; its the business practice, it gives equally spaced data in time for all instruments, and there is no need to filter noise caused by individual ticks during the day.

A universe consisting of instruments quoted on markets with different currencies arises currency issues. If a portfolio uses Euro, the performance of an instrument will be affected by the movement of the instrument as well as the movement of the instrument's quoted currency. Separating these movements could be useful in order to determine hedging against currency risk, but that is not the purpose of this thesis.
The adjusted daily closing prices have been provided via Bloomberg with consent from Daniel Lobo at Bloomberg Sweden. Having data on 734 past constituents of the index.

To calculate the fraction owned of a single company data on the market sizes at each time point is needed, this was provided from SEB's database. The five day average daily volume of each asset could also be retrieved therefrom.

To maximize the Sharpe ratio there is a need for an accurate quote on the risk free rate. The choice fell upon the three month EURIBOR rate, which is widely used in the business. The one month rate could then easily be calculated, given that $r_0 = 1.02$ for a 2% rate e.g., as:

$$r_{1/12} = \left(r_{3/12}\right)^{1/3}$$

Yielding the historical interest rates displayed in Figure 9.

The P/E-ratios required for the fundamental strategy are provided by Derek Laliberte at ABG Sundal Collier from their database. This data is not the easiest to come by and complete sets were therefore not possible to get, but the 402 (out of a possible 734) provided should be enough for testing. This is a random selection and not one with any intention of choosing extra well performing or extra low risk assets.

Time interval for estimation

Choosing the size of the window of historical data to be used in the parameter estimation is not a trivial task, but one of many aspects. To begin the search a mathematical analysis of the data set at hand is done; in Figure 10 a representation of the autocorrelation function for different lags is visualized. The historical return data series were first removed of any linear trends, then normalized by division by the respective standard deviations. Therefrom, each series' sample autocorelation was determined.

The mean curve shows that even though some series where proven to have significant correlations for different lags, no particular lag stands out as present in all series. The maximum curve witnesses that there are those series that do have significant correlations. Since no single GARCH model (with the same number of coefficients, but with different coefficient values) may be fitted to all series, the return data is assumed independent and identically distributed for all individual series. Assuming otherwise would result in difficulties when modeling the instruments' correlation in space.

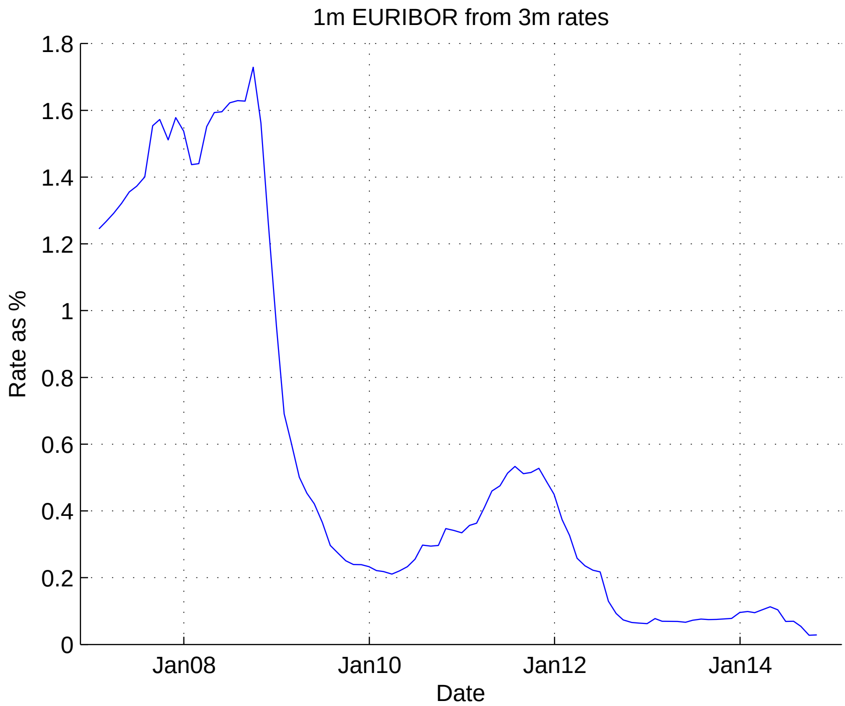

Figure 9: Plot of the historical one month interest rates as calculated from three month EURIBOR.

As the mathematical analysis cannot give a sound answer, reasoning has to be done. It is discussed in Towers & Watson [24] that a sample of size five to ten years back in time from the point of estimation is well suited, which is also supported by DeMiguel, Garlappi and Uppal. [5]. When looking at a short horizon and choosing the universe arbitrarily, a sample size of this proportion might be suitable. However, looking at a universe mastered by an index, where instruments enter and leave often, such a long sample is unreasonable. Largely because of the difficulties of finding that much historical data on the younger constituents. From this perspective a preferred size would be one month of data, so that all new constituents may enter the portfolio. This is not feasible in the context of making good estimations of the future return or the correlation between instruments. E.g. from Section 2.3, there would only be one way of choosing a set for estimation.
A more reasonable size is one year, which has proven successful when used in a test setting. This size is also supported as a reasonable size from a business perspective, by Salla Franzén.

Trivially, for the estimation in every point t_k to be usable and independent of the point t_{k+1} when comparing performance, the interval for estimation may only contain points $\tilde{t} \leq t_k$, so called out-of-sample.

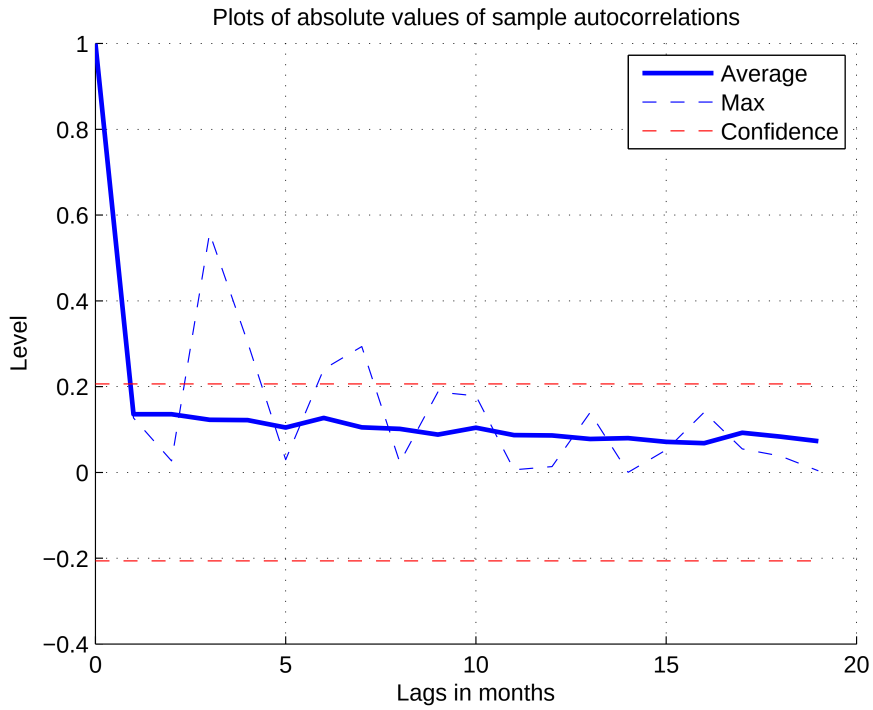

Figure 10: Sample autocorrelations for monthly return data. The 'Average' refers to the mean of the absolute values of each series' sample autocorrelation function. 'Max' refers to the individual series having the largest absolute value of its autocorrelation. 'Confidence' are the 95% confidence intervals. Of course there is a correlation of 1 at lag 0.

Time step

A time step as short as between individual 'ticks' would seem very ambitious, but would not only cause unwanted noise and fluctuations, but also create unnecessary problems. First of all, a stock with high liquidity would generate a very large amount of data, considering a time interval of several years back in time, making the estimation time consuming. Secondly, comparing two data sets where the values are indifferent on the time axis would require some sort of interpolation, which is the case when comparing the close to randomly spread 'ticks' of stock. Thirdly, it is unreasonable to rebalance a portfolio of the investigated size as often as a usage of all 'ticks' would motivate.

A time step of one trading day seems more reasonable for data collection. This would yield a reasonable amount of data, time series with equal time stamps for every stock and a somewhat doable rebalancing scheme. A longer time step will in some sense remove daily fluctuations caused by individual placed orders and give a more stable series reflecting the market's movements.

One might think that the fundamental method causes a special case; data will only be available on a quarterly basis, since it is found in the companies' quarterly reports. That is partially true, but since the price changes more frequently the P/E-ratios can be reevaluated using today's price with the last report's earnings quote. Resulting in time series with the same frequency as the purely price dependent strategies'.

Initial portfolio size

The least amount of money available in the universe at any historical point in time is 2.7 trillion Euro ($2.7 * 10^{12}$ Euro), this sets an upper limit to a portfolio's size. Given that no more than 20% of an asset may be owned, that an outperformance of the cap weighted index of 100% should be feasible, that weights can differ with a factor 10 from the highest to the lowest, and a factor 2 in margin; the resulting initial portfolio size is around 10 billion Euro ($10 * 10^9$ Euro). This will be a reasonable size of the initial capital of the portfolio; large enough to reach the limits of the UCITS-rules, but small enough to not force the portfolio to the limit of being cap weighted.

3.2 Data processing

The price data collected from Bloomberg consists of daily closing prices quoted in Euro that have been adjusted for dividends, meaning; if a dividend d_{i,t_k} is payed at time t_k from instrument X_i, then this is added to the price of the instrument. So an owner of the instrument will see no difference in the portfolio value without the need of adding the dividend payment. The downside being that no adjustment for the cost of the instant reinvestment of the dividends could be made. Assuming that all instruments have payed a percentagewise equal dividend this affect may be neglected when strategies are compared.

The data is inserted into a matrix with instruments $X_i \in U$ as rows and times t_k as columns. The visualized sparsity pattern of missing data for a selection of the price data is shown in Figure 11.

Missing data

No historical sets of data are perfect, there are data missing and there is nothing you can do about it, except filling the holes possible to fill. This is the first step of the data processing.

Looking at Figure 11, when there are a number of prices missing in the beginning of a row, one can safely assume that the instrument did not exist at that time and similarly for the end, that the instrument has ceased to exist. The missing data in the form of points or vertical lines are those in need of special attention. Columns with more than 75% of the data missing are deleted and considered holidays in a significantly large part of the market. The level of 75% were not arbitrarily chosen, but the data indicated a large drop above that level, dates removed included December 24th, January 1st, and those around Easter. The sparsity pattern of the price series after removal of holidays is visualized in Figure 12

The prices still needed to be filled are those not beginning or ending a row, i.e. prices missing due to a trading stop, a national holiday, or something similar. The, perhaps, correct way to fill them in a time series perspective is to interpolate between the data earlier and later in time or with a moving average. A better and more suitable method for financial data is to use the price from the most recent day before. This

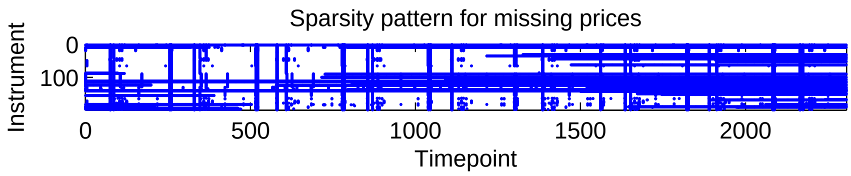

Figure 11: Plot of the sparsity pattern of a selection of the missing prices in the data set.

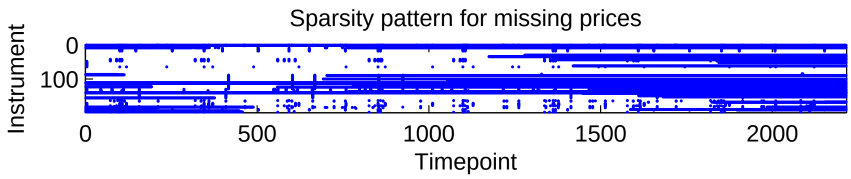

Figure 12: Plot of the sparsity pattern of the missing prices in the data set. Holidays removed.

could be considered correct since the instrument would be traded "after hours" on the most up to date valuation of the instrument, for example on a national holiday. This will also ensure that the model is not dependent on points in the future. The sparsity pattern of the completely filled prices are seen in Figure 13. The price series are now complete and only start and end values are missing.

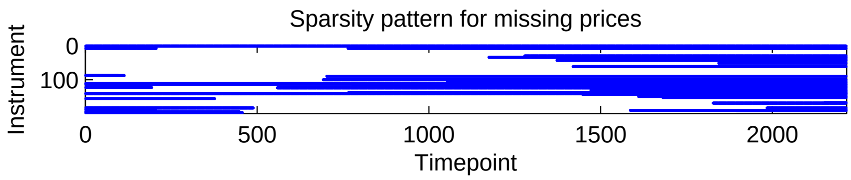

Figure 13: Plot of the sparsity pattern of the missing prices in the data set. Holidays removed. Missing prices filled with the most recent available price.

Returns

The prices are to be used to calculate the returns as the evolution from time t_{k-1} to t_k. The returns are calculated as:

$$r_{i,t_k} = \frac{p^*_{i,t_k}}{p^*_{i,t_{k-1}}}$$

The nature of it is that if price data up until, and including, time t_k is missing for instrument X_i, then the first return possible to calculate is $r_{i,t_{k+2}}$.

The returns are in a latter stage, after filtering, calculated as monthly returns, with the date stamps being the last trade date of each month, $(t_{m,1}, t_{m,2}, ..., t_{m,T_m})$. This is done by simply taking the product of each day's returns, as:

$$r_{i,t_{m,l}} = \prod_{t_{m,l-1} < t_k \leq t_{m,l}} r_{i,t_k}$$

The observant reader soon realizes that this is more easily calculated as:

$$r_{i,t_{m,l}} = \prod_{t_{m,l-1}<t_k\leq t_{m,l}} r_{i,t_k} = \prod_{t_{m,l-1}<t_k\leq t_{m,l}} \frac{p^*_{i,t_k}}{p^*_{i,t_{k-1}}} = \frac{p^*_{i,t_{m,l-1}}}{p^*_{i,t_{m,l}}}$$

This is, however, not possible when the daily return data has been filtered for splits and abnormalities. Adjusting the prices after the return data has been filtered is a more fragile task than calculating monthly returns from daily returns.

Filtering

The data set considered has been adjusted for splits and dividends, hence there should not be any such faults in need of correction. This was tested with a filter that looks for returns within $\epsilon = 0.05$ distance from any of the elements of a set of typical split and inverted split multipliers, e.g. $(0.25, 0.5, 0.75, \ldots, 10, 1/3, 1/6, \ldots)$. For returns within the interval, the return is divided with the matching element. This filter has proven to have great performance when tested on previous sets of data. With the data used in the final version, no points where filtered out, hence the data is already adjusted properly. There are still some outliers left, all daily returns corresponding to a 50% up or down movement are considered wrong and replaced by 1 (”no movement”).

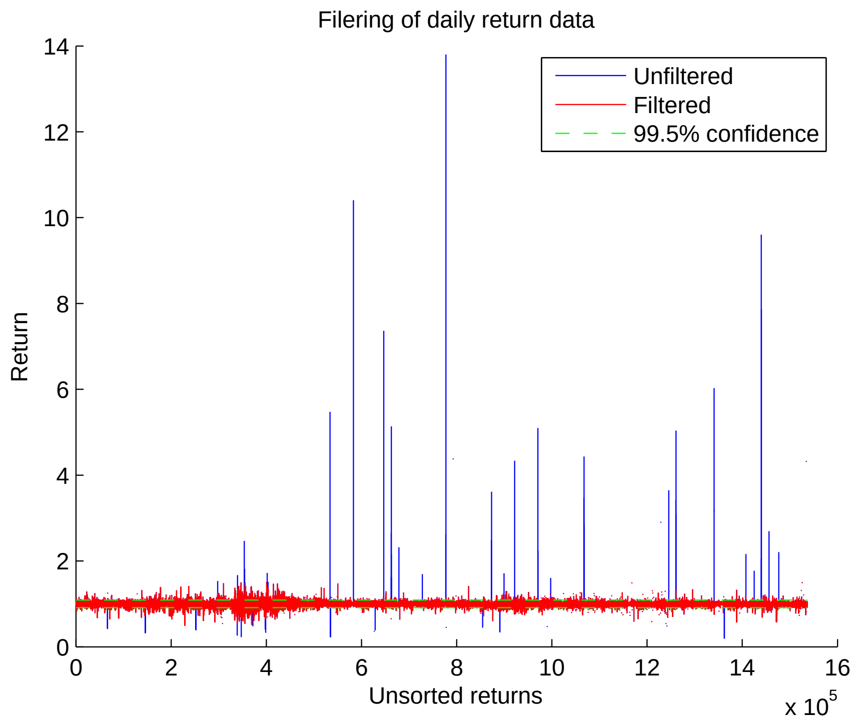

Figure 14: Plot of the unsorted returns, showing data before and after filtering, along with 99.5% confidence interval.

Index indicator

The universe of data, U_D, is far bigger than that of the index, U_I, but they do overlap in all, but a few points. The intersection, $U = U_D \cap U_I$, is the universe in which all points are preceded with at least a year of data and the instrument is in the index. This is motivated by two things; this study assumes that a year of historical data serves well for estimation of the stochastic parameters, and an instrument must be in the index to be allowed in the portfolio.

The sparsity pattern of the universe U is visualized in Figure 15 in the same fashion as before, instruments as rows and dates as columns.

The universe U is represented in the practical applications by a matrix, I_U, of

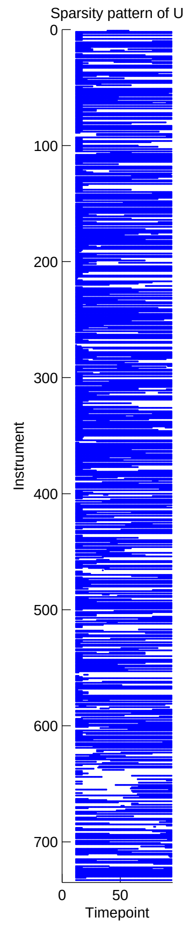

Figure 15: The sparsity of the total investable universe represented by the matrix I_U. Note the much shorter time axis than the section's previous plots, as this is monthly data rather than daily.

ones and zeros, an indicator to when an instrument is allowed in the portfolio. Building this eases the management of the portfolio and its parameters, allowing for a shorter and more dynamic code.

4 Results

All results are measured on the period 1 January 2007 to 31 October 2014. The graphs and tables are produced using Matlab where the methods and algorithms described in previous sections have been implemented.

Recap
Eight strategies have all in all been investigated; a capitalization weighted (CW), where the portfolio closest to one with the weights set linearly proportional to the assets' capitalization divided by the universes capitalization is selected. An equal weighted strategy (EQ), with a portfolio as close to all assets having the same weight is selected. One strategy with maximized future Sharpe ratio (SR), where the weights are the solution to an optimization program maximizing the expected future Sharpe ratio one time step ahead. One strategy with maximized diversification ratio (DR), where the present time diversification ratio is maximized. And one fundamentally weighted strategy using P/E-ratios (PE), where the weights are the squared E/P-ratios, normalized to sum to 1. The latter three strategies have been tried with a constraint on rebalancing costs as well (SRc, DRc, and PEc), where a limit to the level of rebalancing at each time point is set. The portfolio weights have been calculated via optimization with a set of constraints keeping them UCITS compliant.

A universe of the historical constituents of the MSCI Europe develped index is used. Data of return series, average daily volumes, market capitalization etc. was collected and cleansed of abnormalities. Estimations of the covariance matrices at each point in time are carried out via historical simulation, while the expected future returns are based on averages, both use a window of the preceding one year historical data.
End recap

Measures such as: Value at Risk, Expected Shortfall, Sharpe ratio, trading costs, market impact, and distribution between asset capitalizations have been calculated and are presented in this section.

The statistics associated with each of the 8 strategies tested are found in Table 4 below. The averages are calculated as the arithmetic mean of all non `NaN` values in the sequences. The annual returns are displayed in Figure 16 below, with winners and losers highlighted.

The performance without adjustment for trading costs for the complete period is plotted in Figure 17. All 8 strategies are included. The cap weighted benchmark is visualized in green and is found as the third from below.

Figure 18 represents the arithmetic value added by choosing a strategy over CW. In other words, the performance of the cap weighted portfolio subtracted from each other portfolio. The values are represented as percentages of the initial portfolio size of one billion Euro.

The Sharpe ratios for the 8 strategies are plotted in Figure 19.

	CW	EQ	SR	SRc	DR	DRc	PE	PEc
Return %	-11.9	-6.50	14.5	2.70	-19.1	-23.5	14.6	16.8
Trading Cost million Euro	3.0	11	140	61	69	50	27	31
R - C %	-11.9	-6.6	13.2	2.1	-19.8	-24.0	14.3	16.5
Annual return %	-1.6	-0.9	1.7	0.3	-2.7	-3.4	1.8	2.0
-:-post 09-1-30 %	11.7	9.6	13.3	13.6	5.5	5.2	11.6	11.4
VaR %	9.2	9.8	6.5	6.8	4.8	5.3	9.1	9.1
ES %	11.8	12.5	8.8	9.2	6.0	6.8	11.8	11.6
Sharpe ratio average	-0.047	-0.032	-0.012	-0.023	-0.072	-0.074	-0.011	-0.009
Cap avg. billion Euro	42	11	12	12	8.2	8.8	13	13
Liq. Days	29.7	184	897	997	1750	1650	857	933
MI II Perm.	0.14	1.0	2.6	2.0	1.7	1.6	0.98	1.1
MI II Temp.	0.003	0.20	7.0	3.2	4.5	2.7	0.41	0.97

Table 4: Table showing: total return over the period, total trading costs in million Euro (10^6 Euro) associated with managing a 10 billion Euro ($10 * 10^9$ Euro) portfolio, the total return after subtraction of trading cost, the annualized return for the complete period and after 2009-1-30, the Value at Risk and Expected Shortfall as a percentage of the portfolio, average Sharpe ratio, average mean capitalization of assets owned in billion Euro (10^9 Euro), average days to sell complete portfolio with 10% market participation (MI I), total permanent, and average temporary market impact via JP Morgan's model (MI II). For the strategies: cap weighted, equally weighted, maximum Sharpe ratio weighted, maximum Sharpe ratio weighted with rebalancing constraint, diversified, diversified with rebalancing constraint, P/E weighted, and P/E weighted with rebalancing constraint.

Annual return %	2008	2009	2010	2011	2012	2013	2014
CW	-46.03%	27.15%	7.90%	-10.87%	13.24%	16.14%	1.67%
EQ	-49.03%	40.25%	12.90%	-17.92%	16.76%	18.91%	1.86%
SR	-37.26%	26.71%	17.14%	-14.55%	15.98%	20.08%	3.68%
SRc	-43.34%	29.46%	9.41%	-11.12%	15.32%	21.73%	2.84%
DR	-42.60%	11.37%	19.14%	-14.87%	2.73%	18.77%	2.44%
DRc	-44.64%	9.91%	15.79%	-13.81%	2.51%	22.08%	0.69%
PE	-41.60%	41.69%	18.84%	-12.21%	14.86%	21.18%	-5.05%
PEc	-41.60%	43.34%	19.73%	-12.25%	14.87%	21.18%	-5.04%

Figure 16: Return per year for each strategy. Each year's winner (green) and loser (red) is marked. Since the evaluation period ends at 2014-10-31 the returns for 2014 have been scaled according to the squared time method, in order to ease comparison.

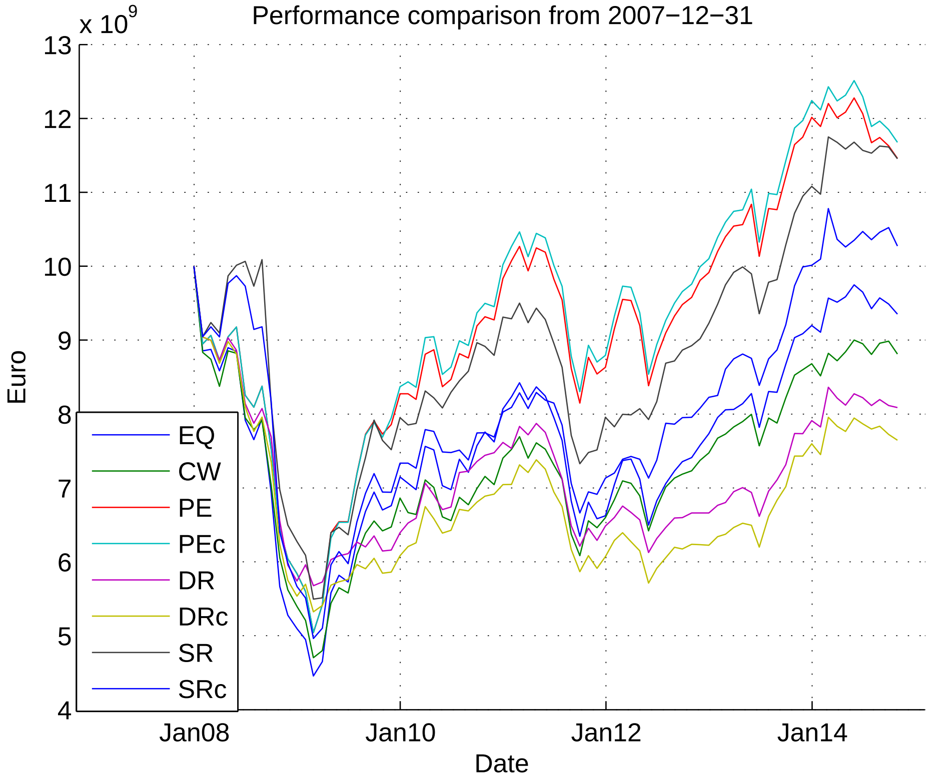

Figure 17: Performance measured as portfolio value at different points in time for all 8 strategies.

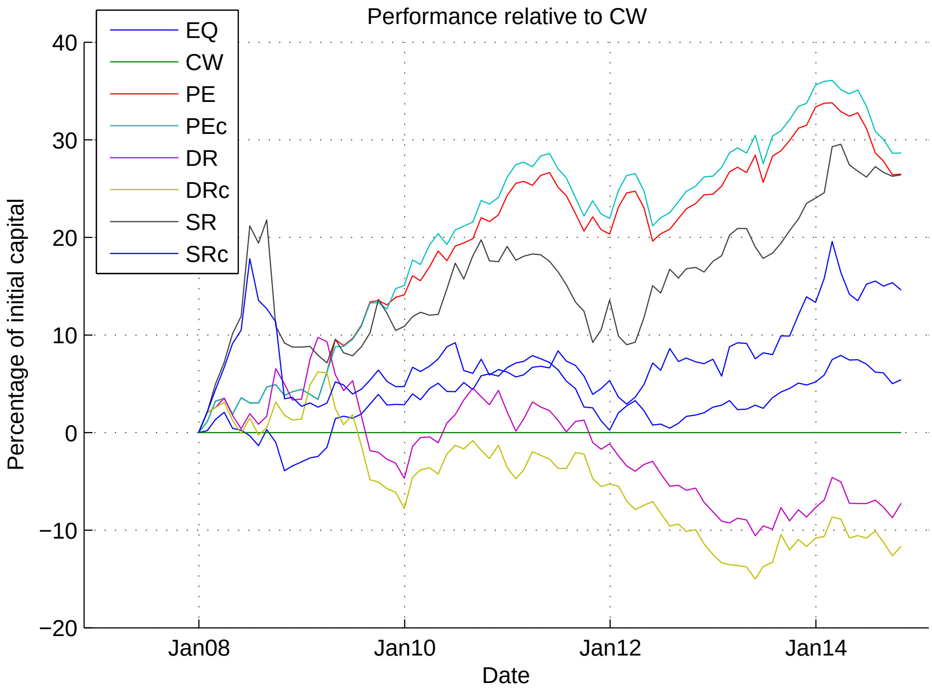

Figure 18: Performance measured as the arithmetic value added by choosing a strategy over the cap weighted. The y-axis is the value added as a percentage of the initial fund capital of one billion Euro.

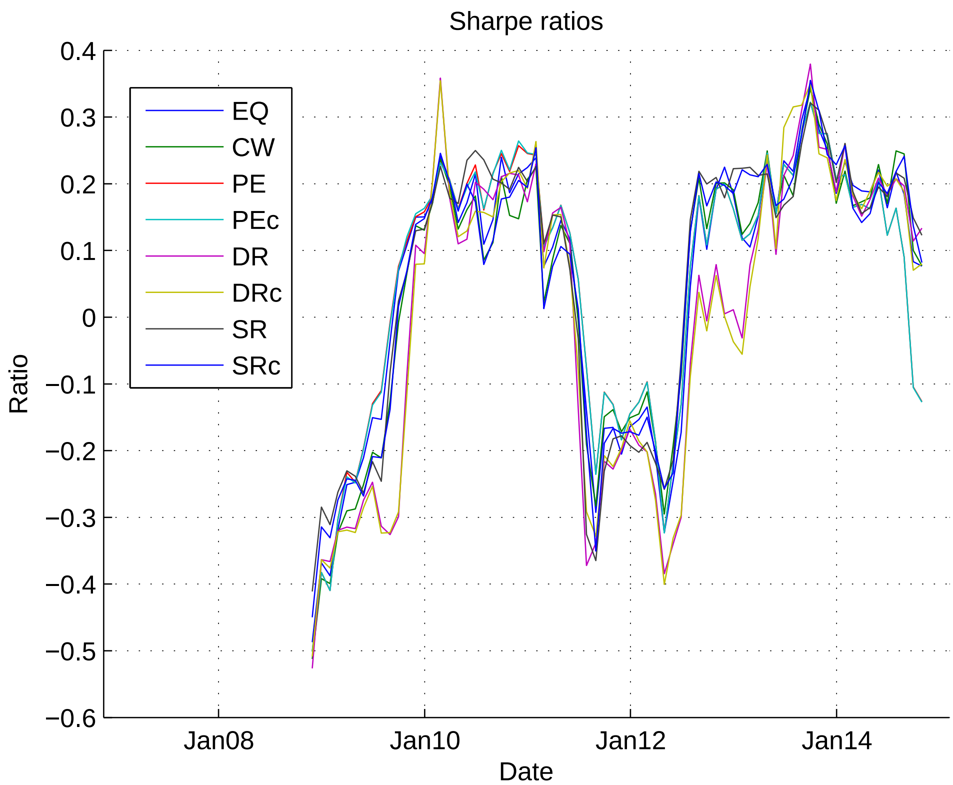

Figure 19: Comparison of actual Sharpe ratios for the different strategies.

All results concerning performance, trading costs, risk, weights distributions, Sharpe ratios, and market impacts for the different strategies are found in Appendix I - Graphs of simulation results, which is also the figures to which the analysis refers.

5 Analysis

5.1 Overall

From Figure 17 the clearly best performing strategy is the fundamentally weighted (PE and PEc), with no obviously larger losses in the market's down movement, but a significant recovery in the latter up movement after 2009. The Sharpe ratio maximizers (SR and SRc) and the equally weighted (EQ) beat the market, but without any signs of great outperformance. The strategy having a hard time keeping up is the most diversified (DR and DRc), performing worse than the market after 2009. Although, close inspection reveals a slightly more controlled drop at the downward movements in 2008 and 2011.

Looking at Figure 19, all strategies' Sharpe ratios follow the same pattern, with small signs of the SR and SRc strategies being among the higher at most times, as predicted by the hypothesis in Section 2.13. The downward spikes can be identified as downward movements in the market, as seen in Figure 17.

The test setting (2008-01-01—2014-10-31); choosing this market might seem odd when proving the strategies' outperformance, since there is a clear downward movement beginning the series which will weaken the performance. However, this thesis is not a sales pitch for an individual strategy, but one aiming at finding weaknesses and strengths in them. In that point of view, the market served superbly.

There are signs of cyclical trends in the market, with drops in late 2008 and late 2011, and peaks at late 2010 and early 2014, as seen in Figure 17. Looking at Figure 18 there is but some evidence of this cyclicality left. The PE and PEc strategies have outperformed on a stable basis over the time period, but the higher volatility and sensitivity to market movements is evident; however this seem to have had little affect during the 2008 drop. The SR and SRc strategies have a similar behavior; even though the lower VaR and ES, the higher return per unit risk has its affects during market drops, as seen in the lowered outperformance. In the very beginning of the sequence, a spike of 20% outperformance by SR is evident, showing that the higher expected return per unit risk might be paying of in bearish times, a drop is however following as the downfall was present throughout the market. This could on the other hand be a strike of pure luck in stock selection. EQ shows the same behavior, but at a level closer to CW. DR and DRc shows a behavior of having a hard time to pick up the pace as the market swings, a type of moment of inertia, being a result of diversification.

5.2 Strategies' performance

Equally weighted

One of the key assumptions (or requirements) for the EQ scheme to work is that the unweighted mean of returns over the investigated period is greater than one (positive return), this is however not the case in the selected market, with an average of 0.9988 return.

Looking at Figure 23 an outperformance of the cap weighted index is present, perhaps an arbitrary spread of the upward movements is the case, rather than an overweight towards the bigger assets. Hence the hypothesis in Section 2.12 cannot be discarded. The findings of De Miguel, Garlappi, and Uppal [5] stating that when $M < 3000$ if $N = 25$ and $M < 6000$ if $N = 50$, with M being number of months that parameter estimations are based on, the equally weighted portfolio should outperform those using the estimates, are proven wrong. In this setting, $M = 12$ at all times and $N \geq 370$ at all times, yet no outperformance. This is assumed to be a result of good parameter estimation.

Sharpe ratio maximizers

SR ans SRc have shown to give a higher level of return when the market is in an upward trend, while having a larger downward movement when the market is in a downward trend. In Figure 19, the visualization of the Sharpe ratios indicate much instability. This is the case when the expected return is smaller than the risk free rate, Definition 2.4 and Figure 9, the ratio will become negative. Along with this follows the problem of optimizing a portfolio using the Sharpe ratio as objective function. If the optimizer cannot find a set of assets with a higher expected return than the risk free rate, the maximizer will minimize the expected loss - seemingly good - while at the same time maximizing the portfolio volatility - not so good. This could explain the vast downward movements when the market is bearish, as the volatility of the selected portfolio and thereby also the movement is increased.
The theory in Section 7 states that the market portfolio - the, in some sense, optimal portfolio - lies along the capital market line. With the possible portfolio μ_p being lower than the uninvestable risk free rate, the theory of finding the tangent point would suggest the point with highest volatility for a set level of expected return, i.e. on the right side of the attainable region. This is precisely the problem of maximizing the Sharpe ratio at bearish times.
Table 4 shows that the average Value at Risk and Expected Shortfall for the Sharpe ratio maximizers are in the lower region, only beaten by the most diversified strategy. The hypothesis of outperformance at a lower level of risk is hence justified.

Most diversified

DR and DRc are those, according to Table 4, with lowest levels of risk, which is in line with the theory. It is also the worst performing, not in line with the hypothesis of outperformance as an effect of choosing assets with high individual volatility. It seems as if the effects of reducing the volatility of the total portfolio in a min vol fashion, Equation (2.25), are higher than the volatility sought via the denominator in Definition 2.5. Looking at just the bearish behavior of the market during 2008 and 2011, the diversified portfolios are both among those with the smallest drops (rougly 16 percentage units less than the equally weighted in total), Figure 17, giving credit to the diversification. This is in line with the slow relative growth during bullish times, Figure 18. The results do however contradict those of Choueifaty and Coignard [4] having shown that the most diversified outperforms in a similar universe.

Fundamentally weighted

Looking at Figures 17 and 18 it is clear that the PE and PEc portfolios outperform in the overall setting. Despite the risk levels in the upper region the strategy does not underperform significantly more than the other strategies during bearish times.
There are three possible scenarios when picking low P/E-ratio assets (excluding the scenarios of the ratio decreasing): 1) the ratio is increased via lower (perhaps expected,

hence the low ratio) earnings, 2) the ratio is increased due to the asset being undervalued and the price therefore increases, or 3) the ratio is stable due to equal movements in earnings and price. The first case yields no change in the portfolio value, but will lower the weight, the second increases the portfolio value while lowering the weight at the same rate, and the last increases the portfolio value while keeping the weight. Looking at the second scenario, this could be the explanation to the fact that the portfolio with restricted rebalancing yields a higher return, but increases the transaction costs Figures 38 and 41. Not rebaancing and keeping the winners will probably result in more growth, hence higher return, but also a larger need of rebalancing at a later stage.

5.3 Winner/loser

The fundamental PE and PEc strategies must be considered the winning strategies, as they give the highest total return with and without transaction costs taken into account, while at the same time keeping the risk measures fairly low. The downside is the need of fundamental asset data, which in some cases may be hard to come by.

The Sharpe ratio maximizers, SR and SRc, are strong performing, with the same attributes as the fundamental strategies. They do, however, not yield the same stable outperformance of the cap weighted strategy and suffer the need of clean high quality return data.

The equally weighted strategy performs at the same level as the cap weighted, but with a larger need of rebalancing and therefore trading costs and management costs. A reasonable choice to the cap weighted.

The evident losers are the most diversified, DR and DRc, with worst performance and mid region costs. The plus side is the stability and low risk, but without performance it gives little consolation to the investor.

5.4 Measures

This section discusses the measures visualized for each strategy in Appendix I; rebalancing requirement and accumulated trading costs, Value at Risk and Expected Shortfall, a selection of weight statistics, the real Sharpe ratio, distribution between asset capitalization, and the market impact measures MI I and MI II.

Trading cost

Naturally, since there is no need for rebalancing due to weight changes, only changes in the index, the cap weighted portfolio will have the lowest trading costs. There is also evidence that the rebalancing restriction on the Sharpe ratio maximizer lowered the costs substantially, with a factor 2, when comparing the results for SR and SRc in Table 4.

In all three cases of the strategies with different trading volumes, the one with higher cost performed better compared to the version with lower trading costs of the same strategy.

There are no evident correlation between the different strategies' trading frequencies over time and the return series. The two possible explanations are that the asset picking might as well be made randomly and the selection will have no affect on return, or that the strategies function stable during both bearish and bullish times, picking well performing assets. The latter is a more reasonable explanation; test portfolios of randomly selected assets in the investigated universe have resulted in almost 100% loss of the initial capital.

As described in Section 5.2, the total accumulated trading cost increased when a restriction on rebalancing was included in the fundamental strategy (PE vs. PEc), different to the maximal Sharpe ratio SRc or most diversified DRc portfolios. An explanation is that the fundamental strategy is designed to perform on a longer horizon than one month, a larger inertia an thus less capability of rebalancing makes for keeping of winners longer. As they are expected to perform over a longer time period, the previous winners are expected to continue winning and therefore shifting the weights even further from the "optimal" according to the PEc strategy, yielding even larger need of rebalancing over time. Whereas in the case of SRc and DRc, rather than selling an asset one day and buying it back the next, a more stable portfolio is held. But with a one month horizon this is not the optimal portfolio.
A simpler explanation might be that the PEc strategy is picking assets with a larger spread in return, and hence less rebalancing increases this spread and thereby the need of rebalancing over time. Whereas the DRc and SRc strategies pick more similar assets in terms of return, hence keeping the spread and need of rebalancing low over time.

Risk
The strategies CW, EQ, PE, and PEc show an increase after the downward movement of 2011, Figures 20, 23, 38, and 41, in their Value at Risk and Expected Shortfall measures, although having similar patterns. This indicates SR and SRc as well as DR and DRc taking the past movement of the assets into account, actively selecting others than those drawing down the market. Hence not showing as large increases in the risk post 2011. With this said, having the resent past as predictor of the near future might not be optimal, considering the SR and DR strategies' vastly different performance; something for the investor to keep in mind in terms of investment horizon.

All strategies show a big increase in risk after 2008. As the VaR and ES measures are based on historical volatility, this validates that the whole market suffered from the recess caused by the subprime crisis.

There is no significant changes in spread between the two measures in any of the eight cases, but there *is* still a spread. This indicates that there is some tail risk involved in all eight portfolios, but it is stable over time. This is because of the very nature of the measures; Value at Risk takes the one selected probability level and in a sense disregards all less probable events. Whereas Expected Shortfall weights all events less probable than the selected level. The latter hence has a greater capability of detecting heavy tails.

The similar behavior in the spread between the two risk measures shown in all strategies could be an indication of the assets' return distrubutions being rather similar. EQ weights smaller capitalized assets much higher than CW, and the other strategies have weightings in between, but still showing a similar spread between VaR and ES, thus a similar total return distribution regardless of the asset selection. This strengthens the belief that a common distribution for all assets may be used.

Weight statistics
The weights statistics shows that the most diversified, DR, strategy violates the condition of owning a maximum of 20% of an asset at two times, Figure 33, and the rebalancing restricted fundamentally weighted, PEc, does so at a couple of times, even breaking the 5-10-40 rule, Figure 42. This is odd, since the DR strategy should aim at diversifying the portfolio as much as possible at all points of rebalancing. Investigation points towards this being an optimization fault that perhaps could have been avoided if another initial point had been used. The same applies to the violations in the PEc strategy.

The weight violations are violations of the UCITS rules, Definition 2.2, and would hence make those portfolios non-compliant at these points in time. A real life application of these strategies would be managed and supervised by a real person, thus avoiding such violations.

No violations making the portfolios impossible are detected, such as owning more than 100% of an asset.

Generally, the "optimized" portfolios, all excluding CW and EQ, tend to overweight one - or a group of - asset, seen in the weights being closer to the limit set by the UCITS rules. As the risk is not increased as vastly, Table 4, in these strategies, one may assume that these bets on single assets are fair and justified.

It may also be said that the CW and EQ portfolios may only face restriction by the size of the assets, due to their nature.

Average asset capitalization
All strategies but CW have similar average asset sizes of around 10 billion Euro (10^9 Euro) in capitalization, with a standard deviation of around 25 billion Euro, while the cap weighted has an average asset size and standard deviation of 40 million Euro. The strategies hence follows more closely to the equally weighted portfolio by selecting assets according to other parameters than the market capitalization, just in line with the definition of a smart beta strategy, Definition 1.1. The large standard deviation in comparison to the mean is a result of a very flat distribution.

It seems that the restriction on rebalancing forces the strategy to pick assets of more similar size at the different times, compare Figures 27 and 30, for SR and SRc respectively.

Comparing EQ and CW, where the main difference is the latter's relative overweight in big assets, Figure 23 gives reason to believe that the smaller assets are more volatile than the bigger. Consider the period of outperformance by the equally weighted portfolio from the beginning of 2009 until the beginning of 2011, the faster drop thereafter, and the slight outperformance from late 2011. The initial drop during the 2008 recession was throughout the market, though, and hence not considered. Connecting the relationship between volatility and return as in the CAPM-theory, Section 2.11, the slight outperformance by the equally weighted portfolio, EQ, is justified.

Market impact
The difference in asset capitalization is evident in the permanent and temporary market impacts caused by rebalancing, compare CW, Figure 22, and EQ, Figure 25. Where trading with a smaller percentage of the total number of stock outstanding of an asset is less noticed.

Comparing the maximum number of days to sell an asset to the average number of days to sell an asset, in all strategies, there is reason to believe that this is caused by one, or a couple of, illiquid assets. Taking this into account when using such a strategy should therefore not make for a large difference than these theoretical findings.

The latter strategies are choosing less liquid portfolios than the CW and EQ strategies. Making them more demanding for the manager in terms of planning sell offs and finding good deals. In reality, this would also impact the price of the assets when interacting with the market, most likely cutting some of the profit. The less liquid portfolios do also hold a larger risk for when the market turns bearish and is perhaps more suited for an investor with a longer horizon, but most importantly, one who can take the punch.

It is evident that the spikes in the "Market impact I" graphs are causing the temporary impacts in the "Market impact II" graphs.

5.5 Possible combination

It is clear that the Sharpe ratio maximizer has succeeded in finding portfolios with high return, while at the same time keeping the risk at a low level. The fundamentally weighted portfolios have performed at a higher rate of return than the Sharpe ratio maximizers, but with a higher level of risk. What if the two could be combined? The easy way would be a linear combination of the two strategies' portfolios, but that would only average the measures. A more sophisticated way could be to use the P/E-ratios to determine the expected return for the Sharpe ratio objective function. Precisely this has been done, and the results are visualized in Figures 44 - 46. At each time the instrument with lowest P/E-ratio has been given the expected return 1.05 and the instrument with highest P/E-ratio an expected return of 0.95. A line has been fitted in between the two and the other instruments could therefrom be given expected returns. A very simple model, with only reasoning as motivation. The averaged Value at Risk is 3.74% and Expected Shortfall is 4.85% for the whole period. The annualized return is 2.8%. This portfolio is hence the best performing, yet having the lowest risk. The strategy show no significant increase in market impact, but violates the UCITS rules at a couple of points.

6 Conclusion

The sorting and filtering procedures for price- and other data were successful. The, sometimes, harsh task of "cleaning" data can be taught to a machine and implemented as algorithms, as shown in Section 3.2.

The distribution plots in Section 2.2 indicate that the historical simulation is valid and the general behavior of the Sharpe ratio and diversification maximizers, indicate a valid approximation of the covariance matrices. A poor estimation of the covariance matrices would make the optimization useless and hence the choice of weights random. The latter, random weight selection, was tried on the dataset and the resulting portfolios did not only perform worse than the cap weighted portfolio, but did in many cases tend to zero in portfolio value. Thus outperformance and stability is a sign of valid approximation.

It is possible to construct an index that follows the market movements, with some tracking error and outperformance, in a systematic way independent of the absolute price or market capitalization of the assets. In all seven cases of portfolios weighted differently than the cap weighted portfolio, the portfolio followed the market movement (the cap weighted portfolio's movement) with some tracking error. As an index portfolio, or a beta portfolio, is to serve as the base in an investment with a sector or region as benchmark, all seven portfolios tested qualifies, see Figure 17.

Regardless of strategy, there is a vast increase in market impact and accumulated trading cost in comparison with the cap weighted portfolio. Naturally, rebalancing a portfolio according to a scheme will generate higher costs than not doing so, but the main issue lies not in the costs themselves, but rather in the gained performance. The vast outperformance in comparison to added costs (see Table 4) in the fundamentally weighted and maximized Sharpe ratio portfolios strengthen the belief of gain from active management, even in a systematic beta fashion.

Of the strategies tested, all but one (equally weighted) yields a lower level of risk in average, taking both measures into account. This is likely a result from poor diversification, selecting more volatile - lesser capitalized - assets without consideration of total portfolio risk.

The suggested strategy is the Sharpe ratio maximizer, but with a smarter way of estimating expected returns. Preferably via the squared inverted P/E-ratios, as this has proven to find future winners, although the strategy presented in Section 5.5 would serve well. P/E-ratios are a better way of determining the future performance of assets, not in terms of absolute numbers, but rather as the expected trends' directions, compared to using historical returns. History can, however, give an indication of the volatility or risk of the assets and may hence be used to optimize the risk adjusted return. The combination of the two makes a splendid pair of input to an asset selection model giving high return at a low level of risk, see Section 5.5.

The assumption made in Section 2.3 of dependence in space, but not in time for the assets somewhat disqualifies the method of using historical data to estimate future return. This could be an explanation for the combined method giving a more sound estimation for the optimization and hence having a higher outperformance.

Bibliography

[1] N. Amenc, F. Goltz and L. Martellini, *"Smart Beta 2.0,"* EDHEC-Risk Institute, Nice, France, 2013.

[2] R. Arnott and E. Kose (2014, Aug.). *What "Smart Beta" Means to Us* [Online]. Available: http://www.researchaffiliates.com/Our%20Ideas/Insights/Fundamentals/Pages/292_What_Smart_Beta_Means_to_Us.aspx

[3] D. Booth and E. Fama, "Diversification Returns and Asset Contributions," *Financial Analyst Journal* , vol. 48, no. 3, pp. 26-32, May. - June, 1992.

[4] Y. Choueifaty and Y. Coignard, "Toward Maximum Diversification," *The Journal of Portfolio Management* , pp. 40-51, Fall 2008.

[5] V. DeMiguel, L. Garlappi and R. Uppal, "Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?," *The Review of Financial Studies* , vol. 22, no. 5, pp. 1915-1953, May, 2009.

[6] ETF.com. (2014, June 2). *How To Run An Index Fund: Full Replication Vs. Optimization* [Online]. Available: http://www.etf.com/etf-education-center/21038-how-to-run-an-index-fund-full-replication-vs-optimization.html

[7] E. Faerber, "Common-Stock Price Ratios," in *All About Value Investing* 1st ed. 2013, ch. 8

[8] A. Ferraris, "Equity Market Impact Models," in *Mathematics at the interface between business and research Stifterverband fÂ¨ur die Deutsche Wissenschaft* , Berlin, Germany, 2008, pp. 14-19.

[9] C. Harvey. (2011). *Closing price* [Online]. Available: http://www.nasdaq.com/investing/glossary/c/closing-price#ixzz3M3IkFOyQ

[10] H. Hult *et.al.* , "Risk Measurement Principles," in *Risk and Portfolio Analysis* New York 2012, ch. 6

[11] J. Kienitz and D. Wetterau, "Constrained Optimization," in *Financial Modelling: Theory, Implementation and Practice with MATLAB Source* Chichester, West Sussex, United Kingdom: 2012, ch. 11, sec. 5.4 Optimization.

[12] L. Kopman and S. Liu, *"Maximizing the Sharpe Ratio and Information Ratio in the Barra Optimizer,"* MSCI Barra Research, 2009.

[13] E. Kose and M. Moroz (2014, May). *The High Cost of Equal Weighting* [Online]. Available: http://www.researchaffiliates.com/Our%20Ideas/Insights/Fundamentals/Pages/229_The_High_Cost_of_Equal_Weighting.aspx

[14] R. Larson. (2014, June). *I'd Choose Emerging Markets, Wouldn't You?* [Online]. Available: http://www.researchaffiliates.com/Our%20Ideas/Insights/Fundamentals/Pages/211_I_Would_Choose_Emerging_Markets_Wouldnt_You.aspx

[15] MathWorks. (2014). *Choosing a Solver* [Online]. Available: http://se.mathworks.com/help/optim/ug/choosing-a-solver.html

[16] MathWorks. (2014). *Constrained Nonlinear Optimization Algorithms* [Online]. Available: http://se.mathworks.com/help/optim/ug/constrained-nonlinear-optimization-algorithms.html

[17] MathWorks. (2014). *fitdist* [Online]. Available: http://se.mathworks.com/help/stats/fitdist.html

[18] MathWorks. (2014). *Writing Constraints* [Online]. Available: http://se.mathworks.com/help/optim/ug/writing-constraints.html

[19] H. Markowitz, ”portfolio Selection,” *The Journal of Finance* , vol. 7, no. 1, pp. 77-91, Mar., 1952.

[20] Ossiam, ”Ossiam ETF,” Stockholm, Sweden, 2014.

[21] A. Pettersen, ”An Investment Strategy Based on P/E ratios.” Bachelor thesis, Ekonomihögskolan, Lunds universitet, Lund, Sweden 2011.

[22] W. Sharpe, ”The Sharpe Ratio,” *The Journal of Portfolio Management* , Fall 1994.

[23] The Investment Association. *Investment Limits* [Online]. Available: http://www.investinginfunds.org/facts-about-funds/how-uk-funds-are-regulated/investment-limits-what-fund-managers-can-invest-in.html

[24] P. Tindall, *”Understanding smart beta,”* Towers Watson, Westminister, London, 2013.

[25] Value Weighted Index. *Equally Weighted Index* [Online]. Available: http://valueweightedindex.com/IndexComparison/EquallyWeighted/

[26] P. Wilmott ”What is Modern Portfolio Theory?,” in *Frequently Asket Questions in Quantitative Finance* , 2nd ed. Chichester, West Sussex, United Kingdom: 2009, ch. 2

[27] V. Zakamulin, ”Sharpe (Ratio) Thinking about the Investment Opportunity Set and CAPM Relationship,” *Economics Research International* , vol. 2011, Article ID 781760, May, 2011.

Appendix I - Graphs of simulation results

Overall

Cap benchmark

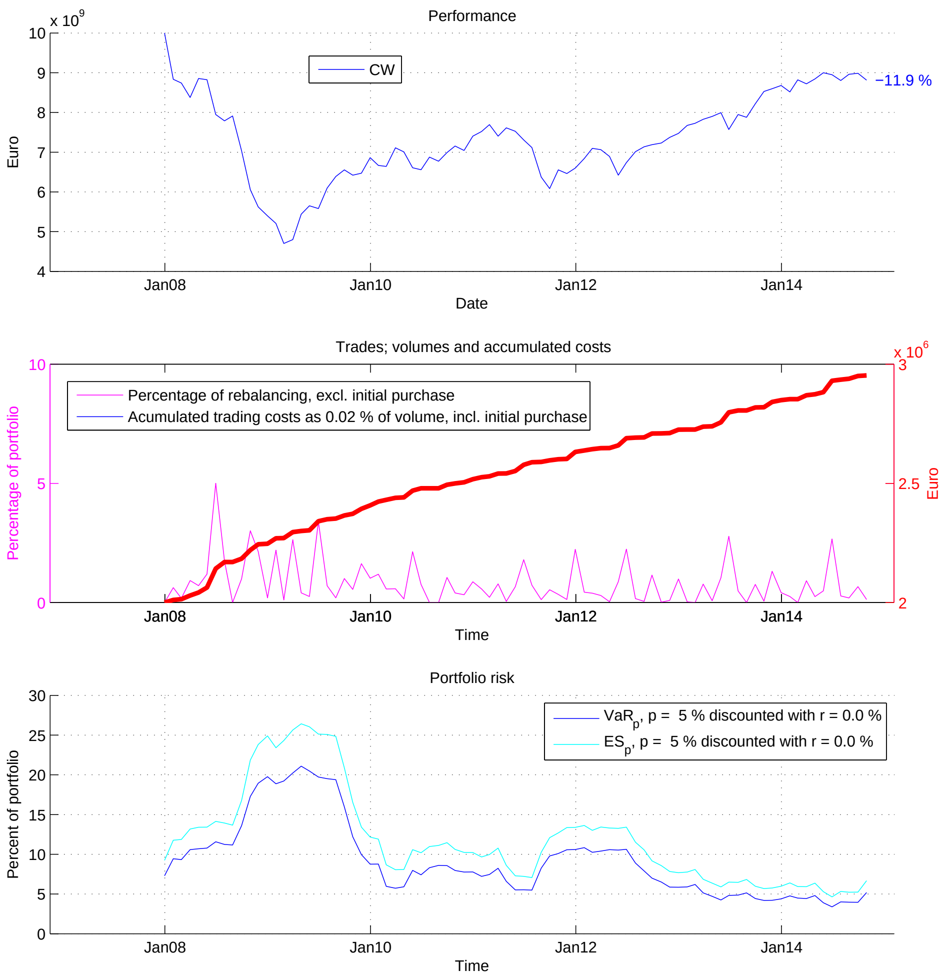

Figure 20: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

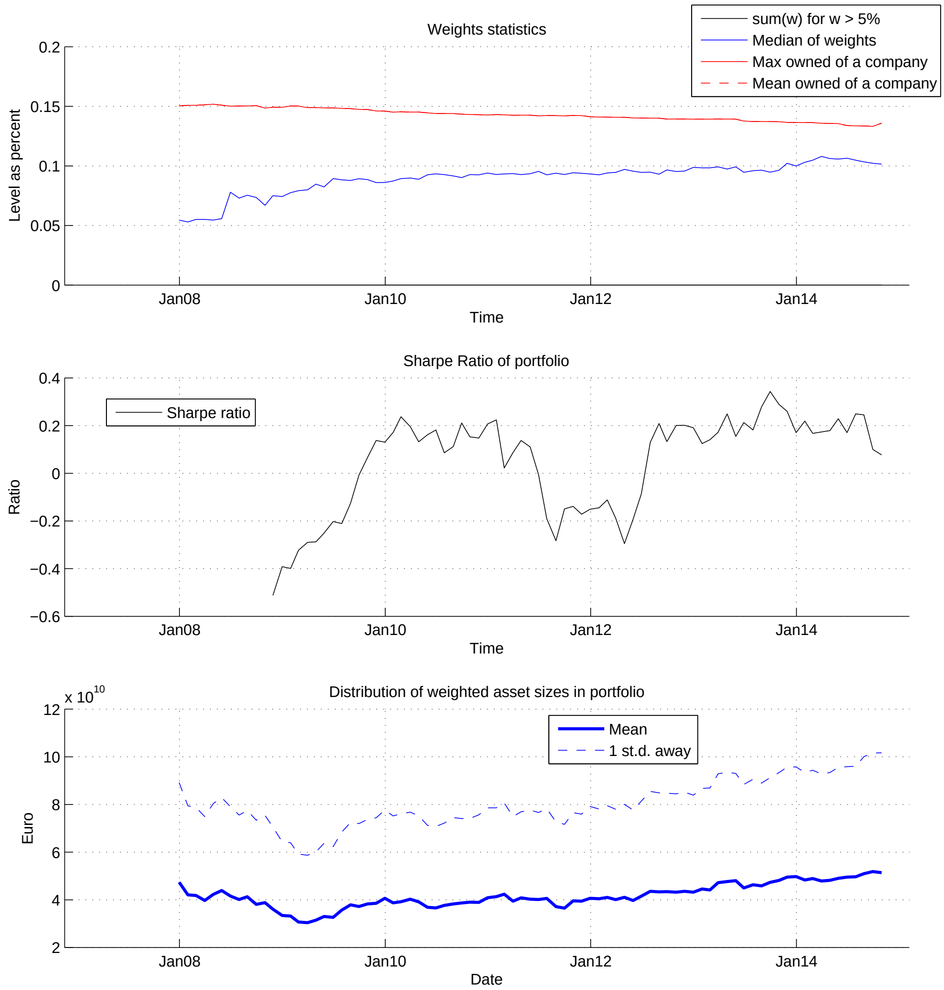

Figure 21: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

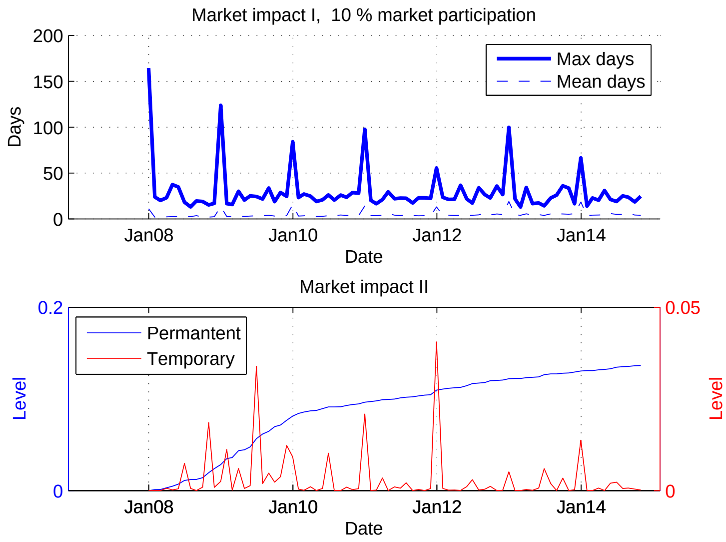

Figure 22: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

Equal weighting

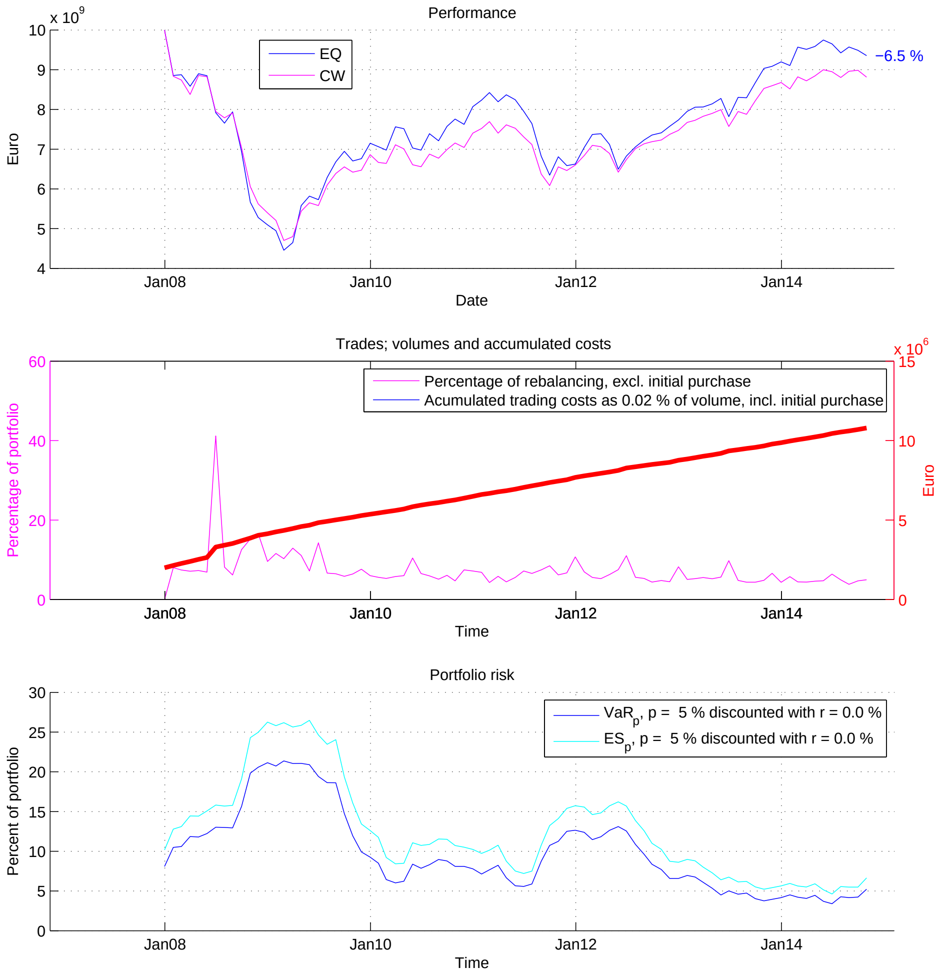

Figure 23: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

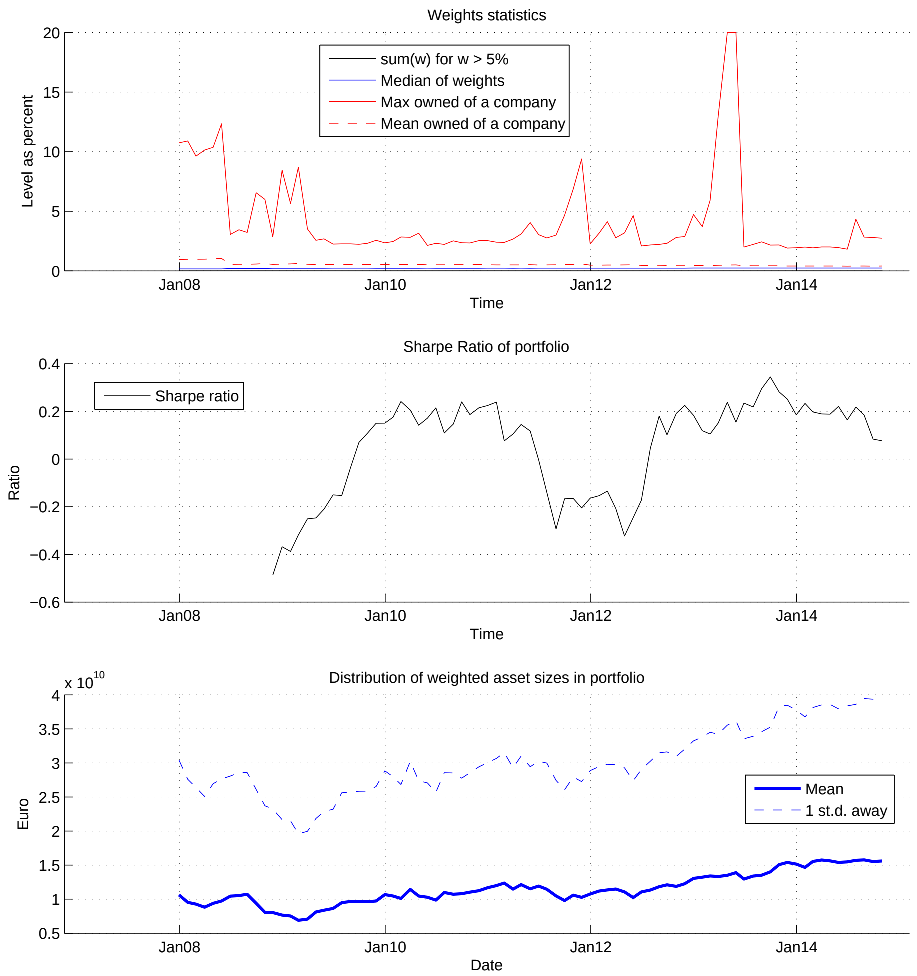

Figure 24: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

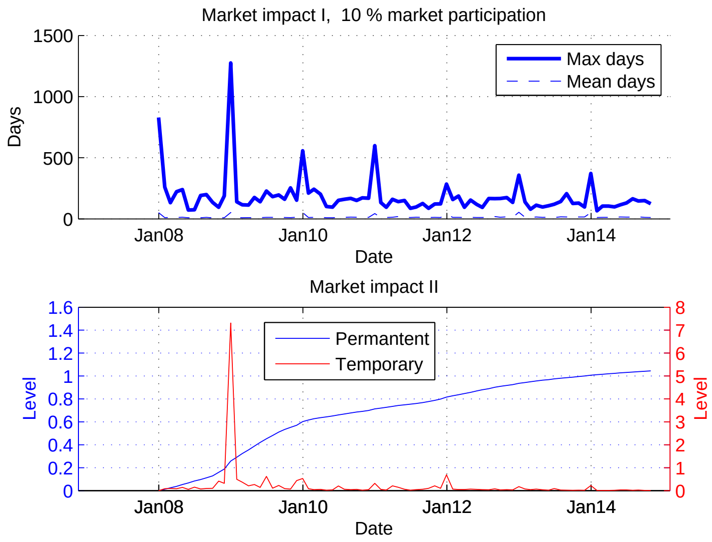

Figure 25: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

Weighted for Maximum Sharpe ratio

Ordinary

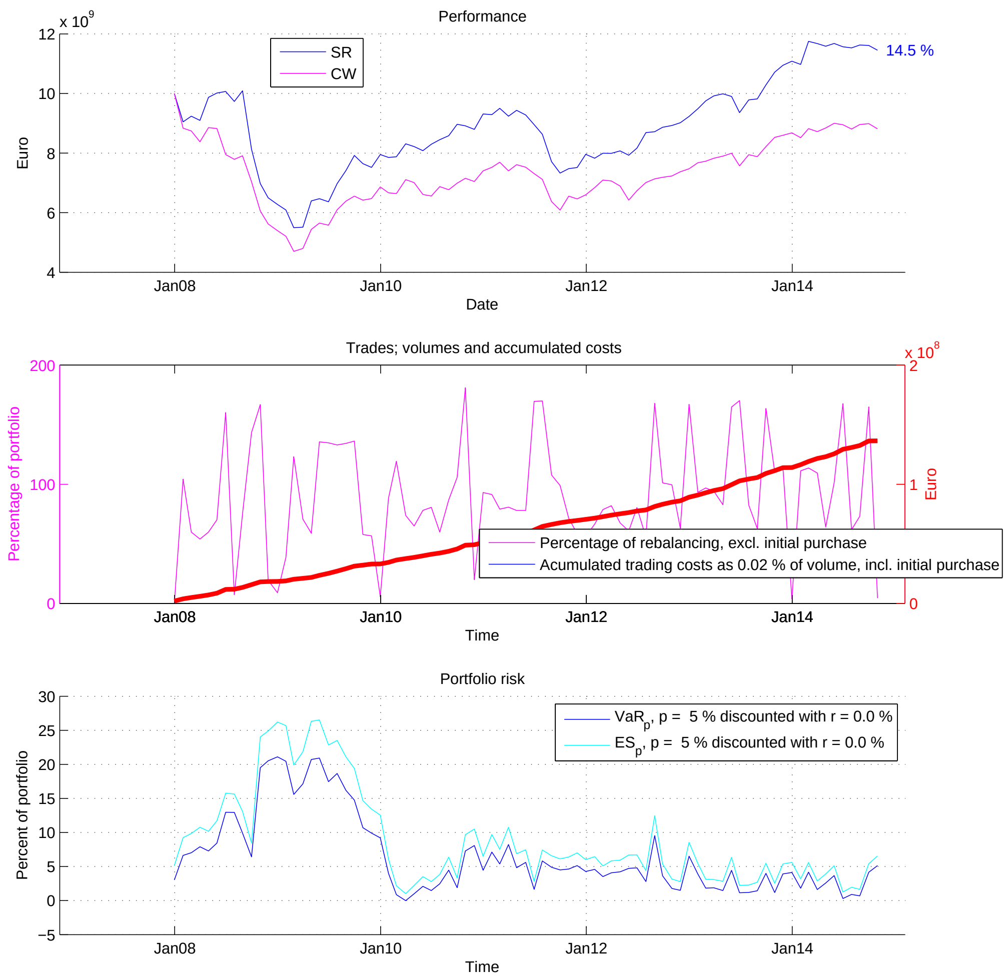

Figure 26: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

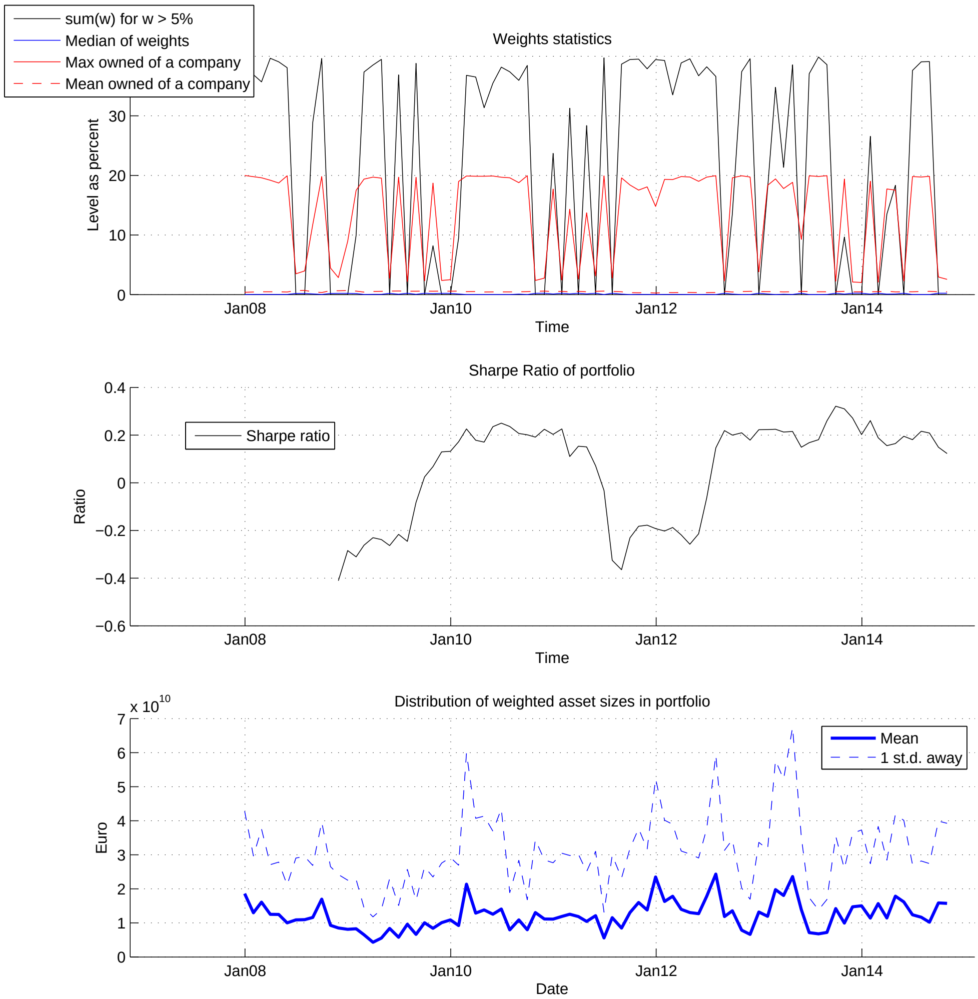

Figure 27: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

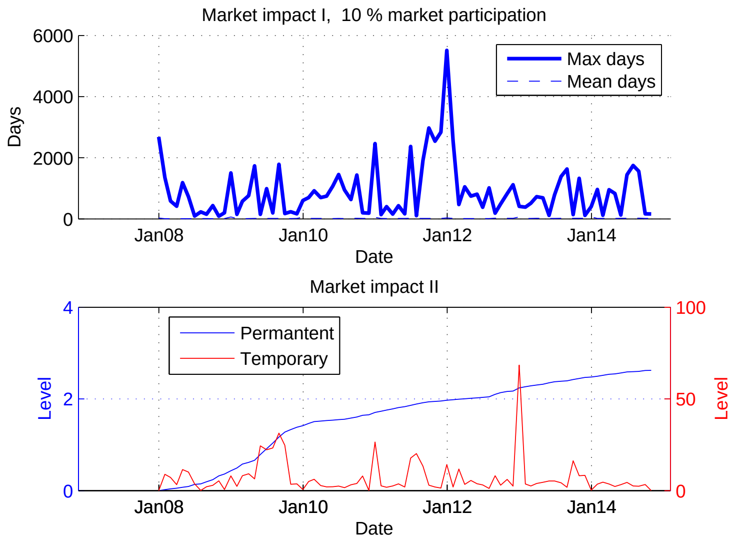

Figure 28: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

With rebalancing constraint

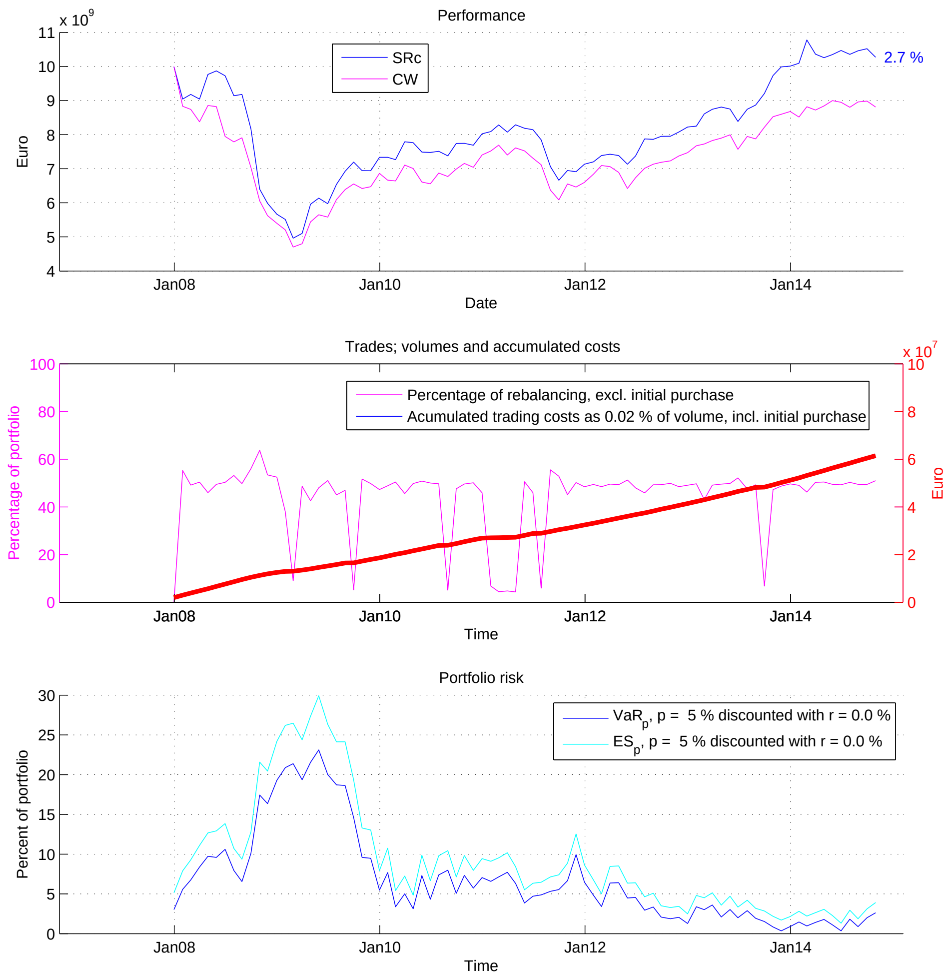

Figure 29: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

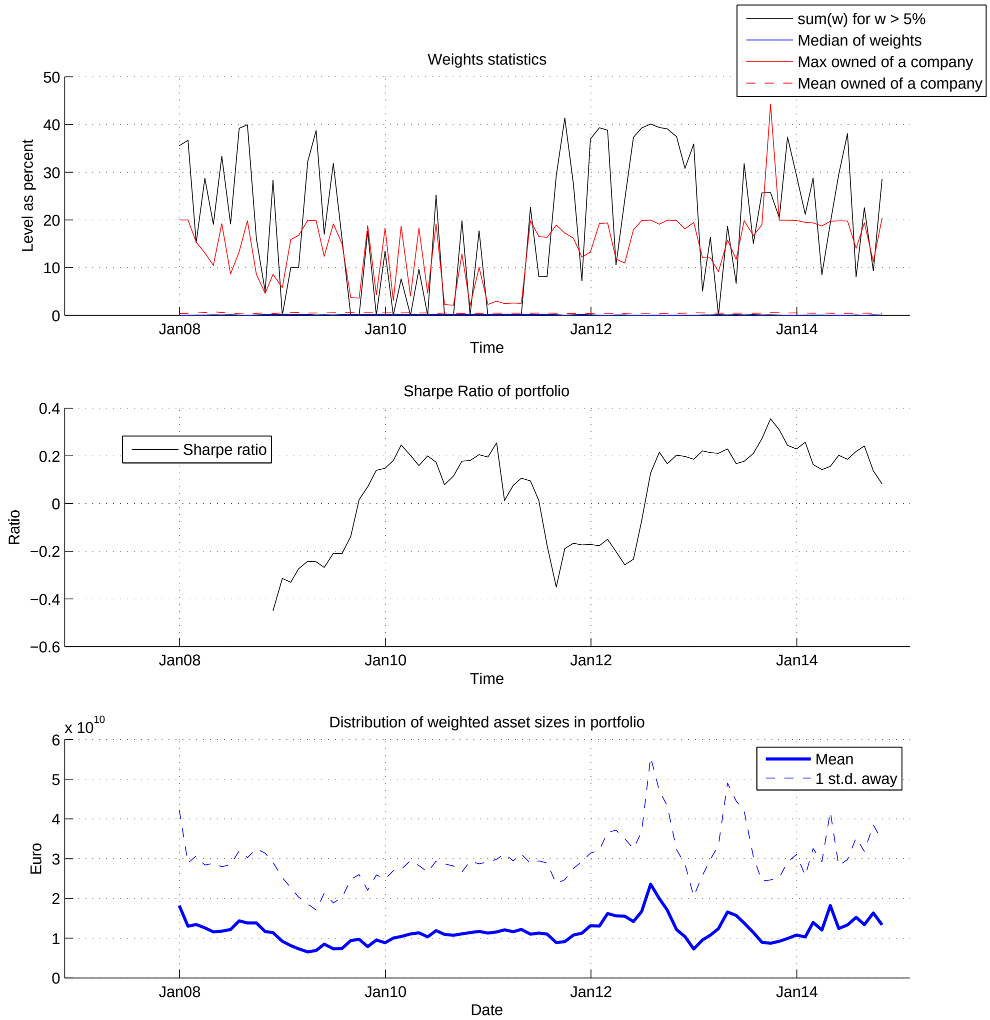

Figure 30: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

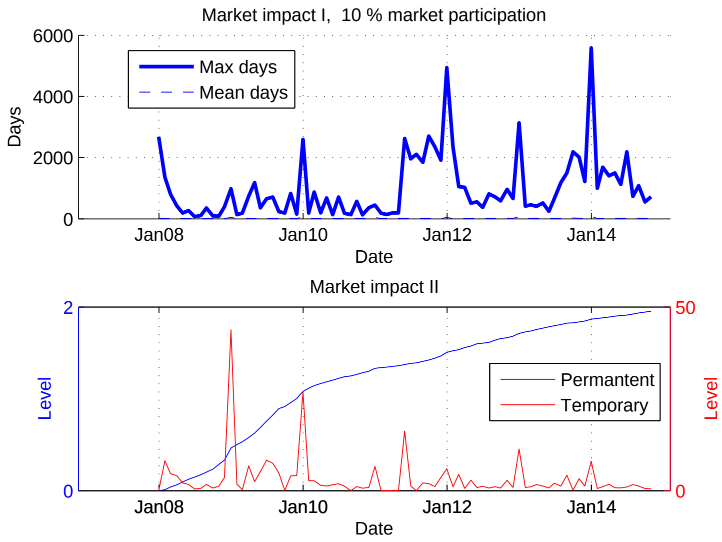

Figure 31: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

Weighted for Maximum diversification

Ordinary

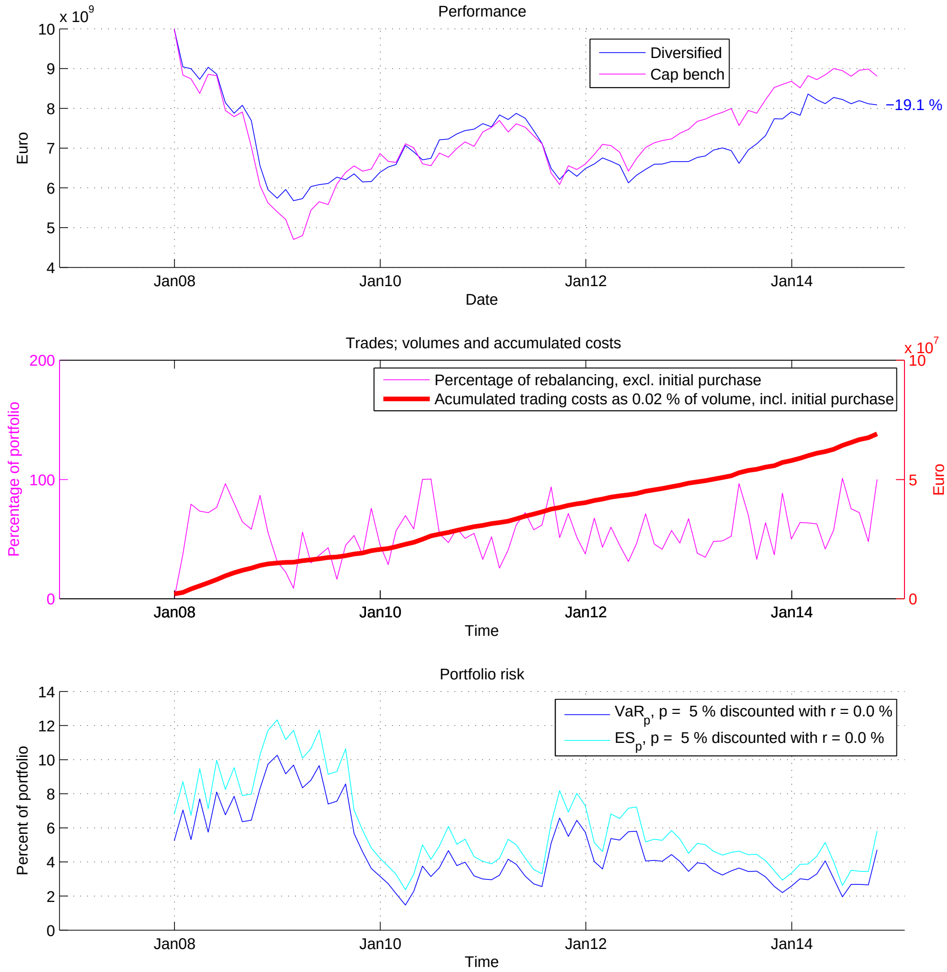

Figure 32: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

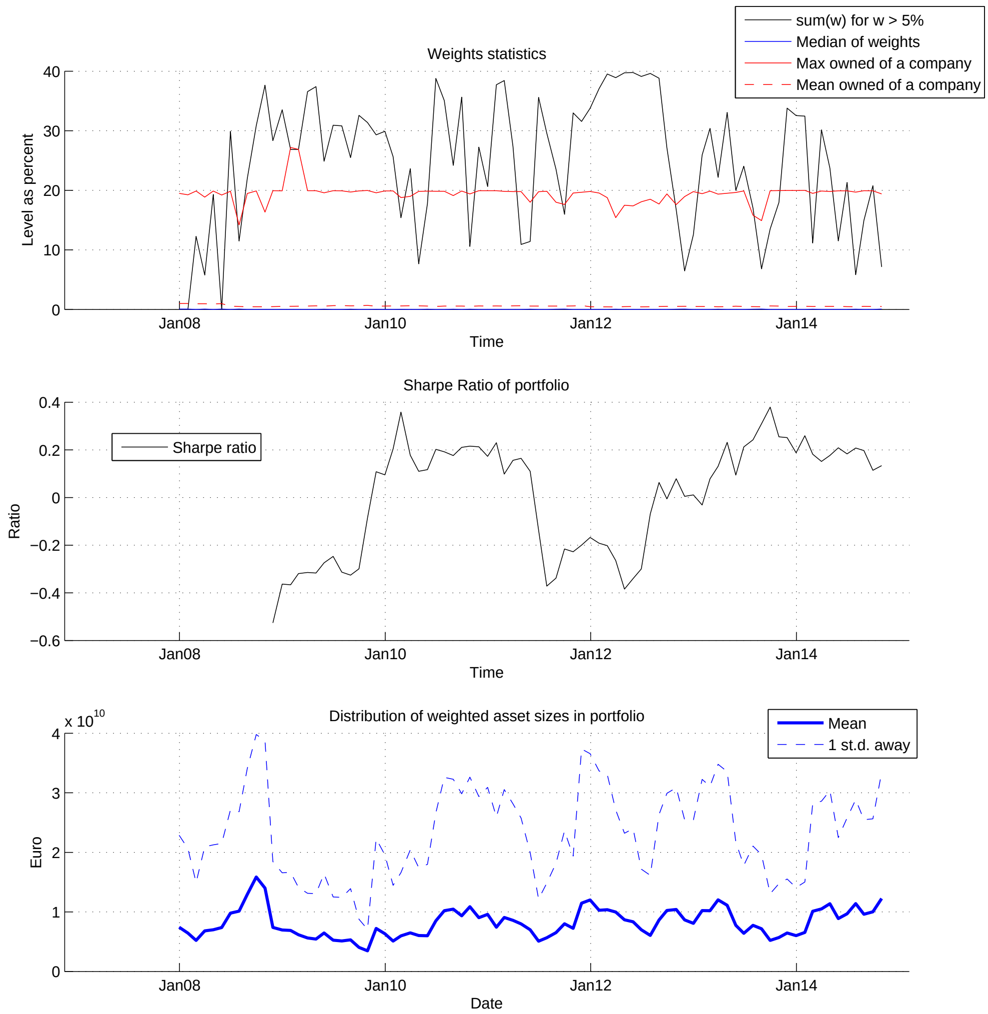

Figure 33: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

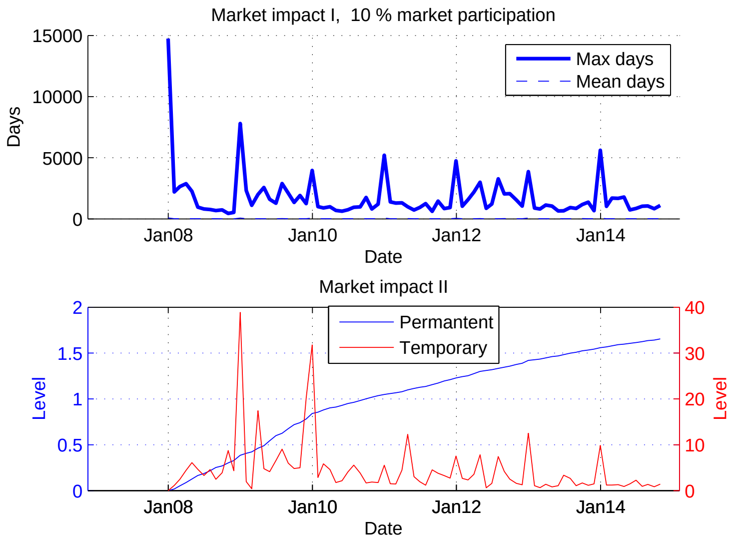

Figure 34: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

With rebalancing constraint

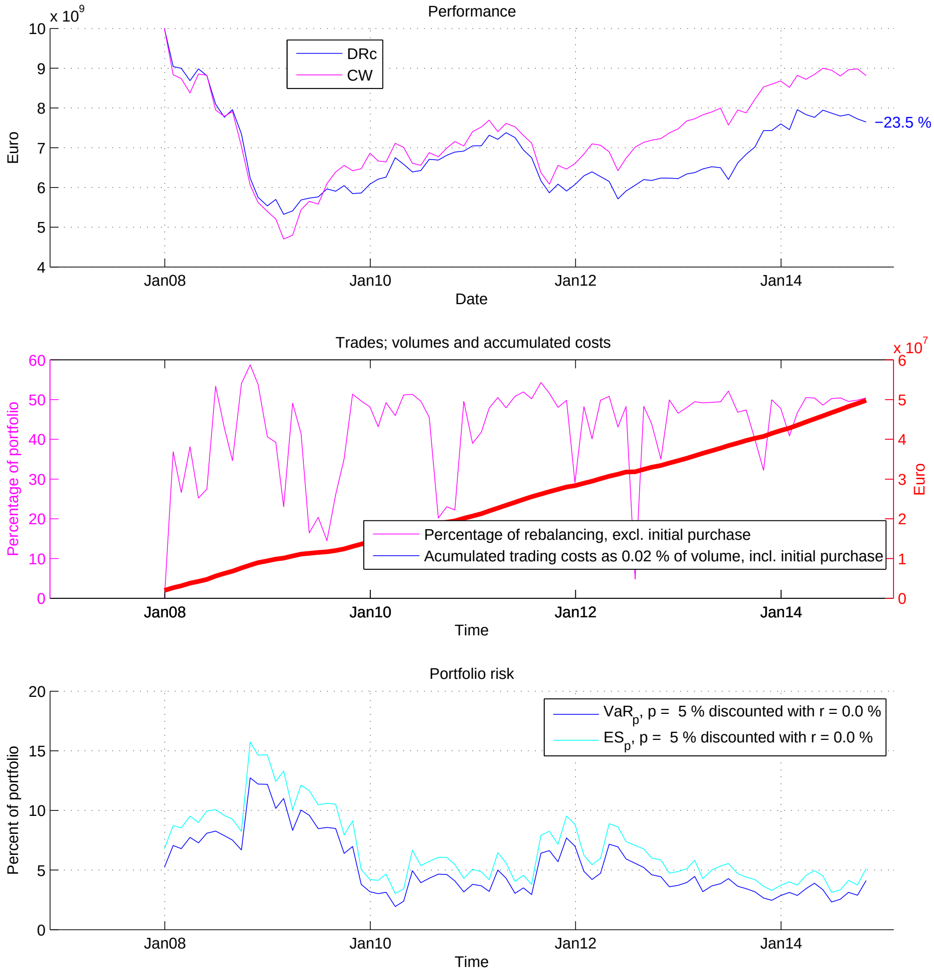

Figure 35: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

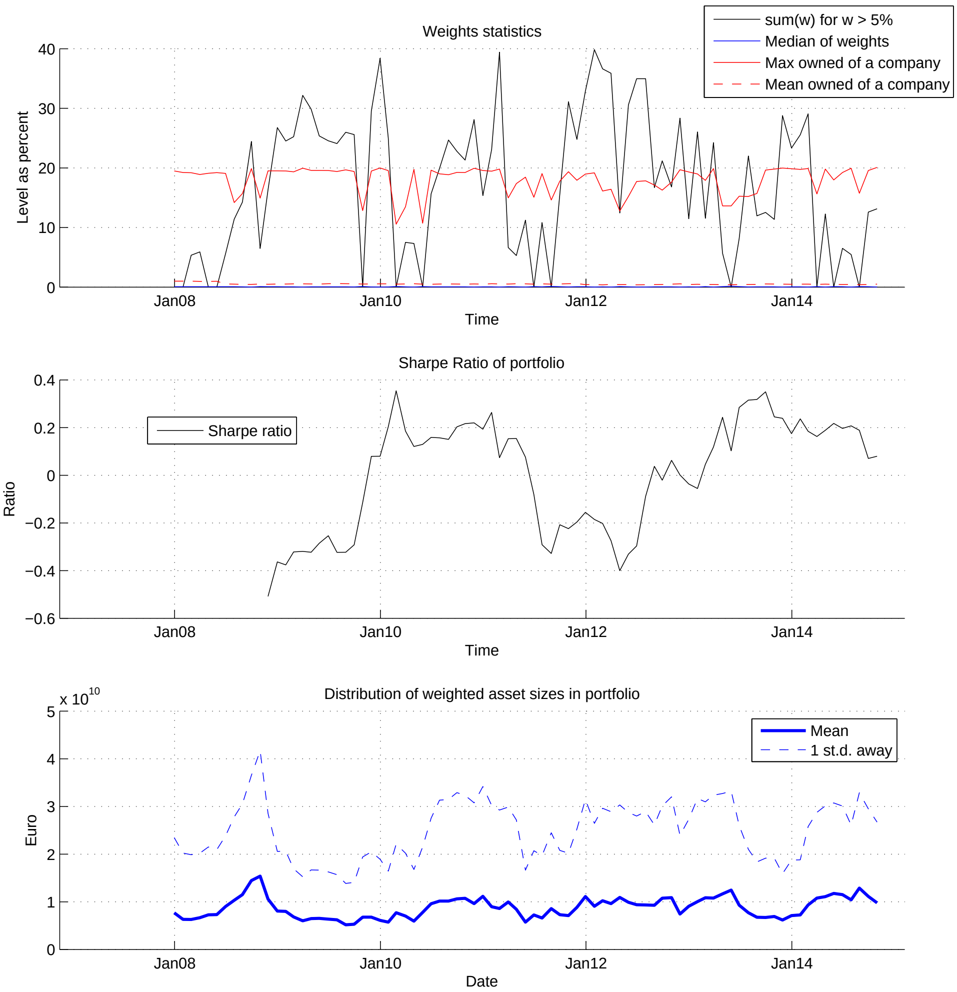

Figure 36: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

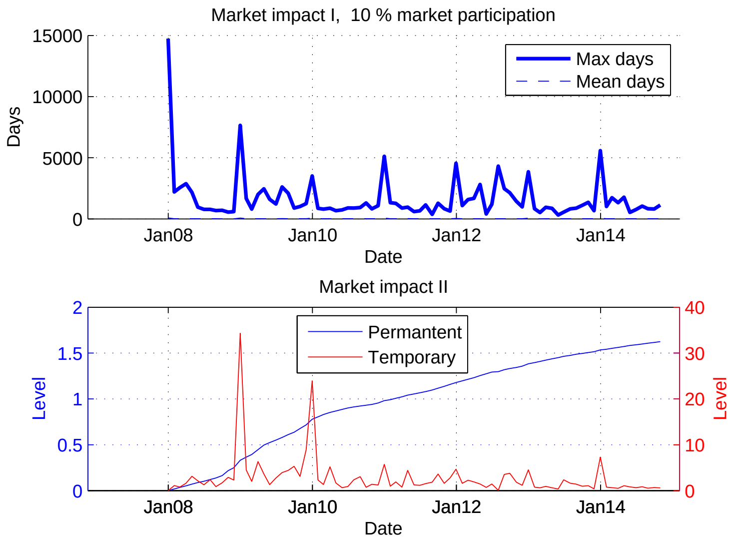

Figure 37: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

Fundamental weighting

Ordinary

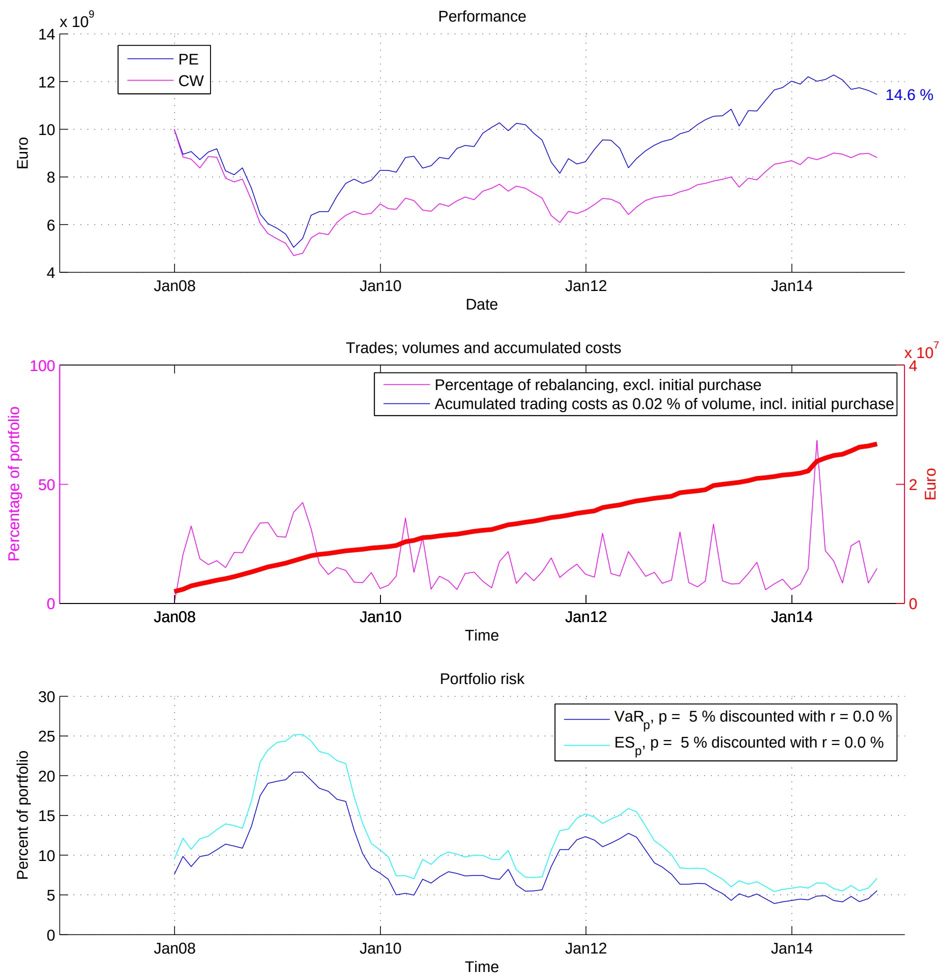

Figure 38: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

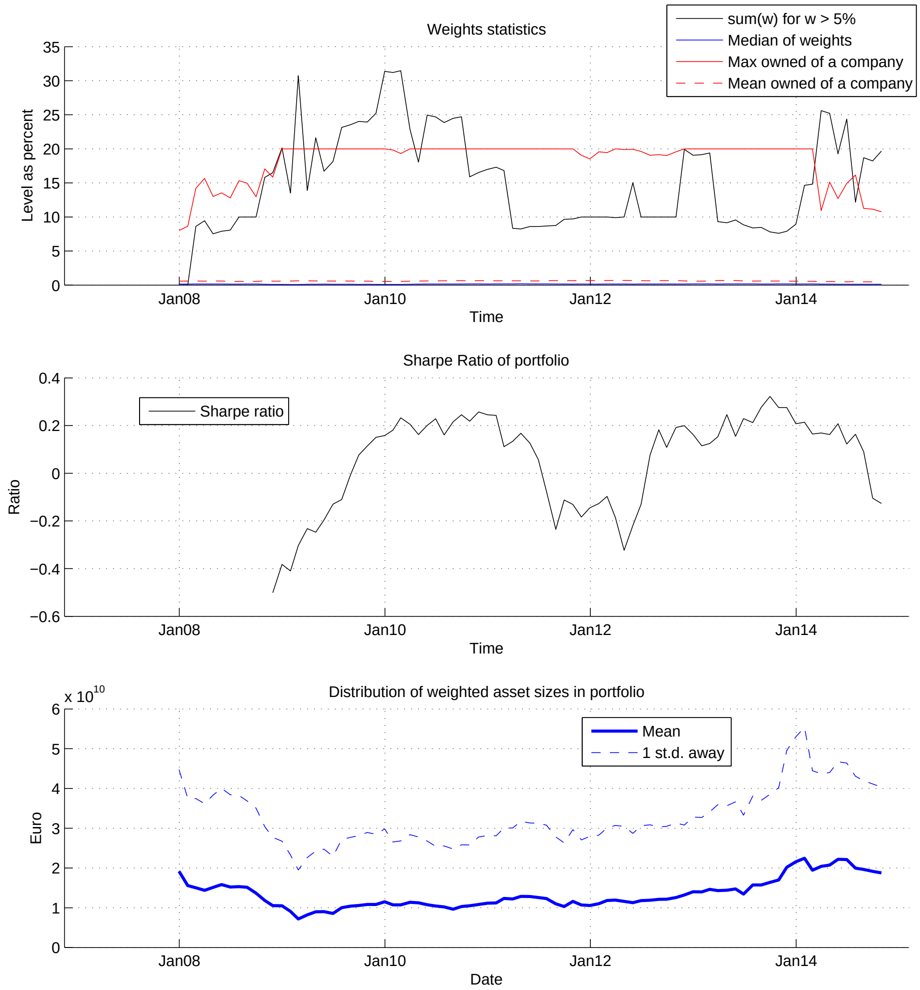

Figure 39: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

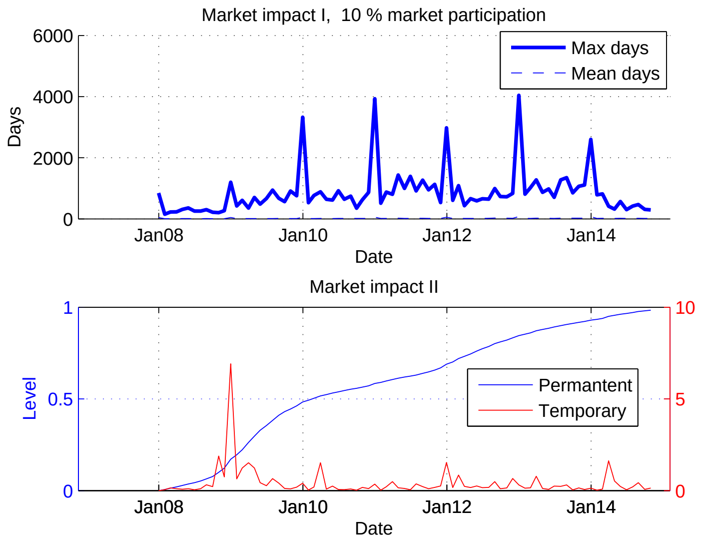

Figure 40: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

With rebalancing constraint

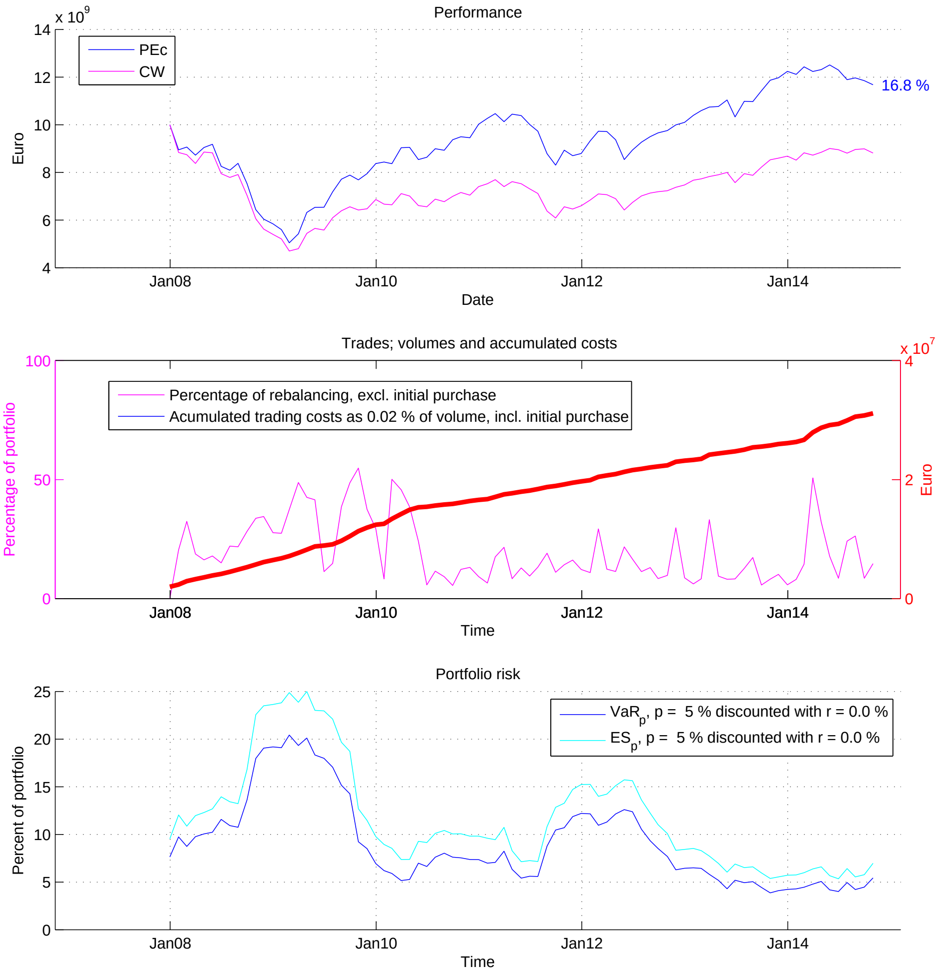

Figure 41: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

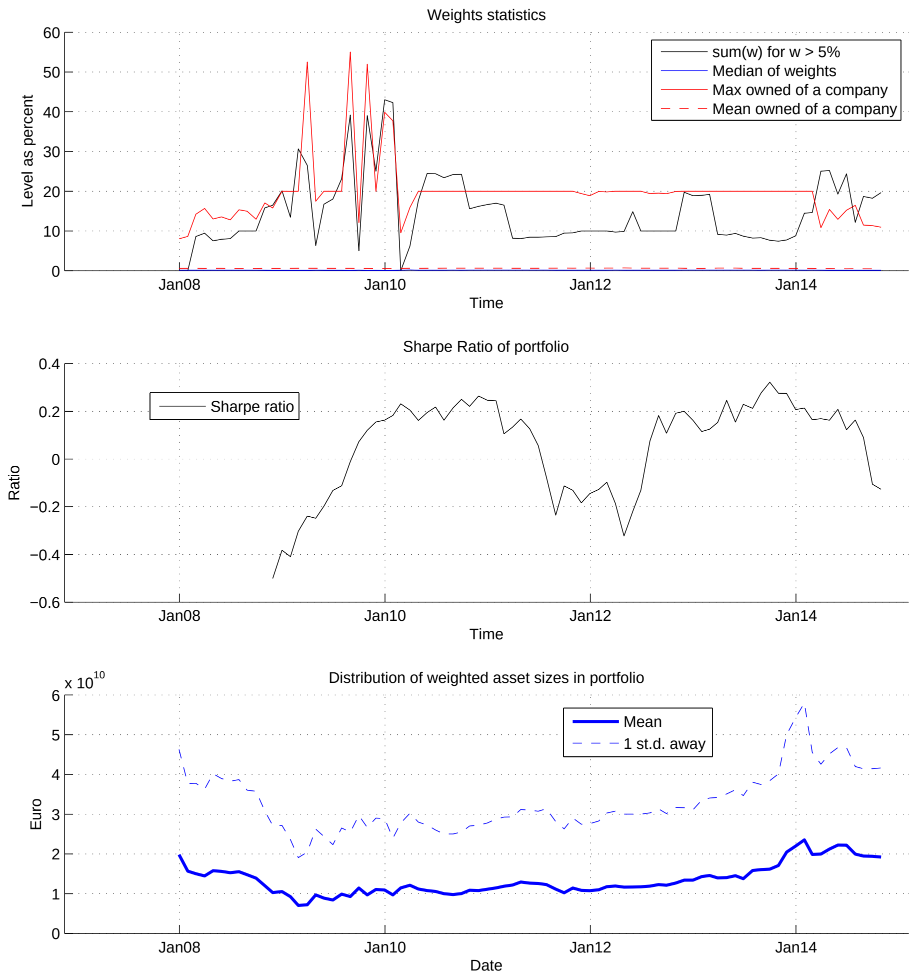

Figure 42: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

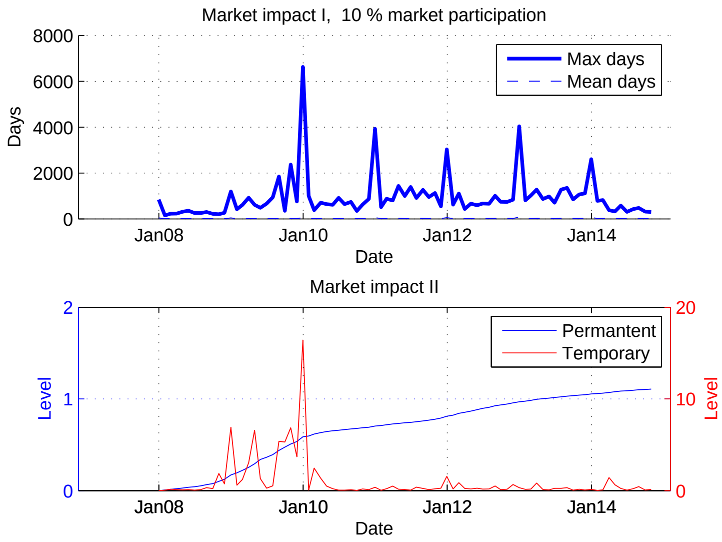

Figure 43: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

Combination of P/E and Sharpe

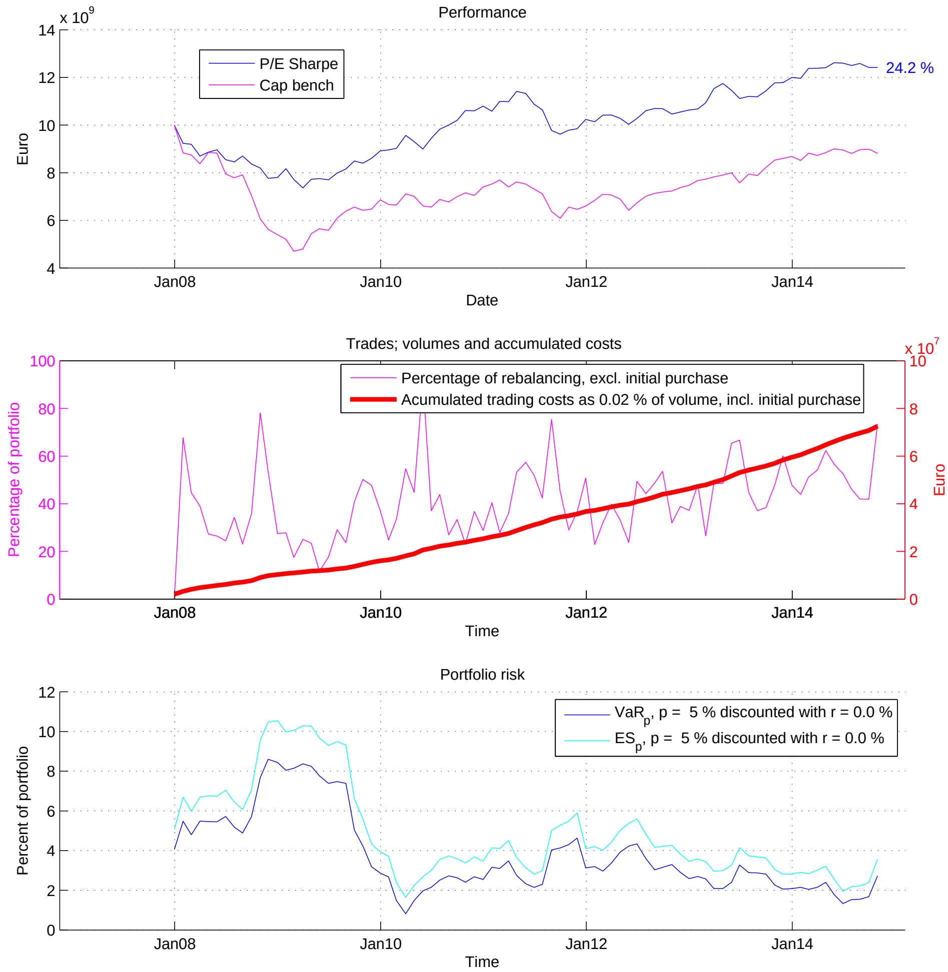

Figure 44: *Top* Performance compared to the capital weighted benchmark. *Middle* Rebalancing volume as percentage of total portfolio and the accumulated cost. *Bottom* Value at Risk and Expected Shortfall as percentage of total portfolio.

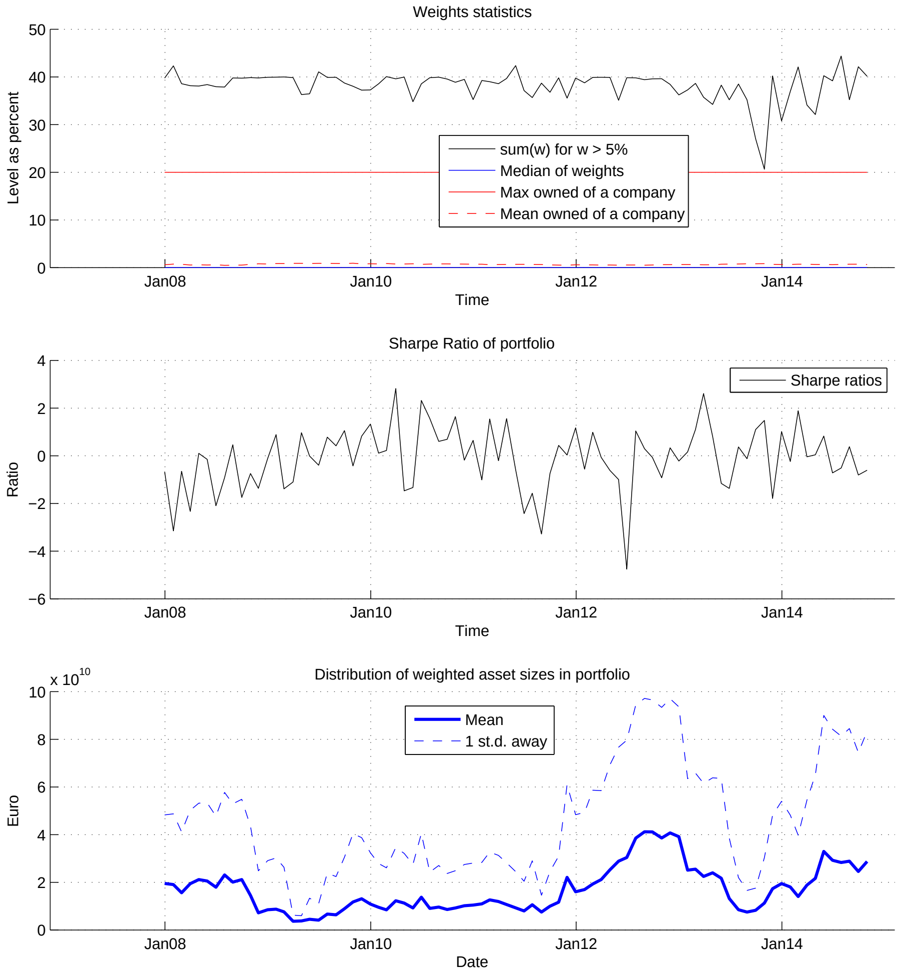

Figure 45: *Top* Statistics for sizes of weights. *Middle* Sharpe ratios at different times, calculated using historically estimated portfolio variances and actual portfolio returns. *Bottom* Weighted mean size of companies (assets) held at each time, along with the level one standard deviation above.

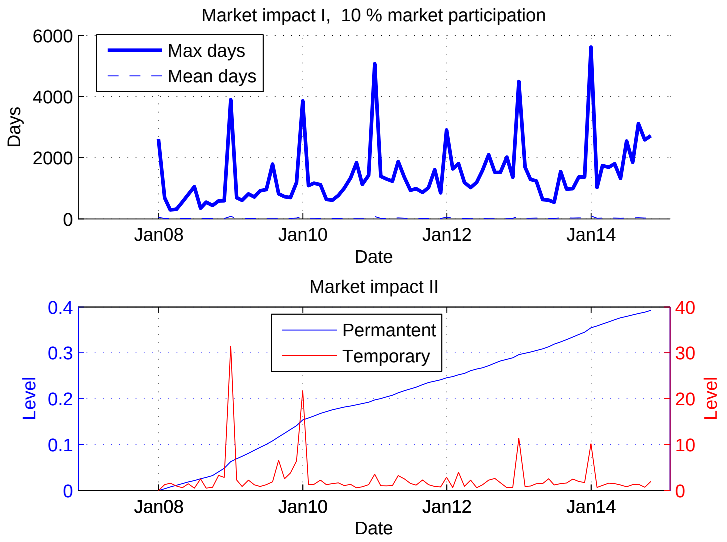

Figure 46: *Top* Days to sell total portfolio with 10% market participation. Maximum time and average time for assets. *Bottom* Total temporary and accumulated permanent market impact as calculated using JP Morgan's model.

Appendix II - Proofs

Proof of Theorem 2.2

Proof. Theorems 0.1 and 0.2 state that we have convexity. We know from Karush-Kuhn-Tucker theory that if x^* is optimal in program (2.12) then x^* satisfies:

$$\begin{aligned} \nabla\left(\frac{A(x)}{\sqrt{B(x)}}\right) &= \sum_{\forall i}\lambda_i^1\nabla g_i(x) \\ \Rightarrow \frac{1}{\sqrt{B(x)}}\left(A'(x) - 0.5B'(x)\frac{A(x)}{B(x)}\right) &= \sum_{\forall i}\lambda_i^1\nabla g_i(x) \end{aligned}$$

with λ^1 being the KKT multipliers corresponding to x^*, then by setting:

$$\begin{aligned} & a = \frac{1}{\sqrt{B'(x^*)}}, \quad b = 0.5a\frac{A(x^*)}{B(x^*)} \\ \Rightarrow \quad & aA'(x^*) + bB'(x^*) = \sum_{\forall i}\lambda_i^1\nabla g_i(x) \\ \Rightarrow \quad & A'(x^*) + cB'(x^*) = \sum_{\forall i}\lambda_i^2\nabla g_i(x) = \nabla\left(A(x^*) + cB(x^*)\right) \\ & c = \frac{b}{a}, \quad \lambda_i^2 = \frac{\lambda_i^1}{a} \end{aligned}$$

Hence x^* is optimal in program (2.13) as well, with λ^2 being the corresponding KKT multipliers. Convexity of the constituents needs to be proved next.

Q.E.D.

Theorem 0.1. *The portfolio mean return denoted $\bar{\mu}^T\bar{w}$ is a convex function.*

Proof. A multi variable function f is convex if its Hessian H_f is positive semi-definite. The Hessian is calculated as:

$$H_f = \nabla^2 f = \nabla^2\left(\bar{\mu}^T\bar{w}\right) = \bar{\bar{0}}$$

The matrix $\bar{\bar{0}}$ has all eigenvalues $\lambda = \bar{0}$ hence non-negative and therefore positive semi-definite. The portfolio mean return is thus a convex function.

Q.E.D.

Theorem 0.2. *The portfolio variance denoted $\sqrt{\bar{w}^T\Sigma\bar{w}}$ is a convex function.*

Proof. A multi variable function f is convex if its Hessian H_f is positive semi-definite. The Hessian is calculated as:

$$H_f = \nabla^2 f = \nabla^2 \left(\bar{w}^T \Sigma \bar{w}\right) = 2 * \Sigma$$

Hence the portfolio variance is convex if the covariance matrix Σ is positive semi-definite, which is always the case (positive definite??). It is easily seen that the commonly used form $\sqrt{\bar{w}'\Sigma\bar{w}}$ will also be convex due to the fact that the argument is always positive and the Hessian positive definite.

Q.E.D.

Considering the analogy with the Sharpe ratio and the slope of the capital market line and the fact that the efficient frontier is convex, it is clear that the Sharpe ratio itself is also a convex function. Hence $\frac{A(x)}{\sqrt{B(x)}}$ is convex in this case.

Σ is positive definite.

Proof of Theorem 2.3

Follows from the proof of Theorem 2.2.

Proof of Theorem 2.1

Proof. Let $V_{i,t_k} = w_{i,t_k} V_{t_k}$ be the desired value of instrument X_i at time t_k in the portfolio as determined by the strategy. Let $V^*_{i,t_k} = V_{i,t_{k-1}} * r_{i,t_k}$ be the actual value of instrument X_i at time t_k in the portfolio, before rebalancing, as a result of the instrument's movement from time t_{k-1}. Then the difference between the two is the value of the rebalancing required, calculated as:

$$\begin{aligned}
c_{i,t_k} =& V_{i,t_k} - V^*_{i,t_k} = \\
=& w_{i,t_k} V_{t_k} - V_{i,t_{k-1}} r_{i,t_k} = \\
=& w_{i,t_k} V_{t_{k-1}} R_{t_k} - w_{i,t_{k-1}} V_{t_{k-1}} r_{i,t_k} = \\
=& V_{t_{k-1}} \left(w_{i,t_k} R_{t_k} - w_{i,t_{k-1}} r_{i,t_k}\right)
\end{aligned}$$

Q.E.D.