UNIVERSIDADE CATÓLICA PORTUGUESA

Pairs trading

Validation
of seasonality in of S&P500

Ricardo Filipe Duarte Matos

Católica Porto Business School

2024

UNIVERSIDADE CATÓLICA PORTUGUESA

Pairs trading Validation of seasonality in of S&P500

Master's Final Assignment – Witten Assignment
Presented to *Universidade Católica Portuguesa*
to obtain a Master's Degree in Finance

by

Ricardo Filipe Duarte Matos

Under supervision of
Prof. Paulo Alves

Universidade Católica Portugues
April, 2024

Agradecimentos

Gostaria de agradecer ao meu pai, à minha mãe e ao meu irmão por todo o apoio que me deram, um agradecimento excecional à minha mãe por todo o seu suporte.

Também agradeço à minha namorada pela ajuda, paciência e todo o amor demonstrado durante todo este tempo e, claro, aos meus amigos, esta tese não seria a mesma sem eles.

Gostaria de expressar o meu mais profundo apreço ao Prof. Paulo Alves por toda a ajuda que me deu durante este projeto e todos os conselhos fornecidos por ele durante este período.

Queria aproveitar este momento para expressar a minha mais profunda gratidão ao Prof. Doutor Pedro Silva pelo seu inestimável apoio e orientação ao longo da minha tese.

Um agradecimento especial e grandioso ao Prof. Paulo Alves por nos fornecer um programa de mestrado de excelente qualidade e prestígio.

Acknowledgements

I would like to thank my father, my mother, and my brother for all the support that they gave me, an exceptional thanks to my mother for all her support.

I also thank my girlfriend for the help, the patience and all the love shown during all this time and of course to my friends, this thesis would not be the same without them.

I would like to express my deepest appreciation to Prof. Paulo Alves for all the help he gave me during this project and all the advice provided by him during this period.

I wanted to take a moment to express my deepest gratitude to Prof. Pedro Silva for him invaluable guidance and support throughout my Thesis.

A special, big thanks to Prof. Paulo Alves for providing us with a master's program of excellent quality and prestige.

Resumo

O principal objetivo deste trabalho é a avaliação do impacto da sazonalidade dos mercados financeiros na implementação de uma estratégia de transação conhecida como *pairs trading*.

O trabalho inicia com um enquadramento teórico e uma revisão das principais estratégias e metodologias desenvolvidas. De seguida, replicamos a estratégia utilizada em Gatev, et al. (2006), com algumas diferenças significativas, utilizando as 500 maiores ações cotadas na *New York Stock Exchange*, mais precisamente o índice S&P500 rebalanceado mensalmente, para um período compreendido entre julho de 2005 até julho de 2023.

O nosso trabalho destaca-se pelos seguintes motivos: em primeiro lugar, é usual uma estratégia de *pairs trading* ter em conta o setor e o respetivo volume da empresa o que no nosso caso não se verifica, pois queremos verificar se esta premissa faz sentido, em segundo lugar iremos aplicar consecutivamente três versões da estratégia de *pairs trading*:_i) uma implementação base que servirá de *benchmark*; ii) uma estratégia de *pairs trading* com sazonalidade (entre maio e setembro não haverá transações); iii) semelhante à segunda, mas no período de sazonalidade iremos transacionar no S&P500.Os resultados destas diferentes estratégias foram de 72.66%, -30.77% e 61%, respetivamente. E um retorno médio por par de 0.07%, -0.04% e 0.08% por transação.

Palavras-chave: *Pairs trading*, Distância mínima, sazonalidade

Abstract

The main objective of this work is to evaluate the impact of financial market seasonality on the implementation of a trading strategy known as pairs trading.

The work begins with a theoretical framework and a review of the main strategies and methodologies developed. Next, we replicate the strategy used in Gatev, et al. (2006), with some significant differences, using the 500 largest stocks listed on the New York Stock Exchange, specifically the S&P500 index rebalanced monthly, for a period from July 2005 to July 2023.

Our work stands out for the following reasons: first, it is common for a pairs trading strategy to take into account the sector and the respective volume of the company, which is not the case in our study, as we want to verify if this premise makes sense. Second, we will consecutively apply three versions of the pairs trading strategy: i) a base implementation that will serve as a benchmark; ii) a pairs trading strategy with seasonality (no transactions between May and September); iii) similar to the second, but during the seasonal period, we will trade on the S&P500. The results of these different strategies were 72.66%, -30.77%, and 61%, respectively. And an average return per pair of 0.07%, -0.04%, and 0.08% per transaction.

Keywords: Pairs trading, Minimum distance, Seasonality

Table of contents

List of figures

List of tables

1.Introduction:

Statistical arbitrage strategies, such as the Pairs Trading, are quantitative investment strategies designed to generate economic returns when market inefficiencies occur (Gatev et al., 2006). They carry a strong message, as proving their validity can lead to the conclusion of market inefficiencies.

The concept of pairs trading was introduced to the American financial markets in the 1980s by Nunzio Tartaglia and his team of physicists, mathematicians, and computer scientists at Morgan Stanley. Their aim was to discover arbitrage opportunities in financial markets. The team utilized sophisticated statistical methods to develop advanced trading programs that automated market entries and exits to profitably exploit market inefficiencies.

This concept gained traction as pairs of stocks in which prices tended to move together were identified. Despite occasional divergence and distancing of prices, they tended to converge again.

In short, the strategy can be explained as follows firstly, the stocks must be correlated with each other, and they should be studied over a 12-month period to verify their correlation, after that when the price divergence between two stocks exceeds a predetermined value, a long position is opened in the stock with the relatively lower price, and a short position is opened in the stock with the relatively higher price. When the prices of the two stocks converge, the open positions are closed, resulting in a profit from either the rise in the price of the stock in which a long position was opened, the fall in the price of the stock in which a short position was opened, or both, it's important to mention that this strategy needs to be complemented by powerful and advanced information systems, as it is practically impossible to execute a pairs trading strategy in real-time.

Applying this strategy, Nunzio Tartaglia's team generated a profit of $50 million for Morgan Stanley. Later, as the team disbanded, the strategy spread and gained popularity, becoming a common investment strategy(Gatev et al., 2006).

In this work, we will address another relevant topic, which is seasonality. We will execute a pairs trading strategy based on three alternatives: 1) a simple pairs trading strategy, 2) a pairs trading strategy to verify the existence of seasonality, and 3) a pairs trading strategy to verify for seasonality and invest in periods when we didn't operate in the S&P 500 as an index. The last strategy aims to fill the gap identified in the second strategy, as only with the same trading period can comparisons be made. We will replicate the (Gatev et al., 2006)method with a different database, rebalancing the S&P 500 monthly. We will work with the top 10 pairs of stocks for 37 simulation periods, each lasting 6 months.

Finally, we will conduct a brief detailed analysis of the results obtained from our strategy, making the necessary comparisons with our hypotheses.

2. Literature Review

2.1 Efficiency Market Hypothesis

Efficient Market Hypothesis (EHM) describes a market where stock prices reflect all available information accurately, making it difficult to predict short-term price movements. EHM comes in three forms: Weak, semi-strong, and strong.

In the weak form, stock prices already reflect all past market data, meaning that analyzing historical trends won’t give you an edge in predicting future prices (Fama, 1965).

(Fama, 1965) argues the semi-strong form suggests that stock prices reflect all publicly available information, including news and financial reports. This form tests how quickly prices adjust to new information.

Taking it a step further, the strong form implies that stock prices reflect even insider or confidential information. The most contentious version raises concerns about whether certain investors have an unfair edge.

According to Fama (1965), stock values move swiftly in response to fresh information, making it very difficult to continually outperform the market. Using the idea of a "random walk," Malkiel (1973) underlined this point further by implying that fluctuations in stock prices are independent and unpredictable.

All things considered, the Efficient Market Hypothesis (EMH) contends that in a market that is efficient, stock prices promptly incorporate all available information, making it difficult for investors to regularly exceed the market.

2.2 Pairs trading

2.2.1 Introduction

Pairs trading originated in the 1980s by Nunzio Tartaglia and his team of mathematicians and physicists at the Morgan Stanley Group. Their main goal was to find arbitrage opportunities in the financial market to develop trading software with the purpose of achieving more profitable outcomes and profiting from market inefficiencies. Pairs trading is a quantitative trading strategy that consists of profiting from market inefficiency (Gatev et al., 2006)The motivation behind this strategy is the pursuit of arbitrage opportunities. Pairs trading consists of taking a long position in one security and a short position in another security in a predetermined ratio.

2.2.2 Motive and Foundations

Essentially, this strategy allows for the automation of market entries and exits to capture market inefficiencies. This strategy is commonly used by hedge funds and investment banks (Vidyamurthy, 2004).

Initially for this strategy, there must exist an analysis period, which is called the pairs formation period, considering pairs from the same sector and with similar or equal transaction volumes (Papadakis & Wysocki, 2007).

A pairs position is opened when the relative prices of the securities in a pair diverge, and secondly, the pairs position is closed to generate a future profit when there is a reversal of this relative price discrepancy. A pairs trading strategy encompasses two stock securities from the same financial sector, so both are exposed to similar systematic risks, and part of that risk can be mitigated by the non-systematic risk, which is the intrinsic risk to a company or sector, by choosing pairs arbitrarily without considering their specific sector. Based on these two elements of price divergence and reversal, we observe that accounting information events can affect the likelihood of observing trades and the profitability of an unconditional pairs strategy (Papadakis & Wysocki, 2007).

Pairs trading strategies can be considered neutral market strategies, as we are transitioning between a short position and a long one.

Sharpe (1987; 1991) mentioned the idea that investment activities can be analyzed as zero-sum bets between different investors. Although the article did not focus solely on pairs trading, it had an impact on the approach and understanding of the market for active strategies. This approach provides a more neutral stance in the market, allowing traders to profit independently of whether the market is rising or falling.

On the other hand, Vidyamurthy (2004) argues that "The specific price of the security is not important (...), it is only important that the prices of the two securities be the same," thus, humans do not like to trade against human nature, which desires to buy stocks after they appreciate, which can lead disciplined traders to take advantage of the mistakes of undisciplined market participants, resulting in higher profits and potentially leading us to an area of investigation to compare the discipline of traders across generations (Schizas, 2011).

2.2.3 Empirical evidences

We should analyze another important topic, the law of one price, as proposed by Blamont et al. (1986). According to their concept, "two investments with the same return in every state of nature must have the same current value." Essentially, this principle dictates that identical products should have the same price, excluding transportation costs, and should be denominated in the same currency. This ensures that no arbitrage opportunity exists.

Consider the following scenario to illustrate this principle: if we exchange Euro for dollars, then dollars for Japanese yen, and finally, Japanese yen back to Euro, the total value at the end of this transaction should be equal to the initial value. Failure to maintain this equality indicates an arbitrage opportunity, albeit momentary, which the market must quickly rectify. Chen and Knez (1995) further extend this argument, suggesting that closely integrated markets should align prices of similar returns closely.

Profiting from pairs trading would serve as evidence that financial markets are inefficient. Shiller et al. (1981) demonstrated that stock prices can be predicted in the long term, challenging the notion of market efficiency. Their seminal works include "Do stock prices move too much to be justified by subsequent changes in dividends?" (1981) and the book "Irrational Exuberance" (2000).

Although financial markets generally reflect all available information and adjust prices automatically, achieving profitability in pairs trading requires a failure of information, causing deviations in prices from the mean and standard deviation (Gatev et al., 2006). In an efficient market, assets with similar historical prices should have similar current prices. Therefore, when prices diverge, it indicates market inefficiency or a violation of the Law of One Price, providing an opportunity for arbitrageurs to apply pairs trading. Gatev et al. (2006) suggest that profits from pairs trading compensate arbitrageurs for enforcing the Law of One Price and ensuring market efficiency. They argue that "If the U.S. equity market were efficient at all times, risk-adjusted returns from pairs trading should not be positive."

Returns of approximately 11% were discovered by (Gatev et al., 2006) for the top-pairs portfolio, or the portfolio with the best performance. They concluded that, when examining the effect on portfolio performance, it is better to initiate positions around financial events, for example, when Coca-Cola announces a tender offer or takeover bid. However, according to Papadakis & Wysocki (2007), there is a higher average return of 20 to 30 basis points (0.02% to 0.03%) when position closure is delayed by 15 days and pairs cross the proximity measure.

Miao (2014) studied a pair trading strategy based on observing prices every 15 minutes using *high-frequency trading*[1] and achieved excessive returns of about 35%.

In conclusion, pairs trading allows for risk reduction by employing diversification and, consequently, potentially increasing profit. Some authors argue that in a world where stocks have similar daily prices over a certain period, they are exposed to similar systematic risks. Being long in one stock implies higher risk, while hedging in the other implies lower risk. Therefore, the average risk converges to a neutral market strategy (Papadakis & Wysocki, 2007).

---

[1] "High-Frequency Trading (HFT) is a type of algorithmic and quantitative trading which is characterized by short holding periods, specifically the use of sophisticated and powerful computing methods to trade securities rapidly. HFT has positions in equities, options, currencies and all other financial instruments that possess electronic trading capability, aiming to capture small profits and/or fractions of a cent of profit on every short-term trade." - (Miao, 2014, p. 97)

2.3 Market risk

The Capital Asset Pricing Model (CAPM), which was first presented in the 1960s, is an essential instrument in contemporary finance theory that lets us forecast how an asset's risk and expected return will interact. The CAPM is an independent development by Sharpe ((1964), Lintner (John, 1965), and Mossin (1966) that expands upon Markowitz's (1952) groundbreaking work.

Vidyamurthy (2004) considers that in the CAPM perspective, the expected return could be dissected into two factors: systematic risk (market) is also called non-diversification risk and non-systematic risk (asset). We can dilute the non-systematic risk by diversifying a portfolio. As the name suggests, non-systematic risk can't be mitigated through diversification investments, It is common caused by specific factors inherent to individual assets, which may bring huge losses to investors. The non-systematic risk is generally composed by the internal financial and operating conditions of the company (Deng & Yuan, 2021).

Thus, it is anticipated that financial markets exhibit trends, and consequently, if stationarity is absent in the financial market, it follows that pairs trading wouldn't be feasible.

The main idea behind CAPM is investors must be compensated by the time value and the risk they incur. The time value of an investment is typically compensated by the risk-free rate, often measured by government bonds. Regarding the compensation for risk, it is measured by the beta of each asset, a metric that reflects their risks. This measure takes into consideration the expected return and the volatility vis-à-vis of the market.

In order to understand both the profitability of the market and its volatility, it's common to use a market portfolio, as a proxy, for instance index S&P 500. This portfolio includes all available assets in the market.

In pairs trading with CAPM, we can conclude that pairs trading strategies are neutral strategies because they are associated with the concept of beta. The beta of a portfolio is a sensitivity measure correlated with the broad financial markets. Based on this we can conclude that all variations in the portfolio aren't explained by the market expected return.

According to Vidyamurthy (2004), when creating a portfolio that includes only long positions in assets, it's expected that this portfolio have a positive beta. This happens because positive market returns result in positive returns for assets, based on this, a portfolio that only has long positions has positive returns, otherwise, a portfolio with only short positions is likely to have a negative beta, therefore, to establish a neutral portfolio it's necessary to combine both positions (short and long) within the same portfolio, resulting in a beta equal to zero. For this reason, we can understand why pairs trading are considered neutral strategies vis-à-vis the market, because they always used a short position and a long position, simultaneously, in each pair.

Throughout the S&P500 examination, we will scrutinize various factors, including the CAPM, in addition to every index that is currently available in the financial market (including the S&P500, S&P100, PSI20, DAX40, and others). Determining whether our approach outperforms the performance of these indexes is one of the conclusions we aim to reach by the project's conclusion. Additionally, we will investigate seasonality because the decision to invest in government bonds or the risk-free rate will depend on seasonal periods. Our objective is to thoroughly investigate and confirm the potential for substantial financial gain.

2.4 Cointegration

In the previous decade, the concept of cointegration was increasingly applied in financial econometrics, in connection to time series analysis and macroeconomics (Alexander et al., 2002).

Cointegration is a very important topic in Pairs Trading - these two concepts are holding hands. Vidyamurthy (2004) provides the most cited work for cointegration-based pairs trading.

Cointegration was first introduced by Clive Granger (1981) and it measures the linear relationship between two assets in a given time frame. Two assets are correlated when their prices are close and move together in the same direction. However, it doesn't necessarily mean that assets have a long-term relationship Vidyamurthy (2004).

Finding cointegrated pairs is crucial when using a pair trading strategy over the time, prices are supposed to move in order to their equilibrium connection, even if occasional outliers are observed. It doesn't imply that cointegration is true in the near run.

It's commonly identified cointegration using statistical techniques, such as the Engle-Granger test (1987), The Johansen (1991) and the Philipps person test. These tests analyze if a long-term relationship exists between the assets, indicating if they share a common stochastic process (Krauss, 2017).

The pairs trading reversal-to-average method in is based on cointegration. The basic theory suggests that if two assets are cointegrated, the spread will tend to return to this equilibrium when it deviates from its long-term average, however, it's important to mention that it could have some outliers (Fama & French, 2004.).

In this study, we will not utilize cointegration process; since we are working with an index, each is S&P500, thus we don't face the issue of selecting and analyze highly correlated assets and their sectors to find the best pairs, instead, our strategy we will use all available combinations for each month and analyze which combination have the highest variance (minimum distance).

2.4.1 Cointegration tests

There are many tests that we could discuss, however, in this topic, I will focus on the most important ones. Cointegration tests is a good measure for finding the right pairs , providing signals about temporary deviations, risk management, the best timing to enter on the trade market and statistical modeling(Krauss, 2017).

Engle (1982) presented the concept of Autoregressive conditional Heteroscedasticity, better known as the ARCH model, these topic was a cornerstone in the field of econometrics particularly in modeling and forecasting time series data where volatility is observed.

The Engle-Granger test is a foundational technique for the analysis of economic and financial time series. This strategy is used to validate the existence og co-integration relationships between time series. This test has a very huge range of applications.

The main methods used for cointegration testing are the Engle & Granger (1987) or the Johansen (1991) test. Even though Vidyamurthy (2004) does not provide empirical results of the cointegration method, it is a framework that can form as a base for subsequent research(Krauss, 2017).

Caldeira & Moura (2013) used cointegration to select pairs on a Brazilian stock index. Using the trading rule proposed by Gatev et al. (1999), they found excess returns of more than 16% per year for the identified pairs, which shows that the cointegration approach can result in large profits.

2.4.2 The Fohansen

The Augmented Dickey-Fuller test (ADF) assumes that the error terms have constant variance and are statistically independent. The PP test, however, which was developed as a generalization of the ADF test, has a milder assumption regarding the error terms (Krauss, 2017).

The regression test takes the following $AR(1)$ form:

$$Yt = a0 + a1Yt{-}1 + et,$$

where the null is that $a1 = 1$ and the alternative that $a1 < 1$. Rejecting the null will indicate that Yt does not have a unit root and is therefore stationary.

2.5 Stochastic

Pairs trading is based on the theory that two assets with strong correlation must diverge for a while before getting the previous position/relationship. Stochasticity is significant in pairs trading as it explains how assets prices can change over time. Stochastic models could be used for forecast potential deviations from the historical relationship between those pairs, which can help traders make more accurate decisions. However, recognizing and keeping in mind the inherent unpredictability and volatility in price changes is essential.

To better understand the behavior of this model in a Pairs Trading strategy, it's necessary to introduce the concept of stationary time series, it's a stochastic process with mean and variance, constant and finite, and an independent autocorrelation in time.

Cointegration and correlation are related, but different concepts. A higher correlation of returns doesn't necessarily imply high cointegration in prices (Alexander, 1999).

The opposite of the description above, stock prices aren't stationary time series because the mean and variance fluctuate over time. Stocks follow tendencies, may exhibit seasonality (as I will study further), and prices are strongly correlated with the past. However, a multivariate model in which stock price series are cointegrated may reveal information about long-term equilibrium in the system (Alexander, 1999).

Another concept must be introduced, random walk is a classic example of a non-stationary time series, in other words, the upcoming price is determined by adding a term of error, thus we can confirm that no tendency or standardizing in the series is verified, since the mean and variance change over time.

Random walk is given by:

$$Y_t = Y_t - 1 + e2$$

And the first difference of a random walk become:

$$\nabla Y_t = e_t$$

And T could take $\{Yt : t = 0, \pm1, \pm2, \dots\}$ is as follows and where e_t is a stationary process.

A simulation of a random walk with 100 observations is found in figure 1 :

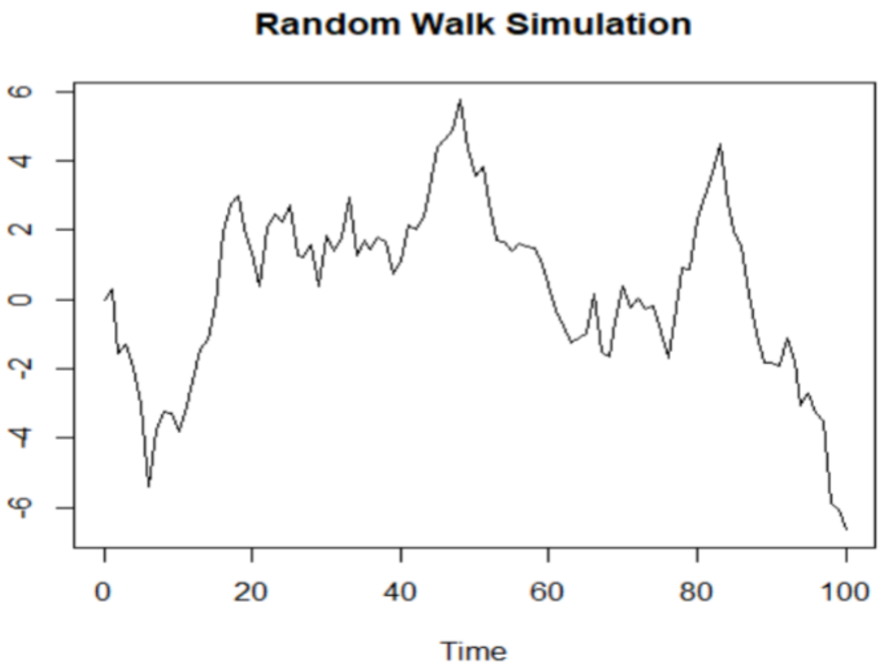

Figure 1 Random Walk Simulation

The Random walk is clearly related to non-stationary time series due to the random and unbalanced nature of this process (Alexander, 1999). Nevertheless, white noise is a concept that is linked with both stationary time series.

When we talk about white noise, we mean a series of Independent and equally distributed observations without any discernible patterns or temporal correlations. This stochastic or random component of a time series is often modeled by this method.

In pairs trading, we usually look for temporary deviations between an asset pair, as approached in its definition initially, the presence of white Noise in the price difference may indicate that these deviations are, usually of a random nature such as political, economic events, financial, social, sentimental and/or speculative not being part of a trend. Thus, white noise is a sequence of uncorrelated random variable with constant mean and variance (Vidyamurthy, 2004).

Finally, the concept of survivorship bias is interesting and important because we haven't used this concept throughout this work. Our strategy involves rebalancing the S&P500 every month. While in reality the S&P500 is rebalance every week, this approach is used to minimize our probability of error, however errors are still possible, for instance a company (let's call the company A) left the S&p500 during the 2nd week of month N and re-enter during the 4th week, at the end of the month we have in our sample the company A, nonetheless in reality we must derecognize the company in our formation period or whether we work on trigger period, we must close the position. Survivorship bias is a concept of analysis where the data are distorted, this means that it does not consider the samples that did not survive in the period studied (Brown et al., 1992).

2.6 Seasonality

Rozeff e Kinney (1976) were among the first to document the existence of patterns of monthly seasonality. The phenomenon was studied in an index of stock in the New York Stock Exchange (NYSE), in 1904 to 1974, and it was possible to conclude that in January the return was slightly higher (average of 3.48% in an average of 0.42% in the following months) (Lobão & Lobo, 2018).

The potential buyers seem to have "sold in May and gone away". According to Bouman & Jacobsen (2002) the month of May signals the start of a bear market, thus, investors are better off selling their stocks and holding cash. Bouman & Jacobsen (2002) argue that have two different ways to return to the market. First "…But remember to come back in September" The second is "But buy back on St. legger day".

Stock returns should be lower during May through September than during the rest of the year, although many Americans tend to be unfamiliar with it as we can see in Bouman & Jacobsen (2002).

Drogalas et al. (2007) examine the concept of "calendar anomalies", basically characteristics of Perfect & Imperfect Markets. There are many explanations for these phenomena, but none of these are statistically significant.

According to Bouman & Jacobsen (2002), stock returns are significantly lower from May until October compared to the remainder of the year. The author argues that there is no evidence to suggest that this effect can be explained by factors such as risk, cross-correlation between markets, or the January effect (Bouman & Jacobsen, 2002).

Bouman & Jacobsen (2002) studied 18 emerging markets from 1987 to 1996 and found a significant seasonal effects. However, they pointed out a lack of evidence for the January effect (Lobão & Lobo, 2018). Similarly, it is concluded that the low performance observed from May until October is not exclusive to developed markets but also occurs in emerging markets. For instance, Claessens (1995) found no statistical evidence of the January effect in a sample of 20 emerging stock markets (Bouman & Jacobsen, 2002).

This phenomenon tends to occur in high-performance months, particularly observed in small-cap stocks, which are companies with lower capitalization (Baker e Wurgler, 2007).Several reasons can be put forward, for instance:

- Tax loss-selling: toward the end of the calendar year, investors sell stocks at a loss to offset capital gains for tax purposes. The strategy starts in December, and it's normal to have negative/less returns on this month. On the other hand, once the new year begins, some investors reinvest in the same portfolio and consequently, stock prices rise.

- Window dressing: Selling poor-performance stocks and buy well-performing stocks at the end of the year, to make their portfolio looks better than it's for clients and stakeholders,
- Investor psychology: The start of the year brings a renewed sense of optimism and confidence about the financial markets, this positive sentiment makes the investors allocate more money into stocks driving up prices.

However, Bouman & Jacobsen (Bouman & Jacobsen, 2002) argue that the returns in the months of March (positive) and June (negative) were statistically significant during the period 1987-1995.

According to Bouman & Jacobsen (2002) While we observe huge returns in most countries from November to April, returns are very close to zero for the remainder of the year. In Europe, average returns over May to October do not exceed 2%; however, during the period from November to April, they exceed 8% in all European countries. As for the Americas, the difference is substantial, with average returns more than 5% higher between November and April than they are during the remainder of the year. Nonetheless, it is important to mention that the second biggest conclusion was that typical explanations, data mining, the January effect, and risk explanations, can be ruled out after examining a number of potential causes for this "sell in May" impact. Several less plausible hypotheses, such as changes in interest rates or volume, are also refuted. Importantly, the authors do not discover that the effect is produced by sector-specific characteristics.

On the other hand, according to Wachtel (1942), forecasts based solely on seasonal movements, without considering traditional cycles and trend analysis, have a high probability of success. Undoubtedly, considering the seasonal curve is essential for developing effective investment strategies.

As mentioned in previous chapters, the primary objective will be to compare a standard pairs trading strategy and a pairs trading strategy based on seasonality. The latter will primarily follow the idea discussed in this topic: buying from November to April and closing all positions by this date while investing in the S&P 500 from May to October.

3. Methodology

3.1 Data

The data used along this project has been imported from DataStream (Refinitiv Eikon/Thomson and Reuters) and consists of data from the top 10 pairs with the highest correlation (minimum distance) from January 1st, 2005 to December 31st, 2022, the data pertain to the S&P500, which was chosen due to it's the index with the highest performance in the world and the most complex one .

From DataStream, we obtained daily returns (including dividends) for the S&P500. We then extracted the monthly decomposition of S&P500. This information allowed us to rebalance the S&P500 at the end of each month and try to mitigate the survivorship bias because the index in reality is rebalanced on a weekly basis.

With this assumption, we can guarantee that the survivorship bias was eliminated. There is only one scenario in which we fail to eliminate the risk, but it is statistically challenging. Assuming a month consists of four weeks, and our rebalance is conducted monthly, we would mislead if, and only if, a company were to hypothetically be in the index during the first week and disappear in the third week, only to reappear in the fourth week. In our rebalance, the company would appear because it was within the index for the entire month; however, in reality, it was out of the index for two weeks. Therefore, we can assume that survivorship bias is virtually eliminated to 100%, in addition if the stock drops off the Index, but continues in the market we continue to use them for our project.

We didn't choose to use the bid/ask prices, as we preferred to conduct the study with adjusted prices. In doing so, we acknowledge a potential limitation in our dataset, as we may have overlooked certain nuances. This could be an area for improvement in future research. Nonetheless, we believe that this oversight does not significantly impact our results, given that it's common for the difference between the Bid and ask price to range from 10 basis points to 50 basis points (100bp = 1%) as well as the commission, we are operate under the assumptions of an ideal scenario without commissions.

3.2 Pairs formation

Utilizing a methodology similar to Gatev et al. (2006), we defined 38 formation periods and 37 simulation periods. However, we made our decisions based on our assumptions: 1) A simple pairs trading strategy, 2) A pairs trading strategy to verify whether exist seasonality, and 3) A pairs trading strategy to verify for seasonality and invest in periods when we didn't operate in the S&P500 as an index. Our study is based on a 6-month trading period. Each simulation period begins immediately after the close of the preceding period.

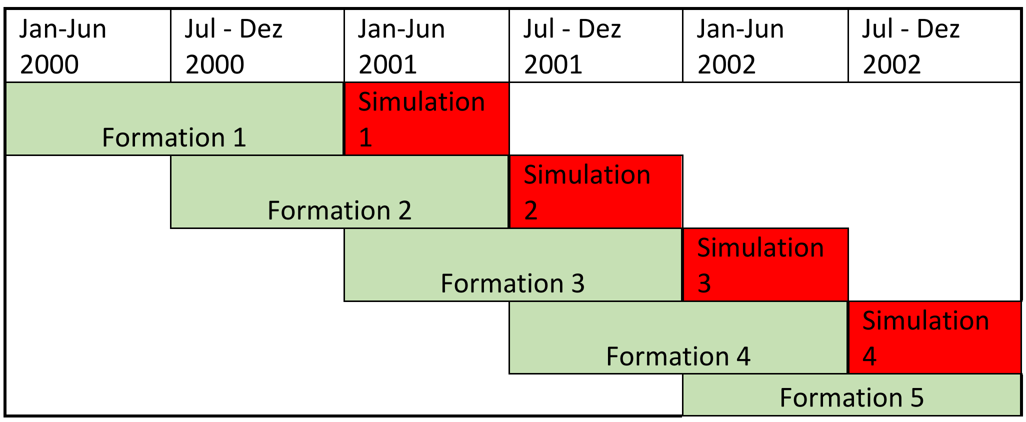

Figure 2 Example of structure of formation period

This methodology allows us to analyze 37 periods instead of 19, which we obtained if we made for each period one year.

Based on our project and assumptions we divided our project into 3 strategies.

The first strategy is a simple pairs trading strategy, where we conduct the formation period for one year and the trading period from January 1st to June 30th and from July 1st to December 31st.For the second trading strategy, we aim to investigate if there is seasonality in the financial markets. We assume that our trading period would occur between January 1st and May 31st, and between November 1st and December 31st. The trading period remains the same as represented above. However, we found it unfair as the trading periods above have an average of 131 trading days, while in our second assumption, the periods were shorter. Therefore, to balance our strategy, in months where there are no trades, we will invest in the S&P 500 index. Thus, in the second strategy, we gain a clear insight into whether there is seasonality compared to the first option, and in the third strategy, we can understand the expected return by equalizing the trading days.

3.3 Minimum distance

In the calculation of the minimum distance, once again followed the methodology of Gatev et al. (2006), where initially, we normalized the daily returns (including reinvested dividends) of each stock to day 1 of each formation period, using the following formula, taken from Papadakis and Wysocki (2007, p. 6).

$$P_t^A = \prod_{\tau=1,t} (1 + r_t^A)$$

Where is the r_t^A return (including reinvested dividends) of stock A at the end of day t, and P_t^A is the respective normalized return. Once the daily returns were normalized, we were able to calculate the minimum distance between each possible pair of stocks. To do so, we calculated the sum of the squared differences of the normalized returns for each pair, for every possible combination of stock pairs, for each day of the formation period. The following formula was utilized,

$$Minimum\ Distance_{A,B} = \sum_{i=1}^{n} \left(\frac{P_{a+1}}{P_a} - \frac{P_{b+1}}{P_b}\right)^2$$

Where $\frac{P_{a+1}}{P_a}$ and $\frac{P_{b+1}}{P_b}$ represents the normalized returns, for each day, of stocks A and B, respectively. Using the R studios program, we obtained a list of possible pairs combinations, where the first pair on the list is the one with the smallest distance between the returns, thus indicating a higher proximity relationship between the prices. The following figure represents the evolution of the normalized returns of one of the selected pairs of stocks during formation period 1.

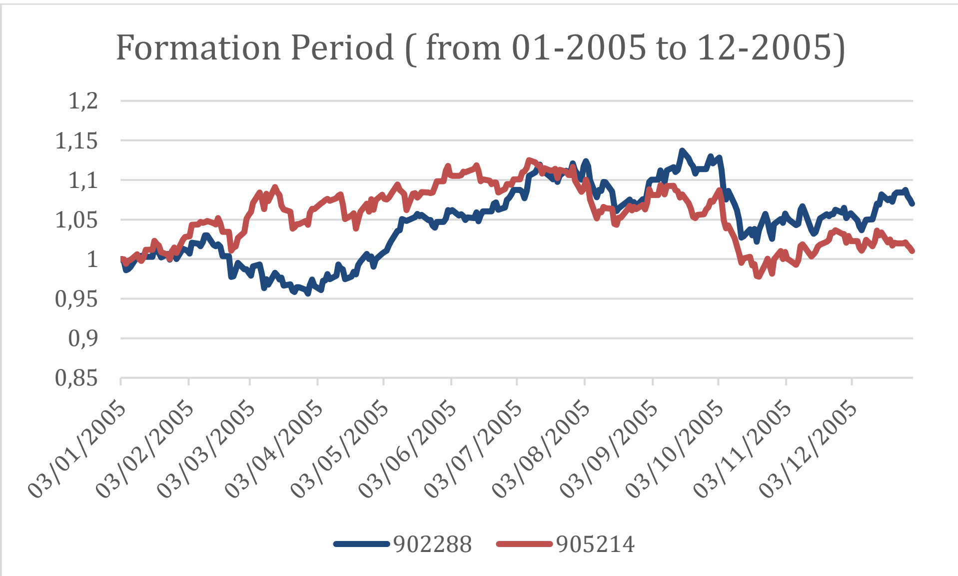

Figure 3 Evolution of formation period for the following pair 902288 – 905214

3.4 Buy and sold signal

An open long-short position is taken whenever the normalized returns of the stocks in each pair, selected during the simulation period (left side of the equation), diverge by more than two historical standard deviations, parameterized during the formation period (right side of the equation).

$$\left| \frac{P_{a+1}}{P_a} - \frac{P_{b+1}}{P_b} \right| > \pm\ 2 \times StDev\left(\frac{P_{a+1}}{P_a} - \frac{P_{b+1}}{P_b}\right)$$

When,

- Whether $\frac{P_{a+1}}{P_a} > \frac{P_{b+1}}{P_b}$, Long position is taken on b and a short position is taken on a;
- Whether $\frac{P_{a+1}}{P_a} < \frac{P_{b+1}}{P_b}$, long position is taken on a and a short position is taken on b.

The simulation period for the pair of stocks from the previous figure is shown in the following figure. As can be seen, we opened a position on day 21 and closed it on day 31 when the prices crossed. We then reopened the position on day 46 and closed it right away on day 47 when the prices crossed again. On the 50th day, we reopened the position, and at the end of the trading session, we were forced to close it. During this particular time frame, these two stocks returned

2.68%(6months).

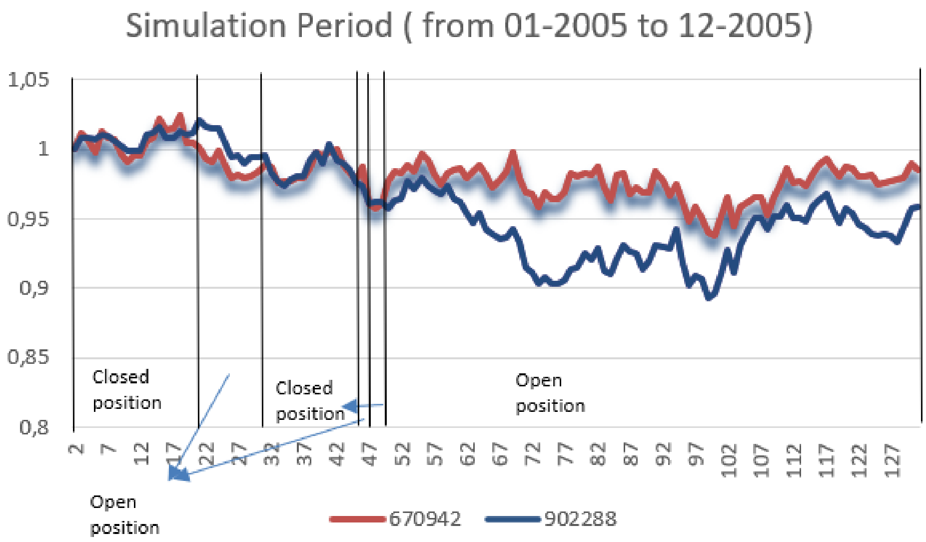

Figure 4 Example of simulation period for the following pair 902288 - 905214

3.5 Excess Returns

Regarding the calculation of returns themselves,

$$\frac{Open\ price_{\ Short} - closeness\ price_{\ Short}}{Open\ price_{Short}} + \frac{closeness\ price_{\ Long} - open\ price_{\ Long}}{Open\ price_{Long}}$$

We used the formula above to find the excess returns. In situations where the returns (normalized) of a pair of stocks, after positions have been opened, do not intersect until the end of the simulation period, the positions are closed on the last day of the period, and profits/losses are recorded.

4. Results

4.1 Results overview

Using our dataset composed of 36 simulation periods during 6 months, we build our strategy and applied it over the following 19 years, started in July 2005 and finished in July 2023.

In this chapter, we will present many important conclusions and some indicators that we consider important.

First of all, we will make a comparison of our three hypotheses. 1) A simple pairs trading strategy, 2) a pairs trading strategy to verify whether exist seasonality and 3) a pairs trading strategy to verify whether exist seasonality and invest in periods when we didn't operate in the S&P 500 as an index.

Results	Without Seasonality	With Seasonality	With Seasonality plus S&p500
Average	2.02%	-0.85%	1.68%
Max	60.23%	41.96%	41.94%
Min	-328.22%	-181.75%	-175.92%
Kurtosis	21.86	17.96	15.87
skewness	-4.28	-3.66	-3.44
Standard Deviation	62.52%	35.74%	36.08%
nº pairs/year	10	10	10
Total nº of trades	1005	632	740
Nº of trades / year	28	18	21
total return	72.66%	-30.77%	61%
Return/trade	0.07%	-0.04%	0.08%

Table 1 Results overview

First, as we can conclude in our project, the possibility to verify seasonality doesn’t exist as we can see by the result (total return), the project 1, had 72.66%, however it’s important to mention the project 2 had 62.88%, and the first project have more transactions than our second hypothesis. Thus, another measure that we need to take into consideration is Return per trade, our last hypothesis has the higher value (total return / nº of trades), and the number of trades in this hypothesis is the same value as the second hypothesis plus the nº of trade that we have to close plus 1 (for each period of trade) for invest in S&P500. Therefore, it’s important to interpret the number 0.08%, due to whether we have in consideration de cost of each transaction it be more favorable our third hypothesis because in average the cost of each transaction is between 0.01% and 0.05%, this assumption will be referred by me in the future improvements stage.

Another topic that deserves our attention is the following measures: minimum, maximum, and standard deviation, as we can perceive that our first hypothesis is riskier and more volatile. Therefore, for an investor with risk aversion, it will be important to take this into consideration, as our third hypothesis has more stable returns and consequently less volatile (lower) returns. For this same reason, it is important to consider this measure as well as, we may have a solution here for investors with more risk aversion.

We can reinforce what was said through the measure of kurtosis, as this measure is greater than zero indicating to us that the distribution of the data has heavier tails than the normal distribution, which means there are more extreme observations in the data (more outliers). If the value were less than zero, it indicates to us that the results have fewer outliers.

The minimum indicator is very high because we have into consideration bad times just like covid-19. There was an economic recession during COVID-19.

We can confirm this, for example, through our first hypothesis where the minimum value is -328%. Regarding our skewness, still for the first hypothesis, it indicates a negative asymmetry in the distribution of data. Therefore, we can conclude that the tail of the distribution extends much more to the left than to the right in order to the mean, which leads us to believe that there is a concentration of values above the mean.

These two indicators together can be interpreted as highly concentrated around the mean with extremely heavy tails and significant negative asymmetry. This may lead us to conclude a highly distorted sample as there is a large number of extreme values and a high concentration of values above the mean. This is an unusual and atypical distribution.

We can reinforce and validate our Assumption based on our third hypothesis, as we have more stable data during the observation period, although still with some volatility.

Another important conclusion, is after excluded bad times just like covid-19 and 2008 crisis:

without Bad Times	Without Seasonality	With Seasonality	With Seasonality plus S&p500
Average	16.87%	9.00%	11.07%
Max	60.23%	41.96%	41.94%
Min	-82.72%	-32.32%	-30.94%
Kurtosis	4.80	0.13	0.04
skewness	-1.83	-0.60	-0.64
SD	28.51%	17.68%	17.31%
nº pairs/year	10	10	10
Total nº of trades	780	484	565
Nº of trades / year	58	36	42
total return	455.44%	242.89%	299.01%
Return/trade	0.58%	0.50%	0.52%

Table 2 Results after excluded bad times

As we can see by the last figure, the indicator average improves more than 8 times, the standard deviation as we expected decreased, and the total return improves more than 6 times in our first hypothesis.

However, it's important to mention that the alternative "with seasonality plus S&P500" is the best because we don't mention the cost of transaction per trade, thus, it's important to consider this kind of costs, in this hypothesis we have fewer trades and a return per trade close than our first hypothesis.

Finally, we will conduct one last analysis, which will involve comparing the different performances graphically.

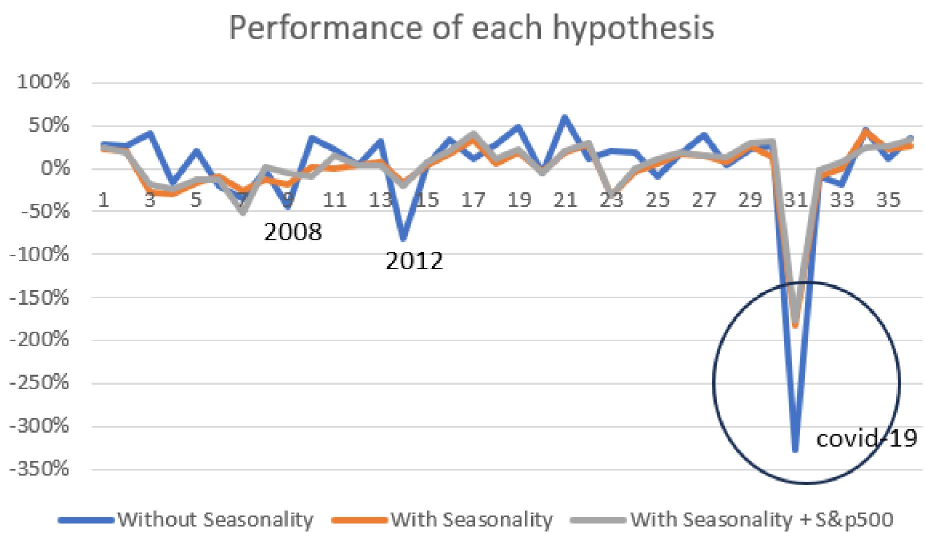

Figure 5 Performance of each hypothesis

As we can see and as indicated in the above graph, during the period of 2008 where there was an economic crash and times of economic recession, we observed a decline in the performance of all 3 hypotheses, with more emphasis on the normal trading strategy hypothesis. Secondly, the year 2012 was a year of downturn, due to economic uncertainty. The US had recently emerged from the recession it had gone through, and there was still no comfort among investors. Another reason was the sovereign debt crisis in Europe and also the elections adding to the economic uncertainty scenario. Another period of downturn, negatively emphasizing the non-seasonal pairs trading strategy, was the year of COVID-19. This graph is very illustrative of what standard risk is because we can perceive that the blue line is the one that deviates from the average, meaning the scenario with the most associated risk. Thus, reinforcing all our conclusions drawn so far.

4.2 Detail of results

Periods		Without Seasonality		With Seasonality		With Seasonality + S&p500		
Ano		N	Return (%)	N	Return (%)	N	Return (%)	Return of S&P500
2005								
	2º half	26	28.43%	10	22.49%	13	23.89%	1.4%
2006	1º Half	35	26.08%	29	20.96%	32	19.1%	-1.86%
	2ºHalf	40	40.59%	15	-27.66%	18	-19.41%	8.25%
2007	1º Half	30	-16.97%	28	-30.47%	31	-24.67%	5.8%
	2ºHalf	42	21.07%	18	-16.05%	21	-12.9%	3.15%
2008	1º Half	17	-20.25%	14	-9.95%	17	-12.73%	-2.78%
	2ºHalf	31	-34.87%	12	-25.59%	15	-51.38%	-25.79%
2009	1º Half	21	-1.56%	20	-13.37%	23	1.35%	14.72%
	2ºHalf	16	-45.42%	10	-17.7%	13	-5.34%	12.36%
2010	1º Half	23	35.17%	19	2.32%	22	-9.79%	-12.11%
	2ºHalf	29	23.31%	12	0.36%	15	14.95%	14.59%
2011	1º Half	26	3.16%	23	4.09%	26	3.76%	-0.33%
	2ºHalf	35	32.37%	14	7.69%	17	3.45%	-4.24%
2012	1º Half	17	-82.72%	12	-16.81%	15	-19.87%	-3.06%
	2ºHalf	28	4.05%	14	2.98%	17	6.66%	3.68%
2013	1º Half	34	33.59%	22	17.58%	25	19.97%	2.39%
	2ºHalf	23	10.88%	13	32.69%	16	41.94%	9.25%
2014	1º Half	41	28.15%	26	5.82%	29	10.45%	4.63%
	2ºHalf	35	48.9%	17	18.63%	20	21.66%	3.03%
2015	1º Half	21	-4.7%	20	-4.57%	23	-4.77%	-0.2%
	2ºHalf	43	60.23%	17	19.04%	20	20.41%	1.37%
2016	1º Half	28	11,00%	23	28.55%	26	30.44%	1.89%
	2ºHalf	16	21.44%	10	-32.32%	13	-30.94%	1.38%
2017	1º Half	23	18.94%	20	-3.04%	23	-0.49%	2.55%
	2ºHalf	17	-10.12%	13	5.31%	16	11.44%	6.13%
2018	1º Half	41	17.27%	33	16.68%	36	19.59%	2.91%
	2ºHalf	29	38.24%	11	15.37%	14	15.49%	0.12%
2019	1º Half	23	3.97%	17	7.98%	20	12.19%	4.21%
	2ºHalf	32	21.76%	18	26.56%	21	29.82%	3.26%
2020	1º Half	34	26.55%	25	12.24%	28	31.29%	19.05%
	2ºHalf	14	-328.22%	7	-181.75%	10	75.92%	5.83%
2021	1º Half	18	-9.88%	14	-9.3%	17	-1.29%	8.01%
	2ºHalf	18	-18.66%	10	0.59%	13	7.91%	7.32%
2022	1º Half	39	45.12%	28	41.96%	31	24.78%	-17.18%
	2ºHalf	24	10.94%	11	22.13%	14	25.65%	3.52%
2023	1ºHalf	36	34.8%	27	25.77%	30	33.95%	8.18%

Table 3 Sample Overview

Please note that in the table above, each transaction corresponds to the opening or closing of a position in a stock. In other words, if we open a long position in A and a short position in B, and then close the long position of A and the short position of B, we have 4 transactions. However, in the table above, one transaction corresponds to the opening and closing of a position, which, in the previous example, corresponds to only 2 transactions.

This table provides valuable insights into returns during periods of crisis and their subsequent behavior. It's evident that from 2012 until the end of 2019, returns showed greater consistency, reaching annual highs of 60% in our pairs trading strategy, making this strategy quite appealing.

Moreover, It’s worth noting that post-2010, the average return has increased compared to the total of our sample, which could be a good topic for future research. Additionally, it is worth highlighting that the average number of transactions remained unchanged.

4.3 Strategy Vs. Market

It’s important to verify the performance of our pairs trading strategy Vs. return of S&P500.

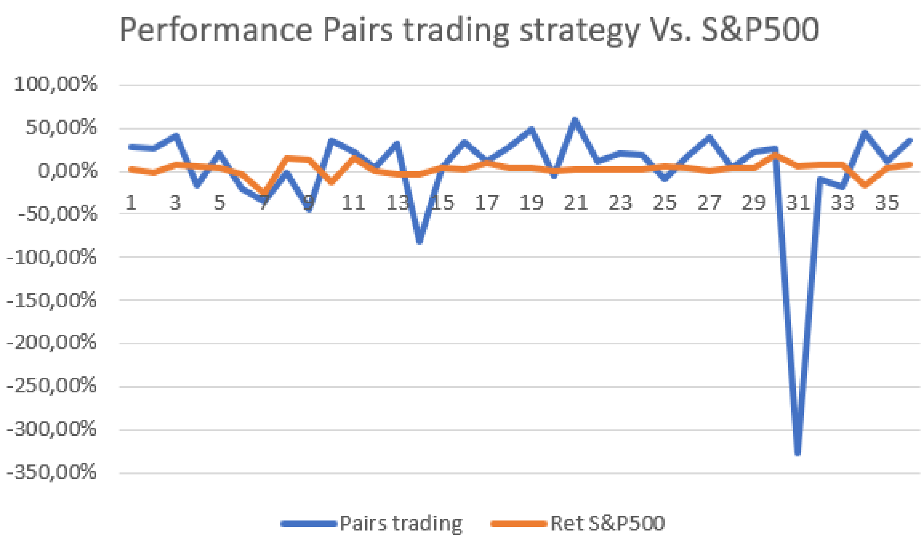

Figure 6 Performance pairs trading strategy Vs. S&P500

After analyzing the graph we can conclude, once again, that S&P500 is more likely than pairs trading strategy for those who are risk-averse.

Based on the data from the above table, we can see that our pairs trading strategy has both a higher and a lower return. Thus, its standard deviation is greater. However, when we consider the total return, we see that the S&P500 has a cumulative return of 91.43%, while our trading strategy has a 72.66% return. This conclusion leads us to assume that for this period, investing in the S&P500 would have been a more reliable choice. Excluding both crisis timeframes, we obtain a return of 455.44 % against 56.12 % with the same terms. Therefore, I can preliminarily conclude that diversification is the methodology worth investing in since the S&P500 increased in percentage during the crisis timeframe.

Contrary to what many authors argue, in general, the pairs trading strategy yields good returns during times of crisis, as we can see throughout our strategy, it was during these periods that we experienced the worst returns (COVID-19 and the subprime crisis of 2008), as they were the years with the most negative returns.

5. Conclusions and improvements for future research

During our project, we applied the Gatev et. Al. (2006) pairs trading strategy with some differences. First of all, we applied our strategy to the Index (S&P500), however we didn’t consider the liquidity volume because all the enterprises in the index have a huge volume of transactions, thus we can assume that when we try to close the position, we don’t have a problem about market operations (buy and sell).

Secondly, we studied the possibility of seasonality in financial markets, thus we will execute a pairs trading strategy based on three alternatives: 1) a simple pairs trading strategy, 2) a pairs trading strategy to verify the existence of seasonality, and 3) a pairs trading strategy to verify for seasonality and invest in periods when we didn't operate in the S&P 500 as an index.

We worked with the top 10 pairs of stocks for 37 simulation periods, each lasting 6 months.

After all simulation for the top 10 pairs of stocks, the average returns were 2.02% , -0.85% and 1.68% , the total return were 72.66% , -30.77% and 61%, the average return for each trade were 0.07% , -0.04% and 0.08%, and for our first strategy has the higher maximum and the lowest minimum as well as the higher standard deviation , almost two times than second strategy (with seasonality).

Regarding the work done by gatev et al. (2006), the returns presented by our strategy were lower, the higher motivation to this conclusion is because we didn’t have in consideration the sector for each pair, we only look at the minimum distance, this could be the reason why we have a huge different from gatev strategy.

About our third strategy, the strategy with higher return per trade (0.08%), and the strategy with lower risk, therefore the lowest highers and the lowest lowers.

However, it’s important mention that alongside our table 1, we can´t conclude if there is seasonality or not, even the standard deviation was higher, by consequently the fluctuation of the returns is higher.

Another important conclusion, is after excluded bad times just like covid-19 and 2008 crisis(table 2), we can concluded that for the measure, return per trade suffer a slight increase along our threes strategies , fixed the values on 0.58% , 0.5% and 0.52%.

On the other hand, we see that the S&P500 has a cumulative return of 91.43%, while our trading strategy has a 72.66% return, based on this we can get our major conclusion from this project, we can't fight with the market, and once again the best strategy for an investment take in consideration, time, money, power, the best choice is buy an index.

For future work, we suggest conducting a pair trading strategy and analyzing the returns after 2010 compared to returns before 2010 to understand the reasons behind any differences. Develop a strategy similar to the one proposed in this project but attempt to group companies by sector and market volume, as is customary in pairs trading strategies.

Another significant topic for future research is to implement a pairs trading strategy where we limit our losses. In our sample, we observe either very negative results or results close to zero. Therefore, it could be an excellent strategy to limit our losses to 15%. This way, we would know that our maximum losses would be 15%, or to simplify, we could suggest that if we experience losses for 7 consecutive days, we will close the position automatically.

Lastly, a significant topic for future research would be to investigate why we have a large number of open pairs that never revert to convergence.

This is of significant importance because those were the pairs that ruined part of our returns and, therefore, we were forced to close those pairs on the last day of the trading period. If, eventually, we can acquire more information about these pairs, perhaps that will help us to improve our returns.

It would also be interesting to test the strategy with intra-day prices, something that is even more realistic, given that nowadays most strategies are applied by computers that trade with high frequency and high precision.

6.Bibliografia

Alexander, C. (1999). *Optimal Hedging Using Cointegration* (Vol. 357, Issue 1758).

Alexander, C., Dimitriu, A., Author, C., Alexander Professor of Risk Management, C., Dimitriu Student, A., Alexander, C., & Dimitriu, A. (2002). *THE BUSINESS SCHOOL FOR FINANCIAL MARKETS*. www.ismacentre.rdg.ac.uk

Baker, M. & Wurgler, J. (2007). Investor Sentiment in the Stock Market. Journal of Economic Perspectives, 21, 129–51.

Blamont, J. E., Young, R. E., Seiff, A., Ragent, B., Sagdeev, R., Linkin, V. M., Kerzhanovich, V. V., Ingersoll, A. P., Crisp, D., Elson, L. S., Preston, R. A., Golitsyn, G. S., & Ivanov, V. N. (1986). Implications of the VEGA balloon results for Venus atmospheric dynamics. *Science*, *231*(4744), 1422–1425. https://doi.org/10.1126/science.231.4744.1422

Bodie, Z.; Kane, A. & Marcus, A. 2010. Investments. New York: McGraw-Hill/Irwin.

Engle, R. F. & Granger, C. W. J. 1987. Co-Integration and Error Correction: Representation, Estimation and Testing. Journal of the Econometric Society, 55:251-276.

Bouman, S., & Jacobsen, B. (2002). *The Halloween Indicator, "Sell in May and Go Away": Another Puzzle* (Vol. 92, Issue 5).

Brown, S. J., Goetzmann, W., Ibbotson, R. G., & Ross, S. A. (1992). Survivorship Bias in Performance Studies. In *Source: The Review of Financial Studies* (Vol. 5, Issue 4). https://about.jstor.org/terms

Caldeira, J. F., & Moura, G. V. (2013). *Selection of a Portfolio of Pairs Based on Cointegration: A Statistical Arbitrage Strategy*. http://ssrn.com/abstract=2196391https://ssrn.com/abstract=2196391Electroniccopyavailableat:http://ssrn.com/abstract=2196391

Deng, X., & Yuan, Y. (2021). A novel fuzzy dominant goal programming for portfolio selection with systematic risk and non-systematic risk. *Soft Computing*, *25*(23), 14809–14828. https://doi.org/10.1007/s00500-021-06226-x

Drogalas, G., A., Athianos, S., P., Bakas, G., & Elekidis, G. (2007). *Seasonalities in stock markets: the Day of the Week Effect*. http://ssrn.com/abstract=2515097Electroniccopyavailableat:https://ssrn.com/abstract=2515097Electroniccopyavailableat:https://ssrn.com/abstract=2515097Electroniccopyavailableat:http://ssrn.com/abstract=2515097

Fama, E. F. (1965). The Behavior of Stock-Market Prices. In *Source: The Journal of Business* (Vol. 38, Issue 1). https://www.jstor.org/stable/2350752

Fama, E. F., & French, K. R. (2004). *The Capital Asset Pricing Model: Theory and Evidence*.

Fountas, S. & Segredakis, K. (2002). Emerging stock markets return seasonalities: the January effect and the tax-loss selling hypothesis. Applied Financial Economics, 12, 291-9.

Gatev, E., William Goetzmann, B. N., Geert Rouwenhorst, K., Gatev Assistant Professor Boston College William Goetzmann Edwin J Beinecke Professor of Finance, E. N., Studies, M., Geert Rouwenhorst Professor, K., Bossaerts, P., Cooper, M., Ingersoll, J. & Jagannathan, R. (2006). *Value Arbitrage Rule Pairs Trading: Performance of a Relative Value Arbitrage Rule*. http://ssrn.com/abstract=141615Electroniccopyavailableat:https://ssrn.com/abstract=141615

Krauss, C. (2017). Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys*, *31*(2), 513–545. https://doi.org/10.1111/joes.12153

Lintner John. (1965). Lintner1965b. *Journal of Finance*.

Lobão, J., & Lobo, C. (2018). *Sazonalidade Mensal e o Efeito Passagem de Ano: Monthly Seasonality and the Turn-of-the-year Effect: New Evidence from the Euronext Lisbon*.

mossin jan. (1966). *Mossin1966*.

Malkiel, B. G. (1973). A random walk down wall street [by] burton G. Malkiel. Norton.

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, 6(3), 96-110.

Papadakis, G., & Wysocki, P. (2007). *Pairs Trading and Accounting Information*.

Schizas, P., Thomakos, D. D., & Wang, T. (2011). *Pairs Trading on International ETFs*. http://ssrn.com/abstract=1958546Tel.:+1

Sharpe, W. F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Sharpe, W. F. (1987). Integrated Asset Allocation. *Financial Analysts Journal*, *43*(5), 25–32. https://doi.org/10.2469/faj.v43.n5.25

Sharpe, W. F. (1991). The Arithmetic of Active Management. In *Source: Financial Analysts Journal* (Vol. 47, Issue 1).

Wachtel, S. B. (1942). *Certain Observations on Seasonal Movements in Stock Prices* (Vol. 15, Issue 2). https://www.jstor.org/stable/2350013

Zhiwu Chen, Peter J. Knez, Measurement of Market Integration and Arbitrage, *The Review of Financial Studies*, Volume 8, Issue 2, April 1995, Pages 287–325, https://doi.org/10.1093/rfs/8.2.287