



UNIVERSIDADE CATÓLICA PORTUGUESA

Pairs Trading

Strategy Refinements

Diogo António Ribeiro Magalhães

Universidade Católica Portuguesa
Abril de 2023



UNIVERSIDADE CATÓLICA PORTUGUESA

Pairs Trading

Strategy Refinements

Trabalho Final na modalidade de Dissertação
apresentado à Universidade Católica Portuguesa
para obtenção do grau de mestre em Finanças

por

Diogo António Ribeiro Magalhães

sob orientação de
Prof. Doutor Paulo Alves

Universidade Católica Portuguesa
Abril de 2023

Agradecimentos

Gostaria de agradecer à minha mãe, ao meu pai e ao meu irmão Bruno por todo o apoio e ajuda que me deram ao longo do meu percurso académico. Agradeço também a todos os meus restantes familiares que sempre me acompanharam e me incentivaram a continuar neste percurso.

Assim, queria agradecer a todos os meus amigos que relevaram ser um grande suporte. Aos meus colegas de mestrado e de trabalho agradeço por todo o apoio e compreensão.

Um último, mas não menos importante, agradeço ao Professor Paulo Alves por todo o seu apoio, orientação e suporte constante durante o desenvolvimento desta dissertação.

Resumo

Na presente dissertação foi aplicada uma estratégia de *pairs trading*, originalmente apresentada por Nunzio Tartaglia nos anos 80, para o período entre 2010 e 2020, aos constituintes do S&P500. Nós replicamos a estratégia usada por Gatev et al. (2006), com algumas modificações, e a nossa estratégia de trading obteve uma média de retorno semestral de 13.133% e um retorno total de 262.658%. A estratégia também apresentou um comportamento significativamente melhor quando os mercados se encontram em declínio.

Este estudo centrou-se em descobrir se o desempenho da estratégia seria positivo em diferentes indústrias. Tendo isso em consideração, criámos diferentes carteiras com os 5 melhores pares de ativos de cada indústria. Além disso, aplicamos estratégias de *stop loss*, com o intuito de melhorar o desempenho da nossa estratégia e diminuir o risco associado.

Os nossos resultados mostram que a estratégia pode não ser aplicável em todas as indústrias, uma vez que existiram carteiras que apresentaram perdas relativamente altas e que a implementação de uma estratégia de *stop loss* pode melhorar o desempenho.

Palavras-Chave: *Pairs trading*; Indústria; *Stop Loss*.

Número de Palavras: 9835

Abstract

In this dissertation, we apply the pairs trading strategy, which was originally presented by Nunzio Tartaglia in the mid-80's, for the period between 2010 and 2020, to the stock's constituents of the S&P500. We replicated the strategy used by Gatev et al. (2006), with some differences, and our trading strategy obtained an average 6-month excess return of 13.133% and a total return of 262.658%. The strategy also proved to have a significantly better performance when the markets were declining.

This study focused on discovering if the performance of the strategy would be positive in different industries. With that in mind, we created portfolios with the top 5 pairs of stocks of each industry and track the results obtained. Besides that, we applied stop losses strategies in order to refine our strategy and decrease the risk associated with the trades.

Our results show that the strategy might not be applied to all of the industry as there are portfolios that presented relatively high losses and that the implementation of a stop loss can improve the performance.

Keywords: Pairs trading; Industry; Stop Loss.

Number of Words: 9835

Table Of Contents

Introduction	1
1. Literature Review	4
1.1. Efficient Market Hypothesis & Market Risks	4
1.2. Pairs Trading Strategy	5
2. Methodology.....	12
2.1. Data Description	12
2.2. Pairs Formation	13
2.3. Closeness Measure Computation	15
2.4. Triggers and trading signals.....	16
2.5. Excess Return Computation.....	18
2.6. Stop Loss Strategy	19
2.7.. Limitations	19
3. Results.....	21
3.1. Unconditional Pairs Trading Strategy	21
3.2. Pairs Trading Strategy in each Industry	22
3.3. Introducing Stop Losses on the Strategy	26
3.4. Detailed Results of the Strategy	31
3.5. Market Neutrality	34
4. Conclusion	36
Bibliography	39

List of Tables

Table 1: Sample Distribution by Industry	13
Table 2: Possible Pairs in each Formation Period	14
Table 3: Results of the Pairs Trading Strategy against the S&P500	21
Table 4: Results of a Pairs Trading Strategy for each Industry	24
Table 5: Pairs Distribution per Industry for the Unconditional Pairs Trading...	25
Table 6: Results of implementing a 10% Stop-Loss on each Portfolio	28
Table 7: Results of implementing a 15% Stop-Loss on each Portfolio	29
Table 8: Results of implementing a 20% Stop-Loss on each Portfolio	30
Table 9: Detailed results of Industrial Pairs Trading Strategy	31
Table 10: Detailed results of trades at the moment of closing position	33

List of Figures

Figure 1: Example of creation of pair formation and trading.....	13
Figure 2: Example of the normalized price returns in a formation period.....	16
Figure 3: Example of the trading period of a pair	17
Figure 4: Returns Evolution of the strategy and the S&P500 Industrials	35

Introduction

Investors have been creating and developing strategies since the beginning of financial markets to reduce risk and improve financial results, to achieve consistent profits. A multitude of strategies is available, including breakout strategies that profit from the identification of emerging trends in the market, and momentum strategies that follow positive trends in the market. Pairs Trading is a popular strategy that will be the main focus of my dissertation.

The Pairs Trading strategy traces back to the early 1980s when a group of mathematicians, physicists, and computer scientists, led by Nunzio Tartaglia at Morgan Stanley, aimed to create and automate new trading systems based on statistical models (Gatev et al., 2006). Since then, the strategy has been adopted by institutional investors, hedge funds, and investment banks. One of the key features that make Pairs Trading attractive is its market neutrality, which enables traders to profit from various market conditions, along with its self-funding ability.

This strategy involves identifying a pair of two financial instruments and monitoring their behaviour over a period of time. If certain conditions are met, the pair is traded. The first condition is that the daily prices of the two instruments have a high correlation, meaning that their prices move in similar ways during the analysis period. Once this condition is met, positions are opened when prices diverge because the strategy assumes that prices will eventually adjust to maintain a similar pattern. When prices converge, the positions are closed. This approach, developed by Nunzio Tartaglia and his team, seeks to generate returns through the divergence and convergence of stock prices for pairs that exhibit similar behaviour. Its attractiveness stems from its market neutrality and self-funding ability, allowing traders to profit from any market condition.

Although “the strategy arguably has low exposure to systematic risk factors because a long position in one stock is effectively hedged by an offsetting position in another stock with similar factor loadings” (Papadakis & Wysocki, 2007). However, the strategy is commonly associated with non-systematic risk, which is inherent in a certain firm or sector, because most of the pairs have stocks from the same sector.

Previous studies have concluded that the correlation between stocks is higher when the stocks are from the same industry than when the stocks are from different industries (Chan et al., 2007). Cavaglia et al. (2000) demonstrate that stock returns are very susceptible to different industry factors, which is a determinant of our research question. As a result, investors may be able to use pairs of stocks within the same industry to reduce or eliminate non-systematic risk. While it may be preferable for investors to choose financial instruments from the same industry, it is unclear whether this approach yields better results for the Pairs Trading strategy. This dissertation seeks to examine whether taking industry risk into account leads to improved performance and to identify the factors that motivate investors to select specific industries for their pairs trading strategies. Cavaglia et al. (2000) also conclude that country factors are important to the stock returns, but, given that our data will only be extracted from Standard & Poor's 500 (S&P500), the diversification of countries is not as relevant.

This dissertation seeks to examine whether taking industry risk into account leads to improved performance and to identify the factors that motivate investors to select specific industries for their pairs trading strategies. As referred, the data used will contain the stocks of the constituents of S&P500 between 2010 and 2020. The analysis will involve two steps: first: we will replicate the methodology introduced by Gatev et al. (2006) to a portfolio an unrestricted portfolio, and then we will emply the same methodology to construct 11 portfolios based on their industry classification. In the second step, we will introduce stop loss limits to all portfolios to assess the impact on strategy performance.

Our findings suggest that among all the portfolios created, the industrials sector exhibited the highest returns and Sharpe ratio. Additionally, we observed that the implementation of stop loss limits per percentage of loss resulted in improved performance for the pairs trading strategy.

1. Literature Review

1.1. Efficient Market Hypothesis & Market Risks

According to the Efficient Market Hypothesis presented by Fama (1970), a market is considered to be efficient if the stock prices reflect all the available information, and therefore all market participants (investors and companies) have access to the same true and transparent information to support their resource allocation decisions. According to the hypothesis, the principle underlying an efficient market is that stock prices react to the arrival of new information.

The Efficient Market Hypothesis posits there are three versions of market efficiency: the strong form, the semi-strong form, and the weak form. The strong form suggests that asset prices reflect all monopolistic information held by investors, the semi-strong form suggests that prices reflect all publicly available information and adjust instantly to reflect new public information, and the weak form suggests that prices reflect all publicly available information.

In addition to the Efficient Markets Hypothesis, there is another important economic concept in the context of our dissertation, namely the “Law of One Price”. The application of this concept to financial markets implies that an asset should have the same price in different markets. Otherwise, there are arbitrage opportunities through the purchase of the asset at a lower price in one market and the sale of the same asset in another market (where its price is higher). According to the Law of One Price, two stocks with the same payoffs and equal risk factors should exhibit a stable relationship with their relative price (Gatev et al., 2006). Once again, if this condition is not met, there is an arbitrage opportunity that should be exploited.

Researchers suggest that in an efficient market, two stocks that present similar historical prices must present similar actual and future prices because they have the same value. Therefore, whenever there is a spread between the stock prices, the market is inefficient, and a violation of the Law of One Price, providing arbitrage opportunities to investors.

Therefore, Gatev et al. (2006) proposed that the returns obtained with the applicability of a pairs trading strategy result from the compensation of the risk of enforcing the "Law of One Price".

1.2. Pairs Trading Strategy

Although Pairs Trading is a well-known strategy and is frequently used by financial institutions and hedge funds managers, there is under-research in the academic literature. Nonetheless, there has been an increase in interest in the study of pairs trading during the past few years, and fresh research contributions have been made in this area. Simultaneously with new methodological advancements and empirical data, these most recent breakthroughs have helped to clarify the underlying principles.

Gatev et al. (2006) is one of the most cited papers and most important papers for the pairs trading strategy since it was the first researcher that presented the strategy academically. The first paper of the researchers was published in 1999 and it had the goal of studying the market and the define the factors and limitations to have in consideration for the reviewing research presented in their final paper. The research of Gatev et al. (2006) had an important impact on the development of future academic literature and improving the strategy because the research conducted makes a very concise study of the strategy and the conditions to apply it, like the composition of the portfolio, the transaction costs,

the industry, the risks of applying this strategy, and even the long and short perspective of the strategy.

In their study, pair trading is implemented in two stages. The first is the formation period (12 months), in which the stock daily prices were analysed and, with that, formed the pairs for the strategy, and the second is the trading period (6 months), in which the researchers trade the pairs that were formed previously. This approach is quite common in research on the Pairs Trading strategy, like the paper (Papadakis & Wysocki, 2007).

In Gatev et al. (2006), to compute the pairs of stocks, the researchers extracted the data from the U.S. market, downloading the prices from 1962 to 2002 of the stocks from the CRSP files. However, the only restriction that all the stocks chosen had was that, if they had one or more days without a trade, they were discarded. After that, Gatev et al. (2006), normalized the returns for each stock, containing the reinvested dividends. Papadakis & Wysocki (2007) also used U.S. stocks to form the pairs in their research and also used a similar approach however the period of tracking the strategy is different (between 1981 and 2006).

Although Papadakis & Wysocki (2007) followed the same strategy as Gatev et al. (2006), when examining the daily returns, the approach used to select the pairs consists in aggregating the stocks in quintiles accordingly to their 12-month return and then they divided the stocks in quintiles again, accordingly to their 6-month return of the first half of the period, which can seem more complicated to apply but, in the end, gives a more concise result. After aggregating the stocks for quintiles, Papadakis & Wysocki (2007) had 25 groups of stocks, because every quintile of the first quintiles was divided into 5 quintiles, and each group of stocks presented a similar co-movement in the stock returns. Accordingly, to them, “this co-movement in stock prices can be partially attributable to the fact that both securities had similar loadings on systematic risk factors” which contributes to the low systematic risk of the strategy. In the end, this approach used by Papadakis & Wysocki (2007) to cluster stocks, can seem more time

dispendious than the regular approach of Gatev et al. (2006), however, it can save some time when computing the closeness measure.

Regarding the excess return of the strategy, both researchers have presented a positive return in the period analysed. Gatev et al. (2006) presented an average monthly return of 1.31% for a portfolio of the best five pairs and an average monthly return of 1.44% for a portfolio of the best twenty pairs. So, according to these results, there is an advantage for investors to use diversification and have a portfolio with multiple pairs. Although Gatev et al. (2006) has admitted that these returns were biased because they normalized the daily returns of the stocks considering the closing price of each day which means that they didn't take into consideration the intraday variation of the price of stocks and the bid-ask bounce. So, to solve the problem of using closing prices, they added the rule of "one day waiting", which consists in opening a position one day after the signal is given and closing the position one day after the convergence of prices. After adding this rule, the average excess return of the strategy dropped to 0.74% on the portfolio with 5 pairs and to 0.89% on the portfolio with 20 pairs.

Papadakis & Wysocki (2007) also used portfolios of 20 pairs of stocks with the smallest value of closeness and applied the "one-day waiting" rule and, for the period of trading, the researchers obtained an average monthly excess return of 0.62%. The main argument given by Papadakis & Wysocki (2007) for the lower return in their strategy is that the strategy contains a more recent period. The results of Papadakis & Wysocki (2007) have proved that there has been a decrease in the profitability of the strategy in recent years since the average monthly return between 1981 and 1993 was 0.87% and the same strategy had an average monthly return of 0.36% between 1994 and 2006.

After applying the strategy developed by Gatev et al. (2006), Papadakis & Wysocki (2007) tried to improve the strategy by studying if accounting information events can affect the profitability of the strategy. So, to begin, they divided the accounting information events into earnings announcements and

clustered analyst forecasts, and they analysed the profitability when a position was either open or closed after the accounting information events. For the full period in analysis, 9% of the trades in a pair occurred after an earnings announcement and, these trades presented an average monthly return of 0.07% however the trades that occurred after a clustered analyst forecast represented 14% of all trades and the average excess return for these pairs was -0.12%. So, “positions *opened* after accounting events are significantly less profitable than pairs positions opened in non-event periods” (Papadakis & Wysocki, 2007).

After that, the researchers analysed the impact of delaying the closing of a pair trade by 5, 10, and 15 days, after an accounting information event unless the date of closing the position occurred at the end of the trading period. The results of this strategy show that by delaying the closing of a pair after an earnings announcement the profitability would increase by 0.08%, 0.17%, and 0.22% for delaying by 5, 10, and 15 days respectively. The same strategy but applied after a clustered forecast analyst would also improve the profitability of the strategy by 0.02%, 0.15%, and 0.31% for delaying by 5, 10, and 15 days respectively. They also computed the profitability of delaying the closing of a pair of non-event dates and the results have shown that delaying the close wasn't profitable for a non-event date. So, the spread in stock prices provoked by an accounting information event can provide an increase in the returns of the strategy when delaying the close of a position.

To better address the problem of the bid/ask bounce, Figueiredo et al. (2017) developed a strategy that uses the bid/ask price instead of the closing price of each stock and then compared the returns obtained with the strategy of Gatev et al. (2006) to understand if the “one-day waiting” rule is capable to eliminate the biased returns. Figueiredo et al. (2017) applied the strategy to the period between January 2017 and June 2016 and created a portfolio of 20 pairs with the closest closeness measure, to replicate the Gatev et al. (2006) strategy. Besides that, it was applied a liquidity restriction on the strategy, in which only stocks that had at

least two hundred trading days in the formation period were available to the strategy. The results obtained in this paper show that without the rule of “one day waiting” the 6 months return of the strategy was 40.90% and, after applying the rule, the returns dropped to 16.39% however, in the strategy of Figueiredo et al. (2017) the 6 months return dropped to 12.89%. So, even though the rule applied by Gatev et al. (2006) significantly reduced the biased returns of the strategy it continues to present biased returns.

Although the strategy of Figueiredo et al. (2017) considers the bid/ask price of stocks, there is still a weakness in the strategy because it doesn't consider the intraday fluctuations of the stock prices. To better understand if the intraday fluctuations can provide better performance to the strategy, researchers like Kim (2011) introduced modified pairs trading using high-frequency data instead of trading prices of stocks because that considers the 30-minute interval of the intraday stock prices. Kim (2011) defended that using high-frequency data can bring advantages to the strategy since investors can have a higher information ratio than with trading prices and investors don't miss intraday signals to open or close positions. The results of Kim (2011) showed that an intraday Pairs Trading strategy can have an average rate of return of 0.68% so, it could be advantageous for investors to consider using high-frequency data. Using this new strategy and high-frequency data could be a better fit for the research question addressed however, the database used in our paper doesn't allow access to the intraday prices for the period in analysis.

Do & Faff (2010) sought to extend the approach and the financial data used by Gatev et al. (2006) to investigate the continued profitability of the Pairs Trading strategy over a period spanning 1963 to 2011 after accounting for trading costs such as commissions and short-selling fees. Their study revealed that while the strategy remained profitable during this period, the inclusion of trading costs significantly reduced its profitability compared to the results obtained without considering these costs. Additionally, their analysis indicated a declining trend

in the profitability of the strategy in recent years. Notably, the researchers found that the Pairs Trading strategy was more effective in situations of strong market downturns, such as the global financial crisis of the late 2000s, which is consistent with those of other scholars in this field.

Andrade et al. (2004) conducted a study on the Taiwan Stock Exchange one year before Gatev et al. (2006) strategy and made two noteworthy conclusions. Firstly, they discovered that pairs trading strategy between the years 1994 to 2002 in this market had an annual excess return of 10.18%. Additionally, they found that the returns generated by this strategy were not explained by exposure to known sources of systematic risk. Andrade et al. (2005) concluded that the pairs trading strategy had a low-risk exposure because the stock in each pair has similar factor loadings.

The choice of the most appropriate approach for analysing the research question concerning the Pairs Trading strategy is a challenging task, given the diversity of available approaches. To address this issue, Krauss (2017) conducted a literature review on the different approaches of pairs trading, with univariate and multivariate strategies. To begin, the first approach introduced was the Distance approach, the strategy introduced by Gatev et al. (2006) and has already been discussed. This approach offers numerous advantages such as ease of application since it is an economic model-free approach, which avoids model misspecifications and can generate statistically significant excess returns. Krauss (2017) highlighted a noteworthy study conducted by Huck (2013) that replicated the distance approach but applied it to the 500 constituents of the S&P500. The main objective of this paper was to discover if, by changing the formation period of the approach, which is commonly 12 months, the profitability of the strategy would change. Given that, the researcher constructed 4 portfolios with different formation period lengths of 6, 12, 18, and 24 months. Unexpectedly, the results have shown a decrease in the excess returns on the 12-month formation periods and a strong positive return in the remaining periods. Nevertheless, Huck (2013)

questioned whether these results were sufficient to challenge the notion of an "optimal" formation period length.

Krauss (2017) discussed three additional approaches to the pairs trading strategy, including the cointegration approach, the time-series approach, and the stochastic control approach. The cointegration approach. Introduced in Vidyamurthy (2004) is commonly known for being more econometrically reliable as the stocks are subjected to cointegration tests in the formation period. This approach involves three major steps: preselecting pairs based on statistical or fundamental similarity measures, measuring the traceability of the stocks with an Engle-Granger cointegration test, and, in the last step, using the Gatev et al. (2006) threshold rule to generate trading signals.

The time-series approach, on the other hand, differs from the previous approaches in that it generally ignores the formation period, relying instead on time-series analysis to generate trade signals. The approach has been criticized for its inflexibility since it is only applicable to a small group of securities. One of the most cited papers on this approach is Elliott et al. (2005).

The stochastic control approach also ignores the formation period, as the previous approach, and aims to identify the optimal investment in the two legs of a pair. In a major study on this approach by Yang et al. (2007) it was found that Gatev et al. (2006) strategy had a superior Sharpe ratio when there is a low rate of mean-regression because the strategy is rarely invested in the spread and. In contrast, with the increasing rate of mean regression, the risk of bankruptcy also increases, limiting the potential for higher profits. The stochastic control approach outperforms the distance approach significantly when the mean regression is higher.

2. Methodology

2.1. Data Description

The first step in implementing a Pairs trading strategy is to define the financial market where the data will be extracted and the period of analysis. The financial instruments could be used from every financial market but, since companies from the same industry and different locations are subjected to local legislations, a pair of stocks from companies of different financial markets will be subject to different risks. Therefore, to mitigate these inherent risks, we have chosen to use data from the S&P500.

The S&P500 is an index that tracks the market capitalization of the 500 U.S. companies that have the biggest market cap. These companies are the best representation of each industry in the U.S. market, and, by that, analyzing the implementation of a Pairs Trading Strategy in them, is a good proxy for the implementation of the strategy in different companies from the same industry.

Having selected the data, the next step is to define the period of analysis. Our analysis covers the period from 2010 to the end of December 2020. All the data was obtained from Refinitiv Eikon, and we constructed a sample with all the constituents of the S&P500 at the beginning of each formation period so, we have a sample of stocks at the beginning of every formation period.

After combining all the samples at the beginning of every formation period, our samples consist of 708 companies, that were used during the period of analysis, as shown below.

Sector	Number of Companies	%
Technology	104	14.69%
Health Care	87	12.29%
Consumer Discretionary	87	12.29%
Financials	86	12.15%
Industrials	83	11.72%
Energy	57	8.05%
Consumer Staples	50	7.06%
Materials	43	6.07%
Communication Services	39	5.51%
Utilities	38	5.37%
Real Estate	34	4.80%
Total	708	100.00%

Table 1: Sample Distribution by Industry

2.2. Pairs Formation

As referred previously, the Pairs Trading strategy involves two critical periods that must be taken into account. The first period, referred to as the formation period, is dedicated to analyzing the stocks' prices and aggregating them into pairs based on their closeness measure. The second period, which lasts for one semester only, involves trading the pairs of stocks by opening and closing positions according to the performance of each pair. Our formation period will last for 12 months, and each formation period will begin every six months after the start of the previous one to avoid overlap with another trading period. During the analysis period, we will have 20 periods of formation and 20 periods of trading, each lasting twelve and six months, respectively. Figure 1 below illustrates the process of creation formation periods and trading periods.

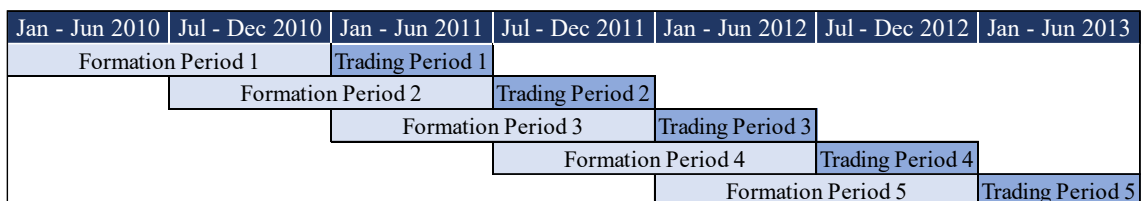


Figure 1: Example of creation of pair formation and trading

To ensure the liquidity of the stocks chosen in each formation period, the first step in the strategy was to eliminate any company that did not present, at least, 200 trading days during the formation periods. However, to guarantee a more realistic approach to the market condition, the number of trading days during the trading period was not filtered, as an investor cannot predict the liquidity of each stock during that period in a real application of the strategy.

Furthermore, another condition applied to the sample when computing all the possible pairs at the beginning of every formation period was that the pairs could not repeat the same combination of companies. For instance, if we have a pair with the stocks AB, we will not need the pair BA. The formula used to calculate the number of possible pairs in each formation period was:

$$\frac{n!}{2*(n-2)!} \quad (1)$$

Where “n” represents the number of firms in the sample. The resulting number of pairs for each formation period is presented in Table 2 below.

Formation Period	Stocks	Possible Pairs
Jan/2010	487	118,341
Jul/2010	486	117,855
Jan/2011	486	117,855
Jul/2011	487	118,341
Jan/2012	489	119,316
Jul/2012	490	119,805
Jan/2013	490	119,805
Jul/2013	489	119,316
Jan/2014	494	121,771
Jul/2014	490	119,805
Jan/2015	484	116,886
Jul/2015	477	113,526
Jan/2016	481	115,440
Jul/2016	486	117,855
Jan/2017	488	118,828
Jul/2017	488	118,828
Jan/2018	487	118,341
Jul/2018	480	114,960
Jan/2019	489	119,316
Jul/2019	488	118,828

Table 2: Possible Pairs in each Formation Period

2.3. Closeness Measure Computation

To compute the closeness measure for the distance between stocks, we followed the methodology presented by Gatev et al. (2006). We used the adjusted (for splits and dividend and/or capital gain distributions) close price for each stock and then we normalized the returns for each formation date with the formula presented by Papadakis & Wysocki (2007):

$$P_t^A = \Pi_{T=1,t}(1 + r_t^A) \quad (2)$$

where r_t^A is the daily return, adjusted for splits and dividend and/or capital gain distributions, at the end of the day t for stocks A and P_t^A is the normalized stock price for stock A on that day.

After that, we computed the distance measure ($Closeness_{AB}$) introduced by Gatev et al. (2006) with the time series of the normalized stock price deviations. This measure is calculated by summing the square deviations of the normalized stock price deviations as shown in the Formula 3 presented below:

$$Closeness_{AB} = \sum_{t=1, \#TradingDays} (P_t^A - P_t^B)^2 \quad (3)$$

The computation of the closeness measure for each formation period was performed using Excel and a Python code developed in the process of the elaboration of our paper. After ordering the pairs of stocks from the lowest $Closeness_{AB}$ value to the highest, we ended up with a list of pairs where the first pair contained the pairs of stocks with the highest proximity relation between the normalized stock prices. In the following figure, we present an example of the normalized price returns of a pair with the lowest Closeness measure in a formation period.

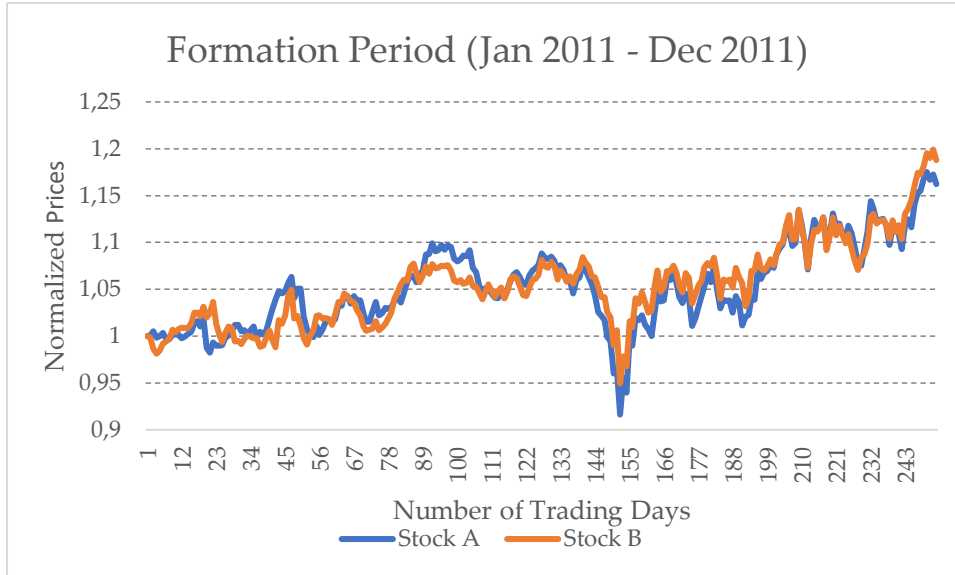


Figure 2: Example of the normalized price returns in a formation period

As it is possible to observe in the figure above, in the formation period from January 2011 and December 2011 the stock prices presented similar variations.

2.4. Triggers and trading signals

Following the conclusion of the formation period, the formation period, the pairs trading strategy continued by adopting the approach proposed by Gatev. et al. (2006). Specifically, positions were established by taking a long position on the undervalued stock and a short position on the overvalued stock, upon the attainment of a trigger point. The trigger is defined as the point at which the difference between the normalized daily stock prices during the trading period exceeds more than two standard deviations computed over the 12-month formation period.

$$Trigger^{AB} = \pm 2 * stdev(P_t^A - P_t^B) \quad (4)$$

When the difference between the normalized prices is higher than the trigger point (4), $|P_t^A - P_t^B| > Trigger^{AB}$, it signals an opportunity to open a position.

The two positions are determined by the comparison of the normalized daily stocks prices as follows:

1. If $P_t^A > P_t^B$ we have a short position on stock A and a long position on stock B;
2. If $P_t^B > P_t^A$ we have a short position on stock B and a long position on stock A.

Once the signal is generated, we only open the position on the following day, adhering to the “one-day waiting” rule introduced by Gatev et al. (2006). Since we are using the adjusted close prices in our sample, opening, and closing a position one day after the signal was given reduced the biased returns provoked by the bid/ask bounce discussed before. Figure 3 presents an example of the trading period of a pair.

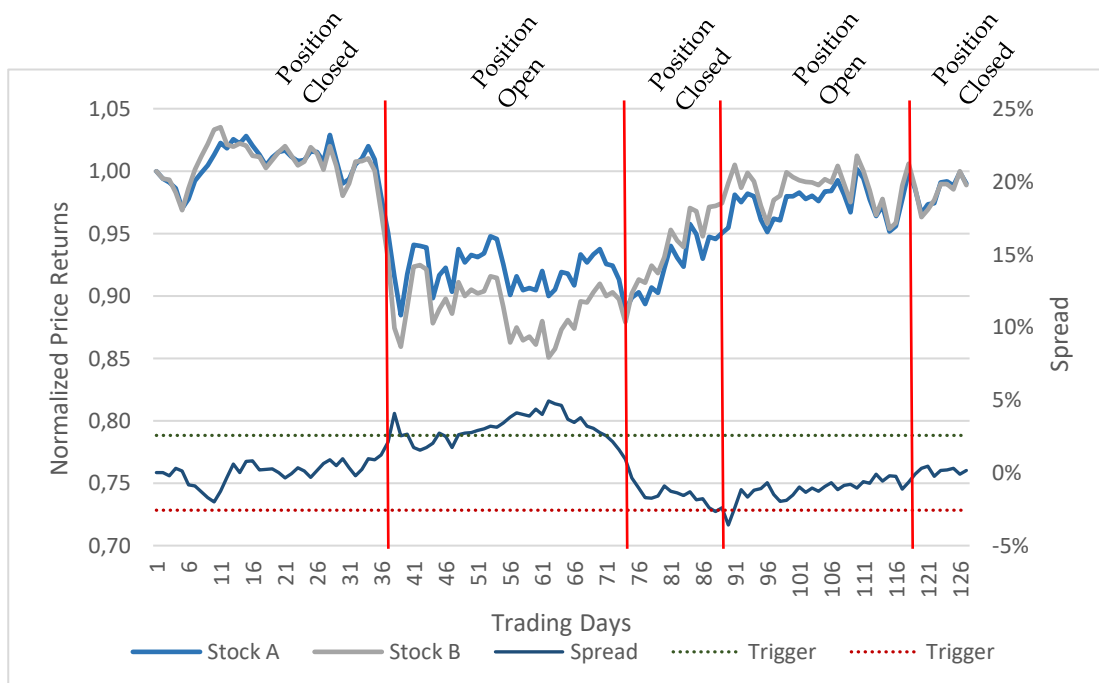


Figure 3: Example of the trading period of a pair

The figure presented above depicts the trading period and demonstrates how trade occurs during this period. The normalized adjusted prices of each stock are represented by Stock A and Stock B respectively. The spread is the difference between the normalized prices of the two stocks, while the two trigger lines

represent the negative and positive values of two historical standard deviations. A trading signal is given whenever the spread breaks out of the trigger level, and a position is opened accordingly.

For instance, in the example provided example, the spread is computed by $(P_t^A - P_t^B)$. When the spread exceeds the positive trigger, it indicates that stock A is overvalued while stock B is undervalued. Therefore, a short position is opened in stock A and a long position in B. The position is then closed when there is a convergence of the normalized prices, which occurs, if $P_t^B > P_t^A$ and, at that point, the spread changes its signal. In the example provided, there was a second trade with a different signal than the first one, but the same methodology was applied. In our strategy, a position is closed when the normalized stock prices converge or on the last day of the trading period, which was not necessary in this example.

2.5. Excess Return Computation

To calculate the return of each opened position, we employed the formula (5), which is given by:

$$\left(\frac{\text{Open Price}_L - \text{Close Price}_L}{\text{Open Price}_L} + \frac{\text{Close Price}_S - \text{Open Price}_S}{\text{Open Price}_S} \right) \quad (5)$$

However, two limitations may arise while computing the returns for our position. The first limitation occurs when a pair does not close until the last day of the trading period, after opening. In this scenario, we will follow the methodology proposed by Gatev et al. (2006) and assume that we close the position on the last day of the trading period. The second limitation arises if a company is delisted from the stock exchange. In such cases, we will use the last available price to compute the returns.

2.6. Stop Loss Strategy

To investigate the performance of the pairs trading strategy in a specific industry or cross-industry approach, this study employs a portfolio construction method. The process involves aggregating stocks based on the Bloomberg Sector Classification and creating a portfolio that consists of five pairs of stocks for each industry. Additionally, the study implements a stop-loss strategy, previously used by Ribeiro (2015), to mitigate risks and enhance profitability. However, unlike Ribeiro (2015), who used two types of stop losses, namely, stop losses per consecutive day and stop loss per percentage loss, this study only utilizes three stop-loss strategies per percentage of loss.

The computation of the stop loss is based on a formula similar to Formula 5, where the closing price in the long position and the open price at the short position is used as if the position was closed on that day. After calculating the excess return, the position is closed on the next day if the percentage of loss exceeds 10%, 15%, or 20%, following the "one-day waiting" rule.

2.7. Limitations

The Pairs Trading strategy analysed in this study has several limitations that can potentially impact its applicability for investors. One major limitation is the lack of consideration for transaction costs when computing the returns of the strategy. Since each trade requires four transactions to be completed, higher transaction costs can significantly impact the strategy's returns. This means that financial institutions with lower transaction costs may have a better-performing strategy than institutional investors. Thus, the returns obtained from this study might be upward biased due to the absence of transaction and financial costs.

Another possible limitation when applying this strategy is related to the liquidity of the market. In this study, we utilized a dollar-based payoff strategy, where one dollar was invested long on the undervalued stock and one dollar short on the overvalued stock. However, the amount of investment is up to the investor, and if the amount invested is significant, the market may not be able to fulfil the investor's order, potentially resulting in market impact costs.

The last limitation of this strategy is associated with the closed price used to compute the returns. In practice, bid-ask spreads may affect an investor's ability to purchase the stock at the triggered signal, potentially leading to suboptimal returns. Additionally, private investors may not have the same level of market access as financial institutions, and delayed trading prices may impact the strategy's performance.

In light of these limitations, potential users of the Pairs Trading strategy should be aware of these issues and attempt to mitigate their impact. This could involve considering transaction costs and market impact costs when computing the returns of the strategy, utilizing a trading algorithm to automate trades and reduce market impact costs, and utilizing real-time or high-frequency data to inform trading decisions.

3. Results

3.1. Unconditional Pairs Trading Strategy

This section presents an analysis of the performance of the unconditional pairs trading strategy originally proposed by Gatev et al. (2006). In order to compare the results obtained from our strategy, we used the S&P500 index as a benchmark for unconditional pairs trading. While a portfolio of 20 pairs would provide more diversification for the investor, the objective of this paper is to assess the impact of the strategy on different industries in our sample, and thus, we opted to create portfolios consisting of the top 5 pairs and compare their performance.

Initially, we computed the top 5 pairs without any restrictions, as implemented by Gatev et al. (2006), and the results of that portfolio are presented in Table 3. Subsequently, we applied the same approach but grouped the stocks by industry to compare their performance (Section 3.2).

	Top 5 Pairs	S&P500
6-month return (Average)	3.092%	6.636%
Total Return	61.83%	132.71%
Maximum	32.53%	21.86%
Minimum	-37.99%	-7.13%
Median	4.68%	6.30%
Kurtosis	1.461	-0.185
Skewness	-0.574	0.132
Standard Deviation	0.154	0.073
Sharpe Ratio (Rf as 0%)	0.201	0.912

Table 3: Results of the Pairs Trading Strategy against the S&P500

It is important to note that the performance of the unconditional pairs trading strategy may vary depending on the period and the stocks selected for the portfolio. In this particular period, the strategy did not outperform the benchmark, but it does not necessarily mean that it is always the case. However, the results suggest that the strategy may not be suitable for all investors as it

presents a higher risk and lower return compared to the benchmark, as we can see by the poor Sharpe Ratio (0.201) against the S&P500.

Although the performance proves to be positive, with an average return of 3.092% in each trading period, the S&P in that period presented an average return of 6.636% for a period of 6 months. Additionally, the negative skewness and high kurtosis suggest that the portfolio has a higher probability of experiencing large losses, which may not be suitable for risk-averse investors. Therefore, it is crucial to consider other factors such as industry classification to improve the pairs trading strategy's performance. The following sections will explore the impact of industry classification and strategy improvements on the strategy's performance.

3.2. Pairs Trading Strategy in each Industry

Before applying the strategy, we aggregated the stocks by utilizing the Bloomberg Sector Classification, as presented in Table 1. Subsequently, we applied the same methodology as before to each industry, yielding the results displayed in Table 4.

The findings presented in Table 4 indicate that the industrial sector outperformed other industries with an average 6-month excess return of 10.4%. On the other hand, the consumer staples sector exhibited the worst performance with an average 6-month excess return of -8.367%, followed by the consumer discretionary sector with an average 6-month excess return of -8.058%. Out of the 11 sectors analysed, only five were able to calculate an average excess return, and only four sectors - industrials, health care, technology, and real estate - demonstrated superior performance compared to the unconditional pairs trading strategy.

Upon comparing all 12 strategies, it was found that the unconditional pairs trading strategy had the lowest average number of days where pairs had an open

position, with an average of 45.63 days. This was in contrast to the Consumer Discretionary sector, which had an average of 63.93 days. However, on average, all the portfolios had a similar number of average open days in a trade.

	Top 5 Pairs	Industrials	Health Care	Consumer Discretionary	Technology	Utilities	Financials	Real Estate	Materials	Communication Services	Consumer Staples	Energy
6-month return (Average)	3.092%	10.401%	7.845%	-8.058%	3.666%	-2.251%	-0.648%	5.058%	-0.860%	2.162%	-8.367%	-0.460%
Total Return	61.834%	208.016%	156.891%	-161.165%	73.317%	-45.018%	-12.959%	101.161%	-17.205%	43.236%	-167.346%	-9.205%
Maximum	32.525%	56.887%	46.784%	65.309%	40.852%	26.140%	53.069%	44.115%	55.103%	39.054%	19.977%	87.023%
Minimum	-37.987%	-27.603%	-56.143%	-93.823%	-34.083%	-37.987%	-82.918%	-37.385%	-50.190%	-71.940%	-50.418%	-48.284%
Median	4.685%	6.913%	12.507%	-10.920%	2.981%	2.456%	5.014%	6.287%	-11.483%	1.285%	-6.881%	-9.908%
Kurtosis	1.461	-0.770	-0.473	-0.672	-0.654	1.071	2.107	1.391	-1.020	1.989	-0.215	0.556
Skewness	-0.574	0.149	-0.589	0.076	0.040	-0.724	-0.861	-0.223	0.367	-0.964	-0.672	0.902
Standard Deviation	0.154	0.238	0.282	0.429	0.211	0.148	0.288	0.171	0.331	0.257	0.193	0.337
Sharpe Ratio (Rf as 0%)	0.201	0.437	0.278	-0.188	0.174	-0.152	-0.023	0.296	-0.026	0.084	-0.434	-0.014
Average number of Trades	8.15	7.60	7.35	6.45	6.55	6.95	7.30	6.25	6.10	6.20	6.70	6.35
Average Return per Trade	0.379%	1.369%	1.067%	-1.249%	0.560%	-0.324%	-0.089%	0.809%	-0.141%	0.349%	-1.249%	-0.072%
Average Time Pairs are Opened (in Days)	45.63	53.27	49.99	63.93	59.27	50.08	53.18	51.98	56.20	59.34	58.26	55.50

Table 4: Results of a Pairs Trading Strategy for each Industry

To gain a deeper understanding of the unconditional pairs trading strategy, we conducted an analysis of the industry classification of each stock pair, as shown in Table 5. The results indicate that only 2 pairs out of the 100 pairs created during the trading period consisted of stocks from different sectors. This is significantly lower than the 20% of mixed sector pairs found in the top 5 pairs portfolio of Gatev et al. (2006). It is worth noting that utilities companies are known to have highly correlated prices, and thus the vast majority of pairs in our unconditional strategy were composed of utilities stocks.

Industry	Number of Pairs
Cross-Industry	2
Consumer Staples	2
Financials	14
Real Estate	4
Technology	1
Utilities	77
Total	100

Table 5: Pairs Distribution per Industry for the Unconditional Pairs Trading

Gatev et al. (2006) found that the utilities sector had the highest excess return, with a monthly average of 1.08%. However, our analysis yielded different results, as the portfolio composed solely of utilities companies had an average 6-month return of -0.648%. Despite the prevalence of utilities pairs in the top 5 pairs portfolio, which might have suggested a similar performance with the utilities portfolio, the strategy performed relatively better due to its diversified nature, resulting in an average 6-months return of 3.092%.

Additionally, it was observed that the industrials portfolio had the highest average 6-months return amongst all the portfolios, even though not a single pair of stocks from the industrial sector was included in the top 5 pairs portfolio. This may be because companies from the industrial sector may have a lower correlation in stock prices, resulting in a higher standard deviation of portfolio returns, which may not be appealing to every investor. However, the industrials

portfolio had the highest Sharpe Ratio (0.437) among all the portfolios, indicating that it had a higher return adjusted to risk than the other portfolios.

3.3. Introducing Stop Losses on the Strategy

Upon examining the performance of each portfolio, we noticed that a significant proportion of them had a negative skewness, which had a detrimental impact on the performance of our strategy. To mitigate our losses, we then implemented three stop-losses at 10%, 15%, and 20%, respectively, and evaluated the performance of the portfolios following the application of the stop-loss mechanism. The outcomes of the application of the stop-losses at 10%, 15% and 20% in our analysis are presented in Table 6, 7 and 8 respectively.

An interesting observation from the analysis of the strategy output is the increase in the number of trades that occurred in the trading period for all portfolios. However, more importantly, it is evident that in almost all portfolios, the performance after the implementation of the stop-loss strategy outperformed the Gatev et al. (2006) methodology. The only exceptions to this trend were the real estate portfolio with a stop loss of 15% (see table 7) and the materials industry portfolio with a stop loss of 20% (see table 8). The energy sector portfolio exhibited the most significant impact on performance, as it initially demonstrated negative performance with the simple strategy, but after the application of the stop-loss strategy, it became one of the top 5 profitable portfolios.

The observed increase in the number of trades per period did not adversely affect the average return per trade due to the application of stop losses, which yielded a higher return for investors, which is evidenced by the increase in return per trade. Additionally, the implementation of stop losses helped mitigate the risk associated with each strategy, as reflected in the improvement of the Sharpe Ratio for most of the portfolio.

Although the application of the stop loss improved the performance of the pairs trading strategy and some portfolios were able to achieve a higher return than the benchmark, namely the S&P500, all of the portfolios exhibited a higher standard deviation of returns than the S&P500. This suggests that although the strategy may be profitable, it poses a higher risk to investors. However, the higher risk associated with a pairs trading strategy does not necessarily lead to higher returns. In fact, the Sharpe Ratio, which measures the risk-adjusted return of a portfolio, was substantially higher for the S&P500 compared to the remaining portfolios.

Given that our research objective was to examine the performance of the pairs trading strategy in both a specific industry and a cross-industry approach, we will proceed with our analysis of the portfolio that has demonstrated superior performance. Specifically, the industrial sector portfolio has consistently exhibited an average excess return that surpasses our benchmark, the S&P500. Even though all strategies applied to this portfolio yielded a profitability increase of approximately 2%, we will continue our analysis with the 20% stop-loss strategy, as it has the highest return and Sharpe Ratio.

Moreover, the 20% stop-loss strategy presents another advantage over the other strategies, since it involves fewer trades. While transaction costs are not considered in our strategy, each trade requires four transactions (i.e., buying the undervalued stock, selling the overvalued stock, and subsequently closing the open positions), thereby implying that an increase in the number of trades would result in higher transaction costs, potentially impacting the strategy's performance.

	Geral	Industrials	Health Care	Consumer Discretionary	Technology	Utilities	Financials	Real Estate	Materials	Communication Services	Consumer Staples	Energy	S&P500
6-month return (Average)	4.079%	12.884%	9.005%	-6.139%	3.989%	-1.857%	0.875%	5.912%	-0.455%	4.188%	-7.692%	5.011%	6.636%
Total Return	81.576%	257.671%	180.091%	-122.772%	79.789%	-37.140%	17.499%	118.236%	-9.097%	83.766%	-153.849%	100.215%	132.710%
Maximum	32.525%	61.750%	52.929%	63.977%	40.852%	26.140%	54.705%	51.172%	56.542%	39.054%	19.874%	168.699%	21.860%
Minimum	-40.176%	-23.176%	-54.739%	-72.211%	-37.571%	-40.176%	-80.398%	-36.536%	-53.779%	-54.312%	-52.120%	-49.418%	-7.130%
Median	5.504%	7.258%	16.745%	-10.151%	3.959%	2.456%	6.232%	6.925%	-10.607%	5.837%	-6.305%	-10.080%	6.300%
Kurtosis	2.700	-0.872	-0.438	-0.955	-0.599	1.795	2.818	1.872	-1.135	0.598	0.666	6.657	-0.185
Skewness	-0.922	0.295	-0.590	0.282	-0.113	-0.701	-0.904	0.102	0.241	-0.570	-0.947	2.256	0.132
Standard Deviation	0.152	0.260	0.288	0.402	0.217	0.140	0.275	0.178	0.351	0.228	0.194	0.468	0.073
Sharpe Ratio (Rf as 0%)	0.268	0.495	0.313	-0.153	0.184	-0.132	0.032	0.333	-0.013	0.183	-0.397	0.107	0.912
Average number of Trades	9.00	9.60	9.55	10.05	8.50	7.80	9.10	7.30	8.30	8.65	8.90	9.65	-
Average Return per Pair	0.453%	1.342%	0.943%	-0.611%	0.469%	-0.238%	0.096%	0.810%	-0.055%	0.484%	-0.864%	0.519%	-
Average Time Pairs are Opened (in Days)	41.44	42.44	38.71	41.39	45.91	44.76	42.91	44.64	41.65	42.91	44.36	36.91	-

Table 6: Results of implementing a 10% Stop-Loss on each Portfolio

	Geral	Industrials	Health Care	Consumer Discretionary	Technology	Utilities	Financials	Real Estate	Materials	Communication Services	Consumer Staples	Energy	S&P500
6-month return (Average)	3.976%	12.969%	8.468%	-6.336%	3.645%	-1.760%	0.736%	5.051%	0.050%	4.183%	-6.963%	3.956%	6.636%
Total Return	79.522%	259.370%	169.354%	-126.716%	72.900%	-35.201%	14.718%	101.020%	1.000%	83.666%	-139.261%	79.117%	132.710%
Maximum	32.525%	69.372%	46.784%	63.397%	40.852%	26.140%	55.913%	44.115%	55.311%	39.054%	19.874%	150.973%	21.860%
Minimum	-40.427%	-23.968%	-55.054%	-75.289%	-36.330%	-40.427%	-82.482%	-37.409%	-52.529%	-57.904%	-51.798%	-49.469%	-7.130%
Median	5.504%	7.180%	17.533%	-10.851%	3.974%	2.456%	5.014%	6.378%	-10.480%	8.872%	-1.720%	-9.567%	6.300%
Kurtosis	2.631	-0.572	-0.466	-0.970	-0.528	1.825	2.884	1.495	-1.138	1.021	0.815	5.908	-0.185
Skewness	-0.910	0.416	-0.650	0.236	-0.066	-0.738	-0.911	-0.219	0.279	-0.722	-1.063	2.075	0.132
Standard Deviation	0.153	0.266	0.283	0.407	0.213	0.141	0.280	0.170	0.355	0.233	0.193	0.429	0.073
Sharpe Ratio (Rf as 0%)	0.260	0.487	0.299	-0.156	0.171	-0.125	0.026	0.297	0.001	0.180	-0.360	0.092	0.912
Average number of Trades	8.50	8.35	8.30	8.45	7.55	7.35	8.05	6.70	7.30	7.25	7.75	8.10	-
Average Return per Pair	0.468%	1.553%	1.020%	-0.750%	0.483%	-0.239%	0.091%	0.754%	0.007%	0.577%	-0.898%	0.488%	-
Average Time Pairs are Opened (in Days)	43.81	48.60	44.51	49.02	51.93	47.43	48.35	48.20	47.18	50.61	50.50	43.72	-

Table 7: Results of implementing a 15% Stop-Loss on each Portfolio

	Geral	Industrials	Health Care	Consumer Discretionary	Technology	Utilities	Financials	Real Estate	Materials	Communication Services	Consumer Staples	Energy	S&P500
6-month return (Average)	3.753%	13.133%	8.428%	-7.199%	3.546%	-1.914%	0.610%	5.173%	-1.166%	4.232%	-7.633%	3.802%	6.636%
Total Return	75.056%	262.658%	168.570%	-143.972%	70.917%	-38.277%	12.209%	103.453%	-23.316%	84.643%	-152.663%	76.047%	132.710%
Maximum	32.525%	71.287%	46.784%	65.453%	40.852%	26.140%	55.912%	44.115%	55.103%	39.054%	19.977%	155.662%	21.860%
Minimum	-39.193%	-23.009%	-56.143%	-85.253%	-36.749%	-39.193%	-81.282%	-35.922%	-53.840%	-58.255%	-53.925%	-49.428%	-7.130%
Median	5.504%	7.166%	17.665%	-10.860%	3.977%	2.456%	5.014%	6.287%	-11.483%	7.939%	-6.289%	-9.908%	6.300%
Kurtosis	2.100	-0.452	-0.428	-0.840	-0.564	1.382	2.673	1.304	-1.160	1.113	0.714	6.375	-0.185
Skewness	-0.771	0.477	-0.670	0.169	-0.047	-0.684	-0.849	-0.155	0.262	-0.748	-0.984	2.157	0.132
Standard Deviation	0.153	0.266	0.286	0.420	0.214	0.143	0.279	0.169	0.351	0.232	0.194	0.437	0.073
Sharpe Ratio (Rf as 0%)	0.246	0.493	0.295	-0.172	0.165	-0.134	0.022	0.306	-0.033	0.182	-0.392	0.087	0.912
Average number of Trades	8.40	8.10	7.75	7.55	6.85	7.25	7.75	6.35	6.60	6.90	7.30	7.40	-
Average Return per Pair	0.447%	1.621%	1.088%	-0.953%	0.518%	-0.264%	0.079%	0.815%	-0.177%	0.613%	-1.046%	0.514%	-
Average Time Pairs are Opened (in Days)	44.31	50.06	47.46	54.76	56.72	48.06	50.17	51.17	52.05	53.47	54.22	48.22	-

Table 8: Results of implementing a 20% Stop-Loss on each Portfolio

3.4. Detailed Results of the Strategy

In this section, we will provide a detailed analysis of the performance of our strategy, which is based on the industrial sector portfolio with a stop loss of 20%. To achieve this, we first divided the returns of the strategy into two categories: (a) the returns of positive and negative trades, and (b) the returns of trades that closed on the last trading day and those that closed before the last trading day.

Trading Period		Global Returns		Positive Trades		Negative Trades		Trades that closed before the last day		Trades that closed on the last day	
		Trades	Return	Trades	Return	Trades	Return	Trades	Return	Trades	Return
2011	1º Half	6	-18.284%	1	6.748%	5	-25.032%	1	6.748%	5	-25.032%
	2º Half	7	43.714%	6	52.257%	1	-8.543%	5	44.406%	2	-0.692%
2012	1º Half	8	17.029%	5	27.123%	3	-10.094%	4	26.195%	4	-9.165%
	2º Half	6	7.526%	4	18.763%	2	-11.237%	2	16.099%	4	-8.573%
2013	1º Half	5	5.978%	3	18.359%	2	-12.381%	2	10.067%	3	-4.089%
	2º Half	6	3.751%	4	26.440%	2	-22.689%	2	18.131%	4	-14.380%
2014	1º Half	9	30.497%	7	31.762%	2	-1.265%	4	23.144%	5	7.353%
	2º Half	10	-23.009%	6	22.899%	4	-45.908%	5	-22.928%	5	-0.081%
2015	1º Half	6	-21.154%	2	3.882%	4	-25.036%	1	3.757%	5	-24.911%
	2º Half	10	18.342%	8	52.643%	2	-34.301%	6	16.351%	4	1.991%
2016	1º Half	10	26.929%	7	51.722%	3	-24.793%	6	51.372%	4	-24.443%
	2º Half	10	47.241%	9	55.599%	1	-8.359%	7	50.209%	3	-2.968%
2017	1º Half	7	-19.873%	3	16.773%	4	-36.646%	3	-5.060%	4	-14.813%
	2º Half	7	25.811%	4	35.514%	3	-9.703%	4	35.514%	3	-9.703%
2018	1º Half	6	4.166%	3	16.584%	3	-12.418%	3	16.584%	3	-12.418%
	2º Half	14	56.887%	11	69.712%	3	-12.826%	10	69.246%	4	-12.359%
2019	1º Half	9	-18.515%	4	16.021%	5	-34.536%	5	-7.842%	4	-10.673%
	2º Half	8	6.806%	5	29.228%	3	-22.422%	3	-0.364%	5	7.170%
2020	1º Half	11	71.287%	8	121.689%	3	-50.401%	8	33.994%	3	37.293%
	2º Half	7	-2.471%	4	22.423%	3	-24.894%	3	-3.246%	4	0.775%
Results		8.10	13.133%	5.20	34.807%	2.90	-21.674%	4.20	19.119%	3.90	-5.986%

Table 9: Detailed results of Industrial Pairs Trading Strategy

Upon reviewing Table 9, it can be inferred that the number of positive trades (5,20) was significantly higher than the number of negative trades (2,90). It is important to note that while positive trades contribute to the excess return of the portfolio, they are not the sole determinant of the strategy's success.

The number of trades that closed before the last trading day (4.20) was also higher than those that closed on the last trading day (3.90). It is expected that the returns of trades that close before the last trading day are positive (19.119%) as the prices of the stocks converge, resulting in profits for the investor. In an unconditional pairs trading strategy, negative returns may only occur when one of the stocks in the pair is delisted. However, since the current strategy has a stop loss implemented when losses exceed 20%, any negative trades that occur before the last trading day can be attributed to hitting the stop loss, prompting the trade to close.

On the other hand, trades that closed on the last trading day resulted in an average 6-month return of (-5.986%). This negative return is expected as the pair is still open, and the normalized prices of the stocks have not yet converged. Consequently, when the pair are forced to close, the difference between the normalized prices is higher than when the positions were opened, leading to losses. Extending the trading period until the prices converge or reach the stop loss may alleviate this issue.

Trading Period		Trades that closed before the last day				Trades that closed on the last day			
		Positive Trades		Negative Trades		Positive Trades		Negative Trades	
		Trades	Return	Trades	Return	Trades	Return	Trades	Return
2011	1º Half	1	6.748%	0	0.000%	0	0.000%	5	-25.032%
	2º Half	5	44.406%	0	0.000%	1	7.851%	1	-8.543%
2012	1º Half	4	26.195%	0	0.000%	1	0.929%	3	-10.094%
	2º Half	2	16.099%	0	0.000%	2	2.664%	2	-11.237%
2013	1º Half	2	10.067%	0	0.000%	1	8.292%	2	-12.381%
	2º Half	2	18.131%	0	0.000%	2	8.309%	2	-22.689%
2014	1º Half	4	23.144%	0	0.000%	3	8.618%	2	-1.265%
	2º Half	3	15.245%	2	-38.172%	3	7.655%	2	-7.736%
2015	1º Half	1	3.757%	0	0.000%	1	0.125%	4	-25.036%
	2º Half	5	36.828%	1	-20.477%	3	15.814%	1	-13.823%
2016	1º Half	6	51.372%	0	0.000%	1	0.350%	3	-24.793%
	2º Half	7	50.209%	0	0.000%	2	5.390%	1	-8.359%
2017	1º Half	2	16.288%	1	-21.348%	1	0.485%	3	-15.298%
	2º Half	4	35.514%	0	0.000%	0	0.000%	3	-9.703%
2018	1º Half	3	16.584%	0	0.000%	0	0.000%	3	-12.418%
	2º Half	10	69.246%	0	0.000%	1	0.467%	3	-12.826%
2019	1º Half	4	16.021%	1	-23.864%	0	0.000%	4	-10.673%
	2º Half	2	18.934%	1	-19.297%	3	10.294%	2	-3.125%
2020	1º Half	5	84.396%	3	-50.401%	3	37.293%	0	0.000%
	2º Half	2	19.475%	1	-22.721%	2	2.948%	2	-2.173%
Results		3.70	28.933%	0.50	-9.814%	1.50	5.874%	2.40	-11.860%

Table 10: Detailed results of trades at the moment of closing position

Continuing, we will analyse the positive and negative returns of trades that closed before the last trading day, as well as the positive and negative returns of trades that closed on the last trading day.

After analysing the table presented, it can be observed that, on average, 0.5 trades are closed before the last trading day and exhibit an average loss of 20%, suggesting that these trades are triggered by the implemented stop-loss strategy. Interestingly, the vast majority of the portfolio's excess return stems from trades that are closed before the last trading day, with an average return of 28.933%.

When considering the trades that close on the last trading day, the results indicate that, on average, 1.5 trades are closed with an excess return of 5.874%. However, as anticipated, the most negative trades of the portfolio typically occur at the end of the trading period, with an average of 2.4 trades per trading period

exhibiting an average return of -11.86%. These negative trades further highlight the need to understand the potential benefits of extending the trading period until prices converge and a higher excess return is attained.

3.5. Market Neutrality

It has been previously asserted that pairs trading is a market-neutral strategy by several authors, including Vidyamurthy (2004) and Papadakis & Wysocki (2007). Additionally, Gatev et al. (2006) noted in their study that the Pairs Trading Strategy performed remarkably well during periods of poor performance in the US market.

To better comprehend the market-neutral nature of the pairs trading strategy, we conducted an analysis of the correlation between the strategy and the broader market. In this study, we utilized the S&P500 index as a benchmark since our research was based on index constituents. However, since our pairs trading strategy focused solely on industrial companies, we opted to compare their returns against those of the S&P500 Industrials, which tracks the performance of the industrial companies within the S&P500 index. The following graph illustrates the performance evolution of the strategy and the index benchmark for each trading period between January 2011 and December 20.

In the following figure, we can observe the evolution of the performance of the strategy and the index benchmark for every trading period from January 2011 to December 2020.

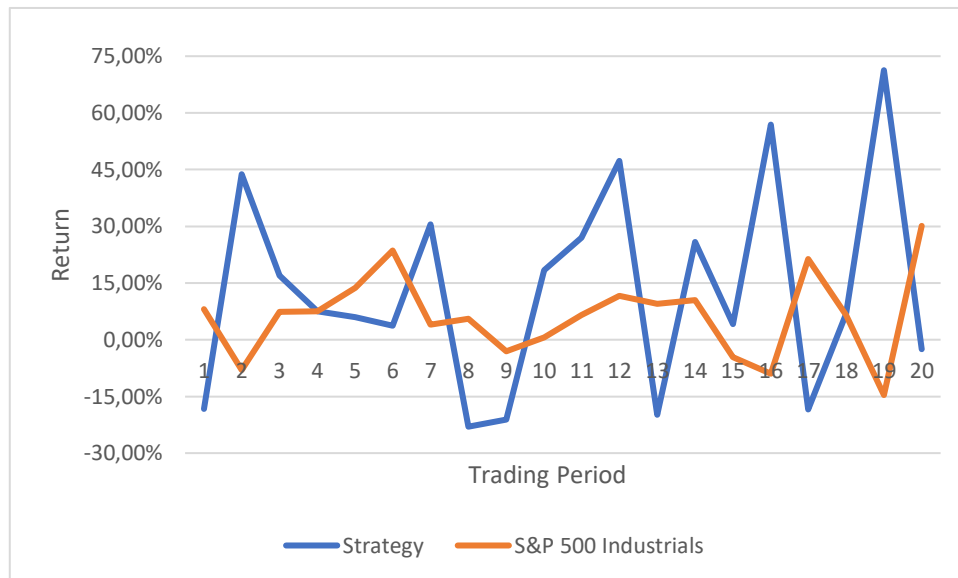


Figure 4: Returns Evolution of the strategy and the S&P500 Industrials

After observing the figure presented above, it becomes apparent that the returns of the strategy are more volatile in comparison to the returns of the S&P500 Industrials. It can also be noted that the returns of the strategy are not correlated with the benchmark returns, and upon computing the correlation coefficient, it is revealed that the correlation coefficient between the returns of the pairs trading strategy and the S&P500 Industrials is -51.08%. Although the results indicate a negative correlation between the market and the strategy, given the limited analysis period and the small number of pairs per formation period, it is not possible to conclude that the strategy is not a market-neutral strategy.

Another observation that can be made from Figure 4 is that, as previously stated by Gatev et al. (2006), during most periods where the market experienced negative returns, the strategy performed exceptionally well. As a matter of fact, the largest returns observed for the strategy occurred when the market had negative performance.

4. Conclusion

In summary, the objective of this study was to enhance and build upon the existing Pairs Trading strategy of previous works, such as Gatev et al. (2006), by investigating its profitability across various market industries. The study also aimed to evaluate the impact of cross-industry portfolios and pairs on the performance of the strategy. By conducting this research, the study aimed to contribute to a better understanding of the Pairs Trading strategy and its effectiveness in different market conditions.

To achieve our research objective, we implemented several modifications to the original strategy presented by Gatev et al. (2006). We restricted our analysis to a sample of companies within the S&P500 index from 2010 to 2020. However, this approach posed a challenge, as the sample of companies varied in each formation period. Besides that, we imposed a liquidity restriction on all stocks by excluding those that had fewer than 200 trading days during the formation period. Additionally, we aggregated stocks by industry, creating 12 portfolios to analyze, consisting of one portfolio with no industry restrictions and 11 portfolios for each industry.

In the last stage of our research, we implemented three stop-loss strategies. Specifically, we used a strategy that halted trading if we experienced a 10% loss, another that stopped at a 15% loss and a third that stopped at a 20% loss. We found that all three strategies improved the performance of the pairs trading strategy. Consequently, we concluded that including these stop-loss strategies in future pairs trading strategies would be beneficial.

While the initial implementation of the unconditional pairs trading strategy presented by Gatev et al. (2006) showed some positive returns, it was not able to outperform the benchmark. The strategy generated an average 6-month excess return of 3.092% and an average of 8.15 trades per trading period. However, the

benchmark performed better than the strategy, indicating that improvements could be made to the original approach.

After aggregating stocks into different industries and implementing the stop-loss strategies in our portfolios, the results indicate that the portfolio consisting of industrial companies produced the most favourable outcome during the analysed period. This particular portfolio exhibited an average 6-month excess return of 13.133% with an average of 8.1 trades per trading period. Despite the heightened volatility of this portfolio in comparison to the one without restrictions regarding industries, its Sharpe Ratio surpassed all other portfolios analysed. Meanwhile, the portfolio with no industry restrictions had the lowest standard deviation, implying that the portfolio encompassing stocks from all industries had the lowest volatility in returns relative to other portfolios. Consequently, we can infer that diversification in pairs trading may assist in mitigating industry risks but may not necessarily lead to the highest-performing portfolio.

In addition, our analysis revealed that the pairs trading strategy generated a positive excess return in only 7 out of the 12 portfolios analysed before applying a stop-loss strategy. Specifically, the negative excess returns were observed in the portfolios consisting of companies from the Consumer Discretionary, Consumer Staples, Energy, Financials, and Utilities sectors. Notably, our findings contradict those of previous studies that reported higher excess returns in the Utilities sector, suggesting a possible bias in our results. Nonetheless, we found that the Utilities portfolio presents a weak performance since 2017, leading to a weaker overall performance of the strategy.

Our analysis confirmed two key characteristics of the pairs trading strategy that were previously highlighted by other researchers. Firstly, we found that the strategy is market neutral, as indicated by a correlation coefficient of -51.08% with the market returns. This finding supports the notion that the strategy can generate returns independent of the market direction. Secondly, our results

indicate that the strategy has the ability to outperform the market during periods of negative market returns. This finding aligns with previous studies, which have noted the strategy's potential to generate positive returns even in a bear market.

Further research avenues could include the investigation of the impact of transaction costs, as well as the implementation of the strategy on bid/ask prices or high-frequency data. Such a study would be valuable in understanding whether our findings may be upwardly biased by employing the "one-day waiting" rule. Additionally, an analysis of the pairs that are compelled to close at the end of each trading period would be crucial for providing a more comprehensive evaluation of the strategy's performance.

Bibliography

- Andrade, S. C., Chang, C., & Seasholes, M. S. 2004. *Uninformed Trading and Asset Prices*. Working paper: U.C. Berkeley.
- Cavaglia, S., Brightman, C., & Aked, M. 2000. The Increasing Importance of Industry Factors. *Financial Analysts Journal*, 56(5), 41–54.
- Chan, L. K. C., Lakonishok, J., & Swaminathan, B. 2007. Industry classifications and return comovement. *Financial Analysts Journal*, 63(6), 56–70.
- Elliott, R. J., Van Der Hoek, J., & Malcolm, W. P. 2005. Pairs trading. *Quantitative Finance*, 5(3), 271–276.
- Fama, E. F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. In *Source: The Journal of Finance* (Vol. 25, Issue 2).
- Figueiredo, L., Cunha, R., & Alves, P. 2017. *Pairs Trading: Aplicação da Distância Mínima*.
- Gatev, E., Goetzmann, W. N., Rouwenhorst, K. 2006. *Pairs Trading: Performance of a Relative Value Arbitrage Rule: The Review of Financial Studies*.
- Huck, N. 2013. The high sensitivity of pairs trading returns. *Applied Economics Letters*, 20(14), 1301–1304.
- Jurek, J. W., & Yang, H. 2007. *Dynamic Portfolio Selection in Arbitrage*.
- Kim, K. 2011. *Performance Analysis of Pairs Trading Strategy Utilizing High Frequency Data with an Application to KOSPI 100 Equities*.
- Krauss, C. 2017. Statistical Arbitrage Pairs Trading Strategies: Review and Outlook. *Journal of Economic Surveys*, 31(2), 513–545.
- Papadakis, G., & Wysocki, P. 2007. *Pairs Trading and Accounting Information*.
- Ribeiro, R. 2015. *Pairs Trading: Strategy Refinements*. Universidade Católica Portuguesa.
- Vidyamurthy, G. 2004. *Pairs Trading: Quantitative Methods and Analysis*. John Wiley & Sons, Inc.