Norges miljø- og
biovitenskapelige
universitet

Master Thesis 2022 30 stp
Faculty of Science and Technology

Exploring Nuances in the Norwegian Equity Market using Pairs Trading

Anders Huse
Indøk Data Science

May 12, 2022

Abstract

In this thesis we explore and provide promising evidence about whether foreign investors have an oversimplified and naive view of the Norwegian equity market. Additionally, results may suggest that some financial factors, especially commodity prices, have a disproportional effect on the Norwegian equity market compared to foreign equity markets.

A new variation of a classical pairs trading framework aided by the field of machine learning is used to explore the nuances of the Norwegian equity market, and how one may be able to profit on these. Results suggest that the strategy performance is closely related to market volatility.

Acknowledgements

Thanks to my supervisors, Jesper and Habib, for giving me great freedom and supporting me throughout the process. Thanks to my father for setting aside time to help out and being an invaluable sounding board in the making of the thesis. Thanks to my twin sister for providing visuals and helping with design issues. Your time is not taken for granted.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Norwegian Equity Market

The Oslo Stock Exchange is the major stock market in Norway, offering the only regulated market for securities trading in Norway. The Oslo Stock Exchange can be further divided in three sub-markets: Osløø Børs, Euronext Growth Oslo, and Euronext Expand Olso. Oslo Børs, constituted by the companies of largest market capitalization (market value of outstanding shares) in the Oslo Stock Exchange, is the market of investigation in this thesis. Insight in the drivers of Oslo Børs is of interest for anyone investing in the Norwegian market, or working in financial institutions.

The Norwegian economy and the Norwegian equity market have historically been dominated by the petroleum industry. Commodity prices such as the price of oil and natural gas have had a large impact on the market. Although petroleum is still a significant part of the market total, the relative share compared to the rest of the market is shrinking. While the relative standing of oil in the Norwegian market is decreasing, that may not be the case of foreign investors presumptions of oil's impact on the Norwegian market. As the combination and the dynamics of the Norwegian market are constantly changing, the views on the Norwegian market by foreign investors may be long-lasting and to some extent sticky.

One of the reasons for proposing this oversimplified and naive view on the Norwegian equity market by foreign investors, is the Norwegian markets relatively small size compared to other global equity markets. Because of its relatively small size, the Norwegian market may not be given as much attention and scrutiny by foreign investors as other, larger exchanges. This, in turn, may lead to foreign investors not having an accurate picture of the nuances of the Norwegian market, and develop a kind of oversimplified view.

1.2 Pairs Trading

As financial markets are getting more efficient, it takes new and more sophisticated approaches to be able to profit from still-existing arbitrage opportunities, if such arbitrage opportunities even still exist. The existence of such arbitrage opportunities have been debated in recent years, as the markets presumably are getting more efficient. This thesis seeks to explore the opportunities for developing a more sophisticated statistical trading model, based on the pairs trading framework.

Pairs trading models the relationship between to financial instruments who have displayed similar development historically, and seeks to exploit opportunities arising when the two instruments deviates from one another [15]. More specifically, one expects the pair to converge in light of a recent divergence. Figure 3.1 in the theory section displays a simple pairs trading scheme of an arbitrary pair. This thesis will go in depth of the causes driving the development of equities constituting a pair, and hopefully be able to give a more qualified guess of the future pair development. As *"Pairs trading questions market efficiency"* [16], this thesis will to some extent assess the market efficiency of the Norwegian equity market.

Pairs trading is chosen to explore the below stated thesis' objectives as several aspects of the strategy coincides with the thesis' objectives, in addition to the strategy's simplicity and ease of interpretability. The relationship between similar securities is at the very core of pairs trading. By exploring these relationships, and the possible drivers behind them, one may be able to extract meaningful insight in what is causing similar Norwegian and foreign equities to showcase short term differences in development. A pairs-trading framework reinforced by new trading rules, aiming at understanding what is driving Norwegian-foreign pair development, is the approach followed in this paper. The reinforcement of a pars trading framework by new trading rules targeting the nuances of the development of the Norwegian market (to be elaborated later), is an approach not seen in previous literature, and may potentially uncover new insight about the Norwegian equity market.

1.2.1 Pairs Trading - Beyond Statistical Modelling

The approach taken in this thesis differ from traditional papers on pairs trading. Where traditional papers focus almost exclusively on optimizing the execution of the trading framework to maximize profits, this thesis puts emphasis on exploring the development of the relationship between two equities constituting a pair, and trying to understand the observed patterns. This alternative approach is chosen to better facilitate exploration of the drivers behind pair relationships and the Norwegian market, as these "drivers" are at the core of the thesis.

1.3 Thesis Objective

Figure 1.1 displays rolling 50 day correlations of the Brent Europe oil price, the ten year Norwegian bond return, and the NOK/USD currency rate with the Oslo Børs Benchmark Index. The time frame of figure 1.1 equals the formation period of the proposed pairs trading framework (to be explained later). The figure displays varying and partly strong correlations of the three key factors to the development of the OSEBX index. These observations showcase that the relationship between the development of potential drivers of the Norwegian market and the Norwegian market itself, are far from stable, and will be subject for exploration later in the thesis.

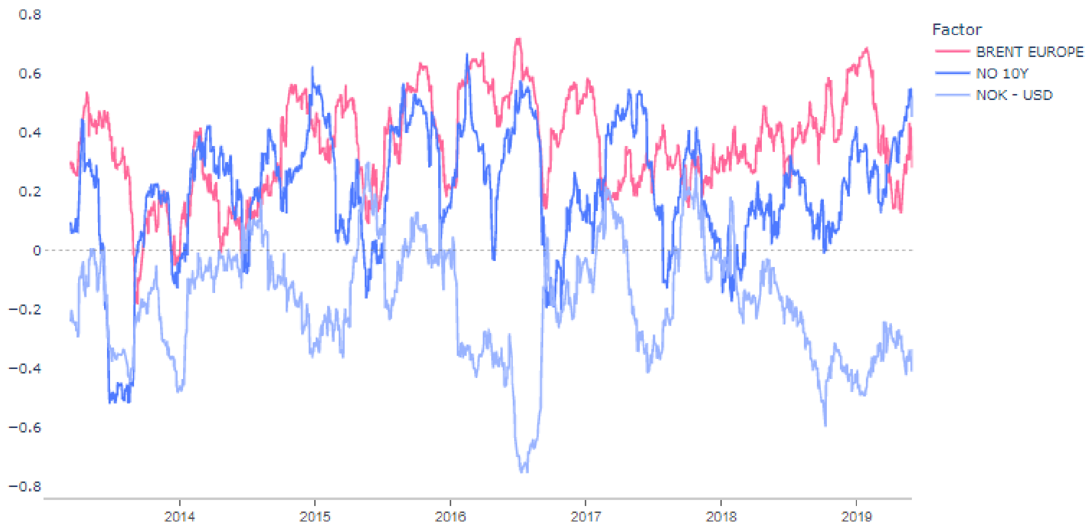

Figure 1.1: 50-day rolling correlation with the OSEBX index for the Brent Europe oil price, the ten year Norwegian bond return, and the NOK/USD currency rate, from 2013 to 2019

How different factors affect the development of the Norwegian equity market may to some extent be attributed to how foreign investors *believe* these factors will affect the market. If foreign investors picture of these factors impact on the Norwegian market are disproportional to the impact the actually have, the reaction on the Norwegian market as a consequence of recent changes in these factors, will be disproportional as well.

The proposed thesis objective focuses on the fundamental drivers or factors of the Norwegian equity market, such as commodity prices and currency rates, and is anchored in the following questions:

“Do foreign investors have an oversimplified or naive view of the Norwegian equity market?”
“Do some financial factors have a disproportional effect on the Norwegian equity market, compared to foreign equity markets?”

1.4 Machine Learning

Recent years have seen an explosive growth in Machine Learning (ML) applications, with financial markets being one of the most prominent, mature and commercialized fields of this innovation. Some recent thesis on pairs trading have made an attempt to reinforce the pairs trading framework by purely predicting the future development of the spread (mentioned in the ”Previous literature” section). Such a prediction is not performed in this thesis. In terms of machine learning, this thesis will focus on an application within the field of unsupervised learning, called clustering. The reason for only applying such a narrow domain of machine learning is not to lose focus on the thesis’ objectives.

The application of clustering within unsupervised learning, is applied in the pair detection phase, to identify equities of similar nature and development. As ML models are often criticized for being ”black boxes” - systems that hide their logic to the user [30], a thorough explanation of the applied clustering models is provided.

”The success of pairs trading, especially statistical arbitrage, depends heavily on the modelling and forecasting of the spread time series. The ability to anticipate the “direction” of this spread is a key point.”[16]

This *“anticipation of the spread”*, or development of the relationship between the two equities constituting a pair, is exactly what this thesis seeks to accomplish with the implementation of additional trading rules.

1.5 Increasing Share of Foreign Investors on Oslo Børs

As the relative share of foreign investors on Oslo Børs is growing, their impact on the trading and market development is also increasing. A thorough understanding of foreign investor behaviour and impact on the Norwegian market becomes of increasing interest and importance.

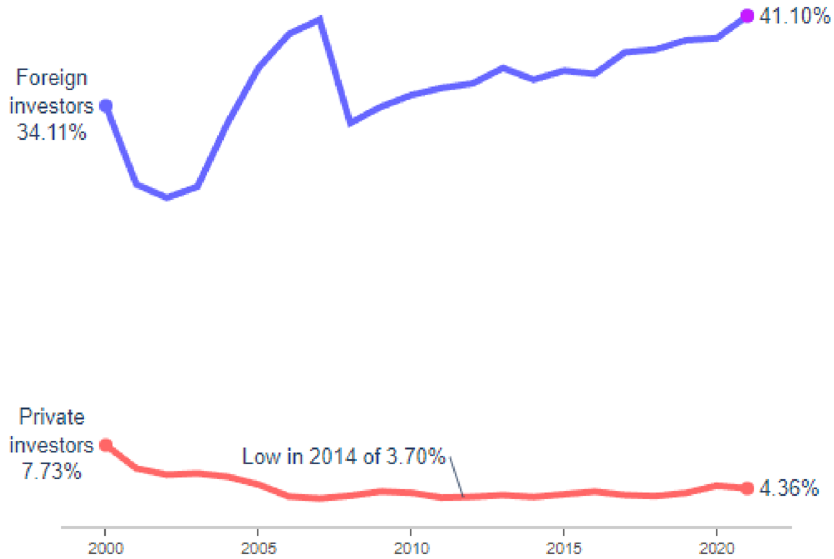

Figure 1.2: Foreign and private investor percentage share of total market value (displayed on the vertical axis) on Oslo Børs since year 2000. The data is obtained from Oslo Børs.

Figure 1.2 visualizes the development of the share of foreign and private investors on Oslo Børs since year 2000, as a percentage of total market value. The percentage of foreign investors have steadily increased over the last two decades, representing 41.10% of the total market value in December 2021. As foreign investors make up such a large and growing part of the total market value, their impact is significant and increasing. In addition to the development of foreign investors, the development of private investors on Oslo Børs has been included in figure 1.2. Although the relative share of private investors has decreased since 2000, it has increased steadily since 2014. The reasoning for including private investors in the visualization is because private (retail) investors will often trade the market more inefficient than well established actors and financial institutions. The confluence of an increasing share of foreign and private (retail) investors on Oslo Børs is especially interesting, as foreign investors' impact increases, and the market may be suspect to increasingly inefficient trading.

The combination of the Norwegian Krone (NOK) being significantly impacted by rising and falling oil prices [6], and the large share of foreign investors in the Norwegian market makes the "Norwegian case" especially interesting. Correlations between the NOK and the oil price may be reinforced by the growing share of foreign investors in the Norwegian market.

1.6 Structure of Thesis

The thesis will start be reviewing some of the most relevant previously written literature about pairs trading, covering the initial (seminal) papers and more recent literature. Then, all theoretical concepts applied in the thesis will be described in detail in the Theory section to facilitate thorough understanding. The emphasis put on the different theoretical concepts will correspond to the relative importance they have in the thesis. After the theoretical foundations have been laid, the approach of the thesis will be presented in the Methods section, covering the entire workflow from data gathering to result inspection. Then, results will be presented and discussed, in the Theory and Discussion sections. In addition to discussing the results, a "Case Studies" sub-section is included within the Results section, to showcase and highlight a handful of interesting case studies. These case studies will be further discussed in the Discussion section. The thesis will round of with thoughts about further work and a conclusion.

Figure 1.3 presents an overview of the workflow of the thesis, from data collection to result inspection. The overview is included to provide an initial idea of how the work with the thesis was conducted. The different elements and terms of figure 1.3 will be elaborated in the following sections.

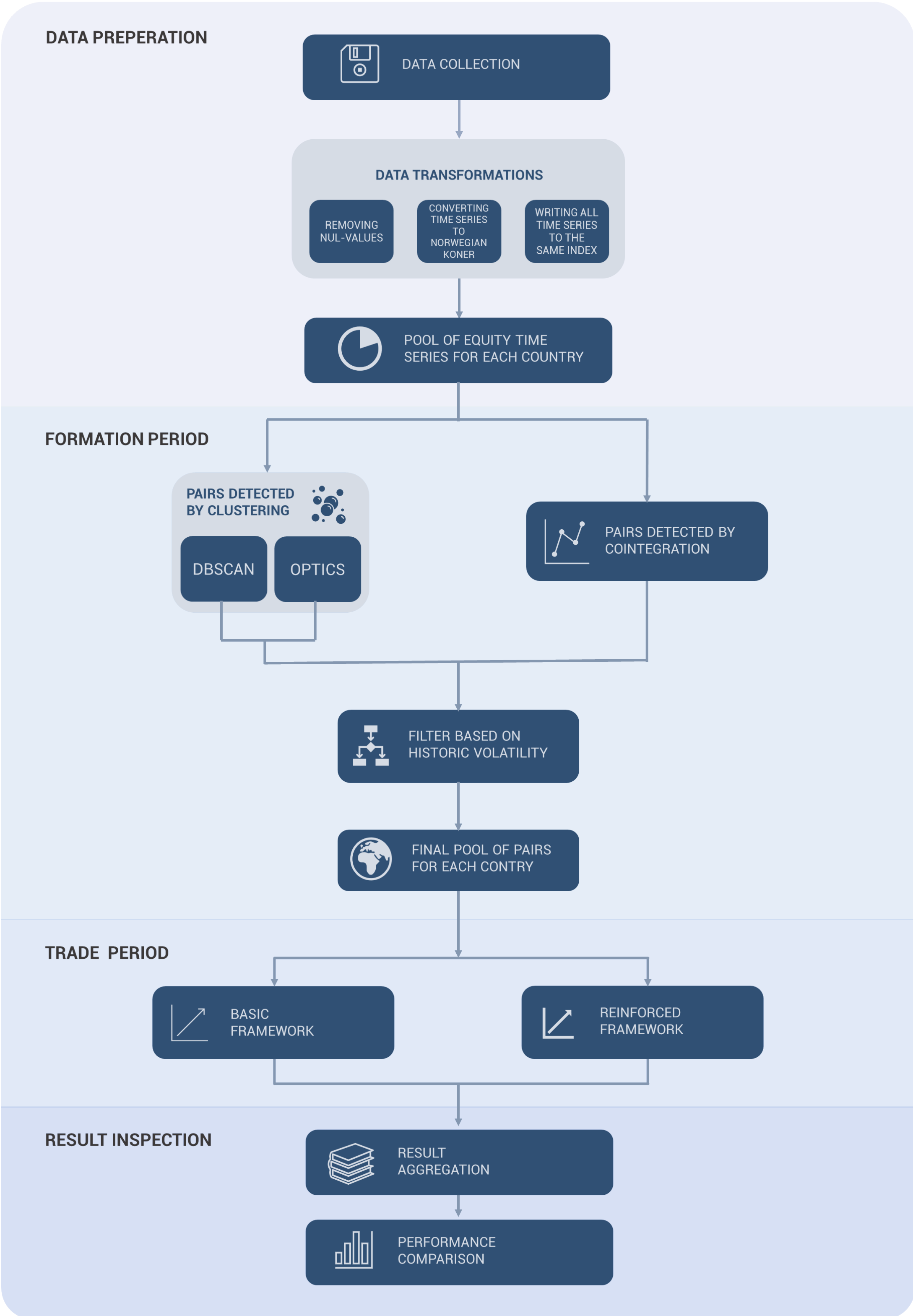

Figure 1.3: Thesis workflow, from data collection to result inspection. After data is collected and transformed, clustering and cointegration calculations are performed in the formation phase. Then, the detected pairs are applied to two pairs trading frameworks over the trade phase, before the results are aggregated and compared.

Chapter 2

Literature Review

The following section reviews relevant literature to put the thesis in "historical" context and provide a compressed timeline of literature on pairs trading. The reviewed literature is presented in two parts: first, some of the initial papers are reviewed, and pairs trading approaches using classical methods as distance and cointegration are assessed. Second, relevant literature regarding recent development in pairs trading approaches, such as the application of machine learning methods as neural networks and clustering, is provided.

The initial pairs trading framework was invented on Wall Street in the mid-1980s by a team led by Nunzio Tartaglia [15]. Tartaglia's program identified pairs of securities whose prices tended to move together and developed trading schemes which took intuition and trader's "skill" out of arbitrage and replaced it with mechanical trading rules. Tartaglia explained the idea behind pairs trading in the following words: "... Human beings don't like to trade against human nature, which wants to buy stocks after they go up not down". This catches the very essence of pairs trading and anchors the strategy as a contrarian strategy (to be elaborated in the theory section).

One of the dominating approaches to detecting pairs for the pairs trading framework was introduced in the 2001 paper by Gatev and Goetzmann [15]: "Pairs trading: Performance of a relative-value arbitrage rule". The 2001 paper is the first significant academic paper on pairs trading and deserves a mention. The paper based the selection of pairs on the distance approach, using daily closing prices from 1962 to 2002. Pairs were detected on a 12-month period and traded in the following 6 months, reporting annual excess returns of 11 percent. Gatev and Goetzmann reported significant returns in the first part of the investigated period, with decaying returns towards the end of the period. The thesis used a variety of risk factors to examine the robustness of the results. In assessing the simplicity of the pairs trading approach, Gatev and Goetzmann wrote: "It is hard to believe that such a simple strategy, based solely on past price dynamics and simple contrarian principles, could possibly make money" [15].

Performing principal component analysis (PCA) and the Ornstein-Uhlenbeck-process on the American stock market, Marco Avellaneda and Jeong-Hyun Lee reported stronger results prior to 2003, in their 2008 paper "Statistical Arbitrage in the U.S. Equities Market", reinforcing the view that profit opportunities from statistical arbitrage decreases as markets mature [5], as suggested by Gatev and Goetzmann. In order to incorporate daily trading volume in the trading signals, Avellaneda and Lee multiplies daily returns by a factor inversely proportional to the trading volume for the past day. This approach accentuates contrarian price signals taking place on low volume and mitigates signals coinciding with days of high volume. The reasoning for this choice of approach is that the authors are less ready to bet against trades occurring on high volume.

In their 2016 paper, Rad, Yew Low and Faff applied several variations of the pairs trading framework, reporting that all strategies performed better during periods of significant volatility, highlighting the cointegration strategy as the superior strategy during turbulent market conditions [27]. Rad, Yew Low and Faff also discussed the seemingly consistent fat left tails of the strategy results, indicating that extreme negative results occur more frequently than corresponding positive ones. As a mean for accommodating unconverged negative trades, stop-loss measures are proposed. The paper attributes the profitability of the strategy profits to investors' under and overreaction to news information [27].

Since the initial exploration of pairs trading on Wall Street, the pairs trading universe has developed rapidly, spanning a variety of different approaches and frameworks. Although the basics are still the same, more sophisticated techniques such as fields of artificial intelligence and machine learning have recently been applied in different aspects of pairs trading.

In the 2020 thesis "Pair Trading on High-Frequency Data using Machine Learning" DaMatta seeks to investigate the application of machine learning (ML) to pairs trading and whether machine learning algorithms can benefit from financial high-frequency data [23]. A recurrent neural network (RNN) is applied in the trading for the purpose of optimizing trade exit rules. The justification of the applied methods is the ability of RNNs to keep a memory of past data and the ability of reinforcement learning to constantly learn (feedback loop) [23]. The strategy shows positive returns, not taking trading costs into account. Some of the proposed portfolios even showed negative returns, disregarding trading costs.

In the 2019 paper “Statistical Arbitrage by Pairs Trading using clustering and ML” it is concluded that pairs trading is still a feasible strategy, but only when machine learning methods are applied. In the thesis, Machine learning is applied in both the formation and trading period. In the formation period, unsupervised learning (clustering) is used to detect stocks of similar characteristics. The clustering method used is Density-Based Spatial Clustering of Applications with Noise (DBSCAN). In the trading phase, a supervised neural network is applied with the objective of trying to predict the next day spread of the pair. If the prediction is in favour of the spread mean-reverting, the trade is executed.

The DBSCAN algorithm is applied in combination with PCA in the 2019 NHH master “Statistical Arbitrage Trading with Implementation of ML”. 67 stocks in the Norwegian stock market were examined over the sample period of 2013 to 2017. The paper concluded that there seems not to be any arbitrage opportunities in the Norwegian stock market, and that pairs trading on OSEBX does not provide excess return nor favourable Sharpe ratio [4].

The paper “Enhancing a Pairs Trading strategy with the application of Machine Learning” applies a recently developed clustering algorithm called OPTICS in combination with PCA to detect pairs in the formation period [31]. To accommodate the challenge of decline periods associated with untimely market positions, a Long short-term memory (LSTM) network is introduced with the objective of forecasting the future development of the spread. The paper concludes that the proposed strategy reduces the average decline period in more than 75 percent of the cases, although at the same time leading to reduced profitability.

Another interesting thesis is ”Pairs trading: the case of Norwegian seafood companies” by Andreas Mikkelsen [25], investigating the performance of a pairs trading strategy on 18 seafood company stocks traded in the Norwegian consumer goods sector, on the Oslo Stock Exchange. Using daily and high frequency data over the period of January 2005 to December 2014, the thesis reported that none of the evaluated strategies had significant profits after accounting for transaction costs. A frequent observation is the high pair non-convergence for both the distance and cointegration approach, reporting that 57% of the pairs never converge. Thus, most of the positions are kept until the trading period ends, and most likely incurs a loss.

Chapter 3

Theory

The following section covers the theoretical concepts applied in the thesis.

3.1 Pairs Trading Fundamentals

Pairs trading, including the concepts of statistical arbitrage, stationarity and mean reversion, is at the very core of the thesis and deserves a thorough explanation.

3.1.1 Statistical Arbitrage

Arbitrage is the practice of taking advantage of different prices for the same product in different markets [3]. An example of a pure arbitrage opportunity is when two stocks of the same company are traded at different exchanges, and the stocks are not traded at the same price over the exchanges. To exploit the arbitrage opportunity, one would buy the relative lowest traded stock, and sell the relative highest traded stock, achieving a profit equal to the difference in prices without taking on any risk at all. Arbitrage opportunities of such obviousness are rare, as financial markets have gotten more efficient over time. Thus, pure, riskless arbitrage opportunities are unlikely to exist in the market, as there is always a risk with engaging in an "arbitrage" trade. A common question is thus if there still exist arbitrage opportunities in financial markets.

Statistical arbitrage refers to arbitrage opportunities where the risk taken is statistically assessed [3]. The core of statistical arbitrage is how to take advantage of temporary mispricings in the market to get a profit, while taking some risk on the way.

Statistical arbitrage covers a broad field of strategies and has a variety of applications.

Common features in the field of statistical arbitrage are [5]:

- The trading signals are systematic and rule based, as opposed to driven by fundamentals.
- The net exposure to the market is neutral, i.e., the market-beta is zero.
- The scheme for generating trading signals is purely statistical.

A relative mispricing follows a deviation from the "law of one price". The "law of one price" is the principle that two similar instruments should be traded at roughly the same price [15]. Whenever one of these instruments deviates from this one price, arbitrage opportunities occur. As Gatev and Goetzmann writes in their 2001 paper, "Pairs trading: Performance of a relative-value arbitrage rule":

"Profits are a compensation to arbitrageurs for enforcing the 'Law of One Price'" [15]

3.1.2 Stationarity and Mean Reversion

A stochastic process is said to be stationary if the mean and variance are time invariant (constant over time). A stationary series is intrinsically mean reverting and fluctuations around the mean should have similar amplitudes [32]. The order of integration is frequently used to refer to whether a process is stationary or not. The order of integration d is a summary statistic which reports the minimum number of differences required to obtain a stationary series [3]. A stationary process is frequently noted I(0) while a non-stationary process is noted as I(1) [32]. Thus, the notion of a stationary process as I(0) refers to the series being stationary, without the need of differentiation.

Stationary processes may be taken advantage of by exploiting temporary deviations from the historical mean, with the expectation that the stationary process will eventually revert back to the historical mean. Stationary processes are seldom found in financial markets, with raw price series often being non-stationary. The concept of pairs trading addresses this challenge, by combining two non-stationary time series into an artificially stationary time series [32]. The resulting time series may thus be treated as stationary series, obtaining the favorable aspects of a stationary process.

The concept of mean reversion is central to statistical arbitrage. The concept assumes that if prices, returns or other economic factors deviates significantly from its mean, they will eventually return to the long-term mean [3]. Thus, a time series is mean reverting if it tends to decrease when the current levels are above the historical mean, and increase if the current levels are below the historical mean. An important distinction of mean reversion is that the concept only applies to changes of a relatively extreme nature, as normal growth and other fluctuations are expected. Detecting the changes of relative extreme nature and separating them from other "expected" fluctuations is a central to the success of pairs trading. Avellaneda and Lee associates the concept of mean reversion to market over-reaction (discussed in subsequent sections) in their 2008 paper:

"The mean-reversion paradigm is typically associated with market over-reaction. Assets are temporarily under-or over-priced with respect to one or several reference securities" [5]

3.1.3 Pairs Trading Inner Workings

Pairs trading is a subbranch of statistical arbitrage, aiming at exploiting temporary mispricing's between two financial instruments. The term "pairs" refers to the simultaneous trading a pair of stocks, assets, portfolios or any two instruments with similar characteristics and who have historically moved together [15]. Pairs trading is performed over a given period of time. This period of time can be compartmentalized into two distinct "periods", one following the other. These two periods are the:

1. Formation period - pairs detection
2. Trade period - trading execution

In short, the formation period is the period when the pairs are formed. This period refers to the part of the time-series on which the pairs have been detected. When picking pairs in the formation period, literature shows a variety of criteria used to restrict the universe of securities to search from and filter out poorly qualified securities for pair detection. Said in other words, criteria to remove stocks not likely to form good pairs. One such criteria can be to exclude securities with low liquidity, as the implementation of the pairs trading strategy requires securities with adequate liquidity. In their 2006 paper, Gatev and Goetzmann screened out all stocks from their database that had one or more days with no trade, thus removing pairs with low liquidity [15]. Saramento an Horta (2021) proposes a unification of pairs selection criteria applied in separate research work [31]. In their approach, a pair is selected if it complies with four different conditions:

1. The pair's constituents are cointegrated.
2. The pair's spread Hurst component reveals a mean-reverting character.
3. The pair's spread diverges and converges within convenient periods.
4. The pair's spread reverts to the mean with enough frequency.

Validating pairs regarding the above conditions, Saramento and Horta aims at enforcing that the pairs' equilibrium persists [31]. The *Hurst component* referenced above measures the degree of mean reversion of a time series [31]. This additional step ensures that undesirable data samples that made it through the cointegration test but are not mean-reverting, are discarded.

The trade period is the period when the pairs trading strategy is performed. As the name suggests, this is the period in which the trading takes place. Modelling the relationship of the two equities constituting the pair may be done in a variety of manners. Some of the most common approaches are using:

- Pair ratio
- Pair spread

The pair ratio is simply defined as one of the equities constituting a pair divided by the other. For two time series S_1 and S_1, the ratio R is defined by the following equation:

$$R = \frac{S_1}{S_2} \tag{3.1}$$

Besides the pair ratio, a pair is associated with a quantity called the spread. This quantity is computed using the quoted prices of the two securities and forms a new time series. The spread represents the relationship between the development of the two equities constituting a pair [14]. The way the spread is defined depends on what technique is used for pair detection. The specific definition of the spread will be elaborated when addressing the different techniques used for pair detection.

Figure 3.1 presents a simple display of the trade phase of a pair trading scheme, using an arbitrary pair. The equities constituting the pair are visualized by the red and blue solid lines. The bottom section of figure 3.1 visualizes the z-score of the pair ratio, with predefined absolute boundaries.

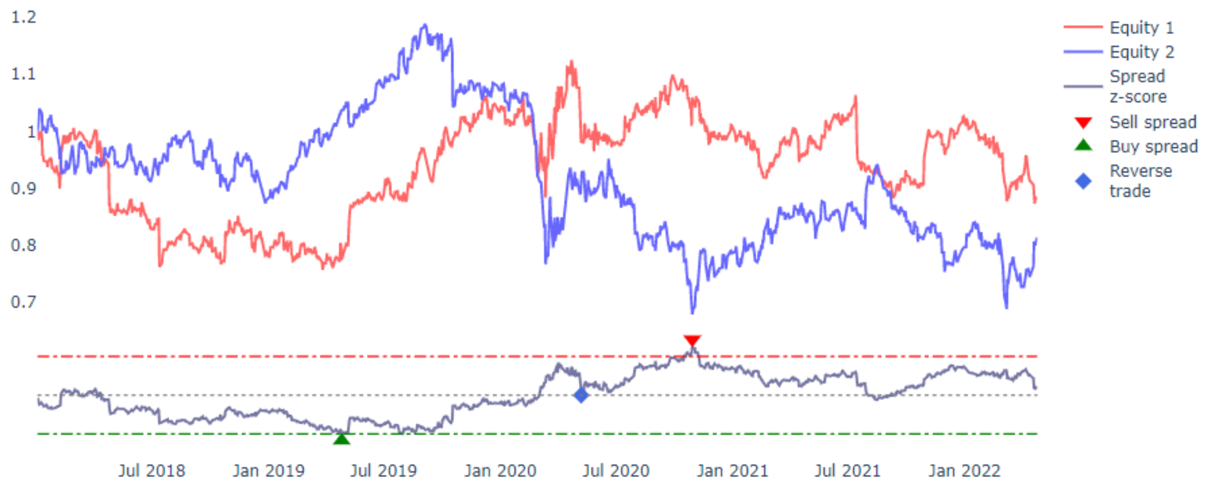

Figure 3.1: Display of the trade phase of a pairs trading scheme for two equities forming an arbitrary pair, and the corresponding ratio z-score. Trade signals for selling, buying and reversing are highlighted.

When the pair ratio crosses through the predefined z-score boundaries, determined based on historical development, one either buy or sell the pair. Buying or selling the pair means buying one equity and selling an equal amount of the other equity. Which equity is bought and sold is based on the how the ratio is composed. When the pair ratio converges back to the historical mean, in the light of a recent divergence, one reverses the trades previously initialized.

3.2 Cointegration

Cointegration and correlation are related, but different concepts. High correlation does not imply high cointegration, and high cointegration does not imply high correlation. A clear distinction between the two concepts is essential, as they should not be confused. While correlation reflects short-term linear dependence in returns, cointegration models long-term dependencies in prices [3]. Correlation does not assess the long-term behavioural relationship between two assets: they may or may not move together over longer time periods, and of such relationships correlation is not an adequate measure. If the spread between to assets is mean-reverting, asset prices are tied together in the long term by a common stochastic trend, and the asset prices are said to be "cointegrated" [2].

Where correlation is based solely on return data, cointegration is based on raw prices. Such raw prices are not normally stationary, they are usually integrated of order 1, de-noted $I(1)$. A set of series are referred to as "cointegrated" if a linear combination of these series results in a stationary series [3]. Two series, x and y, are cointegrated if:

$$x, y \approx I(1), \text{but there exists a constant } \alpha \text{ such that } x - \alpha y \approx I(0) \tag{3.2}$$

In all three equations of section 3.2, c and α are the standard constants of a linear regression, where: $y = c + \alpha x$. ϵ is the error term (the residuals).

Since the series x and y share similar stochastic trends, and since they both are $I(1)$, they never diverge too far from each other (in theory) [4]. All measures of cointegration in the thesis follow the augmented Engle-Granger two-step cointegration test. As the name suggest, the test is divided in two:

1. First, an Ordinary Least Squares (OLS) regression is estimated on the $I(1)$ data.
2. Second, the Augmented Dickey Fuller (ADF) test is applied on the residuals of the OLS-regression, to test for stationarity.

Given two time series, x and y, the Engle-Granger regression is:

$$x_t = c + \alpha y_t + \epsilon_t \tag{3.3}$$

X and y are cointegrated if and only if ϵ is stationary. As measures of cointegration are designed to detect long-run trends in variables, the test will not produce meaningful results or detect stochastic trends if the size of the investigated data period is inadequate.

The Dickey-Fuller (DF) test is an example of a unit root test. Unit root tests are statistical tests of the null hypothesis that a time series is non-stationary against the alternative that the time series is stationary. A simple example of a stationary time series is the process generated by an Auto-Regressive model of order 1, $AR(1)$ [3]. The below equation illustrates an $AR(1)$ model, without a constant term.

$$y_t = \alpha y_{t-1} + \epsilon_t, \text{where } \epsilon \ i.i.d(0, \sigma^2) \tag{3.4}$$

An $AR(1)$ model is only stable if $|\alpha| < 1$, and if such forms a stationary process [3]. The term $i.i.d$ refers to an independent and identically distributed random variable.

3.3 Pairs Trading Selection Framework

There are several possible approaches to pair detection:

1. The distance approach.

2. The stationarity of the price and ratio.

3. Correlation or cointegration between the stock prices'.

4. The copula approach.

Covering all these methods in detail is beyond the scope of this thesis. As the cointegration approach is the approach of choice in this thesis, on background of the ease of implementation and frequent application in relevant literature, it will be described in most detail. The distance approach will also be mentioned, as it may be the simplest approach to pair detection and provides good comparison to the cointegration method. The choice of the cointegration approach in this thesis is aligned with the observations of C. Krauss. Performing a comprehensive review of relevant literature on approaches to pairs trading in his 2017 article, Krauss states that cointegration constitutes a more rigorous framework for pairs trading compared to the distance approach due to the econometrically sound identification of equilibrium relationships [21]. Huck and Afawubo also reports in their 2015 article that while the distance approach generates insignificant excess returns after controlling for risk and transaction costs, the cointegration approach provides a high, stable and robust return [16].

3.3.1 Distance Approach

The distance approach introduced by Gatev et al. [15] proposes a framework for pair selection where pairs are selected to minimize a simple distance criterion. Although Gatev et al. uses the sum of Eucledian squared distances (SSD) the specific distance criterion used varies in literature.

Distance-based approaches provides a simple framework for selecting "good pairs", but in minimizing the distance between two time series, the spread of the resulting pair is often of low variance, limiting the arbitrage opportunities. According to the distance approach, an optimal pair would be a pair with a spread equal to zero over the formation period. This is contradictory to the idea of a potentially profitable pair, optimally having a high spread variance and showing strong mean reversion to facilitate trading opportunities [16].

Following the distance approach, the spread S_t would simply be defined as the difference between the two equities, $S_1$and S_2, constituting the pair [15]:

$$S_t = S_1 - S_2 \tag{3.5}$$

3.3.2 Cointegration Approach

The cointegration approach is based on selecting pairs of equities where the equities constituting the pairs are cointegrated. Given that two securities S_1 and S_2 are cointegrated, then by definition the spread S_t is the series resulting from the following linear combination:

$$S_t = S_1 - \beta S_2 \tag{3.6}$$

β is the cointegrated factor and must be stationary. Assembling the time series in this way is beneficial, since the resulting time series S_t under these conditions is expected to be mean-reverting [32]. The cointegration approach for pairs selection is according to Krauss [21] more rigorous than the previously mentioned distance approach. The reasoning is that selecting pairs based on cointegration results in econometrically more sound equilibrium relationships.

After pairs have been detected using the cointegration framework, the pairs will be traded following a threshold-based trading model. Since the resulting series from equation 3.6 must be stationary, it is expected to have a constant mean over a longer period of time. Thus, when the spread deviates significantly from its mean, action is taken to trade and bet on the mean-reverting of the spread. Upper and lower thresholds are predefined to act as triggers for potential trades. These thresholds are calculated based on the pair development showcased in the formation period. When the z-score of the spread drops below a predefined lower threshold, one goes long the spread. When the z-score of the spread crosses through a predefined upper threshold, one sells the spread short.

The way the spread S_t is defined affects the way positions are set. When the spread is defined following the cointegration approach, one has to decide how the cointegration factor β, is to be handled. Literature shows a variety of different approaches, and there seems not to be consensus on what approach is preferred. This thesis will follow the approach proposed by Dunis et. al. [18], neglecting the cointegration factor to enforce a money-neutral position and invest an equal amount in each instrument constituting the pair [32].

3.4 Contrarian Investing - Overreaction

Overreaction and contrarian investing is a central part of the reason for proposing and exploring the thesis objective, and should therefore be discussed. The concept of overreaction and contrarian investing is summarized by De Bond and Thaler: “The concept of overreaction supports the fundamental concept of a contrarian investment strategy, as overreaction leads to stock prices deviating from the ”true” price, which may be exploited. “ [7].

According to Daniel Kahneman, people seem to make predictions according to a simple matching rule: The predicted values is selected so that the standing of the case in the distribution of outcomes matches its standing in the distributions of impressions [19]. This is an instance of what Kahneman labels the representativeness heuristic, which contradicts the basic statistical principles that the extremes of predictions must be moderated by considerations of predictability.

Pairs trading may be seen as a contrarian investment strategy. Contrarian investment is betting against the market, or investing in contrary to the market, i.e., buying recent “losers” and selling recent “winners” [7]. When buying the stock that has seen a recent decrease in price or selling the stock whose stock price has increased, you are essentially contrarian to the market. The concept of overreaction is central to contrarian investment, often resulting in opportunities to be exploited. The term “overreaction” carries with it an implicit comparison to some degree of reaction that is considered to be appropriate [7]. What is considered an “appropriate reaction” may vary between different market participants.

The concept of overreaction in individual investment has roots back to the renowned investor J.M. Keynes, who stated that: ”... day-to-day fluctuations in the profits of existing investments which are obviously of an ephemeral and nonsignificant character, tend to have an altogether excessive, and even an absurd, influence on the market.” [7]. This *influence* is what is sought to exploit.

The mentioning of pairs trading as a contrarian strategy, pinpoints the reason for the development of the proposed set of new trading rules (mentioned in the introduction). The trading rules are based on the idea of foreign investor overreaction to a handful of fundamental factors central to the development of the Norwegian equity market. The notion of overreaction insinuates that the reaction was larger than what is expected or justifiable [7]. Thus, overreaction will eventually lead to a reversion.

3.5 Machine Learning

Machine learning (ML) is a sub-field of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed and told what to do [8]. The process of model learning, commonly referred to as fitting the model, requires some observations of data (data samples) to be able to explore potential underlying patterns. Although AI may be viewed as complex and non-transparent, the potential learned patterns are nothing more than functions and decision boundaries [9].

The use of ML models enables analyses of massive quantities of data. Data patterns that would be impossible to identify by humans can be accurately extracted using ML models within a short period of time, given the model complexity and available computation power. However, most of the time, accurate results usually require a lot of time and resources.

3.5.1 Supervised and Unsupervised Learning

Machine learning can be further divided into supervised and unsupervised learning. The main difference between supervised and unsupervised learning is the use of labeled data sets; where supervised learning uses labeled data, unsupervised learning does not [8]. Figures 3.2 and 3.3 below aims at showcasing the distinction between the two forms of learning, using animal species as an example.

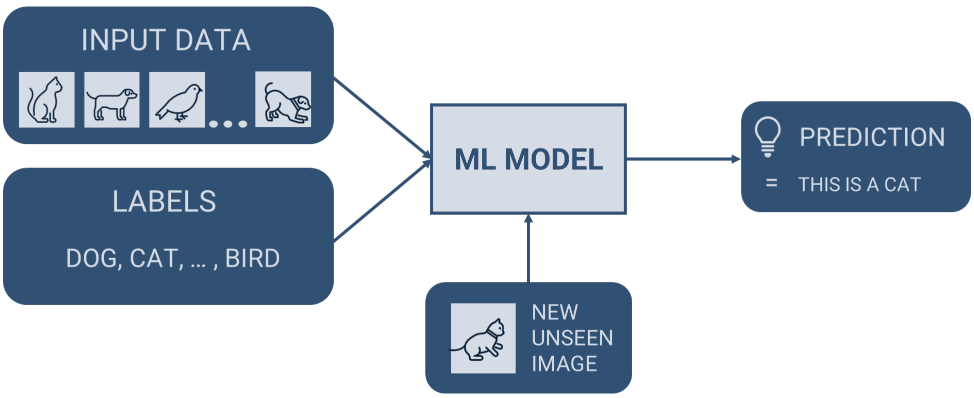

Figure 3.2: Structure of a supervised learning model. The model is provided with input data and corresponding target labels, and produces a prediction when presented with unseen data. Animal images are used as example data.

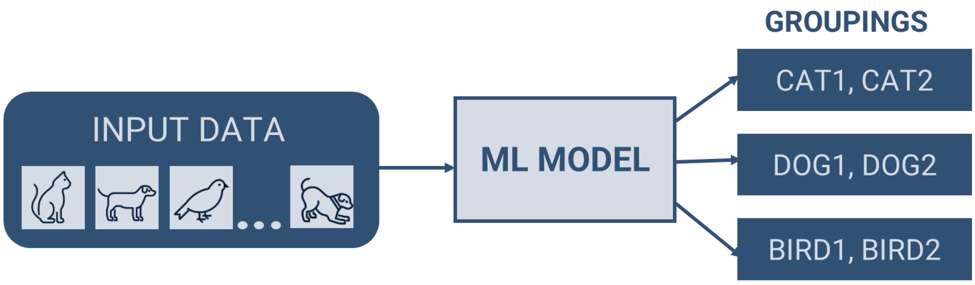

Figure 3.3: Structure of an unsupervised learning model. The model is provided with unlabelled input data and groups the input data according to underlying data structures.

Figure 3.2 visualizes the gist of supervised learning while figure 3.3 visualizes the gist of unsupervised learning. Supervised learning is defined by the use of labeled data. Labeled data sets are used to guide or "supervise" the model in the training phase to help the model classify the target correctly [28]. Having a concrete target label for each data observation it is easy to measure model performance. Supervised learning can be further divided into classification and regression. Classification problems aim at classifying test-data observations into specific labels. Where classification is used to classify observations into discrete labels, regression models are helpful for predicting numerical values based on different data points [8]. The learning process of a regression model aims at understanding relationships between dependent and independent variables [9].

With unsupervised learning there is no need to provide labeled data, the model will figure everything out on its own [9]. This distinction becomes clear when looking at figure 3.2 and 3.3. Unsupervised learning models will learn the inherit structure of the data without using explicitly provided labels, hence "unsupervised". There are several fields of unsupervised learning, of which this thesis applies clustering and dimensionality reduction.

3.5.2 Dimensionality Reduction - PCA

Dimensionality reduction is a widely used data preprocessing technique when performing machine learning, and is the procedure of compressing data from a high number of dimensions or attributes into a lower number of dimensions or attributes, while keeping as much of the variation in the original data as possible [8]. Thus, dimensionality reductions aims at capturing the essence of the given data by finding a compact representation of it.

PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables (p) into a smaller set of linearly uncorrelated variables k ($k < p$), the principal components [32]. The resulting principal components are ordered such that the first principal component accounts for as much the variability in the data as possible. Each succeeding component is then chosen to have the highest variance possible under the constraint that it must be orthogonal (perpendicular) to the preceding component [32].

While some of the variability in the data might be lost when reducing the dimensionality, the reduction of data dimensions yield a variety of advantages [26]:

1. **Decreased training time**. Decreasing the number of dimensions in the data means the training of the data will take less time, thus decreasing the computational resources spent.

2. **Dimensionality reduction accommodates the problem of overfitting**. Data containing too many dimensions or parameters results in more complex models, often too closely adapted to the training data, resulting in poor generalization.

3. **Multicollinearity is addressed**. Multicollinearity occurs when independent variables are highly correlated with other independent variables. Dimensionality reduction combines the correlated independent variables into a set of uncorrelated variables.

4. **Noise in data is reduced**. Noise in the data is reduced by keeping the features explaining the majority of the variability and discarding redundant features.

PCA is used in the preparation of data before the clustering algorithms are applied.

3.5.3 Clustering

Clustering is a sub-field of unsupervised learning, used to extract patterns in the data. The goal of clustering is to find a natural grouping in data so that items (equity time series in this thesis) in the same cluster are more similar to each other than those from other clusters [29]. The number of clusters is data driven; by not specifying the number of cluster beforehand and letting the model figure out the appropriate number itself, no presumptions are introduced in the model [32].

DBSCAN

DBSCAN, short for Density-Based Spatial Clustering of Applications with noise, is the first clustering method applied in this thesis. The general idea behind DBSCAN is to find core samples of high density and expand clusters from them [32]. The DBSCAN algorithm views clusters as areas of high density separated by areas of low density. Thus, clusters detected by DBSCAN can attain any shape. The central component to DBSCAN is the concept of core samples, which are samples in areas of high density. A cluster is therefore a set of core samples, each close to each other (measured by some distance metric) and a set of non-core samples that are close to the core a sample (but are not themselves core samples) [10].

More formally, we define a core sample as being a sample in the data set such that there exist a minimum number of other samples within a defined distance [32]. The number of points within this predefined distance are defined as neighbours of the core sample. This assures that the core samples are located in dense areas. Any core sample is part of a cluster, by definition. The formal definition of a core point is given below.

Definition *core point*: A point q is a core point if it verifies

$$|N_\epsilon(q)| \geq minPts \tag{3.7}$$

where $|N_\epsilon(q)|$ represents the number of points within the ϵ-neighborhood of q, and $minPts$ is the minimum number of points required to form a cluster [32].

Any sample that is not a core sample and is located more than a predefined distance away from any core sample (the same distance used to measure neighbour points), is considered an outlier by the algorithm [32]. Outliers are rejected by the algorithm and are thus not included in any clusters. The distance measure (ϵ), used to assess the neighbourhood of points in a cluster, is one of the central parameters to the DBSCAN algorithm, and is not related to the ϵ in the equations of section 3.2. The formal definition of the application of ϵ is defined below.

Definition ϵ-neighborhood: The ϵ-neighborhood of a point q is defined as

$$N_\epsilon(q) = \{p \in X | d(q, p) \leq \epsilon\} \tag{3.8}$$

where $d(q, p)$ represents the distance between q and p, and X is the set of all points [32].

The above-described parameters and definitions are visualized in figure 3.4.

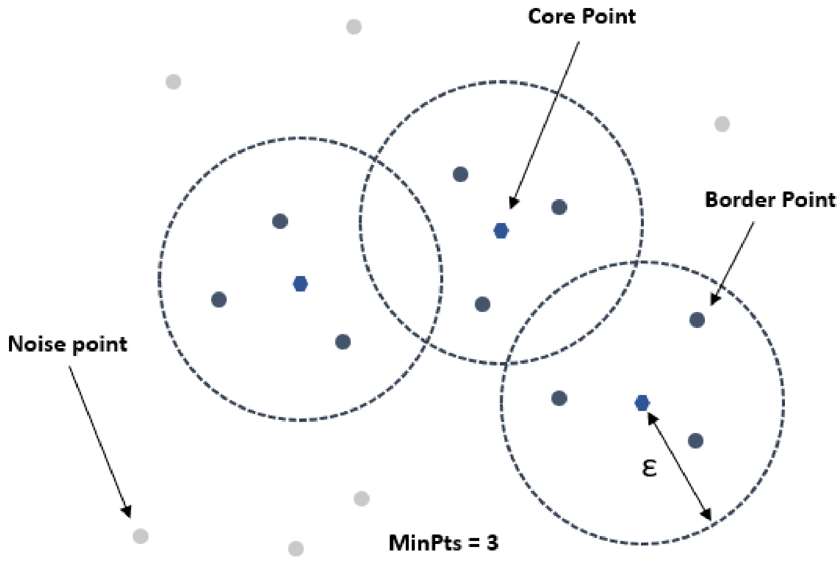

Figure 3.4: Cluster formation using DBSCAN with an arbitrary ϵ and $MinPts = 3$. Blue circles are encapsulating the detected clusters.

While the minimum number of samples needed to form a cluster primarily controls how tolerant the algorithm is towards noise (on noisy and large data sets it may be desirable to increase this parameter), the ϵ parameter is crucial to choose appropriately for the data set and distance function and usually cannot be left at the default value. It controls the local neighbourhood of the points, meaning what points are detected as a part of a cluster [32]. When the distance metric is chosen too small, most data will not be clustered at all. When set too large, the distance metric causes nearby clusters to be merged into one cluster. Eventually the entire data set will be returned as a single cluster, if the distance metric is set high enough. DBSCAN is a good fit for data containing clusters of similar density [32].

OPTICS

OPTICS, short for Ordering Points to Identify the Clustering Structure, is the second clustering algorithm applied is this thesis. OPTICS is closely related to DBSCAN, as it is build on the same functionality. OPTICS can be considered a generalization of DBSCAN that relaxes the distance metric set as the requirement for detection of clusters, from a single value to a single range, thus allowing for detecting clusters of varying density [32]. The key difference between DBSCAN and OPTICS is that the OPTICS algorithm builds a reachability graph, which assigns each sample both a reachability distance, and a spot within the cluster ordering attribute [10]. In addition to the reachability distance, the OPTICS algorithm introduces a concept called *core-distance*. The new concepts are defined in the following, using the same nomenclature as when defining the concepts of the DBSCAN algorithm.

Definition *Core-distance*: Let *minPts-distance(p)* be the distance from a point p to its minPts' neighbour. Then, the core-distance of p is defined as

$$core - dist_{\epsilon,minPts}(p) = \begin{cases} \text{Undefined} & \text{if } |N_\epsilon(q)| < minPts \\ minPts - distance(p) & \text{otherwise} \end{cases} \tag{3.9}$$

The core distance of a point p is the smallest distance ϵ' between p and a point in its $\epsilon - neighbourhood$ such that p is a core point with respect to ϵ [32].

Definition *Reachability-distance*: The reachability-distance of p with respect to o, described as $reach - dist_{\epsilon,minPts}(p, o)$ is defined as

$$reach{-}dist_{\epsilon,minPts}(p, o) = \begin{cases} \text{Undefined} & \text{if } |N_\epsilon(o)| < minPts \\ max(core - distance(o), distance(o, p)) & \text{otherwise} \end{cases} \tag{3.10}$$

The reachability-distance of a point p with respect to a point o can be interpreted as the smallest distance such that p is directly density-reachable from o. This requires that o is a core point, meaning that the reachability-distance cannot be smaller than the core-distance. If that was the case, o would not be defined [32].

The above-defined attributes from equation 3.9 and 3.10 are assigned when the model is fitted to the training data (data from the formation period) and are used to determine cluster membership.

The reachability distances generated by OPTICS allow for variable density extraction of clusters within a single data set, illustrated in figure 3.5 [11]. Figure 3.5 is composed of four smaller plots, of two different types:

1. The top plot (a) combines reachability distances and data set ordering to produce a reachability plot.
2. The three smaller plots (b, c and d, from left to right) below the top plot visualizes

clustering formation for OPTICS and DBSCAN, corresponding to the reachability plot (plot a).

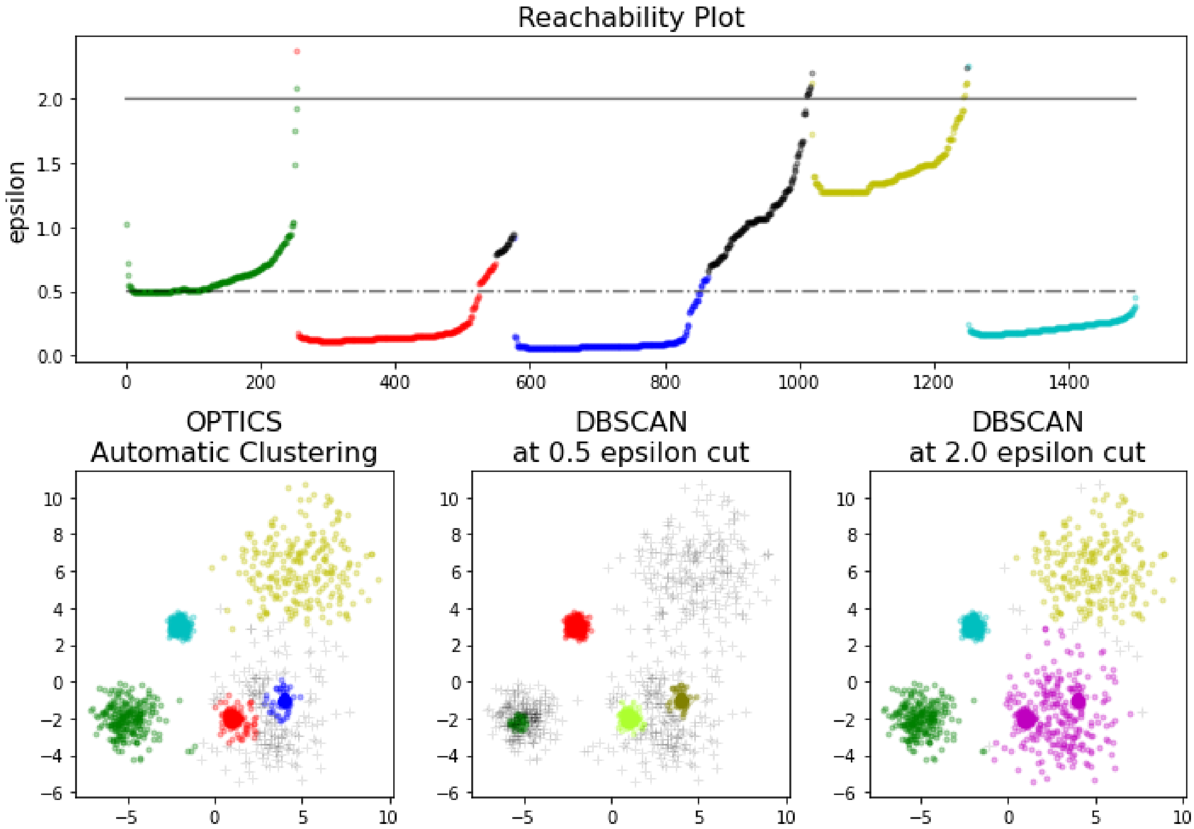

Figure 3.5: Combining OPTICS reachability plot (top plot) and cluster formation visualizations for automatic clustering using OPTICS, and using the DBSCAN algorithm with ϵ equal to 0.5 and 2.0. Source: scikit-learn, Demo of OPTICS clustering algorithm.

As shown in plot a of figure 3.5, combining reachability distances and data set ordering produces a reachability plot, where reachability distance is represented on the y-axis and the ordered points are arranged on the x-axis. Points are ordered such that nearby points are adjacent. Plot a in figure 3.5 detects five clusters of different density, corresponding to the more visual clusters in plot b. Clusters are visualized as different-coloured U-shaped sequences, where clusters with lower y-values (reachability distances) represent denser clusters.

“Cutting” the reachability plot at a single value on the y-axis produces DBSCAN like results; all points above the “cut” are classified as noise, and each time there is a break when reading from left to right on the x-axis (ordering) of the graph, a new cluster is detected [10]. The result of ”cutting” the reachability plot at a single value is visualized in plot c and d of figure 3.5. Setting the reachability distance to a constant value corresponding to the two horizontal lines in the top plot in figure 3.5 results in very different cluster detection.

One caveat of the DBSCAN algorithm is its sensitivity to the specified parameters, specifically to the ϵ parameter. The proper size of ϵ varies from data set to data set, and even after finding the right ϵ for the current data, the DBSCAN algorithm will assume all clusters are of the same density, and will thus not be able to detect clusters of varying density [32]

The ability of the OPTICS algorithm to detect clusters of varying density makes it preferable over the DBSCAN algorithm, especially in the scope of this thesis. There is a natural unbalance of the different sectors of securities included in the initial pool of securities. Assuming the sector of the securities has an impact on the cluster formation, clusters are bound to have varying density. Applying OPTICS to the time series data may thus detect pairs the DBSCAN algorithm would have overlooked.

T-SNE

To illustrate the clusters in a two-dimensional space, the application of T-distributed Stochastic Neighbour Embedding algorithm (t-SNE) is proposed. The t-SNE algorithm is a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data for visualization in a low-dimensional space of two dimensions [32]. Each high-dimensional object is modeled by the algorithm using two dimensions such that similar objects are modelled by nearby points and dissimilar objects by distant points with high probability. Using the above described technique, a two-dimensional map can be attained [32].

3.6 The Norwegian Krone (NOK) as a Petrocurrency

The Norwegian Krone (NOK) is widely recognized as being a petrocurrency. "Petrocurrency" is a notion commonly used for currencies significantly impacted by rising and falling oil prices. In short, a petrocurrency is the currency of an oil producing nation that has significant amounts of oil exports as a percentage of its export portfolio [6]. Given such a large share of exports, the currency will rise and fall in correlation with the price of oil. Examples on other currencies considered petrocurrencies are the Canadian Dollar and the Brazilian Real [6].

Figure 3.6 visualizes the development of the NOK/EUR currency rate and the Brent oil price over the period of 2013 to 2022.

Figure 3.6: Price development of the NOK/EUR currency rate and the Brent Europe spot price over the period of 2013 to 2022.

As Alfred Berg suggests in the article "The Norwegian krone's correlation to the Brent oil price", the correlation between the Norwegian Krone and the Brent oil price is volatile, but gets amplified in periods of high volatility [24], as showcased in figure 3.6. The increased correlation during times of high volatility may be a cause of the inability of investors to successfully focus on all things at all times. Investors may therefore tend to use the Brent oil price as an indicator of where the Norwegian Krone is going in times of high volatility [24]. The majority of the companies in the Norwegian equity market are sensitive to the levels of the Norwegian Krone, but not so much the price of oil. When the volatility is high, and the correlation between the Norwegian Krone and the oil price increases, equities usually not effected by the price of oil may now be increasingly affected by it. This relationship between the Norwegian Krone and the price of oil may have implications on the thesis results, and will be discussed in the discussion section.

Chapter 4

Methods

This section describes the approach followed to explore the stated thesis objectives. Central to the approach are the technical implementation and setup of the thesis, and the making of the new trading rules (to be described in detail later). The programming part of the thesis will thus be subject for thorough examination, and pseudo code will be provided to give the reader a better understanding of some of the code composition. The following subsections will be presented in chronological order to facilitate a sense of how the work with the thesis proceeded, starting with data collection.

4.1 Data

4.1.1 Collected Data

All data collected and used in the thesis will be presented and described in this section. The collected data can be grouped in four categories:

1. Equity time series for each foreign country including the Norwegian market.

2. Time series for chosen equity indices of some of the foreign countries.

3. Time series for the chosen "factors".

4. Native currency rates for all foreign equities .

The first category includes all the equity time series collected for all countries. These equities form the basis for the formation of pairs (the initial pool of equities). For each of the selected countries, time series were obtained for the top 100 equities measured by market size. For Norway, the United States, the United Kingdom, France, Spain, and Germany, the top 200 equities were obtained. A large number of equities were collected to have a sufficient pool for pair comparison. Thus hopefully facilitating diversity in the detected pairs. The second category contains the time series for chosen equity indices for the different countries. The application of the index time series will be described later in the Methods section.

The third category is the obtained financial factors. The term "factors" is used frequently in this thesis and refers to a set of financial factors and ratios which are believed to (to a varying extent) have an impact on the Norwegian market. These factors are at the very core of the thesis, and their impact on the Norwegian market is by and large what is set out to explore. The time series for the equities, equity indices and financial factors are gathered in total return, meaning all cash distributions (dividends) are assumed reinvested. The data used in the thesis is collected from Thomas Reuters Data Stream.

Currency rates for all foreign exchanges were also collected, constituting the fourth category of data. The currency rates were downloaded from Norges Banks website, over the same time interval as the data accessed from Thomas Reuters DataStream.

The time series for all the data was collected over a ten-year interval, starting in 2012 and ending in 2022. The time series consists of daily observations, and all time series are collected in native currency. Performing pairs trading in the Brazilian market, M. Perlin reported that strategies using daily observations outperformed strategies using weekly observations [22]. Perlin argued that the superiority of higher frequencies in the pairs-trading framework is logically consistent, as the objective of pairs-trading is to take advantage of market corrections, and such inefficiency would, as expected, occur more often at high frequencies [22]. This aligns with the choice of collecting the data in daily observations, as opposed to weekly observations, in this thesis.

The below table presents an overview of the included countries with the corresponding number of equities collected. A handful of equity indices were also collected for some of the foreign countries, over the same time period and of the same data granularity as the equity time series. The included indices are listed in the rightmost column of table 4.1.

Table 4.1: Number of equities collected for each foreign country and the corresponding indices obtained. Indices were not gathered for all countries.

Foreign country	**Number of equities collected**	**Index included**
Norway	200	OSEBX
United States	200	NYSE, S&P, NASDAQ
United Kingdom	200	FTSE 100
France	200	CAC 40
Italy	100	MSCI ITALY
Germany	200	
Spain	200	
Canada	100	MSCI CANADA
Switzerland	100	
Australia	100	MSCI AUSTRALIA
Austria	100	
China	100	CSI 300
Japan	100	TOPIX
Hong Kong	100	HANG SENG
Sweden	100	OMX STOCKHOLM
Denmark	100	OMX COPENHAGEN
Finland	100	

Data was collected for the following financial factors, over the same time period and of the same data granularity as the equity time series:

- Crude Oil-WTI Spot
- Brent Europe Spot
- Gold Bullion LBM
- LME-Copper Grade A
- LME-Aluminium
- LME-Copper
- Baltic Exchange Dry Index (BDI)
- RFV Natural Gas
- NYMEX Natural Gas Henry Hub
- Fish Pool Index Spot Salmon NOK/KG
- Silver, Handy&Harman (NY)
- 10 Year bond yield NO
- 10 Year bond yield United States
- 10 Year bond yield United Kingdom

- NOK - USD
- NOK - EUR
- NOK - £

As a total of 17 factors were included, and some of them are suspect to be closely related to each other, visualizing and investigating the relationships between the factors is important. Including several factors of high correlation may result in ambiguous results. Figure 4.1 displays the correlation between the factors as a heat map, calculated over the formation period. The colored bar on the right side of the figure maps color to the magnitude of the correlation.

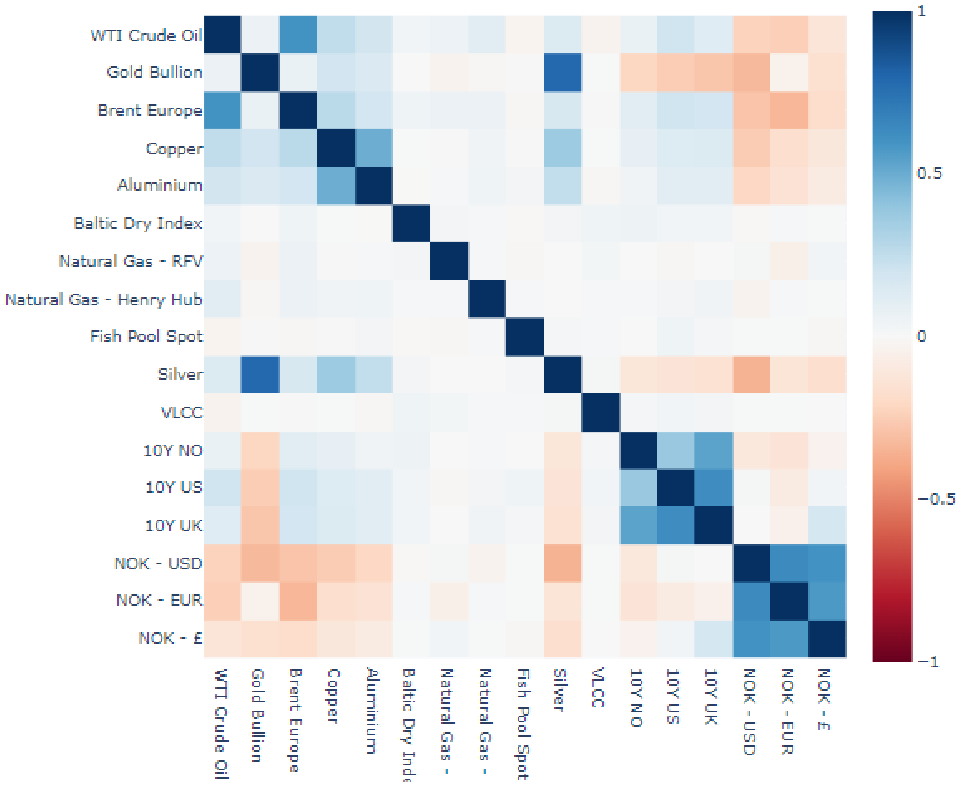

Figure 4.1: Factor correlations between all the proposed factors over the formation period, visualized as a heat map. Blue and red represent positive and negative correlation respectively, and stronger color indicates stronger absolute correlation.

The WTI and Brent oil prices displayed high correlation over the formation period. A decision was made to only include one oil price, Brent Europe, in the final pool of factors for clarity. Figure 4.1 also displays the relative high correlation between the included currency rates, and between the included 10 year bond returns. The highest correlation over the formation period was detected for silver and gold, of 0.78. Although silver and gold showed high correlation over the formation period, both price series were included in the final pool of factors (as they are not perfectly correlated).

4.1.2 Data Cleaning and Preparation

After all data had been collected, the next step was to "clean" the data and write it to the desired format. As the amount of data collected was large, and the processes needed to clean the data are varying, this step of the data handling is subject to a variety of mistakes if not done properly. The process of cleaning the data was thus performed with scrutiny, to make sure all transformations were correctly applied and the resulting data was right. The data preparation process consists of the following steps:

1. Removing nul-values.
2. Writing the time series to Norwegian Kroner.
3. Writing all time series to the same indexing.

The first step of the data cleaning process was removing data with sufficient amount of missing values. This included removing all equities listed later than the starting date of the data interval (2012). Time series containing intermediate nul-values were either forward or backward filled, meaning that the missing value was substituted by the succeeding or preceding value. The filling of nul-values was performed to ensure data continuity.

Next, all time series were written to Norwegian Kroner (NOK). This is a crucial step to facilitate perfect comparison between equity time series. Finally, all time series were written on the same index, as inconsistency in the gathered data resulted in the index not being equal for all time series.

After all the data was written to the desired format, the next step was pair detection in the formation period.

4.2 Formation Period - Detecting Pairs

Figure 4.2 illustrates the process of pair detection for each country compared to the pool of Norwegian equities, from the initial pool of equity time series to the final pool of pairs for each country.

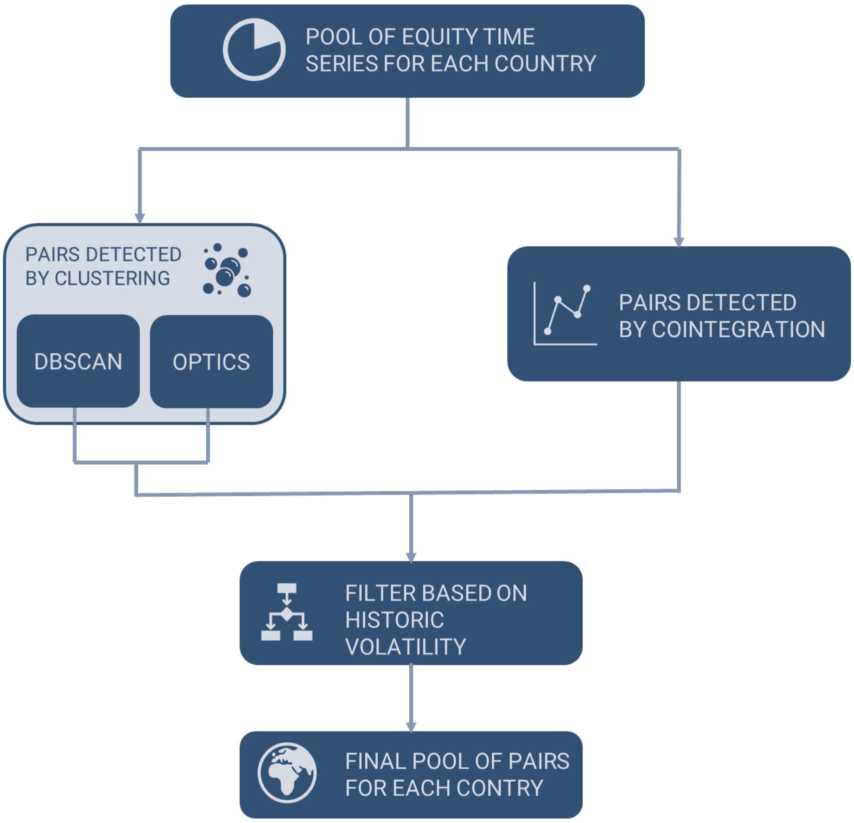

Figure 4.2: Diagram visualizing the pair detection process for each foreign country, compared to the pool of Norwegian equities, starting with the initial pool of equity time series for each country, and resulting in the final pool of pairs for each country.

For each of the foreign countries, the initial pool of equities consists of the cleaned equity time series for the specific country, including the cleaned equity time series for the Norwegian equity market (the OSEBX index). A "formation/trade-phase" split of 70% was applied for all pairs, meaning the first 70% of the data was used for pair detection, and the last 30% were set aside (as unseen test data) for the strategy execution. The detected pairs were thus detected on data over the time period of 2012 to 2019. The time series (collected in daily observations) were re-sampled to weekly observations before being applied to the clustering algorithms. The reason for re-sampling the time series was to make the pair detection by clustering more robust, and less sensitive to daily changes.

After pairs had been detected by either clustering or cointegration, the pairs showing large differences in historical volatility were excluded from the current pool of pairs. The calculations of pair cointegration were performed using the "coint" function of the statsmodels-tsa python module [33]. The statsmodels coint function is based on the the augmented Engle-Granger two-step cointegration test, descried in the theory section.

The filtering based on historical volatility was performed to make sure the equities constituting each pair showcased approximately equal volatility over the formation period. Equities with large differences in volatility were not considered good pairs. Filtering based on volatility is an alternative to beta-adjusting the pairs (investing relatively more in the stock of lowest beta in the pair). As the beta measure is a measure of a stock's volatility in relation to the overall market, the beta measure will be affected by market movements. In contrast, calculating the historical volatility between two equities constituting a pair, will only consider the historical relationship between the two equities. As the beta measure of a stock is prone to change, the historical volatility of two stocks may be seen as a more statistical robust measure to beta, and is thus preferred in this thesis. The final pool of pairs for each country was further filtered to only include pairs composed of one equity from the foreign country of examination and one Norwegian equity, as these are the only pairs of interest. The above-described process was repeated for each foreign country to end up with the final pool of pairs for all countries.

In addition to pairs detected by clustering and cointegration, a set of hand-picked fundamental pairs were included for each of the 20 largest Norwegian equities (excluding seafood companies as they do not have any "good" fundamental foreign pairs), measured by market capitalization. The fundamental pairs were selected based on similarities in company business and structure. Clustering algorithms and cointegration calculations are purely mathematical procedures, and will exclusively detect pairs based on statistical relationships. These approaches may thus overlook good fundamentally connected pairs, not being detected by the mathematical procedures. A saying to keep in mind when working with mathematical procedures is that the algorithms do not see the same things humans do. What may seem like "odd" results to humans make perfect sense to the mathematical procedures, only presented with the numbers. The fundamental pairs are included to nuance the final pool of pairs, and eventually examine whether the pairs detected by clustering and cointegration outperform the fundamental pairs when put through the same proposed pairs trading frameworks under equal conditions.

The hand-picked, fundamental pairs can be found in Appendix A.

4.3 Trade Period - Trading Pairs

After the final pool of pairs had been formed, the pairs were applied to the proposed trading frameworks. The approach followed was to first implement a purely statistical (basic) pairs trading framework, on which the pairs were evaluated. Then, after the pairs had been evaluated on the purely statistical framework, a second, reinforced version of the purely statistical framework was implemented. The new framework seeks to detect the confluence of favourable conditions, resulting in profitable trading signals, in order to place a larger bet with the goal of making a correspondingly larger profit. The new framework is named the "reinforced framework" in this thesis, but could alternatively be named the "nuanced framework", as the framework aims to target and exploit nuances in the development of the Norwegian equity market. This specific reinforced framework represents a new approach to pairs trading, not shown in any previous literature, and will be described in detail below.

The implementation of two frameworks, one basic and one reinforced, is performed to be able to assess the performance of the reinforced framework. In evaluating a pair on both frameworks under the exact same conditions, we may be able to assess whether the application of the new functionality of the reinforced framework yields better performance than a basic pairs trading framework.

4.3.1 Purely Statistical Framework

The first step in the process of implementing the pairs trading framework was building a basic, purely statistical (purely statistical, meaning that all trading signals are based on statistical measures) threshold based model. Following a threshold based trading model, the z-score of the ratio of the pair is calculated. The z-score is a numerical measure that describes a value's relationship to the mean of a group of values, thus assessing the relative magnitude of the pair development compared to historical development [3]. When the absolute value of the z-score crosses through predefined boundaries, trading signals will be triggered. The z-score boundaries correspond to levels where the development of the pair is significantly different from the historical mean. The setting of these boundaries are important for strategy performance. If a narrow boundary is established, many positions are initialized, but profit would be low, while a wide boundary will be highly rewarded by the execution of the strategy [23]. Figure 4.3 visualizes the threshold-based trading model applied in a purely statistical pairs trading framework. The blue line represents the ratio z-score, and the horizontal lines represent boundaries for triggering of trading signals. The magnitude of the ratio z-score is measured on the vertical axis.

Figure 4.3: Threshold-based trading model with red and green horizontal lines corresponding to thresholds for selling and buying the pair, respectively. The blue solid line is the pair ratio z-score (measured on the vertical axis), and the blue shaded area corresponds to z-score levels where trades are reversed.

When the z-score crosses through the red stapled line from below, a short position of the pair will be initialized. The trade execution follows the cointegration approach, described in the theory section. If the z-score crosses through the continuous red line (upper boundary 2) a second, relatively larger short position of the pair will be initialized. The exact opposite actions will be taken if the z-score crosses through the stapled or continuous green lines from above, i.e., long position of the pair will be initialized. The reasoning behind including a double upper and lower trading boundary is to better time the divergence and subsequent convergence of the pair ratio. A single trading boundary may be to general.

Table 4.2 presents the exact thresholds applied to both pairs trading frameworks, corresponding to the thresholds visualized in figure 4.3.

Table 4.2: Z-score thresholds for trade execution. Long and short positions corresponds to the lower and upper boundaries visualized in figure 4.3, correspondingly.

Type of trade	**z-score thresholds**
long position 1	$z \leq -1.2$

Continued on next page

Table 4.2 – continued from previous page

Type of trade	z-score thresholds
long position 2	$z \leq -1.2 * 1.5 = -1,8$
short position 1	$z \geq 1.2$
short position 2	$z \geq 1.2 * 1.5 = 1,8$

The absolute value of the initial threshold was set to a z-score of 1.2, which then was multiplied by a constant (1.5 in this case) to determine the upper boundaries for buying and selling. These numbers were reached after testing different values on data of detected pairs in the formation period, aiming at finding values applicable to most pairs.

4.3.2 Reinforced Framework

After the purely statistical framework had been successfully implemented, the next step was engineering the reinforced framework. The engineering of the reinforced framework was of an iterative nature; adding functionality bit by bit, watching how the model responded, and improving based on the observations. The implementation of the reinforced framework was also highly exploratory. The very idea of what this new framework was hoping to achieve was clear, but not so how it was going to be implemented. After exploring different ideas and possibilities, experimenting with what was possible and what was not, the final trading framework was reached.

The additional functionality of the reinforced trading framework, compared to the purely statistical, is in the form of two algorithmic trading rules, checking for the confluence of favourable conditions and targeting the nuances of the Norwegian market. These "favourable conditions" are closely related to the thesis objectives. The motivation for adding the additional trading rules is to try to understand what is driving the pair development, and in understanding what is driving the pair development being able to filter out the "good" trading signals. The reinforced version of the pairs trading framework is thus an attempt at manifesting the stated thesis objectives in tangible code.

In his book "Noise: A Flaw in Human Judgment", Daniel Kahneman discusses the superiority of simple models over humans: "The combination of personal patterns and occasion noise weights so heavily on the quality of human judgement that simplicity and noiselessness are sizeable advantages. Simple rules that are merely sensible typically do better than human judgement" [20]. This notion that simple trading rules may prove superior to human judgment is part of why the stated thesis objectives are implemented as simple mechanical trading rules in the reinforced framework, free of the inherit noise and bias of human judgment. In the above citation, occasion noise is when the state you are in (personally), for example your mood or outcome of your last decision, has an effect on your judgement [20].

The two trading rules of the reinforced framework are described in detail below:

Trading Rule 1

The first trading rule aims to exploit the idea that a handful of fundamental factors have a disproportional effect on the Norwegian equity market. When those factors change, they seem to affect not only equities fundamentally correlated to the factors, but also equities not fundamentally correlated to the factors. A potential response of a company to the change in a historically low correlated factor is not grounded in company specifics or other fundamentals of the company, and will thus revert to the mean soon after the initial response. As the Norwegian equity constituting the pair seems to overreact to the changes in a handful of factors, this may result in a relative mispricing between the Norwegian and foreign equity constituting the pair.

For each pair, the reinforced model first calculates a list of factors that have historically shown a weak correlation to the Norwegian equity constituting the pair. The historical correlation is based on the average value of 50-day rolling correlations, over a window of the two years leading up to the start of the trade phase (2017 to 2019). The below code block displays the assembly of the *factor_sensitivity* dictionary, grouping factors based on the magnitude of historical correlations with the Norwegian equity of the pair. Only factors mapping to ”LOW” will be included in the list of historically low correlated factors. The *factor_lookback_frame* variable is a data frame containing historical rolling correlations between the Norwegian equity of the pair and all the factors, facilitating simple comparison.

```
factor_sensitivity =  {}
factor_low_corr_list = []

for factor in factor_lookback_frame.columns:

    if factor_lookback_frame.mean()[factor] > threshold_high:
        factor_sensitivity[factor] = 'HIGH'

    elif factor_lookback_frame.mean()[f] < threshold_low:
        factor_sensitivity[factor] = 'LOW'
```

```
            factor_low_corr_list.append(f)

    else:
        factor_sensitivity[factor] = 'MEDIUM'
```

Listing 4.1: Detection of historically low correlated factors. Factor correlations to the Norwegian equity constituting a pair are iterated over in a for-loop and assigned a label corresponding to the magnitude of the historical detected correlation.

For each time point (daily) of the trading phase, the main function checks whether one of the "low correlated factors" has shown an unusually high correlation to the Norwegian component of the pair over a short period of time in the trading phase. These "new" correlations are calculated based on a rolling window of 20 days. Based on the magnitude of the potential detected correlation, the weight of the current day is set to a higher number, with higher correlations corresponding to higher weights. Days of the trading phase where some sort of correlation to historically low-correlated factors is detected will be referred to as "up-weighted days based on trading rule 1" in the rest of the thesis.

When the z-score of the pair ratio breaks through the predefined absolute boundaries, the function will check the current weight of the day and multiply it with a constant, if a correlation with historically low correlated factors has been detected. If the resulting weight for the current time step has a high value it means that the algorithm has detected unusually high correlation with historically low correlated factors for the Norwegian equity of the pair. Periods of high correlation with historically low correlated factors for the Norwegian equity in the pair will thus yield stronger (higher weighted) trading signals if the pair ratio z-scores breaches through the boundaries indicating a trading signal.

Trading Rule 2

The second trading rule, slightly more complicated than the first, aims at exploiting the idea that in periods when the Norwegian market falls sharply it will "drag" all the equities constituting the market down with it. The periods of decline referred to are periods when the Norwegian market is driven down by one or more historically low-correlated factor to the Norwegian equity of the pair, and the Norwegian equity of the pair has seen a relative large decline. An example would be in the case where the Norwegian equity market is driven down by the price of oil, and equities not historically correlated with the price of oil are experiencing negative price development over the same period.

The idea that foreign investors assign a disproportional large effect to the Norwegian stock market, based on the recent development in a handful of financial factors, is based on the notion that the Norwegian market is relatively small in global context, and will thus not be suspect to equally large scrutiny as other foreign markets. A handful of financial factors will thus "dominate" the development of the Norwegian market. This results in the development of the Norwegian market being disproportionately affected by a handful of fundamental factors, compared to other foreign exchanges. Based on the above argument, an additional trading rule assessing whether a negative reaction of the Norwegian equity market to the change in a factor, is larger than for other foreign stock exchanges, is added to the previously described trading rules of trading rule 2. This is the application of the collected equity indices, presented in table 4.1. The recent development of a handful foreign equity indices is compared to the recent development of the Norwegian market. A signal is only triggered if the Norwegian market has seen a larger decline than the average of the foreign exchanges.

The confluence of the Norwegian market dragging the Norwegian equity of a pair down, driven by one or more historically low-correlated factors to the Norwegian equity of the pair, *and* the negative reaction of the Norwegian market being larger than for other foreign exchanges, is especially interesting. Days of the trading phase were such conditions are detected will be referred to as "up-weighted days based on trading rule 2" in the following.

The second trading rule is constituted of three separate checks, which all need to be passed in order for the weight of the day to be updated. The three checks constituting trading rule 2 are listed below:

1. The Norwegian market shows an unusually high correlation to a historically low correlated factor to the Norwegian equity of the pair.
2. The Norwegian equity of the pair has over the same period as the Norwegian market fallen relatively much, compared to historical levels.
3. The Norwegian markets decline is of higher magnitude than the average of other foreign stock exchanges, over the same period in time.

The magnitude of the decline of the Norwegian market, the foreign market, and the Norwegian component of the pair is measured for each time step as the rolling 20-day sum of price changes. The third check of trading rule 2 examines whether the rolling 20-day sum of price change for the Norwegian market is below a predefined quantile of the rolling 20-day sum of price change for all the foreign stock exchanges. The foreign stock exchanges included in the third check are:

- S&P500
- NYSE
- FTSE 100
- MSCI Canada
- MSCI Australia
- CAC 40
- Topix
- Hang Seng

In the final implementation of the reinforced framework, trading rule 1 and 2 are both applied for each day of the trade phase. The highest possible weight for a day occurs during the confluence of both trading rules being passed as True, resulting in a double update of the weight of the corresponding day. Both trading rule 1 and trading rule 2 allow for including a user-defined number of previous days for weight update of the current day. This means that if the current day do not trigger any signals, but the previous day(s) has, the weights are still updated. This functionality is included as the signals are based on rolling statistics, and only updating weights based on the exact day the signal was detected may be too narrow. In the final implementation of the reinforced framework, the number of previous days included was set to 1, to mitigate the noise in the resulting weight updates.

Although both trading rules are being passed, it is not given that what has been checked for is actually what is driving the market dynamic. This may result in spurious trading signals. Performing correlation-calculations on 20-day windows may also generally yield spurious correlations; two instruments moving similarly by pure chance, as the standard error is high when the number of observations are low [3].

4.3.3 Weight Updates

The above-described trading rules are executed for each day in the trade phase. Every single day in the trade phase is assigned a weight, corresponding to the previously described "up-weighted days" based on either trading rule 1 or 2. The weight of the current day is a multiplier defining how large a potential trade on that day will be in size, compared to a "base-amount", should a trade be triggered. The weight of the day is initialized to one, before the trading rules have been performed. If the first trading rule is triggered, the weight of the day will be multiplied by a floating number. The size of which the weight of the day is multiplied by depends on the strength of the detected correlation. Stronger detected correlations results in higher weight updates.

After trading rule 1 have been performed, the algorithm moves on to the second trading rule. If the second trading rule is triggered, the weight of the day will be multiplied by a floating number as well. Figure 4.4 illustrates the performed weight updates during a run the reinforced trading framework. WEIGHT (0) represent the initial weight for each traded day, and WEIGHT(1) represent the weight after both trading rules have been performed.

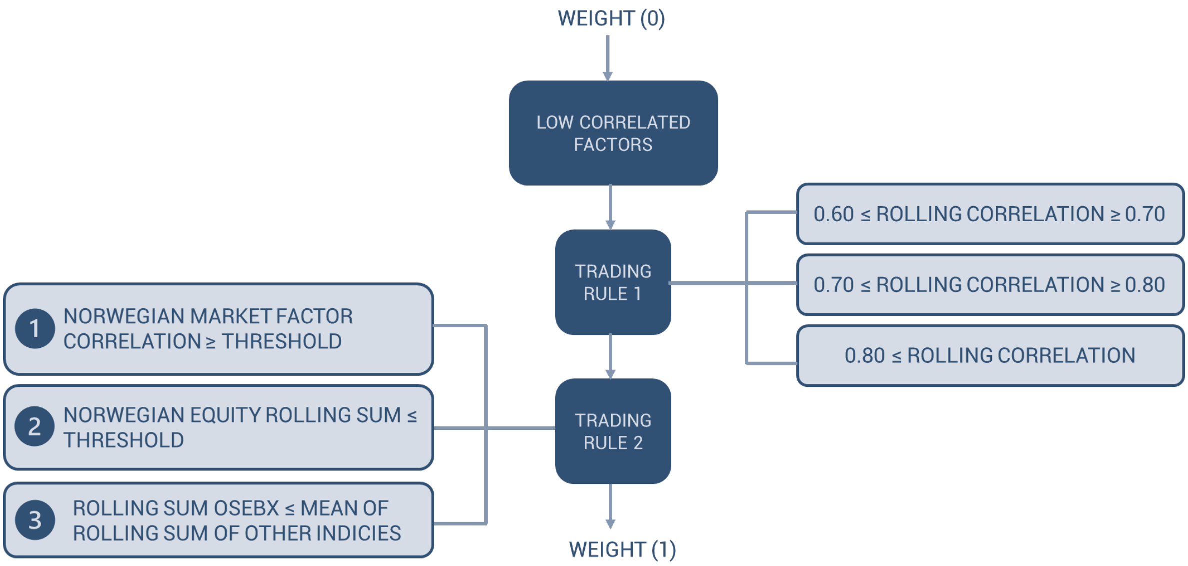

Figure 4.4: Weight updates for the reinforced trading framework, starting with the initial weight of the day and resulting in the weight after both trading rules have been performed. The light blue boxes, adjacent to trading rule 1 and trading rule 2, visualize the checks performed within each of the two trading rules.

Whereas each of the checks constituting trading rule 1 results in a different weight update based on the size of the detected correlation, the checks constituting trading rule 2 must all coincide for the trading rule to be passed. The two proposed trading rules are independent of each other, and so the outcome of the first trading rule will not affect the outcome of the second trading rule. There are four possible scenarios for the resulting weight of a day after the two trading rules have been performed.

1. None of the trading rules are triggered, resulting in an unchanged day-weight
2. Trading rule 1 is triggered but trading rule 2 is not triggered, resulting in a single update of the day-weight
3. Trading rule 1 is not triggered but trading rule 2 is triggered, resulting in a single update of the day-weight
4. Both trading rules are triggered, resulting in a double update of the day-weight.

The strongest possible weight updates will occur when both trading rules are triggered on the same day, resulting in the day-weight having two subsequent updates.

4.4 Implementation in Code

4.4.1 Trading Frameworks

The functionality of the two above described pairs trading frameworks takes place in one single function. The function contains all the functionality to perform a fully valid trading phase. The function has two distinct "modes". The first mode executes a plain, pure statistical pairs trading framework. The second mode executes a trading scheme following the reinforced framework. Feeding the exact same pair data to two different modes of the function, and comparing the corresponding results, yields perfect comparability between the purely statistical and reinforced framework.

The process of putting the detected pairs through the proposed frameworks were done subsequently for each foreign country, with corresponding detected pairs. The function is designed as a for-loop, iterating over all the points (days) of the trade phase from start to end.

4.4.2 Pair Class

To collect all the important functionality for exploring pairs in one place and making it easily accessible, a "Pair" class was implemented. The class follows pythons object-oriented framework, utilizing the *class* functionality. Ever time a new pair was to be explored, an instance of the "Pair" class was instantiated. The Pair class enabled quick and seamless calling of methods, such as calculating and plotting the spread and ratio of the pair, concatenating the pair and desired factors to a common data frame, adjusting the pair for currency, and calculating cointegration statistics for the pair. Once the class was written, all the desired information about a pair was easily accessible, making the pair exploration process efficient. The below code block provides an overview of the Pairs class without showing the exact code, including some of the parameters and essential functionality.

```
class Pair:

    def __init__(self,sec_1, sec_2, start, end, granularity='1d'):

        self.sec_1 : str = sec_1
```

```
        self.sec_2 : str = sec_2
        self.start : str = start
        self.end   : str = end
        self.granularity : str = granularity

    def plot_pair_norm(self, ...) -> go.Figure:
        """Plots the normalized price series of the pair"""

    def ratio(self, ...) -> pd.Series:
        """Calculates the ratio of the pair. Plot if specified"""

    def spread(self, ...) -> pd.Series:
        """Calculates the spread of the pair. Plot if specified"""

    def coint(self, ...) -> tuple:
        """Calculates cointegration statistics of the pair"""

    def to_NOK(self, ...) -> pd.Series:
        """Writes the foreign time series to Norwegian Kroner """
```

Listing 4.2: Overview of Pair class with attributes and methods. Methods are only described by the functionality they perform.

4.4.3 Interactive Dashboards

The first part of the thesis included extensive exploration of possible pair combinations and clustering formations. The large number of equities, and possible parameter combinations for the clustering algorithms, resulted in a high number of possible variations when performing the clustering. To make the tedious clustering exploration process faster, more effective, and more visual, interactive dashboards were frequently used as a tool of exploration. As the collected equities from the different countries are inherently different, there is no "one fits all" combination regarding the clustering parameters. In making the clustering exploration process interactive it was easier to observe the results of changes in parameter values, and adapt the clustering algorithm to each of the different countries. All data was cleaned and prepared prior to entering the dashboards.

The dashboards are written in Dash, an open-source dashboard library for visualization. Figure 4.5 displays an overview of the front end of the final dashboard used for exploring the DBSCAN and OPTICS clustering algorithms. As displayed in the figure, the left sidebar allows for easily changing and combining different parameters. In the specific dashboard visualized below, France is chosen as the country of examination, and the DBSCAN algorithm is used for cluster detection.

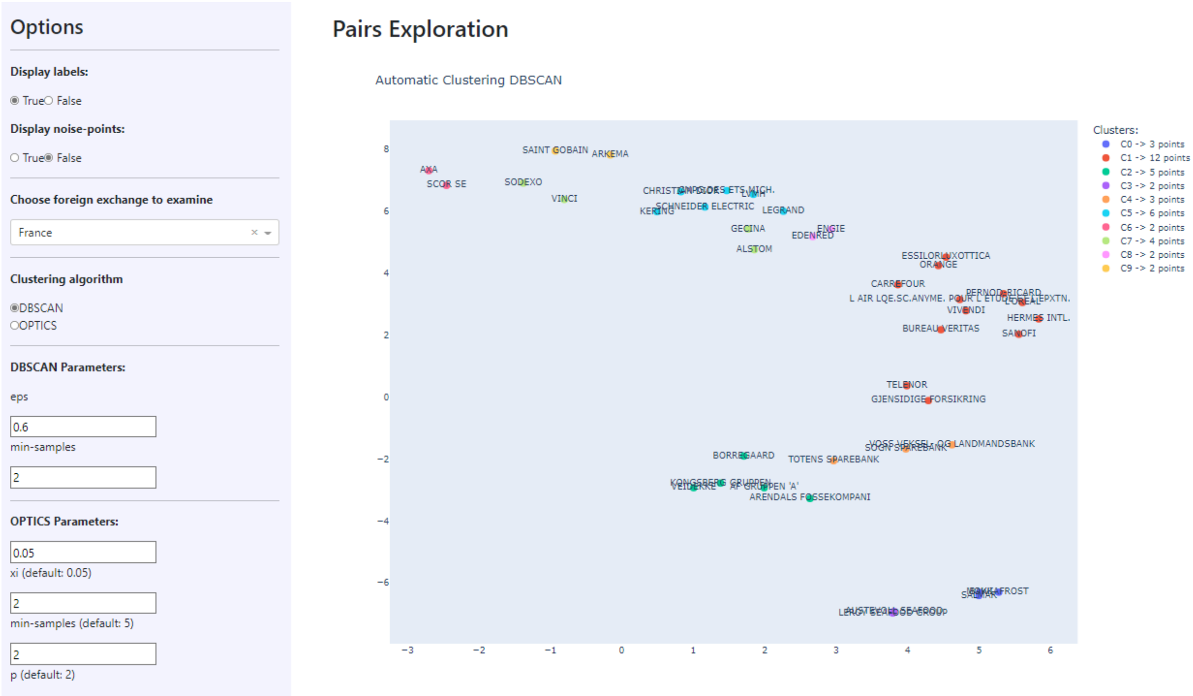

Figure 4.5: Overview of dashboard used for cluster exploration. The left side-bar allows for changing clustering parameters interactively, while the main plot represents the clustering results.

The above figure visualizes the main dashboard used for exploration of the clustering of pairs. The dashboard allows for choosing between the two proposed clustering algorithms and interactively changing the corresponding parameters. Additionally, the dashboard allows for changing the desired country with corresponding equities being compared to the equities of the Norwegian market. This way, the process of testing combinations and watching the results of different parameter combinations was highly visual and instructive.

Chapter 5

Results

In this section the results of the thesis are presented. The results are focused around the difference in performance between the two approaches and the difference in return between the normal trades (trades with no additional weight) and weighted trades (trades occurring on weighted days). The distribution between the different types of trades is also reported. Since the main goal of the thesis is to explore nuances in the Norwegian equity market rather than solely focusing on strategy returns, a range of different results are provided, beyond returns. Different representations of the results are presented to give a more nuanced view of the results and facilitate thorough discussion. Since the specific approach taken to pairs trading in thesis is not seen in any previous literature, comparable results of good quality are low in number. The results and discussion sections will thus mainly focus on the results reported in this thesis, and highlight using relevant literature where it is found appropriate.

When comparing the performance of the proposed frameworks, two terms are frequently used; strategy return and trade return. Strategy return refers to the aggregated returns of all trades initialized (both normal and weighted), over a full trade period cycle, for either the reinforced or the basic strategy. Where the strategy return includes both normal and weighted trades, the term "trade return" distinguishes between normal and weighted trades. Trade return is the mean of all trades of one type, either normal or weighted, for one complete run of the reinforced framework. The reason for reporting the trade returns is to facilitate a "cleaner" comparison between normal and weighted trades, thus being able to assess whether the weighted trades outperformed the normal trades.

5.1 Transaction Costs

Transaction costs needs to be taken into account for the presented results to be realistic. At least three types of costs emerge when executing a pairs trading scheme: commissions, fees for short selling, and implicit market cost [12]. Based on various literature, the basic one-way (either buying or selling) cost for a single trade is set to 25 basis points (0.25%). One complete trade in the pairs trading framework is composed of four smaller trades. One first buys one stock and sells another, before one has to reverse both these trades upon trade reversion. Thus, the basic one-way cost needs to be multiplied by 4 for each complete trade. The resulting total transaction cost for each complete trade is thus set to 100 basis points (1.00%), which will be subtracted from the return of each performed trade.

5.2 Pairs Formed

The below table presents an overview over the number of foreign pairs formed with each Norwegian equity. The table only includes the top 20 Norwegian equities measured by the number of pairs formed. The pairs included in the table count are constituted of pairs detected by cointegration and either of the proposed clustering algorithms, DBSCAN and OPTICS.

Table 5.1: Norwegian equities with the corresponding number of foreign pairs formed, sorted in descending order by the number of pairs formed. The table displays the top 20 Norwegian equities regarding number of pairs formed.

Norwegian equity	**Number of pairs formed**
Selvaag Bolig	46
Mowi	29
Atea	22
Olav Thon Eiedom	18
Tomra Systems	18
Bakkafrost	18
AF Gruppen	17
Yara International	14
Kitron	14
Gjensidige Forsikring	13
Orkla	12
Dnb Bank	11
Contextvision	11

Continued on next page

Table 5.1 – continued from previous page

Norwegian equity	**Number of pairs formed**
Bouvet	11
Borragaard	7
Photocure	7
Leroy Seafood Group	6
Salmar	6
Veidekke	5
Norway Royal Salom	5

Table 5.1 shows that Selvaag Bolig formed the highest number of pairs. Olav Thon Eiendom, another real estate company formed a high number of pairs as well, with 18 total pairs. An interesting observation is that equities in the seafood sector are frequently represented in table 5.1, meaning equities in the seafood sector have formed a relative high number of pairs compared to other equities from other sectors.

The below table presents a handful of pairs detected by either cointegration or clustering, that would not be expected to form "good" pairs based on a fundamental view. Although the pairs presented in table 5.2 may not be expected to form profitable pairs, the resulting strategy returns are high. Due to the large number of total pair comparisons, the chance of detecting spurious relationships is high.

Table 5.2: A variety of "unexpected pairs", detected by either cointegration or clustering with the corresponding reinforced strategy return.

Pair	**Reinforced strategy return**
NIKE "B" - Selvaag Bolig	51%
Moody's - Selvaag Bolig	24.96%
Stora Enso - Storebrand	45%
CAE - Mowi	52%
Pfizer - Mowi	51%
Pfizer - Dnb Bank	29.52%
Altia Consultors - Bakkafrost	37%
Bank of America - Atea	10%
Fiskars "A" - Yara International	10%
Lundin Energy - Salmar	7%

Continued on next page

Table 5.2 – continued from previous page

Pair	Reinforced strategy return
Cisco Systems - Mowi	19.41%
Ansys - Selvaag Bolig	23.56%
Meta Platforms A - Bakkafrost	31.62%

5.3 Performance Comparison

All results in this section are presented including transaction costs. The below table presents the relative distribution of positive, neutral and negative strategy returns for both strategies, for all pairs. The returns labeled "Neutral" are pairs going through the proposed pairs trading framework without a single executed trade.

Table 5.3: Proportion of positive, neutral and negative strategy returns for the reinforced and basic trade functions, averaged over all traded pairs.

Strategy	Return	Relative distribution
Reinforced	Positive	50.9%
	Neutral	21.8%
	Negative	27.2%
Basic	Positive	50.7%
	Neutral	21.8%
	Negative	27.5%

Both versions of the proposed framework resulted in a positive return for over half of the proposed pairs, meaning that the majority of the traded pairs resulted in a profit. Table 5.3 also shows that the reinforced strategy resulted in more profitable runs, and less negative runs, than the basic strategy. For 27.2 and 27.5 percent of the pairs, trading the pair over the trade phase resulted in a loss, for the reinforced and basic strategy respectively. Figure 5.1 displays the return distribution over all the traded pairs, for the reinforced strategy. The single highest positive return is excluded in the below figure to increase the readability.

Figure 5.1: Return distribution of the reinforced strategy, for all detected pairs. Single strategy returns are visualized as vertical blue bars, with magnitude corresponding to the values of the y-axis, where 1.0 corresponds to a strategy return of 100 percent.

The relative distribution between positive, neutral and negative strategy runs presented in table 5.3 are reflected in the above figure. The number of positive returns are visually larger than the number of negative returns. There are more extreme observations with positive returns compared to extreme observations with negative returns, and the extreme positive returns seem to be larger in magnitude.

Figure 5.2 gives a more nuanced view of figure 5.1. The different strategy returns are plotted alongside each other as histograms, with additional boxplots for each of the distributions (the single most extreme positive strategy return is omitted as it reduces the visibility of the distributions). The "box" of the boxplot shows the quartiles of the return distribution while the whiskers (horizontal lines extending from the box) extend to show the rest of the distribution, except from points determined to be outliers, visualized as dots along the horizontal plane [34]. The blue distributions correspond to the reinforced strategy and the red distributions correspond to the basic strategy.

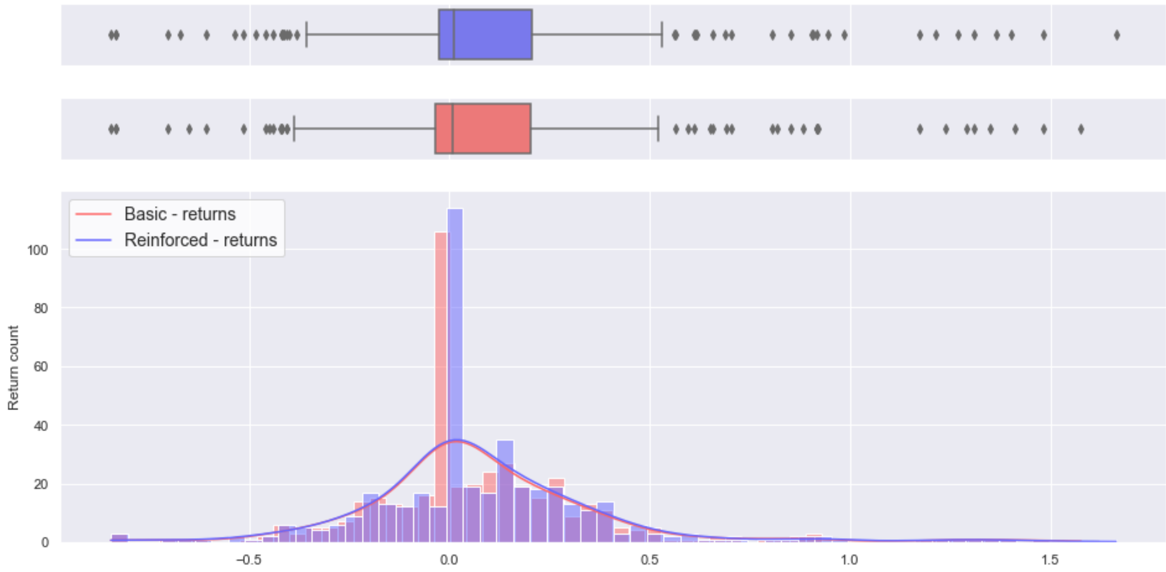

Figure 5.2: Strategy return distribution as histograms and box plots. The two top plots represent the strategy return distribution as boxplots for each of the two strategies, while the strategy return distribution for both strategies are plotted above each other in the bottom plot, as histograms with corresponding kernel density estimation distributions.

From the boxplots one can see that there are fairly many extreme events, both negative and positive. Investigating the histograms carefully, one see that the reinforced strategy has a slightly "fatter tail" to the right than the basic strategy, meaning it has slightly more positive returns. It is interesting to assess the distribution skew. The reinforced and purely statistical strategies have skew of 1.33 and 1.34, correspondingly. Both distributions are positive (right) skewed, meaning they have a long tail to the right, being more exposed to extreme positive events than negative.

Rad, Yew Low and Faff discussed the seemingly consistent fat left tails of the proposed strategies in their paper "The profitability of pairs trading strategies: distance, cointegration and copula methods" [27], indicating that extreme negative results occur more frequently than corresponding positive ones. This observation is in contrast to the results visualized in figure 5.2, where the right tails of the distributions are in fact fatter than the left tails, and the skew of both distributions are positive, meaning that extreme positive outcomes do occur more often than corresponding negative ones.

Table 5.4: Mean weighted return and median return for the reinforced and basic strategy, aggregated over all pairs.

Strategy	Mean return	Median return
Reinforced	9.96%	10.91 %
Basic	9.63%	10.18%
Difference	0.33%	0.73%

As displayed in table 5.4 both versions of the strategy, the basic and reinforced version, had a positive mean and median return aggregated over all the traded pairs. The reinforced version scored 33 basis points better than the basic version on average measured by weighted mean, and 73 basis points better measured by median.

Table 5.5: Mean weighted strategy return of the reinforced and basic strategy grouped by country of the foreign equity constituting the pairs. All results are reported in percentage.

Country	Reinforced	Basic	Difference
Denmark	29.01	27.10	1.92
Spain	28.37	29.01	-0.65
UK	26.04	21.81	4.23
Austria	20.70	19.88	0.83
Canada	17.18	16.88	0.31
China	14.74	14.08	0.66
Australia	13.41	12.87	0.54
Hong Kong	12.69	11.75	0.94
Italy	8.28	8.63	-0.35
Sweden	6.71	5.88	0.83
Switzerland	6.68	6.39	0.28
US	6.20	5.70	0.50
Finland	4.60	3.67	0.94
Japan	1.03	0.81	0.23
Germany	0.08	0.12	-0.03
France	-1.35	-0.72	-0.63

Table 5.5 shows that the proposed strategies yielded a positive mean weighted strategy return, grouped by country of the foreign pair, in most of the cases. The results are sorted in descending order by the reinforced strategy return. Pairs formed with foreign equities from France performed the worst, with a mean weighted strategy return of -1.35 and -0.72. Pairs formed with foreign equities from Denmark performed the best, with a mean weighted strategy return of 29.01 percent. Pairs from Spain, the UK, Austria and Canada also performed well, with mean weighted strategy returns above 17 percent. Inspecting the rightmost column one can see that the reinforced strategy outperformed the purely statistical strategy in most of the countries. The difference in performance between the reinforced and the purely statistical strategy was greatest for pairs formed in the UK and Denmark, with respectively 4.23 and 1.92 basis points on average for all pairs.

Table 5.6: Mean weighted strategy return of the reinforced and basic strategy, grouped by sector of the Norwegian equity constituting the pair, and the corresponding number of pairs formed for each sector. All results, except the number of pairs, are reported in percentage.

Sector	**Num. pairs**	**Reinforced**	**Basic**	**Difference**
Oil	18	51.50	47.23	4.27
Real estate	65	13.93	13.91	0.03
Industry	95	9.75	8.20	1.55
Tech	51	8.93	7.92	1.01
Renewable	24	8.50	8.19	0.31
Seafood	65	7.94	6.62	0.47
Bank	34	6.25	6.59	-0.34
Health care	10	-24.83	-25.14	0.31

Table 5.6 displays the mean weighted strategy return for both proposed strategies grouped by the sector of the Norwegian component of the pair. The sector of the Norwegian component of the pair is determined based on the industry filters provided at Oslo Børs' web pages [13]. The results are sorted in descending order by the reinforced strategy return. Inspecting the table it is clear that pairs formed with Norwegian companies in the sectors of oil and real estate performed the best. All sectors had positive average returns except for the health care sector. The proposed pairs with Norwegian equities in the health care sector yielded an average strategy return of negative 24.83 and negative 25.14 percent. This reflects the low number of pairs formed within this sector, as the aggregate strategy return of the pairs is more dependent on single observations and the

unpredictable nature of health care companies.

Table 5.7: Mean weighted strategy return of the reinforced and basic strategy for the fundamental pairs picked for the top 20 Norwegian equities by market capitalization.

Norwegian Equity	**Num. pairs**	**Reinforced**	**Basic**	**Difference**
Equinor	6	-12.21	-12.24	0.04
DNB	7	-24.29	-23.86	-0.43
Telenor	8	2.50	1.94	0.57
Hydro	1	5.84	8.03	-2.20
Aker BP	6	17.32	17.32	0
Yara International	3	38.95	38.95	0
Gjensidige	7	10.20	9.31	0.88
Orkla	7	-1.52	-1.79	0.27
Schibsted "A"	6	17.03	18.02	-0.99
AF Gruppen "A"	9	-17.40	-16.22	-1.17
Arendals Fossekompani	1	46.18	45.70	0.48
Atea	6	-9.34	-9.83	0.49
Bouvet	6	-21.59	-21.59	0
Borragaard	2	-50.76	-48.24	-2.52
Frontline	1	0	0	0
Kongsberg Gruppen	8	16.53	14.35	2.18
Nordic Semiconductor	9	12.06	11.23	0.84
Protector Forsikring	7	65.05	64.59	0.46
Storebrand	7	30.18	28.94	1.25
Subsea 7	6	70.52	70.16	0.37
Veidekke	9	12.28	-10.99	-1.28
Walenius Wilhelmsen	1	88.13	89.27	-1.15

Table 5.7 displays the number of fundamental, hand-picked pairs included for each of the 20 largest Norwegian companies by market capitalization. Table 5.7 shows varying results. Some of the fundamental pairs resulted in a large profit, while others resulted in significant losses.

5.4 Trade Comparison

Where the above section examined the differences in strategy return, this part will examine the normal and weighted trades performed during the reinforced framework. The comparison of trade returns are even more interesting than the comparison of strategy returns, as the comparison of trade returns better assesses the relative performance of the implemented trading rules of the reinforced framework, triggering the weighted trades. As table 5.8 displays, the weighted trades outperformed the normal trades by 2.58 percent per trade.

Table 5.8: Mean weighted return for the normal and weighted trades, aggregated over all pairs applied to the reinforced framework.

Trade type	**Mean return**
Weighted	11.98%
Normal	9.40%
Difference	2.58%

The below table gives and overview over the number of pairs formed with each proposed foreign country, and the number of trades for pairs in each country, both normal and weighted trades.

Table 5.9: Number of normal and weighted trades over all runs of the reinforced strategy, grouped by the foreign country constituting the pair.

Country	**Num. pairs**	**Normal trades**	**Weighted trades**	**W/N ratio**
Austria	19	285	22	7.7%
UK	33	471	68	14.4%
Australia	35	494	56	11.3%
Canada	30	660	72	10.9%
Hong Kong	17	360	29	8.1%
China	23	330	29	8.8%
Spain	34	294	25	8.5%
US	54	1040	104	10%
Sweden	19	373	48	12.9%
Finland	17	319	17	5.3%
Italy	27	473	42	8.8%
Japan	31	655	65	9.9%

Continued on next page

Table 5.9 – continued from previous page

Country	Num. pairs	Normal trades	Weighted trades	W/N ratio
Germany	14	254	24	9.4%
France	15	302	29	9.6%
Denmark	19	336	37	11%
Switzerland	39	808	104	12.9%

The number of formed pairs was greatest for the US, reflected in the high number of normal and weighted trades. The rightmost column displays the ratio of the weighted trades to the normal trades for each country. The ratio seems to be around 10 percent for all countries, with the UK, Sweden, Switzerland, Australia and Canada having the highest ratios.

Table 5.10 displays the average number of trades and up-weighted days based on trading rule 1 and 2 for all tested pairs. The number of trades and up-weighted days corresponds to the number of trades and days for one single run of the proposed reinforced framework. The table showcases the relative low number of weighted trades compared to normal, non-weighted trades, and the low number of up-weighted days based on trading rule 1 compared to up-weighted days based on trading rule 2.

Table 5.10: Average number of normal and weighted trades, and detected up-weighted days based on either trading rule 1 or 2

Measure	Average
Number of trades	18.3
Number of weighted trades	1.9
Up-weighted days based on trading rule 1	50.1
Up-weighted days based on trading rule 2	7.6

The below table reports the mean weighted trade return for the different types of trades of the reinforced strategy. The results of table 5.11 are sorted in descending order with respect to the weighted trade returns and grouped by the country of the foreign equity composing the pair.

Table 5.11: Mean weighted trade returns for normal and weighted trades, grouped by country of the foreign equity of the pair. Results are reported in percentage.

Country	Weighted trades	Normal trades	Trade return difference
Denmark	44.91	24.67	20.25
UK	43.89	19.73	24.16
Austria	38.58	18.19	20.39
Spain	23.83	28.98	-5.15
Hong Kong	21.68	11.09	10.59
Canada	19.47	16.48	2.99
China	19.01	13.61	5.40
Australia	16.69	12.60	4.09
Finland	14.79	3.29	11.50
Sweden	9.71	5.67	4.04
US	8.30	5.52	2.78
Italy	7.03	8.59	-1.56
Switzerland	5.56	7.00	-1.44
Japan	0.94	1.06	-0.11
Germany	-1.32	0.34	-1.66
France	-5.87	-0.35	-5.53

The difference in trade return shows varying results. Although varying, the weighted trades outperforms the normal trades in most of the cases. The weighted trades performs especially well in the UK, Austria and Denmark, while showing weaker performance in Spain and France.

Similar to table 5.11, table 5.12 reports the mean weighted trade return for the different types of trades of the reinforced strategy. The results of table 5.12 are sorted in descending order with respect to the weighted trade returns and grouped by the sector of the Norwegian equity composing the pair.

Table 5.12: Mean weighted trade return for normal and weighted trades, grouped by sector of the Norwegian component of the pair. Results are reported in percentage.

Sector	Weighted trades	Normal trades	Trade return difference
Oil	76.38	44.85	31.53
Industry	18.18	7.41	10.77
Technology	15.74	7.04	8.70
Real estate	14.80	13.79	1.01
Renewable	13.50	7.84	5.66
Seafood	11.68	5.75	5.93
Bank/Financial	3.23	6.87	-3.64
Health care	-19.80	-26.21	6.40

Table 5.12 displays the superior performance of the weighted trades, over normal trades, in all sectors except the bank/financial sector. Except from the sectors of Health care and bank/financial, the weighted trades displayed good performance, especially for pairs with Norwegian companies in the oil sector, with a weighted mean of 76.38 percent per trade, and a trade return difference of 31.53 percent per trade.

Figure 5.3 visualizes the number of trades over the entire trade phase, aggregated for each day of the trade phase for all pairs. Blue bars are "normal" non-weighted trades while red bars represent weighted trades (by any weight beyond normal base weight). The y-axis measure the number of trades while the x-axis represent the time interval of the trade phase. The bar chart is layered, meaning the bars are plotted above each other (not on top of each other).

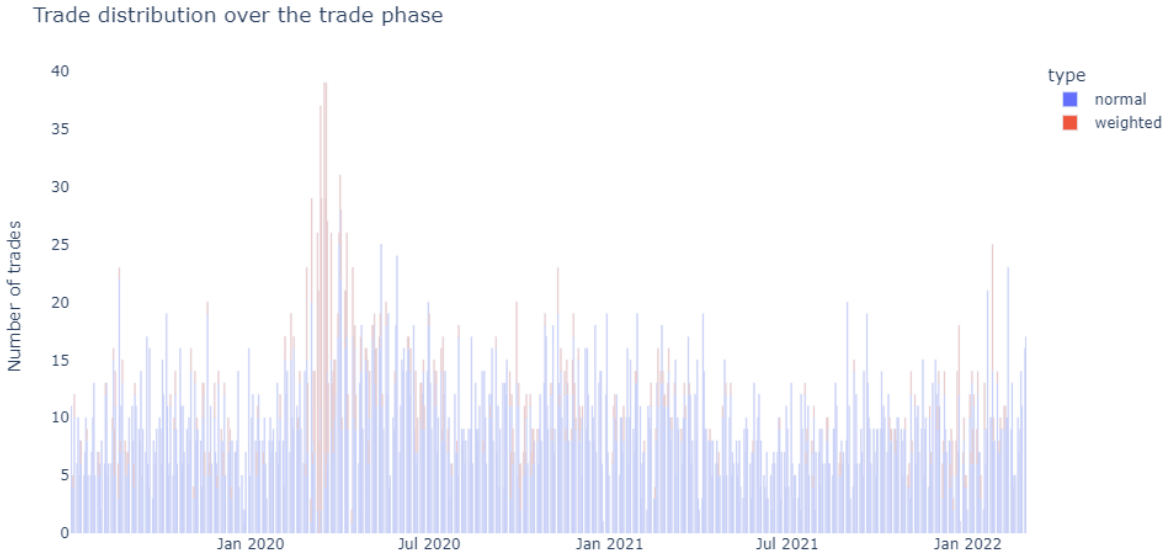

Figure 5.3: Number of performed normal and weighted trades for the reinforced strategy over the course of the trade phase, for all traded pairs.

The number of weighted trades show a significant spike between January and July 2020, during the initial global outbreak of the Covid-19 pandemic. The initial months of 2022 triggered a relative high amount of trades as well, both normal and weighted.

5.5 Factors

Table 5.13: Number of triggered trades grouped and summed by the factor triggering the trade, based on the "top three factors" for all pairs.

Factor	Number of triggered trades	Percentage of total
Copper	3 682	20.72
Brent Europe	3 523	19.82
Silver	2 719	15.30
10Y US bond return	1 751	9.85
10Y NO bond return	1 705	9.59
Gold	1 044	5.87
Aluminium	1035	5.82
VLLC rates	487	2.74
Natural Gas - RFV	445	2.50
10Y UK bond return	431	2.43
Baltic Dry Index	366	2.06

Continued on next page

Table 5.13 – continued from previous page

Factor	Number of triggered trades	Percentage of total
Fish Pool Spot	299	1.68
NOK - USD	90	0.51

Table 5.13 displays the number of triggered trades stemming from the development of the different proposed factors. The two factors triggering the most trades were Copper and the Brent Europe oil price, each totalling over 3 000 triggered trades. The return of the ten year US and Norwegian government bond also triggered a relative high amount of trades compared to the rest of the proposed factors. The list of factors displayed in table 5.13 only contains the factors with the highest relative number of triggered trades.

Table 5.14: Top three factors of high influence on triggering trade signals in the trade phase for all pairs, grouped and summed by sector of the Norwegian equity of the pair.

Sector	Top three factors	Number of trades triggered per factor
Seafood	Brent Europe	746
	10Y NO	684
	10Y US	418
Bank	Brent Europe	416
	Silver	371
	Copper	240
Tech	Brent Europe	678
	Copper	488
	Aluminium	245
Oil	Copper	267
	VLCC	112
	10Y NO	107
Real estate	Silver	859
	Gold	682
	Brent Europe	460
Renewable	Copper	641
	Silver	259
	Aluminium	87

Continued on next page

Table 5.14 – continued from previous page

Sector	Top three factors	Number of trades triggered per factor
Industry	Copper	1164
	Silver	1060
	10Y US	589
Health care	Copper	233
	10Y US	126
	Brent Europe	112

Table 5.14 gives a more nuanced picture of the numbers presented in table 5.13, displaying the number of triggered trades for each sector. In displaying the number of trades triggered per factor in this way, one can easily see what factors had the largest influence on the pairs in different sectors. The distribution of the number of triggered trades per factor also showcases the nuances between the different sectors, which is what was sought to explore in this thesis. An important note is that the factors presented for each of the sectors in table 5.14, are factors showing low correlation to the Norwegian equity of the pair historically. The large number of trades triggered based on historically low correlated factors is very interesting, as it suggests that these factors do have an impact on the trading of the proposed pairs. The Brent Europe oil price and Copper is to be found in the top for almost all sectors, reflecting the high numbers reported in table 5.13. An interesting observation is the top three factors reported for pairs in the seafood sector, namely the Brent Europe oil price and the 10 year bond return for Norwegian and US government bonds. This will be commented on in the Discussion section.

5.6 Case Studies - Timing of Trades

Rather than solely focusing on and comparing strategy and trade performance, it is interesting to examine the trading patterns of the pairs put through the reinforced version of the pairs trading framework. Investigating the timing of the trades gives valuable insight and makes the interpretation of the results more nuanced. As the focus of the thesis is not solely optimizing the algorithms to obtain high returns, but rather being a tool for exploring nuances in the Norwegian equity market, the examination of the trading patterns is perhaps more interesting than reading the results of the strategies. It is important to note that for all pairs, the foreign equity is in the numerator and the Norwegian equity is in the denominator of the ratio calculation, presented in equation 3.1.

The pairs included in this case studies section are only a small part of the total pool of detected pairs. The chosen pairs are included as they mostly display successful trading phases, and provide useful insight in the execution of the reinforced trading framework. There are many pairs showing poor trade phases which are excluded in this section on purpose, as they do not provide much useful insight. The reader should keep in mind that the pairs displayed in this section do not make up a perfect picture of all pairs.

The below figures represent one complete run through the reinforced pairs trading framework. The colored vertical lines visualizes the detection of up-weighted days, being detected by either of the proposed trading rules. Yellow lines correspond to up-weighted days based on trading rule 1 while magenta colored lines correspond to up-weighted days based on trading rule 2. The confluence of either type of these up-weighted days (vertical lines) and detected trading signals will result in larger trades (weighted trades), and is what we are most interested in examining. The blue line represents the z-score of the pair ratio, measured on the y-axis, while the circles of varying size represents the executed trades. Larger circles corresponding to a trades of larger weight. The legends on the right side of the plot represent the different states of the trading signals, and the x-axis represents the duration of the trading period, from left to right.

5.6.1 Gjensidige Forsikring - Allianz

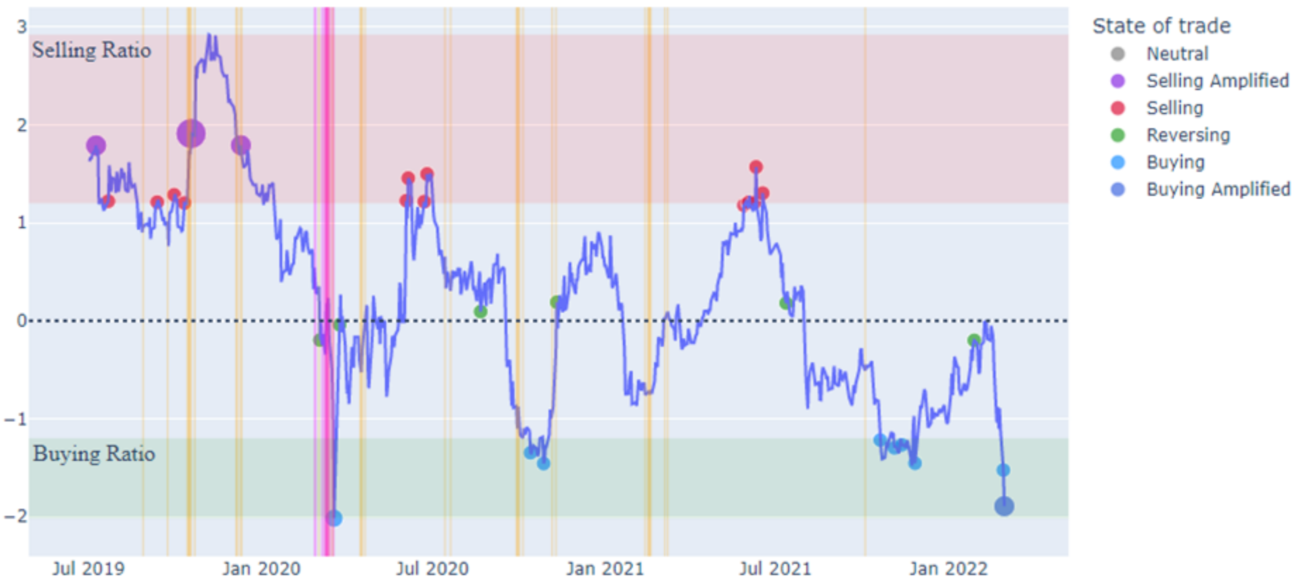

Figure 5.4: Result of the performed trade phase of Gjensidige Forsikring - Allianz using the reinforced framework. Trades and up-weighted days of various kind are visualized as colored circles and vertical lines, correspondingly. The ratio z-score is represented by the blue continuous line.

Table 5.15: Return data and statistics of the performed trade phase of Gjensidige Forsikring - Allianz, using the reinforced framework.

Metric	Value
Top factors with counts	Brent Europe: 22 Copper: 8 Silver: 6
Number of trades	25
Number of weighted trades	2
Number of up-weighted days based on trading rule 1	41
Number of up-weighted days based on trading rule 2	7
Strategy return	Reinforced: 11.83% Basic: 11.31% difference: 0.53%
Weighted order returns	4.0: 19.25% 1.5: 21.99%
Weighted order dates	2019-10-18 2020-03-18
Weighted order aggregated	Mean: 20% Median: 20.62%
Normal order aggregated	Mean: 12.04% Median: 11.89%
Difference trade return	7.96%

Looking at figure 5.4, the pair displays promising behaviour: the z-score subsequently breaches through the selling and buying boundaries and the zero-line, resulting in profitable trades. This is reflected in the strategy and trade returns, displayed in table 5.15. The best trading signal occurs when the z-score just breaches through negative two, between January and July 2020. This trade can be found in table 5.15, inspecting the "Weighted order returns" row, which maps the weight of the weighted trades to the corresponding return. The trade was triggered 2020-03-18, was given a weight of 1.5, and yielded a return of 22%. The period of ratio divergence before the buying signal of the above described trade is placed, triggers a lot of vertical magenta lines (trading rule 2), meaning that the period leading up to the trading signal, and maximum ratio divergence, is potentially driven by the checks constituting trading rule 2. A sharp divergence followed by an equally sharp convergence results in a highly profitable trade. It is these trades we seek to detect and exploit, and the above pair showcases promising behaviour. The factors detecting the up-weighted days, can be found in the top row of table 5.15. The Brent Europe oil price represented the largest influence, followed by Copper and Silver.

5.6.2 Schibsted A - Viviendi

Figure 5.5: Result of the performed trade phase of Schibsted A - Viviendi using the reinforced framework. Trades and up-weighted days of various kind are visualized as colored circles and vertical lines, correspondingly. The ratio z-score is represented by the blue continuous line.

Table 5.16: Return data and statistics of the performed trade phase of Schibsted - Viviendi, using the reinforced framework.

Metric	Value
Top factors with counts	Copper: 40 Silver: 35 Gold: 2
Number of trades	35
Number of up-weighted trades	5
Number of up-weighted days based on trading rule 1	60
Number of up-weighted days based on trading rule 2	10
Strategy return	Reinforced: 15.09% Basic: 14.52% difference: 0.57%

Continued on next page

Table 5.16 – continued from previous page

Metric	Value
Weighted order returns	1.5 11.62% 3.0: 14.48% 4.0: 22.55% 4.0: 22.17% 4.0: 21.53%
Weighted order dates	2020-03-11 2020-03-16 2020-03-17 2020-03-31 2020-04-20
Weighted order aggregated	Mean: 19.75% Median: 21.53%
Normal order aggregated	Mean: 18.03% Median: 14.75%
Difference trade return	1.72%

As previously mentioned, and as will be clear when examining the different case studies in this section, the period around the initial outbreak of the Covid-19 pandemic is subject to a lot of detected up-weighted days and thus that specific period has a significant impact on the pair performance.

Figure 5.5 visualizes the trading phase of another fundamental, hand-picked pair: Schibsted A and Viviendi. As with Gjensidige and Allianz, the z-score of the pair ratio seems to follow a favourable pattern, oscillating around the zero-line. Inspecting table 5.16 one can see that the trading of the pair resulted in five up-weighted trades, all taking place in the interval from 03-11 to 04-20 of 2020. These weighted trades are easily spotted in figure 5.5.

Something to keep in mind while interpreting the charts in this case studies section is that the signals detecting up-weighted days are "lagging". The correlation-calculation is based on a rolling 20 day window and thus the exact day the signal is triggered is the end of the corresponding rolling window. Since the colored lines are only visualized for the exact day when the signal is detected, the rolling window forming the basis of the calculation is not visualized.

From the start of 2022 and on, the pair ratio experienced a continuous increase. This may be seen in context with the stock price of Schibsted falling noticeably in the same period. Since Schibsted is in the denominator of the pair ratio calculation, the pair ratio will increase if the stock price of Schibsted falls. Several up-weighted days are detected the start of 2022 and on, but no weighted trades are executed, as the horizontal lines in this period do not coincide with any trading signals. This is a caveat of the strategy execution: although several up-weighted days are detected in a period, weighted trades are not executed unless the trades coincide with the detection of an up-weighted day.

5.6.3 Mowi - Vinci

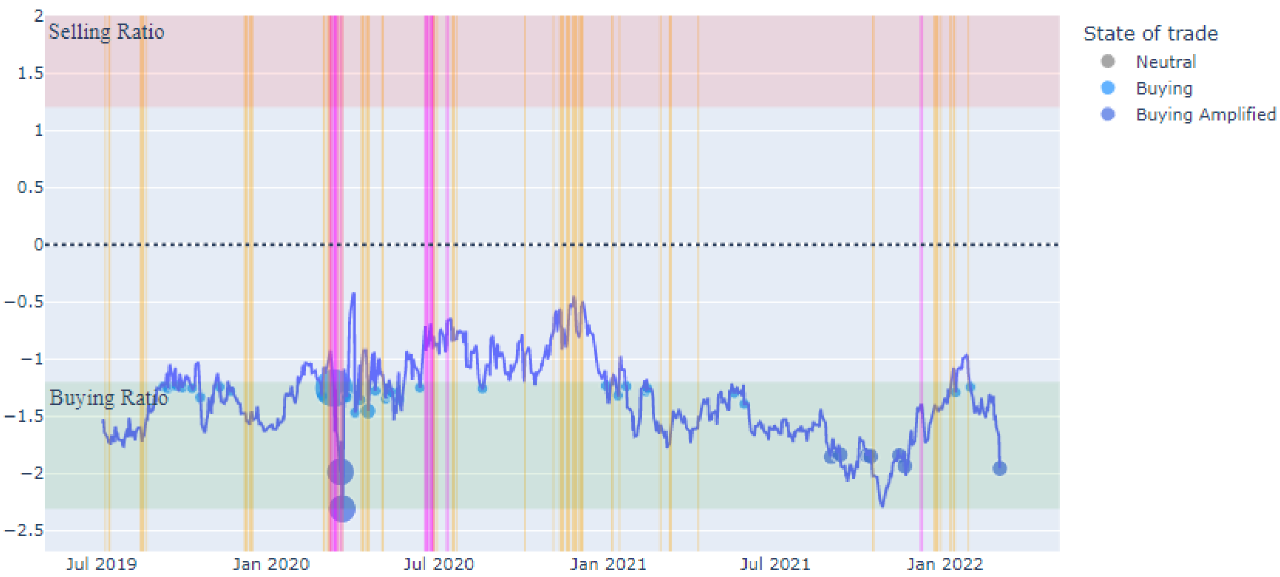

Figure 5.6: Result of the performed trade phase of Mowi - Vinci using the reinforced framework. Trades and up-weighted days of various kind are visualized as colored circles and vertical lines, correspondingly. The ratio z-score is represented by the blue continuous line.

The next case study is of Mowi and Vinci. The case study is included to showcase what may happen when the development of the pair in the trade phase does not follow the development detected in the formation period, resulting in misplaced boundaries for the ratio z-score. Misplaced z-score boundaries have a negative impact on the trading of a pair, as the boundaries are not fitted to the development of the pair in the trade phase and will thus not detect trading signals appropriately. This is a common challenge when performing pairs trading, and will be commented on in the continuation of the case studies in the Discussion section. Since one does not have access to the "test data" in real time, the boundaries of the z-score are placed on background of the development in the formation period. Thus, one cannot be sure whether the development observed in the formation period will continue on in the trade phase, or not; one simply makes a qualified guess of what the pair development in the trade phase will look like, based on the pair development in the formation period.

Table 5.17: Return data and statistics of the performed trade phase of Mowi - Vinci, using the reinforced framework.

Metric	**Value**
Top factors with counts	Brent Europe: 26 10Y NO: 14 10Y US: 13
Number of trades	40
Number of up-weighted trades	4
Number of up-weighted days based on trading rule 1	87
Number of up-weighted days based on trading rule 2	15
Strategy return	Reinforced: -6.43% Basic: -6.63% difference: 0.2%
Weighted order returns	-
Weighted order dates	2020-03-09 2020-03-16 2020-03-18 2020-04-15 2020-04-20
Weighted order aggregated	-
Normal order aggregated	-
Difference trade return	-

Table 5.17 presents the result of the strategy execution of Mowi and Vinci. Although both strategies yield negative returns, the reinforced strategy scores better than the purely statistical strategy, with 0.2% per trade on average.

A common observation when running the reinforced pairs trading framework on the detected pairs, is that the Norwegian seafood companies seem to detect a higher number of up-weighted days compared to companies from other sectors, and thus also a higher number of triggered weighted trades. With 87 detected up-weighted days based on trading rule 1 and 15 detected up-weighted days based on trading rule 2, Mowi and Vinci is well above the average number of detected up-weighted days for all pairs, of respectively 50.1 and 7.6, listed in table 5.10. The number of up-weighted trades for Mowi and Vinci of four, is twice the average of up-weighted trades for all pairs.

5.6.4 International Bus - Olav Thon Eiendom

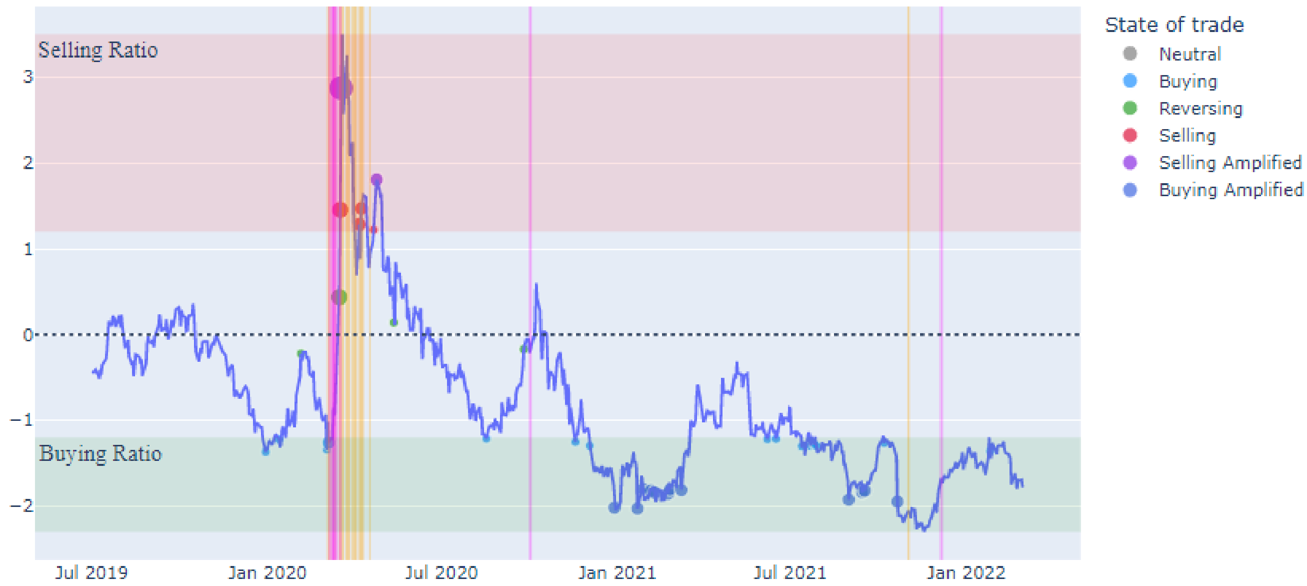

Figure 5.7: Result of the performed trade phase of International Bus - Olav Thon Eiendom using the reinforced framework. Trades and up-weighted days of various kind are visualized as colored circles and vertical lines, correspondingly. The ratio z-score is represented by the blue continuous line.

The last case study is of International Bus and Olav Thon Eiendom. As the Norwegian real estate companies have performed well in comparison to pairs detected with Norwegian companies from other sectors, this case study is provided to comment on why this might be.

A common pattern for the Norwegian real estate companies included in this thesis, Olav Thon and Selvaag Bolig, is that the strategy performance is heavily dependent on trades initialized in the period around the initial outbreak of the Covid-19 pandemic (more so than pairs from other sectors). Investigating figure 5.7 one can see that the above observation holds in the case of International Bus and Olav Thon Eiendom. The calculated ratio of the pairs formed with Olav Thon and Selvaag Bolig often rockets during this period in time, experiencing extreme z-score deviations of several standard deviations (in absolute terms). This often results in trades being up-weighted in the extremes of the ratio deviation, often giving high returns upon ratio convergence. The weighted order returns and weighted order dates of table 5.18 showcases the above-described pattern. The trade of largest weight, 6.6, corresponds to the large purple circle around three standard deviations of the ratio z-score in figure 5.7. This trade yields are return of 37.33%, which is the highest return of any of the performed trades for the given pair.

Since the number of weighted trades is low, some ”extreme” trades as the one described above, heavily influences the mean value of the weighted trades (since the mean is calculated as a weighted average). Thus, it is in some cases interesting to compare the mean to the median value. The median value, ordering the returns and returning the middle observation, addresses the skew of the returns. From table 5.18 one can see that the median value of the weighted trade returns is significantly lower than the mean value, showing that the majority of the weighted trade returns are in the lower range. Since some single extreme trades of high weight may dominate the contribution to the mean value calculation, the mean measure might, in such cases, give a misleading representation of the return distribution.

Although the median measure is in some cases a useful supplement to the mean measure, the median measure may be misleading when the number of weighted trades is low, as it is for most pairs. As an example, calculating the median of two observations does not make sense. Since most pairs have very few weighted trades over the trade phase, the median measure will in most cases give a misleading representation of the trade return distribution, and is thus omitted in the reporting of results in the results section.

Table 5.18: Return data and statistics of the performed trade phase of International Bus - Olav Thon Eiendom, using the reinforced framework.

Metric	**Value**
Top factors with counts	Silver: 21 10Y NO: 4 Gold: 2
Number of trades	33
Number of up-weighted trades	5
Number of up-weighted days based on trading rule 1	28
Number of up-weighted days based on trading rule 2	8
Strategy return	Reinforced: 8.78% Basic: 4.64% difference: 4.14%

Continued on next page

Table 5.18 – continued from previous page

Metric	Value
Weighted order returns	2.0: 28.56% 2.2: 15.06% 2.2: 17.65% 3.3: 17.43 6.6: 37.33%
Weighted order dates	2020-03-05 2020-03-17 2020-03-17 2020-04-06 2020-04-08
Weighted order aggregated	Mean: 26.56% Median: 17.65%
Normal order aggregated	Mean: 19.92% Median: 18.17%
Difference trade return	6.64%

Chapter 6

Discussion

6.1 Performance

6.1.1 Strategy Performance Comparison

Pairs Detected by Clustering and Cointegration

The reinforced version of the pairs trading framework outperformed the basic version overall, regarding the average strategy return for all pairs, visualized in table 5.4. With a mean strategy return of 9.96%, the reinforced trading scheme outperformed the basic scheme with 0.33% per trade over all proposed combinations of pairs. This indicates that the implemented changes (in forms of two trading rules) in the reinforced pairs trading framework did result in higher returns. This improvement from the basic framework may to some extent reinforce the thesis objective that foreign investors have a simplified view of the Norwegian equity market. The observed positive returns for both strategies may possibly also challenge the results reported by Andersen and Tronvoll (2018) and Mikkelsen (2017), that there seems not to be any arbitrage opportunities on the Norwegian stock market [4], and that none of the evaluated strategies had significant profits after accounting for transaction costs [25].

The reader should be aware that the duration of each trade is neglected in the presented results. The trade phase stretches from 2019 – 2022 and the duration of the performed trades may thus be in the interval of a few days up to the entire trade phase. In most cases the proposed pairs show several trades over the trading period, but there are some pairs displaying a very low number of trades. These pairs are often subject to one or more unconverged trades, resulting in the pair ratio being inefficiently traded. As displayed in table 5.10, the average number of trades per strategy run for all pairs is approximately 20.

A more nuanced display of the average trade return per strategy is given in table 5.5, grouping the return by country of the foreign pair. Inspecting the table, it is evident that the strategy performance varies over the different countries, and some countries even resulting in a negative mean weighted strategy return over all pairs formed for that country. Although it may be tempting to look at the two left-most columns in table 5.5, representing the mean weighted strategy return over all traded pairs for the foreign country, the right-most column is the one to put most emphasis on. Even though pairs formed with Spanish equities yields the second highest returns, the difference in strategy return is negative. The right-most column displays the average difference in strategy return for all pairs for each country. Taking a quick look at the column shows that the reinforced trading framework outperformed the purely statistical framework in most countries. It is also interesting to see in what countries the difference in returns between the two frameworks are largest: in this case, the United Kingdom and Denmark, with 4.23 and 1.02 percent per trade per strategy correspondingly. One would thus have benefited most on applying the reinforced framework on the observed pairs from the United Kingdom and Denmark, compared to applying the purely statistical framework. This observation may suggest that the pairs formed with foreign equities from the United Kingdom and Denmark are suspect to conditions favouring the reinforced framework, and that the nuances in the development of the pairs for these countries, may to some extent be overlooked.

Table 5.6 further nuances the results by grouping the pairs by the sector of the Norwegian company of the pair. This specific way of presenting the results may be more interesting compared to grouping the pairs by country, as it differentiates between the sectors of the Norwegian equity market, and the sector differences are more directly related to the thesis objectives and market nuances. In addition to providing the mean weighted strategy return for all pairs over the respective sectors, table 5.6 displays the number of pairs formed within the different sectors. Comparing the sectors, the Industry sector detected most pairs with a total of 95, followed by Real estate and Seafood companies.

The high number of pairs formed with Norwegian seafood companies is interesting, as the Norwegian seafood companies does not have many good fundamental foreign pairs. The relative high number of detected pairs may be a result of the cointegration and clustering procedures, detecting pairs solely based on mathematical relationships. As the Norwegian seafood companies lack good foreign pairs, the detection of viable foreign pairs from other industries may provide useful insight. Pairing Norwegian seafood companies with each other, Mikkelsen (2017) reported no significant returns of any of the pairs [25].

As the detected pairs with Norwegian seafood companies in this thesis resulted in positive returns overall, it may be interesting to ask whether the search for viable pairs to Norwegian seafood companies in foreign countries, based on statistical approaches, may yield more profitable pairs then pairing the seafood companies with each other. Although the foreign pairs are not in the seafood business, the equity prices have displayed similar development over the formation period, and in several cases the trade phase as well, with Norwegian seafood companies. As the detection of these seafood/non-seafood pairs is a direct result of the mathematical procedures applied during pair detection, one should be careful to trade on any of these pairs, as detected relationships may be spurious. On the other hand, the relative high amount of pairs detected with Norwegian seafood companies, and the displayed similarities in development over, give or take, a ten year period, provides evidence that Norwegian seafood companies may actually have good foreign pairs based on statistical properties.

A more nuanced view of foreign pairs formed with Norwegian seafood companies is provided in table 5.1. Companies such as Mowi, Bakkafrost, Leroy Seafood Company, Salmar and Norway Royal Salmon are all represented among the top 20 Norwegian companies regarding the number of pairs formed with foreign equities. Mowi detected the second highest number of pairs overall, with a total of 29 foreign pairs.

Going back to table 5.6, the difference in return between the reinforced and the purely statistical framework over the different sectors display some interesting results. Beginning in the lower end of the table one can see that trading the detected "health-care-pairs" would have yielded terrible results. The health care companies were for a while considered not to be included in the initial pool of equities, as health care companies are sensitive to news and company specific factors. In the end, the health care companies were decided to be included in the study, and the results speak for themselves.

Out of the remaining sectors in table 5.6, pairs formed with Norwegian oil and real estate companies performed best, with average strategy returns of 51.50 and 13.93 percent for the reinforced framework respectively. Pairs including Norwegian seafood companies yielded the sixth highest returns on average, with a mean weighted strategy return of 7.94 over all pairs. Investigating the difference in strategy return in table 5.6, one can see that two of the top three sectors regarding strategy return, oil and industry, are also found in the top three of the highest difference in strategy return. This may suggest that nuances in the development of the pairs within the oil and industry sector may to some extent be overlooked.

Fundamental Pairs

Table 5.7 presents the results of the hand-picked fundamental pairs for each of the top 20 Norwegian equities by market capitalization, and as stated in the results section, the results are varying. The fundamental pairs are not detected by either cointegration or clustering. As the number of fundamental pairs for each Norwegian equity is low, and the results being of varying nature, one should be careful to draw any conclusions based on the presented results. On the other hand, the results presented in table 5.7 may be interesting to explore in more detail, as some of the proposed fundamental pairs yield favourable returns.

Regarding the difference in strategy return for the fundamental pairs for each Norwegian equity, fundamental pairs picked for Storebrand and Gjensidige, resulted in the highest return difference. These two companies are involved in the insurance business and share similar fundamental characteristics. Actually, the fundamental pairs picked for these two companies are the same, listed in table A.1 in the appendices. As the pairs composed with the same foreign companies for Storebrand and Gjensidige results in favourable returns, and show a relative high difference in strategy return, these pairs may be worth investigating in more detail. Although the fundamental pairs formed for Storebrand and Gjensidige showed promising results, a larger portion of the other proposed fundamental pairs ended up showing negative results. The fundamental pairs formed for Equinor and DNB resulted in a mean weighted strategy return for the reinforced strategy of -12.21% and -24.29% respectively.

The difference in performance between the pairs detected by clustering or cointegration and the hand-picked fundamental pairs may be interesting. The pairs detected by cointegration and clustering showed relatively lower and more stable returns than the proposed fundamental pairs. This may in part be due to the large number of pairs detected by cointegration and clustering compared to those formed from the hand-picked foreign fundamental equities. The relative stronger performance of the cointegration and clustering pairs compared to the fundamental pairs may to some extent suggest that cointegration and clustering approaches are superior in detecting pairs suitable for a pairs trading framework, more so than pairs solely selected on a fundamental basis. This could also underline the importance of having a large number of pairs when implementing a pairs trading strategy, and that diversification is a central element.

6.1.2 Trade Performance Comparison

The only difference between the two proposed trading frameworks are the additional trading rules. The reinforced trading framework is simply a continuation of the basic framework, as both frameworks are built on the same base code. This slight difference means that the reason for the reinforced framework outperforming the basic framework, is that the weighted trades in the reinforced framework were superior to the normal trades. Separating the performed trades over all runs of the reinforced framework into normal and weighted trades, and examining and comparing the trades correspondingly, may thus be interesting. A display of the difference in returns between the two trade types is provided in table 5.11 and 5.12.

By Country

Observing the rightmost column in table 5.11 one can see that the weighted trades outperform the normal trades in most of the cases. For pairs formed with foreign equities from Denmark, the United Kingdom and Austria, the difference between the two trade types, in favor of weighted trades, is above 20 percent per trade. Investigating the difference in performance between the different trade types provides clearer comparison than comparing the differences in strategy return. The relative high differences in trade return between the two trade types reported for Denmark, the UK and Austria, agrees with the results presented in table 5.5. The UK and Denmark display the largest difference between the reinforced and basic strategies, of 4.23 and 1.92 percent per trade, respectively. Austria, on the other hand, does not display an equally large difference between the two proposed strategies, with 0.83 percent per trade. This may be seen in context with the relative proportions of weighted and normal trades for each country, reported in table 5.9

As The UK and Denmark have relatively high proportions of weighted to normal trades of 14.4 and 11 percent, respectively, the impact of the weighted trades in these two countries will be relatively large. The difference between the reinforced and basic strategy for pairs formed with foreign countries from Austria on the other hand, with a weighted to normal trade ratio of only 7.7 percent, will be less than for the UK and Denmark, as the weighted trades does not have an equally large impact.

As mentioned above, table 5.9 compares the two trade types to the total number of triggered trades, by country. The reason for including this table is to provide insight into how efficient the pairs in the different countries are traded, and display how large impact the resulting weighted trades will have. Countries with a relatively large portion of weighted trades compared to normal trades may represent more favourable conditions to be exploited, than countries with relatively few weighted trades.

By Sector

Even more interesting than comparing the differences in trade performance between the countries of the foreign equity of the pair, is assessing the difference in trade performance between pairs in different sectors of the Norwegian equity market. Sectors with larges differences between the returns of weighted and normal trades may suggest that the conditions targeted with the additional trading rules of the reinforced framework do occur in pairs from these sectors. Looking at table 5.12, one see that pairs formed in the oil sector resulted in the largest difference in trade returns of 31.53 percent per trade. This is by far the largest difference in trade return for all sectors, and suggest that pairs formed with Norwegian oil companies present more favourable conditions than pairs in other sectors of the Norwegian market, when applied to the proposed reinforced framework.

Overall, the results presented in table 5.12 show positive values for the difference between weighted and normal trades. This observation is promising, as the implemented trading rules in the reinforced framework seem to reap benefits. The sector nuances will be further discussed in the *Sector Nuances* section below.

Trade Distribution and the Price of Oil

The frequency of performed weighted trades is directly related to the thresholds set in the implementation of the new trading rules. "Looser" thresholds loosens the requirement for a weight update, while "strict" thresholds results in fewer weight updates. When building the framework and testing it on the detected pairs, different combinations of the thresholds were experimented with, with the goal of finding an as close to optimal combination as possible. Setting the thresholds too loose would result in a high degree of noise in the detected signals, decreasing the confidence in the resulting up-weighted trades. This would be destructive to the thesis as patterns would be hard to spot and it would be challenging to separate noise from signals. On the other hand, operating with too strict thresholds would possibly overlook good trading signals, and result in the ratio of weighted trades to normal trades approaching zero.

The market conditions targeted with the implemented trading rules of the reinforced framework do not occur with a high frequency. When these conditions do occur, on the other hand, they seem to trigger weighted trades resulting in higher returns than the normal trades. The thresholds of the new trading rules were thus set relatively strict, to reduce the noise in the triggered signals and cultivate detection of pure signals. A 10%

ratio of weighted trades to normal trades seems like an appropriate balance.

In addition to examining the trade performance, an examination of the trade distribution over the trade phase is also interesting. The below figure is identical to figure 5.3 presented in the theory section, but now including the development of the Brent Europe oil price.

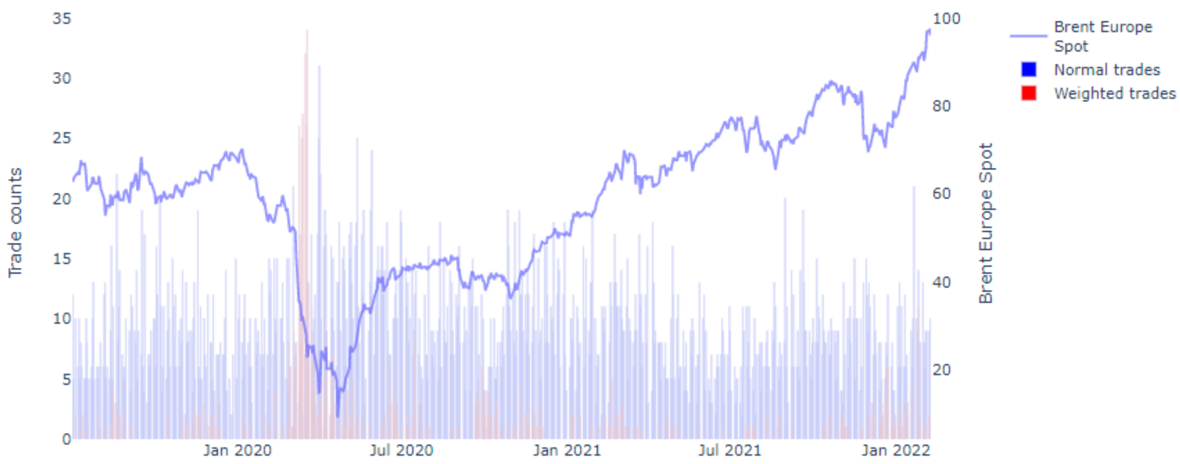

Figure 6.1: Number of normal and weighted trades for all pairs over the trade phase, including the development of Brent Europe oil price. Trade counts are measured on the left y-axis while the levels of Brent Europe are measured on the right y-axis.

The addition of the Brent Europe oil price gives a more nuanced interpretation of the development in the number of triggered trades. The confluence of a spike in the number of weighted trades and a sharp decrease in the oil price during the initial outbreak of the Covid-19 pandemic underlines the view that the Norwegian equity market is highly influenced by changes in the oil price. This pattern is exactly what is targeted with the inclusion of the additional trading rules in the reinforced pairs trading framework.

6.2 Sector Nuances

One of the most interesting studies in this thesis is the comparison of results between pairs in different sectors of the Norwegian market. Where investigating the results of a single pair may be too narrow in the big context, aggregating results for pairs according to sectors may put the results in context and facilitate the possibility for pattern observation.

table 5.14, perhaps the most informative table in the thesis, lists the top three factors having the highest impact on the detection of up-weighted days, aggregated for all pairs in a given sector. The table showcases the difference between sectors regarding factor sensitivity. It is important to state that the factors listed in the "Top three factors" column are factors showing a historical low correlation with the Norwegian equity constituting the pair. Investigating the top three factors of pairs where the Norwegian equity in the pair is in the seafood sector, one can see that the "seafood-pairs" have been sensitive to the development of Brent Europe and the Norwegian and American 10 year bond returns. From table 5.6 we have seen that the "seafood-pairs" showed positive returns over the trade phase in addition to the reinforced framework yielding higher returns than the purely statistical framework, meaning that the weighted trades performed better than the normal, non-weighted trades (as observed in table 5.12). Since, as we just observed from table 5.14 the majority of the up-weighted trading signals for the "seafood-pars" were triggered based on the development in the oil price and return of Norwegian and American 10 year bonds, and the weighted trades performing better than the normal trades, the detection of up-weighted days based on the above mentioned factors resulted in positive returns. This may suggest that the detected Norwegian seafood pairs are sensitive to changes in the development of the oil price and Norwegian bond returns, and that trading upon signals based on these factors for the detected pairs, yields positive returns. The results may be due to chance, but are certainly interesting.

One interpretation of the above observations could be: Norwegian seafood companies are sold down due to falling oil prices alongside the rest of the Norwegian market. However, the seafood companies do not have a direct exposure to oil prices. As observed in figure 3.6, the Norwegian currency rate often weakens as the oil price drops. As most Norwegian seafood companies sell their products in foreign currency, a weaker Norwegian currency rate equals larger profits. The possible confluence of Norwegian seafood companies being sold down due to falling oil prices and a weakening of the Norwegian currency rate is thus very interesting, as the fundamental conditions for the seafood companies have actually improved amid a period of stock price decrease.

The above argument regarding the factor sensitivity for the seafood-pairs can be loosely applied to pairs in all sectors for a set of factors (referred to below as $factor_1, factor_2, factor_3$), if the following is true:

1. The reinforced pairs trading framework yield higher returns than the purely statistical framework, as a consequence of the weighted trades performing better than the normal trades.

2. The relevant factors have been influential in detecting the up-weighted days of the

pair; $factor_1, factor_2, factor_3$ show up in the "Top three factors" of table 5.14.

Investigating table 5.6 once more, we see that pairs for all sectors listed in the table show positive values for the difference in strategy return, except for pairs in the *bank/financial* sector. Following the above argument, this implies that trading upon signals detected by changes in the top three factors listed in table 5.14 for pairs in all sectors except *bank/-financial*, yields positive returns. This may further suggest that pairs in the sectors of technology and real estate (in addition to seafood as shown above) are sensitive to changes in oil price. This is in part what is set out to explore in this thesis. The oil price, in this case Brent Europe, seems to have a significant impact on the detected pairs.

In addition to the Brent Europe oil price, the development of the price of Copper, Silver and to a lesser degree Aluminium, is frequently observed among the "Top three factors" in table 5.14. This suggests that commodity prices in general are influential in determining the development of the proposed pairs, for the given sectors (except *bank/financial*). The relative proportion of triggered trades for each of the "Top three factors" in each sector addresses the relative influence of the factors to the development of the pairs in the respective sectors. For seafood-pairs, for example, Brent Europe has a higher influence on the pairs compared to the 10 year bond returns.

6.3 Case Studies Continued

A common pattern observed in the presented case studies in the results section, is that the detection of up-weighted days (vertical lines) often seem to coincide with the "steepest" areas of the ratio z-score, meaning that the days when the pair ratio is subject to large changes are frequently detected as good trading days. This is promising, as a sharp widening of the pair ratio often results in the best trading signals. As high volatility often implies high uncertainty, executing what in retrospect may be seen as good trades, might be challenging in real-time. A higher confidence in possible trades, in the light of a sharp widening of the pair ratio is thus valuable, as it to some extent decreases the uncertainty of the trade.

The case study of Schibsted and Viviendi displays another pattern frequently spotted when examining different pairs: in periods of sharp ratio divergences, several up-weighted days based on trading rule 2 are detected right before the z-score approaches the maximum divergence, subsequently followed by several up-weighted days based on trading rule 1, forming a "belt" of yellow lines. Such a pattern seem to represent highly profitable trades (if executed), and is often what we are looking for. Several subsequent detected up-weighted days implies a longer period of correlation with historically low correlated factors, and the longer the pair is traded "inefficiently", the higher the chances are for the pair ratio to converge back to the mean.

In the case of Mowi and Vinci, it may seem like the z-score boundaries are set too wide for the development showcased in the trade phase. The pair ratio consequently lies below the historical mean, and since the initialized trades never cross the zero-line, all initialized trades are reversed on market price at the last day of the trade phase. A non-converging pair, such as Mowi and Winci may often result in poor profits, as discussed by Mikkelsen (2017) [25], reporting that 57% of the detected pairs never converged. R. A. Matta (2020) also discussed the challenge of poorly placed boundaries in the trade phase, reporting that loss of cointegration between stocks affected the profitability of the strategy negatively [23]. One can easily imagine Mowi and Vinci being a good pair if the z-score boundaries were better fitted to the pair development of the trade phase. The weighted trades initialized around March and April of 2020 would have resulted in significant profits if reversed in subsequent weeks of the trade initialization.

In retrospect of the strategy execution, it is easy to come up with reasons for the strategy under-performing by looking at charts such as figure 5.6. This is a common caveat when performing analysis of historical data and should be avoided. The reason for addressing the misplacement of the z-score boundaries in the Mowi-Vinci case, and the presence of possible profitable trades, is that the detection of up-weighted days provide new insight, not present in any historical literature on pairs trading. Even though the z-score boundaries may be misplaced, if one in real-time observes the pair development and several up-weighted days based on either trading rule 1 or 2 are detected in the light of a period of sharp ratio divergence, one may "break the rules" and place a trade. This is of course easier said than done, but the argument showcases the opportunities the detection of up-weighted days provide.

The high number of detected up-weighted days for Mowi and Vinci, reported in table 5.17 may suggest that the Norwegian seafood companies are traded more inefficient than companies from other sectors, as the seafood companies seem to be relatively more sensitive than companies from other sectors to changes in historically low-correlated factors. Figure 6.2 visualizes the executed trade phase for Mowi and Fabege, showcasing the relative high number of detected up-weighted days; the number of vertical lines is dense compared to other pairs. Other than the high number of performed trades for Mowi and Fabege, the executed reinforced strategy over the trade phase looks to be a perfect fit for a pairs trading scheme, with adequate volatility and several divergences followed by subsequent convergences.

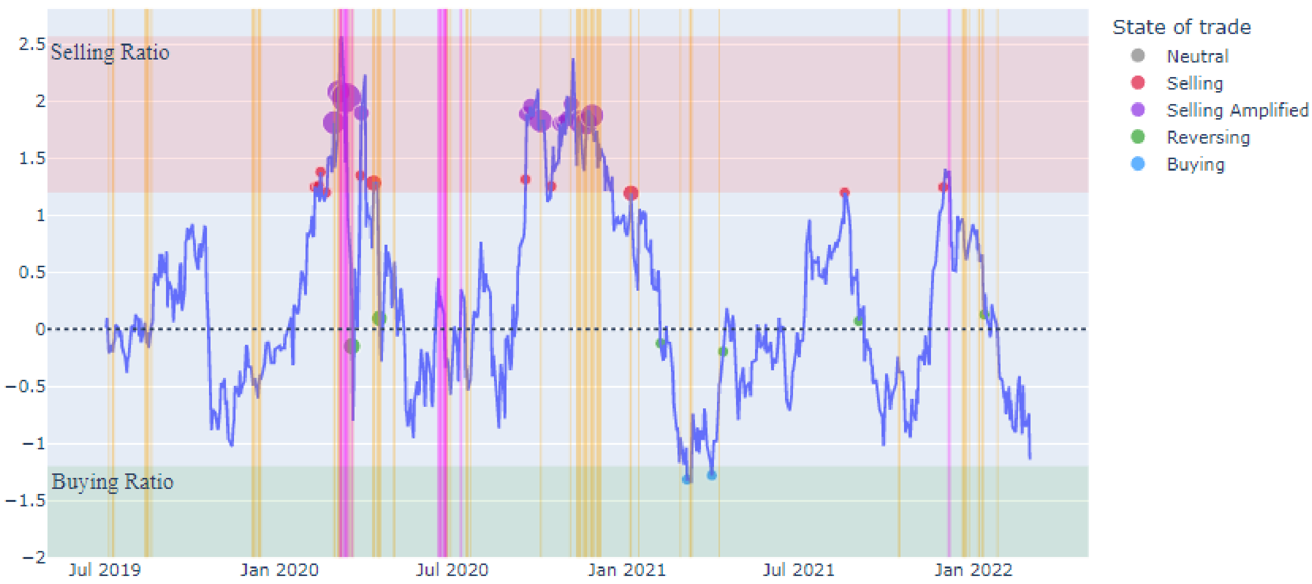

Figure 6.2: Result of the performed trade phase of Mowi and Fabege using the reinforced framework. Trades and up-weighted days of various kind are visualized as colored circles and vertical lines, correspondingly. The ratio z-score is represented by the blue solid line.

6.4 The Effect of Volatility

Figure 6.3 is yet another variation of figure 5.3 presented in the theory section. In addition to the number of trades aggregated for each day of the trade phase, the development of the VIX index over the same period is included. The VIX index (volatility index) is an index measuring the market's expectation of future volatility, based on options of the S&P500 Index [17].

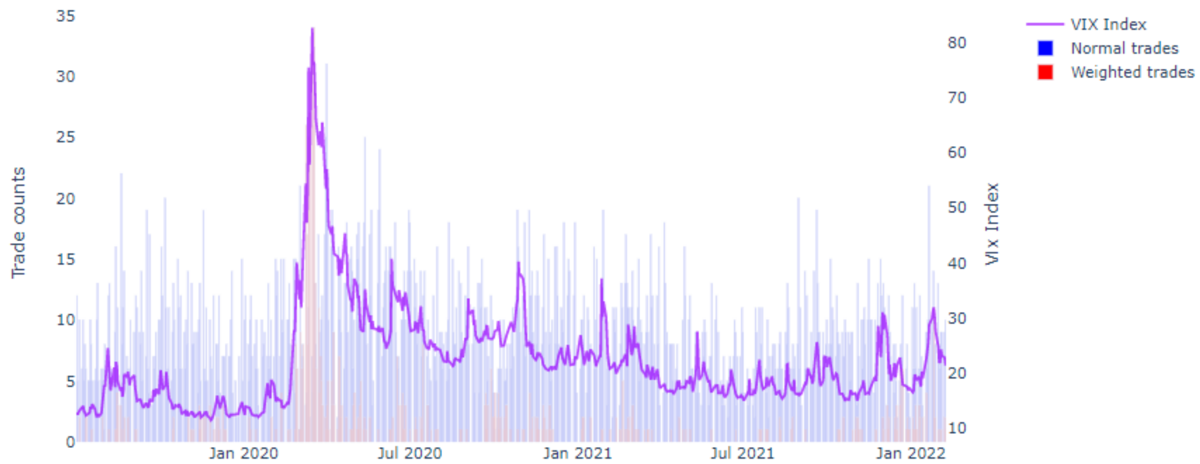

Figure 6.3: Number of normal and weighted trades for all pairs over the trade phase, including the development of the VIX index over the same period. The trade count is measured on the left y-axis while the levels of the VIX index are measured on the right y-axis.

Figure 6.3 showcases the relationship between market volatility, measured by the VIX index, and the number of weighted and normal trades triggered in the reinforced framework, as the levels of the VIX index displays a positive relationship with the number of triggered trades. Only visualizing the weighted trades, figure 6.4 makes the relationship between market volatility and the number of triggered, weighted trades, even clearer.

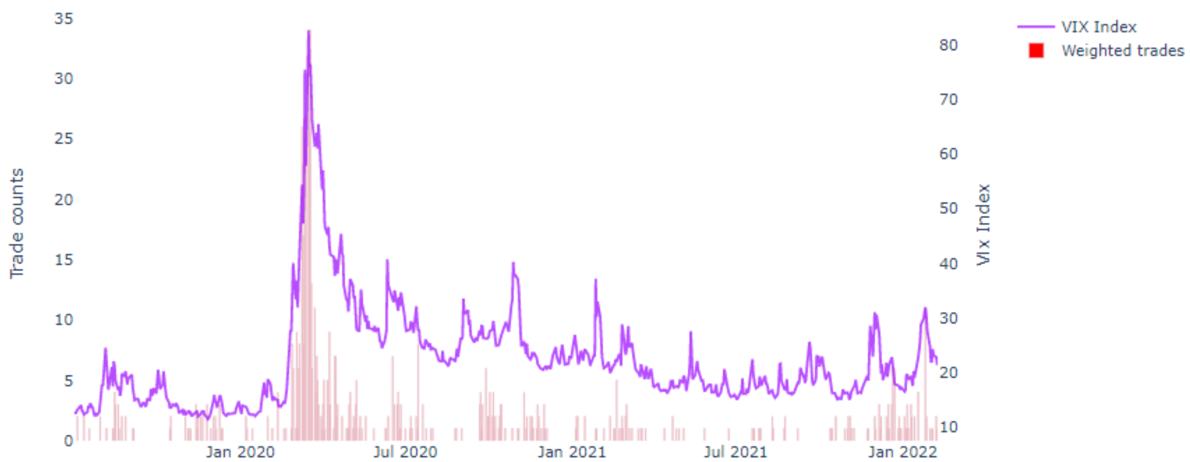

Figure 6.4: Number of weighted trades for all pairs over the trade phase, including the development of the VIX index over the same period. The trade count is measured on the left y-axis while the levels of the VIX index are measured on the right y-axis.

The trading rules triggering the weighted trades aim at detecting periods when the Norwegian market could be inefficiently traded. As the highest number of weighted trades coincide with high levels of the VIX index, and high volatility is often synonym with markets being less efficiently traded, it is tempting to suggest that the proposed trading rules do in fact detect periods where the Norwegian market is inefficient. This is natural as the nuances disappear (from foreign investors) when the volatility increases. If this is true, the proposed trading rules are able to detect periods when opportunities for statistical arbitrage are higher than normal.

The above observation regarding the effect of market volatility agrees with previous literature on pairs trading, observing that more often than not, pairs trading strategies performs better during periods of high volatility. Analysing pairs trading profitability in the United Kingdom from 1979 to 2012, Bowen and Hutchinson (2014) provide evidence for significant out-performance of pairs trading during global financial crisis [1]. Low and Faff (2016) also reported that all their strategies performed better during periods of significant volatility [27]. The reexamined study of Do and Faff (2010) found that the pairs trading strategy performs strongly during periods of prolonged turbulence, namely the 2000-02 bear market and the 2007-09 global financial crisis [12].

6.5 Further Work

As this thesis has been of the exploratory kind (exploring a new approach to pairs trading) several ideas and thoughts have been surfacing during and towards the end of the period. A section regarding thoughts and possibilities around further work is thus included to give the reader insight in some of the conducted thought-processes and experiments.

Predicting the Levels of the Norwegian Market

One of the initial ideas was developing a supervised ML-model aiming at predicting the levels of the Norwegian market, given the levels of foreign equity markets and a handful financial factors. Technically speaking, the model would train on the historical development of a handful of foreign stock exchanges, some financial factors and the Norwegian market, and predict the current levels of the Norwegian market on background of the patterns detected in the historical data. In predicting the current levels of the Norwegian market for one or two days in the future, given the current levels and development of the foreign exchanges and financial factors, one could make a more qualified guess about whether the Norwegian market is ”overreacting” to some external drivers compared to other foreign stock exchanges. If such an ”overreaction” had been the case, one may say that the Norwegian market is currently less efficiently traded (suspect to noise) than foreign markets, and use this information in the pairs trading scheme to perform a more qualified guess about the future development of the pair; if the Norwegian market is currently inefficiently traded it may affect the Norwegian-foreign pair relationship, as the Norwegian equity of the pair might be affected by the "overreaction" of the Norwegian market as a whole.

This proposal is closely related to the third check in trading rule 2, were we check if the Norwegian market has fallen relatively much compared to other foreign exchanges. The predictive performance of the model after exploring the opportunity for some time, was not found good enough to add value to the pairs trading framework.

Predicting the "Goodness" of a Possible Trade

Another conducted ML-experiment was trying to predict the "goodness" of a possible trade in a pairs trading framework, by assessing the current combinations of the proposed financial factors. To reformulate the above as a question: If the ratio z-score of a pair breaches through one of the predefined boundaries, indicating a possible trade, is it possible to predict whether or not this specific data point (day, in this case) is going to be a profitable trade or not? The problem was designed as a supervised ML problem, where the training data was the historical development of the proposed financial factors, and the target data would be a binary number (0 or 1), indicating whether the current data point would have been a profitable trade or not. A point was set to "1" if the pair ratio converged to the historical mean over a subsequent amount of days, and "0" if not. The exploration of such a model was conducted by the notion that a certain combination of financial factors, possibly driving the Norwegian market, would yield better trading signals than other.

Given the limited time frame of the thesis, the exploration of both of the above described models were ended. Although none of the models were successfully developed, the concepts would be interesting to explore further.

In general, a more extensive search for foreign pairs would be interesting, as only a subset of foreign markets and the corresponding largest equities by market capitalization were included in the thesis. As smaller companies are often not subject to the same scrutiny as larger companies, the possible opportunities of including additional, smaller companies, in the initial pool of equities would be interesting to explore.

Chapter 7

Conclusion

At the beginning of the thesis, the idea that foreign investors might have an oversimplified and naive view of the Norwegian equity market was outlined. The objective was introduced on the notion that *if* foreigners lack nuance in their approach to trading the Norwegian market, it is even more important to shed light on the nuances of the Norwegian market. This becomes of increasing importance as the relative share of foreign investors in the Norwegian market is ever increasing.

The implementation of the new trading rules in the reinforced framework, aiming at exploiting periods when the Norwegian equity market may be inefficiently traded, resulted in better performance than a purely statistical pairs trading framework. With a mean strategy return of 9.96 percent (including trading costs) the reinforced framework outperformed the purely statistical framework with 33 basis points per trade. Comparing the different kind of trades, it was also evident that the weighted trades outperformed the normal trades overall, with 2.58 percent per trade on average. The success of the reinforced framework and the weighted trades may support the idea of foreign investors having an oversimplified view of the Norwegian market. As the thesis objective is of qualitative nature, no conclusion is drawn based on the observed results. On the other hand, the results and observed patterns are certainly promising and motivates further exploration.

Table 5.14 sheds light on the sector nuances of the Norwegian market. The high number of trades triggered by the Brent Europe oil price, copper and silver of 55.84 percent of total trades, may provide evidence that some factors do in fact have a disproportional effect on the Norwegian market.

Figure 6.3 and 6.4 visualizes the close relationship between market volatility and the triggering of trades. The increase in the number of triggered, weighted trades, during periods of high volatility, supports the view that the nuances disappear (from foreign investors) when market volatility increases.

Bibliography

[1] M. C. Hutchinson A. Brown. “Pairs trading in the UK Equity Market: Risk and Return”. In: (2014).

[2] C. Alexander. “Cointegration and Asset Allocation: A New Fund Strategy”. In: (2001).

[3] C. Alexander. *Market Models - A Guide to Financial Data Analysis*. John Wiley Sons, LTD, 2001.

[4] H. Andersen and H. Tronvoll. “Statistical Arbitrage Trading with Implementation of Machine Learning”. In: (2018).

[5] M. Avellaneda and J. Lee. “Statistical Arbitrage in the U.S. Equities Market”. In: (2008).

[6] J. G. Baldwin. *The Currencies Most Affected by Falling Oil Prices*. URL: https://www.investopedia.com/articles/investing/102015/currencies-most-affected-falling-oil-prices.asp. (accessed: 05.02.2022).

[7] W. F. M. De Bondt and R. Thaler. “Does the Stock Market Overreact?” In: *The Journal of Finance* 40.3 (), pp. 793–805.

[8] A. Burkov. *Machine Learning Engineering*. True Positive Inc., 2020.

[9] J. Delua. *Supervised vs. Unsupervised Learning: What’s the Difference?* URL: https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning. (accessed: 25.02.2022).

[10] scikit-learn developers. *2.3. Clustering*. URL: https://scikit-learn.org/stable/modules/clustering.html#optics. (accessed: 12.02.2022).

[11] scikit-learn developers. *Demo of OPTICS clustering algorithm*. URL: https://scikit-learn.org/stable/auto_examples/cluster/plot_optics.html. (accessed: 13.02.2022).

[12] B. Do and R. Faff. “Does Simple Pairs Trading Still Work?” In: (2010), pp. 83–95.

[13] Euronext. *Stocks Oslo*. URL: https://live.euronext.com/nb/markets/oslo/equities/list. (accessed: 10.05.2022).

[14] G.Vidyamurthy. “Pairs Trading: Quantative Methods and Analysis”. In: (2004).

[15] E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst. “Pairs Trading: Performance of a Relative Value Arbitrage Rule”. In: (2006).

[16] N. Huck and K. Afawubo. “Pairs trading and selection methods: is cointegration superior?” In: *Applied Economics* 47.6 (Feb. 2015), pp. 599–613.

[17] Cboe Exchange Inc. *VIX Volatility Suite*. URL: https://www.cboe.com/tradable_products/vix/. (accessed: 25.04.2022).

[18] J.Rudy et al. “Statistical Arbitrage and High-Frequency Data with an Application to Eurostoxx 50 Equities”. In: (2010).

[19] D. Kahnemann. *Thinking, Fast and Slow*. 1st ed. Penguin Group, 2012.

[20] D. Kahnemann, O. Sibony, and C. R. Sunstein. *Noise: A Flaw in Human Judgment*. 1st ed. Little, Brown Spark Hacette Book Group, 2021.

[21] C. Krauss. *Statistical arbitrage pairs trading strategies: Review and outlook*. Tech. rep. 09/2015. Nürnberg, 2015. URL: http://hdl.handle.net/10419/116783.

[22] M.Perlin. “Evaluation of pairs-trading strategy at the Brazilian financial market”. In: (2008).

[23] R. Matta. “Pairs Trading on High-Frequency Data using Machine Learning”. In: (2020).

[24] D. Messelt. *The Norwegian krone’s correlation to the Brent oil price*. URL: https://www.alfredberg.com/the-norwegian-krones-correlation-to-the-brent-oil-price/. (accessed: 05.02.2022).

[25] A. Mikkelsen. “Pairs trading: the case of Norwegian seafood companies”. In: (2017).

[26] R. Pramoditha. *11 Dimensionality reduction techniques you should know in 2021*. URL: https://towardsdatascience.com/11-dimensionality-reduction-techniques-you-should-know-in-2021-dcb9500d388b. (accessed: 20.02.2022).

[27] H. Rad, R. K. Yew Low, and R. Faff. “The profitability of pairs trading strategies: distance, cointegration and copula methods”. In: *Quantitative Finance* 16.10 (2016), pp. 1541–1558.

[28] S. Raschka and V. Mirjalili. *Python Machine Learning*. 3rd ed. Packt Publishing, 2019, p. 3.

[29] S. Raschka and V. Mirjalili. *Python Machine Learning*. 3rd ed. Packt Publishing, 2019, p. 353.

[30] K. D Roe et al. “Feature engineering with clinical expert knowledge: A case study assessment of machine learning model complexity and performance”. In: 15 (2020).

[31] S. M. Saremento and N. Horta. “Enhancing a Pairs Trading strategy with the application of Machine Learning”. In: (2021).

[32] S. M. Sarmento and N. Horta. *A Machine Learning based Pairs Trading Investment Strategy*. Springer Nature Switzerland AG, 2021.

[33] statsmodels-developers. *statsmodels.tsa.stattools.coint*. URL: https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.coint.html. (accessed: 12.03.2022).

[34] M. Waskom. *seaborn.boxplot*. URL: https://seaborn.pydata.org/generated/seaborn.boxplot.html. (accessed: 20.03.2022).

Appendix A

Fundamental Pairs

Table A.1: Hand-picked fundamental foreign pairs, for each Norwegian equity.

Norwegian equity	Fundamental foreign pairs
Equinor	BP
	Shell
	Eni
	Total Energies
	ExxonMobil
	Chevron
DNB	Nordea
	Svenska handelsbanken
	Danskebank
	SEB skandinaviska enskilda
	BNP Paribas
	Lloyds Banking Broup
	Barclays
Telenor	Tele2
	Telia
	Deutsche Telekom
	Telefonica SA
	Vodafone Group
	TDC Group
	Swisscom
	BT Group

Continued on next page

Table A.1 – continued from previous page

Norwegian equity	Fundamental foreign pairs
Norsk Hydro	Alcoa
	Kaiser Aluminum
Aker BP (and DNO)	Lundin Energy
	Harbour Energy
	Hess Corporation
	Marathon
	Murphy Oil
	ConocoPhillips
	MEG Energy
Yara	CF Industries
	Nutrien Ltd
	ICL Group
	Mosaic
Gjensidige (Protector, Storebrand)	Topdanmark
	Sampo
	Tryg
	Allianz
	Zurich Insurance Group
	Admiral Group
	Talanx
Orkla	Danone
	Nestle
	Suedzucker
	Tate & Lyle
	AAK
	Kraft Heinz
	Premium Brands
	Unilever
Shibsted A	Sanoma
	Informa
	Lagardere
	Pearson
	JCDecaux
	Vivendi

Continued on next page

Table A.1 – continued from previous page

Norwegian equity	Fundamental foreign pairs
AF Gruppen A (Veidekke)	Skanska
	Balfour Beatty
	Hochtief
	NCC AB
	YIT OYJ
	Peab
	ACS Actividades
	Vinci
	Royal Boskalis Westminister
Arendals Fossekompani	Verbund
Atea (Bouvet)	Capgemini
	Atos
	Sopra Steria
	Computacenter
	Tietoevry
	Accenture
Borregaard	Solvay
	Johnson Matthey
	Element Solutions
	Chemours
Frontline	Tsakos Energy Navigation
	Teekay Tankers
Kongsberg Gruppen	Rheinmetall
	Leonardo SpA
	General Dynamics
	Lockheed Martin
	Northrop Grumman
	Raytheon
	BAE Systems
	Thales

Continued on next page

Table A.1 – continued from previous page

Norwegian equity	Fundamental foreign pairs
Nordic Semiconductor	TSMC
	STMicroelectronics
	Intel
	Qualcomm
	Broadcom
	Nvidia
	Advanced Micro Devices
	Lattice Semiconductor
	Realtek Semiconductor
Subsea 7	Saipem
	Oceaneering
	Helix
	Schlumberger
	Baker Hughes
	Halliburton
Wallenius Wilhelmsen	DFDS
	AP Moller-Mærsk

Appendix B

Clustering Results

Presented below are a handful of clustering results from applying the DBSCAN algorithm on the equity time series for some foreign countries. Clusters including both Norwegian and foreign equities are encapsulated in colored circles, as these are the clusters of interest. As the names of the equities forming the different clusters may be challenging to read from the figures, the Norwegian and foreign equities of the clusters encapsulated in colored circles, are listed in a subsequent table.

The clustering results are visualized in two dimensions, using the t-SNE algorithm proposed in the Theory section.

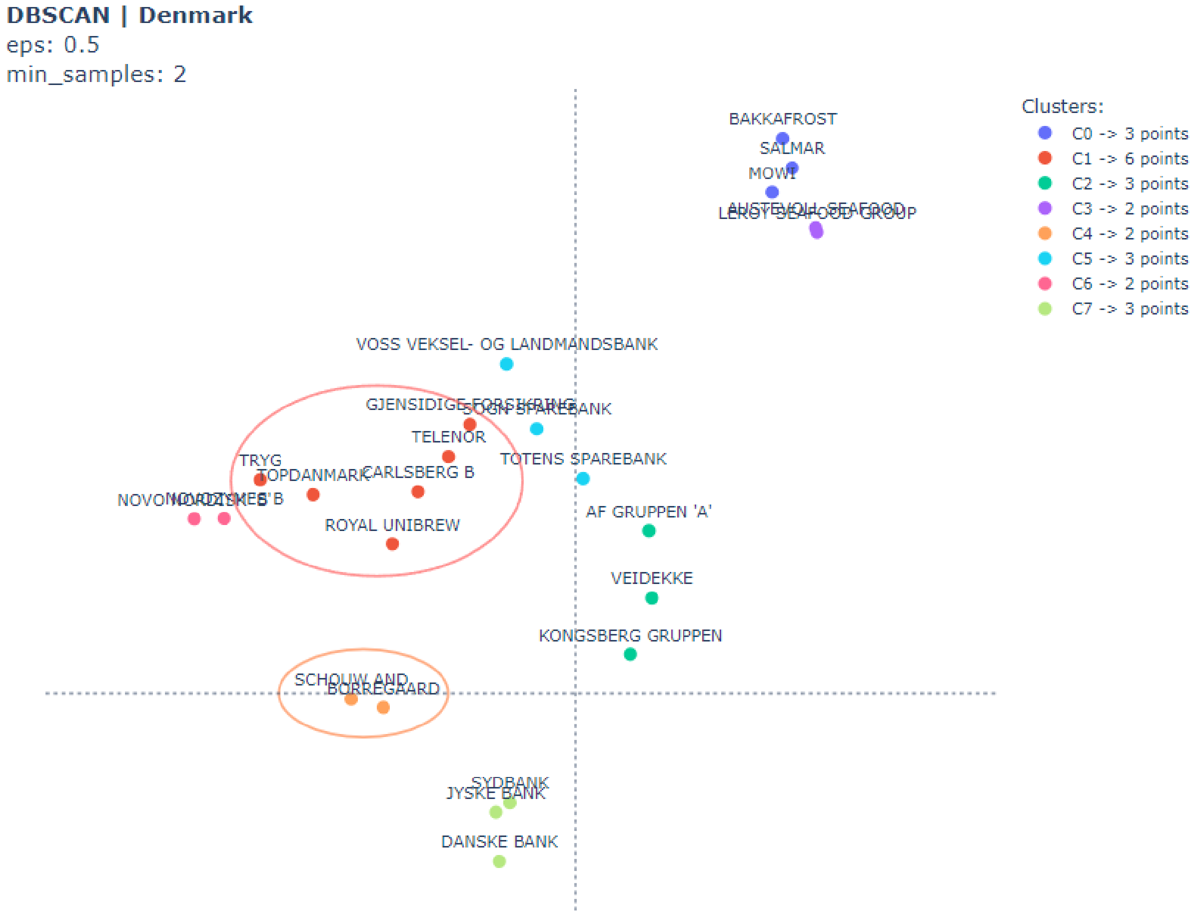

Figure B.1: Clustering results Denmark. Clusters containing Norwegian *and* foreign equities are encapsulated in colored circles.

Table B.1: Clustering results Denmark, corresponding to the clusters encapsulated in figure B.1.

Cluster number	Norwegian equities	Foreign equities
C1	Telenor Gjensidige Forsikring	Carlsberg B Tryg Royal Unibrew Topdanmark
C4	Borregaard	Schouw and

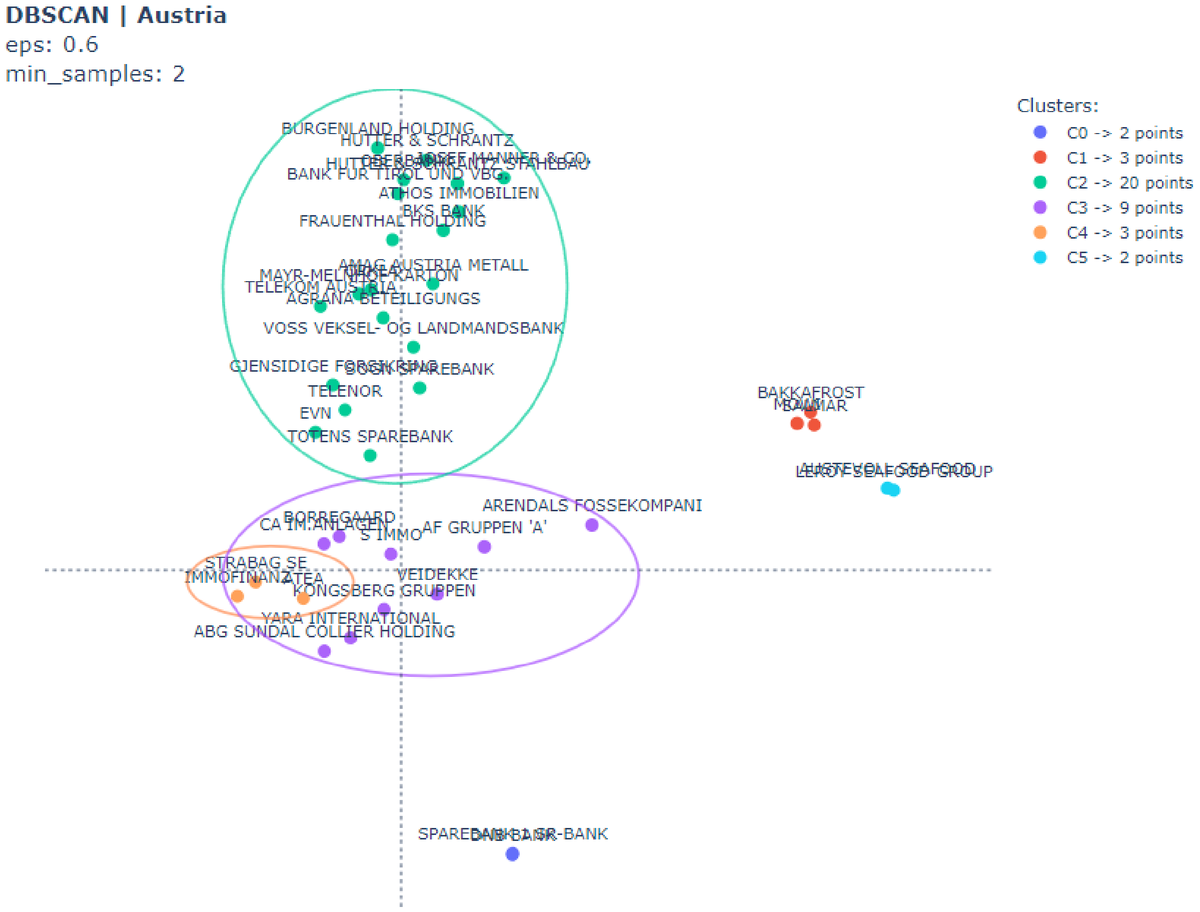

Figure B.2: Clustering result of Austria. Clusters containing Norwegian *and* foreign equities are encapsulated in colored circles.

Table B.2: Clustering results Austria, corresponding to the clusters encapsulated in figure B.2.

Cluster number	**Norwegian equities**	**Foreign equities**
C2	Telenor Gjensidige Forsikring Orkla Sogn sparebank Totens sparebank Totens sparebank	Evn Oberbank Telekom austria Agrana beteiligungs Amag austria metall Bank fur tirol und vbg. Bks bank Mayr-melnhof karton Athos immobilien Burgenland holding Frauenthal holding Hutter schrantz Josef manner & co.
C3	Yara international Abg sundal collier holding Af gruppen A Arendals fossekompani Borregaard Kongsberg gruppen Veidekke	Veidekke S immo
C4	Atea	Strabag SE Immofinanz

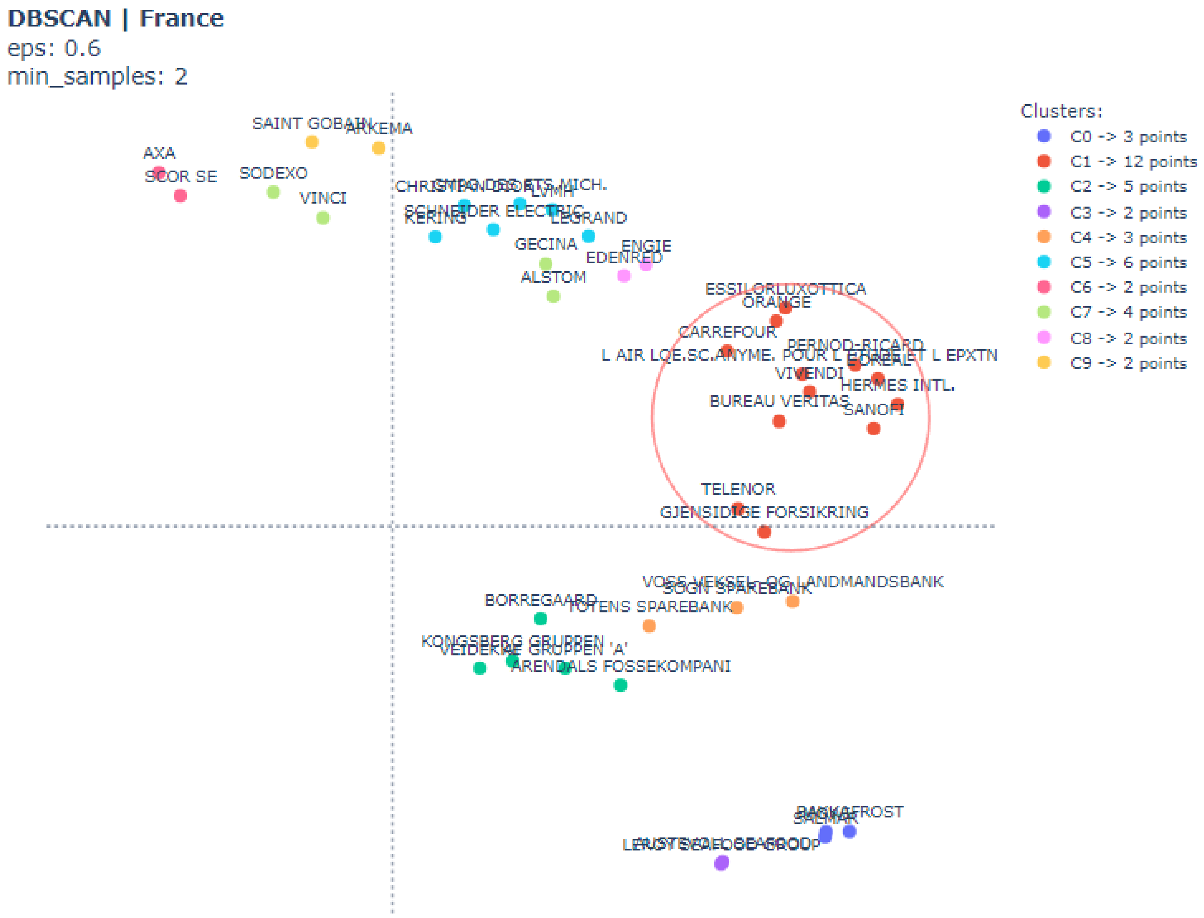

Figure B.3: Clustering result of France. Clusters containing Norwegian *and* foreign equities are encapsulated in colored circles.

Table B.3: Clustering results France, corresponding to the cluster encapsulated in figure B.3.

Cluster number	Norwegian equities	Foreign equities
C1	Telenor Gjensidige Forsikring	L'oreal Hermes intl. Sanofi Essilorluxottica L air lqe.sc.anyme. pour l etude et l epxtn. Pernod-ricard Orange Carrefour Bureau veritas Vivendi

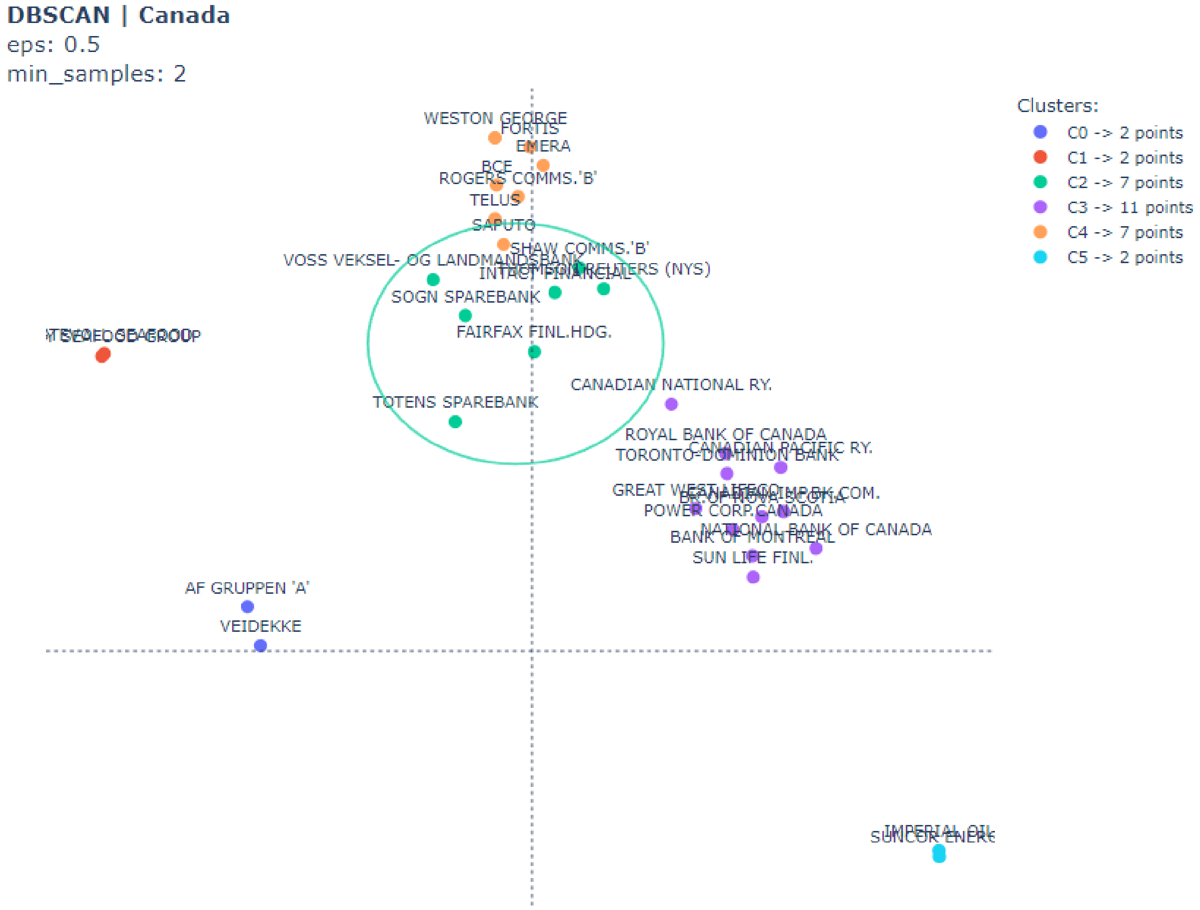

Figure B.4: Clustering result of Canada. Clusters containing Norwegian *and* foreign equities are encapsulated in colored circles.

Table B.4: Clustering results Canada, corresponding to the cluster encapsulated in figure B.4.

Cluster number	Norwegian equities	Foreign equities
C2	Sogn sparebank Totens sparebank Voss veksel- og landmandsbank	Thomson reuters Intact financial Fairfax finl.hdg. Shaw comms.'b'

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway