Pairs Trading

Evaluation of profitability and risks on the Swedish stock market

Master Thesis in Finance

Marc Schurer
Lund University, Sweden

Pavel Lisev
Lund University, Sweden

Supervisor: Lu Liu
Department of Economics
Lund University, Sweden

27th May 2015

Abstract

The turbulent market environment experienced over the last decades has attracted the broad interest of institutional and retail investors towards non-directional and absolute return investment strategies. The scope of this paper mainly concerns the investigation of whether a pairs trading strategy based on the cointegration approach generates excess returns on the Swedish equity market or fails to meet initial expectations. This is enhanced by an assessment of the volatility exposure relative to an investment in the corresponding benchmark. In this research, a comprehensive analysis of the pairs trading strategy is performed by (1) implementing a long-term rolling window backtest applied on the OMX, (2) a corresponding scenario analysis of the Swedish stock market including three different market environments, (3) an investigation of different in-sample pairs selection criteria and their respective impact, (4) an extended analysis of the strategy on the EUROSTOXX50 and DAX30 to support the robustness of the obtained outcomes. The empirical results suggest that the pairs trading technique is in fact profitable and superior in terms of return and risk relative to its benchmarks.

KEYWORDS: cointegration, statistical arbitrage, pairs trading, market-neutrality, rolling window backtest, OMX, DAX30, EUROSTOXX50, mean reversion, risk management.

Acknowledgment

We would like to take this opportunity to express the much-felt gratitude to our supervisor Lu Liu, PhD. Without her guidance and valuable insight this paper would not have been possible. We would also like to give expression to our gratefulness to the fellow students Stamatoula Pappa, Axel Olsson, Sepehr Sharafuddin and Jakob Melin from the Master of Finance program at Lund University for their much-appreciated comments on our work.

The programing for the models used in this thesis was made in MATLAB and is available upon request.

Pavel Lisev
Marc Schurer

Contents

Chapter 1: Introduction

The integral development and implementation of a lucrative trading strategy commences with the "comprehensible sifting of investment options" (Baronyan et al., 2010, p.116). There is a broad spectrum of possibilities designed to exploit given market inefficiencies in the form of trading strategies. Nevertheless, the majority fail to meet initial expectations concerning their actual performance as there exists a significant discrepancy between theoretically backtesting a trading idea and its implementation in a real market environment.

The successive events following the global financial meltdown have triggered abrupt alteration in investor psychology in conjunction with profound and abiding structural changes. The turbulent market environment has prompted investors to demonstrate higher concern regarding overall portfolio diversification and risk minimization, which in turn has stimulated the emergence of non-directional and absolute return investment strategies (Rittereiser and Kochard, 2010). The term 'strategy' can comprehensively be defined as "a set of rules and conditions that determine such factors as what generates a signal and how to enter, manage, and exit the trade" (Logan, 2008, p.100). Hence, the thorny issue of constructing a reliable investment technique that could capitalize on versatile market conditions imposes the necessity to thoroughly investigate pairs trading as a special category of short-horizon contrarian strategies. The pairs trading technique is genuinely used to take advantage of constant shifts in the market setting and benefit from explicit violations of the law of one price (Do and Faff, 2012). More specifically, pairs trading exploits mean reversion which is "a transient but recurring phenomenon" (Hillebrand, 2003, p.1). As a consequence, pairs trading has drawn the attention of the academic field as an opportunity to build a trading approach which overcomes some of the shortcomings inherent to the ordinary investment strategies.

This paper aims to investigate whether this state-of-the-art investment technique is profitable on the Swedish stock market on a risk-adjusted basis. A thorough investigation of this particular equity market is not only essential due to its relatively large size in Scandinavia and inherent efficiency in terms of processing and pricing relevant market information, but also because pairs trading based on the cointegration method has not been extensively explored in the academic work yet. Another theoretical motivation is reflected by the fact that albeit some positive feedback from the academic field, many studies fail to include transaction costs and thorough risk considerations when applying pairs trading. On the practical side, most studies only concentrate on a single market thereby implicitly questioning the robustness of their results. This study attempts to fill the theoretical void by performing a solid backtest using data ranging from 2004 until 2014 including transaction costs and a broad array of sophisticated risk measures. Additionally, this pairs trading strategy is evaluated through several stress-testing periods encompassing major financial turmoils. Compared to previous research, this study shows uniqueness with respect to the choice of markets, the focus on market scenarios containing crucial financial events and the development of a new criterion for the in-sample selection of pairs. The application of an extensive list of risk measures further contributes to the individuality of the paper.

Various kinds of pairs trading techniques have been applied since the mid-1980's by both hedge funds and investment banks worldwide (Muslumov et al., 2010) in order to gain profits from relative misvaluation of narrowly integrated financial instruments (Figuerola-Ferretti et al., 2014). Pairs trading can be generally associated with the term statistical arbitrage and their relationship is often described as ancestor-successor. This imposes the necessity to clarify exactly what is meant by statistical arbitrage which includes a diverse set of possibilities of how to implement those kind of investment strategies. Their mutual specifications are: (i) trades are performed systematically on a rule-driven basis in contrast to investment strategies relying on fundamental analysis; (ii) the trading portfolio ideally maintains market-neutrality with a zero beta against the market, and (iii) abnormal returns are accumulated via the implementation of a purely statistical techniques (Avellaneda and Lee, 2008). In this paper the terms statistical arbitrage and pairs trading are used interchangeably.

This holistic view shows the need to be explicit about exactly what is meant by the concept of 'pairs trading'. The initial idea of this particular strategy is found upon the fact that assets whose trading is influenced by mutual economic forces can be used to undertake trades according to any interim price divergence from their long-term equilibrium state. The nature of a pairs trading strategy is highly determined by the way of opening two contrary positions in the market - hence, a long position is initiated on the undervalued stock in line with an offsetting short position on the overvalued equity instrument (Lin et al., 2006). However, a pairs trading strategy involves substantial risks such as a sudden deterioration of the statistical connection between two interrelated stocks causing the price series to diverge instead of reverting back to their long-term mean. This fact should not be neglected as the bailout of formerly hedge fund Long Term Capital Management (LTCM) in 1998 shows (Andrade et al., 2005).

In order to scrutinize the stability of pairs trading as a viable and lucrative investment strategy, arbitrageurs should be well versed with the central factors supplementing this statistical arbitrage technique. However, academics demonstrate inconsistency in the way they carry out their studies. It is striking that most of the studies differ from each other by the methods they use, data frequencies, signal generations or trade initiation and execution. This incompleteness concerning the domain of pairs trading has imposed the necessity to undertake a research objective aiming to procure an intuitive understanding of pairs trading profitability. The increasing number of fat-tail events shortens the interval between financial crisis more and more. This background truly justifies the question "Can statistical arbitrage thrive on the Swedish stock market, thus providing excess returns?" Accordingly, the OMX30 Stockholm index is considered to be an appropriate benchmark to measure abnormal returns. Hence, this study attempts to answer this question by evaluating the following hypothesis:

H_1 : *equity pairs trading generates excess returns in the Swedish stock market*

Despite how a pairs trading strategy performs, it's inherent risks shall not be neglected. Therefore, this imposes the necessity to also evaluate the following hypothesis:

H_2 : *equity pairs trading is exposed to less risk than the benchmark*

The hypotheses above are used as reference to carry out the given study. The approach of this paper is closely linked to the methodology of Caldeira and Moura (2011) who apply the cointegration method to select tradable pairs according to their in-sample Sharpe-Ratios.

The key results in this paper are as follows: for all of the observed equity markets, it can be concluded that pairs trading is overall profitable, even when transaction cost are taken into account. From 2004 to 2014, pairs trading on the Swedish equity market shows a sound annual return of 8.18% with a fairly low volatility of 5.25%, thus providing a Sharpe-Ratio of 1.2. The generation of market-neutral returns is confirmed by applying pairs trading on different market scenarios. Although this results differ in their magnitude, they still collectively confirm the profitability of this trading strategy. The robustness of the empirical findings are supported by additionally evaluating the strategy on the German Stock Index (DAX) and the European Stock Index (EUROSTOXX). The profitability of pairs trading on te DAX varies slightly from the results of the Swedish equity market, providing a lower Sharpe-Ratio of only 0.96. However, during the observed time-period the strategy still yields a return 7.33% with a volatility of 5.62% on an annual basis. In contrast, for the EUROSTOXX the picture is more clear cut. On average, on the European market pairs trading provided considerably high annual returns of 12.19% with a significantly low risk of 5.84% thus yielding a Sharpe-Ratio of 1.75. In all three equity markets the strategy outcompetes the equivalent benchmark investments.

This paper proceeds in five sections covering the main features of pairs trading. After a short introduction outlining the general aspects of the subject, section 2 provides a detailed review of recent studies on pairs trading across different markets and asset classes. Section 3 presents the methodology and data. Section 4 outlines the empirical results and their analysis. The final section concludes with a detailed summary of the main empirical findings and their most appropriate interpretation.

Chapter 2: Literature Review

The purpose of this chapter is to review the literature on different pairs trading approaches. It begins by exploring the broad array of methods used for implementation purposes. This is followed by a detailed evaluation of cardinal studies assessing strategy profitability in disparate market settings and broad asset classes.

The elusive domain of pairs trading as a "short-term speculation strategy" (Gatev et al., 2006, p.797) has attracted the broad interest of the academic field. Despite this fact, the versatile investment technique is still considered a relatively new concept both in public debate and the academic research. In the recent literature, a consensus has been reached regarding the possibility of exploiting perseverative inefficiencies in financial markets which reflects the fundamental idea of statistical arbitrage (Perlin, 2009). However, several scholars approach and shed light on the topic of pairs trading from different angles in order to evaluate the profitability and applicability of the investment strategy. A comprehensive overview of the most renowned research articles investigating the concept of pairs trading in combination with their research coverage, implementation methodology and empirical findings is presented in Appendix A.

The several ways of how to construct a unique set of rules for training and consequent pairs trading contributes to the discrepancy in the literature covering the topic of statistical arbitrage. The clear variation in implementation methodologies is reduced by a common trait inherent to the vast majority of studies, namely the idea of an existing long-term equilibrium between the relative pricing of two financial time-series. As a consequence of the above statement, assets with similar characteristics must generate identical returns. The scholars are unanimous that significant diversions from the equilibrium steady state are triggered by market under- or over-reactions. Thus, the pairs trading technique relies on the assumption that "these deviations are temporary and will be corrected over time" (Bogomolov, 2013, p.1411).

In general, the implementation of a pairs trading strategy can be build upon three main approaches, namely the distance approach, the stochastic spread technique as well as the cointegration method. All of the aforementioned procedures share a similar internal framework: the composition of pairs relies upon ex-post analysis of historical data. The latter is complemented with various specifications governing the threshold levels triggering signal initiation, execution and consequent closure, which are in turn contingent on the development of the spread process (Bogomolov, 2010).

An important distinctive characteristic inherent to the distance method concerns the formation of pairs based on the minimum-distance principle. The foregoing criterion is build upon a two-stage procedure starting with the implementation of a cumulative total returns index for each security individually and consequently tracing a matching partner for each single stock. The unique pairs combination can thus be obtained by "finding the security that minimizes the sum of squared deviations between the two normalized price series" (Gatev et al., 2006, p.803).

When the focus is switched towards the stochastic spread method, Elliot et al. (2005) propose to model the spread between two assets as a mean-reverting process (Gaussian Markov chain) based upon market observations. By comparing actual market values and predictions of the spread, long- and short-positions are initiated or unwind respectively. However, as a consequence of the frequent trading signals generated by the model, the stochastic spread method is subject to increased transaction costs which may erode the accumulated excess returns. Despite this apparent disadvantage with respect to the other two methods, it can be noted that the stochastic spread technique is the only procedure that has a defined exit-strategy identified by certain thresholds marks. This is an important characteristic as the remaining approaches neglect the possibility of increased price divergence. This deviation from the equilibrium relative stock price relationship may cause consecutive losses (Bogomolov, 2010). Nevertheless, it is worthwhile to mention that although the stochastic spread method gained some attention in the academic area (Tianyong et al., 2013), it still remains a fairly undeveloped field with insignificant number of backtests and performance evaluation results.

Another strand of the literature domain represented by Caldeira and Moura (2011), Vidyamurthy (2004) and Alexander (1999) has emphasized on the cointegration approach. The authors endeavour to exploit a cointegrating relationship between two assets in order to explore the performance of a pairs trading strategy. To accomplish this, different methods can be recognized. Although studies such as Caldeira and Moura (2011) and Dunis et al. (2010) mainly focus on the Engle-Granger 2-Step approach to model and test for a cointegrating relationship, others like Huck and Afawubo (2014) and Dunis and Ho (2005) instead provide the Johansen's test in combination with a vector error correction model (VECM) as a further alternative. Nevertheless, both methods testify whether a long-term steady relationship between two time-series can be recognized. Do et al. (2006) and Lim and Martin (1995) claim that despite its merits, the Engle-Granger 2 step procedure exhibits two cardinal shortfalls. Firstly, the residuals from the regression analysis may be subject to varying set of statistical properties due to the result's sensitivity to the correct sequence of variables. Secondly, the generation of spurious regressions may be a consequence of the lack of cointegration between two series. Puspaningrum et al. (2011) state that in order to overcome the impediments inherent to the Engle-Granger approach, the Johansen's procedure applies vector error correction model that allows all the variables used in the regression specification to be endogenous. Haque and Haque (2014) stress upon this fact as the residuals generated from the estimated equation play a crucial role in pursuing a pairs trading strategy. However, due to its practical applicability, the Engle-Granger 2-Step approach remains the preferred model in the literature. Alexander (1999) justifies the above statement by its straight-forward execution and superiority of the minimum variance criterion over the maximum stationarity measure employed by Johansen in risk management considerations.

It is worthwhile to mention that the existing literature exhibit common ground in the case related to the separation of the sample into a formation and a trading period. Hence, the tractability of the strategy is contingent upon the historical behavior of the data series at hand. The training period can further be specified as those observations that determine the selection of pairs based upon the different approaches of pairs trading. Similarly, the testing period formation rules vary according to the research objective and the empirical model specification. Mori and Ziobrowski (2011) posit that it may be arduous to measure to what extent accumulated profits can be explained by the actual mean reversion of prices or the bid-ask-bounce effect respectively. To take the latter into account, they follow a trading rule proposed by Gatev et al. (2006) that consists of initiating trading positions one day following the divergence.

As a result of the uncertainty involved as to which position (long or short) will generate higher excess returns, Caldeira and Moura (2011) neglect the rebalancing of positions once the portfolio is formed. Hence, even if the market-neutrality condition is violated incited by sharp price movements, rebalancing is not performed. It can be observed that the majority of literature follows the methodology of Perlin (2009) who calculates the strategy's returns contingent on their individual weight in the simulated portfolio. Hence, this approach assigns each trading position equal exposure. Although the time-intervals covering the formation and trading period may be arbitrarily selected, the majority of previous research studies have shown a certain consistency by choosing a 12 month formation period which ceases at the commencement of the trading time-frame (Pizzutilo, 2013).

The calculation of portfolio excess returns for the pairs trading investment strategy is conceptualized with the implementation of committed and fully invested capital metrics. The former scales the returns using the portfolio pairs formed in the training period, whereas the latter divides returns to all the pairs engaged in a position during the formation period (Bowen et al., 2014). Furthermore, it is evident that most of the scholars selected daily equity closing prices as the most suitable data for the evaluation of a pairs trading strategy due to their relative frequent intervals and hence greater observation coverage.

In general, the academic field agrees upon the fact that in order to comprehensively evaluate a trading strategy and thus build a robust system, the models should be backtested encompassing periods of different market regimes on different markets. Muslumov et al. (2010) extend this notion by applying a rolling window technique which shifts the pairs construction and testing periods by a period of one month until the entire sample period is covered. By correctly implementing the procedure, a comprehensive vector of overlapping abnormal returns can be obtained. In contrary, attention on special sub-periods might bias the results but on the other hand may act as a stress-test revealing how a certain trading strategy is performing during crisis times.

Gatev et al. (2006) perform an unrestricted stock pairs selection (disregarding sector dependencies) and allow for a one-day delay in the execution and closure of the trading positions. In the period of 1962 through December 2002, they document pairs trading to be profitable on the US equity market with daily excess return of around 0.75% and 0.89% for the top 5 and top 20 portfolio of pairs applying the minimum historical distance metric. Furthermore, the authors report Sharpe-Ratios in the range of 0.39 to 0.59 respectively. Do and Faff (2012) follow the same methodology employed by Gatev et al. (2006) and confirm their results on the US equity market using a sample ranging from 1963 until 2009. At second sight, these results contrast with the findings of Bogomolov (2010) who argue that the cointegration, distance and stochastic spread approach yield valid abnormal returns ranging from 5% up to 12% per annum for the Australian equity market. Nevertheless, when transaction costs are taken into consideration the accumulated results are significantly diminished away, driving two of the methods into unprofitable range while leaving one technique with minimal proceeds.

Every investment strategy is subject to a certain kind of transaction costs. In the case of pairs trading, transaction costs may increase due to the initiation of contrary positions and their consequent closure upon their reversal. It can be identified that every market bears different trading costs, varying in their magnitude. Furthermore, transactions induce a fairly specific cost, as every bank or retail investor is obliged to a different pricing structure on the short- and long side of their positions. An explicit consideration of the transaction outlays diminishing the excess returns is apparent in the research conducted by Caldeira and Moura (2011), Gatev et al. (2006) and Broussard and Vaihekoski (2012).

The range of transaction cost varies from 100 BP to 165 BP per round trip and pair respectively. The implementation of transaction costs as factor eroding the profitability of pairs trading undertaken by Deaves et al. (2013) has been confirmed by other authors such as Caldeira and Moura (2011). Irrespectively of transaction costs they found pairs trading to be profitable with an annual return of around 16% accompanied with a annual volatility of about 12% yielding a annual Sharpe-Ratio of 1.3. The authors complement the above metrics with the inclusion of Maximum Drawdown statistic, representing a maximum loss of around 24%. Andrade et al. (2005) show similar results observing the Taiwanese equity market as they report annually abnormal returns of 10% with a standard deviation of around 9%. In contrast to the previously mentioned studies, Deaves et al. (2013) conduct a study evaluating the Canadian equity market based upon monthly abnormal returns. Although the discussed paper shows an inconsistent frequency of returns, by scaling the results it can be reported that the top 5 portfolio yielded around 16% per annum on committed capital with similar risk of 16%.

Tianyong et al. (2013) found notable differences for the Chinese equity market in comparison to the previously mentioned studies across the globe. They report a significantly low annual return of almost 3.5% while the strategy is exposed to an annual volatility of 0.378% yielding a fairly decayed Sharp-Ratio of around 1.1. In contrast to previous studies, it is worthwhile to note that they confirm the market-neutrality of the strategy by a Beta coefficient of 0.03 in line with the results of Perlin (2009). Despite the fact that it is claimed by the majority of scholars that pairs trading is found upon a market-neutral concept and thus plays a crucial role, these studies fail to justify their hypothesis. As an exception, Perlin (2009) performs a detailed regression analysis regarding the market sensitivity represented by the respective beta coefficient and the abnormal returns measured through the Jensen's alpha.

A comprehensive assessment of pairs trading as a contrarian strategy involves the comparison of disparate asset classes in order to correctly evaluate which market segment generates highest excess returns. Mori and Ziobrowski (2011) document a decrease in strategy risk due to the use of REITs in adverse market conditions as compared to common stocks. This results in augmented profit levels. In conjunction with their research findings, Alsayed and McGroarty (2012) state that the implementation of a pairs trading strategy with American Depository Receipts results in a low volatility of strategy returns as well as a high frequency of small quickly-decaying mispricings which can be exploited. Bianchi et al. (2009) reports statistically significant abnormal returns as a result of the implementation of pairs trading techniques on commodity futures. Furthermore, their findings suggest that those excess returns are not contingent on ordinary market risk factors and hence are not interrelated with the classic contrarian investing style. Marshall et al. (2013) state ETFs as another major asset class which should be considered due to a large temporary dispersion during the Flash Crash of 2010. This gave raise to an increased statistical arbitrage in those markets during that time. By a variation of the asset class in combination with a tighter time-frame Nath (2003) document the applicability of pairs trading on intraday high-frequency data covering the US Treasury securities. Their study also documents positive excess returns in the observed market.

Analyzing the literature domain investigating pairs trading, the innovative contribution that every study provides can be acknowledged. However, in order to compare the different results every single drawback of each contribution has to be weighted up. Accordingly, Perlin (2009) summarizes that perhaps the most serious weakness of pairs trading-related research is generally reflected by the fact that the majority of the above frameworks fail to address liquidity risk which is inherent in every trading strategy. This may have a negative impact on realized profit and loss of the respective trading strategy. Hence, as slippage follows a random process, any deviations from the optimal price cannot be quantified. Notwithstanding slippage, Broussard and Vaihekoski (2012) provide a study of the Finnish equity market which is considered by the authors as an illiquid market but still document its profitability.

Panyagometh (2013) documents that a pairs trading technique supplemented to the traditional value investing will provide the investor with an edge to enhance portfolio returns. A cross-reference to other similar literature demonstrates that the profitability of the investment strategy is significantly affected by firm-specific news and internal accounting practices. According to Papadakis and Wysocky (2007), the yield accumulated by the pairs trading technique is vastly eroded if positions are initiated in a short time-frame immediately after the occurrence of firm-related accounting event as compared to pair positions triggered in non-event periods. Moreover, sizable abnormal returns are generated by the introduction of lag when positions are unwind. The delayed closure aims at reducing the accounting information impact, thus signalling that the pairs trading performance is intensified in periods when no significant information is revealed to investors. As a continuation of the above statement and in conjunction with previous research, market participants have the tendency to under-react to information events related to changes in dividend policies, earnings estimates and analyst recommendation revisions. Hence, the incremental returns obtained from the implementation of the technical arbitrage procedure are significantly affected by actions developed on a firm level (Papadakis and Wysocky, 2007).

In terms of risk, it can be observed that most of the considered studies lack proper risk management techniques, hence those market positions remain fairly unprotected against idiosyncratic shocks causing further divergence losses. Caldeira and Moura (2011) develop a rudimentary stop-loss threshold to cut off losses from evolving even further. Thus, an open position is closed whenever the 7% loss mark is reached.

Perhaps the most serious disadvantage of the risk management approach undertaken by Caldeira and Moura (2011) is that they fail to incorporate additional margin requirements and borrowing constraints. Nonetheless, Pizzutilo (2013) shows that there is also the possibility to implement a risk management system by adapting the model itself via the usage of various threshold levels which could act as a loss protection.

The empirical outcomes developed in this paper demonstrate an evident disparity between strategy performance generated on the Swedish equity market and pairs trading profitability as documented in previous academic research. With an annual return level of 8.18% and corresponding volatility of 5.25%, the strategy underperforms the results accumulated by Caldeira and Moura (2011). The authors cover the Brazilian equity market and document an average annualized return of 16.39% accompanied by an annualized volatility comprising 12.4%. The Sharpe-Ratio of 1.34 confirms the superiority of the strategy when performed on the Brazilian market even after adjusting for trading costs. The pattern is reversed when pairs trading is compared to the statistical arbitrage technique developed by Bogomolov (2010) and Tianyong et al. (2013). The respective annual return figures range from 0.5% for the former study to 3.44% for the latter. Comparing Sharpe-Ratios, it is evident that a pairs trading technique built on the Swedish equity market (1.20) outcompetes similar strategies performed on the Chinese (1.09) and Australian markets (0.40). When the analysis is performed encompassing different market segments, the robustness of the results is reduced. Alsayed and McGroarty (2012) investigate strategy performance on the UK stock market in conjunction with American Depository Receipts. The authors document an annualized return of 0.8% which represents a reduction of 8.10 percentage points when contrasted to the return level obtained by performing a long-term backtest on the Swedish equity market.

This section has revealed the key aspects of the disparate pairs trading implementations discussed in literature and the methods defining them. The holistic perspective on the literature review aims to investigate the uniqueness and inherent merits of all the approaches. This is accomplished without trying to define any superiority among the methods. In conclusion, it is evident that academic scholars clearly prefer the distance approach due to its straightforward application. Hence, the cointegration approach remains a fairly unexplored field of interest that justifies further investigation.

Chapter 3: Methodology

3.1 Data description and sample selection

The overall sample period encompasses the time frame beginning in *01.01.1999* and ending in *31.12.2014* thus consisting a total number of *4172* observations. The sample data is used to explore whether profitability of the cointegration pairs trading technique persists even in periods of global financial turmoils. Then, smaller sub-samples are formed in order to provide a comprehensive evaluation regarding the performance of the proposed pairs trading strategy during crisis times. Distinct economic events with historical significance have been used to select meaningful sample periods in order to backtest the proposed pairs trading strategy. Hence, the sample selection includes the new economy bubble in the early 2000's, the major financial crisis at its peak in 2008 as well as the recent European credit crisis starting in 2010.

In order to carry out an appropriate analysis of the most influential market events, the aggregated data is subdivided in an *"in-sample"* and *"out-of-sample"* testing periods. Accordingly, in-sample data is used to determine unknown variables such as pairs to trade within the trading strategy, whereas out-of-sample data is utilized to backtest and evaluate the pairs trading strategy onwards. As a consequence, three different scenarios have been generated using the sub-divided sample data:

- Scenario 1 refers to the new economy bubble, consisting an in-sample selection ranging from *01.01.1998* until *31.12.1999* and an out-of-sample period spanning from *03.01.2000* until *31.12.2001*.
- Scenario 2 refers to the most recent financial crisis, encompassing an in-sample period beginning in *01.03.2005* until *01.03.2007* and an out-of-sample time-frame covering the period from *02.03.2007* until *02.03.2009*.

- Scenario 3 refers to ongoing European credit crisis, including an in-sample period ranging from *03.03.2009* until *31.01.2011* and an out-of-sample period consisting of observations ranging from *01.02.2011* until *31.01.2013*.

The pairs trading strategy is analyzed based on daily close-prices of the Swedish stock market index (OMX30) and its equity constituents respectively. Due to data availability only 29 stocks were used for trading. The index itself acts as a benchmark to evaluate the trading performance and risk in an appropriate way. Furthermore, the middle rate of Sweden's 90-day Treasury Bill has been obtained as a relevant riskless proxy. To extend the analysis, different European equity markets are additionally considered: German stock market index (DAX30) and European stock market index (EUROSTOXX50). The latter encompasses the performance of the 50 most highly valued companies in terms of market capitalization in 12 Eurozone countries. Due to limited data availability, only 29 stocks are considered as constituents of the DAX30 index. The index itself acts as a benchmark to evaluate the trading performance and risk in an appropriate way. Accordingly, the three month depository middle rate of Germany and the 3 month offered rate of the EURIBOR is used as their corresponding riskless proxies. Hence, the risk-free rates are used to calculate all the performance evaluation metrics such as Sharpe- and Treynor-Ratios. To be congruent, the time-frame of the risk-free rate should exhibit the same sample size as the backtested return series of the observed equity prices. Consequently, annual figures were scaled for daily calculations. All relevant data can be accessed by either Bloomberg or Datastream to ensure traceability.

3.2 Cointegration Approach in Pairs Trading

The model implemented in this paper is constructed on the basis of cointegrated prices rather than correlation between asset returns. According to Alexander et al. (2002), cointegration is induced by tied asset prices that are co-moving in the same direction over the long-run, thus sharing a common stochastic trend. The evidence revealed by the authors suggests that portfolio optimization built upon the cointegration technique is not exposed to the instabilities inherent to the correlation method over the long-term. Due to the fact that correlation builds upon returns rather than raw prices, this measure is considered to be short-term and unstable thus exhibiting "short memory processes" (Alexander et al., 2002, p.6). Hence, it can be argued that pairs trading methodology based on cointegrated financial instruments should demonstrate signs of increased effectiveness in the distant horizon.

The term statistical arbitrage is commonly defined as an effort to trade short-term deviations from the long-term market equilibrium (Jacobsen, 2008). In this paper, the selection of pairs to be traded depends upon whether a cointegrating relationship can be revealed between two assets. This study is built upon the methodology presented in Caldeira and Moura (2011). The authors posit that the presence of cointegration provides the opportunity to construct a portfolio consisting of a linear combination of two stocks in order to transform two non-stationary processes into a stationary one. This link is used by cointegration tests to identify a mutual long-term trend between two time-series. Any short-term deviations from a long-term steady state may be exploited to profit from market misperceptions through opening a long-position in the undervalued stock (buy) and engaging in a short-position in the overvalued stock (sell). As a result, the selection of an appropriate pair of stocks plays a crucial role in trading mean-reversion through a market-neutral trading strategy.

As Caldeira and Moura (2011) suggest, the pairs trading algorithm used in this study is basically twofold. In the first part of the elaborate methodology, all possible combinations of pairs within the original pool of assets are tested for cointegration. During this process, pairs are identified based on which linear combinations show a solid forecast component that exhibits no correlation with the overall market movement. In the second part, those pairs that have already been identified during the formation period are used to form an equally weighted portfolio of 10 pairs. The portfolio consisting of 20 stocks is then traded in the out-of-sample period to obtain all the necessary performance- and risk-metrics. Figure 3.1 provides an exemplary overview of the algorithm used in this study.

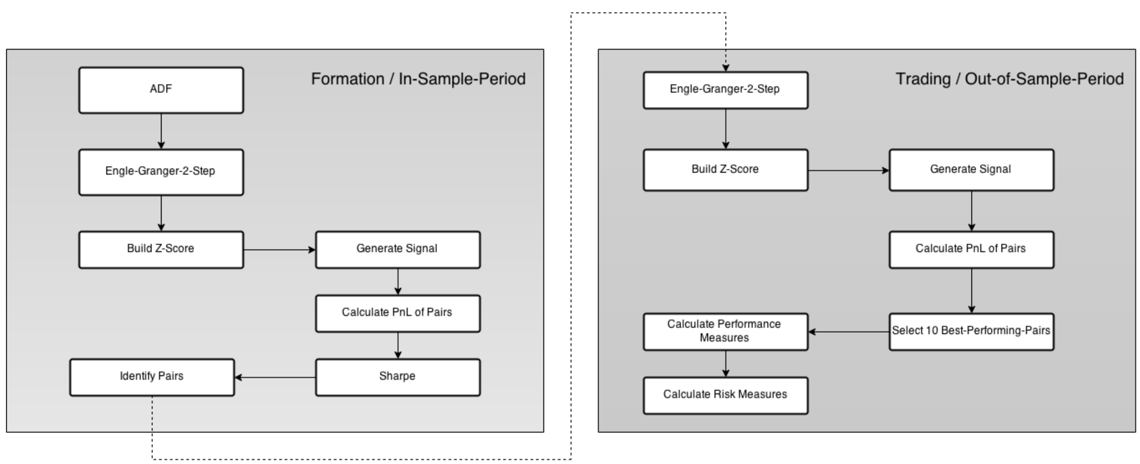

Figure 3.1: Overview of Pairs Trading Algorithm

3.2.1 Pairs Formation and Trading Process

The backtesting of the trading strategy requires the sub-division of the overall time-frame in two distinct intervals, namely the formation and trading periods. In this paper, these concepts are interchangeably associated with the terms in-sample and out-of-sample periods. Within the in-sample period, the algorithm handles the selection of tradable pairs on the basis of several consecutive filter mechanisms:

1. The Augmented Dickey-Fuller test is performed on the whole sample as a prerequisite test used to confirm that the time-series at hand contains an unit root.
2. The cointegration test is performed on all possible variations of those assets which have already passed the unit root test, acting as a second filter.
3. Afterwards, the residuals generated from the assets which exhibit a cointegrating relationship are used in order to build a Z-score for each individually.
4. The implemented Z-score serves as an indicator to generate a trading signal, which is used to backtest the performance of the strategy in the form of profit and loss. This is the basis for the calculation of their respective Sharpe-Ratios.
5. As a result, candidates of pairs used for future trading can be identified by their best performing Sharpe-Ratios or an alternative performance tracking metric.

In the out-of-sample period, the pairs that have been selected as suitable trading candidates are engaged in the same algorithmic procedure as in the formation process. In contrast to the previously described in-sample period, the real strategy performance is evaluated on the basis of the decoupled sub-sample which follows the formation period. Hence, the trading period lays the foundation for further assessment of performance and risk levels.

In order to fully encompass different business cycles, the research conducted in this paper implements backtests of different market scenarios. Moreover, to gain an insight of the long-term performance of the proposed trading strategy, a rolling window procedure is employed. This method continuously shifts the formation- and trading-period forward in time with a constant magnitude. Thus, a period which have been used for out-of-sample testing is consequently handled as a in-sample-period. Despite the constant shift of the sample intervals, the length of the training- and testing periods remains unchanged.

At variance with the procedure undertaken by Gatev et al. (2006), who implement a formation period of 6 months and a trading period of 12 months, this paper employs an in-sample window of 500 trading days and an out-of-sample window consisting of 250 trading days. Figure 3.2 depicts the rolling window backtest procedure. In addition to the pairs-selection method proposed by Rudy et al. (2010) and Caldeira and Moura (2011), this paper investigates further alternatives instead of solely making use of the Sharpe-Ratio. Hence, in this paper the in-sample simulations for the formation of portfolios are based on additional measures such as the MAR, Treynor- and Sortino-Ratio.

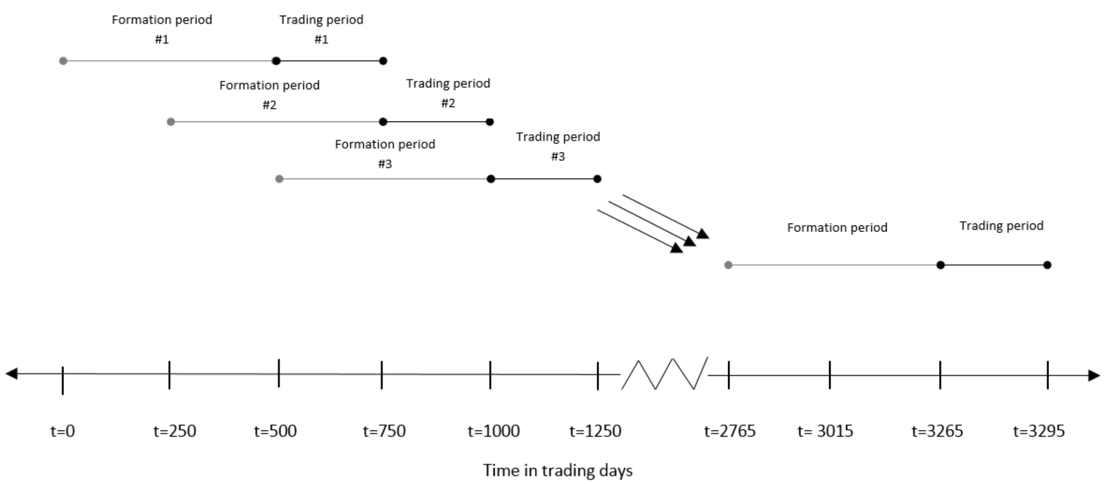

Figure 3.2: Overview of Rolling Window Procedure adapted from Andrade et al. (2005)

3.2.2 Unit Root and Stationarity

The existing literature highlights the importance of differentiating between variables that are either stationary or non-stationary and thereby justifying their most appropriate treatment (Brooks, 2008). In the context of statistical arbitrage, the correct identification of the unit root property is an essential precondition in order to verify the cointegration regression and consequently select a pair of stocks with similar historical dynamics (Miao, 2014).

In general, a time series process is considered stationary if the first two moment conditions and autocovariances are constant. Moreover, if the series demonstrates independence of time regarding mean and variance, rather than the entire distribution, the process is observed as weakly, or covariance stationary (Veerbek, 2004). The above interpretation contrasts with the concept of non-stationarity which is associated with the existence of unit root and time-varying first two moment conditions. Kozhan (2010) argues that a variable can be identified to follow a non-stationary process by either pursuing a visual analysis of the time-series and its correlogram or a quantitative analysis of formal statistical tests. In this study, the Augmented Dickey Fuller (ADF) test is therefore applied in order to verify the non-stationarity specification of each OMX constituent time series. This stage needs to be passed in order to proceed with a cointegration test at a later stage.

3.2.3 Augmented Dickey Fuller

The existing literature provides various methods of how to perform statistical tests such as Augmented-Dickey-Fuller, Philips and Peron and KPSS test in order to determine whether a financial time-series can be identified as a stationary or non-stationary one. Breitung (2002) further specifies that the original Dickey-Fuller test (Dickey and Fuller, 1979) follows a parametric model specification which uses an auto-regressive representation of the observed time series. In general, the initial Dickey-Fuller test builds upon the assumption that the disturbance terms follow a pure random process (white noise) and hence are uncorrelated. This basic underlying assumption, however, suffers from the fact that this is not absolutely the case in reality as there could be signs of serial correlation. The following Augmented Dickey-Fuller (ADF) test relaxes the predefined conjectures by taking some forms of serial-correlation into account (Greene, 2003)

In the academic literature, the Augmented Dickey-Fuller test appears as a prominent statistical test due to the fact that it is "asymptotically valid under much less restrictive assumptions" (Davidson and MacKinnon, 2003, p.610). The regression specification of the ADF-test is represented by the following:

$$\triangle y_t = \psi y_{t-1} + \sum_{p}^{i=1} \alpha \triangle y_{t-i} + u_t \tag{3.1}$$

where $\triangle y_t = y_t - y_{t-1}$

ψ is the coefficient representing unit root process

α is an intercept constant called a drift

u_t is an iid residual term

The ADF test incorporates n-additional lagged values of the dependent variable, namely the regressor $\triangle y_{t-n}$, in order to incorporate "any dynamic structure" (Brooks, 2008, p.329) inherent to the explained variable. The inclusion of additional explanatory variables enhances the regression robustness and ensures that the disturbance terms u_t are not auto-correlated anymore (Gujarati, 2004).

However, difficulties arise when an attempt is made to implement the ADF-test in practice since the optimal number of lags should be explicitly identified. These assumptions would seem to suggest that the inclusion of too many lags will augment the coefficient standard errors. On the other hand, the implementation of too few lags might not remove the serial auto-correlation issue, thereby affecting the regression results (Brooks, 2008).

Despite the inconsistency regarding the choice of an appropriate lag length, the academic literature still perceives the ADF-test as one of the most recognized methods when it comes to performing statistical arbitrage strategies. In the context of pairs trading, the described methods are applied to test whether the observed time-series is integrated of order one $I(1)$. This acts as a prerequisite to perform various cointegration tests which will be explored in detail in the next section (Caldeira and Moura, 2011).

3.2.4 Engle-Granger 2-Step Procedure

A thorough definition clarifying the concept of cointegration is provided by Engle and Granger (1991). The authors posit that if two time-series x_t and y_t are mutually integrated of order one $I(1)$, the series exhibit a cointegrating relationship under the condition that the linear combination of two processes $z_t = m + ax_t + by_t$ $I(0)$ is stationary. In this case, the two non-stationary series share a common stochastic trend. The economic intuition behind the concept of cointegration is mainly reflected by the analysis of whether a given pair of time-series exhibits any empirically meaningful relationship in order to model a long-run equilibrium (Veerbek, 2004).

The cointegration testing performed in this study is build upon a two-step procedure developed by Engle and Granger (1987). The two-step technique is considered by (Alexander, 1999) to be the favoured methodology among others due to (1) its straight-forward implementation, (2) advantages when it comes to risk management considerations and (3) practicability while choosing the dependent variable. Accordingly, the test is performed in the following way:

In the first stage, it is determined whether a long-run (equilibrium) relationship exists by observing the following OLS regression equation:

$$\mu + u_t = \left(P_t^A\right) - \gamma\left(P_t^B\right) \tag{3.2}$$

where P_t^A is price of stock A at time t
P_t^B is price of stock B at time t
γ is cointegration coefficient
μ is an intercept constant called a drift
u_t is an iid residual term

After a regression of the dependent variable P_t^A on the independent variable P_t^B is performed, the estimated regression residuals u_t are consecutively tested for stationarity through a properly conducted ADF-test (Wang, 2009). The respective null hypothesis H_0 states that the estimated residual series is non-stationary and hence allows to conclude that a given time-series pair is not cointegrated. If H_0 can be rejected though, the residual series is found to be stationary and thereby the observed time-series is in fact cointegrated (De Boef and Granato, 1999). In the second stage, the regression residuals from the preliminary step are included as an additional explanatory variable in the following error correction model:

$$\triangle P_t^A = \beta_0 + \beta_1 \triangle P_t^B + \beta_2 \hat{u}_{t-1} + \nu_t \tag{3.3}$$

where $\hat{u}_{t-1}$ contains residuals from 1st step
ν_t is the new disturbance term

Equation 3.3 shows that the error-correction term takes into account the distinct tendency of cointegrated variables to converge to a common stochastic trend. The error-correction model is characterized by regression coefficients for the effect and causal variable expressed in first-difference.

3.2.5 Z-score model

The portfolio formation and consequent trading initiation scrutinized in this paper rely heavily on advanced technical indicators. Following the testing period, the strategy implements a dimensionless Z-score used to generate trigger signals detecting abnormal relative price deviations between the two time series. The divergence is measured relative to a steady historical relation between the two assets.

The Z-score is defined by Pelletier (2007) as the distance, measured in terms of standards deviations, between the spread and its mean. From a statistical perspective, this indicator oscillates freely and can take on positive and negative values relative to its mean. The Z-score is obtained by subtracting the sample mean from the spread and consequently dividing the accumulated result by the sample standard deviation:

$$z_t = \frac{\varepsilon_t - \mu_\varepsilon}{\sigma_\varepsilon} \tag{3.4}$$

where z_t measures the distance to long-term mean
ε_t is the value of the price spread at time t
μ_ε is the 250-day mean of the out-of-sample price spread
σ_ε is the standard deviation

The value of the spread ε_t at time t is calculated as:

$$\varepsilon_t = P_t^L - \gamma P_t^S \tag{3.5}$$

where ε is the value of the spread at time t
γ is the cointegrating coefficient
P_t^L is the price series of long position at time t
P_t^S is the price series of short position at time t

3.2.6 Trading Rules and Trading System

The previously described Z-score model mainly represents an indicator function which takes on disparate values that can be used to generate trading signals. Different Z-score levels demand distinct trading actions. According to many papers in the field of pairs trading, the general rules which are set for opening and closing positions share certain common characteristics since the threshold levels used in these studies lay in an aggregate interval. In the undertaken research, basic trading rules are used following the ones proposed by Caldeira and Moura (2011). They specify trading rules when exactly to initiate and unwind long- and short positions respectively.

According to Figure 3.3, a position will be opened if the value of the Z-score breaches for the second time the level of 2 Standard deviations from above or from below. In this situation, a misvaluation of stocks is consequently detected by the Z-s core implying that the portfolio of pairs is overvalued. Hence, an investor should sell its portfolio short by selling Stock A and buying Stock B simultaneously. In the opposite case, if the Z-score

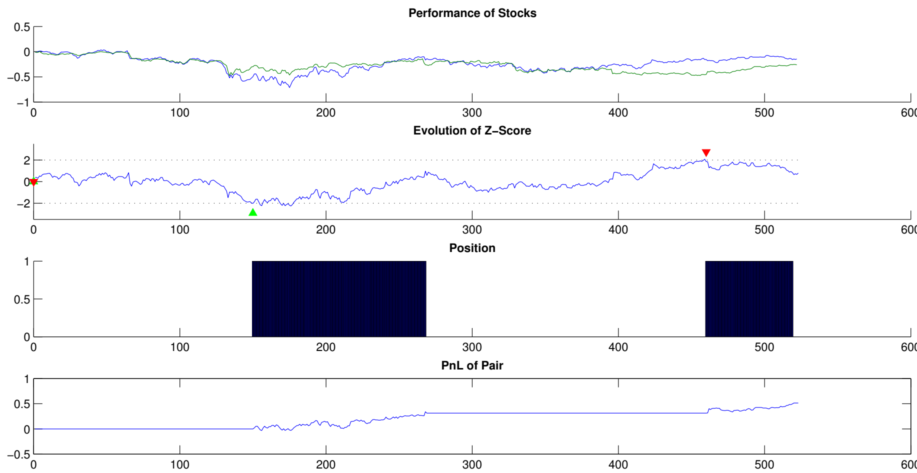

Figure 3.3: Overview of Trading Methodology

breaks the level of -2 Standard deviations, an investor's portfolio is below its long-run equilibrium value. In this case, stock A needs to be bought and stock B should be sold at the same time. In the same way how the long- and short-positions are opened, any open positions must be closed at some time. Hence, an existing short position will be closed if the Z-score falls below 0.75 standard deviations. Moreover, an active long position will be closed if the Z-score falls even further below a level of 0.5 standard deviations.
The described trading rules can be further specified in the following way:

Buy to Close if $\varepsilon_t < 2.00$
Sell to Close if $\varepsilon_t > 2.00$
Close short position if $\varepsilon_t < 0.75$
Close long position if $\varepsilon_t > 0.50$

The frequency of trade initiation and closure heavily depends on the chosen threshold level. If this trigger mark is placed too low, this will result in an increased number of open positions and hence augmented trading costs. In contrast, if the threshold standard deviation is set on a relative high position, the whole testing period might end without opening any trades. The above statements confirm the necessity proposed by Avellaneda and Lee (2008) to test and select the cutoff levels empirically. Despite the chosen threshold level, the algorithm developed in this paper closes all open positions at the end of the testing period even if this is executed at a loss. The assumption of different market scenarios implemented in the methodology requires that the distinct trading periods do not overlap and hence no positions can be left open after the end of a given testing interval.

3.3 Calculations of Pair Returns

In order to calculate the returns of a portfolio including a pair of assets, this paper mainly follows the approach of Caldeira and Moura (2011). The authors claim that instead of predicting the behavior of P_t^L and P_t^S, it is sufficient to only forecast the difference $ln\left(P_t^L\right) - ln\left(P_t^S\right)$. As a consequence, they derive the following equation:

$$r_t^P = ln\left(\frac{P_t^L}{P_{t-1}^L}\right) - \gamma ln\left(\frac{P_t^S}{P_{t-1}^S}\right) + 2ln\left(\frac{1-C}{1+C}\right) \tag{3.6}$$

where γ is the cointegration coefficient

C represents transaction cost in basis points

In equation 3.6, shortselling is taken into account by including a minus sign between the two return variables. A value of $\gamma = 1$ implies that the investor is able to benefit from the pairs trade. When γ approaches a value of zero though, the investor is forced to solely invest in the long position. In contrast to Caldeira and Moura (2011), trading positions in this research are hold until the end of out-of-sample period without engaging in any stop-losses.

3.4 Transaction cost

In order to correctly assess the performance of the pairs trading technique, the magnitude of the transaction charges should be taken into consideration. The research of this paper is mainly conducted on the Swedish equity market represented by the OMX30 index. Hence, the incorporated transaction costs should fit the Scandinavian market environment and its main asset class. Due to its generalized framework, the study of Thapa and Poshakwale (2010) is used as a cross-reference when deciding on the most appropriate level of trading costs. The authors document the pricing of trading activities for a variety of countries. When the focus is switched towards a specific country, it becomes apparent that the total amount of transactions costs is decomposed into commission, fees and slippage. Nevertheless, the authors fail to integrate the crucial rental cost for short-positions as revealed by Caldeira and Moura (2011) in their seminal study. For the Swedish market, Thapa and Poshakwale (2010) estimated a commission of 18.13 BP, fees of 0.47 BP and slippage cost of 12.39 BP as appropriate figures, leading to transaction cost per pairs trade consisting of 31 BP. When these results are complemented with an additional rental cost of 20 BP, as proposed by Caldeira and Moura (2011), the transaction figure comprises a total of approximately 50 BP one way for both trades. When the long-term backtest is extended to the additional stock markets, the transaction cost change slightly. The respective figure for the German market is 47.52 BP per round trip. The value for the European equity market is extrapolated to be the same due to the heavy influence of German constituents on the value of the index. The resulting figures are used as a basis for all the backtests conducted in this study.

3.5 Performance and Risk measures

A comprehensive evaluation of an existing trading strategy demands that the aspect of risk should not be neglected due to the fact that every rate of return is associated with a certain level of risk exposure. In order to identify whether the underlying strategy is profitable, the performance and risk of the statistical arbitrage strategy is evaluated against a predefined benchmark, namely the Swedish stock-market index OMX30. As a consequence, the performance of the proposed pairs trading strategy is assessed through several advanced risk measures such as Maximum Drawdown, Mean Absolute Deviation, Value at Risk and Conditional VaR at 95% confidence level, Sharpe-Ratio, MAR, Treynor- and Sortino-Ratio. The Sharpe-Ratio is computed by dividing the mean excess return over the corresponding standard deviation of the portfolio. The Treynor metric focuses instead on the portfolio beta in the denominator. The Sortino-Ratio includes only the downside deviation of the strategy returns. The MAR-Ratio is obtained by dividing the compound annual growth rate over the Maximum Drawdown of the in-sample period. The analysis and implementation of explicitly defined risk measures enable the comprehensive evaluation of the pairs trading strategy.

The profit and loss of every pair is calculated as a cumulative sum of both stock returns on the long- and short side contingent on the respective trading signal. During the testing period, only a small fraction of pairs is actually traded due to thorough design of the implemented filter algorithms. In this context, the profit and loss of the portfolio is obtained by evaluating the fully invested returns of open positions on the actual used capital. This contrasts with the method of calculating returns on the basis of committed capital. This procedure takes into account the opportunity cost that financial institutions inevitably bear if the strategy is not exposed to any market risk and hence no positions are initiated (Gatev et al., 2006).

An analysis of the main findings and of the principal issues and suggestions which have arisen in this discussion are provided in the next chapter.

Chapter 4: Empirical Results

The subsequent chapter presents the main empirical findings concerning the pairs trading methodology performed on the Swedish stock market. A long-term backtest provides a general resolution whether the trading strategy is profitable over a longer horizon or falls short of expectations. However, the most recent turmoils in the financial markets justified the need of evaluating the performance of the pairs trading strategy in different competitive market environments covered by the previously presented three scenarios. Furthermore, as a result of the discussion concerning how to improve the trading strategy, different selection criteria are identified and evaluated. Finally, in order to assess the robustness of the empirical findings obtained for the Swedish stock market, the strategy is constructed and consequently tested on different asset markets. The basic parameters of the pairs trading strategy are not altered throughout all the different backtests, ensuring that every result is obtained in a consistent manner.

4.1 Longterm Backtesting Results

A long-term backtest allows the comprehensive evaluation of the strengths and shortcomings that the strategy bears. As a result, an extensive backtesting window of 10 years is chosen. All simulated portfolios exhibit a starting value of 100.000 SEK at the beginning of the out-of-sample period. Moreover, the appropriate pairs for trading are determined by the best performing in-sample Sharpe-Ratios. The result of this extensive backtest is shown in Figure 4.1.

The graphical representation demonstrates that after a short period of underperformance, the strategy begins to outperform the benchmark with the onset of the global financial crisis. After reaching a portfolio-value peak during November 2007, the benchmark's exposure to an increased volatility following the inception of the financial downturn triggered the accumulation of substantial losses. Hence, adjustments in market risk levels increase the possibility of relative price divergence between different time-series which

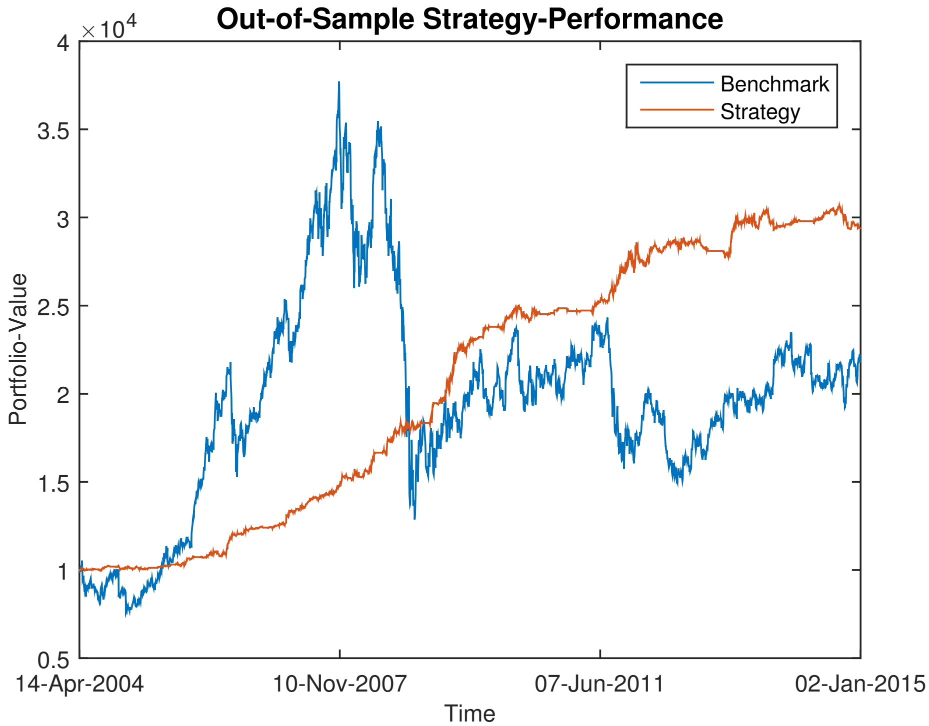

Figure 4.1: Longterm Backtest with in-sample Sharpe-Ratio as selection criteria, incl. transaction cost

can be exploited by trading a pair of assets. This market setting seems to provide an ideal environment for the strategy to increase its performance. The persistence of positive monthly strategy returns throughout the long-term testing period documented in Figure B.2 confirms the above statement. However, it is worthwhile to mention that the strategy and corresponding benchmark do not move in the same magnitude, leading to a clear underperformance of the strategy during the first part of the backtesting window which is considered a time interval of robust price appreciation and sustainable economic expansion. Nevertheless, independent of the huge drawdowns of the index, the equity line of the pairs trading strategy continuously sustains its upward movement and thus demonstrates its profitability even in the long term. The latter is confirmed by an annual return of 11.29% excluding transaction costs. After taking trading costs into account, the annual return diminishes to 8.18% respectively. This result confirms the market-neutral characteristic assigned to the investigated pairs trading strategy. The sensitivity of portfolio returns relative the overall market conditions remains relatively low throughout the extended backtest window as documented in Figure B.3 (presented in Appendix B).

In comparison, a buy-and-hold investment in the benchmark yields a much lower annual return of only 5.92%. Nevertheless, when it comes to out-of-sample risk levels, the impact of transaction costs remains fairly insignificant. From this perspective, the volatility of the trading strategy is substantially lower with around 5.25% as compared to the benchmark's risk of 28.05%. As the volatility rises slightly, the maximum drawdown of the strategy marginally extends as well. In this case, an investment in the trading strategy is exposed to a maximum drawdown of around 4.70%. Furthermore, the trading strategy exhibits a Sharpe-Ratio 1.20 whereas the benchmark's Sharpe-Ratio is 0.14. As a result, the tested trading strategy outperforms by far the benchmark on a risk-adjusted basis. In this case, transaction cost demonstrate a certain degree of impact, as the Sharpe-Ratio of the strategy (1.90) is much higher excluding transaction costs. In terms of risk management, the strategy provides an edge to investors when observing the Value-at-Risk and Conditional-Value-at-Risk figures. According to the former metric, an investor expects with a 95% confidence level that the loss incurred for the next day will not exceed 114.56 SEK. This contrasts with the fairly high Value-a-Risk for the benchmark index of 550.67 SEK. Since the Value-at-Risk metric does not provide evidence about the magnitude of the expected loss, another risk measure is introduced. In order to include the possibility of fat-tail events into account, the Conditional-Value-at-Risk measure is used. The value for this metric comprises 199.23 SEK with the inclusion of appropriate transaction costs. The preceding value is significantly smaller than the figure obtained by the OMX30 (972.54 SEK). As a consequence, an individual is better off by investing rather in the strategy than in the benchmark. The analysis of the third and fourth moments reveals further characteristics of the pairs trading strategy. A Skewness of 1.75 with the Sharpe-Ratio as a selection criterion is implying that instead of living off a small number or large gains, this pairs trading strategy relies upon an increased number of smaller profits. Hence, the strategy is exposed to a decreased risk of large losses but this is obtained at the expense of on average outperforming the expected PnL. Moreover, a high number of 22.70 for the Kurtosis indicates the presence of fat tails, implying higher probability of extreme events.

4.2 Scenario Results

In order to confirm the robustness of an already backtested trading strategy, it is often necessary to further investigate distinct sub-sample periods separately from their overall testing period. Thereby, a dense selection of different sub-intervals provides a compressed perspective on a particular market environment and its impact on the strategy. A comprehensive summary of the main in-sample results can be found in Table 4.1, accompanied by the corresponding out-of-sample statistics displayed in Table 4.2.

The following scenario results are solely carried out on the basis of the Sharpe-Ratio as an in-sample selection criterion:

Table 4.1: Results of in-sample backtest on constituents of OMX30 for different scenarios incl. transaction cost

	Scenario 1[a]		Scenario 2[b]		Scenario 3[c]	
Criterion: Sharpe	**Strategy**	**Benchmark**	**Strategy**	**Benchmark**	**Strategy**	**Benchmark**
CAGR	35.18%	41.53%	24.97%	58.98%	27.12%	15.03%
Return (p.a.)	28.93%	33.40%	21.35%	44.49%	24.04%	14.06%
Risk (p.a.)	11.82%	36.97%	5.74%	28.14%	6.32%	25.17%
Sharpe (p.a.)	2.14	0.81	3.30	1.50	3.75	0.55
Max.Drawdown	5.81%	41.24%	1.89%	29.88%	2.19%	19.95%
MAD	2409.98	1636.81	1493.02	4715.41	1306.64	693.88
VaR_{t+1}	181.93	651.54	76.26	610.29	59.92	323.00
$CVaR_{t+1}$	274.88	997.63	99.90	1075.49	98.86	452.12
Kurtosis	10.95	7.92	8.55	7.64	14.49	6.08
Skewness	1.25	0.00	1.11	-0.47	2.19	-0.08

[a] *represents data covering Technology bubble (01.01.1998 to 31.12.1999)*
[b] *represents data covering Global financial crisis (01.03.2005 to 01.03.2007)*
[c] *represents data covering European credit crisis (01.03.2009 to 31.01.2011)*

Table 4.2: Results of out-of-sample backtest on constituents of OMX30 for different scenarios incl. transaction cost

	Scenario 1[a]		Scenario 2[b]		Scenario 3[c]	
Criterion: Sharpe	**Strategy**	**Benchmark**	**Strategy**	**Benchmark**	**Strategy**	**Benchmark**
CAGR	18.29%	3.33%	17.85%	-14.73%	8.33%	-7.39%
Return (p.a.)	16.15%	3.16%	15.76%	-15.32%	7.67%	7.37%
Risk (p.a.)	11.54%	46.97%	5.46%	46.56%	3.97%	24.32%
Sharpe (p.a.)	1.06	-0.02	2.26	-0.40	1.60	-0.36
Max.Drawdown	6.60%	53.67%	2.56%	65.88%	1.96%	38.70%
MAD	1413.97	1677.64	1021.65	2253.83	463.55	960.62
VaR_{t+1}	131.39	554.29	65.14	349.49	38.04	214.66
$CVaR_{t+1}$	218.73	675.02	95.21	528.95	55.82	327.92
Kurtosis	23.40	4.63	6.08	14.49	7.67	4.80
Skewness	1.88	0.24	0.64	-1.25	0.98	-0.25

[a] *represents data covering Technology bubble (03.01.2000 to 31.12.2001)*
[b] *represents data covering Global financial crisis (02.03.2007 to 02.03.2009)*
[c] *represents data covering European credit crisis (01.02.2011 to 31.01.2013)*

In addition, several graphs (represented by the set of Figures from B.4 until B.12) further support the previously shown results. They can be found in the Appendix B.

Scenario 1 - Technology Bubble

The in-sample period of the first scenario represents data of the rally of the stock market until the end of 1999 towards the beginning of the technology bubble. Concerning the in-sample data-set, the strategy performs fairly well on a risk-adjusted basis yielding a Sharpe-Ratio of 2.14 outcompeting the benchmark which generates a Sharpe-Ratio of 0.81 as evidenced by looking at Table 4.1. In terms of risk, the strategy still exhibits a satisfactory performance with a volatility of 11.82% in contrast to the risk of the benchmark of 36.97%. It is evident that the increased annual risk inherent to the OMX30 index induces a downward push in the Sharpe-Ratio metric. As a contrast to the previous results, the trading strategy underperforms with a 10-year compound annual growth rate (CAGR) of 35.18% or an annual return of 28.93% its corresponding benchmark. The index yields a buy-and-hold return (CAGR) of 41.53% or an annual return of 33.40% respectively.

The out-of-sample period of the first scenario reflects the prevailing market conditions and investor sentiment after the burst of the technology bubble in the early 2000's. By looking at Table 4.2, it becomes apparent that the increased market volatility significantly affects the benchmark performance with returns dropping to a mere 3.16%. The corresponding figure for the strategy hovers around 16.15%. This acts as an additional confirmation of the fact that the plausibility of the investment strategy is increasing during crisis times. In terms of risk, the strategy also significantly outcompetes the benchmark with an annual volatility of 11.54% in contrast to a benchmark volatility of 46.97%. Hence, the annual risk that an investment in the strategy is exposed to, remains fairly constant across in-sample and out-of-sample periods. This is complemented by observing the strategy's maximum drawdown of only 6.60% in comparison to a much larger maximum loss of 53.67% of a buy-and-hold investment in the benchmark. However, from a risk-adjusted perspective, the strategy performs satisfactory with a Sharp-Ratio of 1.06 in contrast to a significantly small and negative Sharp-Ratio of -0.02 for a benchmark investment. As a result, the strategy demonstrates a fairly smooth development of its profit and loss equity line even during times of an economic downturn caused by the technology bubble. An illustrative graphical representation of in- and out-of-sample strategy performance complemented by monthly strategy returns for Scenario 1 is represented by the set of Figures from B.4 until B.6 found in Appendix B.

Scenario 2 - Global Financial Crisis

The in-sample period of the second scenario encompasses the run up towards the most recent global financial crisis. With an annual return of 44.49%, the benchmark clearly outperforms the investigated strategy which only attained an yearly return of 21.35%. The above results could be attributed to the prevailing confidence among investors about the expansionary state of the business cycle and the continuation of the upward trend witnessed in the years following the arrival of the new millennium. Nevertheless, when the focus is switched towards the risk metrics, the advantage of the strategy becomes evident. An investment in the strategy is exposed to much less risk (5.74%) as compared to an equivalent buy-and-hold investment of the benchmark with a annual volatility of 28.14%. This is confirmed by looking at a superior in-sample Sharpe-Ratio of around 3.30 in comparison to the benchmark's Sharpe-Ratio of merely 1.50. Furthermore, it is evident that the strategy exhibits a considerably small maximum loss of only 1.89% in comparison to its benchmark of around 28.14%.

The out-of-sample time-frame is built upon more severe market conditions. The inception of the global financial meltdown may has triggered a divergence between the strategy results and the OMX30 composite index development. An investment in the benchmark lost around -15.32% during the out-sample period, whereas the strategy performs exceptionally well yielding a positive annual return of 15.76%. As a result of the financial downtrend, the benchmark investors exhibit a considerably high volatility of around 46.46% and a fairly large maximum loss of 65.88%. Nevertheless, the strategy exposes the investors only to an annual volatility of around 5.46% and a maximum drawdown of around 2.56% at the time. Hence, the strategy performance remains profitable in an extreme market environment such as the aftermath of the global financial crisis. This is confirmed by the predominant number of positive monthly strategy returns for all OMX constituents over the Scenario 2 out-of-sample period as exhibited by Figure B.9 (found in Appendix B).

Scenario 3 - European Credit Crisis

The in-sample period of the third scenario relates to the sluggish recovery and rebound from the depths of the financial recession. The slowdown of the economic activity strains the in-sample performance of the benchmark, leading to an annual return of 14.06% accompanied by a risk of 25.17%. In contrast, an investment in the strategy yields a significantly larger return of 24.04% and provides a fairly reduced amount of risk of around 6.32%. This figures are confirmed by the substantial disparity between the Sharpe-Ratio of the strategy (3.75) and the Sharpe-Ratio of the benchmark index (0.55). In terms of maximum loss, an investment in the strategy is far less volatile (2.19%) than an equivalent investment in the benchmark (19.95%).

The simulated trading period reflects the developments of the European credit crisis until the end of January 2013. The out-of-sample results extend the trend revealed by the in-sample figures. Although the observed time-series exhibit an economic downturn, the pairs trading strategy still maintains a stable performance. Out-of-sample strategy performance as documented by Figure B.11 found in Appendix B, is indicative of the increased strategy robustness when general market conditions deteriorate. Considering the corresponding risk measures, it becomes evident that the trading strategy is exposed to much less volatility (3.97%) in comparison to its benchmark (24.32%). The maximum drawdown also remains on a fairly low level (2.19%). As most of the figures are coherent, the superiority of the trading strategy is further confirmed by the strategy's high Sharpe-Ratio of 1.6 in relation to even a negative Sharpe-Ratio of -0.36 for the benchmark investment. The sustained resilience of the pairs trading strategy to abrupt market changes confirms its claimed advantages in comparison to long-only strategies. It can be recognized that in periods when the market hits bottom, the performance of the pairs trading strategy decouples from the market trend. As a consequence, the strategy exhibits a market-neutral property, providing either a hedging opportunity or new investment options.

4.3 Different Selection Criteria

The preceding section is solely built upon the pairs selection methodology proposed by Caldeira and Moura (2011). Previously, only the best in-sample Sharpe-Ratios are used as a reference point when forming suitable pairs for portfolio construction and their consequent out-of-sample testing. In order to improve the strategy, a diverse set of selection criteria is employed and backtested. An extended long-term backtest is then performed for each individual selection technique in conjunction with a thorough analysis with a special focus on their performance impact. Furthermore, it is important to mention that the empirical outcomes are evaluated excluding and including transaction costs.

Table 4.3: Results of 10-year out-of-sample backtest on the constituents of OMX30 excl. transaction cost for different selection criteria

Selection Criterion	**Sharpe**	**MAR**	**Treynor**	**Sortino**	**Benchmark**
CAGR	16.03%	16.55%	18.50%	20.04%	8.11%
Return (p.a.)	11.29%	11.63%	12.89%	13.86%	5.92%
Risk (p.a.)	4.96%	4.69%	5.45%	6.49%	28.05%
Sharpe (p.a.)	1.90	2.08	2.02	1.85	0.14
Max.Drawdown	2.97%	3.30%	4.04%	5.58%	65.88%
MAD	11500.87	12310.02	13723.22	15215.49	5078.17
VaR_{t+1}	148.82	153.09	205.72	240.73	550.67
$CVaR_{t+1}$	257.01	241.68	365.27	491.48	972.54
Kurtosis	22.70	25.80	22.04	27.25	24.06
Skewness	2.03	2.61	1.64	2.25	-1.47

An extensive overview of profit-and-loss figures computed for all the pairs selection criteria in conjunction with rolling alpha and beta parameters is presented in Appendix B (Graphical representations B.13 to B.18).

Table 4.4: Results of 10-year out-of-sample backtest on the constituents of OMX30 incl. transaction cost (51.19BP) for different selection criteria

Selection Criterion	**Sharpe**	**MAR**	**Treynor**	**Sortino**	**Benchmark**
CAGR	11.38%	13.80%	12.76%	16.04%	8.11%
Return (p.a.)	8.18%	9.82%	9.12%	11.30%	5.92%
Risk (p.a.)	5.25%	5.23%	5.39%	6.53%	28.05%
Sharpe (p.a.)	1.20	1.52	1.35	1.45	0.14
Max.Drawdown	4.70%	4.01%	4.70%	5.77%	65.88%
MAD	7293.61	9929.26	7538.89	10749.68	5078.17
VaR_{t+1}	114.56	137.83	136.29	187.97	550.67
$CVaR_{t+1}$	199.23	233.47	234.23	365.17	972.54
Kurtosis	22.32	23.23	28.82	27.71	24.06
Skewness	1.75	1.94	1.94	1.99	-1.47

Initial observations suggest that there may be a link between cumulative strategy performance and the choice of appropriate selection mechanism. As a consequence, the initial selection criterion is extended by MAR, Treynor- and Sortino-Ratio. The replacement of the Sharpe-Ratio with the Sortino metric increases annual return by additional 2.57 percentage points from 11.29% to 13.86% excluding trading costs. The appreciation in profitability is accompanied by a slight increase in the annual risk levels though, generating an out-of-sample volatility of 6.49% as compared to 4.96% for the Sharpe-Ratio. The respective annual return values for the MAR and Treynor fall in between the preceding two metrics, with figures hovering around 11.63% for the former and 12.89% for the latter. Portfolio allocation and trading initiation based on the Treynor-Ratio generates slightly higher risk levels as compared to the MAR with figures marginally below the 4.70% mark for the latter and lightly below the 5.50% for the former. It can be inferred that the discrepancy between the volatility rates for the Sharpe, Sortino and Treynor-Ratio can be attributed to the risk measure used in the denominator when computing the ratios. The first metric relies on the standard deviation as a performance standardisation tool, the second is built upon the downside volatility exposure, thus eliminating and not penalizing the upward variation, whereas the Treynor is only contingent upon the portfolio beta. Another important fact to consider is that the upswing in annual return levels following the replacement of the Sharpe-Ratio ratio with the Sortino metric comes at the expense of an incremental movement in the Value-at-Risk and Conditional Value-at-Risk measures. An investment in a portfolio constructed using the Sortino-Ratio is subject to and increased probability of large losses. It can be deducted with 95% confidence level that an investment in this portfolio will not incur a loss greater than 240.73 SEK during the subsequent trading session. Nevertheless this value is greater with 92.09 SEK than the exposure maintained when the Sharpe-Ratio is used as a selection criterion. The change in CVaR levels is even more pronounced. The maximum loss that could be incurred during the subsequent trading day significantly increases from 257.01 SEK to 491.48 SEK. In the worst case, the portfolio is exposed to an additional loss of 234.47 SEK during the next trading day.

The general conclusion that follows is that the Sortino-Ratio demonstrates superiority when the strategy return characteristics is evaluated independently. Nevertheless, when risk levels are taken into account, the pairs-selection algorithm based on the MAR-Ratio appears as more acceptable for a risk-averse investor due to an improved risk-return trade-off. The reduced annual risk exposure contributes to the surge in the level of the Sharpe-Ratio from 1.85 to 2.08 per annum.

When the comparison is performed relative to the benchmark index, it is apparent that regardless of the implemented selection criteria, the strategy outcompetes the OMX30 market index. The figures suggest that an investor which invests in the benchmark is exposed to a significantly larger risk in relation to its obtained return. That is confirmed by observing a low Sharpe-Ratio of 0.15 for an index investment in comparison to an equivalent investment in the pairs trading strategy yielding a Sharpe-Ratio of above 1 independently of the chosen selection criterion. The high volatility level of 28.05% related to a benchmark investment signifies a disproportion between generated excess returns and acceptable risk levels. Hence, an investment in the strategy which is based on the Sortino-Ratio is exposed to 21.56 percentage points less risk than the index investment. The disparity with the other selection criteria is even more indicative. The pattern is reversed when investigating the annual return levels though. A profound investigation of the remaining metrics confirms the superiority of the pairs trading investment technique in the long-run when applied on the OMX30 index. The 65.88% maximum drawdown incurred by the market index demonstrates a considerable potential for large losses. The corresponding maximum drawdown figures for the pairs trading strategy instead range from 2.97% to 5.58%, depending on the chosen portfolio selection technique. The above results reveal the advantage of pairs trading as a non-directional strategy, not directly tied to the constant market fluctuations and investor sentiment. The previous statement is reinforced by scrutinizing the Value-at-Risk and Conditional Value-at-Risk equivalents. The respective figures for the benchmark comprise a total of 550.67 SEK and 972.54 SEK. The MAR-Ratio as a selection criterion is exposed only to a VaR of 153.09 SEK and a maximum fat-tail loss of 241.68 SEK given by the CVaR. Due to the internal specification of market-neutrality inherent to the pairs trading technique, the strategy provides a way on how to exploit given market inefficiencies without intensifying the risk exposure. For further confirmation of the results presented above, the strategy performance is evaluated after taking trading costs into consideration.

The inclusion of transaction costs induces a modest impact on the strategy performance. The most significant downward pressure forces the annual return levels to decrease. The highest decline is borne by an investment strategy based on the Treynor-Ratio. Aggregate returns drops from 12.89% to 9.12%, thus revealing the severe impact trading costs induce on retail and institutional investors. Despite the sharp reduction in profitability after implementing the costs of executing the transactions, the returns accumulated by the pairs trading strategy still remain significantly positive. This is valid for all the selection criteria. Moreover, the long-term out-of-sample backtest encompassing the transaction costs of 100 basis points per round trip, unveils the negligible impact of trading fees on the risk measures built-in the pairs trading strategy. The highest maximum drawdown dispersion comprises 1.73 percentage points which is reflected by the Sharpe-Ratio, whereas the highest VaR and CVaR divergence consists of 69.43 SEK for the MAR-Ratio and 131.04 SEK for the Treynor-Ratio respectively. The data reported here support the assumption that transaction costs do play a major role in statistical arbitrage. Nevertheless, this paper comes to the conclusion that after accounting for brokerage fees, commissions, market impact (slippage fees) and costs of borrowing the short positions, pairs trading on the Swedish market still remains profitable.

4.4 Analysis of Strategy Performance on Different European Equity Markets

The preceding sections presented a comprehensive overview of the pairs trading profitability and risk exposure on the Swedish stock market. The implementation of a long-term backtest procedure in combination with market scenario analysis provides a broader perspective on the importance of statistical arbitrage as a portfolio diversification tool. When employed on the Swedish equity market segment, a thorough comparison with the relative benchmark index (OMX 30) confirms the hypothesized outperformance of the strategy on a risk-adjusted basis. Notwithstanding, the usage of a single equity market as a basis for strategy performance evaluation might be biased, thus leading to erroneous conclusions. A much more holistic approach would identify how the investment technique performs not only on different market settings, but also on different equity markets. One considerable advantage of drawing conclusions based on multiple stock markets is the elimination of selection bias, thus leading to an increased robustness of the accrued empirical results. A detailed interpretation of long-term backtest strategy results performed relative to the German DAX30 index and the European EUROSTOXX50 index is the focus of this section. The complementary figures for profit and loss and their respective monthly strategy returns for the German and European stock markets are given in Appendix B (Figures B.19 to B.22)

Table 4.5: Results of 10-year out-of-sample backtest on the constituents of EUROSTOXX50 incl. transaction cost (47.52BP)

Selection Criterion	**Sharpe**	**MAR**	**Treynor**	**Sortino**	**Benchmark**
CAGR	17.42%	15.53%	12.61%	14.82%	14.43%
Return (p.a.)	12.19%	10.96%	9.02%	10.49%	10.24%
Risk (p.a.)	5.84%	6.40%	4.62%	6.05%	27.80%
Sharpe (p.a.)	1.75	1.40	1.52	1.40	0.30
Max.Drawdown	4.65%	6.75%	4.38%	7.84%	79.60%
MAD	11428.09	9664.45	6179.31	8222.78	6384.94
VaR_{t+1}	197.23	177.19	120.83	170.33	927.95
$CVaR_{t+1}$	338.64	300.46	187.57	301.99	1641.87
Kurtosis	38.03	106.57	25.26	54.05	29.56
Skewness	2.94	5.49	2.11	3.32	-1.68

Table 4.6: Results of 10-year out-of-sample backtest on the constituents of DAX30 incl. transaction cost (47.52BP)

Selection Criterion	**Sharpe**	**MAR**	**Treynor**	**Sortino**	**Benchmark**
CAGR	10.14%	10.59%	12.49%	8.31%	3.98%
Return (p.a.)	7.33%	7.64%	8.94%	6.06%	2.96%
Risk (p.a.)	5.62%	5.19%	5.08%	5.54%	25.98%
Sharpe (p.a.)	0.96	1.09	1.37	0.74	0.04
Max.Drawdown	20.20%	17.90%	6.12%	20.22%	62.13%
MAD	4692.79	4974.37	6950.68	3641.72	3032.12
VaR_{t+1}	114.71	111.56	133.12	87.44	351.64
$CVaR_{t+1}$	206.79	195.32	214.47	172.39	578.15
Kurtosis	21.01	20.86	29.52	28.45	14.66
Skewness	1.13	0.87	1.93	0.07	-0.51

A detailed comparison of all the empirical results reveals that the strategy generates the highest abnormal returns when performed on the European equity market using the Sharpe-Ratio as a selection mechanism. In contrast to the equivalent benchmark index though, the achieved return (12.19%) looses its significance due to a fairly high performance (10.24%) of the EUROSTOXX50 benchmark itself. An important fact to consider is that the strategy performance is evaluated on the mentioned markets using different selection criteria as a reference point when constructing portfolios. In terms of risk considerations, all of the proposed strategies provide a significant edge to investors when executed. The risk figure for an investment in the strategy is substantially low with 5.84%, whereas the respective risk exposure for EUROSTOXX50 exceeds the 27% mark on an annual basis. When the focus is switched towards the German stock market, it is evident that there is a sharp drop in the annual return of the strategy represented by 8.94% for the best performing selection criteria. The latter figure represents a 5.98 percentage points increase relative to the DAX30 benchmark, but a 3.25 percentage points decrease when compared to the top-performing selection mechanism under the EUROSTOXX50 scenario. The high liquidity of the German equity market in combination with increased competition among arbitrageurs leads to a reduction of return levels, thus restricting investors from fully exploiting the potential of this trading strategy. According to this, it can be argued that markets which show lower returns have already been used in the past to apply the proposed strategy. This implies that the Swedish and the European equity market still provide room for application of the trading algorithm. In the case of the EUROSTOXX50, the size of the asset market could be also partly responsible for the success of this trading strategy, as the chances to find good performing pairs increase with the number of assets monitored.

This section has reviewed the robustness of the initially presented empirical results and their most appropriate interpretation. The next section concludes this research by summarizing the main findings and provides suggestions for further research.

Chapter 5: Concluding Remarks

The central conclusion following the main empirical findings is that the results of the pairs trading strategy can be considered significant and conclusive. Hence, if the trading algorithm can generate abnormal returns continuously, over a long period of time, statistical arbitrage can be used as a loss protection and portfolio diversification mechanism. The main empirical outcomes of long-term backtest confirm both of the predefined hypotheses, namely that the trading strategy generates excess returns and is exposed to less risk than the benchmark. Hence, it can be concluded that the strategy is in fact profitable when applied on the Swedish stock market OMX, using the Sharpe-Ratio as a selection criterion. In comparison to the benchmark, the strategy still performs exceptionally well on a risk-adjusted basis even during crisis times yielding an annual return of 11.29% with a volatility of 4.96% excluding transaction costs. Furthermore, we found that transaction costs, liquidity issues and size of the equity market have an significant impact on the strategy's profitability. Including trading costs the annual return of the strategy changes to 8.18% with an annual risk exposure of 5.25%. However, the performance of trading strategy still remains superior to its benchmark with an annual return of 5.92% and a risk of 28.05%. The analysis of the long-term backtest regarding its rolling Beta coefficient confirms the market-neutrality property by oscillating between a range of +0.040 and -0.035. Moreover, the robustness of the empirical results are confirmed by applying the pairs trading strategy further on EUROSTOXX50 and DAX30. It can be found that the overall results of EUROSTOXX50 are outstanding whereas the results of the DAX imply a lack of performance due to Sharpe-Ratio below 1, but still outcompeting its respective benchmark. Finally, it was documented that changes in the selection criterion clearly have an impact on the strategy's performance. However, the empirical results remain inconclusive regarding the superiority of a specific ratio for all markets. Thus, the MAR-Ratio remains the best performer on the Swedish market with a Sharpe-Ratio of 1.52. Moreover, the Treynor-Ratio generates the highest results on the German equity market measured by a Sharpe-Ratio of 1.37. After all, the Sharpe-Ratio as a selection criterion shows the best overall performance with a historically achieved Sharpe-Ratio of 1.75.

Taken together, these findings suggest that the pairs trading technique shows profitable results and superior risk exposure across all markets. This paper contributes to the existing literature by analysing different pairs selection criteria and their impact on trading performance. Furthermore, an isolated analysis of different market scenarios and an extended evaluation of distinct equity markets has been conducted to support the robustness of the empirical results generated by the initial long-term backtest. The implementation of appropriate risk measures further provides a thorough understanding of the associated risk and return trade-off every investor faces. However, this empirical findings need to be interpreted with caution because all the results are based on continuously re-balancing the portfolio everyday due to implementation purposes. Nevertheless, the drawn conclusions of the presented results still remain significant.

As a suggestion for future research, this framework could be used in a more realistic setting, employed by a quarterly or yearly re-balancing of the portfolio. Furthermore, this approach relies on the assumption that the cointegration relationship that is determined in-sample will persist during the whole trading period. As a result, there is a need for more adaptive trading rules taking into account the fact that the cointegration relationship may not be stable over time. Bootstrapping might also be considered a value-adding technique to be implemented in the future. This method consists of comparing strategy performance based on randomly selected pairs and a pre-defined pairs selection algorithm. The bootstrapping procedure can enhance the robustness of the obtained empirical results because it allows random testing with constant pairs replacement. Since the method of pairs trading in this paper appears to be unrestricted in terms of industry criteria, it would be interesting to evaluate pairs trading in a context where pairs are solely build upon stocks belonging to the same sector and their consecutive impact on trading performance.

Appendix A: Literature Review

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Gatev et. al. (2006), "Pairs Trading: Performance a Relative - Value Arbitrage Rule"							
Daily data over the period 1962 through December 2002	U.S. market	Equity	Distance approach	Pairs trading as a contrarion investment strategy based on contegrated prices, seeking-term deviations from long-term steady state	**Mean excess return** Top 5 0.00745 Top 20 0.000895 **Standard deviation:** Top5 0.02101 Top20 0.1527 **Sharpe ratio:** Top 5 0.35 Top 20 0.59 **VaR at 5%:** Top 5 -0.00710 Top 20 -0.00398	a) Initation of trading positions one day following the dvergence in order to account6 for bid-ask bounce b) Implementation of committed and fuly invested captial measurees for excess return evaluation c) Testing the profitability of pairs trading in broad sector domain d) Evaluate strategy performance by splitting the sample period encompassing distinct business cycle periods e) Explore the portfolio systematic risk by regressing the excess returns on Ibbotson and Fama-French risk factors	a) Lack of logical rules in the perfomaed technical analysy avoiding data snooping b) Pairs formation based on firm that might not close econmic substitutes leading to increased probability of further divergence c) The study does not address the obstacles or costs involved in shorting in the U.S. market

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Caldeira and Moura (2013), "Selection of a Portfolio of Pairs Based on Cointegration: A Statistical Arbitrage Strategy"							
Daily equity closing prices, adjusted for dividends and splits, encompassing the period starting January 2005 and ending October 2012 summing up to 1.992 observations	Brazilian market	Equity	Cointegration methodology	Identification of statistical arbitrage (pairs trading) strategy via the implementation of cointegration techniques and relying on relative pricing to explore mean-reversion behavior	**Average annualized return:** 16.39% **Annualized volatility:** 12.42% **Annualized Sharpe ratio:** 1.34% **Maximum Drawdown:** 24.49% **Remark :** *Positions are closed after 50 days in order to avoid profitability decrease	a) Selection of trading pairs for portfolio composition based on "the best in-sample Sharpe ratios" b) Implementation of: 2-step cointegration procedure, dimensionless Z-score as a trading trigger, stop-loss constraint, no restriction on stock pairs selection and transaction costs c) Avoidance of portfolio rebalancing after trade initiations d) Testing of strategy performance during the recent global financial crisis e) Closure of open positions at the end of the trading period even if executed at a loss	a) Necessity to implement metrics evaluating the stability of the cointegration parameters

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Haque and Haque (2014), "Pairs Trading Strategy in Dhaka Stock Exchange: Implementation and Profitability Analysis"							
Daily data comprising a total number of 243 observations encompassing a sample of 20 stocks	Bangladesh market	Equity	Cointegration technique	Development of financially lucrative pairs trading technique based on "short-term movements away from the general long-run equilibrium"	**Pair 1 out-of-sample data:** Return (%) 14.19% Duration (days) 22 Annualized return (%) 804% Exit date 02/02/2011 **Pair 2 out-of-sample data:** Return (%) 1.79% Duration (days) 20 Annualized return (%) 38% Exit date 22/7/2012 **Pair 3 out-of-sample data:** Return (%) 5.46% Duration (days) 24 Annualized return (%) 124% Exit date 02/02/2011	a) Research conducted on the most highly liquid and actively traded stocks in the Bangladesh market b) Utilization of Jensens test of cointegration for the purpose of detecting pairs with steady long-run equilibrium c) Estimation performed via the Vector error correction model (VECM), rather than the more general Vector autoregression model (VAR)	a) Liquidity risk due to strategy construction in developing market b) Restricted stock availability in Dhaka Stock Exchange c) Extreme historical volatility in Bangladesh stock market d) The study neglects the impact of transaction costs on strategy performance
Andrade et. al. (2005), "Understanding the profitability of pairs trading"							
Daily data ranging from 1994 to 2002	Taiwanese market	Equity	Distance approach	Uninformed demand shock as determinants of the profitability and risks associated with pairs trading relative value investment strategy	**Annualized excess return:** 10.18% **Average number of traded pairs:** 19.30 **Annualized standard deviation:** 9.15% **Annualized Sharpe ratio of returns:** 1.11	a) Pairs trading gains as a "compensation for providing liquidity in markets with limited risk bearing capacity" b) Factor loadings as a cardinal cause for relative price divergence	a) The study implicitly alludes but does not extends the possibility of exploring actual trades of arbitrageurs taking the other side of the transactions

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Broussard and Vaihekoski (2012), "Profitability of pairs trading strategy in an illiquid market with multiple share classes"							
Daily data over the period 1987 through the end of 2008	Finnish market	Equity	The squared summation of normalized price deviations (SSD) approach	"Investigate the practical issues of implementing the self-financing pairs portfolio trading strategy"	**Mean excess return / Standard deviation / Sharpe ratio** **Equally-weighted committed:** 0.00497/0.03015/ 0.16 **Equally-weighted fully invested capital:** 0.00989/ 0.06252/ 0.16 **Value-weighted committed:** 0.00323/ 0.03103/ 0.10 **Value-weighted fully invested:** 0.00834/ 0.05618/ 0.15 **Market portfolio:** 0.00600/ 0.08281/ 0.07	a) Document the profitability of pairs trading in an illiquid market setting b) Implement twelve month formation period, combined with six month trading period c) Usage of two discrepant weighting schemes for portfolio return calculation d) Restriction on the quantity of evaluated pairs e) Conduct in-depth subperiod analysis and deploy distinct trigger signals and trading costs f) Pair formation encompassing distinct "shares or stock classes of a companys equity"	a) "The unique institutional characteristic regarding common and preferred shares" may induce selection bias during the formation period due to similar price deviation specifications b) The chosen sample period is mainly built around market downturn conditions which may alter the strategy inherent characteristics

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Yuksel et. al. (2010), "Pairs Trading with Turkish Stocks"							
Daily closing prices adjusted for dividends and stock splits, ranging from January 1st 1990 to March 31st 2007	Turkish market	Equity	Distance approach	Statistical arbitrage as an equity trading strategy identifying relative mispricing by performing in-depth analysis depicting arbitrage forces on an emerging market setting	**"Excess returns on the top 20 pairs portfolio for threshold level of 2.0 std deviations"** Fully invested capital (no waiting) 0.00451 Committed capital (no waiting) 0.00442 Fully invested capital (one-day lag) 0.00448 Committed capital (one-day lag) 0.00439 **"Summary statistics of monthly excess returns on pairs portfolios" (no waiting)** Excess return (fully invested) top 5 0.00642 top 20 0.00451 Standard deviation top 5 0.06706 top 20 0.05425	a) Explicit proof documenting excess return superiority for shorter trading horizon b) Strategy performance evaluated for different values of the parameters characterizing it (length of formation period, length of trading period, threshold level) c) Statement categorizing weaker performance of the arbitrage forces and pairs trading strategy in developing market domain due to high transaction costs d) Positive risk-adjusted return attributable to short component of the pairs portfolio e) Implementation of rolling window approach	a) Presence of relatively high transaction costs b) Weak explicit statement as to why the accumulation of short-lived positive excess returns is not explained by transaction costs, systematic risk and short-term mean reversion c) Insufficient support regarding the idea of "diversification benefits from combining multiple pairs in the portfolio"

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Tianyong et. al. (2013), "Profitability of Pairs Trading Tactics in Chinas Stock Market"							
Closing price data for the period starting in January 7th 2003 and ending in December 31st 2008 providing 1452 individual stock observations	Chinese market	Equity	Single Linkage method	Statistical arbitrage as a cardinal tool for over/undervalued stock portfolio composition aiming to exploit temporary mispricing fluctuations	**Strategy statistics:** Threshold value d= 2.0 stdev Annual return 3.44Annual st deviation 0.378 Annual Sharp ratio 1.0911 Beta 0.03 Jensens 0.000	a) Test pairs trading profitability in a setting not encompassing mature markets b) Detailed analysis of the three methods (cointegration, stochastic spread, single linkage) researching and evaluating the pairs trading technique c) Document augmented strategy performance (lower fluctuations) by considering proper trigger values for position initiations and closures d) Carry into effect two disparate methods for pairs trading performance evaluation (Bootstrap simulation; excess return collation with properly weighted static benchmark portfolio)	a) Lack of comparison between the results obtained for emerging market setting and values characterizing mature market trading performance

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Perlin (2008), "Evaluation of pairs-trading strategy at the Brazilian financial market Perlin"							
Daily data over the period 1962 through December 2002	Brazilian markett	Equity	Distance method (Minimum squared distance rule)	"Verify the performance and risk of pairs-trading in the Brazilian financial market for different frequencies of the database"	**Pairs trading performance for daily frequency:** Threshold value d= 2.0 stdev **Excessive return:** Long positions (Short positions (Total (number of days in the market 36.81 **Beta:** Value 0.055 Probability 0.196 **Jensens (alpha):** Value 0.000 Probability 0.641	a) Provide "the specific details about choosing pairs and defining the threshold value in pairs tradin" b) Compare pairs trading return performance with a nave strategy of buy-and-hold (sell-and-unhold) and the approach of randomly generated market entries by implementing Bootstrap simulation c) Implement equally weighted simulated portfolio assigning each trading position similar weight d) Training period build on the assumption of "moving window composed of approximately 2 years of data for all frequencies"	a) "The framework used in the study did not allow for liquidity risk of the strategy"

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Deaves et. al. (2013), "Pairs trading in Canadian markets: Pay attention to inattention"							
Stock return data and accounting data ranging from July 1983 to December 2011 "Passing this double screen are 2,723 securities"	Canadian market	Equity	Distance approach	Pairs trading as an effective "return-enhancement strategy"	**Monthly average excess returns for full sample:** **Average excess return (fully invested)** top 5 0.0246 top 20 0.0153 **Average excess return (committed capital) :** top 5 0.0137 Top20 0.0071 **Average excess return (one-day waiting)** top 5 0.0056 top 20 0.0077 **Standard deviation:** top 5 0.0478 Top 20 0.0233	a) Excessive performance of pairs trading in volatile market environment, when "investor attention becomes strained" b) Assess strategy performance for distinct sub-periods in order to clarify if gains are dependent on trading period horizon span c) Measure risk factor exposures for long and short positions using Fama-French three factor model d) Consideration of transaction costs as a factor attenuating pairs trading strategy performance	a) "Limits to Arbitrage" encompassing fundamental risk and noise trader risk is implicitly mentioned, but no regression tests are performed to verify the relationship b) Vague and short explanation of the "ostrich effect" ("a tendency for investors to lose focus when markets decline")

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Bogomolov (2010), "Pairs Trading in the Land Down Under"							
Daily closing prices comprising 3455 trading days (January 1st 1996- November 22nd 2010)	Australian market	Equity	Distance approach; Cointegration method; Stochastic spread method	Three disparate approaches to pairs trading-detailed information and in-depth analysis	**Monthly excess returns of:** **Using distance pairs trading strategy (+/- tc; top 5/top 20 pairs)** **Mean:** 0.0095/0.0063/ 0.0080/0.0050 **Standard deviation:** 0.0178/0.0128/ 0.0171/0.0124 **Sharpe ratio** 0.54/0.49/0.47/0.40 **Using cointegration (+/- tc; top 5/top 20 pairs)** **Mean** 0.0105/0.0048/ 0.0087/0.0032 **Standard deviation** 0.0313/0.0209/ 0.0306/0.0204 **Sharpe ratio** 0.34/0.23/0.28/0.16 **Using stochastic spread process method (+/- tc; top 5/top 20 pairs)** **Mean** 0.0038/0.0045/ 0.0005/0.0017 **Standard deviation** 0.0192/0.0123/ 0.0194/0.0119 **Sharpe ratio** 0.20/0.36/0.03/0.14	a) Thorough investigation concerning the "different approaches to pairs trading" b) Implementation of two-stage procedure to sift out stocks with sufficient liquidity specification c) Pairs trading as a naturally-leveraged strategy d) Focus on strategy reaction to Global Financial Crisis	a) Disregard commissions and margin requirements b) Short-selling restrictions imposed by the ASIC at the end of the trading period

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Ferretti et. al. (2014), "Pairs Trading and Relative Liquidity in the European Stock Market"							
Daily closing prices ranging from 2000 to 2009	European market Euro Stoxx 50	Equity	Distance method implementing VECM for pairs trading strategy construction "within demand and supply framework"	"Linear price discovery equilibrium model" connecting presence of pairs trading technique with cointegrated asset classes linked via price spread	**Average profit (absolute value):** Before transaction costs 70.27 After transaction costs 63.18 **Average volatility:** Before transaction costs 43.13 After transaction costs 43.19 **Sharpe ratio:** Before transaction costs 1.63 After transaction costs 1.46 **Cumulative profits** Before transaction costs 72.70 After transaction costs 65.37	a) Implement two consecutive trigger signals in order to initiate and close a position b) Data sample covering pre-crisis market sentiment as well as post 'Lehman era' in order to assess strategy performance in disparate market states	a) Fragmentary analysis on 'common factor exposures that drive prices to parity'

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Pizzutilo (2013) "A Note on the Effectiveness of Pairs Trading for Individual Investors"							
Daily opening, closing and maximum prices, corrected for dividends and corporate actions, Jan 2002-July 2003	Italian market	Equity	Distance approach	Constraints confining the strength of pairs trading as strategy generating abnormal returns for an individual investor	**Tading statistics** **Mean :** Top 5 6.69% Top 20 4.77% **Standard deviation:** Top5 0.779 Top20 0.572	a) Short-selling restrictions and trading costs included as factors eroding pairs trading profitableness b) Internally built restrictive mechanisms controlling for data snooping and robustness of results c) Model driven stop-loss constraints d) Rolling windows for selection and trading periods	a) Extensive Back-testing procedures not employed by the model, thus omitting variables related to "length of the selection period, subset of equities to pair, the metric to choose pairs, the entry and exit rules, the length of the trading period, etc."

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Do and Faff (2012), "Are Pairs Trading Profits Robust to Trading Costs"							
Daily data over the period July 1963 through June 2009	U.S. market	Equity	Distance approach	Algorithm based on direct trading costs as pivotal determinants of pairs trading strategy implementation	**Pairs trading excess returns after trading costs (best performing portfolios (24, 25, 27))** **Mean:** 0.0035 0.0025 0.0027 **Standard deviation:** 0.01 0.02 0.01 **Sharpe ratio:** 0.32 0.14 0.23	a) Roundtrip commissions combined with market impact and short-selling fees as central elements in trading strategy effectiveness b) Empirical investigation based on "a range of pairs portfolios" aiming to improve the baseline algorithm of Gatev et. al. c) Limit the assessed sample to most liquid stocks d) The number of zero crossing metric (NZC) in conjunction with the sum of squared differences in the normalized prices (SSD) as to increase filter robustness e) Document declining trend in pairs trading profits	a) Vague and underdeveloped reproduction of the double-sort contrarian strategy analyzed by (Hameed, Huang, Mian, 2010) used for comparison with pairs trading performance

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Bowen et. al. (2014), "Pairs trading in the UK equity market							
High frequency daily data covering the period 1979-2008	UK market	Equity	Distance approach	Pairs trading as an investment technique built on short-horizon speculative initiations	**Full sample period statistics** **Average annual returns (including transaction costs):** Top 5 pairs 6.50% Top 20 8.88% FTSE All-Share Index 7.80% **Average Annual Standard Deviation:** Top 5 pairs 5.05 Top 20 pairs 4.30 FTSE All-Share Index 16.53 **Recession period average monthly return:** Top 5 pairs 1.64% 0.35 Top 20 pairs 1.32% FTSE All-Share Index 1.55-%	a) Divide the observation window into two sub-sample time frames covering disparate business cycle conditions b) Implementation of risk factors assessing strategy performance c) Formation and Execution (Trading) periods remain constant throughout the study framework	a) Anomalies causing relative prices to diverge, thus triggering position initiations are taken at predefined level and not varied according to disparate thresholds

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Fifield et. al. (2005), "An analysis of Trading Strategies in Eleven European Stock Markets"							
Daily closing prices over 10-year period January 1991-December 2000	Eleven European markets	Equity	Non-parametric approach	"Analyze the predictive ability of technical trading rules and examines their implications for market efficiency"	**Analysis of descriptive statistics for daily returns covering best and worst performing countries** **Mean:** GRE 0.0270 TUR 0.0182 FIN 0.0869 IRE 0.0544 **Standard deviation:** GRE 1.8290 TUR 3.2930 FIN 2.5330 IRE 0.9488	a) Broad focus not limited to one specific trading rule b) Impact of stock exchange inherent specifications on trading rules profitability c) Lower average performance associated with higher risks and return distributions demonstrating signs of skewness and leptokurtosis d) Significant dependence between daily stock returns in almost all of the markets	a) Grouping countries according to geographical location is not adding value to trading strategy performance similarities

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Mori and Ziobrowski (2011), "Performance of Pairs Trading Strategy in the U.S. REIT Market"							
Daily common data combined with REIT data covering the period January 1987 to December 2008	U.S. market	Equity and REIT segments	Distance approach	Detailed investigation of "the effectiveness of pairs trading in the U.S. REIT market and its time-series dynamics compared to those in the U.S. general stock market"	**Summary of pairs trading profits for the overall study period** **No Delay** **Average monthly excess return:** REIT 0.00951 Stock 0.00483 **Standard deviation:** REIT 0.01226 Stock 0.02278 **Risk-adjusted excess return:** REIT 0.78 Stock 0.21 **One-day delay** **Average monthly excess return:** REIT 0.00377 stock 0.00205 **Standard deviation:** REIT 0.00979 Stock 0.02197 **Risk-adjusted excess return:** REIT 0.39 Stock 0.09	a) The selection process of suitable pairs enhanced by the inner stock homogeneity specification inherent to REITs b) REIT strategy performance deterioration following the burst of the technological bubble due to additionally imposed structural changes c) Transparency involved in REITs market, contributing to pairs formation d) REITs asset class experiences lower relative price divergence, which imposes the necessity to implement lower trigger levels as measured by standard deviation disparity	a) The high-frequency daily data implemented for back-testing procedures might cause bid-ask bounce effect on trading profits

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Alsayed and McGroarty (2011), "Arbitrage and the Law of One Price in the Market for American Depository Receipts"							
Intra-day contemporaneous data spanning from January 1st 2007 to December 31st 2009	UK market	Equity combined with ADR	Liquidity filtering combined with the product of individual stock price with ADR ratio for distinct companies in order to arrive at a mispricing used for trading	Pairs trading as a technique enforcing the Law of One price, evaluated by applying pairs combination between two disparate market segments	**Results for retail investors taking into account transaction costs** **Annualized return:** 0.89% **Daily return volatility (bp):** 1.77 **1% Daily VaR (bp):** -2.44 **Maximum Drawdown:** 5.82%	a) Pivotal study applied to high-frequency pairs trading as a mechanism fitted to ADR, thus exploiting the Law of One Price b) Implementation of the longest historical high-frequency period used for backtesting considerations c) Inclusion of exchange traded only ADRs and ordinary shares (without preferred shares) as to guarantee maximum level of liquidity d) Implementation of an extensive array of transaction costs hindering pairs trading performance e) ADR reveal greater resemblance to their "underliers" as compared to common stock	a) Gains and losses incurred by filtering out illiquid price observations not extensively weighted against each other b) Sudden shift in currency pairs exchange rate might bias the results

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Nath (2003), "High-Frequency Pairs Trading with U.S. Treasury Securities: Risks and Rewards for Hedge Funds"							
Data covering the period January 1994-December 2000	U.S. market	Treasury Securities	Distance approach	Implementation of pairs trading as a statistical arbitrage strategy and the examination of "risk-reward tradeoffs of pairs trading in the secondary market for U.S. Treasury securities"	**Strategy comparison** **Mean:** 1.44, 5.07, 2.41, 2,05, 1.43, 1.87, 1.32, 1.41, 1.41, 1.29, 1.29 **Standard deviation:** 0.17, 89.43, 24.66, 4.83, 4.68, 4.70, 4.52, 2.56, 4.68, 4.21, 4.22 **Sharpe ratio:** -, 0.04, 0.04, 0.13, 0.00, 0.09, -0.03, -0.01, 0.00, -0.04, -0.04 **Values presented in the order Risk free, SP 500, Salomon, P1505Z, P1505W, P1510Z, P1510W, B1505Z, B1505W, B1510Z, B1510W:** ***P=pairs trading, B=benchmark, 15=percentile opening trigger, 5=percentile stop-loss trigger, W=with transaction costs, Z=without transaction costs:**	a) Conceptualize risks involved in pairs trading as a form of an arbitrage strategy b) Pairs selection based on price as opposed to return c) Pairs trading performed via the selection of the most liquid securities in the market place d) Creation of duration-matched benchmark via the implementation of strips data is used for comparison e) Implementation of risk control metrics	a) Allude the possibility of premature closure of the open position but no consequent testing is performed

Sample Period	Market	Asset Class	Methodology	Main focus	Statistics	Innovative Contribution	Implementation restrictions
Bianchi et. al. (2009), "Pairs Trading Profits in Commodity Futures Markets"							
Daily commodity futures closing prices spanning the period January 1990-August 2008	Multiple market locations where commodities are traded	Commodity futures	Modified distance approach encompassing commodities specifications	Pairs trading as a compensation for enforcing the Law of One Price in the commodities futures market	**Committed capital portfolio (one-day delay)** **Mean excess return:** Top 5 pairs 0.0152 Top 15 pairs 0.0134 All pairs 0.0138 Benchmark index GSCI 0.0056 Benchmark index MSCI World 0.0013 **Standard deviation:** Top 5 pairs 0.0321 Top 15 pairs 0.0303 All pairs 0.0286 Benchmark index GSCI 0.0549 Benchmark index MSCI World 0.429 **Sharpe ratio:** Top 5 pairs 0.04740 Top 15 pairs 0.4409 All pairs 0.4836 Benchmark index GSCI 0.1015 Benchmark index MSCI World 0.0297	a) Robust presence of cointegration revealed in unrelated commodities facilitating strategy construction b) Regression analysis exploiting the connection between LOP and commodity futures returns c) Comparison undertaken between nave investment strategy built on fundamental analysis and pairs trading in commodities market segment d) Deep investigation of triggers generating strategy excess returns e) One-day delay accommodating "non-synchronous effects in global futures markets" f) No restriction (limit) imposed on trading period duration and hence "no need to recalculate new set of pairs"	Limits to arbitrage as a non-risk-free investment strategy in a general setting and in commodities markets in particular. The issue is augmented with different time-zone characteristics of the markets where the strategy is executed, thus restricting the full use to exploit temporary mispricings due to execution limitations (Commensurate level of volatility accompanying strategy returns making them non-risk-free)

Appendix B: Further Information

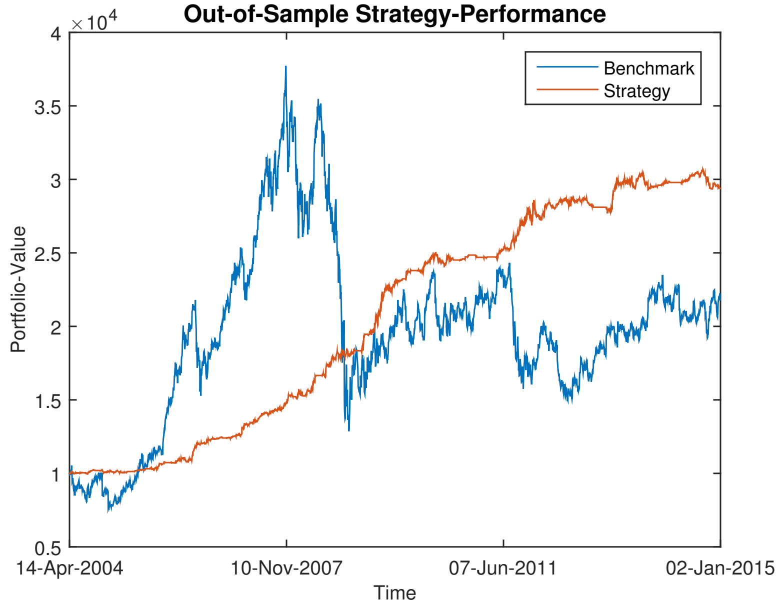

Figure B.1: Profit-and-Loss of the Strategy applied on the constituents of OMX

Longterm Out-of-Sample Backtest from 2002-01-05 until 2014-12-31

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2002	0	0	0	0	0	0	0	0	0	0	0	0
2003	0	0	0	0	0	0	0	0	0	0	0.77	-0.34
2004	-0.02	-0.5	2.12	-0.09	-0.02	-1.17	0.28	0.44	-0.86	0.66	0	0.79
2005	0.31	0.73	-0.41	1.86	2.09	0.43	-0.19	2.32	0.86	-2.26	6.14	4.32
2006	0.6	1.89	0.24	0.37	0.2	0.27	0.61	0.83	4.21	1.42	2.06	1.12
2007	1.92	0.5	1.45	0.03	2.32	3.99	-0.32	1.29	-0.89	1.88	6.5	0.27
2008	-0.76	4.47	2.97	0.93	0.53	0.56	0.91	0	5.94	1.23	4.99	3.38
2009	5.69	-0.62	1.84	0.73	1.23	1.64	0	-0.07	2.03	1.79	0.36	-1.99
2010	1.17	0.09	-0.42	0.24	0.64	0.48	-0.63	0	0.02	0.09	0.32	1.69
2011	0.32	2.13	1.56	3.34	3.1	-1.32	-0.32	0.57	2.32	-1.06	0.94	0.78
2012	-0.18	1.11	-0.65	0.39	-1.23	-0.99	0	0	-0.51	3.56	2.75	0.56
2013	-0.19	-1.31	2.51	-0.62	-2.27	0.5	0.21	0.83	0.04	-0.23	0	0.5
2014	0.41	0.96	-0.46	0.31	-3.03	1.12	-1.17	0	0	0	0	0

Figure B.2: Monthly Returns for the Strategy applied on the constituents of OMX

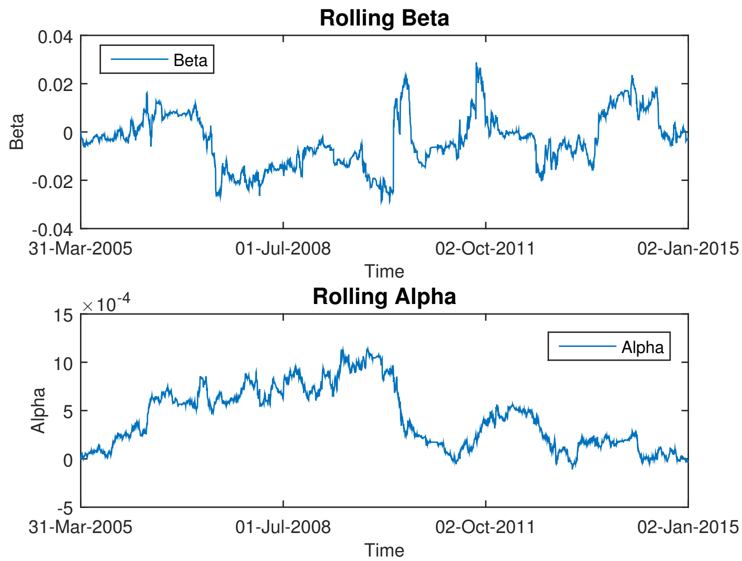

Figure B.3: Alpha/Beta Parameters of the pairs trading strategy for constituents of OMX

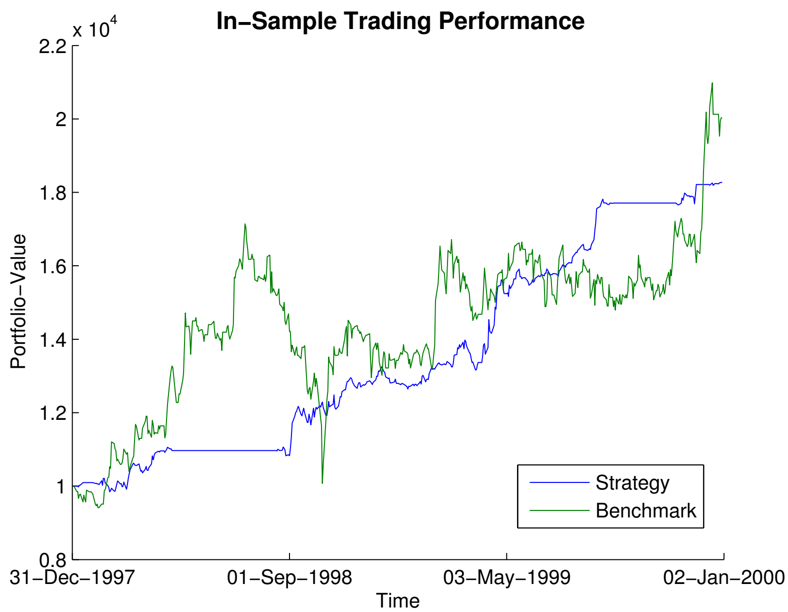

Figure B.4: In-Sample Strategy Performance of Scenario 1 for constituents of OMX

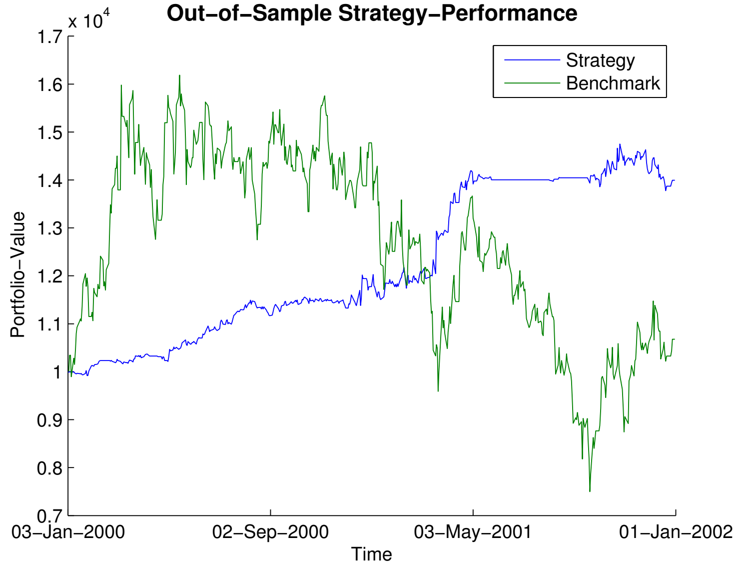

Figure B.5: Out-of-Sample Strategy Performance for Scenario 1 for constituents of OMX

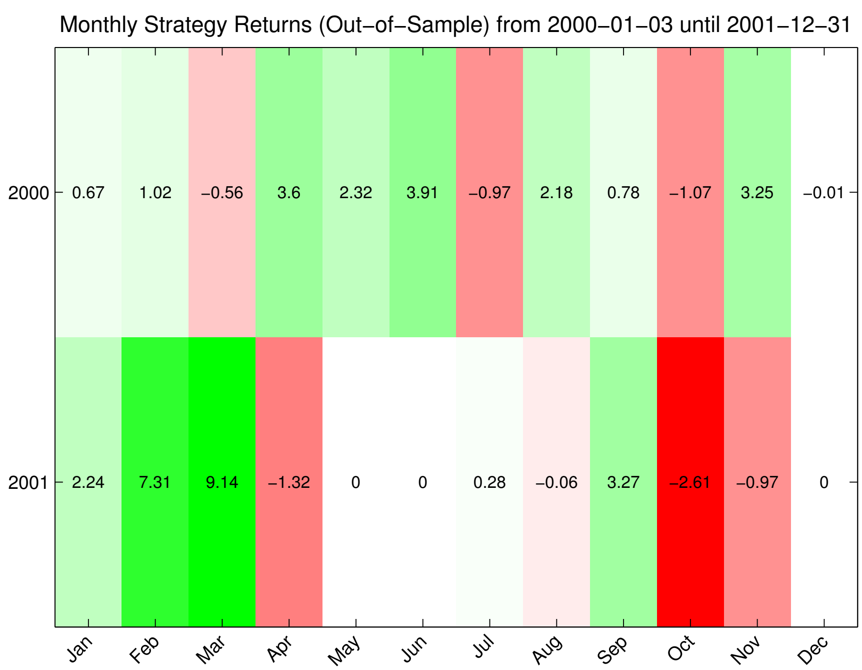

Figure B.6: Monthly Strategy Returns of Scenario 1 for constituents of OMX

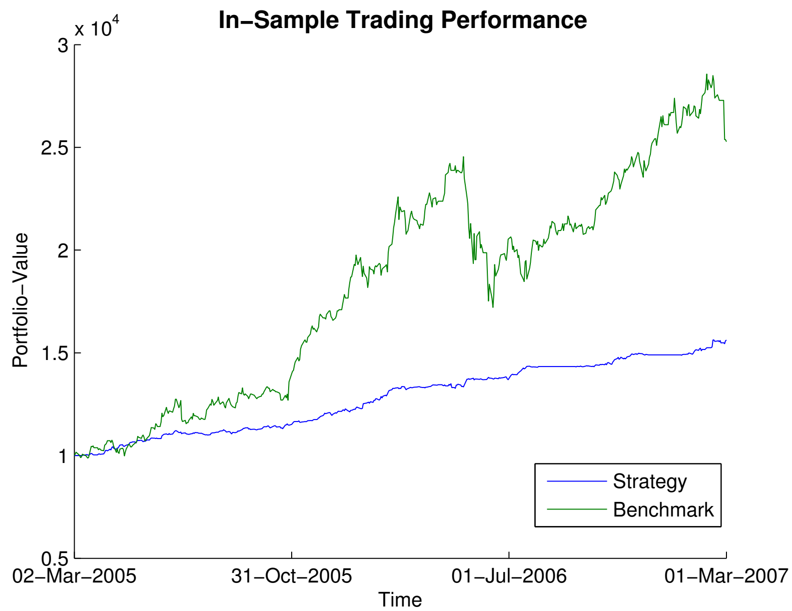

Figure B.7: In-Sample Strategy Performance of Scenario 2 for constituents of OMX

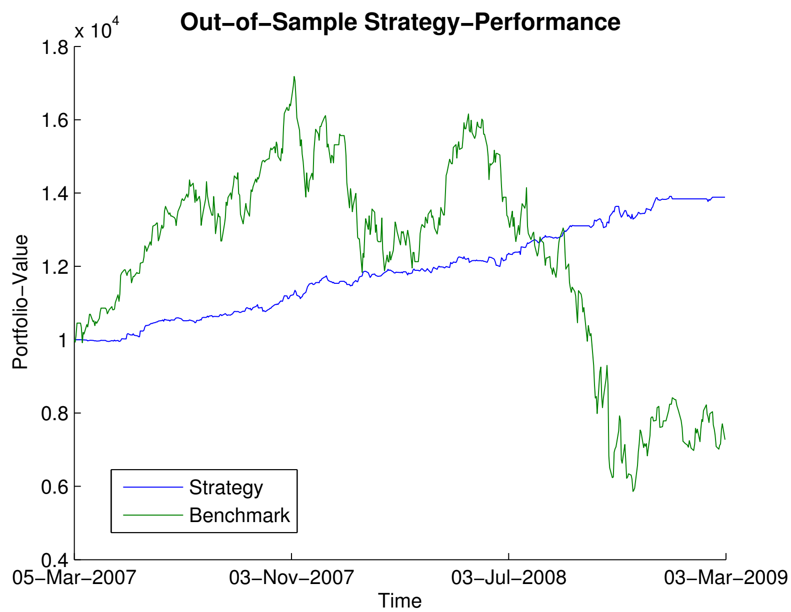

Figure B.8: Out-of-Sample Strategy Performance of Scenario 2 for constituents of OMX

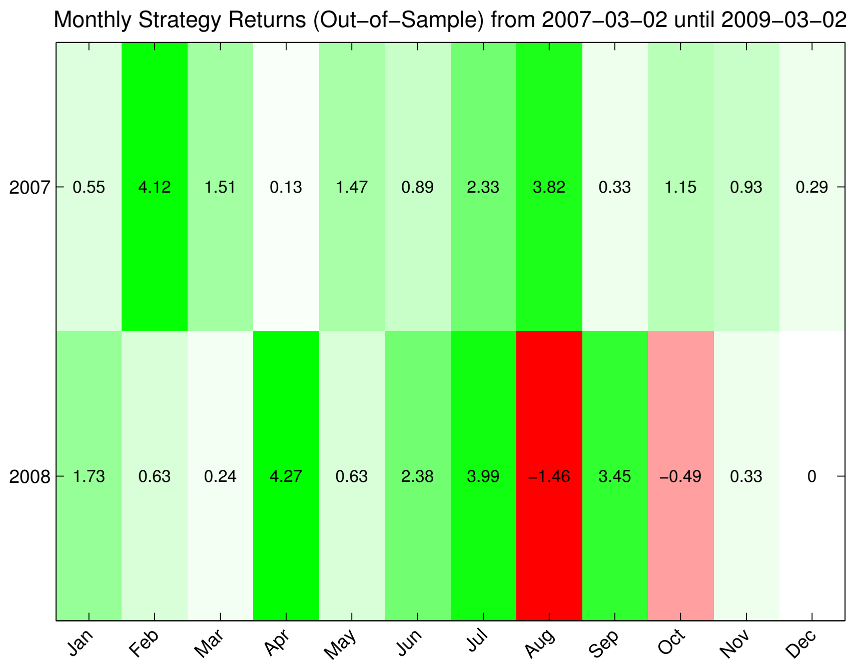

Figure B.9: Monthly Strategy Returns of Scenario 2 for constituents of OMX

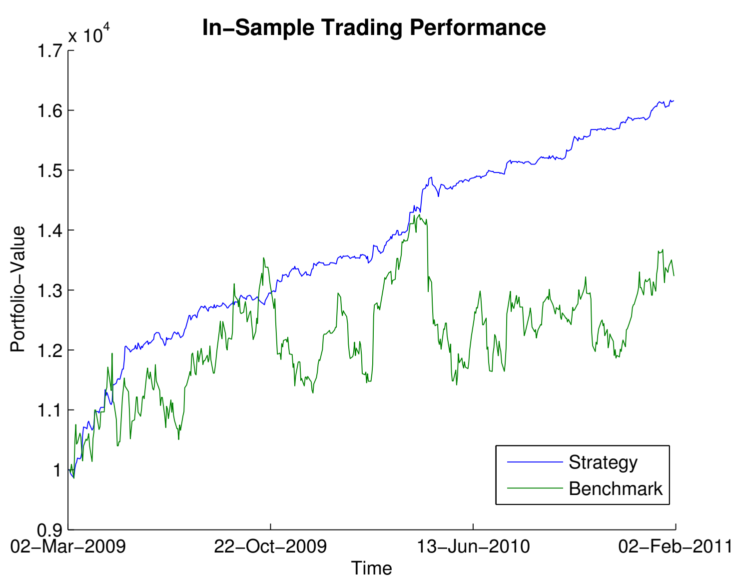

Figure B.10: In-Sample Strategy Performance of Scenario 3 for constituents of OMX

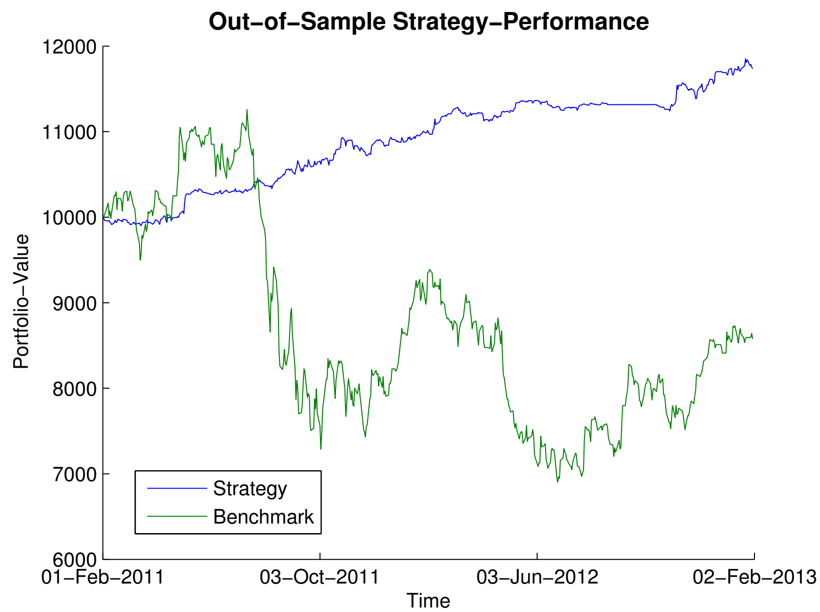

Figure B.11: Out-of-Sample Strategy Performance of Scenario 3 for constituents of OMX

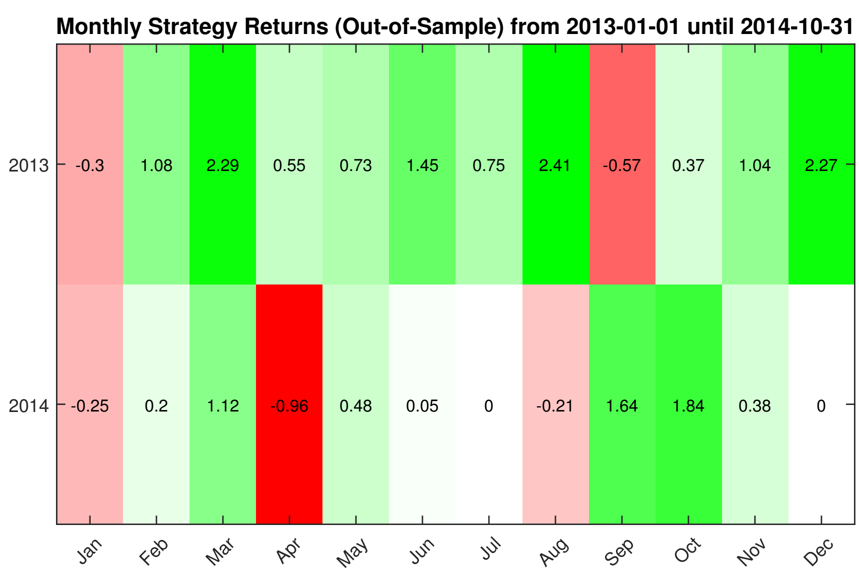

Figure B.12: Monthly Strategy Returns of Scenario 3 for constituents of OMX

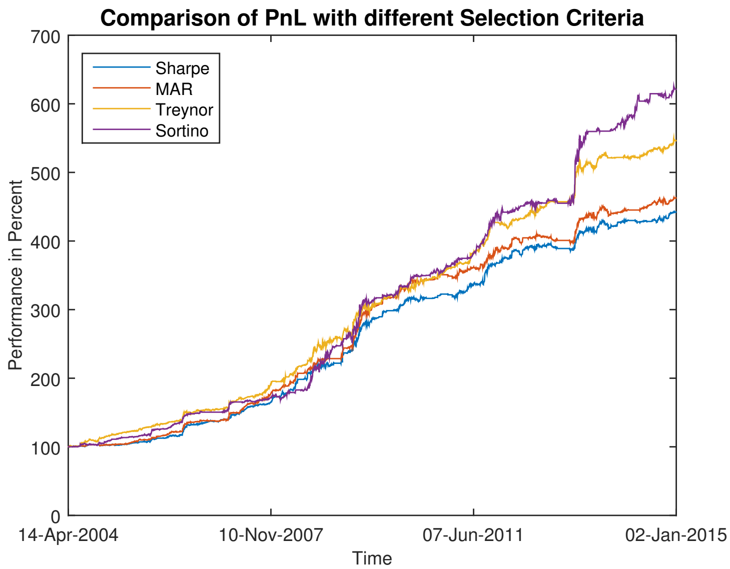

Figure B.13: PnL Overview of different Pairs-Selection-Criteria on OMX excl. Transaction cost

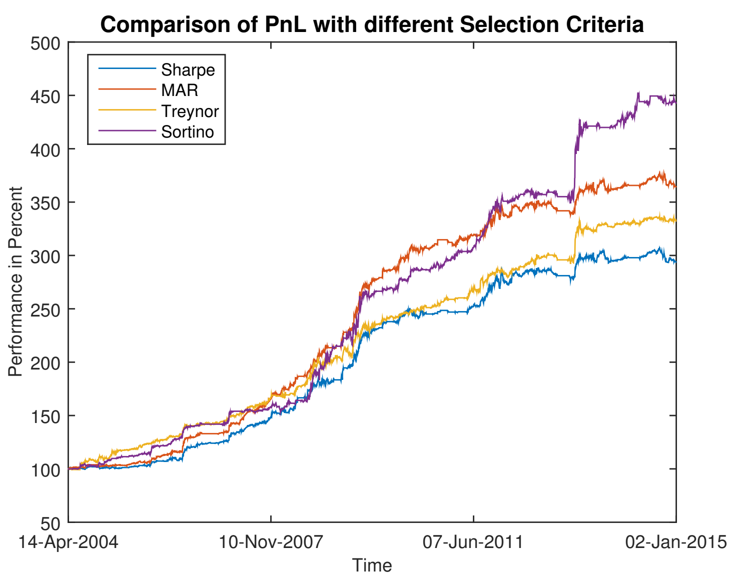

Figure B.14: PnL Overview of different Pairs-Selection-Criteria on OMX incl. Transaction costs

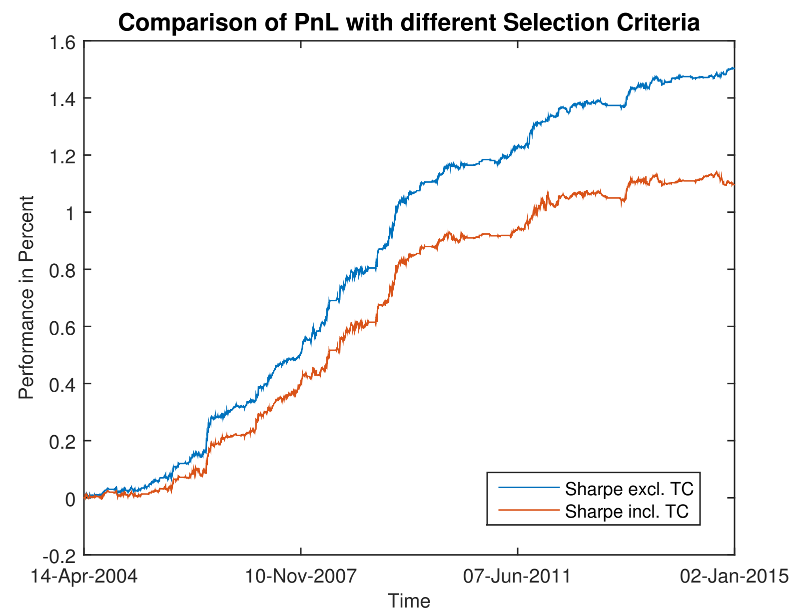

Figure B.15: Impact of Transaction cost on trading performance applied on constituents of OMX

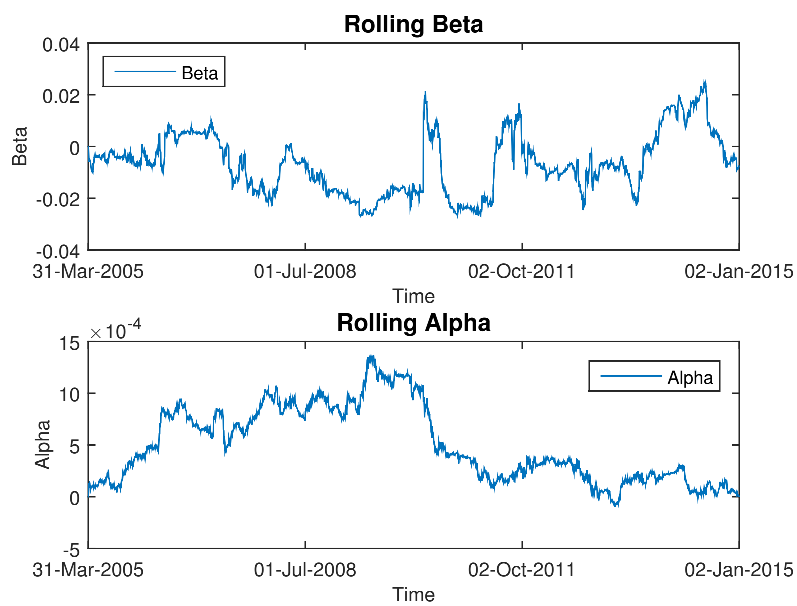

Figure B.16: Rolling Alpha and Beta using MAR-Ratio applied on constituents of OMX

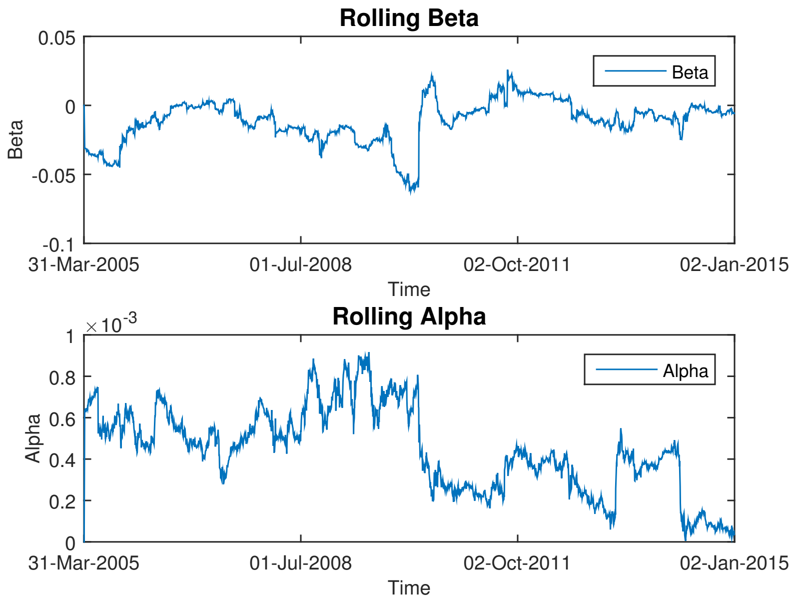

Figure B.17: Rolling Alpha and Beta using Treynor-Ratio applied on constituents of OMX

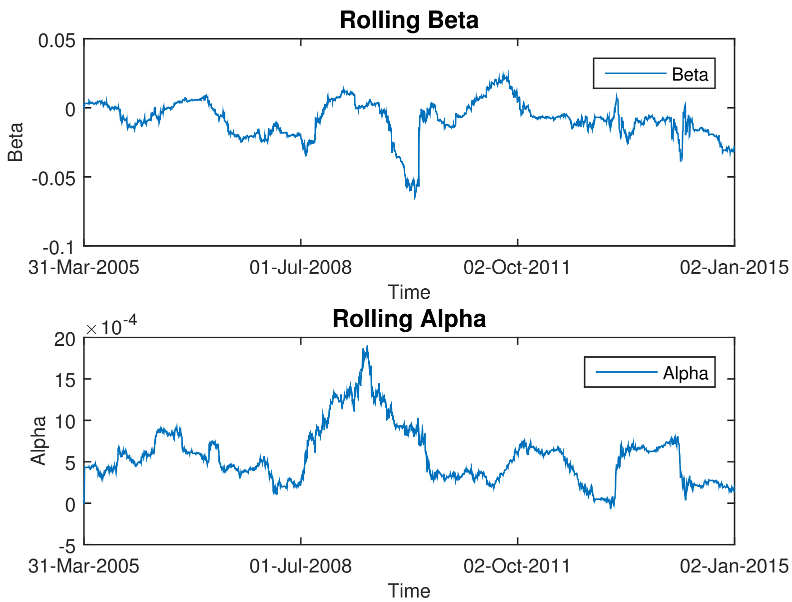

Figure B.18: Rolling Alpha and Beta using Sortino-Ratio applied on constituents of OMX

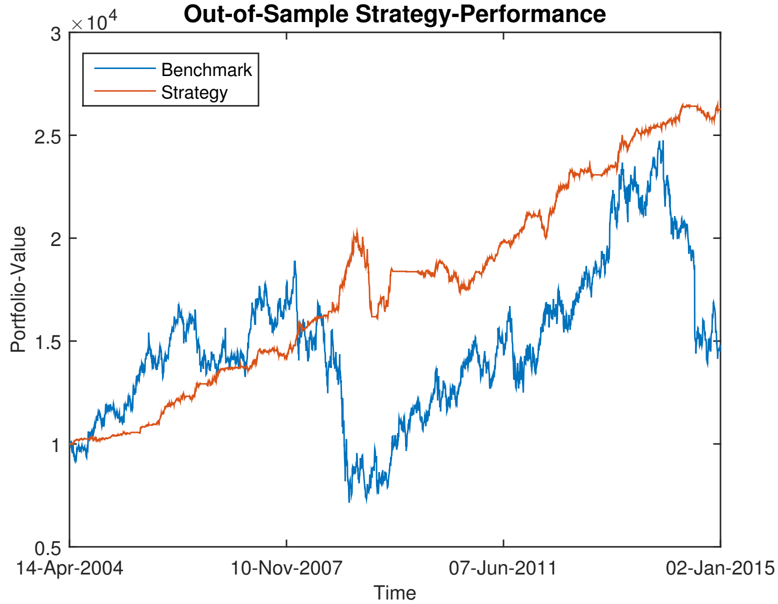

Figure B.19: Profit-and-Loss of the Strategy applied on the constituents of DAX30 using Sharpe-Ratio

Longterm Out-of-Sample Backtest from 2002-05-16 until 2014-12-31

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2002	0	0	0	0	0	0	0	0	0	0	0	0
2003	0	0	0	0	0	0	0	0	0	0	0.23	2.03
2004	-0.16	0.28	0.36	-1.41	1.54	-0.31	1.73	-0.22	0.71	0.17	0.49	0
2005	2.5	0.73	0.46	1.23	3.58	3.27	0.51	1.25	2.9	-1.02	0.4	3.93
2006	0.44	0.51	1.58	2.59	0.53	0.34	0.24	0.16	0.15	-0.69	2.36	4.23
2007	-0.14	-0.29	0.41	-1.81	-0.36	3.41	0.23	6.47	-1.13	1.68	1.67	0.18
2008	0.42	1.35	0.15	6.88	4.14	3.15	4.31	-1.57	-3.89	-10.01	-4.75	5.43
2009	-2.42	10.1	-0.33	0	-0.1	0.01	0	-0.5	0.76	-1.1	2.4	0.67
2010	-1.33	-2.19	-0.01	-2.37	-0.93	2.6	1.92	-0.03	1.92	1.94	0.62	3.23
2011	-0.11	0.31	1.81	1.34	3.43	-0.5	-0.34	-2.18	-0.64	4.64	3.63	1.43
2012	0.38	1.74	0.33	0.57	0.29	-0.95	0	0.36	0.77	0.59	2.52	2.3
2013	-0.84	2.02	0.68	0.05	0.65	0.29	0.2	0.63	0.71	0.01	0.06	2.33
2014	0.29	-0.13	-0.59	-0.85	-0.88	0.39	1.42	0	0	0	0	0

Figure B.20: Monthly Strategy Returns for constituents of DAX30 using Sharpe-Ratio

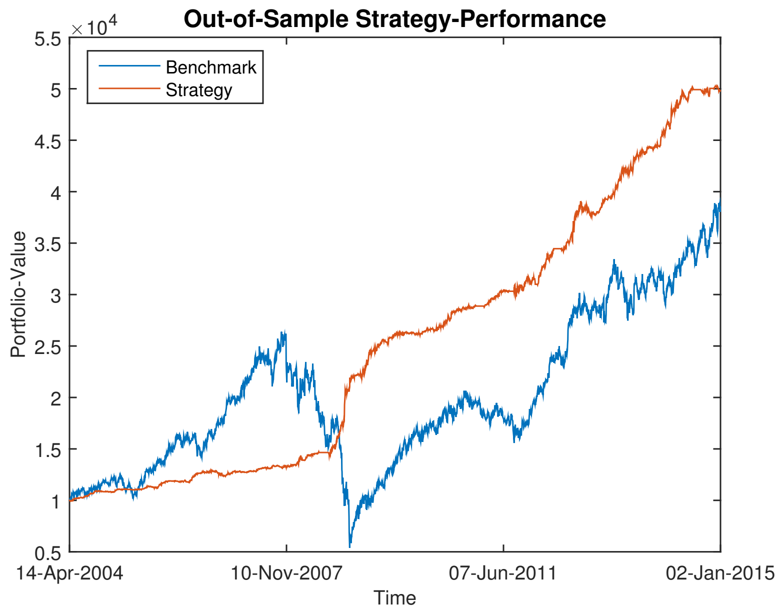

Figure B.21: Profit-and-Loss of the Strategy for constituents of EUROSTOXX50 using Sharpe-Ratio

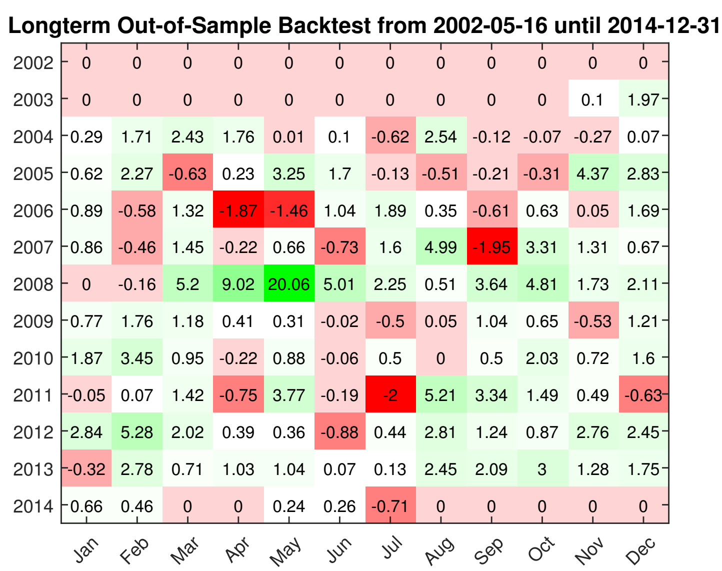

Figure B.22: Monthly Strategy Returns for constituents of EUROSTOXX50 using Sharpe-Ratio

Bibliography

Alexander, C. (1999). Optimal hedging using cointegration. *Phil. Trans. Royal Society Lond. A 357, pp. 2039-2058.*

Alexander, C., Giblin, I., and W., W. (2002). Cointegration and asset allocation: A new active hedge fund strategy. *Journal of Financial Risk Management, Volume 15, pp. 65-89.*

Alsayed, H. and McGroarty, F. (2012). Arbitrage and the law of one price in the market for american depository receipts. *Journal of International Financial Markets, Institutions and Money, Volume 22, Issue 5, pp. 1258-1276.*

Andrade, S., Di Pietro, V., and Seasholes, M. (2005). Understanding the profitability of pairs trading. *UC Berkeley Working paper.*

Avellaneda, M. and Lee, J. (2008). Statistical arbitrage in the U.S. equities market. *Quantitative Finance, Vol. 10, Issue 7, pp. 761-782.*

Baronyan, S., Boduroglu, I., and Sener, E. (2010). Investigation of stochastic pairs trading strategies under different volatility regimes. *The Manchester School, Vol. 78, Issue Supplement s1, pp. 114-134.*

Bianchi, J., Drew, M., and Zhu, R. (2009). Pairs trading profits in commodity futures markets. *Proceedings of Asian Finance Association 2009 International Conference, University of Queensland, Hilton Brisbane, Brisbane, Queensland, pp. 1-26.*

Bogomolov, T. (2010). Pairs trading in the land down under. *Finance and Corporate Governance Conference Paper.*

Bogomolov, T. (2013). Pairs trading based on statistical variability of the spread process. *Journal of Quantitative Finance, Vol.13, No. 9, pp. 1411-1430.*

Bowen, D., Hutchinson, M., and O'Sullivan, N. (2014). Pairs trading in the uk equity market. *European Journal of Finance, Forthcoming.*

Breitung, J. (2002). Nonparametric tests for unit roots and cointegration. *Journal of Econometrics, Vol. 108, pp. 343-363.*

Brooks, C. (2008). *Introductory Econometrics for Finance.* Cambridge University Press, 2nd edition.

Broussard, J. and Vaihekoski, M. (2012). Profitability of pairs trading strategy in an illiquid market with multiple share classes. *Journal of International Financial Markets, Institutions and Money, Vol.22, pp. 1188-1201.*

Caldeira, J. and Moura, G. (2011). Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. *Rev. Bras. Financas (Online), Rio de Janeiro, Vol. 11, No. 1, March 2013, pp. 49-80.*

Davidson, R. and MacKinnon, J. (2003). *Econometric Theory and Methods.* Oxford University Press, 1th edition.

De Boef, S. and Granato, J. (1999). Testing for cointegrating relationships with near-integrated data.

Deaves, R., Liu, J., and P., M. (2013). Pairs trading in canadian markets: Pay attention to inattention. *Canadian Investment Review.*

Dickey, D. and Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, Vol. 74, pp. 427-431.*

Do, B. and Faff, R. (2012). Are pairs trading profits robust to trading costs. *The Journal of Financial Research, Vol. XXXV, No. 2, pp. 261-287.*

Do, B., Faff, R., and Hamza, K. (2006). A new approach to modeling and estimation for pairs trading.

Dunis, C., Giorgioni, G., Laws, J., and Rudy, J. (2010). Statistical arbitrage and high-frequency data with an application to Eurostoxx 50 equities. *Liverpool Business School, Centre for International Banking, Economics and Finance.*

Dunis, C. L. and Ho, R. (2005). Cointegration portfolios of european equities for index tracking and market neutral strategies. *Working Paper of Liverpool John Moores University.*

Elliot, R., Van der Hoek, J., and W., M. (2005). Pairs trading. *Quantitative Finance, Vol. 5, No. 3, pp. 271-276.*

Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica, Vol. 55, Number 2, pp. 251-276.*

Engle, R. F. and Granger, C. W. J. (1991). *Introduction to Long-Run Economic Relationships: Readings in Cointegration*. Oxford University Press, New York, 1st edition.

Figuerola-Ferretti, I., Paraskevopoulos, I., and Tang, T. (2014). Pairs trading and relative liquidity in the european stock market. *Working Paper*.

Gatev, E., Goetzmann, W., and Rouwenhorst, K. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies, Vol. 19, No. 3, pp.797-827*.

Greene, W. (2003). *Econometric Analysis*. Prentice Hall, Pearson Education, Inc., Upper Saddle River, New Jersey, 5th edition.

Gujarati, D. (2004). *Basic Econometrics*. The McGrawHill Companies, 4th edition.

Haque, S. and Haque, A. (2014). Pairs trading strategy in dhaka stock exchange: Implementation and profitability analysis. *Asian Economic and Financial Review, Vol. 4, No. 8, pp. 1091-1105*.

Hillebrand, E. (2003). A mean-reversion theory of stock-market crashes. *CeVis Center for Complex Systems and Visualization, University of Bremen*.

Huck, N. and Afawubo, K. (2014). Pairs trading and selection methods: is cointegration superior? *Applied Economics, Vol. 47, No. 6, pp. 599-613*.

Jacobsen, B. (2008). Demonstrating error-correction modelling for intraday statistical arbitrage. *Applied Financial Economics Letters, 4:4, 287-292*.

Kozhan, R. (2010). *Financial Econometrics with eViews*. Ventus Publishing ApS, 1st edition.

Lim, G. and Martin, V. (1995). Regression-based cointegration estimators. *Journal of Economic Studies, Vol. 22(1), pp. 3-22*.

Lin, Y., McCrae, M., and Gulati, C. (2006). Loss protection in pairs trading through minimum profit bounds: A cointegration approach. *Journal of Applied Mathematics and Decision Sciences Volume 2006 pp. 1-14*.

Logan, T. (2008). *Getting Started in Candlestick Charting*. John Wiley and Sons, Inc., Hoboken, New Jersey, 1st edition.

Marshall, B., Nguyen, N., and Visaltanachoti, N. (2013). ETF arbitrage: Intraday evidence. *Journal of Banking and Finance, Vol. 37, Issue 9, pp. 3486-3498*.

Miao, G. J. (2014). High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. *International Journal of Economics and Finance; Vol. 6, No. 3*.

Mori, M. and Ziobrowski, A. (2011). Performance of pairs trading strategy in the U.S. reit market. *Real Estate Economics, Vol. 39, No. 3, pp. 409-428.*

Muslumov, A., Yuksel, A., and Yuksel, A. (2010). Pairs trading with turkish stocks. *Middle Eastern Finance and Economics, Issue 7, pp. 38-54.*

Nath, P. (2003). High frequency pairs trading with U.S. treasury securities: Risks and rewards for hedge funds. *Available at SSRN: http://ssrn.com/abstract=565441.*

Panyagometh, K. (2013). Pairs trading and value investing strategies. *International Journal of Business and Economics Perspectives, Vol.8, No. 2.*

Papadakis, G. and Wysocky, P. (2007). Pairs trading and accounting information. *Boston University and MIT Working Paper.*

Pelletier, R. (2007). Improve your trading through correlation analysis. *CSI Technical Journal, Volume XXV, Number 3, pp. 1-7.*

Perlin, M. (2009). Evaluation of pairs-trading strategy at the brazilian financial market. *Journal of Derivatives and Hedge Funds, Vol. 15, No. 2, pp. 122-136.*

Pizzutilo, F. (2013). A note on the effectiveness of pairs trading for individual investors. *International Journal of Economics and Financial Issues, Vol. 3, No. 3, pp. 763-771.*

Puspaningrum, H., Lin, Y., and Gulati, C. (2011). Finding the optimal pre-set boundaries for pairs trading strategy based in cointegration technique. *Journal of Statsitical Theory and Practice, Vol. 4(3), pp. 391-419.*

Rittereiser, C. and Kochard, L. (2010). *Top Hedge Fund Investors. Stories, Strategies and Advice*. John Wiley and Sons, Inc., Hoboken, New Jersey, 1st edition.

Rudy, J., Dunis, C., Giorgioni, G., and Laws, J. (2010). Statistical arbitrage and high-frequency data with an application to Eurostoxx 50 equities. *University of Liverpool, Available at SSRN: http://ssrn.com/abstract=2272605.*

Thapa, C. and Poshakwale, S. (2010). International equity portfolio allocations and transaction costs. *Journal of Banking and Finance, Issue 34, pp. 2627-2638.*

Tianyong, H., Ming, D., and Liang, W. (2013). Profitability of pairs trading tactics in chinas stock market. *Innovation Conference (SIIC), Suzhou-Silicon Valley-Beijing International.*

Veerbek, M. (2004). *A Guide to Modern Econometrics*. John Wiley and Sons, 2nd edition.

Vidyamurthy, G. (2004). *Pairs Trading Quantitative Methods and Analysis*. John Wiley and Sons, Inc., Hoboken, New Jersey., 1st edition.

Wang, P. (2009). *Financial Econometrics*. Routledge, 2nd edition.