UPPSALA
UNIVERSITET

Statistical arbitrage

Can a pairs trading strategy beat a buy-and-hold strategy?

André Aho & Simon Löw

Bachelor's thesis

Department of Statistics

Uppsala University

Supervisor: Lars Forsberg

2022

Abstract

In this thesis, the aim is to investigate whether a pairs trading strategy on Swedish stocks can generate a higher risk-adjusted return compared to a buy-and-hold strategy on a benchmark index. The benchmark index is the OMX Stockholm Benchmark-index (OMXSBPI), which is an index that should reflect the Swedish market in general. With a statistical focus, a trading algorithm is built which is then evaluated on data between the years 2018 to 2021. The statistical concepts this thesis is based on are stationarity and cointegration and it is the Augmented Dickey-Fuller test that forms the basis for being able to test these concepts. The risk-adjusted return for the strategy is evaluated using the popular measure Sharpe ratio, which is then compared to the Sharpe ratio for the OMXSBPI-index. The results obtained in this study can not confirm that the pairs trading strategy is better than a buy-and-hold strategy on the OMXSBPI-index in terms of risk-adjusted return. One indication, however, is that the strategy seems to perform better in conditions when the market is declining. In 2018, the index went down by 7.7060 while the strategy went up by 7.5100 percent. As it is data for only one year, it is not possible to determine whether it is due to chance or a potential edge of the strategy.

Keywords: Algorithmic trading, Quantitative methods, Pairs trading, Cointegration, Sharpe ratio

The code for R will be given upon request: simonlw23@gmail.com & info.andreaho@gmail.com

Contents

1 Introduction

As computers become faster and also able to handle more data, there has been a steady increase in algorithmic trading within the financial markets. Hedge funds, institutions and individual investors are among some of the participants (Zhao and Palomar, 2018). Algorithmic trading uses pre-defined trading instructions and automated execution systems to place orders in the market, with as little human interaction as possible.

The rise of algorithmic trading has allowed even more complex models and increased the speed of execution in particular, where often the aim is to create an edge in the market. Trying to beat the market in terms of performance is a topic that has been debated for years and a disputed one. There are indeed few strategies supported by research that actually perform better than the overall market. Partly because many strategies in fact perform worse than the overall market but probably also because those who have found a profitable strategy rarely share their knowledge. A potential edge however, is not necessarily achieved by a unique strategy but from elaborated trading parameters, efficient execution and a serious risk management (E. P. Chan, 2013).

There are many different strategies in the scope of algorithmic trading. A common strategy, in both research and practical application, is pairs trading. Pairs trading is a kind of statistical arbitrage where the aim is to find discrepancies in relative value between pairs of assets closely related.

The capital asset pricing model (CAPM), introduced by William Sharpe, states that in order to receive higher returns it is necessary to take on more risk (Sharpe, 1991). A common measure of risk-adjusted returns are the Sharpe ratio, invented by the same William Sharpe (Brooks, 2019). In this study pairs trading as an automated trading strategy will be applied to the Swedish stock market. Furthermore, it is of interest to investigate if it is necessary to take on more risk in order to receive higher returns. This is measured by comparing the Sharpe ratio of the strategy with the Sharpe ratio of a buy-and-hold strategy on the OMX Stockholm Benchmark-index (OMXSBPI), a benchmark index that is intended to reflect the overall market in Sweden. This leads to the following question,

can an automated pairs trading strategy on Swedish stocks generate a higher Sharpe ratio than a buy-and-hold strategy in the broad benchmark index OMXSBPI?

First, an introduction to pairs trading will be given to the reader. Then a necessary financial theory is presented followed by the statistical theory of a more technical nature. Next, the data and methodology are described. Lastly, the results are presented, followed by a discussion.

2 Pairs trading

The following section will give a background of the pairs trading strategy.

2.1 Background

Pairs trading was first introduced in the early 1980s by Gerry Bamberger, at the time working for Morgan Stanley (E. P. Chan, 2013). The emergence of pairs trading as an automated trading strategy, however, took place a few years later. Nunzio Tartaglia and his team of mathematicians, computer scientists and physicists, also working for Morgan Stanley, developed pairs trading into an algorithmic trading strategy that had great success within the bank. With advanced statistical methods and automated trading, the strategy looked for arbitrage opportunities between stock pairs which were simultaneously bought and sold in order to generate profits. Despite its success, the team was split up in the end of 1980s and the creators behind the strategy brought the knowledge along. (Vidyamurthy, 2004)

Pairs trading received a lot of attention during the 1990s, when the hedge fund Long Term Capital Management applied the strategy. Among the people behind the hedge fund were Nobel-prize winning economists Robert Merton and Myron Scholes. The fund performed a remarkable 40 percent per annum in the first two years, the third year the fund performed 27 percent which was approximately the same as the US stocks the same year. In 1990 the financial crisis hit Russia and there was a clear outflow from risky asset to more stable investments. Long Term Capital Management took a big hit during this period and in order to avoid a global meltdown, the Federal reserve Bank of

New York announced a rescue package. (Lowenstein, 2000)

Despite its successes and setbacks over the years, pairs trading is still a popular strategy for market participants and research.

2.2 Application

The idea with pairs trading is rather simple. Asset prices are often considered completely random and any attempted prediction will most likely be inaccurate. A successful prediction is more due to chance than skill. Instead of predicting the direction of individual assets, pairs trading aim at finding pairs of assets with similar behaviour where profits occur from the relationship between them, rather than the direction of them individually.

There are several ways to apply a pairs trading strategy. Some well-known methods are

- the time series method
- the stochastic method
- the cointegration method,

to name a few. (Krauss, 2015) They all go under the collective name of statistical arbitrage. In this thesis, the cointegration method will be applied as Krauss (ibid.) argues it is a rigorous method for pairs trading.

The cointegration method is based on finding pairs of assets that are cointegrated, these are usually found among assets with similar characteristics. To ensure that the pairs are in fact cointegrated, statistical tests appropriate for that purpose are performed. A combination of cointegrated assets are created in such a way that the linear combination is not considered fully random anymore. The general idea when looking for similar assets is that they should be similar in such a way that there exists some kind of theoretical relationship between them. The theoretical relationship is often determined by sectors in the cases where stocks are applied. The assets do not need to be moving identically through time, as it is hard to find, but rather hold a relative co-existing relationship. Furthermore, when a pair of similar assets are found, the strategy looks for occasions when the prices of the two, which usually moves in the same pattern, diverge. When this occurs, the strategy simultaneously buy the cheaper asset and sell the more expensive asset and then profits as the two prices converge to equilibrium. (Vidyamurthy, 2004) The strategy is a so called market neutral strategy and it will be further described in the following sections.

3 Financial theory

The following section presents the theory used to describe the financial aspects of the pairs trading strategy.

3.1 Capital asset pricing model

The capital asset pricing model (CAPM) was introduced in the 1960s by William Sharpe. The purpose of the model, at least in part, is to describe the relationship between the expected return of a single asset and the systematic market risk. Since its introduction, the model has been both praised and criticized but it can not be denied that it has made a significant impact on the financial industry. (Jagannathan and McGrattan, 1995) The mathematical expression of the model is

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f), \tag{1}$$

where $E(R_i)$ is the expected return of asset i, β_i is the systematic risk, R_f is the risk-free rate and $E(R_m)$ is the expected return of the overall market. The expression $(E(R_m) - R_f)$ is often referred to as the the market risk premium. (Brooks, 2019)

Equation (1) is also often referred to as the security market line (SML) and it is the graphical representation of CAPM. More specifically it is a linear relationship between asset returns and the systematic risk. (Sharpe, 1991)

For a given level of systematic risk, it is possible to see the expected return of an individual asset. The slope of the line is given by β. Intuitively, the slope gives an idea of how volatile an asset is in relation to the overall market. With a $\beta > 1$, the individual asset will be more volatile compared to the overall market and with a $\beta < 1$, the individual asset will be less volatile compared to the overall market. (Vidyamurthy, 2004)

Sometimes CAPM can instead be written as

$$r_{i,t} = \beta r_{m,t} + e_{i,t}, \tag{2}$$

where $r_{i,t}$ is the return of asset i at time t, $r_{m,t}$ is the return of the overall market at time t, β is still the systematic risk and $e_{i,t}$ is the residual return at time t. (Vidyamurthy, 2004)

A pairs trading strategy consist of both long and short positions. A long position means the purchase of an asset and where a profit occurs when the asset increases in value. A short position, on the other hand, means first selling an asset which is not owned, with the objective of buying it back at a lower price. A profit from a short position occurs when the asset decrease in value. In order to sell something that is not owned, requires the asset to be borrowed first.

The combination of both long and short positions can lead to something called market neutrality. By combining long and short positions in the pairs trading strategy, the aim is to get rid of the β-coefficient in Equation (2). By removing the β-coefficient, or the systematic risk, only the residual return remains. Getting rid of beta completely is difficult in practice but it should, however, be close to zero. According to CAPM the residual return is uncorrelated with the overall market, which means that market neutrality for asset r_i is obtained. Furthermore, CAPM assumes that the expected value of the residuals return equals zero, which leads to a mean-reverting behaviour. (ibid.) This mean-reverting behaviour fits well with a pairs trading strategy.

Sometimes the term cash-neutral strategy is used instead of market-neutral strategy, this is because the income generated by the short position is used to finance the long position, which means that the net cash value is close to zero. In other words, the strategy is self-financing. (ibid.) This can be beneficial from, for example, an accounting perspective.

3.2 Sharpe ratio

The Sharpe ratio, denoted S_r, is a measure of risk-adjusted performance. It was introduced by William Sharpe in 1966 and is today a widely used measure when evaluating the performance of an investment portfolio (Brooks, 2019). The mathematical expression is

$$S_r = \frac{r_p - r_f}{\sigma_p}, \tag{3}$$

where r_p is the return of a specific portfolio, r_f is the risk-free return and σ_p is the standard deviation of the portfolio. The expression $r_p - r_f$ is also referred to as the excess return. (E. P. Chan, 2021) The Sharpe ratio is maximized when the excess return of the portfolio is high and the standard deviation of the portfolio is low. In other words, a high Sharpe ratio is desirable.

E. P. Chan (ibid.) highlights one common issue when calculating Sharpe ratios, even among professional market participants, that is whether the risk-free rate should be deducted from the return of a cash neutral strategy. The simple response is no. As described in the previous subsection, the cash neutral strategy is self-financing. In practice this means that the finance cost is trivial as it equals the spread between debit interest rates and credit interest rates. Due to this, the risk-free rate can be ignored in the back-test.

3.3 Z-score

The pairs trading strategy is a long-short strategy, sometimes also referred to as a spread trading strategy. This means that both long and short positions being taken simultaneously, but it is also possible to see it as the spread between the two being traded. The spread can mathematically be expressed as

$$z_t = stockA_t - \beta stockB_t, \tag{4}$$

where z_t is the spread, $stockA_t$ is the log-price of the first stock at time t, $stockB_t$ is the log-price of the second stock at time t and β is the coefficient retrieved by running a regression on the pairs of log-stock-prices, also called the hedge ratio. (E. P. Chan, 2013).

Zhao and Palomar (2018) argues that in the implementation of a pairs trading strategy there are an investment budget constraint to consider. The investment budget constraint can be divided into two different constraints, the first one is the cash neutral constraint and the second one is the net budget constraint. The cash neutral constraint refers to the fact that the strategy is self-financing, in other words that the net cash value is zero. The net budget constraint refers to the fact that the net positions in the strategy are separated from zero, instead they equals the budget which is normalized to one.

Equation (4) can be written in terms of a portfolio instead

$$w = \begin{bmatrix} 1 \\ \beta \end{bmatrix}, \tag{5}$$

where w is the portfolio. Furthermore, the profit that occurs from each trade in a pairs trading strategy is often quite small. In practice, leverage may be preferable in order to get higher returns. However, in order to make the results of the strategy comparable, it is preferred to restrict the leverage to 1. This can be written as $||w||_1 = 1$. (Palomar et al., 2019) The result is a normalized portfolio that can be written as

$$w = \begin{bmatrix} 1/(1+\beta) \\ -\beta/(1+\beta) \end{bmatrix}. \tag{6}$$

The meaning of this is nothing but the leverage being split between $stock_A$ and $stock_B$ based on the β-coefficient. In other words it is the normalized cash weights invested in each asset. This leads to the core of the strategy, which is the z-score. The z-score is essential as it generates the trading signals and it is the trading signals that determine when a trading positions is taken. Zhao and Palomar (2018) The z-score is expressed as

$$z_t{}^{score} = \frac{z_t - E[z_t]}{\sqrt{var[z_t]}}, \tag{7}$$

where $E[z_t]$ is the expected value of the spread and $var[z_t]$ is the variance of the spread.

4 Statistical theory

Under this section, the theory needed to comprehend the method is presented. Every subsection should be seen as a building block for understanding the statistical tests.

4.1 White noise

White noise refers to a rudimentary stochastic process that is always stationary. This is described with equation

$$Y_t = e_t, \tag{8}$$

where the term e_t defines the white noise of a time series. Continuing this reading, the term will be seen in many time series equations, referred to as the residual. An assumption of white noise is that it is independently identically distributed (iid). (Cryer and K.-S. Chan, 2008).

A question remains, how would the white noise be identified in the time series? It will be identified trough the autocorrelation function. When both the autocorrelation-, and the partial autocorrelation function (ρ) are equal to 0, for all t on a set significance level, it indicates that only white noise is being observed (Asteriou and Hall, 2016).

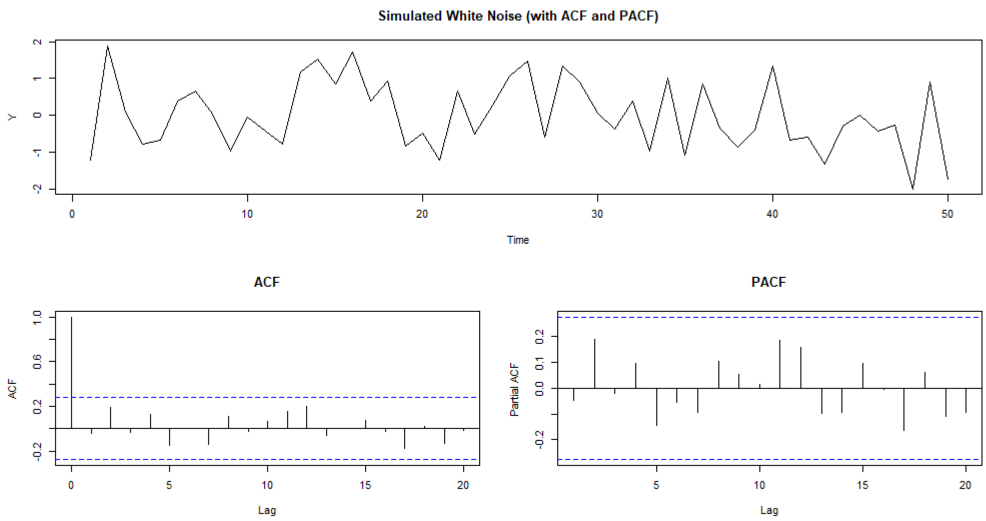

Figure 1: Simulated white noise with proof (Autocorrelation functions)

In Figure (1) assumptions of both stationarity and white noise is visualized. The mean reversion and constant variance can be observed in the plot. The autocorrelation functions are zero on a 5 percent significance level, where the dotted line indicates significance.

4.2 Stochastic processes and random walk

A stochastic process $\{Y_t : t \in T\}$ is a sequence of random variables, taking on values from a space state S and is indexed by a set T, which can be of either a discrete nature, $T = \{0, 1, 2, ...\}$ or of a continuous nature, $T = [\,0, \infty)$ (Wasserman, 2004). A sequence of random variables when indexed by time is also sometimes called a time series process. The collection of time series data is one realization of the stochastic process. (Wooldridge, 2018)

One of the most fundamental stochastic processes is the simple random walk. The random walk is of a discrete nature where the value of the next period in time is derived from the current period in time, in addition to an independent (or at least uncorrelated) error term. (ibid.) A time series $\{Y_t : t = 1, 2, ...\}$ with independent, identically distributed random variables $\{e_t : t = 1, 2, ...\}$, each with mean zero and variance σ_e^2 is given by

$$\begin{aligned} Y_0 &= 0 \\ Y_1 &= e_1 \\ Y_2 &= e_1 + e_2 \\ &\vdots \\ Y_t &= e_1 + e_2 + ... + e_t. \end{aligned} \tag{9}$$

Preferably, it can be written as

$$Y_t = Y_{t-1} + e_t, \tag{10}$$

where e_t can be seen as the magnitude of each step and Y_t is the location of the random walk at time t. (Cryer and K.-S. Chan, 2008)

4.3 Stationarity

Stationarity may refer to different types of stationarity. Strict stationarity can be confirmed if all points Y_t have the same joint distribution for all combinations of t and k in Y_{t-k} (where t = time and k = lag) (Cryer and K.-S. Chan, 2008). This thesis will, however, mainly focus on weak stationarity, here referred to as covariance stationarity since it focuses on the second moment. According to Asteriou and Hall (2016), covariance stationarity indicates three main concepts about the data used for time series. Firstly, it exhibits mean reversion, meaning that the time series holds a constant mean over time. Secondly, the variance is finite and, just like the mean, does not change over time. Thirdly, that the autocorrelation function decreases over time. The main concepts, given the data Y_t, described in simple mathematical terms follows

$$E(Y_t) = \mu, \; for\; all\; t, \tag{11}$$

$$Var(Y_t) = \sigma^2, \; for\; all\; t, \tag{12}$$

$$Cov(Y_t, Y_{t+k}) = \gamma_{0,k} \;\; , \; for\; all\; t \;\; and\; for\; all\; k \neq 0. \tag{13}$$

If a time series is not stationary the results can not be interpreted in the same way as usual and will therefore not be valid. The time series is then spurious. (ibid.) Since this thesis focuses on a mean reverting strategy, stationarity is a main concept to keep in mind. It will however be used in a way where a spurious time series is allowed.
(E. P. Chan, 2021). This will, however, be elaborated under "Cointergration" in section 4.6.

4.4 Unit root

A unit root is a feature which might appear in stochastic processes. To fully understand the concept, the following equation is considered

$$y_t = \mu + \Psi(B)e_t, \tag{14}$$

where $\Psi(B) = \sum_{i=0}^{\infty} \psi_i B^i$, Ψ is the weight, B is the backshift operator and e_t is the white noise at time t. Equation (14) is also called the infinite moving average, MA(∞), which is a general class for all time series that are stationary. Time series that are stationary are the weighted sum of the current and past stochastic "disturbances". (Montgomery et al., 2015) Furthermore, the weights can be assumed to follow an exponential decay pattern, meaning that the disturbances far back in time will have less impact on a time series, compared to newer disturbances. By setting $\psi_i = \varphi^i$ and $|\varphi| < 1$, the decay pattern is guaranteed. Equation (14) can then be written as

$$y_t = \mu + \sum_{i=0}^{\infty} \varphi^i e_{t-i}. \tag{15}$$

From this, the following holds

$$y_{t-1} = \mu + e_{t-1} + \varphi e_{t-2} + \varphi^2 e_{t-3} + ..., \tag{16}$$

the result of combining Equations (15) and (16) is

$$\begin{aligned} y_t &= \mu - \varphi\mu + \varphi y_{t-1} + e_t \\ &= \delta + \varphi y_{t-1} + e_t, \end{aligned} \tag{17}$$

where $\delta = (1 - \varphi)\mu$.

This process is the first-order auto-regressive process, commonly referred to as AR(1). It is an auto-regressive process as y_t regress on y_{t-1}. (ibid.) There are three different scenarios for an AR(1) process.

If $|\varphi| < 1$, the effect of past disturbances will decay as time passes, going towards its equilibrium. The time series is stationary.
If $|\varphi| = 1$, the effect of past disturbances influence the current value equally, this is called a unit root process. The time series is non-stationary. This is a random walk.
If $|\varphi| > 1$, the effect of past disturbances will have bigger impact as time passes. The time series will be explosive. Montgomery et al. (ibid.) mentions this process is of little practical interest however.

As the AR(1) process is non-stationary when $|\varphi| = 1$, there is only one unit root. It becomes stationary by taking the first difference and the process is therefore integrated of order 1, $Y_t \sim I(1)$.

A generic AR(p) process (where p = [1,2,3...]) can be integrated of higher order than 1, which means it takes more than the first difference to make the process stationary. A general mathematical expression is given by

$$\Delta^d y_t = \delta + \varphi \Delta^d y_{t-1} + e_t, \tag{18}$$

where Y_t is the stochastic process, integrated of order d. It becomes stationary by taking the d:th difference, $Y_t \sim I(d)$. (Asteriou and Hall, 2016)

To discover the existence of unit roots in the process, one could perform the augmented Dicker-Fuller test.

4.5 Augmented Dickey-Fuller test

The augmented Dickey-Fuller (ADF) test is an extension of the simple Dickey-Fuller test for unit roots. In the simple Dickey-Fuller test, the error term is assumed to be white noise, which is unlikely in practice. Therefore the augmented Dickey-Fuller test add lags, p, as additional regressors until the error term indeed becomes white noise. (ibid.) With too few lags the errors will suffer from auto-correlation but with too many lags the power of the test will suffer instead (Wooldridge, 2018). There are three different ADF tests.

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + u_t, \tag{19}$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + u_t, \tag{20}$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + u_t, \tag{21}$$

where γ equals $(\varphi - 1)$. Equation (20) considers an intercept, a_0, meanwhile Equation

(21) considers both an intercept a_0 and a non-stochastic time trend $a_2 t$. Furthermore, $\sum_{i=1}^{p} \beta_i \Delta y_{t-i}$ is the sum of the lags together with their coefficients and u_t is the error term. (Asteriou and Hall, 2016)

If the lags of the series, y_{t-1}, does not provide any information regarding the change in y_t then the time series is non-stationary. In other words, a unit root is present. Alternatively if the lags of the series, y_{t-1}, does provide some information regarding the change in y_t then the time series is stationary. Thus the test have the following hypotheses

$$H_0 : \gamma = 1$$
$$H_a : \gamma < 1,$$

where the null hypothesis implies a unit root is present and the alternative hypothesis implies a unit root is not present. (ibid.) The test statistic takes the following form

$$ADF_{obs} = \frac{\hat{\gamma}}{\hat{\sigma}_\gamma}.$$

The test follows a non-standard distribution where a sufficiently negative value leads to the rejection of the null hypothesis.

4.6 Cointegration

Referring to the section 4.3 "Stationarity", it was mentioned that a different kind of stationarity will be used. To find cointegration, a linear combination of two difference stationary variables is calculated. Instead of looking at a moving average or an auto-regressive function, a time series regression is needed to test cointegration between both X_t and Y_t. (Dolado et al., 1990) Consider a time series with the model

$$Y_t = \beta X_t + z_t, \tag{22}$$

now presume that both $X_t \sim I(d)$ and $Y_t \sim I(d)$ (are integrated of the same order), then in general

$$z_t = Y_t - \beta X_t \sim I(d). \tag{23}$$

There are, however, special cases where $z_t \sim I(d-b)$, where b $>$ 0. Then the constant β operates in such a way that the cumulative, of all X_t and Y_t, long run components cancel out; thus, $\beta \neq 0$. The constant β is therefore called the cointegrating vector. (Engle and Granger, 1987)

This should be interpreted as when both Y_t and X_t are integrated of the first order I(1), and the cointegration vector is decided so that $z_t \sim I(0)$, the difference between X_t and Y_t is I(0) or CI(1, 1) (*Cointegrated*). (ibid.)

5 Data

Historical data are a necessity in order to evaluate a pairs trading strategy. It is required for both the statistical tests, as well as the back-test of the strategy. There are various sources for export of financial data and in a variety of price ranges. Some examples are Bloomberg, Refinitiv and Yahoo! Finance, which are in order of price of the products respectively. Bloomberg is the most expensive and Yahoo! Finance is free to use. Throughout this thesis, Yahoo! Finance will be used. Yahoo! Finance was founded in 1997 and it provides financial data on companies and covers nearly 80 different markets. The financial data includes, among others, price data, descriptive information on companies and financial statements. The data provider behind Yahoo! Finance is ICE Data Services. (Yahoo, 2022) There are several reasons to why Yahoo! Finance is used. First and foremost it is free to use which makes it available to many. Furthermore, the data are automatically adjusted for splits and dividends (E. P. Chan, 2021). Without adjustments there would be incorrect moves in historical data, for example a 2:1 split does not mean that a company has lost half its market capitalization, leading to bias in both the statistical tests, as well as the back-test.

This thesis is based on data from the Swedish stock market. In popular parlance, there is often talk about "Stockholmsbörsen". The official name, however, is Nasdaq Stockholm and since 2007 it is a part of Nasdaq Inc (Nasdaq, 2022a). On Nasdaq

Stockholm, there are about 350 companies divided into small cap, mid cap and large cap. It is not possible to include all companies from the three lists, this is mainly due to the fact that not all of them are within the scope of short selling. The stocks that are possible to short at one of the major brokers in Sweden, namely Avanza, are included while the stocks that are not in the scope of short selling are excluded. It is not relevant to include stocks that are not possible to short as it is not practically possible to build a pairs trading strategy based on these. As of 1st of May 2022, there are 149 Swedish companies in the scope of short selling at Avanza (Avanza, 2022). Data for these 149 companies are retrieved from Yahoo! Finance and the data includes company names, dates, sectors and adjusted closing prices. The 149 stocks are divided into 10 different sectors, these are

- Industrials
- Communication Services
- Financial Services
- Healthcare
- Technology
- Energy
- Basic Materials
- Real Estate
- Consumer Defensive
- Consumer Cyclical,

the split of the 149 companies in scope are not equal among the 10 different sectors. However, there are at least two stocks within each sector which makes it possible to build at least one pair. The energy sector includes only two stocks while the rest of the sectors are significantly larger. A detailed table of each company and the 10 sectors can be found in Appendix A (tables A1 to A10).

A risk of having small sectors in terms of number of stocks, such as the energy sector, is that there may be times when cointegrated pairs can not be found in all sectors. It is preferred to have cointegrated pairs in more than one sector at a given point in time. Otherwise, in the emerge of sectors-specific news there can be a severe hit on the profitability of the strategy. For the purpose of this thesis, the split among the sectors are considered satisfactory.

The OMX Stockholm Benchmark-index (OMXSBPI), an index managed by Nasdaq Inc, is used as a benchmark. The index includes a selection of large and frequently traded stocks from most sectors. The purpose of the index is to be a representation of Nasdaq OMX Stockholm and its performance, calculated as price return. The index is re-balanced twice a year and it is composed in such a way that it is possible to replicate. (Nasdaq, 2022b)

To avoid the restriction of short selling on Swedish stocks, an alternative would be so called contract for differences (CFD). An CFD is a derivatives contract which allows an investor to speculate in both price directions of an underlying asset but without trading the underlying asset directly (ESMA, 2013). With CFDs it may be possible to gain access to a larger range of Swedish stocks.

Furthermore, a disadvantage with the choice of data is that Yahoo! Finance is affected by survivorship bias. The reason for this is because Yahoo! Finance does not include stocks that, for some reason, have been delisted. (E. P. Chan, 2021) In the case of the pairs trading strategy, it is not considered to be a problem as a long and short strategy is not affected to the same extent. The back-test would be overestimating the long position but at the same time underestimating the short position. In the case of a directional strategy, a long- or short-only strategy, it could be a problem to consider. (Kelliher, 2022) The benchmark in this thesis is a buy-and-hold strategy on the OMX Stockholm Benchmark-index, which is a directional strategy. Therefore it may be affected by survivorship bias. However, the area of interest is not the performance of the index but rather the performance of the pairs trading strategy. The use of a benchmark index is just to get a perspective on the strategy and a benchmark index whose performance is better than what it should be in reality is a disadvantage of the strategy rather than an advantage. Therefore, the strategy is not emphasized better than it is, as the result will be more conservative.

By choosing another data source, it would be possible to get data that are free from survivorship bias. However, these data are usually expensive according to E. P. Chan (2021). Thus, it is a cost consideration each trader must make.

6 Method

In the following section, the methodological application of the theoretical framework is presented. This shows how the study, in large, is designed and how results are reached.

6.1 Time frame

With the universe of stocks available as described in the data section, the first step is to decide a time frame for the statistical tests, as well as the back-test. The data are split into two, where the first subset is used as a training period and the second subset is used as a test period. The training period forms the basis of the statistical tests as well as the decision of trading rules and signals whereas the test period is used as a validation on unknown data. With the help of the test period, it is possible to investigate whether the strategy developed during the training period has an edge. There is no golden rule for how to split the data but the choice is somewhat arbitrary (Brooks, 2019). In this thesis, there is a 75/25-split between training data and test data, which means that 75 percent of the data are devoted to the training period while 25 percent are devoted to the test period. In other words, for every one year of test data, there are three years of training data.

Furthermore, the training and test period will follow a rolling window, which means that there will basically be four training periods followed by four test periods. The reason for this is to avoid that the back-test, or test period if preferred, consists of excessively old data that may no longer be relevant. Therefore, the first training period will run between 2nd of January 2015 and 29th of December 2017 and will be the basis of the first year of the back-test, which will run between 2nd of January 2018 and 28th of December 2018. The second year of the back-test, which will run between 2nd of January 2019 and 30th of December 2019, will instead be based on the second training period which will run between 4th of January 2016 and 28th of December 2018. The remaining two years are structured in the same way. A visualisation of the split is shown below.

Figure 2: Timeline of training and test periods

It is possible that pairs that once were cointegrated, cease to be so. If this happens, there is a risk that the performance of the strategy will deteriorate. A rolling window can hopefully reduce the risk of this happening.

6.2 Stock selection

Once the time frame is decided, the next step is to find cointegrated pairs that can be used in the back-test. The universe of stocks, as mentioned in the data section, consists of 149 Swedish stocks that are in the scope of short selling at Avanza. The 149 stocks are divided into 10 different sectors and for each sector the maximum number of possible pairs are given by

$$Pairs = \frac{n(n-1)}{2}, \tag{24}$$

where n is the number of stocks in each sector. In reality, the number will likely be smaller. This is due to the requirement of each stock being non-stationary individually, which may not be the case for every stock. Another reason is that not all 149 stocks are listed on the exchange during the first test period. Therefore, the number of stocks available increases over time. In order to find the true number of pairs within each sector, there are statistical tests to be done.

There will be some restriction implemented in the stock selection. In each sector, a stock may be included in only one pair. The reason for this is twofold and will be illustrated from both a theoretical and practical problem. Illustratively, stock A is cointegrated with stock B and stock C. The problem of having a stock in several constellations of pairs is the fact that it may be bought in one constellation and shorted in the other constellation simultaneously. If the long and short positions cancel each other out, it means in practice that no position is taken in stock A. It can also be the case that the long and short position does not completely cancel each other out, but only the ratio is affected, which makes the whole scenario even more difficult to handle. Having combinations of three or more stocks are of high practical value as it opens up for even more trading opportunities, on the other hand, the complexity of the strategy increases significantly as the number of dimensions increases. As such a complex model is considered to be outside the scope of this thesis, each stock is limited to one pair.

In case of a stock being cointegrated with more than one other stock, the selection is based on the strength of the cointegrated relationship. As the selection is based on the training period and not the test period it is not possible to know in advance how this restriction affects the results and is therefore considered as an acceptable approach.

Furthermore, pairs consisting of the same company but different share classes will not be taken into consideration. The reason is because the spread between the two share classes will most likely be small, which results in few or no trading opportunities. From a profitability perspective, these pairs are probably not the best to trade, especially not when taking transaction costs into account.

6.3 Testing for stationarity

The first statistical test is the augmented Dickey-Fuller (ADF) test, which is applied to each stock individually. The ADF-test is used to test if a stochastic process is considered stationary. In the case of the individual stocks, the requirement is for them to be non-stationary and possess a unit root. All statistical tests throughout this thesis is tested on a 5 percent significance level.

Each stock for which a unit root is present will be saved to the universe of stocks possible being cointegrated. The stocks for which a unit root is not present, will not be examined further and therefore also excluded for the current training period. The reason for this is because stocks that are stationary are creating a linear independence, which is not desired as the aim is to find stocks that are stationary as pairs.

There are other tests to see if a time series is stationary, for example the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. The main difference between the KPSS-test and the ADF-test is the establishment of the hypotheses. The ADF-test tests the null hypothesis of a unit root being present and therefore the alternative hypothesis is that a unit root is not present. The KPSS-test on the other hand, test the null hypothesis of a unit root not being present while the alternative hypothesis is that a unit root is present. (Brooks, 2019) This leads to the fact that they have different type I and type II errors. A type I error means rejecting a true null hypothesis, while the type II error means not rejecting a false null hypothesis (Wasserman, 2004).

The type I error of an ADF-test means rejecting the null hypothesis of a unit root being present, although a unit root actually is present. In the case of a KPSS-test, the type I error means rejecting the null hypothesis of a unit root not being present, although there is no unit root. In other words, the ADF-test incorrectly rejects non-stationarity meanwhile the KPSS-test incorrectly rejects stationarity. When applying the statistical test on the individual stocks they are, as mentioned above, required to be non-stationary in order to be examined further. In the case of ADF, the type I error will lead to removing stocks from the universe even though they should be kept while in the case of KPSS, the type I error will lead to keeping stocks in the universe even though they should be removed. When taking this into consideration, the ADF-test is preferable as it is a more conservative approach.

6.4 Testing for cointegration

Once all the non-stationary stocks are found, the next step is to test all the possible pairs in each sector. This will be done with the ADF-test as well but with a slightly different approach than before. First, the time series of the two stocks will be regressed on each other, creating a linear combination of the two. When the linear regression model is created, the ADF-test will be applied to its residuals. Equation (22) can be written as

$$stockA_t = \beta stockB_t + e_t, \tag{25}$$

where $stockA_t$ is the log-price of the first stock in the pair at time t and $stockB_t$ is the log-price of the second stock in the pair at time t. β is the hedge ratio and e_t is the residual at time t (Kelliher, 2022).

As the pairs trading strategy in this thesis is built on the idea of pairs reverting back to equilibrium, stationarity of the residuals is essential. It is thus the opposite of what is desired for the individual stocks. Therefore, for this application it is a good sign to reject the null hypothesis in favor of the alternative hypothesis as it means that the stocks are cointegrated. The cointegrated stocks are kept and constitutes the universe for which the back-test can be based on.

In this application of the ADF-test, a type I error and type II error take on a different meaning. The type I error for the ADF-test imply that two stocks are considered to be cointegrated even though they are not, which can have a negative impact when running the trading algorithm. As the KPSS-test tests opposite hypotheses, the type I error would be considered more conservative. However, in previous literature, E. P. Chan (2021) and Fabozzi et al. (2010) mention that the ADF-test is a common test in quantitative trading and thus it will be used in this thesis.

Another test for cointegration that is widely used is the Johansen's test. Johansen's test allows to test if more than two stocks are cointegrated and then it is possible to create combinations of three or more stocks. (E. P. Chan, 2013) This increases the complexity of the cointegration test drastically and is therefore outside the scope of this thesis. Due to this, Johansen's test will not be considered as the best test.

6.5 Back-test

A common disclaimer associated with financial data reads "past performance is no guarantee of future results", which is true. However, any trading idea should be tested on historical data before trading with real money. It is true that a strategy that has been profitable historically is necessarily not profitable in the future, but it is the best way to get an indication of how well a strategy may perform. A strategy that does not perform well on historical data is probably not worth spending neither money nor time.

The back-test of the pairs trading strategy in this thesis will run over a four year period split in four, the trading years are between 2018 and 2021. The trading year of 2018 will be based on statistical tests with data between 2015 and 2017, the trading year of 2019 will be based on statistical tests with data between 2016 and 2018 and so on. This means that cointegrated stocks traded in the first year will not necessarily be traded in the following years. Some stocks that are cointegrated in the first period might lose the cointegration properties, while stocks that are not cointegrated in the first period may be cointegrated later.

The trading algorithm is built on three important components, which are implemented in the back-test. These are

- divergence thresholds
- convergence thresholds
- stop-loss thresholds

and will be explained in detail one by one.

6.5.1 Divergence thresholds

When the z-score is at or above the high threshold, the spread is shortened. By shorting the spread, stock A is being shorted and stock B is being bought. Opposite position occurs when the spread is at or below the low threshold, then a long position of the spread is taken. The long position of the spread means stock A is being bought and stock B is being shorted. To clarify, when a long or short position is taken in the spread, in practice it is still both a long and short stock position that is taken simultaneously. However, it is easier to visualize and interpret the positions in terms of the spread and thus it will be described as such in the text that follows.

The divergence threshold is described either as a high divergence threshold or as a low divergence threshold. These are defined as below

- High divergence threshold: 1.00
- Low divergence threshold: -1.00.

The area between the thresholds can be described as a corridor and it can be argued that a narrower corridor is preferable in order to not miss any trading opportunities. On the other hand it would mean that the strategy trade more frequently, which comes together with higher transaction costs. It can also be argued that a wider corridor is better due to fewer trades and thereby lower transaction costs, on the other hand it would mean that even more trading opportunities are lost. Hence, the above thresholds are arbitrarily set but it is considered a good compromise between trading opportunities and transaction costs.

6.5.2 Convergence thresholds

Once a trading position is taken, whether it is being long the spread or being short the spread, it is important to know when to exit the position. Ideally, once a long or short position is taken the spread should revert back to equilibrium. As the spread returns, a profit grows as the spread approaches equilibrium. Maximum profit occurs when it is exactly at equilibrium, or in other words, when the spread is zero. It is important to realize that a profit occurs earlier than that, even though it is not the maximum. However, it is not certain that the spread will always reach zero. Therefore, there is a risk of only taking profits as the spread reaches exactly its equilibrium. It is preferable to use, so called, convergence thresholds as well. When the z-score, regardless of the direction of the position, reaches or exceeds the thresholds, the position is closed and a profit is realized. The thresholds for closing profitable trades are as follows

- High convergence threshold: 0.20
- Low convergence threshold: -0.20.

As well as for the divergence thresholds, some trade-offs are required. Having the convergence threshold at zero will result in maximum profit of a particular trade but there is a risk of losing the profit if the spread never reaches equilibrium. By using a corridor the profit will never be maximized for the individual trade but on the other hand more profits will arise in cases when the spread widens before it reaches its equilibrium. The convergence thresholds are arbitrarily chosen as well but it is considered a good compromise between the size of each profit and the total number of profits.

6.5.3 Stop-loss thresholds

The idea of pairs trading is, as mentioned before, that the spread should return to equilibrium. To assume that the spread will always do so is not realistic. There are many different explanations for a spread not returning to equilibrium. The most likely explanation is the fact that the pair of stocks have lost their cointegration. Another explanation is as simple, or as complex, as that the market is often governed by psychological factors that cannot always be explained by mathematical models.

If the spread continues to widen after a position is taken, a loss occurs. The fact that the spread does not return to equilibrium does not necessarily mean that it will never do so. There are two aspects to consider from a risk perspective, the first one is the loss that occurs when the spread widens and the second one is the opportunity cost. As a trader, it is important to manage risk, many times more important than having a unique strategy. Because if the losses become big enough, naturally the trader will sooner or later be out of the game. The opportunity cost is important to take into consideration as well.

Due to this, the strategy also consists of stop-loss thresholds. These are as follows

- High stop-loss threshold: 2.50
- Low stop-loss threshold: -2.50.

Similarly as for the divergence threshold and the convergence threshold, there are trade-offs to consider. Too tight stop-losses will result in positions being closed at a loss too often, perhaps in scenarios when the spread returns to equilibrium shortly afterwards. Too wide stop-losses entails a risk with unnecessarily large losses. Like the other thresholds, the stop-loss thresholds are arbitrarily chosen but they are considered reasonable. They are considered reasonable because when the z-score crosses these thresholds, it is in an area of low probabilities and can be a clear indication that something is wrong.

It can be argued that there should be a stop-loss considering time as well. That is, if a position does not return towards equilibrium within a certain time frame, it should be closed. This is primarily about an opportunity cost. On the other hand, it can be argued that one should not intervene in a strategy if it is really not necessary. The latter approach will be used in the back-test, this to avoid closing positions due to impatience and which then turns out to be profitable.

6.6 Performance evaluation

In the following subsections, measures for evaluating the strategy are presented. Mainly, how the returns are calculated and how the Sharpe ratio is calculated.

6.6.1 Calculate returns

By using the spread of the pair, explained by Equation (4), the returns of each pair can be calculated. Firstly, the first difference is taken on the spread in order to detrend the time series (ΔS). Important to highlight, the detrending is only applied to the calculations of the daily returns and not on the signal generation. Secondly, the trading signals (TS) can take on the values -1 (short position), 0 (no position) and 1 (long position). These are lagged to match the length of the spread. (Aronson, 2007) To get the return of the strategy (Re), the formula

$$Re_t = \Delta^1 S_t * TS_{t+1} \tag{26}$$

is used. The next step is then to calculate the profit and loss (PnL). This is simply done by taking the cumulative sum of the return at day t (Re_t), so that

$$PnL_t = Re_t + Re_{t-1}. \tag{27}$$

With the purpose of making the strategy as realistic as possible, transaction costs need to be deducted from the returns. For every trade, regardless if long or short, a commission of 0.05 percent is taken into account. Given it is two stocks that are traded simultaneously the commission is multiplied by two, meaning a commission of 0.10 percent for each time a position is opened or closed. A round trip thus involves a transaction cost of 0.20 percent.

6.6.2 Calculate Sharpe ratio

As for the returns of the strategy, the Sharpe ratio will be calculated for each year individually as well as for the total period. As justified, the risk-free rate is excluded from the calculations of the strategy. Therefore Equation (3) is written as

$$S_r = \frac{r_p}{\sigma_p}, \tag{28}$$

where r_p is the return of the strategy and σ_p is the standard deviation of the strategy.

The returns are expressed as daily returns in Equation (26). The annualized return is achieved by multiplying the daily returns with the square root of the number of days for the period, the same applies for the standard deviation of the strategy, which initially is expressed as standard deviation of daily returns. To be able to do this transformation, it is necessary to assume that the daily returns are iid.

The Sharpe ratio of the strategy will then be compared to the OMXSBPI-index. However, as the OMXSBPI-index is considered to be a buy-and-hold strategy, which is not a self-financing strategy, the risk-free rate should not be excluded from the calculations. Frennberg and Hansson (1992) argues that the official discount rate of the Swedish Riksbank is considered to be a good approximation of the risk-free rate. During the period between 2018 and 2021 the discount rate was negative or zero and therefore the risk-free rate will be set to zero in the calculations of the Sharpe ratio for the index.

7 Results

In this section, results of the study will be presented in text. The results are, however, also presented in tabular form, which can be found in Appendix A (Tables (A11) to (A16)). A visual representation of the results can be found in Appendix B.

7.1 Finding stationary stocks

While stock prices tend to often be non-stationary, the tests for stationarity shows it is not always the case. By running an ADF-test at the 5 percent significance level, the following results can be presented. In the first period (2nd of January 2015 to 29th of

December 2017, found in Figure (2)), 16 stocks are found to be stationary. For the second period, only four stocks are stationary. The third and fourth period, have 7 and 13 stationary stocks respectively. Once the stationary stocks are identified, they are removed from the remaining tests of that period.

7.2 Finding cointegrated pairs

Following the test for stationarity of stocks, a linear regression is run on all remaining (non-stationary) stocks, creating pairs as shown in Equation (24). Once again, an ADF-test, on both the 5 percent and 1 percent significance level, is used to test for cointegration. This time it is the residuals of the regression being tested. During the first period, a total number of 46 pairs are found to be cointegrated. The second period yields 74 pairs, the third yields 51 pairs and the last period contains 64 pairs. From those a total number of 15 cointegrated pairs are chosen for each period. The 15 pairs are chosen based on two criteria: low significance level, and an (as possible) equal division between sectors. Pairs used for the back-test can be found in Appendix A, in Tables: (A11), (A12), (A13) and (A14).

7.3 Results from the back-tests

In the following subsections the results of the back-test is presented in text format. To find tables, see Table (A15) and (A16). Visuals is represented by all figures in Appendix B.

7.3.1 The aggregated trading period

The strategy generates two years with a positive return and two years with a negative return. The first year is clearly the year with the best performance, followed by the year with the worst performance. The third year is almost flat in terms of performance and the last year have the largest draw-down but recovers towards the end of the year, it ends with a negative return nevertheless. The total performance for the strategy between 2018 and 2021 is 4.4648 percent, which can be seen in Figure (3) below. The standard deviation of daily returns for the whole period is 0.2274 percent.

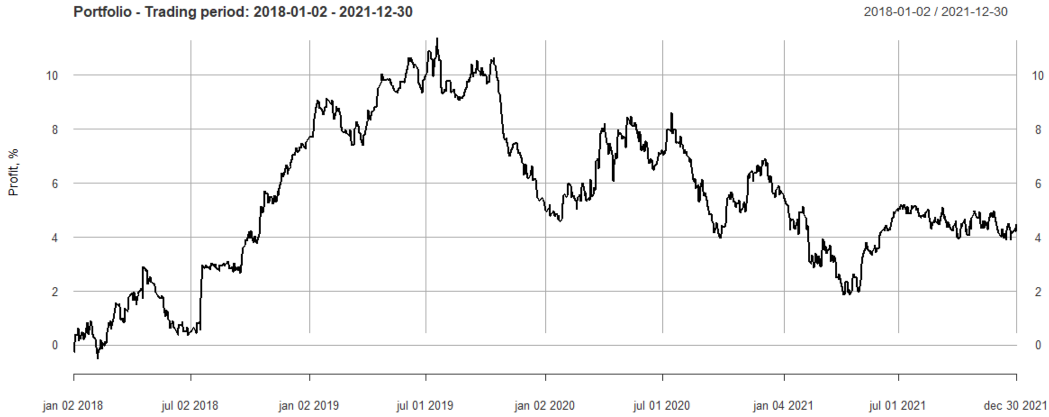

Figure 3: Strategy: all trading periods combined (2018-2021)

The OMXSBPI-index have a total performance of 59.6235 percent and the standard deviation of daily returns is 1.1829 percent. Which can be seen in Figure (4) below.

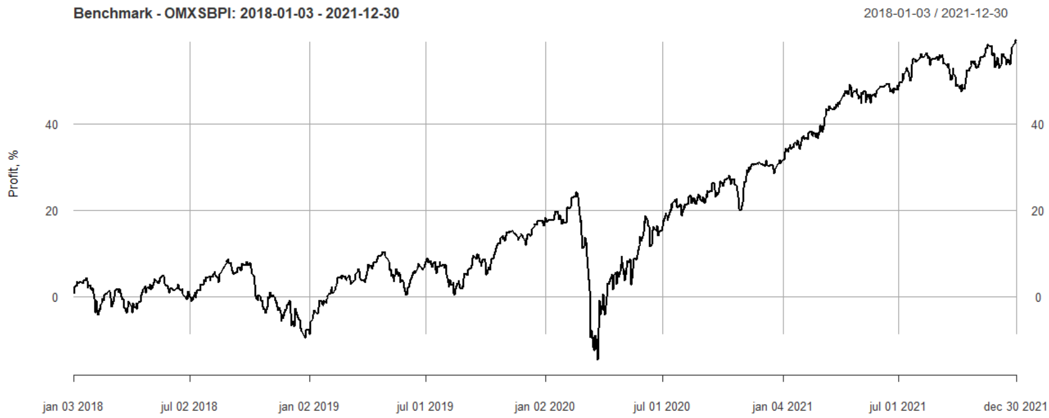

Figure 4: Benchmark index - OMXSBPI: all trading periods combined (2018-2021)

The annual Sharpe ratio for the whole period is 0.6196 for the strategy, which can be compared with a Sharpe ratio of 1.5955 for the OMXSBPI-index.

7.3.2 The trading year of 2018

The return for the trading year of 2018, which runs between 2nd of January and 28th of December, is 7.5100 percent after transaction costs. The standard deviation of the daily returns for the strategy is 0.2143 percent. By looking at Figure (B1) in Appendix B, it can be seen that some of the pairs have positive returns while some of the pairs have negative returns. The best performer is the pair with CellaVision (CEVI.ST) and

Swedish Orphan Biovitrum (SOBI.ST) with a return of 49.7226 percent. The performance for the pair is driven by three periods, as illustrated in Figure (B1), for the rest of the year the performance is flat. The worst performer is the pair with BillerudKorsnäs (BILL.ST) and Rottneros (RROS.ST) and the return for the pair is -12.3343 percent. The negative return is driven by a decline in the beginning of the trading period and remains flat thereafter. The spread hit the stop-loss and do not return during the period. Despite the fact that three pairs yields more than 20 percent, the overall performance of the strategy is neutralized by six pairs with negative returns.

The maximum draw-down of the period is 2.5333 percent and the maximum draw-down duration is from trading day 77 to 112, corresponding to 35 days. There are nine months with positive returns and three months with negative returns. The year consists of 141 days with positive returns and 111 days with negative returns. This can be compared to the maximum draw-down for the OMXSBPI-index being 18.2641 percent and which lasts 83 days.

The Sharpe ratio for the strategy is 2.2077 while the Sharpe ratio for the OMXSBPI-index is -0.5228. However, a negative Sharpe ratio has no intuitive interpretation. The Sharpe ratio for the OMXSBPI-index is based on a return of -7.7060 percent and a standard deviation of daily returns as 0.9341 percent.

7.3.3 The trading year of 2019

The strategy yields -2.1675 percent in 2019, after transaction costs, and the standard deviation of the daily returns is 0.2137 percent. The trading period is from 2nd of January to 30th of December. Figure (B2) shows that the majority of the pairs have a negative return, the worst performer is the pair with Bilia (BILI-A.ST) and Kindred Group (KIND-SDB.ST) with a return of -22.5433 percent. The best performer is the pair with Addnode Group (ANOD-B.ST) and Lagercrantz Group (LAGR-B.ST) with a return of 28.7243 percent. The other pairs with positive returns yields 10 percent or less. Most of the pairs are traded throughout the whole year but the pair with Bure Equity (BURE.ST) and Swedbank A (SWED-A.ST) has a draw-down of 9.9487 percent early in the year and is not traded anymore thereafter.

The maximum draw-down of the strategy during 2019 is 6.0351 percent and the maximum draw-down duration was 111 days. There are five months with positive returns and seven months with negative returns. On a daily basis, there are 125 days with positive returns throughout the year.

The Sharpe ratio is -0.6455 for the strategy. For the same period, the OMXSBPI-index has a Sharpe ratio of 1.8324, which comes from a return of 24.2506 percent and a standard deviation of the daily returns as 0.8438 percent. The maximum draw-down for the OMXSBPI-index is 10.0673 percent, during the trading days 81 to 103.

7.3.4 The trading year of 2020

The pair with Dustin Group (DUST.ST) and I.A.R Systems Group (IAR-B.ST) is the best performer during the trading period of 2020, which runs between 2nd of January and 30th of December, and yields a return of 43.1831 percent. This can be seen in Figure (B3). For the same period, the pair with Astra Zeneca (AZN.ST) and SECTRA (SECT-B.ST) is the worst performer with a return of -22.6041 percent. The negative return occurs in the first half of the year and remains flat for the second half of the year. The majority of all pairs have a negative return but due to four pairs having a return of more than 20 percent each, the overall strategy have a positive return in 2020. After transaction costs, the return of the strategy is 0.3876 percent and the standard deviation of daily returns is 0.2326 percent.

The Sharpe ratio for the strategy is 0.1050 in 2020 and the Sharpe ratio for the OMXSBPI-index is 0.4441. The OMXSBPI-index has a return of 12.4184 percent and the standard deviation of the daily returns is 1.7648 percent.

The strategy yields 125 days with positive returns. On a monthly basis, the strategy has six months with positive returns and six months with negative returns. The maximum draw-down is 4.6462 percent and the maximum draw-down duration is 53 days. Meanwhile, the maximum draw-down for the OMXSBPI-index is 38.9016 percent between the trading days 33 to 56, accounting for 22 days.

7.3.5 The trading year of 2021

The last trading period runs between 4th of January and 30th of December. The return of the strategy during the period is -1.2654 percent after transaction costs and the standard deviation of the daily returns is 0.2126 percent. The return for the OMXSBPI-index is 27.7290 percent for the same period and with a standard deviation of 0.9377 percent.

By looking at Figure (B4), the best performer is the pair with Addtech (ADDT-B.ST) and Lifco (LIFCO-B.ST) with a return of 13.8799 percent. The worst performer is the pair with Biotage (BIOT.ST) and Swedish Orphan Biovitrum (SOBI.ST) with a return of -14.2428 percent. The worst pair have its decline early and remains flat for the most of the year. There are almost a 50/50-split between pairs with positive returns and pairs with negative returns.

The maximum draw-down for the strategy is 3.6452 percent and the maximum draw-down duration is 65 days. There are four months with positive returns and eight months with negative returns. The number of days with positive returns are 126 out of a total of 253 days. For the OMXSBPI-index, the maximum draw-down is 8.9202 percent between trading day 151 and 189.

The Sharpe ratio for the strategy is -0.3742 and the Sharpe ratio for the OMXSBPI-index is 1.8740 during the same period.

8 Analysis

When comparing the returns between the strategy and the OMXSBPI-index during the period 2018-2021, the benchmark is a clear winner. A performance of 59.6235 percent compared to 4.4648 percent leaves no doubt. In terms of risk-adjusted returns the benchmark is a clear winner as well, with a Sharpe ratio of 1.5955 compared to the strategy with a Sharpe ratio of 0.6196.

The year of 2018 is a bad year for the OMXSBPI-index. Between 2019 to 2020 there is a clear bull market, except for the Covid-19 crash of 2020 but which recovers rather quickly. The total performance for the four years leads to the conclusion that it is a good period for the stock market, especially for a buy-and-hold strategy. The performance of 4.4648 percent in four years for the strategy is a track record that is unlikely to attract attention. Especially not when the benchmark yields 59.6235 percent. Of course, it is important to remember that the strategy is market neutral and should therefore not be affected whether it is a bull or bear market, at least not theoretically. However, taking the Sharpe ratio of 0.6196 into consideration as well, it can not be overlooked that a better strategy would probably have been found elsewhere.

The strategy is evaluated during a market for which the majority of the time is considered bullish, where a buy-and-hold strategy usually performs well. It would be interesting to see how the market-neutral strategy performs during a longer period of time, consisting of both bullish and bearish markets. Since the beta coefficient, or the market risk, is neutralized, the strategy should be able to generate positive returns even in a falling market. This is not possible for a buy-and-hold strategy. This could change the results obtained. An indication that the strategy may perform better in a falling market can be seen by studying the year of 2018. The strategy yields 7.5100 percent, compared to the OMXSBPI-index which yields -7.7060 percent. In terms of Sharpe ratio it is 2.2077 for the strategy and -0.5228 for the OMXSBPI-index. However, this is only a year and thus it is difficult to determine whether it is a coincidence that the strategy performs well and significantly better than a buy-and-hold strategy or whether there is an actual edge in the strategy. Another observation is that the strategy seems to handle the Covid-19 crash better, this can be seen in Figure (3) and Figure (4). While the OMXSBPI-index have a draw-down of 38.9016 percent, the strategy stands up fairly well. The strategy has a draw-down as well but it is recovered within the same month, meanwhile the OMXSBPI-index recovers in roughly three months.

There are several possible reasons for the modest performance of the strategy. It may be that the convergence thresholds, divergence thresholds and stop-loss thresholds do not have the best values. Other values may yield higher returns without the risk of the strategy being proportionally higher. Another possible reason is that the windows are being to long, especially the test windows. The test periods are one year each, which is quite long as much can happen in a year. This applies to the cointegration especially, as stocks that are cointegrated may cease to be so.

By studying the figures displaying the performance of each pair, it is clear that there are quite many pairs in each period affected by a decline and then remain flat for the rest of the period. This is the result of the spread hitting the stop-loss and then not returning to equilibrium. Either the spread continues to widen or it constantly oscillates at a higher level. Whatever the reason, it is an indication that the cointegration is lost. Implementing a function that allows a position to be opened only when the spread is returning to equilibrium, reduces the risk of trading a spread where the cointegration is lost. Of course, it is not possible to determine in advance whether the spread is actually reverting. A precaution can be defined as the spread being smaller than a five days average, to get an indication that the spread might be on its way back. Shorter trading periods and more frequent tests for cointegration would probably also reduce the risk of trading pairs for which the cointegration is lost. Another precaution is to apply several tests for cointegration, especially tests that set up hypotheses in different ways. Since the tests for cointegration is performed on a 5 percent significance level, there is always a risk that the outcome is incorrect. In the case of committing a type I error, the strategy trade pairs that in fact are not cointegrated.

Based on the results of the cointegration tests, there are sectors where no cointegrated pairs can be found. The first year of the back-test include no pairs from the energy sector nor the consumer defensive sector while the second year of the back-test include no pairs from the healthcare sector nor the energy sector. The last two years have no pairs in the energy sector. Furthermore, there are several sectors with limited number of cointegrated pairs, which leads to that the number of pairs in the other sectors are being limited as well. The reason for this is to avoid a skewed distribution and overweight in some sectors. Sector constraints may be a reason behind the modest performance.

It should not be forgotten that another set of pairs would probably generate a different result. There are 15 pairs selected for each year and with certain restrictions. For example, a stock may be included in only one pair and no dual class stocks are traded against each other. A more sophisticated trading model which can take into account infinitely many pairs and constellations would probably generate better returns. As previously mentioned, however, such a complex model is outside the scope of this thesis.

An important aspect, regardless of the possible explanations to the poor performance of the strategy and all the possible improvements, it is not certain that an updated version would yield higher returns without also increasing the risk. In other words, it is not certain that the Sharpe ratio will increase. Finally, it should not be overlooked that the strategy actually have some pairs that generates good returns. If it would be possible to minimize the bad trades, the strategy would probably look much better.

9 Conclusion

The aim with this study is to investigate if a pairs trading strategy applied on the Swedish stock market can generate higher Sharpe ratio than a buy-and-hold strategy on a benchmark index, in this thesis the OMXSBPI-index which should reflect the market as a whole. The question asked in the introduction is,

can an automated pairs trading strategy on Swedish stocks generate higher Sharpe ratio than a buy-and-hold strategy in the broad benchmark index OMXSBPI?

Between the period 2018 and 2021, results confirming the question can not be achieved. The pairs trading strategy perform worse than the buy-and-hold strategy on the OMXSBPI-index, in terms of both absolute return and Sharpe ratio. However, the performance of the strategy is evaluated and compared mostly in a bull market and as the strategy is market neutral, it could have an edge in a declining market when a buy-and-hold strategy loses money. The year of 2018 and the Covid-19 market crash in 2020 gives an indication that this may be true but the time period is too short to be able to draw any conclusions. It is also a question of whether the shortcomings that have arisen in the implementation of the strategy have a negative impact on the performance. For example, the fact that some sectors have no cointegrated pairs throughout the whole period and the limitations in the strategy due to complexity, could both have negative impact on the performance.

From a statistical perspective, it may be preferable to test cointegration with several methods. This is to ensure that pairs actually are cointegrated. The biggest risk to a pairs trading strategy is probably to trade pairs where cointegration is not achieved, in the case of this thesis it would mean committing in a type 1 error. Furthermore, there are more sophisticated tests that make it possible to investigate whether combinations of more than two stocks are cointegrated. This is of high practical value and generates more trading opportunities.

Finally, the strategy did not outperform the buy-and-hold strategy on the

OMXSBP-index and is not suitable as a stand-alone strategy. On the other hand, it can be a complement to a broader portfolio of strategies.

10 Further research

A proposal to further studies is to develop a more sophisticated algorithm, which can take into account more dynamic parameters but also the combination of more than two stocks. Furthermore, it would be interesting to investigate longer time periods, consisting of both positive and negative markets environments.

References

Aronson, David (2007). *Evidence-Based Technical Analysis : applying the scientific method and statistical inference to trading signals*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Asteriou, Dimitrios and S. G. Hall (2016). *Applied econometrics*. 3rd Edition. London: Palgrave Macmillan.

Avanza (2022). *Blankningsbara värdepapper*. `https://www.avanza.se/kundservice/kundservice/fragor-svar/handel-vardepapper/blankning/blankningsbara-vardepapper.html`. Accessed: 2022-05-01.

Brooks, Chris (2019). *Introductory econometrics for finance*. 4th Edition. Cambridge: Cambridge University Press.

Chan, Ernest P. (2013). *Algorithmic Trading: Winning Strategies and Their Rationale (Wiley Trading Series)*. Hoboken, New Jersey: John Wiley & Sons, Inc.

– (2021). *Quantitative Trading How to Build Your Own Algorithmic Trading Business*. 2nd Edition. Hoboken, New Jersey: John Wiley & Sons, Inc.

Cryer, Jonathan D. and Kung-Sik Chan (2008). *Time Series Analysis: With Application in R*. 2nd Edition. New York: Springer Science + Business Media.

Dolado, Juan J., Tim Jenkinson, and Simon Sosvilla-Rivero (1990). "COINTEGRATION AND UNIT ROOTS". In: *Journal of Economic Surveys* Vol: 4.Num: 3, pp. 249–273.

Engle, Robert F. and C. W. J. Granger (1987). "Co-Integration and Error Correction: Representation, Estimation, and Testing". In: *Econometrica* 55.2, pp. 251–276.

ESMA (2013). *The European Securities and Markets Authority - Contracts for difference (CFDs)*. `https://www.esma.europa.eu/sites/default/files/library/2015/11/2013-267.pdf`. Accessed: 2022-05-17.

Fabozzi, Frank J., Sergio Focardi, and Petter N. Kolm (2010). *Quantitative equity investing: techniques and strategies*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Frennberg, Per and Björn Hansson (1992). "Computation of a monthly index for Swedish stock returns 1919–1989". In: *Scandinavian Economic History Review* 40.1, pp. 3–27.

Jagannathan, Ravi and Ellen R. McGrattan (1995). “The CAPM Debate”. In: *Quarterly review - Federal Reserve Bank of Minneapolis* 19.4.

Kelliher, Chris (2022). *Quantitative finance with Python : a practical guide to investment management, trading, and financial engineering*. Boca Raton, Florida: CRC Press.

Krauss, Christopher (2015). *Statistical arbitrage pairs trading strategies: Review and outlook*. IWQW Discussion Papers 09/2015.

Lowenstein, Roger (2000). *When genius failed : the rise and fall of Long-Term Capital Management*. New York: Random House.

Montgomery, Douglas C., Cheryl L. Jennings, and Murat Kulahci (2015). *Introduction to time series analysis and forecasting*. 2nd Edition. Hoboken, New Jersey: John Wiley & Sons, Inc.

Nasdaq (2022a). *Om Oss*. `http://www.nasdaqomxnordic.com/omoss?languageId=3`. Accessed: 2022-05-13.

– (2022b). *Overview - OMXSBGI*. `https://indexes.nasdaqomx.com/index/overview/omxsbgi`. Accessed: 2022-05-14.

Palomar, Daniel P., Ziping Zhao, and Rui Zhou (2019). “Optimal Mean-Reverting Portfolio With Leverage Constraint for Statistical Arbitrage in Finance”. In: *IEEE transactions on signal processing* 67.7, pp. 1681–1695.

Sharpe, William (1991). “CAPITAL ASSET PRICES WITH AND WITHOUT NEGATIVE HOLDINGS”. In: *The Journal of Finance* 46.2, pp. 489–509.

Vidyamurthy, Ganapathy (2004). *Pairs trading : quantitative methods and analysis*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Wasserman, Larry (2004). *All of statistics: a concise course in statistical inference*. 1st Edition. New York: Springer.

Wooldridge, Jeffrey M. (2018). *Introductory econometrics : a modern approach*. 7th Edition. Boston: Cengage.

Yahoo (2022). *Exchanges and data providers on Yahoo Finance*. `https://help.yahoo.com/kb/finance-for-web/SLN2310.html?locale=en_US`. Accessed: 2022-05-10.

Zhao, Ziping and Daniel P. Palomar (2018). “Mean-Reverting Portfolio With Budget Constraint”. In: *IEEE transactions on signal processing* 66.9, pp. 2342–2357.

Appendix A (Tables)

Basic Material - All Stocks
BillerudKorsnäs
Boliden
Gränges
HEXPOL B
Holmen B
Lundin Mining Corporation
Rottneros
SCA B
SSAB A
SSAB B
Stora Enso R

Table A1: All stocks used from sector: Basic Material

Communication Services - All Stocks
Embracer Group B
LeoVegas
Millicom Int. Cellular SDB
Modern Times Group B
Paradox Interactive
Tele2 B
Telia Company

Table A2: All stocks used from sector: Communication Services

Consumer Cyclical - All Stocks
Autoliv SDB
Betsson B
Bilia A
Bonava B
Bulten
Clas Ohlson B
Dometic Group
Electrolux B
Evolution
Haldex
Hennes & Mauritz B
JM
Kindred Group
Mips
New Wave B
Nobia
Pandox B
SkiStar B
Thule Group

Table A3: All stocks used from sector: Consumer Cyclical

Consumer Defensive - All Stocks
AAK
AcadeMedia
Axfood
Cloetta B
Duni
Essity B
Midsona B
Scandi Standard
Swedish Match

Table A4: All stocks used from sector: Consumer Defensive

Energy - All Stocks
International Petroleum Corp.
Lundin Energy

Table A5: All stocks used from sector: Energy

Financial Services - All Stocks
Bure Equity
EQT
Handelsbanken A
Handelsbanken B
Hoist Finance
Industrivärden A
Industrivärden C
Intrum
Investor A
Investor B
Kinnevik B
Latour B
Lundbergföretagen B
Nordea Bank Abp
Ratos B
Resurs Holding
SEB A
Swedbank A
TF Bank
Öresund
Arion Banki SDB

Table A6: All stocks used from sector: Financial Services

Healthcare - All Stocks
Ambea
Arjo B
AstraZeneca
Attendo
BioGaia B
Biotage
CellaVision
Elekta B
Getinge B
Humana
Medicover B
RaySearch Laboratories B
SECTRA B
Swedish Orphan Biovitrum
Vitrolife

Table A7: All stocks used from sector: Healthcare

Industrials - All Stocks
ABB Ltd
Addtech B
AFRY
Alfa Laval
Alligo B
AQ Group
ASSA ABLOY B
Atlas Copco A
Atlas Copco B
Beijer Alma B
Beijer Ref B
Bergman & Beving B
Bravida Holding
Bufab
Byggmax Group
Electrolux Proffesional B
Epiroc A
Epiroc B
Fagerhult
Husqvarna B
Indutrade
Instalco
Inwido
Lifco B
Lindab International
Loomis
Mekonomen
Munters Group
NCC B

Industrials - All Stocks (continued)
NIBE Industrier B
Nolato B
Peab B
SAAB B
Sandvik
Securitas B
Skanska B
SKF B
Svedbergs B
SWECO B
Systemair
Trelleborg B
Troax Group
Volvo A
Volvo B

Table A8: All stocks used from sector: Industrials

Real Estate - All Stocks
Atrium Ljungberg B
Castellum
Catena
Fabege
Fast Balder B
Hufvudstaden A
Nyfosa
Platzer Fastigheter Holding B
Sagax B
Sagax D
Wallenstam B
Wihlborgs Fastigheter

Table A9: All stocks used from sector: Real Estate

Technology - All Stocks
Addnode Group B
Dustin Group
Ericsson B
Hexagon B
HMS Networks
I.A.R Systems Group
Lagercrantz Group B
Proact IT Group
Tietoevry

Table A10: All stocks used from sector: Technology

Cointegrated pairs - 2018	
Stock A - 2018	**Stock B - 2018**
BillerudKorsnäs	Rottneros
Fagerhult	Husqvarna B
Bufab	Trelleborg B
AQ Group	SWECO B
Atrium Ljungberg B	Hufvudstaden A
Lagercrantz Group B	Tietoevry
Ericsson B	I.A.R Systems Group
Addnode Group B	Proact IT Group
Modern Times Group B	Tele2 B
Millicom Int. Cellular SDB	Telia Company
Betsson B	Bulten
Nordea Bank Abp	SEB A
Industrivärden C	Investor B
Bure Equity	Öresund
CellaVision	Swedish Orphan Biovitrum

Table A11: Cointegrated pairs 2018

Cointegrated pairs - 2019	
Stock A - 2019	**Stock B - 2019**
SSAB A	Stora Enso R
NCC B	Systemair
Byggmax Group	Svedbergs B
Bufab	Sandvik
Addtech B	SKF B
Castellum	Fabege
Hexagon B	Tietoevry
Addnode Group B	Lagercrantz Group B
Modern Times Group B	Tele2 B
Bilia A	Kindred Group
Betsson B	Thule Group
Handelsbanken B	Lindab International
Bure Equity	Swedbank A
Duni	Swedish Match
AAK	Axfood

Table A12: Cointegrated pairs 2019

Cointegrated pairs - 2020	
Stock A - 2020	**Stock B - 2020**
Gränges	Stora Enso R
Inwido	Systemair
Atlas Copco B	SKF B
Platzer Fastigheter Holding B	Wihlborgs Fastigheter
Hexagon B	Tietoevry
Dustin Group	I.A.R Systems Group
Embracer Group B	Telia Company
Bulten	New Wave B
Bilia A	Kindred Group
Investor B	Swedbank A
Industrivärden A	TF Bank
Handelsbanken A	Kinnevik B
Axfood	Duni
Biotage	RaySearch Laboratories B
AstraZeneca	SECTRA B

Table A13: Cointegrated pairs 2020

Cointegrated pairs - 2021	
Stock A - 2021	**Stock B - 2021**
Boliden	Lundin Mining Corporation
Lifco B	Sandvik
Alligo B	Volvo B
Addtech B	Lifco B
Wallenstam B	Wihlborgs Fastigheter
Catena	Fast Balder B
Addnode Group B	Lagercrantz Group B
Modern Times Group B	Tele2 B
Hennes & Mauritz B	Pandox B
Bonava B	New Wave B
Intrum	Resurs Holding
Bure Equity	Investor A
AAK	Axfood
Elekta B	Humana
Biotage	Swedish Orphan Biovitrum

Table A14: Cointegrated pairs 2021

Results from all periods: Strategy					
	2018	**2019**	**2020**	**2021**	**Combined**
Standard Deviation	0.002143	0.002137	0.002326	0.002126	0.002274
Mean Return	0.000298	-0.000088	0.000015	-0.00005	0.000045
Sharpe Ratio (Annually)	2.207685	-0.645500	0.104963	-0.374224	0.619607
Profit / Loss	0.075100	-0.021670	0.003876	-0.012654	0.044648
Max Draw-down (Max DD)	0.025333	0.060351	0.046462	0.036452	0.095015
Period Start (Max DD)	77	134	134	1	386
Period End (Max DD)	112	245	187	66	817
Total No. of Days	252	247	252	253	1004

Table A15: Results from all periods. Strategy - 2018-2021 & combined results of all years

Results from all periods: Index (OMXSBPI)					
	2018	**2019**	**2020**	**2021**	**Combined**
Standard Deviation	0.009341	0.008438	0.017648	0.009377	0.011829
Mean Return	-0.000310	0.000986	0.000495	0.001114	0.000597
Sharpe Ratio (Annually)	-0.522790	1.832375	0.444149	1.874012	1.595472
Profit / Loss	-0.077060	0.242506	0.124184	0.277290	0.596235
Max Draw-down (Max DD)	0.182641	0.100673	0.389016	0.089202	0.389016
Period Start (Max DD)	165	81	33	151	530
Period End (Max DD)	248	103	56	189	553
Total No. of Days	249	246	251	249	998

Table A16: Results from all periods. Index - 2018-2021 & combined results of all years

Appendix B (Visuals)

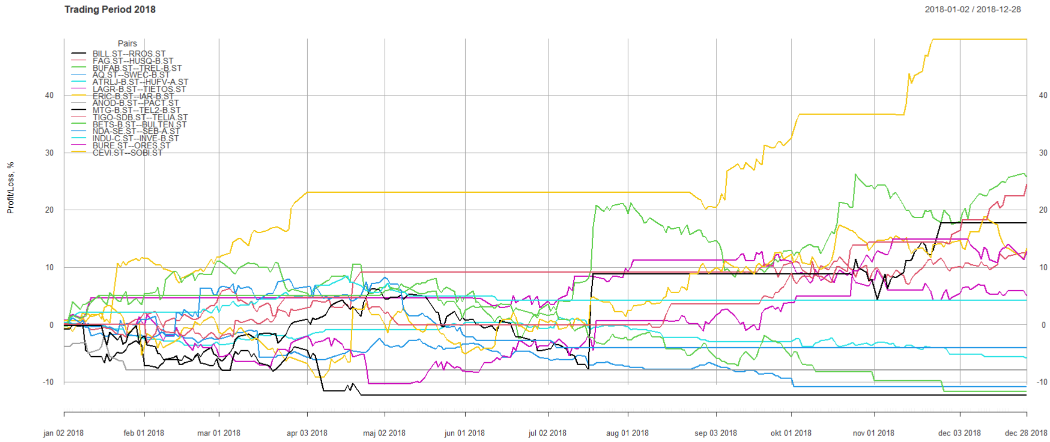

Figure B1: Trading period - 2018 (return for each pair)

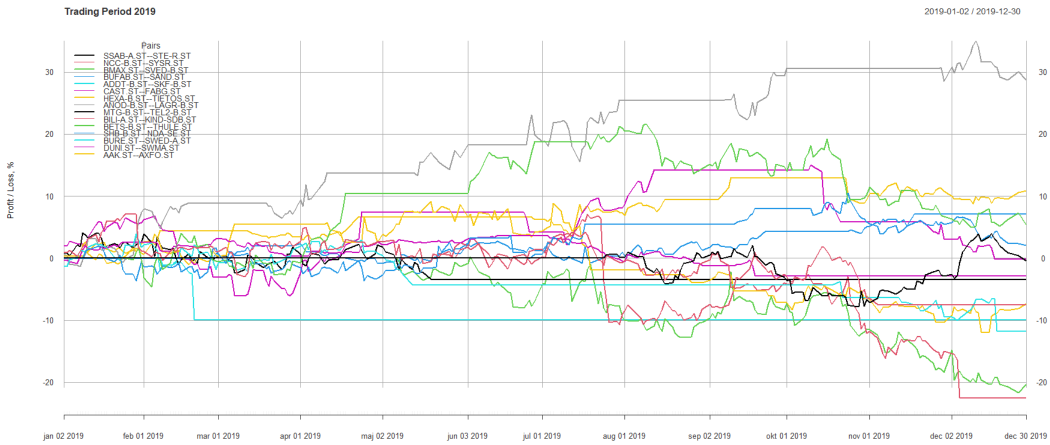

Figure B2: Trading period - 2019 (return for each pair)

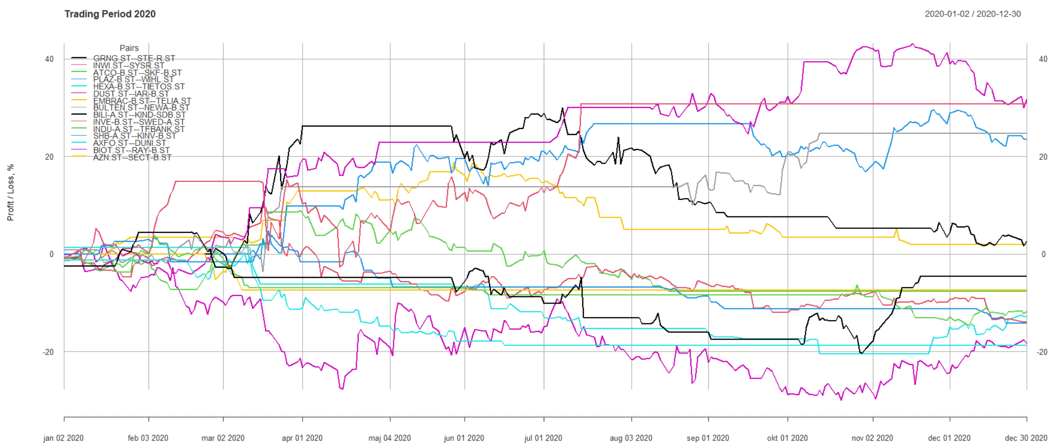

Figure B3: Trading period - 2020 (return for each pair)

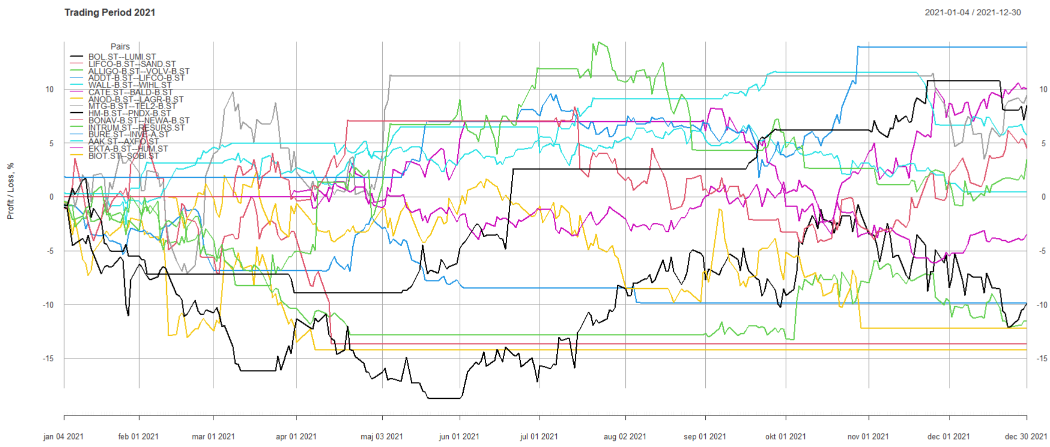

Figure B4: Trading period - 2021 (return for each pair)

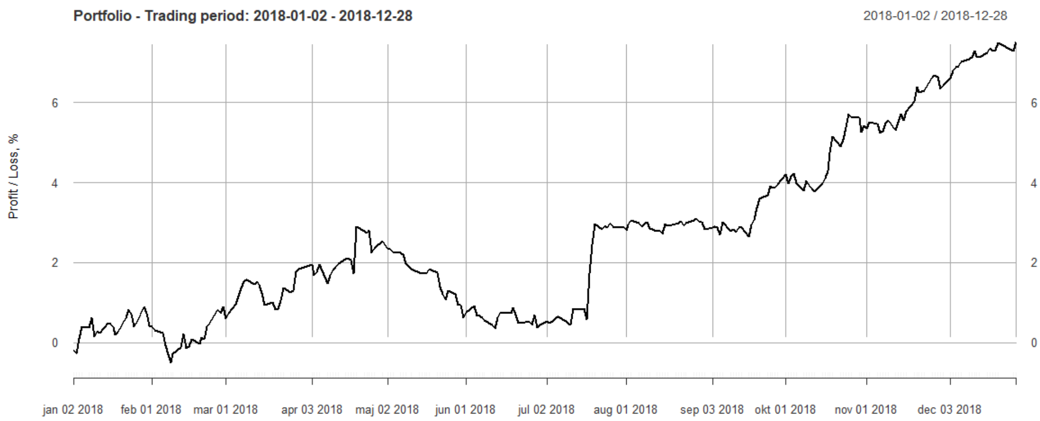

Figure B5: Trading period - 2018 (return for the strategy)

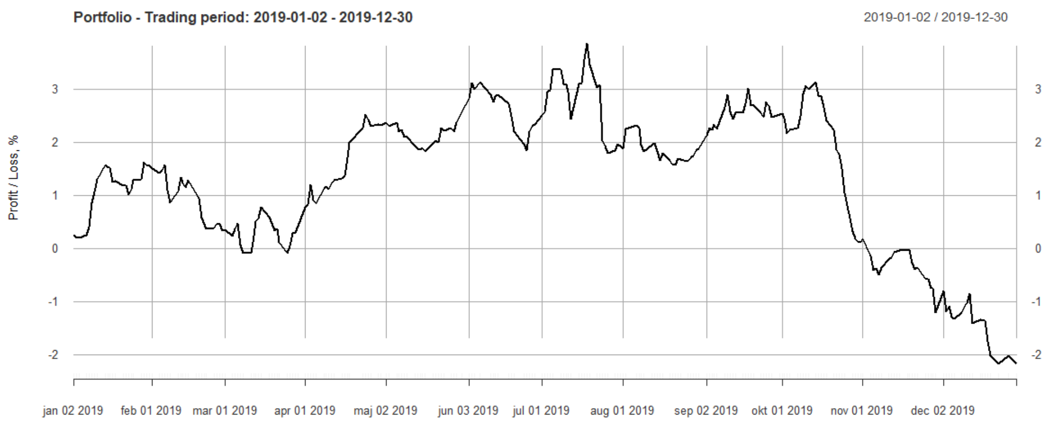

Figure B6: Trading period - 2019 (return for the strategy)

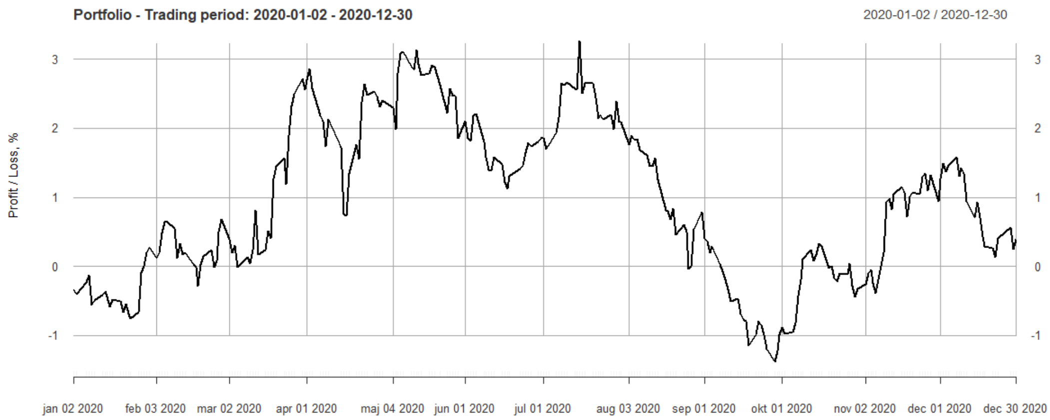

Figure B7: Trading period - 2020 (return for the strategy)

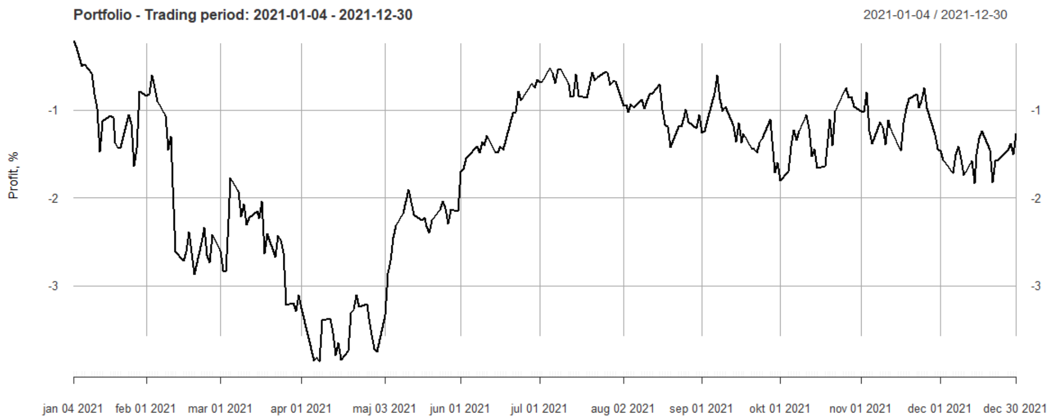

Figure B8: Trading period - 2021 (return for the strategy)