LUT
Lappeenranta
University of Technology

LUT School of Business and Management

Business Administration Bachelor Thesis

Financial Management

Performance of the Passive Aggressive Mean Reversion trading algorithm with survivorship-bias free data on the Helsinki Stock Exchange

30.05.2018

Author: Toivo Toikka

Supervisor: Jan Stoklasa

ABSTRACT

Author:	Toivo Toikka
Title:	Performance of the Passive Aggressive Mean Reversion trading algorithm with survivorship-bias free data on the Helsinki Stock Exchange
School:	School of Business and Management
Degree programme:	Financial Management
Supervisor:	Jan Stoklasa
Keywords:	Trading strategy, Investment strategy, Algorithmic Trading, Machine Learning

In this thesis, we are studying Performance of the Passive Aggressive Mean Reversion (PAMR) trading algorithm with survivorship-bias-free data from the Helsinki Stock Exchange. This algorithm uses passive aggressive online learning algorithm and mean-reversion trading strategy to its advantage. We are interested could we accept this strategy as a viable trading strategy and how survivorship biased data affects to results.

PAMR's performance was simulated with two datasets. One dataset contained historical close prices of companies from OMXH25-index in 2008 and another dataset contained historical close prices of companies from OMXH25-index in 2018. The study period was 12.03.2008 - 09.03.2018.

The algorithm performed well when comparing to OMXH25-index, but survivorship biased data gave better results than survivorship bias-free data. From this, we can make a conclusion that when using survivorship biased data results are overly confident in the study period.

1 INTRODUCTION

Algorithmic trading is shaking financial industry and being one of the biggest trends in the money management industry. For example Larry Fink The chief executive of BlackRock started a radical restructuring of its equities unit by sacking seven fund managers and shifting billions of dollars they used to manage to a Systematic Active Equities, a $100bn computer-powered “quantitative” investment unit. (Robin Wigglesworth 2018)

Algorithmic Trading is natural evolvement of three developments in the financial industry. The first is that the financial system is becoming more complex over time which has created demand for algorithmic trading. The second development is the breakthroughs in the development of quantitative modeling of financial markets pioneered over the past three decades. For example Harry Markowitz (1952) mean variance analysis, William F. Sharpes (Sharpe 1964) capital asset pricing model and many others. Algorithmic trading is one of the many intellectual progeny that they have fathered. Lastly the fast evolvement of computer technology has created new ways to operate in financial markets. (Andrei A. Kirilenko, Andrew W. Lo 2013) One interesting advancement is using machine learning in trading and investing strategies. In this thesis we are testing one algorithmic trading strategy which uses machine learning on its advantage. It is called Passive Aggressive Mean Reversion.

Bin Li et al. (2012) proposed a strategy called Passive aggressive mean reversion or PAMR. It is a combination of two ideas: Mean reversion of price and passive aggressive online learning algorithm. Online learning is machine learning where model adapts in real time when data comes available. (Dochow 2016) This makes the online learning methods quite interesting alternative for financial markets. Mean reversion refers to investment strategy where investor believes that price will revert back to mean price. (Chan 1988)

PAMR has two stages: Failure or success. Each timestep algorithm either passively holds portfolio or at failure algorithm aggressively rebalances portfolio with mean reversion strategy, then holds this until next failure. It is called passive aggressive because it passively holds when successful or actively rebalances weights when failing. We are interested in how this strategy performs in Helsinki Stock Exchange if it would be applied.

When studying algorithmic trading strategy we should consider few things. Firstly we are considering what data we use. Data has crucial affects to results. Bin Li et al. (2012) did not use survivorship bias free data. Which meant that in their study, data contained only survived companies and PAMR did not have possibility to choose poorly performing of failed stocks. They agree that it was a limitation and could contribute for too optimistic results. In order to show the effect of survivorship bias, we are using two datasets in this thesis. One that contains survivorship bias and one which does not and then compare results. These datasets are used as a research material.

Secondly we are interested in methods of generating information. PAMR was simulated with historical prices and its performance evaluated with financial performance metrics and then compared to different benchmarks. Because we are studying PAMR's performance in Helsinki stock exchange we are simulating PAMR's performance with stock from OMXH25 and compare its performance to OMXH25-index. If strategy performs better than benchmark we accept that strategy.

The goal of this thesis is to study PAMR's performance in Helsinki Stock Exchange, how survivorship bias and survivorship bias free data affects results and make a statement that could we accept PAMR as a viable investment strategy in Helsinki Stock Exchange.

Structure of thesis is following. In Literature review section we review relevant studies considering the evolution of trading and investment strategies, algorithmic trading, mean reversion of prices and online learning solutions for portfolio selection.

In theory section, we review the theory of online learning, PAMR, and modern portfolio theory. In chapter four we review research material. In chapter five we review research methods. In chapter six we present results and in chapter seven we make a conclusion.

2 LITERATURE REVIEW

Trading strategies are widely studied in the academic field and in markets. In literature, there is no clear consensus for terms. Sometimes it is hard to make a difference between trading and investment strategy. Secondly, names for technical implementations of strategies are mixed up in literature. For example the meaning of electronic trading, traditional investing, quantitative investing, algorithmic trading and high-frequency trading overlap often. That might be caused by that algorithmic trading became introduced to markets in 2003. (Chaboud et al. 2014) That is why we are suggesting a definition for algorithmic trading which is used in this thesis.

When algorithmic trading being relatively new subject scientific studies differ from practice. Secondly, information about trading strategies is usually found in textbooks which do not necessarily follow standards of scientific papers. Lastly, detailed implementations of Algorithmic Trading systems are usually kept secret and protected by important intellectual property rights. Algorithmic strategies are considered as a key component of traders business models. (Gomber, Gsell 2009) In this section, we review the history of trading and investment strategies, studies about algorithmic trading, mean reversion and online portfolio selection algorithms.

2.1 Trading and Investment Strategies

Technical analysis was one of the earliest techniques that came in the early 1910's for generating trading signals. In the Technical analysis, traders tried to identify recurring patterns in security prices and generating trading signals from that. It was working strategy for first half of the 20th century when markets were relatively inefficient compared to market today. The pace of market was so slow that it was possible to analyze supply and demand from price charts. (Aldridge 2009)

Technical analysis has been criticized in the literature for its highly subjective nature. Finding meaningful shapes from historical price data is in eyes of viewers. People usually tend to come up with meanings from data that does not actually exist. For example, sometimes individuals make judgments and decisions that are successful by luck. These individuals will be susceptible to an illusion of skill and to overconfidence. (Kahneman, Klein 2009)

Lo, Mamaysky and Wang published a study in 2000 where they noticed that it is possible that certain technical patterns can be found with nonparametric kernel regression. In 2000 they said that human judgment is still better to find visual patterns, but advancements of statistical learning theory have had successful applications of fingerprint recognition, handwriting analysis, and face recognition. Might be that technical analysis is next frontier for pattern recognition models. (Lo, Mamaysky & Wang 2000) Today Machine learning is used in financial markets for finding patterns to trade on. This act is called as “Data mining”. (Chang 2016)

Another popular investing and trading technique was fundamental analysis. Fundamental analysis originated in the 1930s. The Idea was to create fundamental signals from financial statements data and trade and invest from that information. Traders noticed that future cash flows, such as dividends, affected market price levels. (Graham and Dodd (1934) as cited in Aldridge 2009) Over the years fundamental analysis expanded to include economic variables. (Aldridge 2009, Abarbanell, Bushee 1997)

Abarbanell and Bushee (1997) studied fundamental analysis and argue that fundamental analysis does not completely impound the information that investors perceive is contained in the fundamental signals. One explanation is that fundamental signals may capture information about the firm that has little to do with near-term earnings. A second explanation supported by their study is that analysts forecast revisions fail to completely impound the information in the fundamental signals about future earning changes. Analysts Forestercast errors reveal that analyst inefficiency takes the form of generalized underreaction. Also, Abarbanell and Bushee mention that macroeconomic variables such as inflation and GDP condition some of the relations between the fundamental signals and future earnings, revisions, and forecast errors. (Abarbanell, Bushee 1997)

Some selected technical models and some fundamental models were adopted by quantitative traders. Quantitative traders extended the precision of models with statistical techniques and often sped up the calculation of the relevant values with computers. Fair values of equities following an earnings announcement were recomputed on the fly, enabling quantitative traders to reap the profits at the expense of fundamental traders practicing analysis in Excel spreadsheets. (Aldridge 2009) For example, Qing Li et. al, (2014) found in their research that the media influence of financial news on stocks exists and can be quantified using natural language processing techniques in Computer Science. Fundamental information can enrich the knowledge of investors and affect their trading activities. Meanwhile, news article sentiment may lead to emotion fluctuations of investors and interfere with their decision making. (Li et al. 2014)

We can notice that speed became the most obvious aspect of quantitative traders competition. Whoever was able to run quantitative model the fastest was the first to identify and trade on a market inefficiency and was the one to capture the biggest gain. This lead to a competition of equipment. Traders started developing faster ways to trade. Traders started to analyze data with computational methods to have a competitive edge in the market. This lead to the development of Algorithmic Trading. (Aldridge 2009) It has gone so far that high-frequency traders transmit information with microwaves that can travel faster than light. (Budish, Cramton & Shim 2015)

Traders had always had an advantage when being faster. That is why high-frequency trading and algorithmic trading is a natural evolution of the securities markets instead of a completely new phenomenon. (Gomber et al. 2011, Aldridge 2009)

When we observe trading and investing from the perspective of game theory, we can notice that market practitioners are also analyzing how other players act in the market and use this information in their decision process. (Arthur 1995)

2.2 Algorithmic Trading

Irene Aldridge (2009) describes algorithmic trading as the basis of all computer-assisted trading processes. Algorithmic trading can refer to a variety of algorithms spanning order-execution processes as well as high-frequency portfolio allocation decisions.

Chan (2014) and Chaboud et al. (2017) describe Algorithmic trading simply and briefly as a strategy where computer program analyzes data, creates trading signals from it and then automatically submits sell and buy orders to a broker. The Difference between electronic trading and algorithmic trading is that electronic trading refers to the ability to transmit the orders electronically as opposed to transmit the orders in person. Algorithmic trading generates trading orders that are usually executed with electronic trading. (Aldridge 2009)

For purpose of this thesis, we define algorithmic trading as follows: Algorithmic trading is trading where decision making is done by an algorithm. This algorithm processes inputs with set logic and then gives buy and sell orders as an output.

The declining costs of technology have led to widespread adoption of computer technology throughout financial industries. The resulting technological change has revolutionized financial markets and the way financial assets are traded. Algorithmic trading has lowered the costs of trading and increased the informativeness of quotes by contributing more to a discovery of the efficient price than human trading. Algorithmic trading consumes liquidity when it is cheap and provide liquidity when it is expensive. (Hendershott, Jones & Menkveld 2011 Hendershott, Riordan 2009)

There is also negative talk and criticism towards algorithmic trading. For example, Gregory Scopino compares algorithmic traders who use machine learning in their strategies to “replicants” from Blade Runner. Replicants are androids who look like human and are not allowed on the earth. In this case, regulators are trying to find those market manipulating algorithmic traders who disguise themselves to the vast amount of traders. (Scopino 2015) Safety of Artificial intelligence is taken seriously in the academic field. When comparing replicants to statistical learning models is not so far fetched. Power of self-learning models is that it can develop its own methods to maximize set reward. AI will do anything to maximize set goal, which can lead to a situation where AI’s methods have a negative impact on human society. How can we be sure that AI’s actions are in line with its creator's interests? (Bostrom 2014, Vladeck 2014)

One problem of algorithmic trading is illegal and questionable market manipulation strategies which are executed with HFT systems. Generally, these market manipulation strategies are usually based on analyzing and manipulating order book to confuse other algorithmic traders and/or human traders. For example Spoofing strategy where high-frequency trader intentionally distorts order book by making fast orders that are not executed to change other traders view of supply and demand. (Aldridge 2009)

Another case that can describe the attitude towards algorithmic trading is when Doug Kass said “I say kill the quants and their technology before they kill us,” (Doug Kass as cited in CNBC (2012)). Doug Kass said this after Knights Capital Group accidentally lost 440 million dollars in 30 minutes. Knights Capital Group blamed this massive mistake on a software glitch. (Matthew Philips 2012).

Algorithmic trading has been criticized for creating flash crashes. Flash crashes are situations where automated trading crashes market prices in minutes. (Golub, Keane & Poon 2012). There have been many situations of flash crashes. For example on May 6, 2010, in the course of about 30 minutes, U.S stock market indices, stock-index futures, options, and exchange-traded funds suddenly dropped by five percent, followed by a rapid rebound. (Kirilenko et al. 2011)

Kirilenko et al, (2011) in their analysis they believed that High-Frequency Traders exhibit trading patterns inconsistent with the traditional definition of market changes.

This activity comprises a large percentage of total trading volume but does not result in a significant accumulation of inventory. This can be mistaken for liquidity to Fundamental Traders and when rebalancing their positions, High-Frequency Traders may compete for liquidity and amplify price volatility. (Kirilenko et al. 2011)

In the end algorithmic trading is part of a much broader trend in which computer based automation has improved efficiency by lowering costs, reducing human error. But regulatory framework which supposed to oversee financial markets are not necessarily catched up this development. (Andrei A. Kirilenko, Andrew W. Lo 2013)

2.3 Mean Reversion Of Price

Mean reversion of price in finance is a theory which states that prices of securities return to their long-term mean. It has been studied extensively in studies about Efficient Market Hypothesis. The weak form of Efficient Market Hypothesis states that: A financial market is efficient when market prices reflect all available information about value. (Malkiel, Fama 1970) This means that there should not be autocorrelation in time series. Price changes should be a random walk. (Fama 1995)

Eugene F. Fama and Kenneth R. French studied mean reversion of prices by examining autocorrelations of returns in 1926 - 1985 period. They noticed that there is no autocorrelation in daily changes but a large fraction of 3-5 year returns was negatively autocorrelated. At the time of the article, they acknowledge that studying price changes more they need better statistical techniques. (Fama, French 1988)

Mean reverting property of returns makes contrarian stock selection strategy possible. In contrarian stock selection strategy stock is bought when returns are lower and sold or shorted when returns are greater than some mean value. These strategies rely upon markets overreaction. Investors use this market inefficiency to gain when prices revert back to fundamental values. (Chan, 1988) Amos Tversky and Daniel Kahneman findings support this in their study of rationality. They argue that rational market hypothesis assumes too much for peoples rational thinking. (Tversky, Kahneman 1986)

Merton criticizes studies about rational-market hypothesis by saying that other economic and financial theories rely too much on this theory. (Merton, 1985)

The scientific dialogue above can be described as a battle of Efficient Market Hypothesis (EMH) and Behavioral Finance. Andrew W. Lo describes a new framework called Adaptive Market Hypothesis (AMH) where financial econometrics can co-exist together with behavioral models. (Lo 2005)

Adaptive Market Hypothesis can be viewed as a new version of EMH. Where AMH consist of the following ideas:

1. Individuals act in their own self-interest
2. Individuals make mistakes
3. Individuals learn and adapt.
4. Competition drives adaptation and innovation.
5. Natural selection shapes market ecology.
6. Evolution determines market dynamics

EMH and AMH have similar starting points but the two paradigms part company in 2 and 3. Inefficient markets, investors do not make mistakes because the market is always in equilibrium. In AMH framework, mistakes occur frequently, but individuals learn from mistakes and adapt accordingly which ultimately leads to market equilibrium by this process. (Lo 2005)

According to AMH, Mean-Reversion strategy can be seen as a process that eliminates over and undervaluation which corrects prices closer to their efficient values.

2.4 Online Portfolio Selection

Can computer learn to predict the stock market has always been an interesting question. Online algorithms are machine learning solutions for an algorithmic problem where the entire input is not available from the beginning. When market conditions constantly change online algorithms became an interesting method to solve portfolio selection problems. Online learning algorithm constantly tries to adapt its predictions each step where data comes available. (Dochow 2016)

In this thesis, we are studying passive aggressive mean reversion algorithm (PAMR). PAMR is one of online portfolio selection algorithms and thus in this chapter, other online portfolio selection algorithms and PAMR are reviewed separately.

2.4.1 Review of online Portfolio Selection

Online portfolio selection has been extensively studied across several research communities, including finance, statistics, artificial intelligence, machine learning and data mining. These algorithms can be generally divided into five classifications: Benchmarks, Follow-the-Winner, Follow-the-Loser, Pattern-Matching Approaches and Meta-Learning Algorithms. (Li, Hoi 2014)

Benchmarks algorithms are strategies that are used to evaluate other strategies. For example Best Hold Strategy. In this strategy, historical returns of best possible hold position are compared to other algorithms performance. This strategy is hypothetical and it can not be implemented because it has intended lookahead but it sets a scale for performance evaluation. (Li, Hoi 2014)

Follow the winner algorithms try to track best constant rebalancing portfolio BCR or the best experts provided. These experts are fictional investors that use different strategies. (Dochow 2016) For example, Universal Portfolios were algorithm tracks different Experts and their wealth which is diversified to set assets. Generally, those assets which are more beneficial for experts should generate periodically more wealth in future. (Cover 2011)

Follow-The-Loser algorithms are based on believing that assets are mean reverting. Meaning that Follow-The-Loser algorithms allocate wealth from high performing assets to low performing. (Dochow 2016) For example, Anti Correlation Algorithm which transfers wealth from the high-performance stock to the anti-correlated low performing stock. This algorithm can also be used for short selling when a stock is outperforming other stocks. Especially if this strong performance is anti-correlated with other stocks performance. Short selling is selling where a trader can sell securities without owning them. Idea is to find assets that have distinctly different performance than other assets. This indicates a counter movement of performance in the future (Borodin, El-Yaniv & Gogan 2004)

Pattern-Matching algorithms try to predict the next market distribution based on a sample of historical data and explicitly optimizes the portfolio based on sampled distribution. Meta-Learning Algorithms refers to algorithms that combine multiple strategies. (Li, Hoi 2014)

2.4.2 Review of Passive Aggressive Mean Reversion

In this study, we are studying Passive Aggressive Mean reversion (PAMR) which is according to Li and Hoi (2014) classified as the Follow-The-Loser algorithm. Li et al (2012) proposed Passive Aggressive Mean Reversion or PAMR. This strategy is based on two ideas, mean reversion relation of financial markets and using online learning to balance between actively rebalancing or holding current portfolio. PAMR has one sensitivity parameter which is set before trading starts. It is called sensitivity parameter and its purpose is to adjust how much PAMR is willing to take loss before rebalancing weights and how strasticly it changes weights of portfolio.

PAMR was studied by backtesting algorithms performance over time. The performance was measured by following industry standard performance metrics:

Sharpe ratio, Maximum drawdown, Calmar ratio and Annual percent yield. (Li, Hoi 2014)

It was able to perform better than market benchmarks when comparing performance metrics. One notion of its performance is that it behaved differently with different dataset. Li et al (2012) backtested this algorithm with stocks from NYSE, S&P500, MSCI, TSE and DJIA. They also tested PAMR's performance with different sensitivity parameters. They noticed that DJIA and S&P500 returns did not grow when sensitivity parameter got closer to 0.

Data that they used was survivorship biased when containing only companies that haven't got bankrupted or merged. They also state that this could have contributed to excellent results. (Li, Hoi 2014) Therefore I intended to extend the scope of analysis and also use surviorbias free data.

3 THEORY

3.1 Performance Evaluation of Algorithmic Trading Strategy

Performance evaluation is to make a statement about the behavior of an algorithm. In literature algorithms performance is studied by simulating its behavior with historical data. And measuring its performance with different metrics from machine learning and finance community. Machine learning community are usually interested in the quality of learning and accuracy of predictions. Finance community is interested in metrics of returns, risk, and risk-adjusted performance. (Dochow 2016)

Simulating trading strategies is also called backtesting. It is a testing using historical data that determines the performance of strategy if it had actually been employed. While backtesting does now allow one to predict how a strategy will perform under future conditions, its primary benefit lies in understanding the vulnerabilities of a strategy through a simulated encounter. (Treleaven, Galas & Lalchand 2013)

Backtesting algorithm with historical close prices has its limitations. For example, close prices might not be that price algorithm would actually execute and market effect is not taken into account. (Chan 2017)

A market effect can be caused by a number of things. One of them is due to fact that order book is filled with buy and sell orders or bids and asks. Bids and asks are market-making traders orders that are listed in the order book. Each bid and ask order has price and quantity of underlying security. When a trader takes these orders straight it acts as a market maker and its order is filled with these orders. If the size of the market taking order has a greater amount of securities that lowest bid or ask it will use next bid or ask to fill that order. This will raise the price of sell orders or lower the price of buy orders thus creating a situation where market taker is not paying the lowest price but the average price of those bids or asks that fill market takers order. It is possible to simulate this but in this study, we are limited to use only close prices.

(Aldridge 2009) Another market effect that it is very hard to simulate is traders that reacts to other traders actions. Hypothetically algorithms actions could lead to a different outcome which is not shown in historical market prices.

3.2 Online Learning

Online learning is a method of machine learning where data becomes available in sequential order and then it is used to update the best predictor of future data at each step. Online algorithms work in a situation where decisions must be made with incomplete knowledge of the future, that is why it suits well in financial analysis. (Karp 1992) The online learning model makes no statistical assumptions on how a sequence of inputs and rewards is generated. (Auer 2010)

The online learning model is formalized as follows. In each trial $t = 1, 2, ...,$ the learner:

1. Receives input $x_t \in X$
2. Makes prediction $y_t \in Y$
3. Receives response $z_t \in Z$
4. incurs loss $l_t = l(y_t, z_t)$

Where $l : Y \times Z \rightarrow R$ is some loss function. The performance of a learner up to trial T is measured by its accumulated loss $L_T = \sum_{t=1}^{T} l_t$. (Auer 2010)

3.3 Passive Aggressive Mean Reversion Strategy for Portfolio Selection

Passive Aggressive Mean Reversion or PAMR is portfolio selection strategy. This strategy is based on two ideas, mean reversion relation of financial markets and using online learning to balance between actively rebalancing or holding current portfolio. (Li et al., 2012) Portfolio selection can be divided into two stages. The first stage is where observation is made and experience gathered to estimate what future will be for set securities. The second stage is where the portfolio is collected with current knowledge to maximize returns to certain risk constraint. (Markowitz, 1952)

Bin Li et al, 2012 based this portfolio selection strategy to Crammer et al. 2006, suggested passive aggressive binary classification algorithm. PAMR has two formulas. First is loss function which triggers optimization when portfolio loss is greater than sensitivity threshold. The formula of the loss function is following:

$$l_\epsilon(b; x_t) = \begin{cases} 0 & b_t \cdot x_t \leq \epsilon \\ b_t \cdot x_t - \epsilon & otherwise, \end{cases}$$

Where b_t is portfolio vector at timestep t assets and x_t price relative vector at time t and ε sensitivity parameter. (Li et al., 2012, Crammer et al. 2006) We want always be in a situation where portfolio vectors sum is alway 1. That is why we optimize that vector. Optimization problem of weights is following:

$$b_{t+1} = argmin_{b \in \Delta_m} \tfrac{1}{2}\left\|b - b_t\right\|^2 \;\; s.t \;\; l_\varepsilon(b;\; x_t) = 0 \;\; ,$$

Where $\Delta_m = \{b : b \in R_+^m, \sum_{i=1}^{m} b_i = 1\}$. Formulation attempts to find an optimal portfolio by minimizing the deviation from the last portfolio b_t under the condition of satisfying the constraint of zero loss. If portfolio return is below the threshold, ε. When the loss is nonzero, PAMR aggressively updates the solution by forcing it to satisfy constraint $l_\varepsilon(b_{t+1}; x_t) = 0$.

The procedure of PAMR’s backtest algorithm can be set to eight parts. Which is described in Annex 1.

In the first part portfolio vector is initialized. Which describes how portfolios wealth is divided for each security. Second part is for-loop where weights are calculated for each step t.

Third section creates stock price relatives vector x_t. Price relative vector is vector which contains relative price changes. It is calculated as current price divided by previous price.

In part four suffer loss l_e^t is calculated by multiplying portfolio vector b_t with price relative vector x_t. This is also return of portfolio. If this return is greater than sensitivity parameter ε algorithm will rebalance. This is done because of mean reversion strategy. Like described before in mean reversion strategy wealth is transferred from over performing- to underperforming stocks. Suffer loss does not mean loss of investments but online learning algorithms logic.

In part five τ is calculated which is used in part 6 for portfolio update. $\left\| x_t - \overline{x_t} 1 \right\|^2$ in the denominator describes market quadratic variability if it is high τ will be small and if variability is low τ is higher. Idea is to manage risk and return when updating portfolio. If τ is high portfolio will be updated more aggressively in part "Update portfolio". If loss function l_e^t has been 0 PAMR will not update portfolio.

In the 7th part of "for loop" new weights are normalized by projecting them to simplex domain to make sure that portfolios total wealth invested is exactly the money available. Finally when each time step is calculated for loop and program is ended.

Input: ε:sensitivity parameter;
Procedure:

1. $Initialize\ b_1 = (\frac{1}{m}, ..., \frac{1}{m})$
2. $for\ t\ =\ 1, 2, 3, ..., n\ do$
3. Receive stock price relatives : $x_t = (x_{t1}, ..., x_{tm})$
4. Suffer loss : $l_\varepsilon^t = max\{0, b_t \cdot x_t - \varepsilon\}$
5. Set Parameters : $\tau_t = \frac{l_\varepsilon^t}{\|x_t - \overline{x}_t 1\|^2}$
6. Update portfolio : $b_{t+1} - \tau_t(x_t - \overline{x}_t 1)$
7. Normalize portfolio : $b_{t+1} = arg\ min_{b \in \Delta_m} \|b\ - b_{t+1}\|^2$
8. $End\ for$

END

Annex 1, PAMR Backtest Algorithm (Li et al., 2012)

3.3 Modern Portfolio Theory

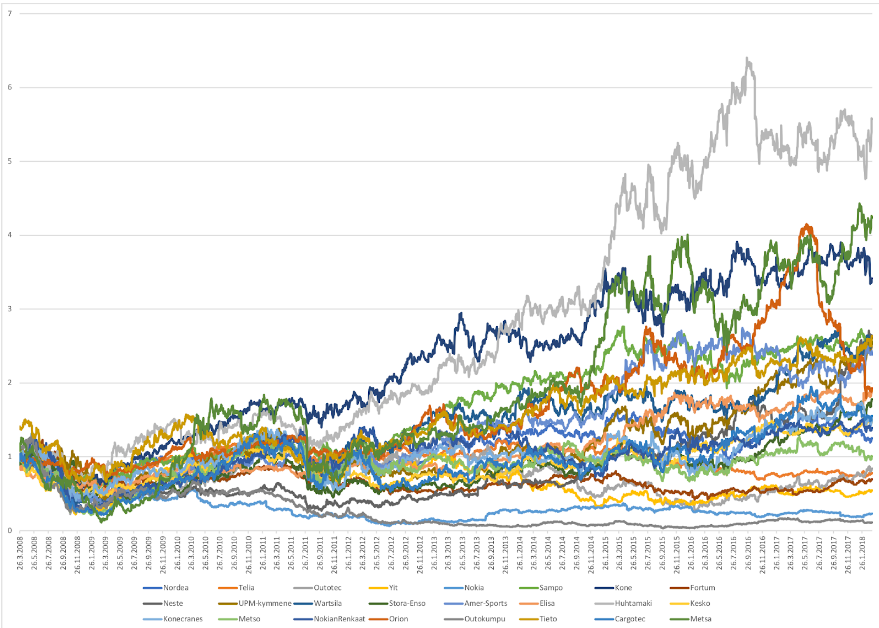
Harry Markowitz is usually described to be father of modern portfolio theory (MPT). His fundamental theorem of mean variance portfolio theory which suggest that it is possible to maximize return to given risk. Markowitz suggested that investor could take an advantage from knowing how securities co-moved with each other. (Elton, Gruber 1997)

Later William F. Sharpe (1966, 1967) developed still used metrics for mutual fund performance. Maybe the most famous of them is Sharpe ratio which idea is to measure how much portfolio can generate return to given risk. Another is market model where return of stock is measured also by its sensitivity with the whole market movements. That is why beta is usually used to describe market risk.

4 RESEARCH MATERIAL

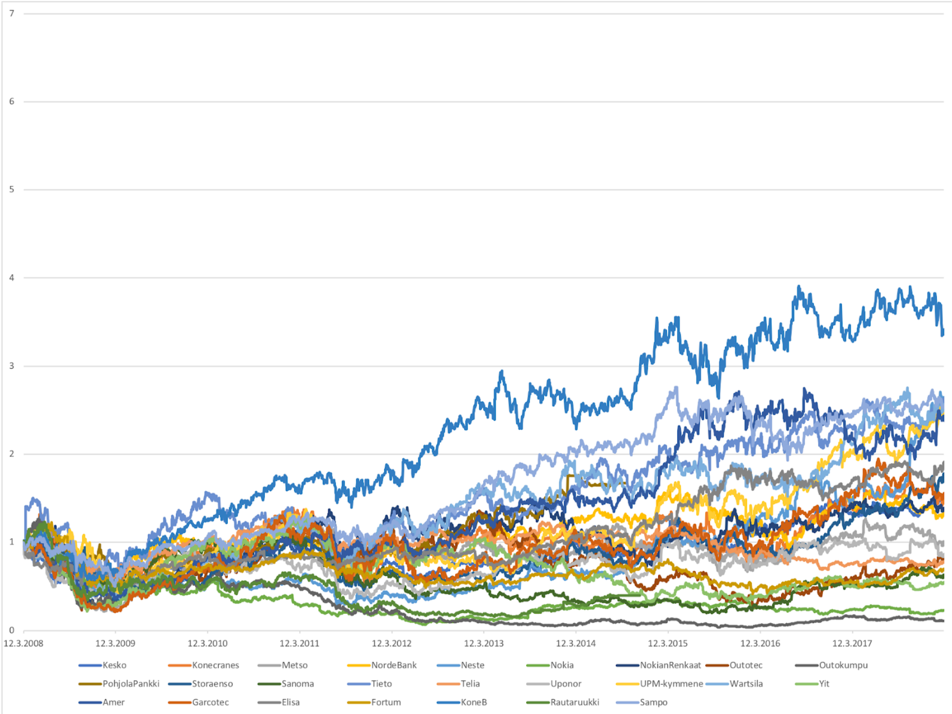
This study is limited to underlying stocks of OMXH25 in early 2018 and 2008 in the time period of 12.03.2008 - 09.03.2018. Prices are recorded at close prices. Underlying stock are listed in appendix.

OMXH25 in early 2018 contains all the stocks that are in there early 2018. When taking the historical sample of those stocks it creates lookahead bias because it will have all well-performed stocks in given historical time period. These prices are indexed and plotted in GRAPH 1.



GRAPH 1. Indexed prices of OMXH25 stocks in 2018

OMXH25 in early 2008 contains all the stocks that are in there early 2008. This dataset does not have lookahead bias because it contains also poorly performed stocks and companies that left stock exchange entirely. These prices are indexed and plotted in GRAPH 2.



GRAPH 2. Indexed prices of OMXH25 stocks in 2008

Those stocks that left exchange where: OP-pohjola which in 2014 bought it's share back and left public securities exchange, in 2014 Rautaruukki merged with SSAB and Uponor and Sanoma was dropped from OMXH25. Later Huhtamäki, Metsä group, and Orion were added. Data was gathered from Datastream service and it contains 2608 daily observations.

5 RESEARCH METHODS

5.1 Empirical Testing of PAMR

We are simulating PAMR's performance with historical close prices of two datasets. One that has survivorship bias and another which has not. Simulation time period is 12.03.2008 - 09.03.2018. We assume that close prices do not differ from actual prices algorithm would get and there is no market effect. PAMR needs predetermined sensitivity parameter as an input. We are running the model for multiple values of the sensitivity parameter.

As a Benchmark we are using the OMXH25 index. We chose this because we think that strategy which outperforms market index is considered as an acceptable strategy.

5.2 Performance Metrics

In this study we are interested in three things from the view of investor: Measures of return on investment, measures of risk, measures of risk-adjusted performance. Performance metrics are calculated with a risk-free rate of Finland's 10-Year Government Bond Yield at 02-15-2018. The rate was 0.87%.

5.2.1 Annual Percentage Yield

Annual percentage yield is a number that tells you how much you'll earn with compound interest over the course of one year. Annual percentage yield or APY is calculated in following way.

$$APY = \left(\frac{W_t}{W_0}^{\frac{1}{Number\ of\ Years}} - 1\right) ,$$

Where W_t is wealth of last timestep and W_0 wealth at start. (Dochow 2016)

5.2.2 Sharpe Ratio

Sharpe Ratio is risk corrected performance metric where Portfolios returns are divided by standard deviation. This standard deviation is commonly used to describe risk.

Sharpe ratio is risk adjusted performance metrics. It is calculated with following formula:

$$SR = \frac{(r_i - r_f)}{\sigma_i},$$

Where r_i is portfolio's return, r_f risk free profit and σ_i portfolios volatility. (Sharpe 1966) This is important performance metric to evaluate performance. Sharpe ratio can be used to compare different investment.

5.2.3 Beta

In statistics Beta or β means regression analysis slope which describes relations of two variables. In finance β is used to describe market risk by underlying assets relation to the market indexes movement. Market index is usually index that holds same assets that investment strategy is done.

In this study β is measured by backtests correlations with benchmark. It is used to describe market risk. Beta is calculated as follow:

$$\beta = \frac{Covariance(r_i, r_m)}{Variance(r_m)},$$

Where r_i is return of portfolio and r_m markets return. We are expecting Beta to be higher than 1. This is due to fact that algorithm is trying to catch only undervalued securities which should have faster rise of price than market in general. Second thing is that returns are generated only from long positions [Buying security and holding it]. This means that algorithm is only capable to generate return from growing stocks. If we would also short sell [sell a stock that is borrowed] and profit from falling prices, expected Beta would be lower than 1.

5.2.4 Maximum Drawdown

Maximum drawdown is biggest possible downward movement from peak to slump. This is commonly used to describe risk and how stable performance is. It is calculated as follow:

$$MDD = max_{t=0,\ldots,T}\left(max_{\tau=0,\ldots,t} W_{\tau} - W_{t}\right),$$

Where W is wealth. (Dochow 2016)

5.2.5 Jensen's Alpha

Jensen's alpha or alpha in finance is used to determine the abnormal return of a performance theoretical expected return.

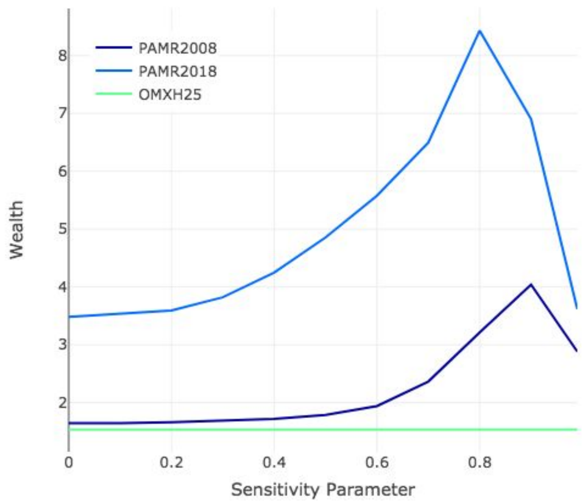
$$\alpha = R_i - (R_f + \beta_{iM}(R_m - R_f)),$$

Where R_i is the realized return of investment, R_m the return of market, R_f the risk-free rate of return and β_{iM} investments beta to the portfolio. (Jensen 1968) We are expecting Alpha to be positive. Positive alpha would describe abnormal or excess return that could describe returns coming from mean reversion strategy.

6 RESULTS

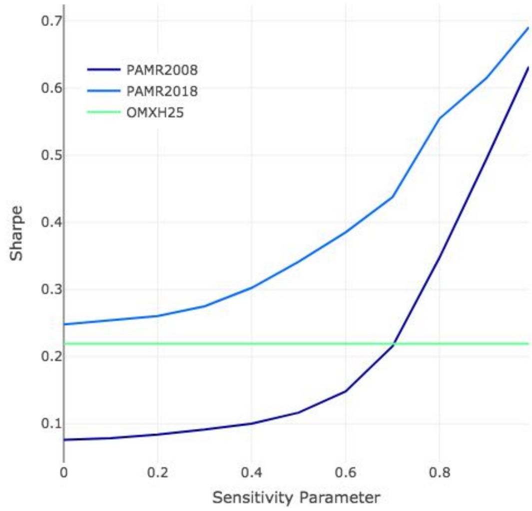
In this chapter, we will refer PAMR backtested with the dataset from 2008 and 2018 stocks as PAMR2008 and PAMR2018. PAMR2008 contains stocks that left stock exchange or merged with other companies. PAMR2018 contains only stocks that were on OMXH25 at early 2018.

In GRAPH 3. we plotted sensitivity parameters effect to achieved total wealth. We can see that total wealth generated for both datasets rises when sensitivity parameter grows, but at some point between 0.8 and 0.99, it starts to decline sharply. This is due to the fact that when sensitivity parameter grows algorithm won't rebalance portfolio so dramatically and returns are generated mostly from portfolio returns and that is why wealth starts to get closer to index.



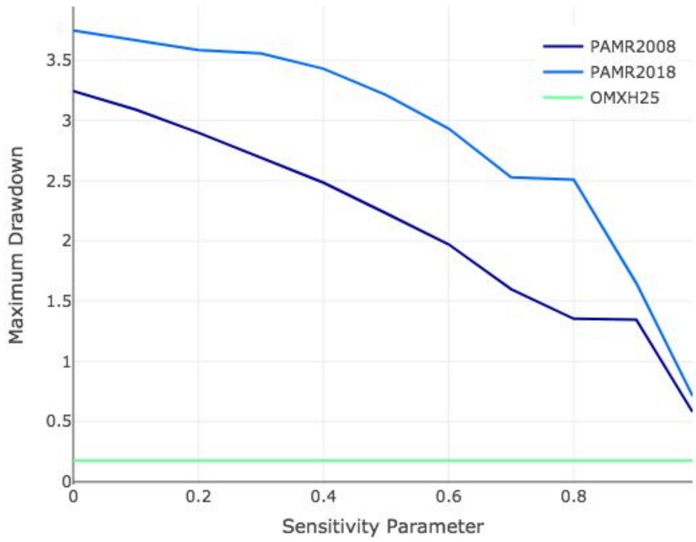
GRAPH 3. Sensitivity parameter to wealth generated.

In GRAPH 4 we illustrated sharpes sensitivity to sensitivity parameter. Sharpe rises when sensitivity parameter grows. This happens because when sensitivity parameter rises PAMR diversifies weights more broadly. This decreases the variance of the whole portfolio. PAMR2008 can't perform better OMXH25 with all sensitivity parameters but after 0.7 it reaches same level as OMXH25



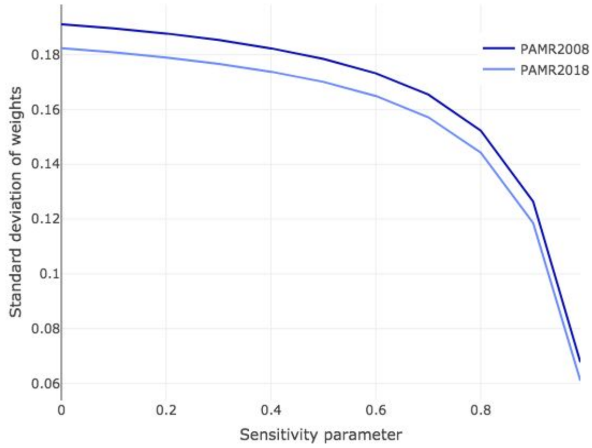
GRAPH 4. Sensitivity Parameter to Sharpe

In GRAPH 5 sensitivity parameter is plotted with the maximum drawdown. Maximum drawdown declines also when sensitivity parameter grows.



GRAPH 5. Sensitivity Parameter to Maximum Drawdown

In GRAPH 6 We can examine the behavior of this algorithm by plotting sensitivity parameter to the standard deviation of weights. Like expected algorithms weights are spread more broadly when sensitivity parameter grows.



GRAPH 6. Sensitivity Parameter to Standard deviation of weights

From these graphs, we can notice four interesting values for sensitivity parameters. When sensitivity is 0 algorithm rebalances every time the previous portfolio fails, 0.8 where PAMR2018 makes biggest returns, 0.9 where PAMR2008 performs best and lastly 0.99 which gives best Sharpe for both strategies.

In GRAPH 7 we have PAMR2018 achieved wealth plotted over time with different sensitivity parameters described above. Also, we added the mean performance of all sensitivity parameters between 0 and 0.99. We can notice that it fails between 2011 and 2016 with sensitivity parameter of 0. But with 0.8 it makes great results.

GRAPH 7. Wealth generated of PAMR2018 with different sensitivity parameters and mean wealth generated of all sensitivity parameters.

In next GRAPH 8, we can see how PAMR2008 performed over time with different sensitivity parameters. It does not perform as well PAMR2018 does. It is probably due to fact that PAMR2018 uses data from nicely performed companies and PAMR2008 does not. Best achieving sensitivity parameter is different than PAMR2018.

GRAPH 8. Wealth generated of PAMR2008 with different sensitivity parameters and mean wealth generated of all sensitivity parameters.

We can notice that both datasets work quite well, but with sensitivity, parameter 0 algorithm runs to problems with both datasets in between 2011 - 2016. In GRAPH 9. we plotted mean strategy for both datests and OMXH25-index. From this we can interpret that both sets seems to have similar behaviour.

GRAPH 9. Wealth generated by mean strategy with PAMR2008, mean strategy with PAMR2018 and wealth generated by OMXH25.

In TABLE 1. and TABLE 2. We collected performance metrics for different sensitivity parameters, mean performance of all sensitivities and performance of OMXH25. Like previously discussed we were interested in strategies that can perform better than the index. When looking TABLE 1. we can notice that 0.8 0.9 and 0.99 performed better than the index (OMXH25) when comparing Sharpe and annual percentage yield. On the other hand, the maximum drawdown was greater with all versions PAMR than OMXH25.

	0	0.8	0.9	0.99	mean	OMXH25
sharpes	0.08	0.35	0.49	0.63	0.19	0.22
MaxDD	0.60	0.44	0.39	0.22	0.46	0.18
Beta	1.88	1.77	1.84	1.17	1.69	1.00
APY	6.83%	16.78%	20.39%	15.10%	11.28%	5.84%
Alpha	-0.35	0.48	2.07	1.26	0.34	N/A
Standard Deviation	0.721	0.590	0.575	0.275	0.573	0.201

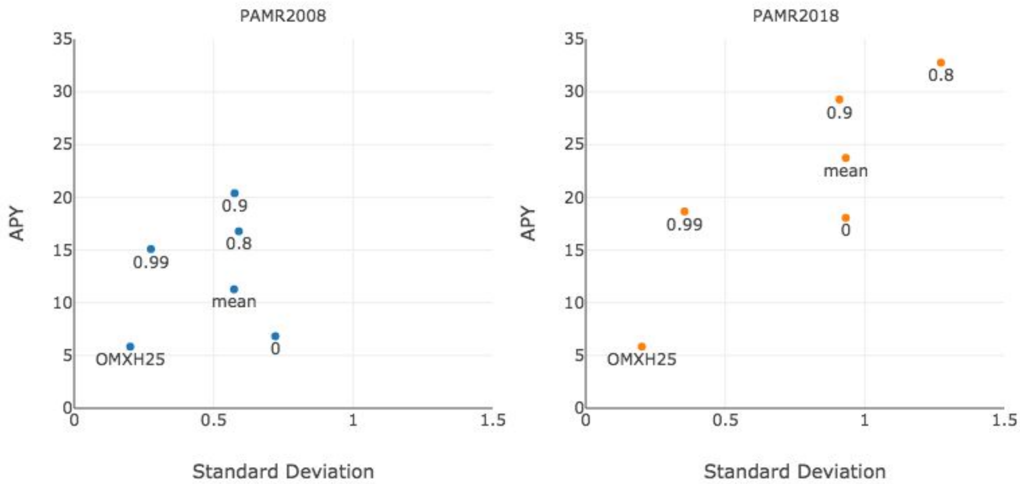
Table 1. Performance metrics of OMXH2008 data

	0	0.8	0.9	0.99	mean	OMXH25
sharpes	0.25	0.55	0.62	0.69	0.40	0.22
MaxDD	0.77	1.35	0.90	0.32	0.81	0.18
Beta	2.57	3.76	2.92	1.49	2.85	1.00
APY	18.05%	32.77%	29.28%	18.66%	23.73%	5.84%
Alpha	1.13	5.46	4.37	1.83	2.46	N/A
Standard Deviation	0.932	1.273	0.909	0.354	0.932	0.201

Table 2. Performance metrics of OMXH2018 data

We tested mean of sensitivity parameters between 0 and 0.99. Rather guessing which sensitivity parameter is right, one strategy could be run multiple algorithms with all sensitivity parameters at once and collect mean returns. With a stock of OMXH25 in 2018, this strategy performed better than the OMXH25 index with all metrics except beta. Beta was over 1 which means that algorithms market risk is high. However Beta being high it means that it was able to generate a substantial excess return. With 2008 stock mean strategy did not perform better than OMXH25 when looking at Sharpe.

In GRAPH 9. PAMR2008 and PAMR2018 standard deviation and annual percentage yields are plotted. From these graphs, we can see if there is any dominant strategy. We can notice that OMXH25 index being always less risky option than PAMR if measuring risk by standard deviation of returns. But according to this graph, there is no strategy that dominates the OMXH25 index.



GRAPH 9. Annual Standard Deviation of returns to Annual Percent Yield

7 CONCLUSION

In this study, we studied Passive Aggressive Mean Reversions performance in Helsinki stock exchange and how survivorship biased data affects results. The algorithm was simulated with two datasets. One with stocks from OMXH25 early 2018 and OMXH25 stocks in 2008. The time period of 12.03.2008 - 09.03.2018.

When comparing sensitivity analysis of wealth and sensitivity parameter with Bin Li et al’s (2012) results we can notice that PAMR with underlying stocks of OMXH25 seems to behave the same way as in their study PAMR worked with stocks of S&P500. Both cases total wealth did not get better when sensitivity parameter gets closer to 0. Rather there was a sweet spot between 0 and 1. This might be caused due to fact that both indexes are constructed by same methodology. We can make the following conclusion from this: By balancing with portfolio returns and mean reversion strategy creates a potential for excess returns. We assume that positive Jensen's Alpha explains returns generated from the mean reversion property of price. These results support that the mean reversion strategy is profitable between study period.

We can notice that using data that had only survived stocks performed better. And it is something to consider when testing trading strategies. In the dataset of 2018, there is less risk to invest in declining stock when we know that it will survive and be in OMXH25 in future. This affected greatly to returns and risk. That is why we decline all PAMR’s that have used survivorship biased data.

Because we can’t determine optimal sensitivity parameter beforehand we decline all the strategies which use preset sensitivity parameter. In the end, we are left with PAMR strategy which used 2008 data and all sensitivity parameters. This strategy did not perform better than OMXH25 index when considering Sharpe, but it outperformed Index with all other performance metrics. When comparing returns relations to standard deviation we can notice mean strategy or index does not dominate each other. From this, we can make a conclusion that this strategy could be a suitable solution for an investor that is looking to maximize return. From this point mean strategy can be accepted. But if investor is looking for maximizing Sharpe-ratio strategy declined because OMXH25-index gives better Sharpe ratio.

When looking wealth graphs it seems that there are two periods 2008 - 2011 and 2016 - 2018 where mean reversion strategy seems to work well. But between 2011-2016 algorithm fails with both datasets. It raises a question that is market condition different those times that affected performance. Bin Li et al (2012) were interested in methods to estimate market conditions where mean reversion strategies work. Because like in their backtests same happened in this thesis. There are periods where mean reversion strategy is not optimal. Studying a way to estimate mean reversion and how to adaptively change sensitivity parameter parameter goes beyond this thesis, but that subject is a great continuum for my master thesis.

REFERENCES

Abarbanell, J.S. and Bushee, B.J. (1997), "Fundamental analysis, future earnings, and stock prices", Journal of Accounting Research, Vol. 35, No. 1, pp. 1-24.

Aldridge, I. (2009), High-frequency trading: a practical guide to algorithmic strategies and trading systems translated by Anonymous John Wiley and Sons.

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J. and Mané, D. (2016), "Concrete problems in AI safety", arXiv preprint arXiv:1606.06565, .

Andrei A. Kirilenko and Andrew W. Lo (2013), "Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents", The Journal of Economic Perspectives, Vol. 27, No. 2, pp. 51-72. available at: https://www.jstor.org/stable/23391690.

Arthur, W.B. (1995), "Complexity in economic and financial markets: Behind the physical institutions and technologies of the marketplace lie the beliefs and expectations of real human beings", Complexity, Vol. 1, No. 1, pp. 20-25.

Auer, P. (2010), "Online Learning", in C. Sammut & G.I. Webb(Eds.), Encyclopedia of Machine Learning Springer US, Boston, MA, pp. 736-743.

Borodin, A., El-Yaniv, R. and Gogan, V. (2004), "Can we learn to beat the best stock", "Can we learn to beat the best stock", Advances in Neural Information Processing Systems, pp. 345.

Bostrom, N. (2014), Superintelligence, 1. ed. ed., translated by Anonymous Oxford Univ. Press, Oxford.

Budish, E., Cramton, P. and Shim, J. (2015), "The high-frequency trading arms race: Frequent batch auctions as a market design response", The Quarterly Journal of Economics, Vol. 130, No. 4, pp. 1547-1621.

Chaboud, A.P., Chiquoine, B., Hjalmarsson, E. and Vega, C. (2014), "Rise of the machines: Algorithmic trading in the foreign exchange market", The Journal of Finance, Vol. 69, No. 5, pp. 2045-2084.

Chan, K.C. (1988), "On the contrarian investment strategy", journal of Business, , pp. 147-163.

Chang, E.P. (2016), Machine Trading : Deploying Computer Algorithms to Conquer the Markets translated by Anonymous John Wiley & Sons, Incorporated, Somerset.

Cover, T.M. (2011), "Universal portfolios", The Kelly Capital Growth Investment Criterion: Theory and Practice World Scientific, pp. 181-209.

Crammer, K., Dekel, O., Keshet, J., Shalev-Shwartz, S. and Singer, Y. (2006), "Online passive-aggressive algorithms", Journal of Machine Learning Research, Vol. 7, No. Mar, pp. 551-585.

Dochow, R. (2016), Online algorithms for the portfolio selection problem translated by Anonymous Springer.

Elton, E.J. and Gruber, M.J. (1997), "Modern portfolio theory, 1950 to date", Journal of Banking & Finance, Vol. 21, No. 11-12, pp. 1743-1759.

Fama, E.F. and French, K.R. (1988), "Permanent and temporary components of stock prices", Journal of political Economy, Vol. 96, No. 2, pp. 246-273.

Golub, A., Keane, J. and Poon, S. (2012), "High frequency trading and mini flash crashes".

Gomber, P. and Gsell, M. (2009), "Algorithmic trading engines versus human traders: Do they behave different in securities markets?", CFS Working Paper, Vol. 2009/10 available at: https://www.econstor.eu/handle/10419/43202.

Hendershott, T., Jones, C.M. and Menkveld, A.J. (2011), "Does algorithmic trading improve liquidity?", The Journal of Finance, Vol. 66, No. 1, pp. 1-33.

Jensen, M.C. (1968), "The performance of mutual funds in the period 1945–1964", The Journal of finance, Vol. 23, No. 2, pp. 389-416.

Kahneman, D. and Klein, G. (2009), "Conditions for intuitive expertise: a failure to disagree.", American psychologist, Vol. 64, No. 6, pp. 515.

Karp, R.M. (1992), "On-line algorithms versus off-line algorithms: How much is it worth to know the future?", "On-line algorithms versus off-line algorithms: How much is it worth to know the future?", IFIP Congress (1), pp. 416.

Kirilenko, A., Kyle, A.S., Samadi, M. and Tuzun, T. (2011), "The flash crash: The impact of high frequency trading on an electronic market", Available at SSRN, Vol. 1686004.

Lee Brodie (2012), Aug 2,-last update, Doug Kass: Kill the Quants & Their Technology Before They Kill Us. Available: https://www.cnbc.com/id/48470448.

Li, B. and Hoi, S.C. (2014), "Online portfolio selection: A survey", ACM Computing Surveys (CSUR), Vol. 46, No. 3, pp. 35.

Li, B., Zhao, P., Hoi, S.C. and Gopalkrishnan, V. (2012), "PAMR: Passive aggressive mean reversion strategy for portfolio selection", Machine Learning, Vol. 87, No. 2, pp. 221-258.

Li, Q., Wang, T., Gong, Q., Chen, Y., Lin, Z. and Song, S. (2014), "Media-aware quantitative trading based on public Web information", Decision Support Systems, Vol. 61, pp. 93-105.

Lo, A.W. (2005), "Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis".

Lo, A.W., Mamaysky, H. and Wang, J. (2000), "Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation", The journal of finance, Vol. 55, No. 4, pp. 1705-1765.

Malkiel, B.G. and Fama, E.F. (1970), "Efficient capital markets: A review of theory and empirical work", The journal of Finance, Vol. 25, No. 2, pp. 383-417.

Markowitz, H. (1952), "Portfolio selection", The journal of finance, Vol. 7, No. 1, pp. 77-91.

Matthew Philips (2012), Knight Shows How to Lose $440 Million in 30 Minutes, Industrial/Management ed. ed., Bloomberg Finance LP, New York.

Merton, R.C. (1969), "Lifetime portfolio selection under uncertainty: The continuous-time case", The review of economics and statistics, , pp. 247-257.

Peter Gomber, Björn Arndt, Marco Lutat and Tim Uhle (2011), "High Frequency Trading
" available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1858626.

Robin Wigglesworth (2018), BlackRock bets on algorithms to beat the fund managers.

Scopino, G. (2015), "Do Automated Trading Systems Dream of manipulating the Price of Futures Contracts-Policing Markets for Improper Trading Practices by Algorithmic Robots", Fla.L.Rev., Vol. 67, pp. 221.

Scott Rasmussen and Eric Loewen (2017), Exceeding Customer Expectations EQES, Inc, Glen Ellyn.

Sharpe, W.F. (1994), "The Sharpe Ratio", Journal of Portfolio Management, Vol. 21, No. 1, pp. 49-58. available at: http://search.proquest.com/docview/195581284.

Sharpe, W.F. (1966), "Mutual fund performance", The Journal of business, Vol. 39, No. 1, pp. 119-138.

Sharpe, W.F. (1964), "Capital asset prices: A theory of market equilibrium under conditions of risk", The journal of finance, Vol. 19, No. 3, pp. 425-442.

Treleaven, P., Galas, M. and Lalchand, V. (2013), "Algorithmic trading review", Communications of the ACM, Vol. 56, No. 11, pp. 76-85.

Tversky, A. and Kahneman, D. (1986), "Rational choice and the framing of decisions", Journal of business, , pp. S278.

Vladeck, D.C. (2014), "Machines without principals: liability rules and artificial intelligence", Wash.L.Rev., Vol. 89, pp. 117.

Appendix 1.

List of stocks in datasets.

OMXH25 2008	OMXH25 2018
Nordea	Nordea
Telia	Telia
Outotec	Outotec
Yit	Yit
Nokia	Nokia
Sampo	Sampo
Kone	Kone
UPM-kymmene	UPM-kymmene
Wartsila	Wartsila
Stora Enso	Stora Enso
Amer Sports	Amer Sports
Elisa	Elisa
Kesko	Kesko
Konecranes	Konecranes
Metso	Metso
Nokian Renkaat	Nokian Renkaat
Outokumpu	Outokumpu
Tieto	Tieto
Cargotec	Cargotec
Fortum	Fortum
Neste	Neste
Removed later after 2008	**Added later after 2008**
OP-Pohjola	Huhtamäki

Sanoma	Metsä Group
Uponor	Orion
Rautaruukki	