NOVA IMS Information Management School

# MEGI

Master Degree Program in
**Statistics and Information Management**

# Pairs Trading using a Cointegration Approach

## An Empirical Investigation using US Stock Market data

Samuel Tomé Alexandre Lourenço

Master Thesis

presented as partial requirement for obtaining the Master Degree in Statistics and Information Management

**NOVA Information Management School**
**Instituto Superior de Estatística e Gestão de Informação**
Universidade Nova de Lisboa

**NOVA Information Management School**
**Instituto Superior de Estatística e Gestão de Informação**
Universidade Nova de Lisboa

# Pairs Trading using a Cointegration Approach

## An Empirical Investigation using US Stock Market data

by

Samuel Tomé Alexandre Lourenço

Master Thesis presented as partial requirement for obtaining the Master's degree in Statistics and Information Management, with a specialization in risk management.

**Supervised by**

Prof. Doutor Jorge Miguel Ventura Bravo, PhD

NOVA IMS & Université Paris-Dauphine PSL

June, 2024

# STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

[Lisboa, 2024]

# ACKNOWLEDGEMENTS

First of all, I want to thank all my family, who had supported me all along this journey and opened my eyes to reality, when it was necessary. Primarily I want to thank my brother, who has been a role model for me since I was born, supporting me in every step that I took. Then, a thank to my sister Rute, who is a remarkable student and a very kind and resilient person, made me a better person in that aspect too. Thank my sister Luísa who every day shows me what is the will to live, despite every adversity. Lastly, I thank my parents who gave me the best education they could and made me who I am today. Thank you to my father who is my idol, being the most hardworking person I know, and to my mother, the kindest person I know, who is always there for me.

Secondly, I want to thank all my friends and my girlfriend, who supported me in this thesis, giving me the strength to not give up, and comprehension in the moments that I could not be present in the last months.

Finally, thanks to my supervisor, Professor Doutor Jorge Bravo, who helped me along the way and was patient and flexible with me.

# ABSTRACT

This dissertation will study the return of a pairs trading strategy in the US stock market. Pairs trading is a market-neutral trading strategy that exploits the price movements between two historically correlated securities. The purpose of this thesis is to evaluate the performance of the cointegration approach in this market-neutral strategy and to assess if it has an edge over other strategies. This research employs a quantitative approach, using historical price data of 50 securities from various sectors over 8 years. The study applies a cointegration approach to identify pairs and execute trades, with the approach done with the Engle-Granger method. Performance metrics such as returns, Sharpe ratio, and maximum drawdown are analyzed. The findings in this study indicate that pairs trading can generate consistent returns with relatively low volatility in stable market conditions. However, the strategy's performance deteriorates during periods of high market volatility. This study observed that the pairs trading strategy had better returns with cointegrated pairs and contributed to the existing literature on the cointegration approach by offering a complete evaluation of pairs trading performance and returns with US stock market data.

# KEYWORDS

Pairs Trading ; Cointegration ; Least Squares ; Market-neutral strategy ; Statistical Arbitrage

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS AND ACRONYMS

**AR** – Autoregressive

**PP** - Phillips-Perron Test

**ADF** - Augmented Dickey-Fuller

**PGFF -** Pantula, Gonzales-Farias and Fuller Test

**ERSD -** Elliott, Rothenberg and Stock DF-GLS Test

**JOT** - Johansen's Trace Test

**SPR** - Schmidt and Phillips Rho Test

**OLS -** Ordinary least squares

**LS** – Least Squares

**EMH** - Efficient Market Hypothesis

# 1. INTRODUCTION

A pairs trade is a trading strategy that involves matching a long position with a short position in two stocks with a high correlation. Pairs Trading is an investment strategy directed at exploiting market inefficiencies that is known as a “market-neutral” strategy. The first empirical work on this strategy was made by Gatev et al. (1998) where it was documented this strategy and its impacts on the market efficiency theory. The work was later updated, in 2006, and became one of the most cited ones regarding pairs trading. This research focuses on exploring the efficacy of pairs trading in present-day markets, addressing its relevance in the current financial markets, marked by increased volatility and uncertainty.

This approach was first developed by Nunzio Tartaglia's quantitative team at Morgan Stanley during the decade of 1980. In 1987 this team reported a profit of over 50 million dollars, however, after two years they split up the group. Since then, it has continued to be a significant statistical arbitrage method that hedge funds use, as well as common traders (Vidyamurthy, 2004). Gatev et al. (1998), Vidyamurthy (2004), Whistler (2004), and Reverre (2001) provide historical context and analysis on pairs trading. Additionally, pairs trading is considered a form of Statistical Arbitrage, offering opportunities to exploit market anomalies, having as key concepts the statistical arbitrage as referred before, cointegration and mean reversion. This strategy can be applied to various equilibrium relationships in financial markets, including market-neutral portfolios with both long and short positions, as Nicholas (2000) discussed. Pairs trading gained popularity due to its potential for generating returns with reduced risk exposure. Understanding these initial assumptions and historical developments is important to comprehend the complexities of pairs trading and its application in today's markets.

Markets are not always in equilibrium, they tend to converge to rational levels over time, providing opportunities for traders to capitalize on deviations. This strategy involves taking a long position in one security that is undervalued while simultaneously taking a short position in another overvalued security. If the two securities belong to the same sector, a ratio of unity may be applied. This ratio is chosen to create a market-neutral portfolio, one with zero correlation to overall market movements, often referred to as a spread. The spread is modeled as a mean-reverting process using market data, enabling to forecasting of its behavior. When the value reaches a Z-score of 1 or -1 the trade is initiated, aiming to profit when the spread returns to its expected level. Nevertheless, we don´t know if this strategy is effective or not in current markets and in the US stock market. Therefore, how effective is pairs trading as a market-neutral strategy in the current market environment, particularly in the US stock market and in the context of increased market volatility? The main goal of this thesis is to evaluate the performance of pairs trading strategies in recent financial markets with the cointegration approach and to analyze the impact of market volatility on the profitability of pairs trading. This study aims to contribute to the existing empirical research by addressing gaps in the literature regarding the current applicability of pairs trading with the cointegration approach with US stock market data, providing present-day insights into its performance among new market dynamics. The scope of this research involves the analysis of pairs trading strategies with cointegration approach from 2016 to 2024, focusing on major US stock market data. The study delimits itself to equity pairs, excluding other asset classes such as commodities and currencies.

This research employs quantitative methodology, utilizing historical price data to identify and analyze pairs trading opportunities. The research end includes statistical tests for cointegration, with this thesis being divided into several parts. Firstly, it will review the pairs trading theory, such as the different strategies and approaches, the efficient market hypothesis, and arbitrage concepts. Then I will describe the data source used, and the pairs selection, with a detailed description of the pairs correlation analysis, cointegration, and other criteria. Then the trading strategy will be defined, such as trading rules, entry and exit, and risk management. In the next part of the thesis, it will be discussed the empirical results, such as the profitability of the strategy and backtesting results.

# 2. LITERATURE REVIEW

Compared with other topics, Pairs Trading is a very recent subject in literature. The first work on this theme was done by (Gatev, et al 1998) who researched an analysis of this strategy and the impacts on the efficient market hypothesis. In this approach, he used the distance approach to find tradeable pairs. The strategy was based on trading signals that were generated when the prices diverged more than two standard deviations from the mean observed in the formation period.

The Efficient Market Hypothesis (EMH) considers that stock prices reflect all available information, making it impossible to consistently achieve higher returns than average market returns on a risk-adjusted basis (Fama, 1970). Pair trading, as a form of statistical arbitrage, confronts the Efficient Market Hypothesis by exploiting temporary inefficiencies in stock price movements between two stocks (Krauss, 2017).

Various studies have examined the profitability of pair trading in different markets, often with mixed results. For instance, (Engelberg, Gao, and Jagannathan, 2009) found that pair trading leads to significant abnormal returns in the U.S. stock market, while (Bowen, Hutchinson, and O'Sullivan, 2016) demonstrated the importance of market environments in the performance of pair trading strategies in European equities.

Statistical methods such as correlation and cointegration are vital in identifying and evaluating pair trading opportunities. The application of these methods helps in understanding the long-term equilibrium relationships and short-term dynamics between paired stocks (Vidyamurthy, 2004). Techniques such as the Johansen cointegration test and the Engle-Granger two-step method are particularly prominent in academic and practical applications (Alexander, 2009).

Regarding the different approaches for pairs trading, there is no consensus on the most reliable approaches.

Cointegration was first intended in an article, by two econometricians, (Engle and Granger, 1987). The cointegration approach in pairs trading is an advanced statistical strategy used to identify pairs of stocks that exhibit a stable, long-term relationship despite short-term fluctuations. Cointegration, unlike simple correlation, indicates that the price series of two assets move together in such a way that their spread is stationary, meaning it has a constant mean and variance over time. (Granger, Clive W. J. 2004). Gatev et al. (2006) demonstrated that pairs formed using the cointegration method exhibited consistent profitability over different market conditions, highlighting the method's robustness.

As I mentioned earlier, pairs trading gained significant academic attention starting with the publication by Gatev et al. (1998). This method employs the distance approach to identify tradeable pairs of stocks. To find the minimum distance between stocks, a cumulative return index is constructed with dividends reinvested. The stock prices are then normalized, and different indices are tested against each other to find pairs with the minimum sum of squared distances. This process uses daily data over a 12-month formation period followed by a 6-month trading period. Trading signals are generated when prices diverge more than two standard deviations from the mean observed in the formation period. Gatev et al (1998). based their research on daily data from 1962 to 1997 for all liquid US stocks, demonstrating annualized excess returns of up to 12% and strength to transaction costs, with no correlation to S&P 500 returns.

Despite its success, the profitability of pairs trading has declined over time. Do and Faff (2012) extended the observation period to 2009 and showed decreasing returns, particularly since 2002. This decline is attributed to increased hedge fund trading activity exploiting statistical arbitrage opportunities and a higher number of pairs deviating from equilibrium and not returning.

Other approaches have also been explored. Vidyamurthy (2004) introduced a cointegration-based method, finding it more effective than correlation for identifying pricing anomalies influenced by shared underlying factors. Rad, Low, and Faff (2016) expanded on this with a copulas approach, though it was less effective in terms of return.

Elliott et al. (2005) introduced a stochastic approach using an Ornstein-Uhlenbeck process to model the spread, which reverts to the mean. This method estimates the time for the spread to return to the historical mean and uses a Kalman filter to estimate the true state of the spread, providing a continuous time model suitable for forecasting.

As stated before, pairs trading has been a topic of interest in literature since the mid-1980s, gaining practical attention at Morgan Stanley. However, the methodologies remained proprietary until the 2000s. Gatev, et al. (1998) empirical work documented pairs trading strategies and their implications for market efficiency. Their work, updated in 2006, became highly cited and provided the basis for further studies. The strategy involves a formation period of 12 months to identify equity pairs using the distance method, followed by a trading period where positions are opened when prices diverge by more than two standard deviations and closed when they converge.

Studies like Perlin (2007) and Papadakis and Wysicki (2008) have further explored pairs trading across different markets and conditions, showing varied profitability and factors influencing returns. Engelberg, Gao, and Jagannathan (2009) added informational events to the strategy, finding that differences in how quickly new information is incorporated can affect profitability.

Recent research, such as Franco (2014) and Ribeiro (2015), confirms a significant decrease in pairs trading performance in the last decades, particularly during high transaction cost periods. However, they also suggest that the strategy remains profitable under certain conditions, such as during market crises or when limited to same-industry pairs.

Overall, pairs trading remains a robust and intriguing strategy, though its profitability has been challenged by market changes and increased competition from hedge funds. Literature continues to evolve with new methodologies and innovations, keeping the academic and trading communities engaged.

# 3. METHODOLOGY

The methodology is based on cointegration approach and least squares regression. The time horizon is 8 years, from 2016 to 2024. There will be analyzed 25 pairs, so 50 stocks from the US stock market. There are some pairs where the time horizon is not the above since some companies made their IPO after that date. The objective here is to demonstrate that the strategy in pairs that are cointegrated has a better performance than the ones that are not.

## 3.1. COINTEGRATION APPROACH

Forecasting stock market prices and returns is particularly difficult, given the nonlinearity, volatility, and complexity of the time series and the generally accepted, semi-strong form of market efficiency (Fama, 1970; Bravo, 2024). Previous research identifies several univariate or multivariate techniques considering the lagged values of the time series, and fundamental and technical analysis factors, considering both statistical learning and artificial intelligence methods, including model combination approaches (Bravo et al. 2020, 2021, 2023; Ashofteh et al. 2021, 2022). Cointegration is an econometric model which replies to the long-run equilibrium between economic time series. If we have two or more time series that are nonstationary, but a linear combination of them is stationary, then they are probably cointegrated (Wei, 2006).

The concept of stationarity is correlated to the properties of these stochastic processes. If the data are assumed to be stationary if the means, variances, and covariances of the series are independent of time, rather than the entire distribution we assume weak stationarity. Nonstationarity in a time series occurs when there is no constant mean or no constant variance, however, the most important one is the unit root.

Regarding the unit root, a sequence that has one or more specific roots that are equal to one is named a unit root process. The simple model that should have a unit root is the AR(1) model.
In the empirical results, for the programming phase in Rstudio, I´ve used the package, egcm, that computes the Engle-Granger cointegration test, and many other unit-root tests, like Augmented Dickey-Fuller (ADF) Test, Phillips-Perron (PP) Test, Johansen's Trace Test (JOT) and others.

Regarding the Engle-Granger method (Engle-Granger, 1987), it is based on the fact that regressing non-stationary series on other series, can lead to misleading results. However, if all variables are found to have a unit root (i.e., they are integrated of order one, I(1), the regression can still be significant if the variables are cointegrated. Then to check for cointegration, we estimate the least squares regression equation (that I will talk about later) and analyze the residuals for a unit root. If the residuals are stationary, I(0) indicates that the variables are cointegrated and share a long-term equilibrium relationship. The Engle-Granger method uses this approach, so it tests for no cointegration by examining the stationarity of the regression residuals. This method uses the ordinary least squares (OLS) to estimate the relationship between the variables and apply unit root tests to the residuals to check for stationarity. If the null hypothesis of a unit root is rejected, it supports cointegration.

Regarding the unit root tests, the Augmented Dickey-Fuller (ADF), one of the most popular ones, developed by (Dickey and Fuller, 1979), tests for the existence of a unit root in a time series sample, which helps determine if the series is non-stationary. The methodology relies on the fact that the ADF test augments the basic Dickey-Fuller test by including lagged differences of the dependent variable to account for higher-order correlation and, secondly, the null hypothesis. Due to the invalidity of the DF statistic in the presence of serial correlation, the test is typically applied in its augmented form. ($H_0$) is that the series has a unit root (i.e., it is non-stationary) and lastly alternative hypothesis ($H_1$) is that the series is stationary.

Interpretation of this test depends on whether the p-value is less than the significance level (e.g., 0.05), if it´s less than the significance level, we reject $H_0$ and conclude that the series is stationary. If the p-value exceeds the significance level, we fail to reject $H_0$ and conclude that the series is non-stationary.

Another test is the Phillips-Perron (PP) Test. It is similar to the ADF test, the PP test checks for a unit root in a time series sample. Regarding the methodology, the PP test adjusts for serial correlation and heteroskedasticity in the errors by modifying the test statistics. It uses non-parametric statistical methods to account for the autocorrelation and the null hypothesis ($H_0$) is that the series has a unit root (i.e., it is non-stationary). Interpretation is the same. Cheung, Y. W., & Lai, K. S. (1997).

Concerning the Pantula, Gonzales-Farias and Fuller (PGFF) Test, is another approach to testing for unit roots and determining stationarity in a time series. The PGFF test involves testing the presence of a unit root by considering both level and trend stationary alternatives. This test can account for structural breaks in the time series data and the null hypothesis ($H_0$) is that the series has a unit root.

The Elliott, Rothenberg, and Stock DF-GLS (ERSD) Test is a more efficient version of the ADF test for testing for a unit root in a time series. It applies a Generalized Least Squares (GLS) detrending procedure to the data before performing the Dickey-Fuller test. The null hypothesis ($H_0$) is that the series has a unit root.

Johansen's Trace Test (JOT) is used to determine the number of cointegration vectors in a multivariate time series. The test is based on a Vector Autoregressive model and assesses the rank of the cointegration matrix. It tests the null hypothesis that the number of cointegration vectors is less than or equal to a given number against the alternative hypothesis of more cointegration vectors. In this test, if the test statistic is greater than the critical value, we reject $H_0$ and conclude that there are more cointegration vectors. If the p-value is less than the significance level, it indicates the presence of cointegration.

Lastly, the Schmidt and Phillips Rho (SPR) Test is another method to test for unit roots in a time series. It is an extension of the Phillips-Perron test and adjusts for potential serial correlation in the error terms and the null hypothesis ($H_0$) is that the series has a unit root.

These tests are designed to check for unit roots (stationarity vs. non-stationarity) and cointegration in time series data. They are essential tools in econometrics for understanding the properties of time series and for building reliable models. Each test has its strengths and weaknesses, and often multiple tests are used together to make a more robust conclusion about the data.

## 3.2. Distance Approach

The distance approach is a popular and relatively straightforward method used in pairs trading popularized in the research of Gatev et al. (2006). It focuses on the price difference or spread between two correlated assets, usually stocks, to identify trading opportunities based on the assumption that the spread will revert to a historical mean. This method leverages the concept of mean reversion, which is central to many pairs trading strategies. In the distance approach, traders identify pairs of stocks that have shown a stable long-term relationship. The basic premise is that while the prices of the two stocks might diverge temporarily due to various market factors, they will eventually converge again because of their underlying correlation.

The approach is split into two phases. There is a formation period in which likewise moving pairs are selected and then a trading period in which trading signals are created and used to take a position in the stocks and consequently to close them.

(Gatev et al., 2006) use daily data from 1962 to 2002 on all liquid US stocks. For every stock, a cumulative total return index is formed and normalized for the 12 months formation period in the formation period

In that research, it is stated that the Top 20 pairs that have the smallest historic distance for the 12-month formation period are now selected for the following 6-month trading period. In addition to that, they also applied sector filter criteria to filter out pairs that belong to the same industry, criteria that I will apply in my selection of pairs. During the trading period, the price series of the securities are normalized again on the first day. Positions are opened when the historic spread of the pairs diverges by more than two standard deviations from their mean and are closed once they revert to the historic mean. This method adds a qualitative element to the quantitative approach.

## 3.3. OTHER APPROACHES

Like the cointegration and distance approaches, the time series approach in pairs trading relies on statistical methods to identify and exploit relationships between two securities. In time series analysis, traders typically use statistical models to forecast future price movements based on historical data. The time series approach can be integrated into pairs trading by focusing on modeling and forecasting the spread between two correlated assets, Patterson, K. D. (2000).

The stochastic control approach refers to a methodology used in financial modeling and decision-making under uncertainty, Mudchanatongsuk, et al (2008). It involves utilizing mathematical techniques from stochastic analysis and control theory to optimize decision-making processes in situations where outcomes are influenced by random variables. This approach is particularly relevant in finance, where market dynamics are often unpredictable and subject to stochastic processes. By applying stochastic control techniques, practitioners aim to devise strategies that maximize desired objectives while accounting for the inherent randomness and uncertainty present in financial markets.

Lastly, the copula approach is a statistical method used to model the dependence structure between multiple variables, particularly in finance and risk management. The objective of the copula approach is to apply the optimal copula between two stock returns and detect relative positions between pairs, Liew and Wu (2013). It involves separating the marginal distributions of individual variables from their joint distribution, allowing for a flexible and comprehensive analysis of their interrelationships. Copulas are functions that link the marginal distributions to the joint distribution, capturing the dependence patterns regardless of the specific distributions of the individual variables. This approach is valuable for assessing and managing various types of risk, such as market risk, credit risk, and operational risk, by accurately modeling the dependency between different factors and assets. Additionally, copulas are used in portfolio optimization, pricing of complex financial products, and measuring systemic risk in financial systems.

## 3.4. SELECTION OF APPROACH

Krauss (2017) conducted a study that compared the performance of various pairs trading strategies. For the distance approach, the annualized return for equities ranges between 7% and 11%. Although Vidyamurthy's (2004) paper, which is the most frequently cited for the cointegration approach, does not provide empirical results, Caldeira and Moura (2013) report an annualized return of 16,38% for the cointegration method in equities despite the research only examine the period from 2005 to 2010 for Brazilian stocks.

Despite this, several indicators suggest that the cointegration approach may be superior to the distance approach. The distance approach involves a simpler methodology, merely measuring the distance between price indices. In contrast, the cointegration approach, which involves regression analysis and stationarity tests to select tradeable pairs, is considered more robust from an econometric perspective, consequently, I will use in this study the cointegration approach.

## 3.5. LEAST SQUARES REGRESSION

As stated before, pairs trading is a representative market-neutral trading strategy that simultaneously, longs an undervalued stock and shorts an overvalued stock. This strategy is a form of statistical arbitrage trading that assumes the movements of the prices of the two assets will be like previous trends. It follows the hypothesis that prices will return to the long-term equilibrium. This strategy started from the idea that arbitrage opportunities exist when the price gap between two assets expands to or past a certain level. It is also based on the belief that historical price movements will not change significantly in the future. Gatev, et al. (2006).

Least Squares Regression is a method used to estimate the relationship between two variables by fitting a line that minimizes the sum of the squared differences between the observed values and the values predicted by the line. Björck, Å. (1990). In the context of pairs trading, it is used to model the linear relationship between the prices of two assets. The basic form of the regression model is:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad (1)$$

Equation (1) defines the simple linear regression model. It is also called the two-variable linear regression model or bivariate linear regression model because it relates the explanatory ($x$) and explained ($y$) variables. The variable $\epsilon$, called the residual term or disturbance in the relationship, represents factors other than $x$ that affect $y$ (Wooldridge, 1996).

The Engle-Granger two-step method directly uses least squares regression as it regresses $y_t$ on $x_t$ using ordinary least squares (OLS) to obtain the residuals $\epsilon_t$. Then it tests the residuals $\epsilon_t$ for stationarity using a unit root test, like the Augmented Dickey-Fuller test. If the residuals are found to be stationary, the series $x_t$ and $y_t$ are considered cointegrated.

Regarding its direct implication in the pairs trading strategy, it´s very important in the step of calculating the spread. So, the mechanism is the following, first, a pair of stocks with similar trends is identified. Second, regression analysis such as ordinary least squares (OLS), total least squares (TLS), and error correction models (ECM) is used to calculate the spread of these stocks. Finally, if the spread hits preset boundaries, investors will open a portfolio that takes a long position on the undervalued stock and shorts the overvalued stock. Subsequently, if the spread reverses to the mean, investors will close the portfolios that are opposite position to the open portfolio. In this case, the investor obtains an arbitrage profit by executing this strategy. However, there is a risk when the spread does not reverse to the mean. In such a situation, investors are at high risk because they cannot close the portfolio.

## 3.6. RISKS

Pairs trading can be relatively low risk if pairs are well-selected and the relationships truly are mean reverting, so that´s because my thesis is based on the cointegration approach. It can also be neutral to overall market movements, focusing only on the relationship between the pair.

By setting a stop-loss boundary, investors can hedge the risk. Many researchers have applied various statistical methods to improve the efficiency and performance of pairs trading. They focused on using the spread as a trading signal. The challenges that I will have here involve model risk (incorrect model or parameters), higher drawdowns that may lead to extended non-realized losses, and the breakdown of historical relationships due to structural market changes, if a black swan event occurs in one of them, like fraud or some similar event like the Enron case that leads to bankruptcy. This risk is why the returns on individual stocks cannot be taken as stationary. However, as we are short in the overvalued stock, this risk is diminished.

## 3.7. DATA

The data consists of fifty stocks from the S&P500, forming twenty-five pairs that will be analyzed. The data is from 01-01-2020 to 30-04-2024 and it was obtained through Yahoo Finance. It was adjusted for splits and other corporate actions. The stocks that I chose were based on the criteria that they must be from the same sector to facilitate cointegration. There are fifty stocks but that shouldn't be a problem as Alexander et al. (2002) prove, efficient long-short hedge strategies can be accomplished with relatively few stocks, therefore this is not a problem. Furthermore, hedge funds usually only use pairs that belong to the same activity sector, and I will do that too. I have chosen market data from the US stock market because it is the stock market with higher liquidity, more data, and more reliable nowadays.

# 4. EMPIRICAL STUDY

## 4.1. Research Questions

In this thesis on pairs trading using the cointegration approach, it is essential to formulate research questions that guide the investigation into the methodology, effectiveness, and application of this trading strategy. Therefore, to start the empirical study, this paper will focus mainly on the following research questions:

How does the pairs trading strategy with the cointegration approach perform, using US stock market data, in the current financial markets environment?

What is the risk-adjusted returns, Sharpe ratios, and other performance metrics for pairs trading using cointegration?

Who provides a better return, cointegrated or non-cointegrated pairs?

## 4.2. Assumptions

First of all, the assumptions made in this work were made to make this a better research and to best serve this work.

The transaction costs were assumed at 0, due to the difficulty of quantifying them. Then, the beta was assumed to be 0, as this is a market-neutral strategy, or a statistical arbitrage, in my point of view, the beta at 0 makes sense due to the absence of risk. With the beta at 0, the CAPM will be the value of the risk-free rate, which I assumed to be the 10-year treasury bond on 30 April 2024, which stands at 4,69%.

Regarding the capital invested, the returns are shown with only 1 unit, so I will not assume in this work that money was invested, I will work with units and percentages.

## 4.3. Cointegration tests

As stated before, there is a very convenient package, egcm, that computes the Engle-Granger cointegration test and many other unit-root tests. In particular, given two series $x(t)$ and $y(t)$, it searches for parameters $\alpha$, $\beta$, and $\rho$ such that

$$y_t = \alpha + \beta x_t + r_t \quad (2)$$

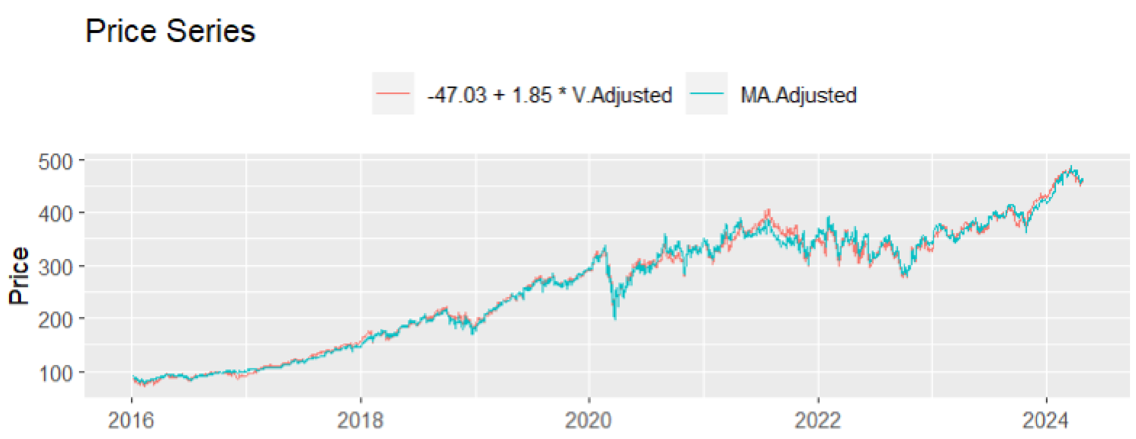
$$r_t = \rho r_{t-1} + \epsilon_t \quad (3)$$

where $r_t$ is the regression residual and $\epsilon_t$ the innovation. If |ρ|<1, then $x_t$ and $y_t$ are cointegrated (i.e., $r_t$ doesn't contain a unit root). Of course, the difficulty is in assessing exactly how much smaller than 1 the value of $|\rho|$ has to be. This is the task of the unit-root test. The most common example of a unit-root test is the Augmented Dickey-Fuller (ADF) test as I stated before in the methodology, which considers a null hypothesis that a unit root is present and an alternative hypothesis that the series is stationary (so a small ~~p~~$p$-value means an indication of strong stationarity). Complementary tests to the Augmented Dickey-Fuller (ADF) approach include for instance Kwiatkowski et al. (1992).

Another interesting quantity to measure mean-reversion is the half-life, defined as the time it takes for the spread to mean-revert half of its distance after having diverged from the mean of the spread.

The egcm library in R is designed for the estimation and testing of cointegrating relationships between pairs of financial time series. egcm stands for "Engle-Granger Cointegration Models," as stated before, which refers to the statistical methodology used to identify cointegrated pairs. This package simplifies the process of finding and testing cointegrated pairs, which is particularly useful in pairs trading strategies.

Let´s check if Visa and Mastercard are cointegrated. In Figure 1 below, the price series shows a regression model of Visa into Mastercard because the stock price of Mastercard is higher. Residual series is the differences in prices, so the spread and the innovations are the changes in residuals.



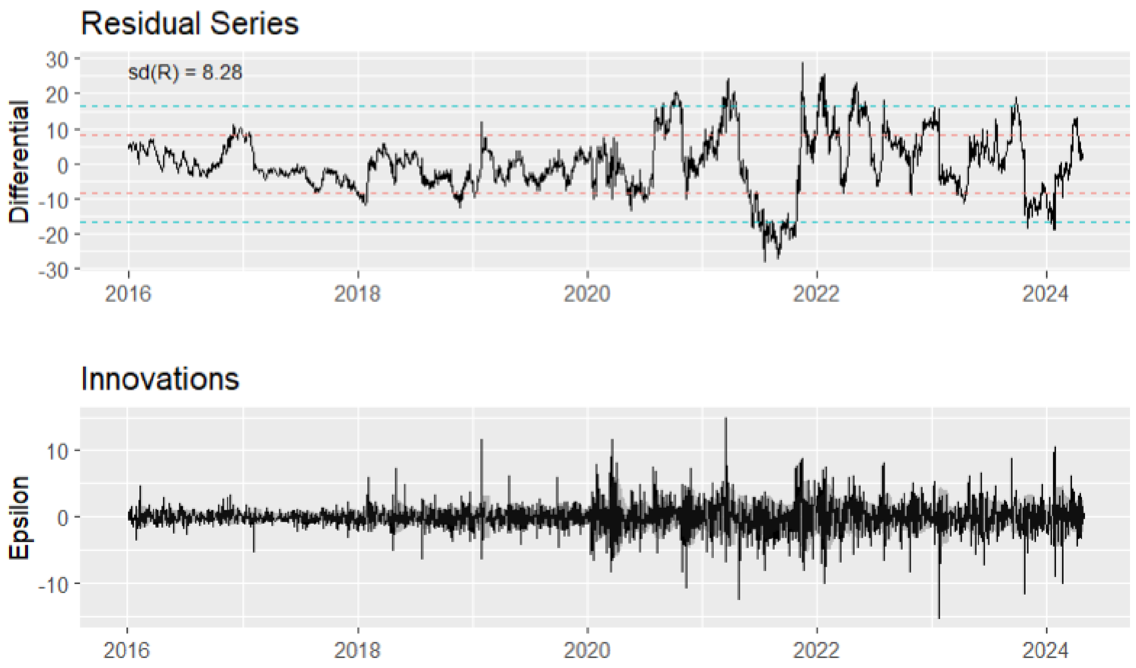



Figure 1: Cointegration Results

In the price series, we can see the historical price movements of Visa and Mastercard from 2016 to 2024. The red line represents the fitted relationship between the two assets using a linear regression model:

$$\hat{y}_i = -47.03 + 1.85 \times V$$

where -47.03 is the intercept $\beta_0$, 1.85 is the slope $\beta_1$ and $V$ is the independent variable.

The blue-green line represents the adjusted price of the second asset, labeled as "MA Adjusted. So, we can see that these two stocks are cointegrated. If they were not, we would see a red text stating "not cointegrated".

The residual series shows the residuals (or differences) between the observed prices and the fitted prices from the regression model. The black line represents the residuals over time. The red dashed lines indicate the standard deviation bands, specifically $\pm 1$ standard deviation (sd) from the mean. The blue dashed lines indicate $\pm 2$ standard deviations from the mean. A note on the graph indicates the residuals' standard deviation (sd) is 8.28. The residuals fluctuate around zero, with occasional significant deviations. Periods, where the residuals cross the standard deviation bands, indicate potential trading signals in a pairs trading strategy, and we will see that in trading the spread.

The innovations illustrate the changes in these residuals, helping to understand the dynamics and volatility of the deviations over time.

## 4.4. LEAST SQUARES REGRESSION FOR PAIRS TRADING

As stated before, in the context of pairs trading, least squares regression is used to model the linear relationship between the prices of two assets. The basic form of the simple linear regression model is especified in equation (1). To implement pairs trading using least squares regression, typically it is used a linear model to determine the hedge ratio, which is the coefficient that minimizes the sum of squared deviations between the prices of the two securities.

Firstly, we must identify the pairs to trade. Then, it is necessary to formulate the spread, as I will mention upfront. After that, the least squares regression will be used to estimate γ , the hedge ratio. The regression model is :

$$y_{1t} = \mu + \gamma y_{2t} + \epsilon_t \tag{4}$$

Where $y_{1t}$ is the observed value of the first variable at time $t$, or the price of stock at time $t$;. $y_{2t}$ is for the second variable and γ is the coefficient that scales $y_{2t}$, in relation to $y_{1t}$, subsequently, the hedge ratio. $\epsilon_t$ is the residual term.

The next step is to calculate the coefficients. To find the optimal values of $\mu$ and γ, we minimize the sum of squared residuals as follows:

$$\min_{\mu,\gamma} \sum_{t=1}^{T} \big(y_{1t} - (\mu + \gamma y_{2t})\big)^2 \tag{5}$$

The next steps are the computation of the spread and then generate the trading signals, as we will see in the next steps of this research.

In Figure 2 below, there are the prices of Visa and Mastercard deployed. We can see the dates and the ticker of both stocks. Then the data is downloaded from Yahoo Finance. The data is checked for missing values, and they are interpolated. Then we have some visual exploration of prices which provides an initial outlook on the content of the data.

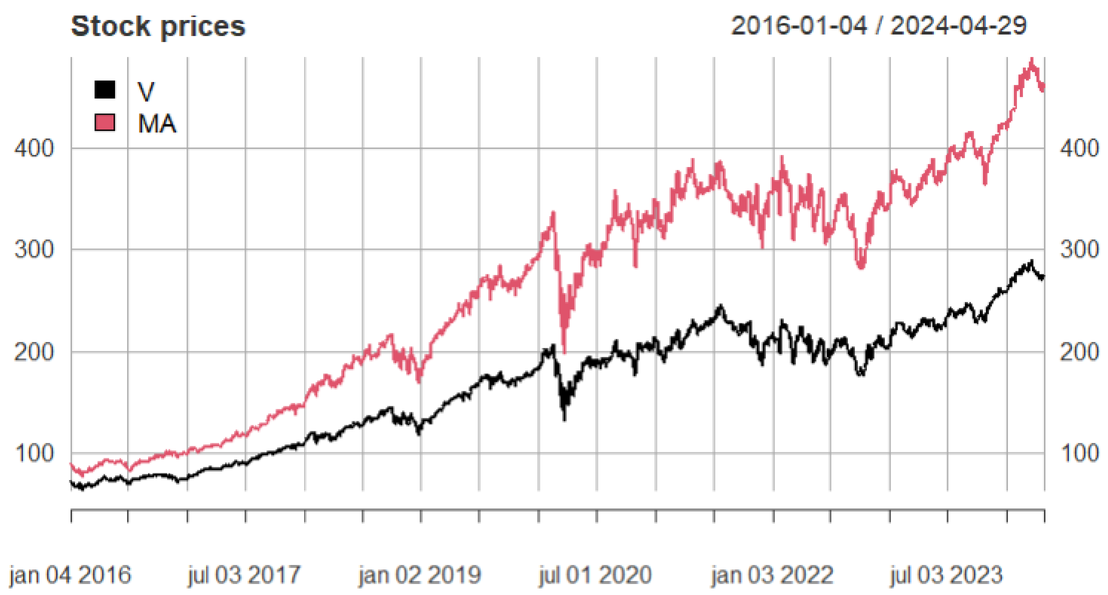



Figure 2: Stock Prices

The prices in this way don´t cross each other, so I must solve this. As we can see in Figure 15 in appendix, I used a least squares regression to estimate $\gamma$ and $\mu$ from a training period and then used it in the test period to test the strategy in training. The log prices of the two securities will be stored in "y1" and "y2". I didn´t use a rolling window cross-validation scheme. Instead, a fixed training and test set approach was used. The training period is fixed as the first 70% of the data, and the testing period is the remaining 30%. This is a straightforward split of the dataset and does not involve moving windows. The approach doesn´t employ a rolling window cross-validation scheme, which would involve repeatedly moving a fixed-length training and testing window across the entire dataset. In the code I have used a simple alternative procedure to split the data into training and testing sets, which is effective, however, does not provide the same level of robustness as a rolling window cross-validation scheme.

I will plot the log-prices of one stock and its regression based on the other and the pairs will now have the same "price" in this scale as we can see in Figure 3.

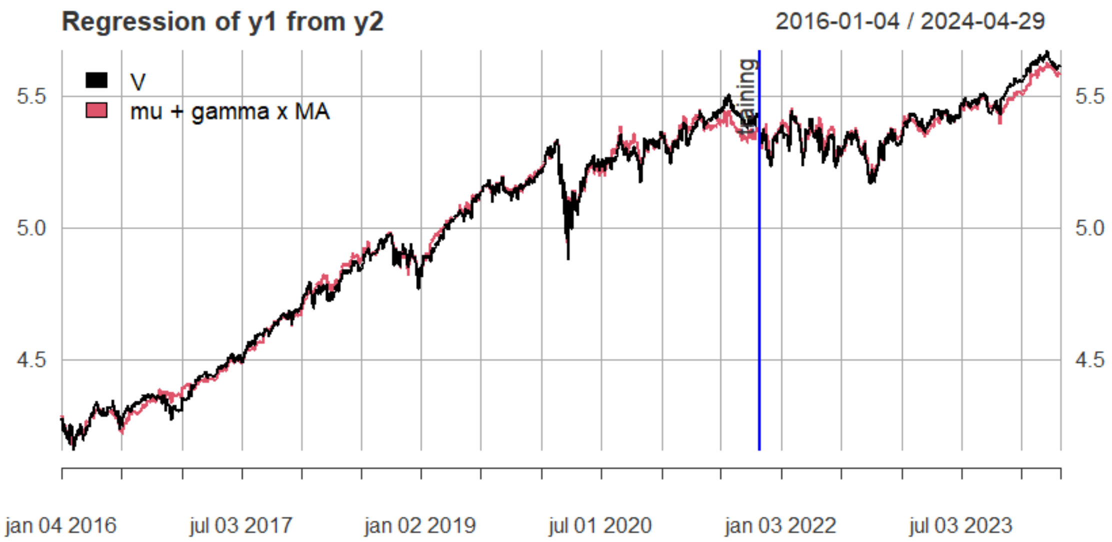



Figure 3: Least Squares Regression

## 4.5. Spread and Z-score

Now, we can plot the spread to observe its mean-reversion property. The spread refers to the difference in price or value between two correlated financial instruments. It is a metric used to identify trading opportunities and manage trades within the strategy. The spread is calculated based on the relationship between the prices of the two assets, which can be modeled using various approaches, such as least squares regression or cointegration.

The price spread between close substitute assets should have a long-term stable equilibrium over time. Hendry & Juselius (2001), use this idea to show that short-term deviations from these equivalent pricing conditions can create short-lived arbitrage opportunities depending on the duration of price deviation. When a deviation in the spread of the long-term equilibrium price relationship is identified to be substantially greater than the slipage1 due to the bid-ask spread, a position is opened simultaneously, buying the relatively undervalued stock and selling the relatively overvalued stock. The position is closed when the prices return to the spread level of long-term equilibrium. The net profit of the operation is the sum of the profits from the long and short positions, calculated as the difference between the open prices and closed prices (ignoring transaction costs). The formula of spread is established as:

$$z_t = y_{1t} - \gamma y_{2t} \tag{6}$$

Where $z_t$ is the spread at time *t*, $y1_t$ is the observed value of the first variable at time *t*. $y2_t$ is for the second variable and γ is the coefficient that scales $y2_t$, in relation to $y1_t$.

Indeed, the spread seems to be mean reverting. However, one can observe that the mean reversion is stronger during the training period than during the test period. This is a bit dangerous, however, the training period is much longer. One thing is for the spread to be mean-reverting during some period and a very different thing is whether the mean-reversion is persistent enough over subsequent periods. If it's not persistent enough, then it is very dangerous to trade the spread.

### 4.5.1. TRADING THE SPREAD

The spread formed as we can see in equation 6, is effectively equivalent to using the portfolio $w = [1 - \gamma]$. In order to make fair comparisons, we will fix the leverage of the portfolio so that $\| w \| 1 = 1$.
The normalized portfolio is weighed using the variable "gamma". The normalized portfolio and spread between both securities are computed. We can see that in Figure 4.

Let's form a more realistic spread based on the normalized portfolio, which can be seen in Figure 4 below.

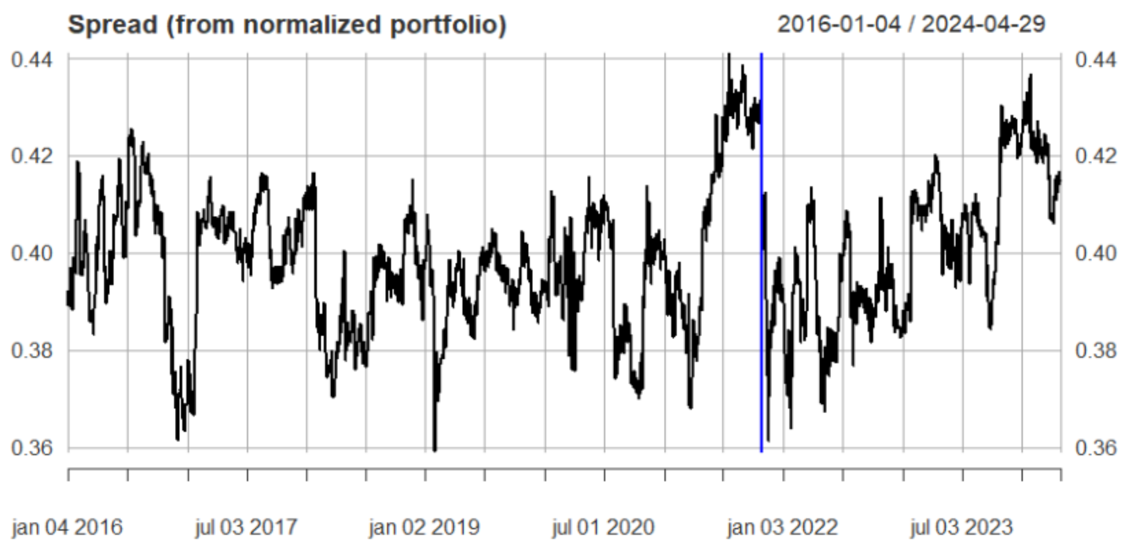



Figure 4: Spread from normalized portfolio

After the input of spread, we need something to create trading signals. The Z-score is a normalized version of the spread, which is useful to generate these trading signals. The z-score helps to standardize the spread between the two instruments, making it easier to identify when the spread significantly deviates from its historical mean. In particular, it is defined as:

$$\text{Zscore} = \frac{zt - E(zt)}{Std(zt)} \tag{7}$$

In Figure 5 below we can see the Z-score plotted into spread. There is a clear indication of when to enter a long or short position based on the spread´s deviation from the mean. I will first compute the Z-score and then define its thresholds, so the “threshold_short” is set to 0.7, meaning a short position is taken when the Z-score exceeds 0.7, and in the “threshold_long” a long position is taken when the Z-score drops below -0.7.

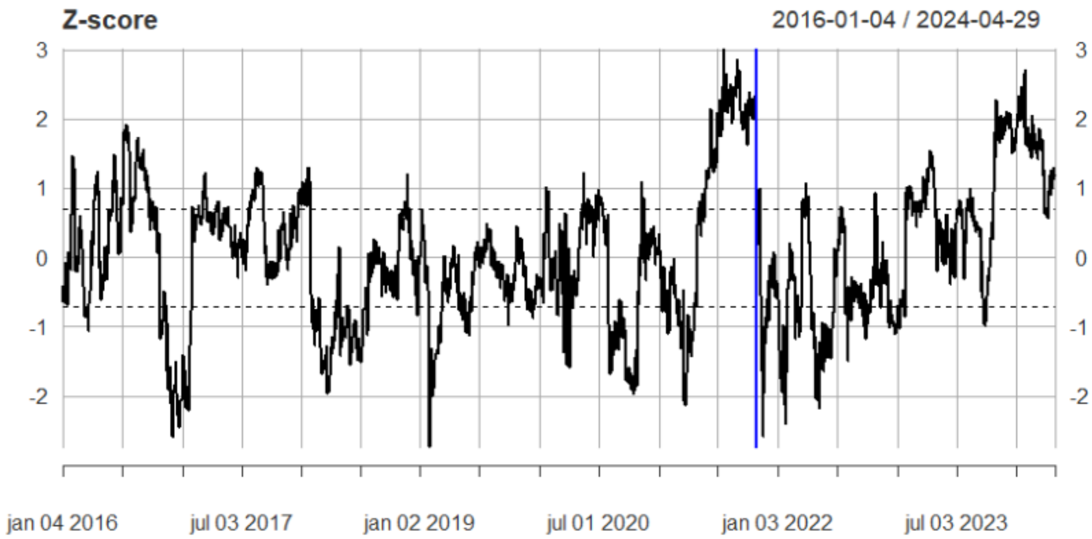


Figure 5: Z-score

It is ready to generate the trading signal. In Figure 6, when Z-score reaches the value of -1, the standard deviation of -1, the undervalued stock will be bought, and the overvalued stock will be sold and vice versa. When it reaches the std of 0, the mean, it closes the position. Regarding the code on initial position, signal [1] < - 0 initializes the first signal to 0 (no position). If the first Z-score is less than or equal to the long threshold, a long position (1) is taken. If the first Z-score is greater than or equal to the short threshold, a short position (-1) is taken.

In the loop part, it goes over each Z-score value to update the signal based on the previous signal and current Z-score. If no position is held (signal[t-1] = 0), the signal is updated based on the current Z-score relative to the thresholds. If a long position is held (signal[t-1] = 1), the position is closed if the Z-score is non-negative, and if a short position is held (signal[t-1] = -1), the position is closed if the Z-score is non-positive.

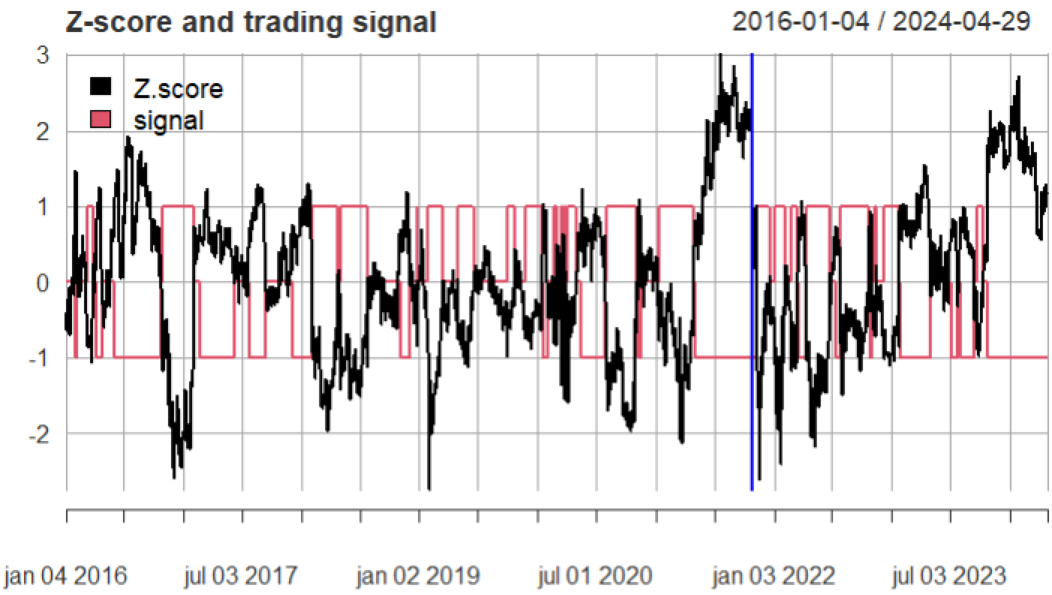



Figure 6: Z-score and trading signal

### 4.5.2. Performance of the strategy

Finally, we can compute the return profit and loss of the strategy. A simple way of doing this is directly from the spread and the trading signal. According to Figure 21, the spread return is computed as the difference between consecutive values of the spread. The traded return is calculated by multiplying the spread return by the lagged signal. The lag(signal) shifts the signal by a one-time step to align the position with the subsequent return. This ensures that the position taken at time t-1 affects the return at time t. In Figure 7 below, it´s possible to analyze this return profit and loss of the strategy, thus the return of the spread between two stocks. Some spikes indicate the volatility in the markets at that moment, as at the beginning of 2019 and March 2020, with the COVID-19 crisis.

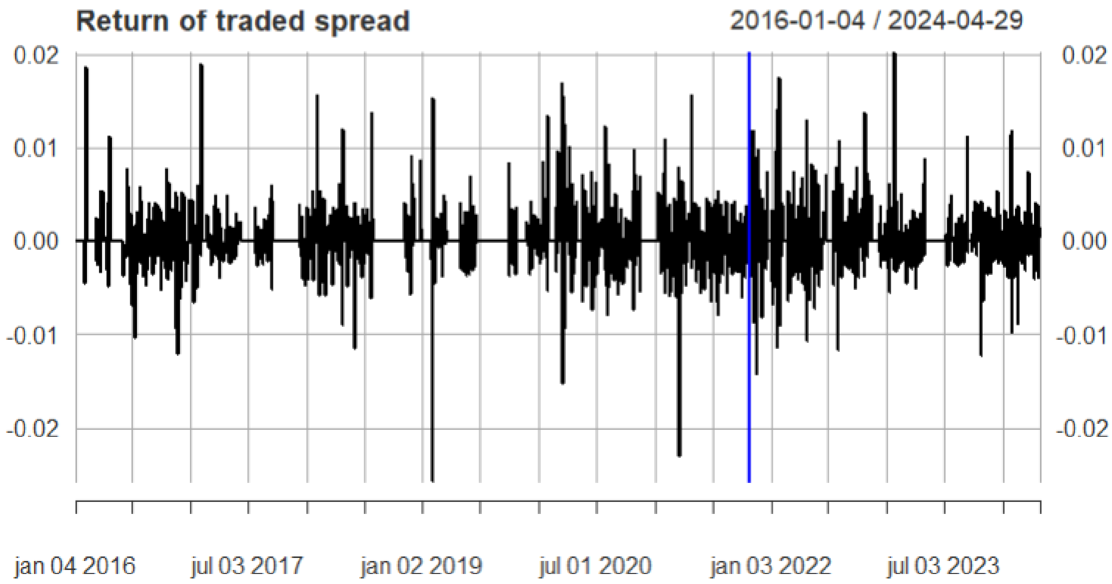



Figure 7: Return of traded spread

Concerning the wealth or cumulative profit and loss, there are two different ways to compute it. With reinvestment or not compounded. The same initial budget of, say, only 1€ is invested every time a new trade. With reinvestment, all the existing wealth at each time is fully reinvested. I will only show the compounded way because as we will see later, it is the most profitable, however, the not compounded is more reliable because returns on indexes like S&P500 are not compounded. In Figure 8 there is the plot of the cumulative profit and loss of the traded spread, showing resilience in the results, even after the end of the training period, showing robust results on this strategy and more specifically in this cointegrated pair.

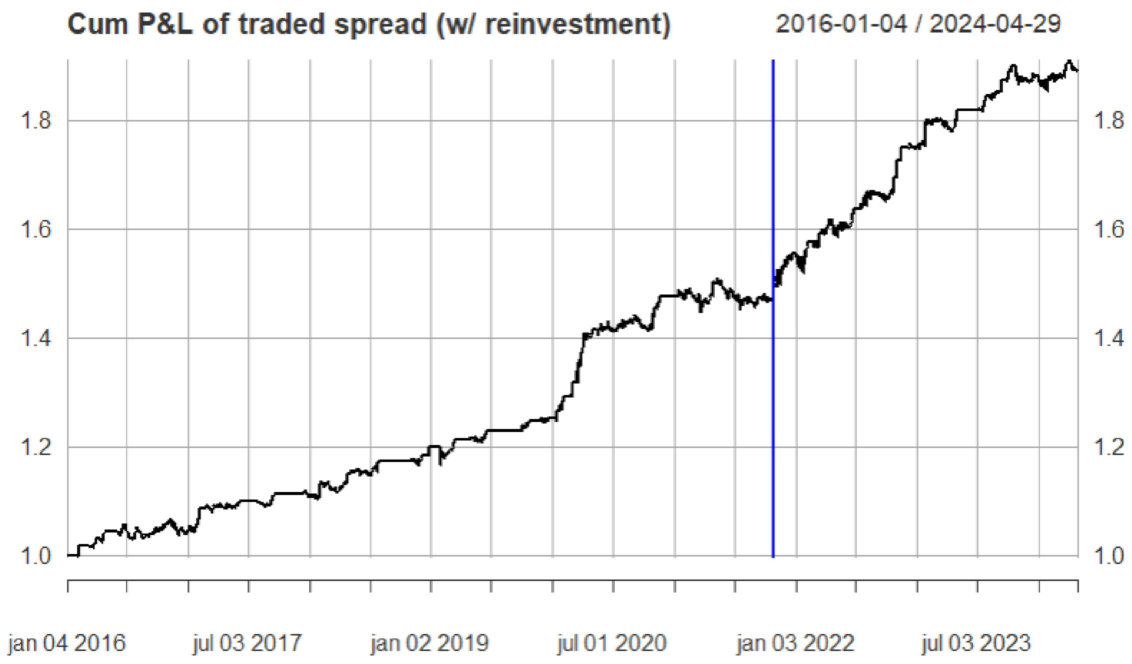



Figure 8: P&L of traded spread with reinvestment

We can observe that wealth increases steadily during the training period and in this case, it continues to rise after that. However, in some cases, it flattens out during the test or out-of-sample period, which is the one that counts.

The same plots can be obtained with the package PerformanceAnalytics, and the argument geometric determines whether it is compounded or not. However, I have used this library to plot the drawdowns of the traded spread returns.

The below Figure 9 shows the drawdown of the strategy. This plot shows part of the risk associated with this strategy. Larger drawdowns indicate higher risk as they can lead to extended non-realized losses in some periods, therefore, a strategy with smaller and shorter drawdowns is considered less risky. It´s notable the correlation between the drawdown and the times that there are not any mean reverting spread. For example, in 2021 we see a big drawdown and the z-score reaches a standard deviation of 3.

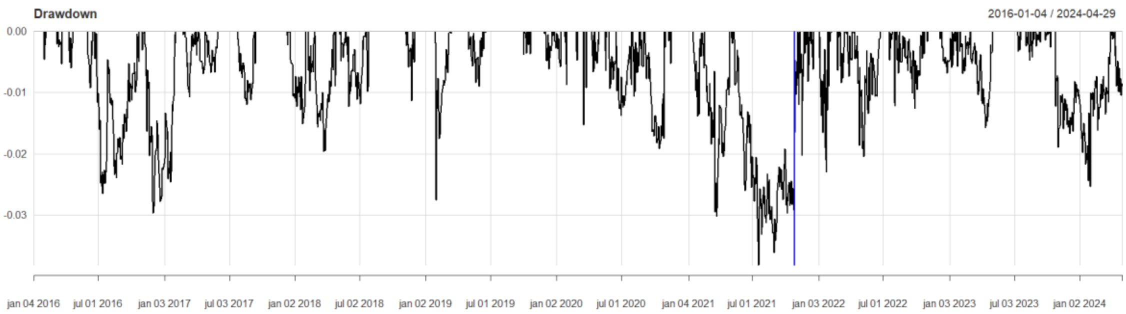



Figure 9: Drawdown

Now I will input the trading signal with the z-score. There is a two-panel plot to visualize the Z-score with trading signals and the cumulative profit and loss of the traded spread.

In Figure 10 we can see the Z-score and trading signal with the cumulative profit and loss of traded spread. This provides a clear view of the strategy's profitability and growth over the selected period. It helps to assess the long-term viability of the trading strategy. We can see the relation between trading signals and the profit and loss of the strategy. In the period from July 2021 to January 2022, only one trading signal was generated, resulting in a time of declining returns.

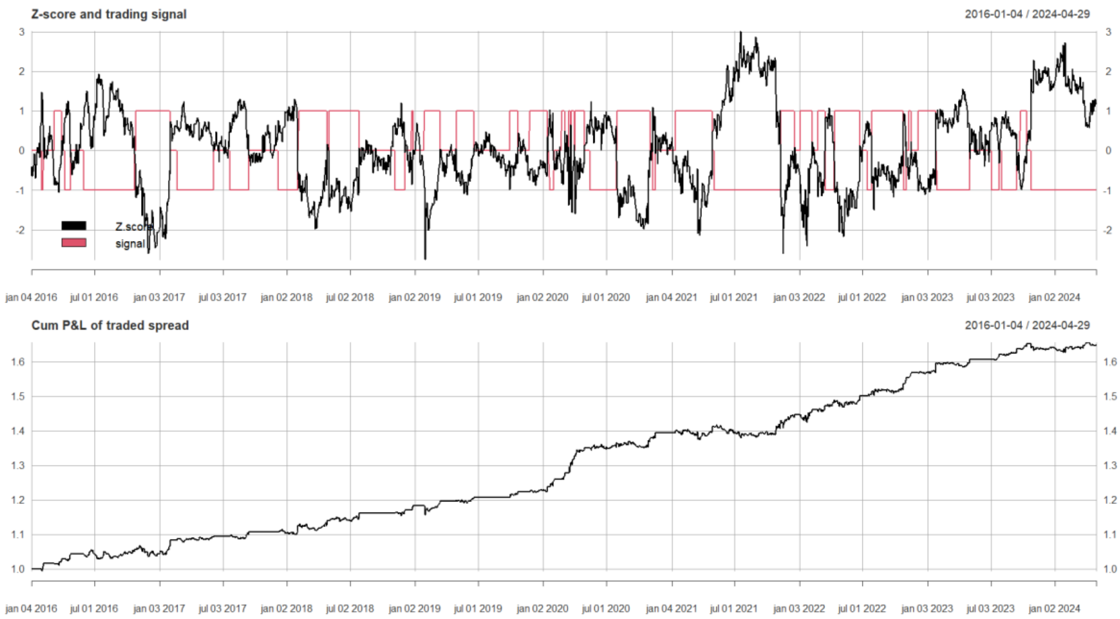



Figure 10: Trading signal and cumulative P&L

A more efficient way to figure out the return and P&L of the strategy is via the equivalent portfolio. The previous way of calculating the P&L directly from the spread can lead to wrong results if $\gamma$ or $\mu$ change over time. If that happens, then the spread may look very nice but it's not realistic because once I am in a trade, shouldn't change $\gamma$ or $\mu$. Let's use now the equivalent portfolio approach. First, I combine the normalized portfolio weights with the lagged trading signal to create the portfolio weights. After that, run the log returns from the log-prices and calculate the portfolio returns by applying the portfolio weights to the log-returns.

The plot shows the Z-score again and the cumulative profit and loss of the traded spread. Figure 11 shows that with this equivalent portfolio approach, the performance of this pairs trading strategy is better, accumulating profits from 1 to over 1.8. Without this approach, the returns were slightly above 1.6, as shown in Figure 10.

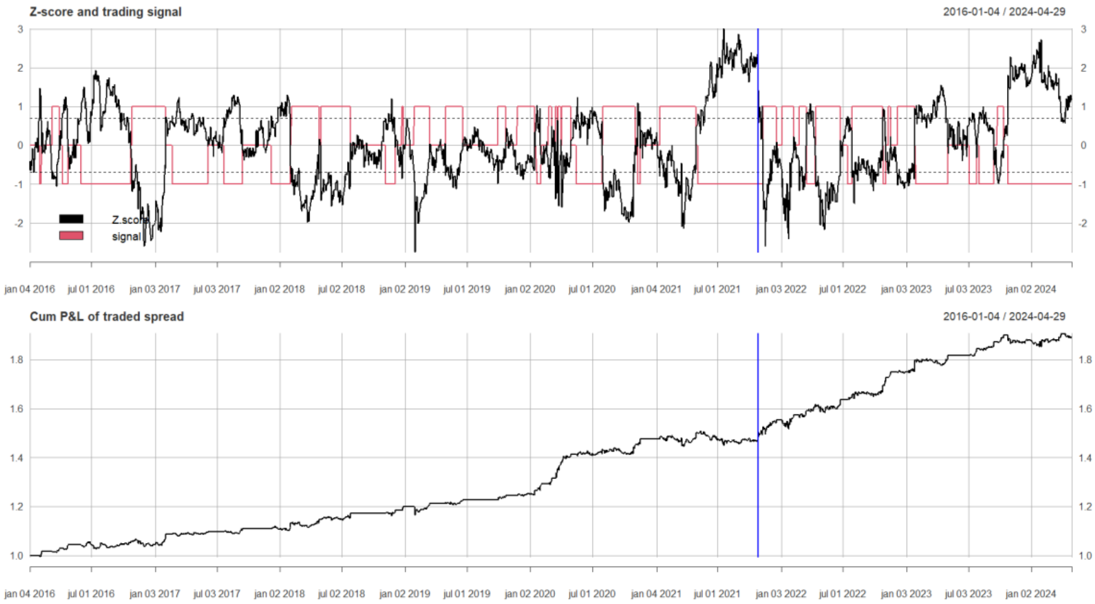



Figure 11: Equivalent portfolio approach code

## 4.6. Risk-Adjusted Performance Metrics

### 4.6.1. Capital Asset Pricing Model (CAPM)

Regarding risk-adjusted performance metrics, I have used some well-known metrics, sharpe ratio, and CAPM. The Capital Asset Pricing Model (CAPM) developed by William Sharpe, John Lintner, and Jan Mossin in the 1960s, represents a pivotal development in financial economics, Sharpe (1964) and Lintner (1965). It is a finance model that establishes a linear relationship between the required return on investment and risk. CAPM extends the principles of portfolio theory by Harry Markowitz, providing a framework to quantify the relationship between systematic risk and expected return for assets, Markowitz (1959), who argues that investors are risk averse and will choose a portfolio by trading off between risk and return for one investment period. Therefore, investors will choose efficient portfolios that minimize the variance of portfolio return, given a specific level of expected return, or maximize expected return, given a specific level of variance.

In theory, the capital asset pricing model is employed to set the investor required rate of return on a risky security given the non-diversifiable firm-specific risk, as the systematic risk will be eliminated in a well-diversified portfolio.

The CAPM equation according to Sharpe and Lintner's assumptions of risk-free borrowing and lending is expressed as follows:

$$E(Ri) \ = \ Rf + \beta i[E\ (Rm) - Rf] \tag{8}$$

Where E($Ri$) is the expected return on asset I, $Rf$ is the risk-free rate, $\beta i$ is the beta of asset I and E($Rm$) is the expected return of the market portfolio

Where market Beta:

$$\beta i \ = \ \frac{COV\ (Ri;Rm)}{\sigma^2\ (Rm)} \tag{9}$$

Where $COV(Ri, Rm)$ is the covariance between the return of the stock ($Ri$) and the return of the market ($Rm$) and $\sigma^2$ $(Rm)$ is the variance of the market return.

#### 4.6.1.1. BETA

The beta of a pairs trading strategy is typically very low and often close to zero. This is because pairs trading is a market-neutral strategy, meaning it is designed to be independent of the overall market movements. The strategy involves taking simultaneous long and short positions in two correlated stocks, with the expectation that their price movements will converge, irrespective of market direction.

The key point of beta in pairs trading is market neutrality since pairs trading aims to hedge out market risk by taking opposite positions in two correlated stocks. This neutralizes the effect of broader market movements on the strategy's performance. Then, the returns of the pairs trading strategy are driven by the relative price changes between the two stocks rather than the overall market movements, the beta of the strategy with respect to the market index is usually very low, close to zero and the short position in one stock offsets the market risk of the long position in the other stock, which contributes to the low beta.

The beta of a pairs trading strategy is usually close to zero, highlighting its market-neutral characteristic. This indicates that the strategy's performance is independent of overall market movements, relying instead on the relative price movements of the paired stocks.

To conclude, I will consider beta as 0 since this is a market-neutral strategy, being the value of CAPM the current risk-free rate. For that I will assume the 10-year treasury rate as of 30 April 2024 as a risk-free rate, that is 4,69% and it will be used as a benchmark to compare the result of the pairs trading strategy.

### 4.6.2. SHARPE RATIO

Roy (1952) was the first to suggest a risk-reward ratio to evaluate a strategy's performance. Sharpe (1966) applied Roy's ideas to Markowitz's mean-variance framework, in what has become one of the best known performance evaluation metrics.

Sharpe ratio has become the 'gold standard' of performance evaluation. Sharpe ratios are greatly affected by some of the statistical traits inherent to hedge fund strategies in general. The Sharpe ratio compares how well an equity investment performs to the rate of return on a risk-free investment, such as U.S. government treasury bonds or bills. When comparing funds or portfolios, investors should consider both absolute returns and risks. While one portfolio or fund could have higher returns, it is only a good investment if those higher returns do not come with additional risk. The formula of the sharpe ratio is the following:

$$Sharpe\ Ratio = \frac{Rp - Rf}{\sigma p} \qquad (10)$$

Where $Rp$ is the portfolio return, $Rf$ is the risk-free rate and $\sigma p$ is the standard deviation of the portfolio.

A good Sharpe Ratio largely depends on the investor's risk tolerance and investment objectives. Generally, any value greater than 1 is considered acceptable, while a ratio higher than 2. is considered very good, and a ratio of 3 or higher is considered excellent. However, it's important to note that a higher Sharpe Ratio may not always be better, as it could be a result of taking on excessive risk. Ultimately, investors should evaluate their portfolios based on their unique investment goals and risk tolerance, using the Sharpe Ratio as one of many tools for making informed investment decisions.

For the Sharpe Ratio calculation, I have used the annual return with reinvestment as the portfolio return, the risk-free rate is the 10-year treasury rate and I have estimated the standard deviation in Rstudio for every pair of stocks.

# 5. RESULTS AND DISCUSSION

## 5.1. COINTEGRATION RESULTS

For instance, let´s check the case of Visa and Mastercard. Based on the provided regression and unit root tests, I need to assess whether there is a cointegration relationship between **MA** and **V** based on the information about the residuals.

### Regression Model and Residuals

The given regression model:

$$\mathrm{MA}[\mathrm{i}] = 1.8454 \times \mathrm{V}[\mathrm{i}] - 47.0284 + \mathrm{R}[\mathrm{i}]$$
$$\text{where } \mathrm{R}[\mathrm{i}] = 0.9660\mathrm{R}[\mathrm{i}-1] + \epsilon[\mathrm{i}], \epsilon[\mathrm{i}] \sim \mathrm{N}(0{,}2.23422)$$

suggests that **MA** is regressed on **V** and the residuals R[i] are supposed to follow an AR(1) process. However, there's a warning that the residuals do not exhibit a perfect AR(1) structure. Now, let's review the unit root tests applied to the residuals that can be seen in Table 1 below:

| V-MA | |
|---|---|
| **Statistics** | **P-Value** |
| -5.618 | 0.00010 |
| -63.894 | 0.00010 |
| 0.963 | 0.00010 |
| -4.203 | 0.00219 |
| -38.439 | 0.00010 |
| -50.648 | 0.00372 |

Table 1: Unit root tests

Regarding the Augmented Dickey-Fuller (ADF), that is the test that I will rely more on since it tests the ability to account for higher-order autocorrelation, has flexibility in modeling different types of time series data, and its well-established presence in the literature and its ease to use, make a good choice for rely on testing for unit roots. It has a test statistic of -5.618 and a p-value of 0.00010, which indicates strong evidence against the null hypothesis of a unit root, suggesting that the residuals are stationary. The P-value is smaller than the most common significance level of 0.05, and it´s even more minor than the 0.01 level. Regarding the statistics, the value of -5.618 is smaller than the critical value of the ADF test for the significance level of 0.01 which is -3.43. The critical values for the ADF test are in the following Table 2, according to Cook (2001). In his research, Cook used several sample sizes. I will use the 80-sample size, even if mine is only 50. For all sample sizes (including I = o) and at all levels of significance, the MK critical values are greater in absolute value than those of Cheung and Lai (1995). Consequently, the MacKinnon values further exaggerate the low power problem associated with ADF tests. However, MacKinnon used to calculate critical values for any sample size, and this has become standard practice in empirical work, so I will use it too.

On the left, we have the significance level, and on the right, the matched critical value for this test.

| Significance Level | Critical Value |
|---|---|
| 1% | -4,076 |
| 5% | -3,466 |
| 10% | -3,159 |

Table 2- Critical values

A more negative test statistic falling in the rejection region means we have statistically significant evidence to conclude that the series is stationary and does not have a unit root.

By focusing on more negative values, these tests help identify when a time series is likely to be stationary, which is crucial for making meaningful inferences and avoiding false regressions in time series analysis.

For the cointegration tests, a p-value less than 0.05 indicates strong evidence against the null hypothesis, allowing me to reject it and I will rely on that value. However, a p-value less than 0.01 indicates even stronger evidence against the null hypothesis. If it´s greater than 0.05 and less than 0.10, the evidence is weak, and I will not have confidence in that data. Regarding the cointegration tests, it´s not usual for a pair of two stocks to pass on every unit root test, so I will accept that a pair is cointegrated if 3 of the 5 tests are below the 0.05 p-value (95% confidence).

Regarding the other tests, the Phillips-Perron (PP) test also confirms the stationarity of the residuals with a statistic of -63.894 and a p-value of 0.00010. Then Pantula, Gonzales-Farias, and Fuller's (PGFF) test unusually shows a statistic of 0.963 but a very low p-value of 0.00010, typically suggesting rejection of the null hypothesis of a unit root, though the interpretation might need context. Elliott, Rothenberg, and Stock DF-GLS (ERSD) test with a statistic of -4.203 and a p-value of 0.00219 also suggests that the residuals are stationary. Johansen's Trace Test (JOT) and Schmidt and Phillips Rho (SPR): Both of these cointegration-specific tests indicate no unit root, supporting the presence of cointegration. Regarding the case of the PGFF test that has a positive statistic, I will disregard that and consider an outlier, regarding the other decent values to reject the null hypothesis.

Therefore, despite the warning that the residuals are not perfectly autoregressive (AR(1)), almost all unit root tests on the residuals strongly suggest that they are stationary. Therefore, it can be concluded that Mastercard and Visa are cointegrated.

Regarding the cointegration tests, we can consider that 8 pairs have some sort of cointegration, so a p-value smaller than the significance level or a statistic smaller than the critical value. We have other pairs that are not cointegrated but have shown a good performance. But in the next step, we will see the results.

## 5.2. RISK-ADJUSTED PERFORMANCE METRICS RESULTS

We have to check the risk of the portfolio. The risk performance metrics for the investment portfolio were calculated for the period from January 1, 2016, to April 30, 2024. These metrics provide insights into the risk-adjusted returns and overall volatility of the portfolio. The key metrics considered include the Sharpe Ratio, standard deviation, and CAPM.

As I stated before, for the Sharpe Ratio calculation, I have used the annual return with reinvestment as the portfolio return, the risk-free rate is the 10-year treasury rate, which stands at 4,69% and then I have estimated the standard deviation in Rstudio for every pair of stocks. The standard deviation was calculated from the spread of this strategy in each pair. The spread is how much the price of two stocks will deviate from each other, so it will de the standard deviation of the strategy. As mentioned before, I have assumed the beta at 0 for this pairs trading strategy, so the value of CAPM will be the risk-free rat, which stands at 4,69% on 30 April 2024. For the cointegrated pairs, in Table 3, we can see that the pair with the best Sharpe Ratio, so, the pair with the best performance and less risk, was Visa and Mastercard (V-MA). The pair with the lowest Sharpe Ratio was the NVO-LLY, with a value of 0,05. This shows the correlation between standard deviation and Sharpe ratio, as this pair had the highest value in standard deviation and the lowest in Sharpe ratio, being a pair with higher risk than the remaining ones.

| Cointegrated Pairs | Annual return with/reinvestment % | Standard Deviation | Sharpe Ratio |
|---|---|---|---|
| PYPL-SQ | 11,20% | 8,16% | 0,80 |
| V-MA | 10,71% | 1,42% | 4,24 |
| ADBE-ADSK | 16,73% | 5,12% | 2,35 |
| SPGI-MCO | 13,25% | 2,05% | 4,17 |
| NVO-LLY | 4,88% | 8,27% | 0,02 |
| KO-PEP | 7,54% | 2,65% | 1,07 |
| Mean | 10,72% | 4,61% | 2,11 |

Table 3: Sharpe Ratio

## 5.3. PAIRS TRADING STRATEGY RESULTS

From 2016 to 2024, there are 6 pairs that I consider fully cointegrated and therefore, they are the pairs that I will first analyze below.

Regarding the returns on this pair trading strategy based on the cointegration approach, the results were very interesting. From 2016 to 2022, as we had interest rates at almost 0%, the return on this strategy is remarkable. However, from 2022 on with the increase in the interest rates to combat inflation, the United States has raised its rates to 5,5%. In Europe, the percentage is 4,25%, so some pairs are not very attractive in my point of view. The benchmark that I will compare this strategy with is the index S&P500.

The pair with the best-annualized return is ADBE-ADSK, Adobe, and Autodesk. The annual return with reinvestment was 16,73%. The total return of S&P500 from 01-01-2016 to 30-04-2024, the range of our strategy, was 149,79%, given an annualized return of 17,97%. With this being said, we can see that the pairs trading strategy, with the cointegration approach, is an asset for every trader who wants to profit from an arbitrage.

In Table 3 below, it is possible to realize that the total annual return with reinvestment for this pairs trading strategy was 10,72%. This value is below the annual return of the S&P500, however, this strategy is almost risk-free, avoiding the drawdowns of the index over the years. The pair with best performance was ADBE-ADSK and the pair with worst was NVO-LLY.

| Cointegrated Pairs Return | Without reinvestment | With reinvestment | Annual return without/reinvestment % | Annual return with/reinvestment % |
|---|---|---|---|---|
| PYPL-SQ | 1,7868 | 1,9335 | 9,44% | 11,20% |
| V-MA | 1,6485 | 1,8925 | 7,78% | 10,71% |
| ADBE-ADSK | 1,9520 | 2,3942 | 11,42% | 16,73% |
| SPGI-MCO | 1,7584 | 2,1041 | 9,10% | 13,25% |
| NVO-LLY | 1,4232 | 1,4064 | 5,08% | 4,88% |
| KO-PEP | 1,5030 | 1,6280 | 6,04% | 7,54% |
| Mean | 1,6787 | 1,8931 | 8,14% | 10,72% |

Table 4- Cointegrated pairs performance

Now let´s compare the results with the rest of the pairs, the non-cointegrated ones, to finally conclude if there is an advantage or not if the pairs are cointegrated. For that, Table 4 shows the performance of the non-cointegrated pairs.

| Pairs trading strategy investment return | Without reinvestment | With reinvestment | Annual return without/reinvestment % | Annual return with/reinvestment % |
|---|---|---|---|---|
| PANW-CRWD | 1,1260 | 1,0096 | 1,51% | 0,11% |
| BMBL-MTCH | 1,7645 | 1,9060 | 9,17% | 10,87% |
| AAPL-MSFT | 1,5186 | 1,6185 | 6,22% | 7,42% |
| JPM-GS | 1,3031 | 1,3187 | 3,64% | 3,82% |
| MRK-ABBV | 1,2509 | 1,2147 | 3,01% | 2,58% |
| BK-FAF | 1,7390 | 1,8721 | 8,87% | 10,46% |
| F-GM | 1,2961 | 1,2389 | 3,55% | 2,87% |
| GOOG-META | 1,6795 | 1,7873 | 8,15% | 9,45% |
| JNJ-PFE | 1,4594 | 1,5277 | 5,51% | 6,33% |
| NVDA-AMD | 1,1729 | 0,9781 | 2,07% | -0,26% |
| AVGO-QCOM | 1,3188 | 1,2692 | 3,83% | 3,23% |
| INTC-MU | 1,4681 | 1,3729 | 5,62% | 4,47% |
| WFC-C | 1,3251 | 1,2977 | 3,90% | 3,57% |
| BKNG-ABNB | 1,7176 | 1,9196 | 8,61% | 11,04% |
| WMT-COST | 1,4191 | 1,4653 | 5,03% | 5,58% |
| XOM-CVX | 1,1179 | 1,0967 | 1,41% | 1,16% |
| NFLX-DIS | 1,7015 | 1,7702 | 8,42% | 9,24% |
| MCD-SBUX | 1,2849 | 1,2797 | 3,42% | 3,36% |
| PM-MO | 1,4333 | 1,4806 | 5,20% | 5,77% |
| Mean | 1,4261 | 1,4433 | 5,11% | 5,32% |

Table 5: Non-cointegrated pair's performance

With the above, we can state that the annualized return with reinvestment, on pairs trading strategy without cointegrated pairs has a rate of 5,32 %, much below the return with cointegrated pairs. The pair with the poorer performance was NVDA-AMD. This pair is one of the pairs that have less cointegration according to my cointegration tests, and that corroborates with my thesis. We have some pairs with good returns like BMBL-MTCH and BK-FAF, however, in general, the overall average returns indicate a lesser benefit from using a non-cointegration pairs trading strategy, with reinvestment providing a slight edge. The overall annualized return for all the pairs in this strategy was 6,62%, showing a solid performance, however below the results of some other empirical works. The sample of cointegrated pairs only has six pairs, however, in my point of view, that is enough data to take an informed opinion on this.

To analyze the benchmark for this trading strategy, I have added the pair with the best results, and the one with the worst, then the cointegrated pairs and non-cointegrated ones. At least, index S&P500 and the mean of all pairs were added. The annual return that is being checked here is without reinvestment, as S&P500 don't have reinvestment, so we have a better view on performance. ADBE-ADSK emerges as a standout performer with high returns and low risk, as evidenced by its high Sharpe Ratio, low standard deviation, and high return. In contrast, NVDA-AMD shows poor performance with low returns, high volatility, and a negative Sharpe Ratio, being the worst pair in terms of risk and performance in this strategy. Cointegrated pairs show strong metrics. All 6 cointegrated pairs have a Sharpe ratio of 2,26 and a standard deviation of only 4,61%. Besides this, we have a strong annual return of 8,14%, being an outrageous value for a market-neutral strategy. The overall mean performance of the pairs trading strategy is moderate, however, behind the S&P500 index, which provides the highest returns but with substantial volatility. With this comparison, we can highlight the importance of pair selection in pairs trading and the contrasting success of different pairs under the strategy, with the cointegrated pairs presenting better results in terms of return and risk.

| Benchmark | Annual Return | Standard Deviation | Sharpe Ratio |
|---|---|---|---|
| ADBE-ADSK | 11,42% | 5,12% | 2,49 |
| NVDA-AMD | 2,07% | 18,32% | -0,27 |
| Cointegrated | 8,14% | 4,61% | 2,26 |
| Non Cointegrated | 5,11% | 8,54% | 0,20 |
| Mean | 5,84% | 7,60% | 0,70 |
| S&P500 | 17,97% | 15,29% | 2,37 |

Table 6: Benchmark

# 6. CONCLUSIONS AND FUTURE WORKS

In this research, I have intended to evaluate the performance of a pairs trading strategy based on the cointegration approach according to (Vidyamurthy, 2004) and to evaluate if there is a relation between returns and cointegration, using data from 50 US stocks over 8 years from 2016 to 2024. By employing the Engle-Granger cointegration test (Engle-Granger, 1987), and other unit-root tests, this thesis identified the pairs that had a long-term equilibrium relationship between prices, meaning that they were cointegrated. The results demonstrate that pairs trading with cointegrated pairs offer better performance compared to non-cointegrated pairs, aligning with the hypothesis that cointegrated pairs yield superior returns. Out of the 25 pairs analyzed, 6 pairs were found to be fully cointegrated.

My findings are consistent with previous studies, such as Gatev et al. (2006) and Vidyamurthy, (2004), which documented the profitability of pairs trading regarding the distance and cointegration approach, with Gatev et al. (2006) reached an average annualized excess returns of 11% for the period 1962-2002, being my research with a smaller time horizon and with a smaller return with no reinvestment in cointegrated pairs, being 8,14% from 2016 to 2024. As Vidyamurthy, 2004, I had chosen pairs within the same industry and with similar market dynamics. While this thesis results align with earlier works in demonstrating significant returns from cointegrated pairs, they also highlight the strategy's sustained relevance in contemporary markets, which have evolved significantly since those earlier studies.

This research challenges the Efficient Market Hypothesis, by illustrating that market inefficiencies, such as deviations in cointegrated pairs stock prices, can be exploited for profit. This study also contributes to the literature on the cointegration approach, emphasizing the importance of cointegration, in the current market environment, as a method for identifying and leveraging market anomalies. The conclusions in this research have practical implications for traders and fund managers seeking to implement pairs trading strategies, such as retail traders also. By using cointegration tests to identify trading pairs, investors and traders can improve their returns and reduce the risk in an uncertain market environment. This study provides a framework for applying pairs trading strategies in today´s situation of the markets, with high interest rates, highlighting its potential to yield better risk-adjusted returns compared to non-cointegrated pairs.

One limitation of this study is the assumption that historical relationships between stock pairs will persist in the future, which may not always be true, due to market dynamics being in change over the time. Additionally, the study don´t assume transaction costs, which would impact the strategy's performance and returns. Gatev et al. (2006) achieved returns of 11% annually net of transaction costs, being my study limited in this aspect. This research contributes to the improvement and application of advanced cointegration techniques to pairs trading within the recent market environment. This study fills a gap in the literature, by providing recent market data from US stock market and results of the strategy on a high interest rate environment, checking into the performance of pairs trading strategies using cointegration, especially considering the increased market volatility over the past decade.

From further research regarding the pairs trading topic, I have some insights. There is a lack of research on returns for European stocks and even more for Portuguese ones. The difficulty here is regarding data and volume. Portuguese stocks have low volume, leading to fewer trades available between pairs. Another research topic that can be evaluated is regarding other approaches. Distance and cointegration approaches are widely studied, with copula and stochastic method approaches being left behind. With the increase in the use of artificial intelligence, the research in the machine learning approach will increase.

Regarding the results, annualized return with reinvestment for cointegrated pairs was 10,72%, slightly below the S&P500's annualized return of approximately 17,97%. However, the strategy's lower risk profile and drawdown mitigation make it an attractive option for risk-averse investors. The pair ADBE-ADSK (Adobe and Autodesk) achieved the highest annualized return of 16,73%, showing the potential of the cointegration approach in pairs trading. The pair NVO-LLY (Novo Nordisk and Eli Lilly) had the lowest performance among the cointegrated pairs, indicating variability in returns even within cointegrated pairs. Regarding the non-cointegrated pairs, they had an overall annualized return of 5,32%, which is significantly lower than the cointegrated pairs. The pair NVDA-AMD (NVIDIA and AMD) was the worst performer in the non-cointegrated group, further supporting the thesis that cointegrated pairs offer more consistent and higher returns. The strategy's performance was notably better in a low-interest-rate environment (2016-2022). The rising interest rates post-2022 affected the attractiveness of some pairs. However, this strategy provided a hedge against market downturns, offering a stable return even during periods of high market volatility.

The research confirms that a pairs trading strategy based on the cointegration approach can deliver solid returns with reduced risk compared to a market index like the S&P500. While the overall returns were slightly lower than the S&P500, the strategy's market-neutral nature and resilience to market fluctuations make it an effective approach for risk-averse traders seeking consistent performance. Consequently, I can conclude that the cointegration approach leads to better returns over time.

# BIBLIOGRAPHICAL REFERENCES


Alexander, C., Giblin, I., & Weddington, W. (2002). Cointegration and asset allocation: a new active hedge fund strategy. Research in International Business and Finance, 16(5), 65-90.

Ashofteh, A., Bravo, J. M. (2021). A Conservative Approach for Online Credit Scor-ing. Expert Systems With Applications, Volume 176, p. 1-16, 114835.

Ashofteh, A., Bravo, J. M., Ayuso, M. (2021). A Novel Layered Learning Approach for Forecasting Respiratory Disease Excess Mortality during the COVID-19 pandemic. Atas da 21ª Conferência da Associação Portuguesa de Sistemas de Informação 2021, CAPSI 2021, Volume 2021-October, Code 183080.

Ashofteh, A., Bravo, J. M., Ayuso, M. (2022). A New Ensemble Learning Strategy for Panel Time-Series Forecasting with Applications to Tracking Respiratory Disease Ex-cess Mortality during the COVID-19 pandemic. Applied Soft Computing, 128, 109422.

Björck, Å. (1990). Least squares methods. Handbook of numerical analysis, 1, 465-652.

Bowen, D. A., & Hutchinson, M. C. (2016). Pairs trading in the UK equity market: Risk and return. The European Journal of Finance, 22(14), 1363-1387.

Bravo, J. M. (2024). Ensemble methods for Stock Market Prediction. In Meo, R. & Silvestri, F. (editors). Machine Learning and Principles and Practice of Knowledge Discovery in Databases. International Workshops of ECML PKDD 2023, Turin, Italy, September 18–22, 2023, Revised Selected Papers, Part V. MIDAS - The 8th Workshop on MIning DAta for financial applicationS

Bravo, J. M. (2020). Longevity-Linked Life Annuities: A Bayesian Model Ensemble Pricing Approach. CAPSI 2020 Proceedings. 29. https://aisel.aisnet.org/capsi2020/29.

Bravo, J. M. (2021). Forecasting Longevity for Financial Applications: A First Experiment with Deep Learning Methods. In: Kamp M. et al. (eds) Machine Learning and Principles and Practice of Knowledge Discovery in Databases. ECML PKDD 2021. Communications in Computer and Information Science, vol 1525. Springer, Cham. 232–249.

Bravo, J. M. (2022). Pricing Participating Longevity-Linked Life Annuities: A Bayesian Model Ensemble approach. European Actuarial Journal, 12, 125–159.

Bravo, J. M., & Ashofteh, A. (2023). Ensemble Methods for Consumer Price Inflation Forecasting. CAPSI 2023 - 23rd Conference of the Portuguese Association for Information Systems (23ª Conferência da Associação Portuguesa de Sistemas de Informação), pp. 317-336. DOI: 10.18803/capsi.v23.317-336

Bravo, J. M., Ayuso, M. (2021a). Linking Pensions to Life Expectancy: Tackling Conceptual Uncertainty through Bayesian Model Averaging. Mathematics, 9(24): 3307, p. 1-27.

Bravo, J. M., Ayuso, M., Holzmann, R., Palmer, E. (2021). Addressing the Life Expectancy Gap in Pension Policy. Insurance: Mathematics and Economics, 99, 200-221.

Bravo, J. M., Ayuso, M., Holzmann, R., & Palmer, E. (2023). Intergenerational Actuar-ial Fairness when Longevity Increases: Amending the Retirement Age. Insurance: Mathematics and Economics, 113, 161-184.

Bravo, J. M., El Mekkaoui, N. (2022). Short-Term CPI Inflation Forecasting: Probing with Model Combinations. In: Rocha et al. (eds) Information Systems and Technolo-gies. WorldCIST 2022. Lecture Notes in Networks and Systems, vol 468. Springer, Cham, pp. 564–578. https://doi.org/10.1007/978-3-031-04826-5_56

Caldeira, J., & Moura, G. V. (2013). Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. Available at SSRN 2196391.

Cheung, Y. W., & Lai, K. S. (1997). Bandwidth selection, whitening, and the power of the Phillips-Perron test. Econometric Theory, 13(5), 679-691.

Cook, S. (2001). Finite-sample critical values of the Augmented Dickey-Fuller statistic: a note on lag order. ECONOMIC ISSUES-STOKE ON TRENT-, 6(2), 31-46.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366a), 427-431.

Elliott, R. J., Van Der Hoek*, J., & Malcolm, W. P. (2005). Pairs trading. Quantitative Finance, 5(3), 271-276.

Engelberg, J., Gao, P., & Jagannathan, R. (2009, November). Anatomy of pairs trading: the role of idiosyncratic news, common information, and liquidity. In Third Singapore International Conference on Finance.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society, 251-276.

Fama, E. F. (1970). Efficient market hypothesis: A review of theory and empirical work. Journal of Finance, 25(2), 28-30.

Gatev, et al. (1998). Pairs trading: Performance of a relative value arbitrage rule. Technical report, Working Papers ysm26, Yale School of Management.

Gatev, et al. (2006). Pairs trading: Performance of a relative-value arbitrage rule. The Review of Financial Studies, 19(3), 797-827.

Hendry, D. F. (1995). Dynamic econometrics. Oxford University Press.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.

Krauss, C. (2017). Statistical arbitrage pairs trading strategies: Review and outlook. Journal of Economic Surveys, 31(2), 513-545.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics, 54(1-3), 159-178.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. The journal of finance, 20(4), 587-615.

Markowitz, H. M. (1991). Foundations of portfolio theory. The journal of finance, 46(2), 469-477.

Mudchanatongsuk, S., Primbs, J. A., & Wong, W. (2008). Optimal pairs trading: A stochastic control approach. In 2008 American control conference (pp. 1035-1039). IEEE.

Nicholas, J. G. (2000). Market neutral investing. Princeton, NJ: Bloomberg Press.

Patterson, K. D. (2000). An introduction to applied econometrics: a time series approach (Vol. 1). New York: Palgrave.

Sharpe, W. F. (1966). Mutual fund performance. The Journal of business, 39(1), 119-138.

Vidyamurthy, G. (2004). Pairs Trading: quantitative methods and analysis (Vol. 217). John Wiley & Sons.

Wei, L., & Keogh, E. (2006). Semi-supervised time series classification. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 748-753).

Whistler, M. (2004). Trading pairs: capturing profits and hedging risk with statistical arbitrage strategies (Vol. 216). John Wiley & Sons.

Wooldridge, J. M. (1996). Introductory Econometrics: A Modern Approach 3rd ed.

# APPENDIX A

| | PYPL-SQ | | V-MA | | ADBE-ADSK | |
|---|---|---|---|---|---|---|
| **Unit Root Tests of Residuals** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** |
| Augmented Dickey-Fuller (ADF) | -2,899 | 0,13613 | -5.618 | 0.00010 | -3,664 | 0,02055 |
| Phillips-Perron (PP) | -22,654 | 0,03519 | -63.894 | 0.00010 | -25,024 | 0,02152 |
| Pantula, Gonzales-Farias and Fuller (PGFF) | 0,988 | 0,04338 | 0.963 | 0.00010 | 0,989 | 0,05725 |
| Elliott, Rothenberg and Stock DF-GLS (ERSD) | -2,995 | 0,01597 | -4.203 | 0.00219 | -3,033 | 0,0143 |
| Johansen's Trace Test (JOT) | -19,739 | 0,06185 | -38.439 | 0.00010 | -15,624 | 0,19575 |
| Schmidt and Phillips Rho (SPR) | -33,325 | 0,01401 | -50.648 | 0.00372 | -28,89 | 0,02589 |

| PANW-CRWD | | BMBL-MTCH | | SPGI-MCO | | AAPL-MSFT | | NVO-LLY | |
|---|---|---|---|---|---|---|---|---|---|
| **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** |
| -1,558 | 0,71967 | -2,395 | 0,32847 | -3,9 | 0,00977 | -1,241 | 0,83804 | -3,406 | 0,04135 |
| -4,629 | 0,76381 | -11,064 | 0,33452 | -45,294 | 0,00123 | -7,561 | 0,53288 | -37,15 | 0,00546 |
| 0,996 | 0,62086 | 0,986 | 0,27946 | 0,969 | 0,00134 | 0,995 | 0,32028 | 0,981 | 0,00673 |
| -0,798 | 0,61468 | -1,596 | 0,27805 | -2,306 | 0,0781 | -1,049 | 0,5007 | -2,684 | 0,03395 |
| -5,442 | 0,96755 | -20,057 | 0,05603 | -30,662 | 0,00495 | -11,318 | 0,51443 | -44,254 | 0,0001 |
| -2,99 | 0,92651 | -6,166 | 0,71848 | 1,569 | 0,99987 | -4,616 | 0,82614 | -11,688 | 0,3762 |

| JPM-GS | | MRK-ABBV | | BK-FAF | | KO-PEP | | F-GM | |
|---|---|---|---|---|---|---|---|---|---|
| **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** |
| -2,034 | 0,50211 | -2,096 | 0,47323 | -2,273 | 0,39057 | -3,686 | 0,01945 | -2,048 | 0,49574 |
| -8,245 | 0,48574 | -8,589 | 0,46715 | -13,704 | 0,19541 | -29,821 | 0,00926 | -8,338 | 0,48073 |
| 0,996 | 0,40222 | 0,996 | 0,38186 | 0,994 | 0,21177 | 0,985 | 0,02218 | 0,996 | 0,48476 |
| -1,434 | 0,34067 | -1,912 | 0,16419 | -1,73 | 0,21792 | -3,498 | 0,00733 | -1,141 | 0,46227 |
| -8,208 | 0,80636 | -12,234 | 0,44343 | -11,835 | 0,47306 | -18,578 | 0,08778 | -15,236 | 0,22084 |
| -6,062 | 0,71896 | -10,637 | 0,42453 | -10,385 | 0,43614 | -29,358 | 0,02426 | -6,354 | 0,69717 |

| GOOG-META | | JNJ-PFE | | NVDA-AMD | | AVGO-QCOM | | INTC-MU | |
|---|---|---|---|---|---|---|---|---|---|
| **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** |
| -2,151 | 0,44741 | -1,695 | 0,65842 | -0,672 | 0,94361 | -1,584 | 0,70978 | 0,686 | 0,99659 |
| -8,454 | 0,47444 | -9,256 | 0,4311 | -2,316 | 0,91346 | -5,552 | 0,69044 | 1,509 | 0,99732 |
| 0,996 | 0,39396 | 0,995 | 0,36028 | 0,999 | 0,84704 | 0,997 | 0,57235 | 1,001 | 0,99385 |
| -1,662 | 0,24618 | -1,537 | 0,29792 | -0,941 | 0,54731 | -1,34 | 0,3797 | 1,23 | 0,99021 |
| -15,726 | 0,19171 | -8,802 | 0,75189 | -33,024 | 0,00275 | -16,98 | 0,14181 | -7,926 | 0,82765 |
| -9,48 | 0,47778 | -8,209 | 0,55888 | -2,528 | 0,95851 | -5,097 | 0,79092 | -6,644 | 0,67555 |

| WFC-C | | BKNG-ABNB | | WMT-COST | | XOM-CVX | | NFLX-DIS | |
|---|---|---|---|---|---|---|---|---|---|
| **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** | **Statistics** | **P-Value** |
| -2,712 | 0,19101 | -2,117 | 0,45953 | -2,145 | 0,45032 | -2,438 | 0,31329 | -1,924 | 0,55317 |
| -9,384 | 0,42424 | -10,031 | 0,3902 | -10,195 | 0,38041 | -8,532 | 0,47025 | -7,522 | 0,53596 |
| 0,996 | 0,46758 | 0,986 | 0,24303 | 0,995 | 0,29086 | 0,997 | 0,60348 | 0,995 | 0,37004 |
| -1,195 | 0,43992 | -2,015 | 0,14009 | -1,333 | 0,38261 | -0,81 | 0,60412 | -1,642 | 0,25461 |
| -8,559 | 0,77487 | -7,64 | 0,85199 | -14,887 | 0,24671 | -9,388 | 0,69665 | -10,371 | 0,60385 |
| -12,801 | 0,32498 | -7,331 | 0,63864 | -6,474 | 0,6882 | -12,989 | 0,31634 | -7,06 | 0,64452 |

| MCD-SBUX | | PM-MO | |
|---|---|---|---|
| **Statistics** | **P-Value** | **Statistics** | **P-Value** |
| -1,8 | 0,60997 | -2,918 | 0,1305 |
| -8,632 | 0,46482 | -16,456 | 0,12084 |
| 0,995 | 0,31055 | 0,992 | 0,14275 |
| -1,702 | 0,22963 | -2,851 | 0,02235 |
| -9,989 | 0,63993 | -12,069 | 0,45568 |
| -4,053 | 0,8671 | -19,296 | 0,11709 |

Table 7: Unit Root test of residuals

| Pairs trading strategy investment return | Without reinvestment | With reinvestment | Annual return without/reinvestment % | Annual return with/reinvestment % |
|---|---|---|---|---|
| PANW-CRWD | 1,1260 | 1,0096 | 1,51% | 0,11% |
| BMBL-MTCH | 1,7645 | 1,9060 | 9,17% | 10,87% |
| AAPL-MSFT | 1,5186 | 1,6185 | 6,22% | 7,42% |
| JPM-GS | 1,3031 | 1,3187 | 3,64% | 3,82% |
| MRK-ABBV | 1,2509 | 1,2147 | 3,01% | 2,58% |
| BK-FAF | 1,7390 | 1,8721 | 8,87% | 10,46% |
| F-GM | 1,2961 | 1,2389 | 3,55% | 2,87% |
| GOOG-META | 1,6795 | 1,7873 | 8,15% | 9,45% |
| JNJ-PFE | 1,4594 | 1,5277 | 5,51% | 6,33% |
| NVDA-AMD | 1,1729 | 0,9781 | 2,07% | -0,26% |
| AVGO-QCOM | 1,3188 | 1,2692 | 3,83% | 3,23% |
| INTC-MU | 1,4681 | 1,3729 | 5,62% | 4,47% |
| WFC-C | 1,3251 | 1,2977 | 3,90% | 3,57% |
| BKNG-ABNB | 1,7176 | 1,9196 | 8,61% | 11,04% |
| WMT-COST | 1,4191 | 1,4653 | 5,03% | 5,58% |
| XOM-CVX | 1,1179 | 1,0967 | 1,41% | 1,16% |
| NFLX-DIS | 1,7015 | 1,7702 | 8,42% | 9,24% |
| MCD-SBUX | 1,2849 | 1,2797 | 3,42% | 3,36% |
| PM-MO | 1,4333 | 1,4806 | 5,20% | 5,77% |
| PYPL-SQ | 1,7868 | 1,9335 | 9,44% | 11,20% |
| V-MA | 1,6485 | 1,8925 | 7,78% | 10,71% |
| ADBE-ADSK | 1,9520 | 2,3942 | 11,42% | 16,73% |
| SPGI-MCO | 1,7584 | 2,1041 | 9,10% | 13,25% |
| NVO-LLY | 1,4232 | 1,4064 | 5,08% | 4,88% |
| KO-PEP | 1,5030 | 1,6280 | 6,04% | 7,54% |
| Mean | 1,6787 | 1,8931 | 5,84% | 6,62% |

Table 8: Pairs trading strategy investment return

| Cointegrated Pairs | Without reinvestment | With reinvestment | Annual return without/reinvestment % | Annual return with/reinvestment % |
|---|---|---|---|---|
| PYPL-SQ | 1,7868 | 1,9335 | 9,84% | 11,20% |
| V-MA | 1,6485 | 1,8925 | 8,11% | 10,71% |
| ADBE-ADSK | 1,9520 | 2,3942 | 11,90% | 16,73% |
| SPGI-MCO | 1,7584 | 2,1041 | 9,48% | 13,25% |
| NVO-LLY | 1,4232 | 1,4064 | 5,29% | 4,88% |
| KO-PEP | 1,5030 | 1,6280 | 6,29% | 7,54% |
| Mean | 1,6787 | 1,8931 | 8,48% | 10,72% |

Table 9: Pairs trading return for cointegrated pairs

| Non Cointegrated Pairs | Annual return with/reinvestment % | Standard Deviation | Sharpe Ratio |
|---|---|---|---|
| PANW-CRWD | 0,12% | 8,72% | -0,52 |
| BMBL-MTCH | 11,33% | 8,66% | 0,77 |
| AAPL-MSFT | 7,73% | 5,87% | 0,52 |
| JPM-GS | 3,98% | 7,89% | -0,09 |
| MRK-ABBV | 2,68% | 10,54% | -0,19 |
| BK-FAF | 10,90% | 8,74% | 0,71 |
| F-GM | 2,99% | 9,96% | -0,17 |
| GOOG-META | 9,84% | 4,61% | 1,12 |
| JNJ-PFE | 6,60% | 4,38% | 0,44 |
| NVDA-AMD | -0,27% | 18,32% | -0,27 |
| INTC-MU | 4,66% | 7,85% | 0,00 |
| BKNG-ABNB | 11,49% | 4,99% | 1,36 |
| WMT-COST | 5,82% | 3,86% | 0,29 |
| XOM-CVX | 1,21% | 10,12% | -0,34 |
| NFLX-DIS | 9,63% | 12,63% | 0,39 |
| MCD-SBUX | 3,50% | 6,80% | -0,18 |
| PM-MO | 6,01% | 6,43% | 0,20 |
| Mean | 5,17% | 7,39% | 0,21 |

Table 10: Sharpe Ration non-cointegrated pairs

```r
library(quantmod)
library(egcm)  #install.packages("egcm")
egcm.set.default.pvalue(0.01)
```

Figure 12: Library´s used

```r
V_prices <- Ad(getSymbols("V", from = "2016-01-01", to = "2024-04-30", auto.assign = FALSE))
MA_prices <- Ad(getSymbols("MA", from = "2016-01-01", to = "2024-04-30", auto.assign = FALSE))
plot(cbind(V_prices, MA_prices), legend.loc = "topleft", main = "prices")

res <- egcm(V_prices, MA_prices)
summary(res)
plot(res)
```

Figure 13: Cointegration Code

```r
library(xts)
library(quantmod)

# set begin-end dates and stock namelist
begin_date <- "2016-01-01"
end_date <- "2024-04-30"
stock_namelist <- c("V", "MA")

# download data from YahooFinance
prices <- xts()
for (stock_index in 1:length(stock_namelist))
  prices <- cbind(prices, Ad(getSymbols(stock_namelist[stock_index],
                                        from = begin_date, to = end_date, auto.assign = FALSE)))
colnames(prices) <- stock_namelist
tclass(prices) <- "Date"
T <- nrow(prices)  # number of days

# interpolate NAs
anyNA(prices)
#> [1] TRUE
prices <- na.approx(prices)

# some visual exploration
str(prices)
head(prices)
tail(prices)
plot(prices, legend.loc = "topleft", main = "Stock prices")
```

Figure 14: Prices Loading

```r
T_trn <- round(0.7*T)  # define the training set
T_tst <- T - T_trn
y1 <- logprices[, 1]
y2 <- logprices[, 2]

# do LS regression
ls_coeffs <- coef(lm(y1[1:T_trn] ~ y2[1:T_trn]))
ls_coeffs

mu <- ls_coeffs[1]
gamma <- ls_coeffs[2]
```

Figure 15: Training set and LS regression

```r
# spread
spread <- y1 - gamma*y2
{ plot(spread, main = "Spread")
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
```

Figure 16: Spread Equation

```r
# normalized portfolio
w_ref <- c(1, -gamma)/(1+gamma)
sum(abs(w_ref))
#> [1] 1
w_spread <- matrix(w_ref, T, 2, byrow = TRUE)
w_spread <- xts(w_spread, index(y1))
colnames(w_spread) <- c("w1", "w2")

# resulting normalized spread
spread <- rowSums(cbind(y1, y2) * w_spread)
spread <- xts(spread, index(y1))
colnames(spread) <- "spread"
{ plot(spread, main = "Spread (from normalized portfolio)")
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
```

Figure 17: Normalized Portfolio and Normalized Spread

```
spread_mean <- mean(spread[1:T_trn], na.rm = TRUE)
spread_var <- as.numeric(var(spread[1:T_trn], na.rm = TRUE))
Z_score <- (spread-spread_mean)/sqrt(spread_var)
colnames(Z_score) <- "Z-score"
threshold_long <- threshold_short <- Z_score
threshold_short[] <- .7
threshold_long[] <- -.7

{ plot(Z_score, main = "Z-score")
  lines(threshold_short, lty = 2)
  lines(threshold_long, lty = 2)
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
```

Figure 18: Z-score code

```
spread_sd <- sqrt(spread_var)
print(sqrt(spread_var))
```

Figure 19: Spread Standard Deviation

```
# we define a function for convenience and future use
generate_signal <- function(Z_score, threshold_long, threshold_short) {
  signal <- Z_score
  colnames(signal) <- "signal"
  signal[] <- NA

  #initial position
  signal[1] <- 0
  if (Z_score[1] <= threshold_long[1]) {
    signal[1] <- 1
  } else if (Z_score[1] >= threshold_short[1])
    signal[1] <- -1

  # loop
  for (t in 2:nrow(Z_score)) {
    if (signal[t-1] == 0) {  #if we were in no position
      if (Z_score[t] <= threshold_long[t]) {
        signal[t] <- 1
      } else if(Z_score[t] >= threshold_short[t]) {
        signal[t] <- -1
      } else signal[t] <- 0
    } else if (signal[t-1] == 1) {  #if we were in a long position
      if (Z_score[t] >= 0) signal[t] <- 0
      else signal[t] <- signal[t-1]
    } else {  #if we were in a short position
      if (Z_score[t] <= 0) signal[t] <- 0
      else signal[t] <- signal[t-1]
    }
  }
  return(signal)
}

# now just invoke the function
signal <- generate_signal(Z_score, threshold_long, threshold_short)

{ plot(cbind(Z_score, signal), main = "Z-score and trading signal", legend.loc = "topleft")
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
```

Figure 20: Z-score and trading signal code

```r
# let's compute the PnL directly from the signal and spread
spread_return <- diff(spread)
traded_return <- spread_return * lag(signal)    # NOTE THE LAG!!
traded_return[is.na(traded_return)] <- 0
colnames(traded_return) <- "traded spread"

{ plot(traded_return, main = "Return of traded spread")
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
```

Figure 21: Return of traded spread code

```r
{ plot(cumprod(1 + traded_return), main = "Cum P&L of traded spread (w/ reinvestment)")
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
```

Figure 22: P&L of traded spread with reinvestment code

```r
library(PerformanceAnalytics)
{ chart.CumReturns(traded_return, main = "Cum P&L of traded spread (w/ reinvestment)",
                   geometric = TRUE, wealth.index = TRUE)
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }

{ chart.Drawdown(traded_return, main = "Drawdown", lwd = 1)
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
```

Figure 23: Drawdown code

```r
par(mfrow = c(2, 1))
plot(cbind(Z_score, signal)["2016::2024"], main = "Z-score and trading signal",
     legend.loc = "bottomleft")
plot(1 + cumsum(traded_return)["2016::2024"], main = "Cum P&L of traded spread")
```

Figure 24: Trading signal with Z-score code

```r
# combine the ref portfolio with trading signal
w_portf <- w_spread * matrix(lag(signal), T, 2)    # NOTE THE LAG!!

# now compute the PnL from the log-prices and the portfolio
X <- diff(cbind(y1, y2))  #compute log-returns from log-prices
portf_return <- xts(rowSums(X * w_portf), index(X))
portf_return[is.na(portf_return)] <- 0
colnames(portf_return) <- "portfolio"

par(mfrow = c(2, 1))
{ plot(cbind(Z_score, signal), main = "Z-score and trading signal", legend.loc = "bottomleft")
  lines(threshold_short, lty = 2)
  lines(threshold_long, lty = 2)
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
{ plot(cumprod(1 + portf_return), main = "Cum P&L of traded spread")
  addEventLines(xts("", index(y1[T_trn])), lwd = 2, col = "blue") }
```

Figure 25: Equivalent portfolio approach

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa