Algorithmic Portfolio Rebalancing - A Test of the Efficient Market Hypothesis

Mattias Lundahl
19372@student.hhs.se

4210 Thesis within Finance
Stockholm School of Economics
Fall 2005

Tutor: Joel Reneby
Discussants: Rickard Blomdahl and Adam Lewenhaupt
Presentation: December 15, 2005, 8.15-10.00, Room 343

Acknowledgements

I would like to thank my tutor Joel Reneby for his valuable feedback on the thesis during my work on it.

I would also like to thank Johan Karlsson of Kaupthing Bank Sverige AB for providing me with numerous time series that I have used in the empirical testing.

Abstract

This thesis aims at testing the Efficient Market Hypothesis (EMH) by implementing and evaluating four distinct algorithms (Universal Portfolio, Exponentiated Gradient, Anticor and Constant Proportion Portfolio Insurance) for automated rebalancing of fixed-asset portfolios based on the past performance of the individual assets included in the portfolio. If the EMH holds, technical analysis such as algorithm based investments should not be able to generate abnormal returns without introducing abnormal risk. The algorithms are implemented according to the articles presenting them, and I perform statistical hypothesis tests to determine whether the algorithms can provide significant positive abnormal return over broad indices.

The results indicate that abnormal returns above broad indices such as the S&P500 and the STOXX are possible. In Monte Carlo simulations, results are statistically significant at the 1% level for three of the algorithms. In tests on actual time series, one algorithm provides statistically significant abnormal return on the 5% level. All algorithms have economically significant abnormal returns for actual time series, indicating that technical analysis can create value despite the assertion of the weak form of the EMH.

Trading costs are also introduced. Two algorithms prove very sensitive to trading costs, one is fairly sensitive and one is almost completely insensitive to trading costs, indicating that this algorithm might be useful even for smaller investors.

Contents

List of Tables

List of Figures

1 Introduction

A common joke in the field of finance is that of two investors walking down a road, when one of them sees a $100 bill. The other one, a firm believer in market efficiency, quickly says "Don't bother picking it up. If it were real, someone else would already have picked it up."

This thesis aims to test the weak form of the efficient market hypothesis by examining four algorithms[1] using only publicly available, numerical information to determine how to invest in baskets of securities. I will test the ability of the algorithms to beat different benchmarks over extended time periods. Historical security data will be used as input to the algorithms, ensuring that realistic conditions are used. To test the algorithms more thoroughly, Monte Carlo simulations will also be employed, primarily using the observed historical distribution of returns.

The term algorithmic portfolio rebalancing refers to a particular class of investment strategies. They use numerical information about financial assets, such as past market prices, to calculate a suggested rebalancing scheme for coming time periods. Only the weighting of wealth among the assets in the portfolio is calculated; the starting choice of which assets to use must be determined exogenously. This class of algorithms is interesting to study because of the possibility to earn higher returns than what is possible using buy-and-hold schemes or similar unsophisticated strategies, while still having a rebalancing strategy that does not require human analysis beyond the initial choice of assets.

The first algorithm studied is the Universal Portfolio algorithm. It was introduced by Thomas Cover [6], who also co-wrote several articles elaborating on different aspects of the algorithm [7], [8], [23], [14]. Kalai and Vempala [16] aimed to reduce implementational and computational complexity. Some further extensions were provided by Blum and Kalai [3]. Intuitively, this algorithm tries to pick up momentum in the assets, allocating more to the assets that have performed best in the past.

The second algorithm, the Exponentiated Gradient algorithm (EG), was introduced by Helmbold et al [13]. While the theoretical properties of this algorithm are weaker than those of Universal Portfolio, the article authors find that in experiments, EG tends to outperform Universal Portfolio. Like Universal Portfolio, EG also tries to benefit from momentum in the assets, moving wealth from assets that have performed poorly to assets with higher returns.

The third algorithm, Anticor, was introduced by Borodin et al [4]. It differs from Universal Portfolio and EG in the sense that it does not try to use momentum effects. Instead, it is developed to benefit from negative correlations and mean reversion in the prices of the assets. The article authors also find that this algorithm tends to prefer different assets compared to Universal Portfolio and EG.

The final algorithm evaluated, the Constant Proportion Portfolio Insurance

[1]The word algorithm refers to a rule, or set of rules, that unambiguously defines a process for solving a particular task.

(CPPI), was introduced by Black and Jones [2] and Perold and Sharpe [24] as an insurance type algorithm, which combines a guaranteed, riskless minimum return with a controlled investment in a risky asset that enables excess returns. There is a conceptual limitation in this algorithm compared to the others used in this thesis; CPPI by construction allows only portfolios consisting of two assets, one risky and one risk free. Extensions to the CPPI algorithm have been devised by Bertrand and Prigent [1] and Boulier and Kanniganti [5].

To determine how useful the results would be in practice, transaction costs will be introduced into the model. This will allow evaluating whether the algorithms are truly usable for investment purposes. I will try to determine the critical transaction cost level at which any benefit from using the algorithms is extinguished.

The main question examined in this thesis is *Does the Efficient Market Hypothesis hold for these algorithms?* If it does, the algorithms should not be able to give higher returns than market benchmarks without also having higher risk.

The rest of the thesis is organized as follows: Section 2 describes previous work on market efficiency and also gives the technical background for the algorithms used, section 3 presents the data series that are used, section 4 describes the methodology, section 5 holds the findings of the tests used, section 6 contains the analysis of the results and section 7 provides conclusions and suggestions for further research. Some of the technicalities regarding the algorithms are presented in appendices to make the core subject of the thesis easier to follow. In the last appendix, a glossary is provided for convenient access to terminology explanations.

2 Theoretical background

2.1 Literature review

The Efficient Market Hypothesis (EMH) in essence builds upon the theory of Homo Oeconomicus, the rational human, which can access information and react to it in a balanced, economically rational manner. A market consisting of large amounts of investors which are assumed to be rational can be assumed to be a rational, efficient market. The EMH has been a prevalent model for the prices of securities in financial markets. It has three common forms [29],

- Weak form: All past market prices and data are fully reflected in securities prices. In other words, technical analysis does not provide a possibility to earn abnormal risk-adjusted returns.
- Semi-strong form: All publicly available information is fully reflected in securities prices. In other words, fundamental analysis does not provide a possibility to earn abnormal risk-adjusted returns.
- Strong form: All information is fully reflected in securities prices. In other words, even insider information does not provide a possibility to earn abnormal risk-adjusted returns.

Fama [10] wrote an influential article about the EMH, in which was defined that 'a market in which prices always "fully reflect" available information is called "efficient"'. Fama has in later works changed the exact wording of the definition to refine its meaning. He has also tested for autocorrelations in returns data[9], finding that there are often positive serial correlations of the first order, suggesting that there are momentum effects.

In [11], Fama provides an overview of the research presented so far in the market efficiency field. Apart from his own previous findings, other authors have also found positive autocorrelations, especially in small stocks. However, Fama and others argue that the predictability of returns is largely obscured by daily variations in stock prices, implying that the actual usefulness of the predictability is small. Over longer time periods, some authors have found anomalies that they argue can be considered irrational deviations, but Fama and others argue that the patterns are not distinguishable from rational time-varying expected returns.

Another article by Fama [12] includes event studies where abnormal returns from stocks are examined to determine whether there are any anomalies that indicate nonefficiency of financial markets. Fama finds that the evidence does not suggest abandoning the EMH when scrutinized. Apparent overreactions of stock prices to new information is found to be about as common as underreactions, and it is about as common for pre-event abnormal returns to continue as it is for them to reverse after the event. In addition, Fama finds that long-term return anomalies are fragile and tend to disappear rather quickly.

The EMH has not been unanimously accepted by all observers and investors in the financial market however. Many attempts have been made to disprove the theory, and Malkiel [22] has written an article that addresses several such attacks. Malkiel personally argues that the EMH does hold and that evidence of the opposite often is built upon questionable methodology, such as e.g. using a time frame with properties which are not representative for longer time periods. He also argues that it is important to distinguish between statistical and economic significance and points out that transaction costs are likely to have a notable impact on the true profitability of suggested investment schemes.

Earlier, Malkiel [20] has also argued that a blindfolded chimpanzee throwing darts at the Wall Street Journal could select a portfolio that would do as well as the experts. In later work [21], he explains that the main point of this claim is to recommend investments into index funds rather than actively managed funds.

Shostak [27] argues that there is reason to use fundamental analysis to earn abnormal returns, and consequently that the EMH does not hold. His main point is that the EMH is largely an equilibrium model that does not represent the real behavior of financial markets in the short run. This does not really contradict most EMH proponents however, since the EMH is typically only argued to hold over longer time periods.

Schleifer and Summers [26] among others point to the stock market crash of 1987 as a sign that financial markets are in fact not efficient. Malkiel's response is that while psychological factors could be held as evidence against the EMH, the crash was also largely explained by a series of negative events for the investment environment. Shiller [25] among others considers the internet bubble at the end of the 1990s evidence against the EMH. Malkiel comments on this by noting the difficulties in correctly valuing equity and that despite the overvaluation of high tech stock, no strict arbitrage opportunities existed, since it was not possible to predict the time that the bubble would burst. Furthermore, he argues that since this type of bubble occurs infrequently, it does not really contradict the idea that an efficient market on average prices securities correctly.

Lo and MacKinlay [17] have found that serial correlations in stock prices are often non-zero, meaning that there in fact is evidence of momentum, and consequently that the random walk model is not fully realistic. Similar findings were presented by Lo, Mamaysky and Wang [18], that used non-parametric statistical techniques to show that some price series patterns studied in technical analysis, for instance head-and-shoulders and double-bottom formations, actually have some predictive power.

Jensen [15] gives support to the EMH, performing the first study on mutual fund performance, in which he finds that active fund managers on average underperform the market by approximately the amount of their added expenses. Malkiel [22] shares this argument by claiming that there is also a substantial survivorship bias, meaning that the true average active fund performance should be even lower.

It should be noted that few observers or investors believe that the EMH strictly holds at all times. Malkiel argues that investors clearly are not rational at all times, and that over certain time windows, apparent inefficiencies can arise and be exploitable for some time. The common view of EMH proponents is that such occurences will be exception rather than rule and will disappear quickly if and when they are discovered.

2.2 Algorithm descriptions

In this section, the background of algorithmic portfolio rebalancing in general and for the four algorithms to be used will be described briefly. The section has been kept short, moving the more technical parts to appendices for interested readers. This is to make the testing of the EMH the focus of the thesis rather than the technicalities of the particular algorithms.

2.2.1 Common notation

The investment period is assumed to be discrete points in time starting at time 0, i.e. $t = 0, 1, 2, \ldots, n$. A total of m assets are assumed to be available for the algorithms to divide wealth among. These assets are assumed to have a price at each of the discrete points in time included in the investment period, with $s_{t,j}$ denoting the price of asset j at time t. Then, the vector $\mathbf{x}_t = (x_{t1}, x_{t2}, \ldots, x_{tm})^{\mathrm{T}}$,[2] where

$$x_{tj} = \frac{s_{t,j}}{s_{t-1,j}},$$

denotes the performance of the investment universe from time $t-1$ to t, since $x_{tj} = (1+r_{tj})$, where r_{tj} is the simple periodic return on asset j from time $t-1$ to time t.

The concept of wealth is defined as a time series that takes real number values S_t at all time points and that starts at time 0 with the value 1, $S_0 = 1$. Values at later points in time denote relative wealth compared to the initial wealth of 1; a wealth of 2 at some point in time denotes twice the initial wealth etc.

Let the weight vector at time t be denoted $\mathbf{b}_t = (b_{t1}, b_{t2}, \ldots, b_{tm})^{\mathrm{T}}$. By definition of a weight vector, $\sum_j b_{tj} = 1$ at all times t. In this notation, b_{tj} is the fraction of total wealth that is invested in the j:th asset at time t. When one time period has passed from the initial state, the wealth will be $S_1 = b_{11}x_{11} + b_{12}x_{12} + \ldots + b_{1m}x_{1m} = \mathbf{b}_1^{\mathrm{T}}\mathbf{x}_1$, and after n time periods, the total wealth, henceforth referred to as *total return factor*, is calculated as:[3]

$$S_n = \prod_{t=1}^{n} \mathbf{b}_t^{\mathrm{T}}\mathbf{x}_t \tag{1}$$

The total return factor will be used throughout the thesis as a performance measure.

The term *constant rebalanced portfolio* (CRP) refers to a rebalancing scheme where the portfolio is only rebalanced in order to keep the fraction of wealth invested in each of the assets in the portfolio unchanged at all rebalancing times. Using equation 1 and the notation $\mathbf{b}$ for the constant weight vector, it is clear that after n time periods, the wealth of such a portfolio will be $S_n = \prod_{t=1}^{n} \mathbf{b}^{\mathrm{T}}\mathbf{x}_t$.

From the definition of CRP, it is clear that ex post, there must for each set

[2] The raised T refers to the transpose operation that shifts row vectors into column vectors and vice versa.

[3] The large Pi character denotes multiplication over the index t, which goes from 1 to n. Each factor in this product is a sum of m products of a weight b_{tj} and a relative price change x_{tj} as shown in the example calculation of S_1.

of assets be a *best constant rebalanced portfolio* (BCRP), which is the portfolio of the CRP class that has obtained the highest wealth after n time periods. Mathematically, its wealth after n time periods can be written

$$S_n^* = \max_{\mathbf{b}} \prod_{t=1}^{n} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t.$$

Clearly, the determination of the BCRP requires that the outcomes of asset prices are known, and thus it can only be determined after the outcomes have been observed. It should be noted that by construction the BCRP will always provide a wealth which is at least as good as the best individual component of the portfolio.[4]

2.2.2 Universal Portfolio

The Universal Portfolio algorithm, introduced by Cover [6], can ensure that the asymptotic performance of the portfolio is comparable to BCRP performance.

With a sequence of market performance vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots$ as defined above, the Universal Portfolio algorithm specifies that weights should be calculated according to the following equation:

$$\hat{\mathbf{b}}_k = \frac{\int \mathbf{b} \prod_{t=1}^{k-1} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t \, d\mu\left(\mathbf{b}\right)}{\int \prod_{t=1}^{k-1} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t \, d\mu\left(\mathbf{b}\right)}, \tag{2}$$

Here, $\mu\left(\mathbf{b}\right)$ should be the multivariate probability distribution of the BCRP weights for the specific asset set used.

Intuitively, this can be interpreted as follows: the weights are calculated as the weighted average of all possible weight allocations, using the CRP performance over all past time periods as weight function. A methodological problem for this algorithm is that the right hand side of equation 2 is not feasible to calculate for all probability distributions.

In this thesis, the Dirichlet distribution will be used. It is limited to the $[0, 1]$ interval for each variable and is thus suitable for weight vectors. Furthermore, Universal Portfolio will only be tested on portfolios consisting of two assets, since the algorithm becomes computationally infeasible for larger portfolios. Thus, the distribution used is the two-variable Dirichlet distribution.

After a long derivation, which is given in appendix A.1, it is shown that the Universal Portfolio weights after n time periods, using the Dirichlet distribution, can be calculated as follows:

$$\hat{\mathbf{b}}_n = \frac{1}{\sum_{l=0}^{n-1} Q_{n-1}(l)} \begin{bmatrix} \sum_{l=0}^{n-1} \frac{l+\alpha_1}{n+\alpha_1+\alpha_2-1} Q_{n-1}(l) \\ \sum_{l=0}^{n-1} \frac{n-l+\alpha_2-1}{n+\alpha_1+\alpha_2-1} Q_{n-1}(l) \end{bmatrix}.$$

[4]From a set of m assets, taking any individual asset is equivalent to a CRP with a weight vector that contains exactly one element that is 1, and the remaining elements 0. Since the BCRP by definition is the best CRP, it will be at least as good as the best one asset portfolio.

In the above expression, $Q_n(l)$ must be calculated recursively using $Q_0(0) = 1$, and the following recursion rules

$$Q_n(l) = x_{n1} \frac{l + \alpha_1 - 1}{n + \alpha_1 + \alpha_2 - 1} Q_{n-1}(l-1) + x_{n2} \frac{n - l + \alpha_2 - 1}{n + \alpha_1 + \alpha_2 - 1} Q_{n-1}(l)$$

for $1 \leq l \leq n-1$, and the corresponding endpoint recursions are

$$Q_n(0) = x_{n2} \frac{n + \alpha_2 - 1}{n + \alpha_1 + \alpha_2 - 1} Q_{n-1}(0),$$

$$Q_n(n) = x_{n1} \frac{n + \alpha_1 - 1}{n + \alpha_1 + \alpha_2 - 1} Q_{n-1}(n-1).$$

These equations specify the Universal Portfolio weights for general values of α_1 and α_2, the parameters of the two variable Dirichlet distribution.

Universal Portfolio has a theoretical guarantee (as shown by Cover [6]) that (with S_n^* denoting BCRP performance and $\hat{S}_n$ Universal Portfolio performance)

$$\lim_{n \to \infty} \frac{1}{n} \ln \left(\frac{S_n^*}{\hat{S}_n} \right) = 0$$

which intuitively means that asymptotically[5], the Universal Portfolio will grow at the same average continuously compounded rate as the BCRP. Note that this requires that the probability distribution used is in fact the true distribution of BCRP weights; the Dirichlet distribution might not be able to actually fulfill this property even asymptotically.

2.2.3 Exponentiated Gradient

This algorithm, introduced by Helmbold et al [13], has slightly weaker theoretical properties than Universal Portfolio. Using the same notation as above,

$$\lim_{n \to \infty} \frac{1}{n} \ln \left(\frac{S_n^*}{\hat{S}_n} \right) = \sqrt{\ln \left(\frac{m}{2r^2} \right)}$$

where m is the number of assets in the portfolio and r is a lower bound on the relative price changes between trading days. This can hypothetically be arbitrarily close to zero, meaning that the right hand side could be large.

It is possible according to Helmbold et al to achieve the same asymptotic performance with Exponentiated Gradient (EG) as with Universal Portfolio, but only using a modified version that is significantly more complex to calculate. However, according to the authors, EG often provides better performance than Universal Portfolio in experiments. Results in their article show that EG can often come quite close to BCRP performance.

The EG algorithm has one parameter, denoted η, which is the so called *learning rate*. In this thesis, this parameter will be held constant over time for each test, although it is technically possible to let it vary over time. A high value of η

[5] As time approaches infinity.

means that EG more quickly reallocates to the strongest performing assets. The weights are calculated according to the following expression

$$b_{t+1,j} = \frac{b_{t,j} \exp\left(\frac{\eta x_{tj}}{\mathbf{b}_t^{\mathrm{T}} \mathbf{x}}\right)}{\sum_{k=1}^{m} b_{t,k} \exp\left(\frac{\eta x_{tk}}{\mathbf{b}_t^{\mathrm{T}} \mathbf{x}}\right)}, \tag{3}$$

which is given without any derivation or formal justification in the source article. One property can be noted: If the initial weight vector has no negative elements, then consecutive weight vectors also cannot have negative elements, since the exponential function is strictly positive. Thus, EG can be compared fairly to Universal Portfolio using the Dirichlet distribution in the sense that neither of them allow short selling of assets.

It should be noted that for EG it is computationally feasible to use portfolios of more than two assets. Since EG and Universal Portfolio are otherwise similar, it is interesting to see whether it is beneficial to use portfolios of more than two assets with EG. If it is not, then this can heuristically be assumed to hold also for Universal Portfolio, since it is similar to EG in conceptual functionality; it also tries to capture momentum and to obtain BCRP performance.

2.2.4 Anticor

The Anticor algorithm was introduced by Borodin et al [4]. The BCRP performance is *not* a strict upper bound for Anticor, as it is for Universal Portfolio and EG. Also, Anticor does not try to allocate to past strong performers, but instead tries to find mean reversion patterns in the asset prices. The algorithm has one parameter, namely the length of the *window*. This is a period of time used to determine correlations between the assets in the portfolio. The weights for coming periods are then determined using the calculated correlations. The name of the algorithm reflects the fact that it works best with negatively correlated assets.

Denote the length of the window w, which should be an integer number of points in time. The *growth rate* of an asset over the length of the window denotes the price of the asset at the end of the window period divided by the price at the beginning of the window period. Anticor will reallocate from asset j to asset k at time t if both of these conditions hold:

- The growth rate of asset j must be higher than that of asset k over the window starting at $t - w + 1$ and ending at t.
- There must be a positive correlation between asset j over the window from $t - 2w + 1$ to $t - w$ and asset k over the window from $t - w + 1$ to t.

If it is assumed that assets are mean-reverting and perform on average equally well over longer time periods, these two conditions together can be taken as an intuitive indication that asset k should outperform asset j over the window going from $t + 1$ to $t + w$.

The full formal description of the Anticor weight calculation is rather complicated, and for this reason it is presented in appendix A.2.

2.2.5 CPPI

The Constant Proportion Portfolio Insurance (CPPI) algorithm was introduced by Black and Jones [2] and Perold and Sharpe [24] and later tested by Bertrand and Prigent [1]. It is relatively common in commercial settings, where it is used as a form of insurance strategy. CPPI uses one risky asset and one risk free asset that increases in value at a continuously compounded risk free rate r.

The base form of CPPI is thus constrained to portfolios consisting of two assets, though it is technically possible to let the risky asset be a portfolio of several risky assets, even a Universal Portfolio, EG or Anticor algorithmic portfolio. Let V_0 denote the initial portfolio value and F_0 the initial *floor*. The purpose of the floor is to provide a guaranteed minimum portfolio value at any point in time. It is necessary to have $F_0 < V_0$. The floor will then be allowed to grow deterministically at the continuously compounded rate r, i.e. the same as the risk free asset. At all points in time t, C_t denotes the *cushion*, which is defined by

$$C_t = V_t - F_t.$$

The total investment in the risky asset, which is called the *exposure*, is denoted by E_t. CPPI prescribes that the exposure should be

$$E_t = mC_t,$$

where m is a parameter called the *multiplier*. The expected payoff function becomes convex if $m > 1$, but the floor is then no longer absolutely guaranteed. If $m \leq 1$ is enforced, the exposure will never be larger than the cushion, so at least the amount that constitutes the floor will be invested in the risk free asset.

Some generalizations of CPPI were introduced by Boulier and Kanniganti [5]. One extension is the moving floor, that is immediately raised following large increases in the value of the risky asset, thus limiting the cushion and consequently the exposure to the risky asset. The moving floor protects against subsequent falls in the price of the risky asset. An obvious disadvantage is that persistent strong performance of the risky asset will not be fully captured, since a higher floor implies a smaller cushion and thus a smaller exposure.

The moving floor version of CPPI also introduces a new parameter which is called p. It denotes the maximum fraction of the total wealth permitted to invest in the risky asset. To keep the guarantee of the floor as minimum wealth, set $p = 1 - f$, where $f = (F/V)_{\min}$, the minimum floor allowed expressed as a fraction of total wealth. The p parameter can also be used to invest more aggressively in the CPPI algorithm: Letting $p > 1$ allows short selling of the risk free asset to invest more in the risky asset, allowing the investor to leverage returns from the risky asset.

Using this version of the CPPI gives the following expression for the floor at each point in time t:

$$F_t^{\text{new}} = \begin{cases} \frac{m-p}{m} V_t & \text{if } mC_t > pV_t \\ F_t^{\text{old}} & \text{otherwise.} \end{cases}$$

The explanation for this can be found in appendix A.3.

In both versions of CPPI, the weights for the two components are given by

$$b_{t,\text{risky}} = \frac{E_t}{V_t} \quad \text{and} \quad b_{t,\text{riskfree}} = 1 - \frac{E_t}{V_t}.$$

The total return factor is calculated as

$$\hat{S}_n = \frac{V_n}{V_0}.$$

2.2.6 Algorithm summary

For convenience, an overview of the algorithms is provided in table 1 below to remind the reader of their respective parameters, desired asset features and any other algorithm features.

Table 1: Overview of the algorithms used and their features

Algorithm	Parameters	Desired asset features	Other features
Universal Portfolio	Distribution of BCRP weights	Momentum	Computationally feasible only for two asset portfolios
EG	Learning rate η	Momentum	Not computationally complex
Anticor	Window length w	Negative correlation, mean reversion	Medium computational complexity
CPPI	Multiplier m, max leverage p	One risk free asset, momentum	The floor is a guarantee level

2.2.7 Trading costs

Trading has additional costs besides the actual purchase prices, affected by relative price changes of the assets. There is also a trading cost, which is often a certain fraction of the total volume traded. This approach is referred by e.g. Blum and Kalai [3] and Black and Jones [2]. Clearly, this will primarily affect algorithms that make large changes to the portfolio weights.

In this thesis, trading costs are implemented as a reduction of a certain number of basis points (hundredths of percentage units) of the total volume of assets traded at each transaction. I will perform tests with a range of trading costs for algorithms that show capability of generating abnormal returns. This will test the robustness of the algorithms in the sense of their ability to provide satisfactory returns even with trading costs and thus how useful the algorithms are for smaller investors.

3 Data

The data used is time series of past price data for a number of different assets of several categories. Categories represented are mainly broad indices and volatility indices. Some commodities and the Swedish T-bill index OMRX are also included.

Broad indices included are STOXX with 3528 daily observations from December 1991 to November 2005, OMX with 1891 daily observations from February 1998 to June 2005, S&P500 with 19562 daily observations from January 1928 to November 2005 and MSCI and MSCI Russia with 1638 daily observations from February 1998 to June 2004.

The volatility indices used are VIX, which is the implied volatility of the S&P500 index, with 5011 daily observations from January 1986 to November 2005, VXN, which is the implied volatility of the Nasdaq composite index, with 1202 daily observations from February 2001 to November 2005 and VDAX, which is the implied volatility of the DAX index, with 3351 daily observations from January 1992 to April 2005. Volatility indices are introduced because these indices are in themselves highly volatile, and their volatility might be useful for generating high returns. In figure 1 below, the time series for the VIX index is presented to demonstrate the characteristics of a volatility index. The very high spike corresponds to October 19, 1987, also known as the Black Monday, when the S&P500 fell by more than 20% in one day. VXN and VDAX have similar characteristics and are not shown.

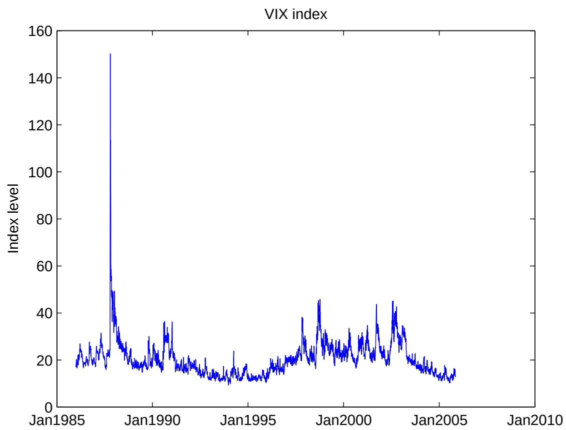

Figure 1: Development of the VIX index

Commodities included are gold and oil, both with 1638 daily observations from February 1998 to June 2004. The OMRX index, with 4023 daily observations from January 1990 to June 2005, is the only T-bill index included because T-bill indices do not differ much from one another in terms of riskyness or returns. The index has very low volatility and a low rate of return that remains nearly constant over time. It is included primarily to serve as the riskfree asset needed by the CPPI algorithm.

4 Methodology

To test the algorithms, experiments will be performed using actual time series of securities prices, but Monte Carlo simulations using the historical distribution of per period returns will also be used.

4.1 Monte Carlo simulation

Values are sampled from a (preferably large) set of empirical data with uniform probabilities for each observation. If the set is a representative sample of the data series in a longer time frame, realistic time series of virtually unlimited size can be generated to simulate what types of behavior to expect if longer actual time series were available. It should be noted that the process is likely to eliminate momentum from the original time series and might worsen performance of the algorithms that rely on momentum.

4.2 Performance measures used

Two performance measures will be used: A comparison of the return of an algorithmic strategy using a broad index such as the S&P500 to the return of the broad index itself, and the Sharpe measure, which will be used to measure economic significance[6] of performance differences.

Formally, the test performed will be to simulate 1000 time series, each of 10 years length, and to calculate the difference between the return of the algorithm and the return of the market index included in the portfolio,

$$X = r_a - r_m$$

where r_a is the geometric average annual return of the algorithmic portfolio and r_m is the geometric average annual return on the broad index which is represented in the algorithmic portfolio. Then, the hypothesis that the average of this stochastic variable is zero is tested against a two sided alternative that it is different from zero. Let μ_X denote the expected value of X. Then the hypotheses tested can be formally written as

$$H_0 : \mu_X = 0$$

$$H_1 : \mu_X \neq 0$$

This is tested using the test statistic

$$t = \frac{\overline{X} - \mu_X}{\sigma/\sqrt{n}}$$

where n is the number of observations, i.e. 1000, and σ is the standard deviation of X. Under the assumption that X is normally distributed, the statistic will follow the Student's t distribution with $n - 1 = 999$ degrees of freedom.

[6]In this thesis, economic significance is used to describe the situation where a performance difference is likely to be considered relevant by an investor, regardless of whether the difference is found to have a statistically significant deviation from zero or not.

Apart from this test, the estimated standard deviation $\hat{\sigma}$ will also be used to test the hypothesis that the return on the algorithmic portfolio when the actual time series are used is larger than the actual return on the index included in the portfolio. Formally, this can be written

$$H_0 : X = 0$$

$$H_1 : X \neq 0$$

The test statistic is then

$$t = \frac{\hat{X} - X}{\sigma}$$

which follows the Student's t distribution with one degree of freedom under the assumption that X is normally distributed. As stated, the standard deviation will be estimated from the 1000 Monte Carlo runs in each case.

The yearly Sharpe measure S_p for a portfolio with yearly return r_p and yearly volatility σ_p when the risk free rate is r_f is calculated according to

$$S_p = \frac{r_p - r_f}{\sigma_p}.$$

4.3 Experimental structure

Each of the algorithms will be tested in turn. Some tests of the sensitivity to the choice of parameters will be performed and the algorithm performance will be compared to the two benchmarks, first using simulated time series according to the description of Monte Carlo simulation and then using the actual time series. The results of the hypothesis tests will be shown and the observed distributions of the X variable will be shown in histograms to determine whether the normality assumption holds or if it is otherwise clear that any discrepancy from normality does not actually affect the validity of the hypothesis tests. For testing economic significance, the results when using real time series will be shown and the Sharpe measure will be calculated.

In case the null hypothesis is rejected for the real time series, an additional test will be run to determine at what trading cost level the result is no longer significant. This is done to determine whether the result is also economically significant, i.e. if it can be exploited not only by agents with low transaction costs but also by e.g. individual investors who typically face higher transaction costs.

4.4 Delimitations

The issue of asset liquidity will not be addressed in this thesis. It may be the case that some of the assets used cannot actually be traded according to the patterns suggested by the algorithms. No analysis will be performed to see effects of limited liquidity. Experiments will implicitly assume that necessary liquidity is present.

The algorithms will only be evaluated using daily rebalancings, since this was found by Lundahl [19] to give the best performance even in the presence of trading costs.

5 Experiments

The testing will include testing the algorithms' performance on simulated data as well as their performance on real asset price data, to find out what the optimal patterns are for the scalable algorithms (e.g. if it seems preferrable to use a particular number of assets) and to find evidence of the dependence of algorithms on the different parameters of the assets used.

5.1 Universal Portfolio

The only parameters of the Universal Portfolio algorithm are the parameters of the probability distribution assumed for the BCRP weights. Using the Dirichlet distributions as discussed earlier, it is interesting to see whether there are significant differences in performance of the algorithm using e.g. the Dirichlet(1/2,1/2) and the Dirichlet(1,1) distributions. This is tested by simulating 1000 runs of the algorithm on two randomly selected assets and calculating the difference between the portfolio total return factor and the arithmetic average total return factor of the two individual assets. This difference can intuitively be viewed as a measure of performance of the algorithm, since it shows any extra performance compared to the simple buy-and-hold portfolio on the same two assets.

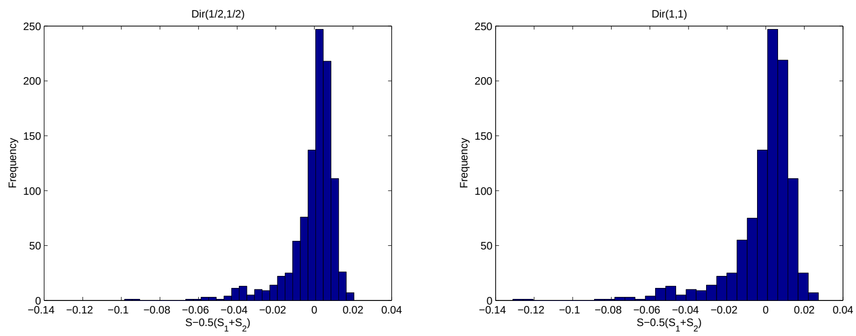

Figure 2: Comparison of Universal Portfolio algorithm using Dirichlet(1/2,1/2) (left) and Dirichlet(1,1) measures (right)

Besides the fact that these particular assets do not seem to provide particularly impressive algorithm return, it is also seems that, having settled for the Dirichlet class of probability measures, the actual choice of its parameters is not very relevant. The histograms are virtually identical. A number of other tests of different Dirichlet parameters indicate the same thing. Also, in some of the source articles, there are also tests of the Universal Portfolio algorithm with respect to the sensitivity to the choice of probability measure, reaching the same conclusion. The probability measure chosen does not appear to be an important concern. Henceforth, this thesis shall for Universal Portfolio use the Dirichlet(1,1) probability measure, which is equivalent (see appendix A.1) to the uniform distribution.

One selection of assets that seems to work well with the Universal Portfolio algorithm is the STOXX index [30] combined with the VIX volatility index [28]. This portfolio has the advantage that the VIX index, reflecting the implicit volatility of the S&P500 index, does not exhibit consistent growth over time but rather stays close to a constant level apart from occasional spikes when the S&P500 becomes highly volatile. This makes it reasonable to compare the algorithmic return only to the return on the STOXX index, since the return on the VIX index is typically close to zero over most periods, with exception only for short periods from the "base level" to the height of a peak as shown in section 3. For a portfolio based on these two assets, the distribution of the X variable is shown in figure 3 below.

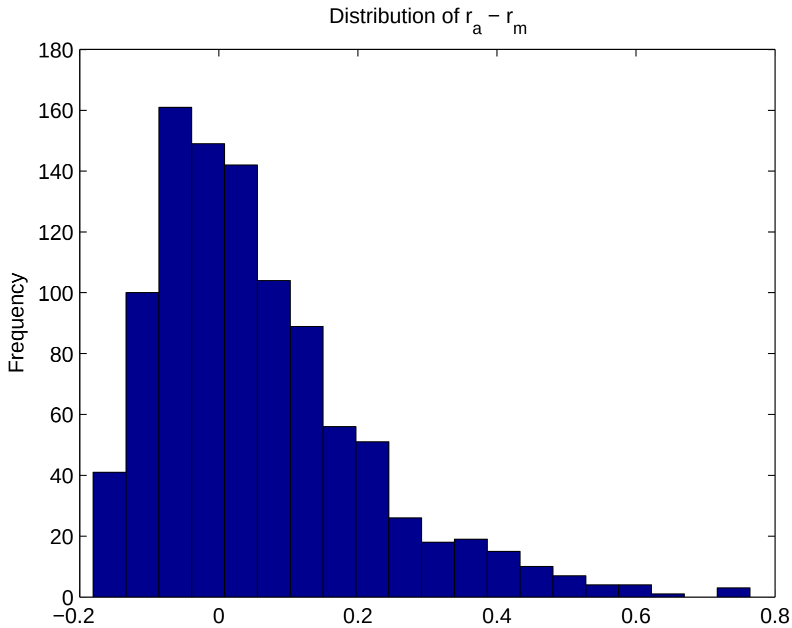

Figure 3: Histogram of the distribution of the X variable for Universal Portfolio on STOXX and VIX

The first hypothesis test gives an observed value of the t statistic that is 11.526. With 999 degrees of freedom, this corresponds to a p value of 0.0000, i.e. the null hypothesis is rejected on the 1% level. There is a significant positive effect from using the Universal Portfolio algorithm on these assets. Similar results are obtained using other assets, for instance the other volatility indices or other broad indices. Combining VIX with S&P500 improves results, as would be intuitively assumed.

It should be noted that the distribution does not look quite normal. In particular, it appears to have a fat upper tail. Other than this tail, the normality assumption appears to be justified. However, the fat upper tail does not contradict the conclusion that the algorithm's return is greater than the return on the STOXX index. Rather, it corroborates this conclusion, showing that in some cases, the average yearly return is much higher using the algorithm compared to not using the algorithm.

Testing on the actual time series of STOXX and VIX, approximately 14 years from the end of 1991 to the end of 2005, gives the result presented in figure 4 below.

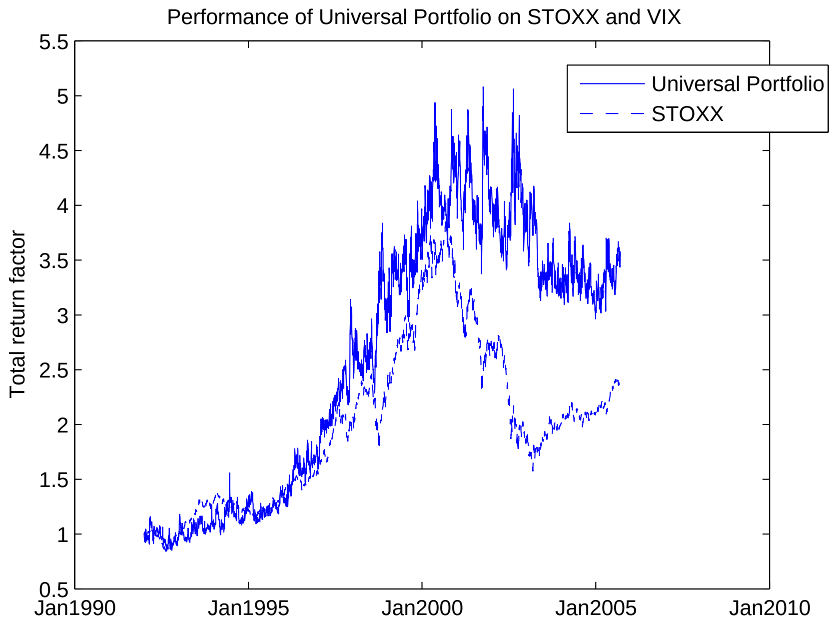

Figure 4: Performance of Universal Portfolio on actual time series compared to constituent index

From the plot, it appears that the return is clearly higher for the Universal Portfolio algorithm but that this comes at the price of higher volatility. It turns out however that the average yearly return is not significantly higher for Universal Portfolio in this case, with a t value of only 0.1885 and one degree of freedom, corresponding to a p value of 0.8814. The null hypothesis that the Universal Portfolio does not give a higher return can not be rejected at any conventional level. In terms of economic significance[7] however, investors are likely to think that earning approximately 250% is clearly better than earning approximately 150%. The Sharpe measure is 0.1805 for the Universal Portfolio and 0.1064 for the STOXX index, strengthening the view that there is an economically significant difference between the Universal Portfolio and the STOXX index.

To test the sensitivity to trading costs, a series of tests on actual time series were run with successively increasing trading costs. For the portfolios with volatility indices, trading costs of 1.5% of the traded volume at each trade brings the performance in level with the underlying broad index. At 1% of the traded

[7]Recall from the methodology section that economic significance is used in this thesis to denote performance differences that investors are likely to find relevant.

volume, there still appears to be economically significant excess return. Large financial institutions will typically face 0.02-0.03% trading costs according to the author's work life experience.

5.2 Exponentiated Gradient

The Exponentiated Gradient (EG) algorithm has one parameter, the learning rate η. The sensitivity to the choice of η will be tested first by calculating returns for different values of this parameter. The time series used are the STOXX and VIX indices also used in the Universal Portfolio experiments.

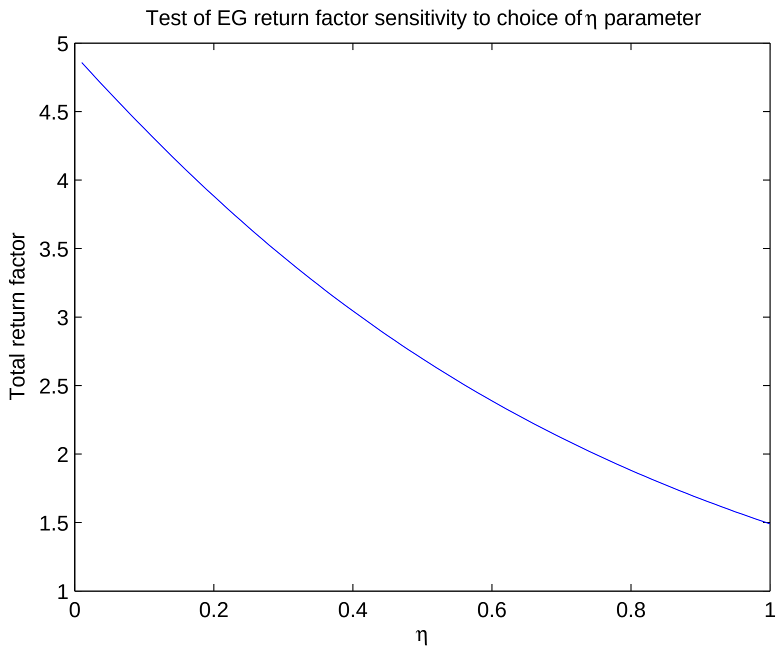

Figure 5: EG total return factor as a function of η

The results are in line with what Helmbold et al [13] find. A low η seems to be the best choice. This corresponds to a rather slow rebalancing process, while a large η means reallocating very quickly to the strong performers in the portfolio. An η value of 0.01 will be used for the remainder of the EG experiments.

Since the EG algorithm is not very computationally complex, it is suitable for tests on portfolios of more than two assets. Experiments show that portfolios of more than two assets do not provide any additional performance compared to portfolios of two assets. In figure 6 is the result of one such experiment. Using a set of N assets, it is possible to construct $2^N - N - 1$ portfolios of at least two assets[8], and in this particular experiment, nine assets were used to test a total

[8]It is possible to construct a total of 2^N portfolios (each of the N assets can either be included or not included, independently of one another) and of these, one is the empty portfolio with no assets and N portfolios have one asset each.

of 502 portfolios.

In the figure, the total return factors of the portfolios are plotted against the *portfolio number*, which is originally the decimal representation of the nine bit binary string used to select the assets from the set of nine assets. Note that in the figure, the original order of the portfolio numbers has been changed to sort the portfolios according to number of constituent assets keeping the portfolios with few components to the left. This is intended to help drawing conclusions about whether there are any benefits when using a particular number of assets.

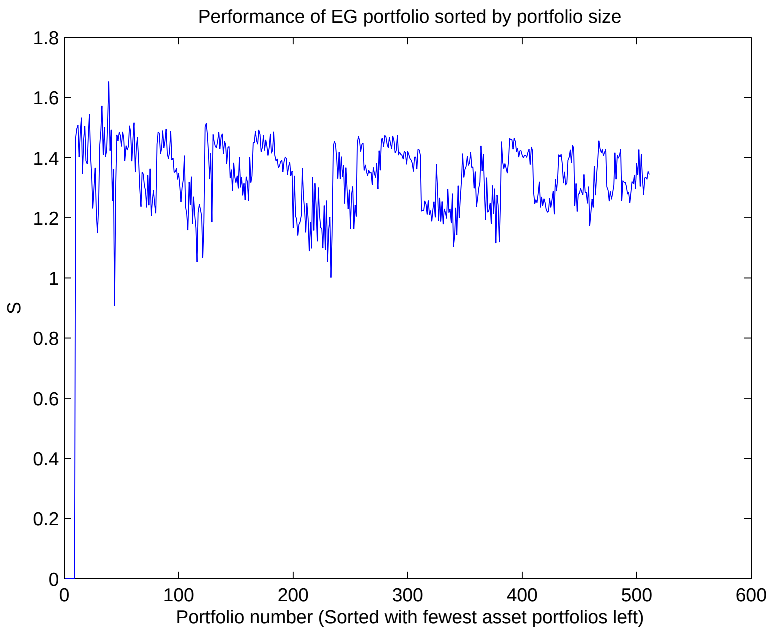

Figure 6: Total return factor for 502 EG portfolios, sorted by number of constituent assets

There seems to be no significant benefit from including more than two assets in a particular portfolio. Return factors appear to be largely unaffected while the volatility of the return appears to decline as more assets are included. The volatility conclusion is based on the fact that there seems to be smaller fluctuations in the total return factor in the right part of the figure, where the portfolios containing several assets are represented. This should not come as a surprise, but rather be seen as a sign of the diversification effect. More assets should imply lower volatility, even if the weights are determined according to an algorithm. The results in figure 6 are also confirmed by several other tests on other sets of assets and over different time periods.

As pointed out earlier, the conceptual similarity between EG and Universal Portfolio can be used to argue that heuristically, the same result could be assumed to hold for Universal Portfolio. For the remainder of the thesis, experiments will be performed on two asset portfolios. Conceptually, an investor that wants to exploit diversification to decrease volatility could do so by dividing investment among several separate algorithmic portfolios rather than creating a single algorithmic portfolio of more assets. In conclusion, focusing on two asset portfolios is not a severe limitation.

When running a 1000 run experiment on the EG algorithm using the STOXX and VIX indices, the observations of the X variable are distributed according to the histogram in figure 7.

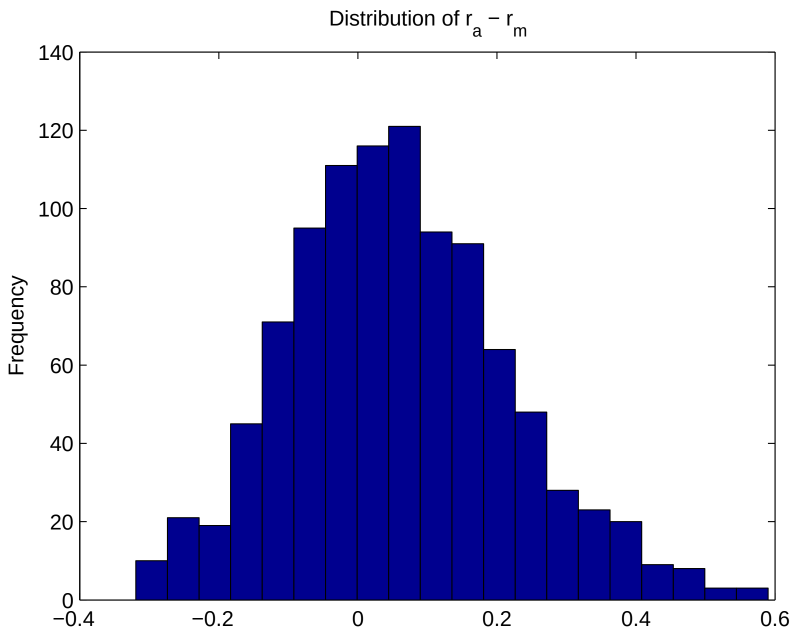

Figure 7: Histogram of the distribution of the X variable for the EG algorithm on STOXX and VIX

The first hypothesis test results in an observed t statistic value of 11.895. The statistic has 999 degrees of freedom, meaning that the observed value corresponds to a p value of 0.0000. The null hypothesis is rejected on the 1% level of significance. There is on average a positive contribution to yearly return from using the EG algorithm on these assets.

The histogram looks similar to the one observed for the Universal Portfolio algorithm. The right tail looks perhaps somewhat less heavy than in the Universal Portfolio test, although it is still heavier than the left tail in this histogram. The shape of the distribution suggests that the normality assumption is not unjustified. There appears to be a skew, but that skew strengthens the rejection of the null hypothesis rather than contradicts it. There appears to be a significant positive effect from using the EG algorithm.

The test performed using the actual time series gives a t statistic of 0.3393, with one degree of freedom, which corresponds to a p value of 0.7918. The null hypothesis of no contributing effect can in this case not be rejected on any conventional level of significance. In figure 8 below, the performance of EG using the actual time series is shown.

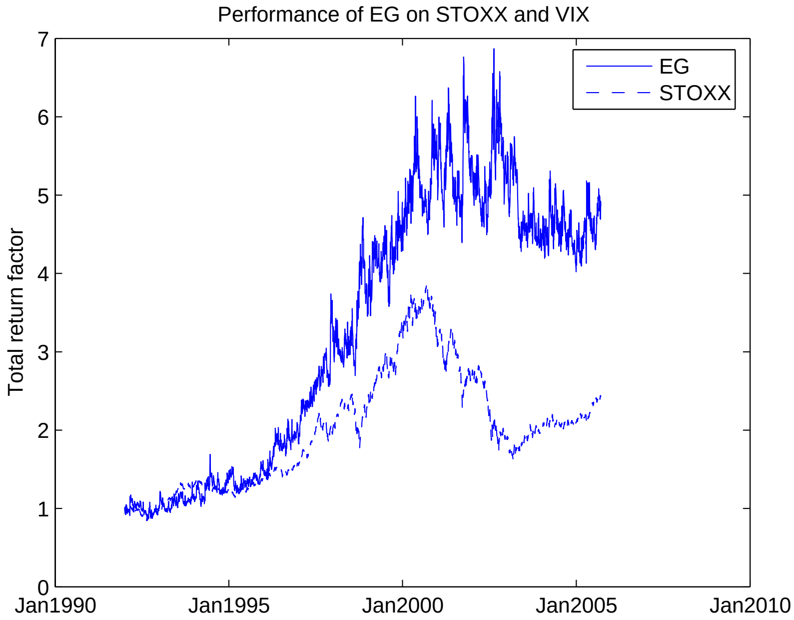

Figure 8: Performance of EG on actual time series compared to constituent index

Even more than in the Universal Portfolio case, there seems to be economic significance in the results presented in figure 8. Almost 400% return should be considered significantly better than almost 150% return. For comparison, the Sharpe measure for this EG portfolio is 0.2825, almost three times the Sharpe measure of the STOXX index which was calculated to be 0.1064 in the previous section. Results are similar for all tested portfolios, especially those containing volatility indices, and the S&P500 and VIX combination again seems to be the best combination. The STOXX and VIX combination is representative of the typical performance of portfolios containing at least one volatility index. Portfolios of only other assets, for instance combinations of broad indices, do not exhibit the same overperformance, but does show an ability to give a return higher than the arithmetic average of the constituent asset returns combined with a lower volatility than an equally weighted buy-and-hold portfolio has.

The EG algorithm with a low learning rate η proves to make very small weight adjustments at each rebalancing and, as a result, the algorithm is almost completely insensitive to trading costs. Even trading costs of 50% of the traded volume still leaves an economically significant abnormal return. It takes trading costs of 100% of the traded volume to bring the portfolio return in level with the underlying index, indicating that the reallocations the algorithm makes on average are return generating.

5.3 Anticor

The Anticor algorithm has one parameter, namely the length of the window which is denoted by w. As a first step, the sensitivity of the algorithm to this parameter will be tested by calculating total return factors for the STOXX-VIX portfolio for a range of different values of w. In figure 9 below, the results of this test can be seen.

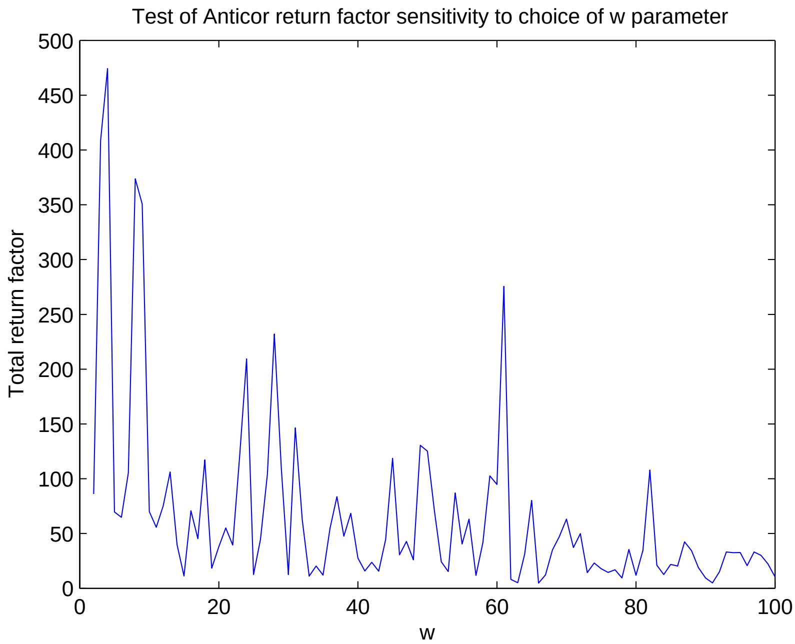

Figure 9: Anticor total return factor as a function of w

As can be seen, the results seem quite sensitive to the choice of w. This is in accordance with what Borodin et al [4] find. Their suggested solution is to instead choose one minimum and one maximum w, calculate the suggested Anticor weights for each w between these two values and take the arithmetic average of the suggested weights, essentially creating a buy-and-hold portfolio with a number of Anticor portfolios as constituents. Very low values of w should be avoided since the suggested weights can differ significantly from suggested weights when a slightly larger w is used (and the result is typically worse for very low w according to Borodin et al).

The minimum value of w is set to 5, and the obtained results are similar to those presented by Borodin et al, i.e. the algorithm is not so sensitive to the maximum value of w as long as it is at least somewhere around 20. In the remainder of this thesis, averaging will be used with $w_{\min} = 5$ and $w_{\max} = 30$.

The first experiment on Anticor using 1000 runs of simulated time series gives the histogram of the X variable distribution presented in figure 10.

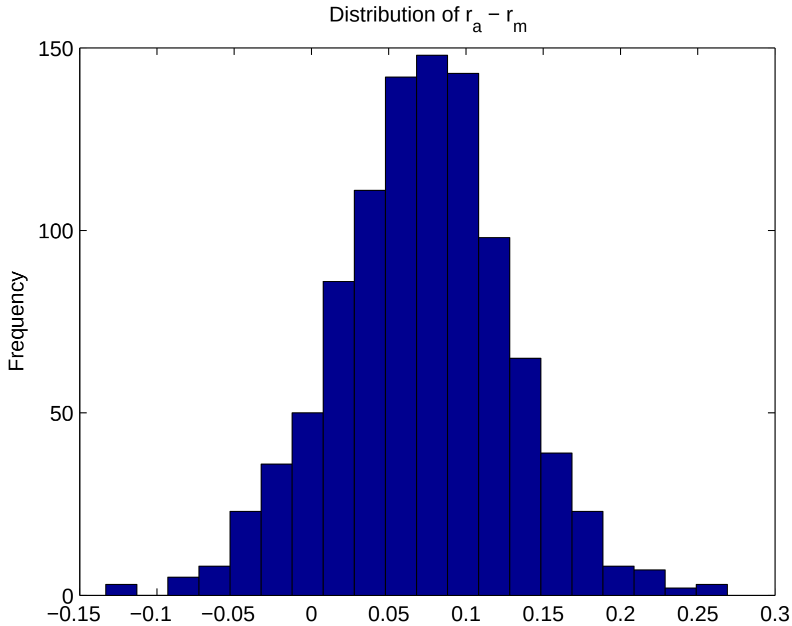

Figure 10: Histogram of the distribution of the X variable for Anticor on STOXX and VIX

Like before, the normality assumption appears from the shape of the distribution histogram to be justified, and the mean appears to be well above zero. The observed value of the t statistic in this first test is 39.481, and with 999 degrees of freedom as before, this corresponds to a p value of 0.0000. The hypothesis test indicates a significant positive contribution from using the Anticor algorithm compared to investing only in the STOXX index. Similar results are seen using other combinations including at least some volatility index.

The portfolio on the actual STOXX and VIX time series over the same time period as before gives the result presented in figure 11 below.

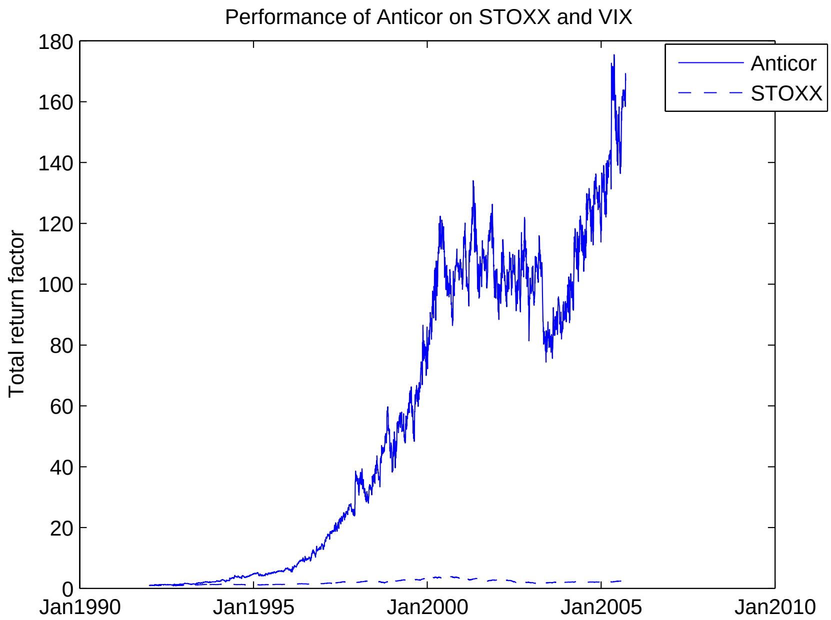

Figure 11: Performance of Anticor on actual time series compared to constituent index

As can be seen from the figure, Anticor manages to provide an abnormal return that is quite impressive. The t statistic has a value of 6.5507 with one degree of freedom, corresponding to a p value of 0.0964. Thus, the null hypothesis can be rejected on the 10% level of significance. As for economic significance, the performance of Anticor is clearly much better than that of the underlying STOXX index. The Sharpe ratio for the Anticor portfolio is 1.5759, much higher than the 0.1064 seen for the STOXX index.

This pair of assets clearly fits the desired characteristics for the Anticor algorithm well. There is high volatility in the VIX volatility index, and it is negatively correlated with the STOXX index. This negative correlation is not perfectly intuitive, but there is at least one possible explanation. The VIX index reflects the implicit volatility of the S&P500 index. Thus, a negative correlation should be expected between the returns on S&P500 and the returns on VIX, since when S&P500 is falling, the volatility tends to increase, and when it is rising, volatility tends to fall. Furthermore, while the STOXX index is made up of European components in contrast to the S&P500 which is made up of American components, there is a correlation between the behavior of the STOXX index and that of the S&P500. This also means that a negative correlation can be found between STOXX and VIX.

For reference, some results when combining VIX with the S&P500 are also given. In figure 12 below are the results when using the actual time series over a period of almost 20 years, from the beginning of 1986 to present day.

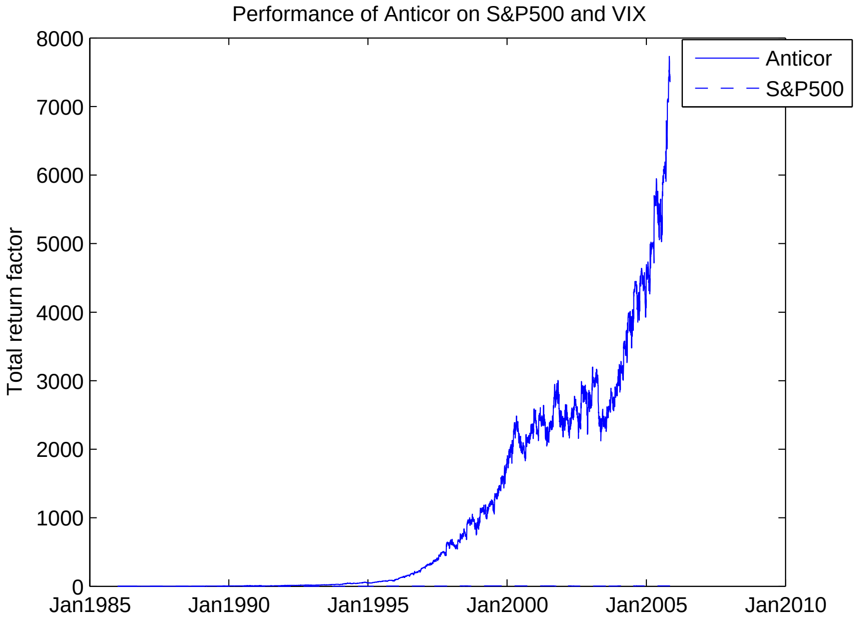

Figure 12: Performance of Anticor on actual time series compared to constituent index

Note that the S&P500 actually is plotted in the graph and normalized to start at 1 in the beginning of 1986 just as the total return factor of Anticor, but due to the performance of the Anticor algorithm, the S&P500 series is not visible since the comparatively small magnitude of its returns makes it blend together with the horizontal axis. It appears that the effect is even stronger when combining the VIX with the S&P500, which is to be expected since the S&P500 is the underlying index whose implicit volatility is reflected by the VIX. A hypothesis test of the same type as above gives a t statistic value of 13.8942, which corresponds to a p value of 0.0457. The null hypothesis can be rejected on the 5% level of significance. The Sharpe measure is 3.4848 for this portfolio, to be compared with 0.1471 for the S&P500. The volatility is 23.86%, which is higher than the 8.61% that the S&P500 exhibits.

The combination of S&P500 and VIX gives the best results with the Anticor algorithm, but in general, all portfolios tested that contain at least one volatility index exhibit behavior similar to the STOXX-VIX combination. Unlike for the Universal Portfolio and EG algorithms, Anticor also outperforms the individual assets in portfolios of only broad indices. The overperformance in this case is not as large as in the portfolios that have volatility indices, but there is a positive effect that is comparable to what Universal Portfolio and EG show in portfolios with volatility indices.

5.3.1 Anticor under extreme conditions

As a test of the Anticor algorithm's behavior under extreme conditions, the performance over 1987 is presented in figure 13 below. The large drop in the S&P500 corresponds to October 19, also known as the Black Monday.

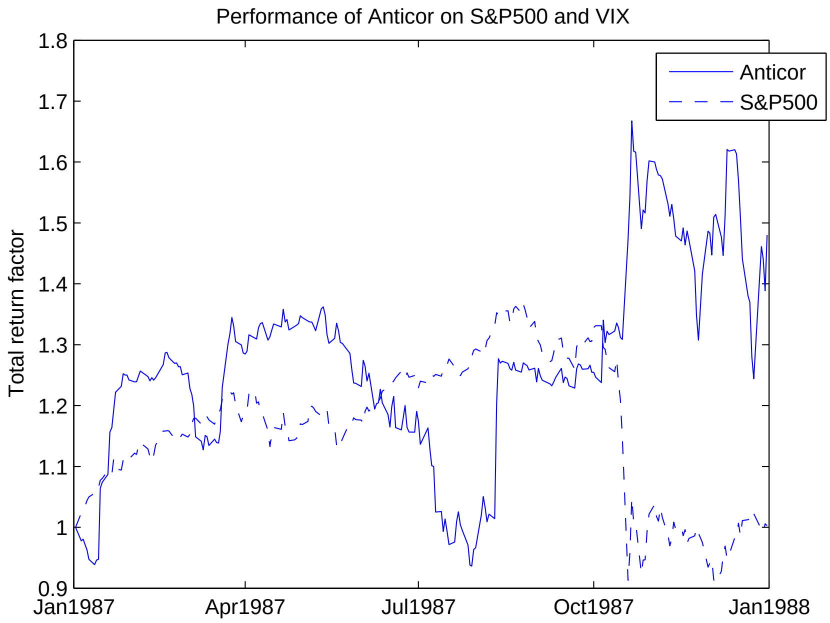

Figure 13: Performance of Anticor under extreme conditions

In this figure, it is especially interesting to note that while there appears to be some volatility at the end of the year, the Anticor algorithm makes a rather large profit on the Black Monday, successfully capturing the sharp spike in the VIX index that occurred on that day and is shown in figure 1 in section 3. The whole year performance is approximately up 40%, to be compared with the S&P500 index whole year return of 0.25%.

While the Anticor algorithm has rather strong performance, it achieves this performance through rather large reallocations at each rebalancing. As a consequence, much of the abnormal returns is lost if trading costs are high. At trading costs of 1% of the traded volume, Anticor has a performance comparable to that of Universal Portfolio or EG at low trading costs. When trading costs are 1.2%, the algorithm performance is brought to level with the underlying index.

5.4 CPPI

The CPPI algorithm will be used in the moving floor version only, since it is a generalization of the basic version. The floor will not be allowed to grow at the risk free rate, since this was found by Lundahl [19] to be undesirable, causing the portfolio to fall through the floor too often. There are two parameters for this algorithm, the multiplier m and the maximum relative investment in the risky asset p. Since the CPPI algorithm is not intended to be an abnormal return generating algorithm like the previous three tested, but rather an insurance algorithm providing a safety level combined with potential upside, the same types of results should not be expected for CPPI. Also, it is specifically constructed for two asset portfolios where one of the assets is reasonably risk free. For this purpose, the Swedish OMRX T-bill index will be used in this thesis since it has very low volatility and a predictable, nearly constant growth rate.

Tests were run to determine reasonable values for the m and p parameters. The p parameter proved to have a rather irregular influence on the resulting total return factor as can be seen in figure 14 below.

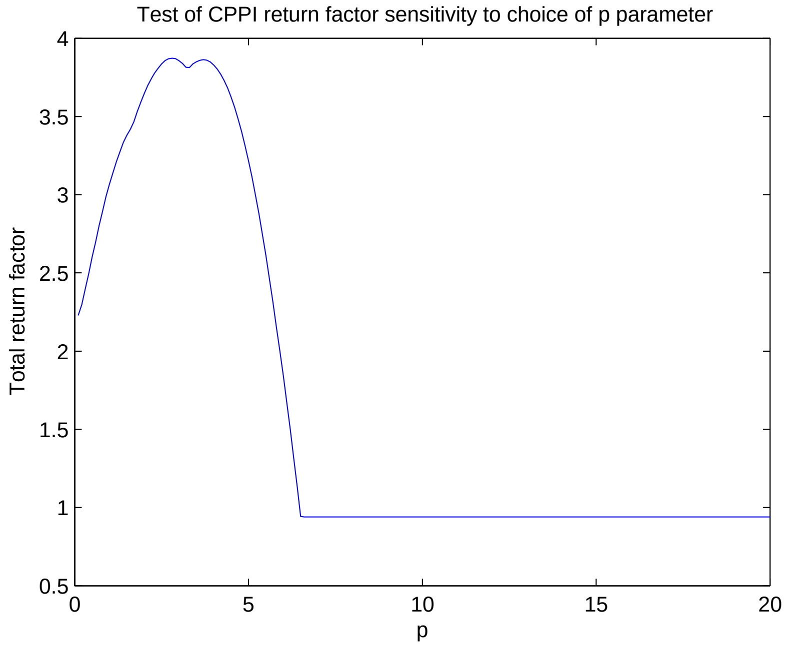

Figure 14: CPPI total return factor as a function of p

Considering the meaning of the p parameter, the above pattern is not very surprising. The parameter denotes the maximum investment in the risky asset as fraction of the total portfolio value. If p is 1, the entire value can be put in the risky asset and if $p > 1$, it is possible to short sell the risk free asset to increase investment into the risky asset. Thus, if p is allowed to be very large, massive short selling can occur, and if the risky asset decreases in value, the portfolio sharply decreases in value since it is highly levered. It appears that $p = 3$ would be a good choice.

For the m parameter, a similar test was made, both before and after the test for p, to ensure that the m value suggested before determining a suitable p provides good results also with the p that was chosen in the test for the p parameter. In figure 15 below, the results of the second test are shown. The results of the first test are very similar and are thus not shown here.

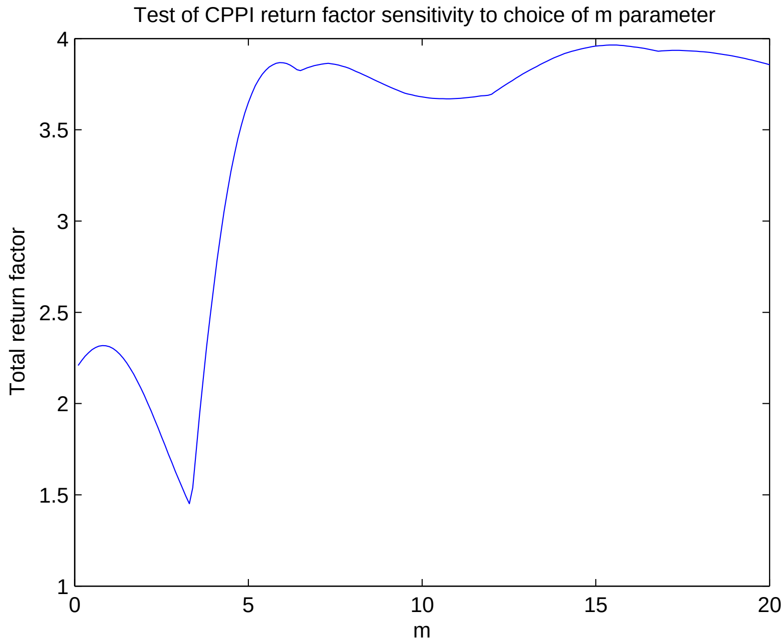

Figure 15: CPPI total return factor as a function of m

It appears that as long as $m > 5$, the actual value does not affect return very much. This is reasonable, since p specifies an upper limit on the investment in the risky asset. Increasing m would increase the suggested exposure to the risky asset, but if it already is larger than the maximum allowed exposure, it does not matter if m is increased. The actual exposure will be the one specified by the value of p. In this thesis, $m = 6$ will be used. Higher values of both m and p could be used to amplify returns, but that would also amplify risks, since higher p values means that it is possible to short sell the riskfree asset more, and higher m values means that the sensitivity for drops in the risky asset is increased, as is shown in appendix A.3.

Using the parameters mentioned above and a portfolio consisting of the STOXX and OMRX indices for the 1000 runs provides the histogram presented in figure 16 below.

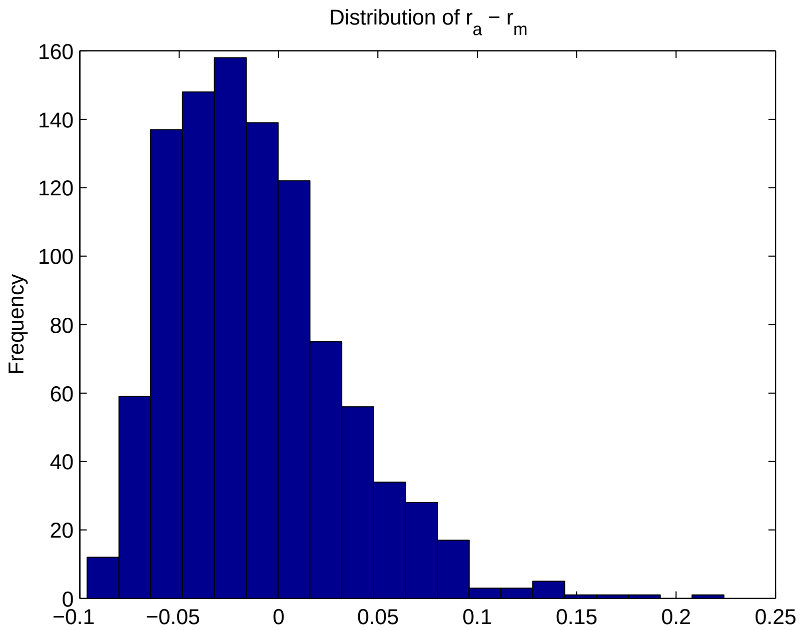

Figure 16: Histogram of the distribution of the X variable for CPPI on STOXX and OMRX

The observed value of the t statistic is -8.4851, corresponding to a p value of 0.0000 with 999 degrees of freedom. There appears to be a significant *negative* effect on return if CPPI is used compared to investing only in the STOXX index. There is however a heavy right tail as for the other algorithms. The shape of the distribution resembles the normal distribution.

When the actual STOXX and OMRX time series are used, with the same time period as the previous experiments, the result is as in figure 17 below. The value of the t statistic in this case is 0.8235 with one degree of freedom, which corresponds to a p value of 0.5614. The null hypothesis can not be rejected on any conventional level of significance. The behavior of the algorithm is similar regardless of what index is used instead of STOXX. The algorithm is not very well suited for other types of assets, so only pairs of OMRX and a broad index have been tested.

The CPPI algorithm is very sensitive to presence of trading costs. Apart from the empirical fact that CPPI proves to make large reallocations at each rebalancing, the sensitivity can intuitively be understood, since an important feature of the algorithm is to raise the floor to secure returns. If trading costs must be paid, there is not room for raising the floor, and subsequent falls will affect the portfolio more. In fact, for trading costs of only 0.20% of the traded volume, the portfolio performance is at the same level as the underlying index.

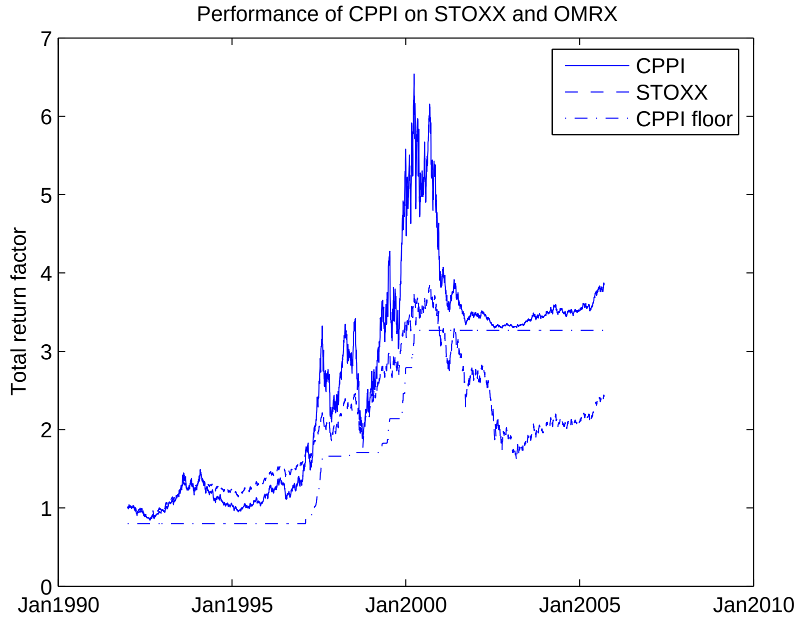

Figure 17: Performance of CPPI on actual time series compared to constituent index

The Monte Carlo simulation technique does not quite capture the momentum characteristic of the STOXX index, meaning that the potential for CPPI to generate abnormal returns is significantly limited. Also, another point of the CPPI algorithm is the guarantee provided by the floor. In appendix A.3, it is shown that if the risky asset does not fall by a fraction larger than $1/m$ over one period, the portfolio will not break through the floor. Thus, as is shown in figure 17 above, CPPI would provide a *guaranteed* return that matches the STOXX level at the height of the IT bubble, provided that STOXX never falls more than $1/6 \approx 0.1667 = 16.67\%$ over one period, i.e. over one day with daily rebalancing. Over the 14 year period used, the maximum daily drop in STOXX was 5.39%, so clearly, falls of more than 16% appear to be rare. Even if the portfolio falls below the floor, it will then allocate fully into the risk free asset and remain invested in it until wealth is above the floor again.

The Sharpe measure for the CPPI portfolio is 0.3885, which is clearly higher than the Sharpe measure for the STOXX index. The average annual volatility is 11.32%, to be compared with 8.03% for the STOXX index. Combined with the added insurance that the portfolio will not fall below the floor unless the STOXX index falls by more than 16% in one day, it appears that the CPPI algorithm is indeed useful in reality despite the negative outcome of the Monte Carlo simulation. The fact that CPPI structures are provided by large investment banks also strengthens this view.

6 Analysis

The experiments for the Universal Portfolio, EG and Anticor algorithms show clearly significant positive results in Monte Carlo simulations. The CPPI algorithm shows a significant negative result in Monte Carlo simulations, but the corresponding result when using actual time series indicates that certain characteristics of the underlying time series that are important to the CPPI algorithm might not be captured by the Monte Carlo simulation.

For all algorithms, real time series results show economic significance, while only Anticor manages to provide statistical significance at conventional levels for these tests. Also, the Sharpe measures observed for the different algorithms indicate that the algorithms manage to provide performance that is superior to the underlying broad indices.

A natural question is why the EMH, which is certainly plausible on an intuitive level, appears to fail to hold even in the weak form. A possible explanation is that while the EMH implicitly assumes perfect distribution of information to allow for enlightened and equally knowledgeable investors, in reality not all investors might have access to all information. While the time series of past price data certainly are available to a large share of investors and technical analysis in many forms is commonly used, the knowledge of particular quantitative techniques might not be widely available. Also, the very existance of the EMH might discourage many investors from even trying to use insights from technical analysis. Even for investors willing to use technical analysis, complexity of implementation might discourage the use of certain algorithms in favor of simpler methods.

It should also be noted that one of the delimitations made in this thesis is that asset liquidity is not researched. While the broad market indices such as the S&P500 and the STOXX are generally known to be highly liquid, volatility indices such as VIX or VDAX might not be widely traded. Futures contracts do exist for these indices, but they might not be widely traded. If liquidity is limited, the actual usefulness of the results, at least for large scale investment purposes, might be significantly reduced.

The algorithmic portfolios might be perceived as more risky by some investors. However, apart from the potential liquidity risk, there is not much evidence of significantly increased risk. Observed volatilities of the algorithmic portfolios are not substantially larger than those of the broad market indices themselves and certainly not larger than typical single stock volatilities, indicating that volatility is not a major concern.

Trading costs is an issue that might erode the economic significance of the results. The experiments show that in particular CPPI is sensitive to trading costs and might thus be best suited for implementation in larger financial institutions. In the presence of high trading costs, Anticor loses most of its abnormal performance, meaning that it might not be suitable for small investors. EG proves to be very resistant to trading costs and might be interesting even for individuals.

7 Conclusions

Despite the plausibility of the EMH, technical analysis seems to improve investment performance. Clearly significant results in large simulations indicate robustness of several purely technical investment strategies. Experiments on actual time series do not consistently provide convincing statistical significance in themselves, but evidence from the execution of hypothesis tests indicate that there is an economically significant effect from using algorithmic techniques. Also, it is clear that abnormal returns need to be very high to be statistically significant in single tests on actual time series. Thus, economic significance can clearly exist even in the absence of statistical significance.

Some concerns about the validity of the results can be raised, and among these, the liquidity issue is arguably the most potentially troubling, possibly along with the issue of trading costs. With low liquidity or high trading costs, scalability and profitability might be limited.

In conclusion, within the delimitations of this thesis, there is strong evidence that the EMH *does not* hold fully considering the performance of these algorithmic techniques. Assuming necessary liquidity and trading costs not exceeding specified levels, there appears to be potential for persistent and economically significant abnormal returns for the algorithms tested in this thesis.

7.1 Suggestions for further research

Limited length of actual time series data gives rise to concerns about sustainability of the results over longer periods of time. While Monte Carlo simulations can reduce this concern, it is not clear that the simulated time series mimic actual time series closely enough to be realistic. For instance, serial correlations might not be captured. As time passes and longer actual time series are possible, tests could be run again to determine in hindsight whether there is any sustainable profitability for any particular algorithmic strategy.

Another interesting topic would be to evaluate the psychological implications of using algorithmic investment strategies. Investment managers might want to add an element of fundamental analysis to create a strategy that does not unconditionally follow the algorithm suggestions but under certain conditions allows non-automated input as support for investment decisions. Helmbold et al [13] expand on this by introducing the concept of side information. Hybrid strategies of this kind is an interesting extension of the algorithmic approach and could provide many further research topics.

The issue of liquidity might be addressed by either introducing significantly increased trading costs for particular assets. Also, if the portfolio wealth needs to be kept small to alleviate liquidity problems, it could be interesting to round algorithm weights to multiples of some percentage[9] to handle problems with large contract sizes and see if the observed patterns remain.

[9] If, for instance, contracts in a particular asset can only be integer multiples of 100 000 and liquidity is low, a portfolio wealth of 1 000 000 could be kept as a means of ensuring that necessary liquidity is available, but in this case, weights for the non-liquid asset could only be integer multiples of 100 000 as a fraction of 1 000 000, i.e. 10%.

References

Literature

[1] P. Bertrand and J.-L.Prigent, *Portfolio Insurance Strategies: OBPI versus CPPI*, Working Paper, Université de Cergy, 2002.

[2] F. Black and R. Jones, *Simplifying Portfolio Insurance*, The Journal of Portfolio Management, 1987 48-51.

[3] A. Blum and A. Kalai, *Universal Portfolios With and Without Transaction Costs*, Annual Workshop on Computational Learning Theory, Carnegie Mellon University, May 1997.

[4] A. Borodin et al, *Can We Learn to Beat the Best Stock*, Journal of Artificial Intelligence Research 21 (2004) 579-594.

[5] J.-F. Boulier and A. Kanniganti, *Expected Performance and Risks of Various Portfolio Insurance Strategies*, 5th AFIR International Colloquium.

[6] T.M. Cover, *Universal Portfolios*, Mathematical Finance, 1(1): 1-29, January 1991. Updated: Stanford University, October 1996.

[7] T.M. Cover and D. Julian, *Performance of Universal Portfolios in the Stock Market*, Proceedings of IEEE International Symposium on Information Theory, Sorrento, Italy, June 2000.

[8] T.M. Cover and E. Ordentlich, *Universal Portfolios with Side Information*, IEEE Transactions on Information Theory, vol. 42, no. 2, March 1996.

[9] E. F. Fama, *The Behavior of Stock-Market Prices*, Journal of Business, Volume 38, Issue 1, 1965, 34-105

[10] E. F. Fama, *Efficient Capital Markets: A Review of Theory and Empirical Work*, Journal of Finance, Volume 25, Issue 2, 1970, 383-417

[11] E. F. Fama *Efficient Capital Markets: II*, Journal of Finance, Volume 46, Issue 5, 1991, 1575-1617

[12] E. F. Fama, *Market efficiency, long-term returns, and behavioral finance*, Journal of Financial Economics, 49, 1998, 283-306

[13] D.P. Helmbold et al, *On-Line Portfolio Selection Using Multiplicative Updates*, Mathematical Finance, 8(4):325-347, 1998.

[14] G.N. Iyengar and T.M. Cover, *Growth Optimal Investment in Horse Race Markets with Costs*, IEEE Transactions on Information Theory, vol. 46, no. 7, November 2000.

[15] M. Jensen, *The Performance of Mutual Funds in the Period 1945-64*, Journal of Finance, vol. 23, 1969, 389-416

[16] A. Kalai and S. Vempala, *Efficient Algorithms for Universal Portfolios*, Journal of Machine Learning Research 3 (2002) 423-440.

[17] A. W. Lo and A. C. MacKinlay, *A Non-Random Walk Down Wall Street*, Princeton University Press, 1999

[18] A. W. Lo, H. Mamaysky and J. Wang, *Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation*, Journal of Finance, 55, 2000, 1705-1765

[19] M. Lundahl, *Evaluation of Automated Portfolio Rebalancing Algorithms*, Master's Thesis, KTH Stockholm 2005

[20] B. G. Malkiel, *A Random Walk Down Wall Street*, New York: W. W. Norton & Co., 1973, 1st edition

[21] B. G. Malkiel, *Reflections on the Efficient Market Hypothesis: 30 Years Later*, The Financial Review 40, 2005

[22] B. G. Malkiel, *The Efficient Market Hypothesis and Its Critics*, Princeton University, CEPS Working Paper No. 91, April 2003

[23] E. Ordentlich and T.M. Cover, *The Cost of Achieving the Best Portfolio in Hindsight*, Mathematics of Operations Research, 23(4):960-982, November 1998.

[24] A. Perold and W.F. Sharpe, *Dynamic Strategies for Asset Allocation*, Financial Analysts Journal, January-February 1988 16-27.

[25] R. J. Shiller, *Irrational Exuberance*, Princeton University Press, 2000

[26] A. Shleifer and L. H. Summers, *The Noise Trader Approach to Finance*, Journal of Economic Perspectives 4, no. 2, 1990 19-33

[27] F. Shostak, *In Defense of Fundamental Analysis: A Critique of the Efficient Market Hypothesis*, Review of Austrian Economics 10, no. 2, 1997,

Internet sources

[28] Chicago Board Options Exchange, Available at [2005-11-19]: http://www.cboe.com/

[29] Investor Home, Available at [2005-11-10]: http://www.investorhome.com/

[30] STOXX Limited, Available at [2005-11-19]: http://www.stoxx.com/index.html

A Appendices

A.1 Derivation of Universal Portfolio weight formulas

As noted, this thesis uses the two variable Dirichlet distribution for Universal Portfolio. To conveniently describe the two variable Dirichlet distribution, the *Gamma function* must first be defined:

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x}\, dx, \quad t \geq 0$$

A property of this function is that if t is an integer, $\Gamma(t) = (t-1) \cdot \Gamma(t-1)$, and it is easily seen by solving the integral that $\Gamma(1) = 1$. Using the Gamma function, the *Beta function* can be written

$$B(t,u) = \frac{\Gamma(t)\Gamma(u)}{\Gamma(t+u)}, \quad t,u \geq 0$$

The two variable Dirichlet probability distribution has two parameters α_1 and α_2 and is characterised by the probability density function

$$f(\mathbf{b}) = \frac{\Gamma\left(\sum_{i=1}^2 \alpha_i\right)}{\prod_{i=1}^2 \Gamma(\alpha_i)} b_1^{\alpha_1 - 1} b_2^{\alpha_2 - 1} = \frac{1}{B(\alpha_1, \alpha_2)} b_1^{\alpha_1 - 1} b_2^{\alpha_2 - 1}.$$

In addition the restrictions $b_1 + b_2 = 1$ and $0 \leq b_1, b_2 \leq 1$ must hold, which implies that the Universal Portfolio algorithm will not be able to short sell any assets (since weights below zero are prohibited).

In the special case where $\alpha_1 = \alpha_2 = 1$, the density function becomes

$$f(\mathbf{b}) = \frac{\Gamma(1+1)}{\Gamma(1)^2} b_1^{1-1} b_2^{1-1} = \frac{\Gamma(2)}{\Gamma(1)^2} b_1^0 b_2^0 = 1,$$

i.e. the uniform distribution (since $\Gamma(2) = 1 \cdot \Gamma(1) = 1 \cdot 1 = 1$). If instead $\alpha_1 = \alpha_2 = 1/2$ the distribution becomes

$$f(\mathbf{b}) = \frac{\Gamma(1/2 + 1/2)}{\Gamma(1/2)^2} b_1^{1/2-1} b_2^{1/2-1} = \frac{\Gamma(1)}{\Gamma(1/2)^2} b_1^{-1/2} b_2^{-1/2} = \frac{1}{\pi} b_1^{-1/2} b_2^{-1/2}$$

since $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(1) = 1$.

Next, a derivation of a closed form solution of the obtained total return factor, $\hat{S}_n$, will be performed for general α_1 and α_2, since it is contained in the expression for the Universal Portfolio weights (2).

First, the general Universal Portfolio weight formula is restated for convenience of reference.

$$\hat{\mathbf{b}}_k = \frac{\int \mathbf{b} \prod_{t=1}^{k-1} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t \, d\mu(\mathbf{b})}{\int \prod_{t=1}^{k-1} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t \, d\mu(\mathbf{b})}, \tag{4}$$

To explicitly calculate the algorithm weights and the total return for Universal Portfolio in the case when portfolios are limited to two assets, it is necessary to find a closed form solution to the integrals in equation (4). Firstly, to reach an expression for $\hat{S}_n$, rewrite the total return factor for an arbitrary CRP after n rebalancing time points:

$$S_n(\mathbf{x}^n, \mathbf{b}) = \prod_{t=1}^{n} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t = \prod_{t=1}^{n} (b_1 x_{t1} + b_2 x_{t2}) = \sum_{J \in \{1,2\}^n} \prod_{t=1}^{n} b_{j_t} x_{tj_t} =$$

$$\sum_{l=0}^{n} b_1^l b_2^{n-l} \left(\sum_{J \in T_n(l)} \prod_{t=1}^{n} x_{tj_t} \right),$$

where $T_n(l)$ is the set of all sequences $J \in \{1,2\}^n$ with l occurences of 1 and $(n-l)$ occurences of 2. Next, let

$$X_n(l) = \sum_{J \in T_n(l)} \prod_{t=1}^{n} x_{tj_t}, \tag{5}$$

which gives

$$S_n(\mathbf{x}^n, \mathbf{b}) = \sum_{l=0}^{n} b_1^l b_2^{n-l} X_n(l).$$

To calculate the integrals in equation (4) above and the realized total return factor of the Universal Portfolio, this expression is integrated over the Dirichlet probability measure that was assumed to hold for the BCRP weights.

$$\hat{S}_n(\mathbf{x}^n) = \int \sum_{l=0}^{n} b_1^l b_2^{n-l} X_n(l)\, d\mu(\mathbf{b}) = \sum_{l=0}^{n} X_n(l) \int b_1^l b_2^{n-l}\, d\mu(\mathbf{b}).$$

By introducing the notation

$$C_n(l) = \int b_1^l b_2^{n-l}\, d\mu(\mathbf{b}), \tag{6}$$

the total return factor expression can be written

$$\hat{S}_n(\mathbf{x}^n) = \sum_{l=0}^{n} X_n(l) C_n(l).$$

Using this result, a closed form expression for the Universal Portfolio weights and also its realized total return factor at time n can be obtained. Continuing from equation (4),

$$\hat{\mathbf{b}}_n = \frac{\int \prod_{t=1}^{n-1} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t \mathbf{b}\, d\mu(\mathbf{b})}{\hat{S}_{n-1}(\mathbf{x}^{n-1})} = \frac{1}{\hat{S}_{n-1}(\mathbf{x}^{n-1})} \begin{bmatrix} \int b_1 \prod_{t=1}^{n-1} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t\, d\mu(\mathbf{b}) \\ \int b_2 \prod_{t=1}^{n-1} \mathbf{b}^{\mathrm{T}} \mathbf{x}_t\, d\mu(\mathbf{b}) \end{bmatrix},$$

the weights can in fact be calculated using a procedure which is very similar to the one just used to calculate the Universal Portfolio return. One can see that

$$\hat{\mathbf{b}}_n = \frac{1}{\hat{S}_{n-1}(\mathbf{x}^{n-1})} \begin{bmatrix} \sum_{l=0}^{n-1} X_{n-1}(l) \int b_1^{l+1} b_2^{n-1-l}\, d\mu(\mathbf{b}) \\ \sum_{l=0}^{n-1} X_{n-1}(l) \int b_1^l b_2^{n-l}\, d\mu(\mathbf{b}) \end{bmatrix} =$$

$$= \frac{1}{\hat{S}_{n-1}(\mathbf{x}^{n-1})} \begin{bmatrix} \sum_{l=0}^{n-1} X_{n-1}(l) C_n(l+1) \\ \sum_{l=0}^{n-1} X_{n-1}(l) C_n(l) \end{bmatrix}. \tag{7}$$

Now, to obtain formulas that can be efficiently implemented in a program, the integrals in equation (6) need to be solved. The Dirichlet density will now prove to be a good choice. Recall that its density function is given by

$$f(\mathbf{b}) = \frac{1}{B(\alpha_1, \alpha_2)} b_1^{\alpha_1 - 1} b_2^{\alpha_2 - 1},$$

meaning that the integrals in equation (6) can be solved relatively easily. Firstly, note that

$$C_n(l) = \int b_1^l b_2^{n-l} d\mu(\mathbf{b}) = \int_0^1 b_1^l b_2^{n-l} \frac{1}{B(\alpha_1, \alpha_2)} b_1^{\alpha_1 - 1} b_2^{\alpha_2 - 1} db_1,$$

where $b_2 = 1 - b_1$.

Take the constant Beta function outside the integral and merge the b_1 and b_2 factors to obtain

$$C_n(l) = \frac{1}{B(\alpha_1, \alpha_2)} \int_0^1 b_1^{l + \alpha_1 - 1} b_2^{n - l + \alpha_2 - 1} db_1.$$

In this integral, we can see that the integrand is essentially a Dirichlet density function. If it were completed with a constant factor containing suitable gamma functions, it would become a true Dirichlet density function. Now, identify the parameters of this distribution, calling them γ_1 and γ_2. These new parameters clearly have the following relation to the original distribution parameters α_1 and α_2:

$$\gamma_1 = l + \alpha_1,$$
$$\gamma_2 = n - l + \alpha_2.$$

The general probability theory definition of a probability measure states

$$\int d\mu(\mathbf{b}) = 1.$$

The $C_n(l)$ integrals in equation (6) can thus be solved. The solution is

$$C_n(l) = \frac{B(\gamma_1, \gamma_2)}{B(\alpha_1, \alpha_2)} \int_0^1 \frac{1}{B(\gamma_1, \gamma_2)} b_1^{\gamma_1 - 1} b_2^{\gamma_2 - 1} db_1 =$$
$$= \frac{B(\gamma_1, \gamma_2)}{B(\alpha_1, \alpha_2)} \int d\mu(\mathbf{b}) = \frac{B(\gamma_1, \gamma_2)}{B(\alpha_1, \alpha_2)}.$$

Substituting the expressions introduced earlier for γ_1 and γ_2 into this gives

$$C_n(l) = \frac{B(l + \alpha_1, n - l + \alpha_2)}{B(\alpha_1, \alpha_2)} = \frac{1}{B(\alpha_1, \alpha_2)} \frac{\Gamma(l + \alpha_1) \Gamma(n - l + \alpha_2)}{\Gamma(n + \alpha_1 + \alpha_2)}.$$

Using equation (6), it is clear that $C_0(0) = \int d\mu(\mathbf{b}) = 1$ regardless of what values are chosen for the α_1 and α_2 parameters. Using the rule that $\Gamma(N+1) = N\Gamma(N)$, $C_n(l)$ can now be expressed recursively as a function of the values of α_1 and α_2, by relating $C_n(l)$ to $C_{n-1}(l-1)$ and $C_{n-1}(l)$.

$$C_n(l) = \frac{l + \alpha_1 - 1}{n + \alpha_1 + \alpha_2 - 1} C_{n-1}(l - 1)$$

$$C_n(l) = \frac{n - l + \alpha_2 - 1}{n + \alpha_1 + \alpha_2 - 1} C_{n-1}(l)$$

For $X_n(l)$ from equation (5), it is clear that when $1 \leq l \leq n-1$, the following recursion can be used:

$$X_n(l) = x_{n1} X_{n-1}(l-1) + x_{n2} X_{n-1}(l).$$

At the endpoints, two specific recursions hold:

$$X_n(0) = x_{n2} X_{n-1}(0)$$

$$X_n(n) = x_{n1} X_{n-1}(n-1).$$

To simplify further, note that in the expressions in equation (7) for the Universal Portfolio weights and also in the adjacent equation for the realized total return factor, $X_n(l)$ and $C_n(l)$ only occur multiplied by one another. Thus, by introducing

$$Q_n(l) = X_n(l) C_n(l),$$

the realized total return factor for Universal Portfolio after n rebalancings can be more conveniently written:

$$\hat{S}_n(\mathbf{x}^n) = \sum_{l=0}^{n} Q_n(l).$$

Note that also the quantities $C_n(l+1)X_{n-1}(l)$ and $C_n(l)X_{n-1}(l)$, occuring in equation (7), can be simplified by using the $Q_n(l)$ notation. Use the $C_n(l)$ recursions to see that:

$$\begin{aligned}
C_n(l+1)X_{n-1}(l) &= \frac{l + 1 + \alpha_1 - 1}{n + \alpha_1 + \alpha_2 - 1} C_{n-1}(l+1-1) X_{n-1}(l) = \\
&= \frac{l + \alpha_1}{n + \alpha_1 + \alpha_2 - 1} Q_{n-1}(l), \\
C_n(l)X_{n-1}(l) &= \frac{n - l + \alpha_2 - 1}{n + \alpha_1 + \alpha_2 - 1} C_{n-1}(l) X_{n-1}(l) = \\
&= \frac{n - l + \alpha_2 - 1}{n + \alpha_1 + \alpha_2 - 1} Q_{n-1}(l).
\end{aligned}$$

Finally, insert the above expressions into equation (7), and the explicit Universal Portfolio weight expression is obtained:

$$\hat{\mathbf{b}}_n = \frac{1}{\sum_{l=0}^{n-1} Q_{n-1}(l)} \begin{bmatrix} \sum_{l=0}^{n-1} \frac{l+\alpha_1}{n+\alpha_1+\alpha_2-1} Q_{n-1}(l) \\ \sum_{l=0}^{n-1} \frac{n-l+\alpha_2-1}{n+\alpha_1+\alpha_2-1} Q_{n-1}(l) \end{bmatrix}.$$

A.2 Description of Anticor weight calculation

To describe Anticor more thoroughly, we must first define

$$LX_{1t} = [\log(\mathbf{x}_{t-2w+1}), \ldots, \log(\mathbf{x}_{t-w})]^{\mathrm{T}}, \qquad LX_{2t} = [\log(\mathbf{x}_{t-w+1}), \ldots, \log(\mathbf{x}_{t})]^{\mathrm{T}}$$

where $\log(\mathbf{x}_{t_0})$ denotes $[\log(x_{t_0 1}), \ldots, \log(x_{t_0 m})]^{\mathrm{T}}$, t_0 is a point in time and m is the number of assets in the portfolio as before. Thus, LX_{1t} and LX_{2t} are $w \times m$ matrices, containing logarithms of relative prices of assets over the windows from $t-2w+1$ to $t-w$ and from $t-w+1$ to t respectively. Denote the jth column of LX_{lt} by $LX_{lt}(j)$ (where l is 1 or 2). Also, let $\mu_l = [\mu_l(1), \ldots, \mu_l(m)]$ and $\sigma_l = [\sigma_l(1), \ldots, \sigma_l(m)]$ denote the vectors of sample averages and sample standard deviations of the columns of LX_{lt} (i.e. the sample averages and sample standard deviations over the two windows for all assets) respectively. Using this notation, the cross correlation matrix can be calculated according to

$$M_{\mathrm{cov},t}(j,k) = \frac{1}{w-1}(LX_{1t}(j) - \mu_1(j))^{\mathrm{T}}(LX_{2t}(k) - \mu_2(k))$$

$$M_{\mathrm{cor},t}(j,k) = \begin{cases} \frac{M_{\mathrm{cov},t}(j,k)}{\sigma_1(j)\sigma_2(k)} & \text{if } \sigma_1(j), \sigma_2(k) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Here, $M_{\mathrm{cor}}(j,k)$ is the sample correlation between the log-relative prices of asset j over the window from $t-2w+1$ to $t-w$ with those of asset k over the window from $t-w+1$ to t. Should $\sigma_1(j)$ or $\sigma_2(k)$ be zero over either of the windows, it means that the logarithm of the relative price of that asset is constant over that interval. Anticor does not reallocate between assets j and k in this case.

The next step is to calculate the *claim matrix*. Here, $\mathrm{claim}_i(j,k)$ denotes an initial approximation of the amount of wealth to reallocate from asset j to asset k.

$$\mathrm{claim}_t(j,k) = \begin{cases} M_{\mathrm{cor},t}(j,k) + A_t(j) + A_t(k) & \text{if } \mu_2(j) > \mu_2(k) \text{ and } M_{\mathrm{cor},t}(j,k) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The $A_i(h)$ quantity in the above expression is calculated according to

$$A_t(h) = \begin{cases} |M_{\mathrm{cor},t}(h,h)| & \text{if } M_{\mathrm{cor},t}(h,h) < 0 \\ 0 & \text{otherwise.} \end{cases}$$

Intuitively, if $M_{\mathrm{cor},t}(j,k) > 0$, one could argue that assets j and k are correlated in consecutive windows, so that a rise in one predicts a future rise in the other, and $M_{\mathrm{cor},t}(h,h) < 0$ shows that asset h is negatively autocorrelated in consecutive windows.

Next, the transfers that are made are calculated:

$$\mathrm{transfer}_t(j,k) = \begin{cases} \mathbf{b}_t(j)\frac{\mathrm{claim}_t(j,k)}{\sum_k \mathrm{claim}_t(j,k)} & \text{if } \sum_k \mathrm{claim}_t(j,k) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

With the transfers calculated, the portfolio weights at time $t+1$ can be calculated using the weights at time t through the following relation:

$$\mathbf{b}_{t+1} = \mathbf{b}_t + \sum_{k \neq j} [\mathrm{transfer}_t(k,j) - \mathrm{transfer}_t(j,k)].$$

A.3 Technical details of the CPPI algorithm

The moving floor version of CPPI takes advantage of increases in the value of the risky asset to put more wealth into the secured floor part of the portfolio. Enforcing the maximum relative exposure determined by p, there can now be an *excess cushion* if the risky asset increases in value. The excess cushion is given by

$$\frac{mC_t - pV_t}{m} = \frac{m(V_t - F_t) - pV_t}{m} = \frac{m-p}{m}V_t - F_t$$

and this should be added to the previous floor, F_t, to limit the exposure to the risky asset. The new floor is related to the old floor as follows:

$$F_t^{\text{new}} = \begin{cases} \frac{m-p}{m}V_t & \text{if } mC_t > pV_t \\ F_t^{\text{old}} & \text{otherwise.} \end{cases}$$

Intuitively, this can be interpreted as 'If the calculated exposure is larger than the maximum allowed fraction of the portfolio wealth, the floor is moved upwards until the exposure becomes exactly the maximum allowed fraction of portfolio wealth'. To see this, note that if $mC_t > pV_t$, we set

$$F_t = \frac{m-p}{m}V_t,$$

giving

$$E_t = mC_t = m(V_t - F_t) = m\left(V_t - \frac{m-p}{m}V_t\right) = m\left(V_t - V_t + \frac{p}{m}V_t\right) =$$

$$m\frac{p}{m}V_t = pV_t,$$

i.e. exactly the maximum exposure allowed.

In the base CPPI version, if the risky asset behaves like a geometric Brownian motion, i.e. with normally distributed logarithmic returns, the CPPI algorithm expected wealth at time T has been shown [5] to be

$$V_T = V_0 + C_0 e^{-k_m}\left(\frac{S_T}{S_0}\right)^m,$$

where

$$k_m = (m-1)\left(\frac{m\sigma^2}{2} + r\right)$$

In the moving floor version, there exists no closed form solution for the portfolio wealth, but the SDE can be expressed as

$$dV_t = \begin{cases} rV_t dt & \text{if } V_t \le F_t \\ (V_t - F_t)dZ_t + rF_t dt & \text{if } F_t < V_t < \frac{m}{m-p}F_t \\ V_t dX_t & \text{if } V_t \ge \frac{m}{m-p}F_t. \end{cases}$$

Here,

$$dZ_t = (m(\mu - r) + r)dt + m\sigma dW_t,$$

$$dX_t = (p(\mu - r) + r)dt + p\sigma dW_t.$$

Another version of the CPPI presented by Boulier and Kanniganti [5] instead focuses on not letting the exposure to the risky asset grow too small in the case of a decline of the value of the risky asset, noting that a CPPI portfolio will follow the floor if it ever reaches it and behave unpredictably in the case of aggressive investing leading to a portfolio value below the floor.

Denote the initial margin M_0. Also, define E_t^* as a lower limit for the exposure at any time t. The margin version of CPPI states that, whenever $E_t < E_t^*/2$,

$$F_t^{\text{new}} = F_t^{\text{old}} - \frac{M_t}{2},$$

$$M_t^{\text{new}} = \frac{M_t^{\text{old}}}{2},$$

$$E_t^{\text{new}} = mC_t = m(V_t - F_t^{\text{new}}).$$

The lower limit can be set in several fashions. This CPPI version will not be used in this thesis, and hence it is only covered shortly here.

There is a maximum limit on how much the risky asset can fall during one investment period before the CPPI portfolio falls through the floor. Denote this critical level expressed as a fraction by the letter d. The CPPI portfolio development over one period can then be approximately written

$$V_{t+1} = E_t \cdot (1 - d) + (V_t - E_t) \cdot 1$$

if the return on the riskfree asset can be approximated to have zero return over one period. Here, $E_t = mC_t = m\,(V_t - F_t)$. Substitute to obtain

$$V_{t+1} = m\,(V_t - F_t) \cdot (1 - d) + (V_t - m\,(V_t - F_t))$$

Expand the above expression to obtain

$$V_{t+1} = mV_t - mF_t - dm\,(V_t - F_t) + V_t - mV_t + mF_t = V_t - dm\,(V_t - F_t)$$

The portfolio has fallen through the floor if and only if

$$V_{t+1} < F_t$$

or equivalently if and only if

$$V_t - dm\,(V_t - F_t) < F_t$$

Rearrange to obtain

$$dm\,(V_t - F_t) > V_t - F_t$$

Finally, divide through by $m\,(V_t - F_t)$ to obtain

$$d > \frac{1}{m}$$

Thus, as long as the risky asset does not fall by a fraction larger than $1/m$ (or slightly less if $p > 1$ and the riskless asset actually increases while the investor is net a short position in it) over one period, the CPPI portfolio stays above its floor. As a corollary, if $m \leq 1$, the portfolio will never break through the floor.

A.4 Glossary

Algorithm - A rule or a set of rules that unambiguously defines a process for solving a particular task. In this thesis, it specifically refers to a process where portfolio weights for a given portfolio are calculated using past asset prices.

Anticor - An algorithm which primarily benefits from negatively correlated and mean reverting assets. It uses serial and cross correlations to determine asset weights.

Best Constant Rebalanced Portfolio (BCRP) - For each set of assets, the BCRP is the CRP which gives the highest total return factor. Its total return factor is a strict upper bound for the total return factors of EG and Universal Portfolio.

Constant Proportion Portfolio Insurance (CPPI) - An algorithm which is conceptually limited to two asset portfolios, where one asset should be risk free and the other risky. It is primarily suitable for insurance purposes and not for return generation. A prominent feature is the floor, which is the guaranteed level of wealth. The algorithm benefits from momentum in the risky asset.

Constant Rebalanced Portfolio (CRP) - A portfolio where weights are kept constant over time. Rebalancings are only made to ensure that the same weights are kept at all times.

Economic significance - In this thesis, the term economic significance is used to describe the situation where a performance difference is likely to be considered relevant by an investor, regardless of whether the difference is found to have a statistically significant deviation from zero or not.

Exponentiated Gradient (EG) - An algorithm which uses momentum effects to generate abnormal returns. It has only one parameter, the learning rate η, and it tries to achieve BCRP performance.

Total return factor - The factor by which initial wealth has changed for any algorithm at any time.

Universal Portfolio - Another algorithm which uses momentum effects. It uses an assumption about BCRP weight distribution to calculate weights and BCRP total return factor is an upper bound for the algorithm's total return factor. Due to computational complexity, it is only feasible for two asset portfolios.