UNIVERSIDADE CATÓLICA PORTUGUESA

Pairs Trading

Cointegration-based methods

Applied to the Cryptocurrency Market

Daniel da Silva Carvalho

Católica Porto Business School

2021

UNIVERSIDADE CATÓLICA PORTUGUESA

Pairs Trading

Cointegration-based methods

Applied to the Cryptocurrency Market

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

por

Daniel da Silva Carvalho

sob orientação de

Paulo Alves

Católica Porto Business School

Universidade Católica Portuguesa

Outubro 2021

Acknowledgements

At first, I would like to acknowledge my family for all the unconditional support given throughout these academic years. My friends and coworkers, specifically Diogo Tavares for helping me with the python code. The academic professors that always guided me in the right direction, particularly Prof. Dr. Paulo Alves, my supervisor, to who I am immensely grateful. Lastly, I want to acknowledge Católica Porto Business School for its high standard of quality and excellence in education.

Abstract

In statistical arbitrage strategies such as Pairs Trading, the use of cointegration tests has been well-established in the field of econometrics and economics. In this dissertation, the cryptocurrency market was selected to implement, and compare the three best-known types of cointegration tests: the Augmented Dickey-Fuller test, Johansen's test, and Phillips Peron's test. Pairs are tested for cointegration over a 3-month and a 6-month window and then traded in a window of the same length. The cryptocurrencies included in the study are 10 cryptocurrencies with the highest market capitalization between January 1st 2020 to July 1st 2021. The performance of each portfolio is compared with their corresponding buy and hold benchmark. Although all portfolios underperformed their buy and hold benchmark, with and without transaction costs, the 6-month trading and testing procedure yielded a higher return compared with the 3-month procedure. Of the three proposed approaches, Augmented Dickey-Fuller test was best at predicting a cointegrated relationship. We can conclude that cointegrated relationships between cryptocurrencies are more likely to hold over longer periods of time and Engle-Granger´s approach employing the Augmented Dickey-Fuller test was the best predictor of cointegration relationships with an excess mean return of 338%.

Keywords: Pairs Trading; Cointegration; Cryptocurrency; Time Series.

About cointegration.

A drunk and his dog leave the bar after a night of drinking. The path of the drunk looks much like a random walk. The dog follows, but is slow and falls behind of its owner. The dog will go this way and that, wherever it's nose leads it. However, whenever she falls behind too much the drunk calls for the dog. The dog is loyal and catches up with the owner. So the drunk and his dog go forth wandering aimlessly at night, together.

by Michael P. Murry

Resumo

Em estratégias de arbitragem estatística como *Pairs Trading*, o uso de testes de cointegração tem assumido grande relevância no campo da econometria e da economia. Nesta dissertação, o mercado de criptomoedas foi selecionado para implementar e comparar os três tipos mais conhecidos de testes de cointegração: o teste Dickey-Fuller Aumentado, o teste de Johansen e o teste de Phillips Peron. Os pares são testados e formados em períodos com a duração de 3 e 6 meses, e, em seguida, executados em períodos de simulação com mesma duração. As criptomoedas incluídas no estudo são as 10 criptomoedas com a maior capitalização de mercado entre 1 de janeiro de 2020 e 1 de julho de 2021. O desempenho de cada carteira é comparado com o portefólio de referência. Embora todas as carteiras tenham apresentado um retorno inferior ao portefólio de referência, independentemente de haver ou não custos de transação, o procedimento de formação e simulação de 6 meses foi o que obteve a maior rentabilidade em comparação com o procedimento de 3 meses. Das três abordagens propostas à cointegração, o teste Dickey-Fuller Aumentado foi a obteve melhores resultados. Deste modo, conclui-se que as relações de cointegração entre as criptomoedas tendem a ser mais consistentes em períodos de tempo mais alargados e, a abordagem de Engle-Granger, aplicando o teste de Dickey-Fuller Aumentado, foi o que melhor conseguiu prever relações de cointegração entre os pares, com um retorno excessivo médio de 338%.

Palavras-chave: Pairs Trading; Cointegração; Criptomoeda; Séries Temporais.

Table of Contents

List of figures

List of tables

1 Introduction

In 2003, Clive Granger and Robert Engle were awarded the Nobel Prize in Economics for an article published in 1987 in one of the most cited academic journals in the field of econometrics – Econometrica. The paper concerned cointegration, which was already introduced by Granger in 1981 but the Econometrica paper in 1987 meant a rapid take-off for the idea. Today, the Engle and Granger (1987) paper are one of the most cited papers in econometric time series research. Granger came up with the concept of cointegration by attempting to prove to his colleague David Hendry that pairs of integrated series could not form a stationary process and in an attempt to prove Henry wrong, Granger discovered that Henry was right and generalized the property to cointegration. The first 1981 article by Granger was rejected for many reasons, such as the need of rewriting the proof of theorems and lack of empirical applications. Granger then started to work with his colleague Engle to perfect the article which led to the revolutionary 1987 article "Cointegration and Error Correction: Representation Estimation and Testing" and to their Nobel Prize award in Economics (Syczewska, 2011).

At the same time, a quantitative analyst named Nunzio Tartaglia at the American bank Morgan Stanley led a team of mathematicians, physicists, computer scientists to research strategies that would detect arbitrage opportunities in financial markets based on quantitative methods in the late 80s. The team's work resulted in a speculative trading strategy, pairs trading, which took advantage of cointegrated relationships of assets in financial markets in order to detect arbitrage opportunities.

This strategy attempts to identify a pair of stocks with similar historical price movements exposed to similar systematic risk over a certain period. Then, whether there is sufficient divergence between prices, a short-long position simultaneously established bets that the pair's divergence is temporary and will eventually converge if they ever diverge (Caldeira & Moura, 2013). Summarily, this strategy consists of buying an undervalued stock and shorting an overvalued stock, in order. Expectedly, the long position will increase in value, and the short position will decline in value. If this happens, and the positions are of equal size, the counterparty will benefit. In this sense, the strategy is often described as market neutral (Caldeira & Moura, 2013). Thereby, the strategy was expected to perform well both when the market was in a bear and a bull phase because the performance depended on the cointegrated relationship between the assets in a pair and not of the movements of the market.

> *Cointegration enables us to combine the two stocks in a certain linear combination so that the combined portfolio is a stationary process. If two stocks share a long-run equilibrium relationship, then deviations from this equilibrium are only short-term and are expected to die out in future periods. -*
>
> *(Caldeira & Moura, 2013, p. 51)*

The market chosen in this dissertation was the recent unexplored financial cryptocurrency[1] market, the reason being, mitigates a lot of the conditions used in pairs trading strategies as industry selection, sector, and country.

The idea for the first cryptocurrency, the Bitcoin, appeared during the world financial crisis in 2008 and cryptocurrencies, originally were supposed to be an alternative for "normal" currencies. All other cryptocurrencies, called Altcoins[2], are very dependent on Bitcoin price. The main idea of cryptocurrencies was to be independent of governments and central banks, because these two institutions, according to the creator of Bitcoin, were the main reason for the world financial crisis in 2008. This solution might have advantages such as reduced transactions fees, it is anonymous, fast, and easy to set up. On the other hand, this anonymity might lead to

[1] Hereinafter, we rely on the following definition according to Merriam-Webster (2018): Cryptocurrency is any form of currency that only exists digitally, that usually has no central issuing or regulating authority but instead uses a decentralized system to record transactions and manage the issuance of new units, and that relies on cryptography to prevent counterfeiting and fraudulent transactions.

[2] Altcoins: all other cryptocurrencies that are not Bitcoin.

the growth of the black market. Over time this idea evolved and now each cryptocurrency can have different purposes. For example, besides being an alternate source for money, some cryptocurrencies thrive on being typical investing commodities, or tokens, used inside virtual platforms.

Prices are based on supply and demand, are not rooted in any material goods, so the rate at which cryptocurrencies can be exchanged for another currency can fluctuate widely. Some economists consider that the cost of producing cryptocurrency, which takes a large amount of electricity, is directly related to its market price. Once someone owns cryptocurrency, they behave like physical gold coins: they have value and can be traded just as typical coins. People can use cryptocurrencies to purchase goods and services online or store them and hope that their value increases over time. Cryptocurrencies are transferred from one personal digital wallet to another. A wallet is a personal database that is stored on a computer drive, on a smartphone, tablet, or somewhere in cloud computing. Wallets are connected with the user by digital code, not by people's name.

The monetary supply is set to be fixed or is produced from solving complex mathematical tasks by miners - the huge network of people who contribute their personal computers to the network. They verify cryptocurrency transactions and keep ledgers safe. Miners are paid for their work by earning new cryptocurrency coins. The number of new cryptocurrencies which they get depends on the complexity of the task. Cryptocurrencies hold a simple data ledger file called the blockchain. Each blockchain is unique to each individual user and his personal wallet. All transactions are logged and available in a public ledger, helping ensure their authenticity and preventing fraud. There are some minor costs connected with using cryptocurrencies. Owners of some server nodes will charge one-time transaction fees every time a person sends money across their nodes and online exchanges will similarly charge when a person cash out in fiat currency.

The main objective of this thesis is to investigate the risk and return of a proposed pairs trading strategy for the cryptocurrency market. The data used contains the daily closing prices of the 10 cryptocurrencies with the highest market capitalization between January 1^{st} 2020 to July 1^{st} 2021, summing up to 547 daily observations. First, the specified select cryptocurrencies will be unit root tested by the Augmented Dickey-Fuller (ADF) test and then paired and tested for cointegration using the ADF, Johansen's (JOE), and Phillips Perron's (PP) test. Second, the selected pairs will form a portfolio, traded and performance of each portfolio will be analyzed, making the appropriate comparisons with a benchmark portfolio, and presenting our results.

Our results seem to indicate that cointegrated relationships between cryptocurrencies are more likely to hold over a longer period time and Engle-Granger´s approach employing the ADF test was the best predictor of cointegration relationships.

The remainder of this dissertation is structured as follows. In chapter 2, existing literature on pairs trading, cointegration, and cryptocurrency markets is reviewed. Chapter 3 describes the three mains chosen cointegration approaches. Chapter 4 explains the methodology used in this research and is followed by a discussion of the results in chapter 5. Finally, chapter 6 concludes this dissertation and expands on shortcomings and possible future research.

2 Literature Review

2.1 Pairs Trading Strategy

Pairs trading is a quantitative trading strategy that exploits financial markets that are out of equilibrium. The strategy is widely used by hedge funds and investment banks (Khandani & Lo, 2007). The idea behind pairs trading is as follows. First, a pair of assets is selected, that is known to historically move together and have some sort of long-run relationship. Using the assumption that the spread, defined as the difference in price between the paired assets, is mean-reverting, deviations from the mean can be exploited. In case of a deviation from the mean of the spread, an investor should take a short position in the overvalued asset and a long position in the undervalued asset. As soon as the spread converges back to its mean, the investor should unwind both positions, resulting in a profit. Even though this may sound like an intuitive approach, sophisticated econometric techniques can be used in all the steps involved in pairs trading.

Pairs trading is a strategy developed by a team of quants of the Morgan Stanley group somewhere in the mid-1980s. It remains an important statistical arbitrage technique used by hedge funds. They found that certain securities were correlated in their day-to-day price movements (Vidyamurthy, 2004). One of the first main articles introducing pairs trading was conducted by Gatev et al. (1999). They introduce the distance method, in which they match stocks into pairs employing the minimum distance between normalized historical prices. They show that a simple trading rule, which is executing a trade when the spread deviates two standard deviations from its mean, yield relatively large annualized excess returns of up to 11% that typically exceed conservative transaction-cost estimates. As stated by Krauss (2015), the distance method is one of the most investigated pairs trading frameworks, since it is relatively easy to implement the framework in practice due to its simplicity, transparency, and nonparametric character. These advantages and their use in practice, make the distance method a good reference point in this research to see how the machine learning techniques perform.

One of the first parameterized methods, that is also used in practice, is the cointegration method as thoroughly explained by Vidyamurthy (2004). Vidyamurthy (2004) selects the paired assets based on the cointegration relationship between two financial instruments. The idea behind this strategy is that two cointegrated assets will follow the same long-term trend and should return to their mean in case of deviations. Using cointegration has the advantage that the choice of a certain pair can be explained statistically while also confirming the desired mean-reversion of the pair. The main methods used for cointegration testing are the Engle & Granger (1987) or the Johansen (1991) test. Even though Vidyamurthy (2004) does not provide empirical results of the cointegration method, it is a framework that can form as a base for subsequent research. Caldeira & Moura (2013) use cointegration to select pairs on a Brazilian stock index. Using the trading rule proposed by Gatev et al. (1999), they find excess returns of more than 16% per year for the identified pairs, which shows that the cointegration approach can result in large profits.

2.2 Cointegration: Fundamental Concepts

In the previous decade, the concept of cointegration was increasingly applied in financial econometrics, in connection to time series analysis and macroeconomics (Alexander & Dimitriu, 2002). It is an extremely powerful technique, which allows dynamic modeling of non-stationary time series. The fundamental observation that justifies the application of the concept of cointegration in the analysis of stock prices is that a system involving non-stationary stock prices in levels can have a common stochastic trend (Caldeira & Moura, 2013).

2.2.1 Stochastic processes and a random walk

A sequence of random variables $\{Y_t : t = 0, \pm 1, \pm 2, \dots\}$ is called a stochastic process and serves as a model for an observed time series. An important stochastic process for modeling financial assets is the random walk. The observed process for a random walk,

$\{Y_t: t = 0, \pm1, \pm2, \ldots\}$ is as follows

$$Y_t = 0 \quad (1)$$

$$Y_t = e_1 \quad (2)$$

$$Y_t = e_1 + e_2 \quad (3)$$

$$Y_t = Y_{t-1} + e_2 \quad (4)$$

And the first difference of a random walk becomes

$$\nabla Y_t = e_t \quad (5)$$

where e_t is a stationary process (Asterious & Hall, 2011). A simulation of a random walk with 100 observations is found in Figure 1 below.

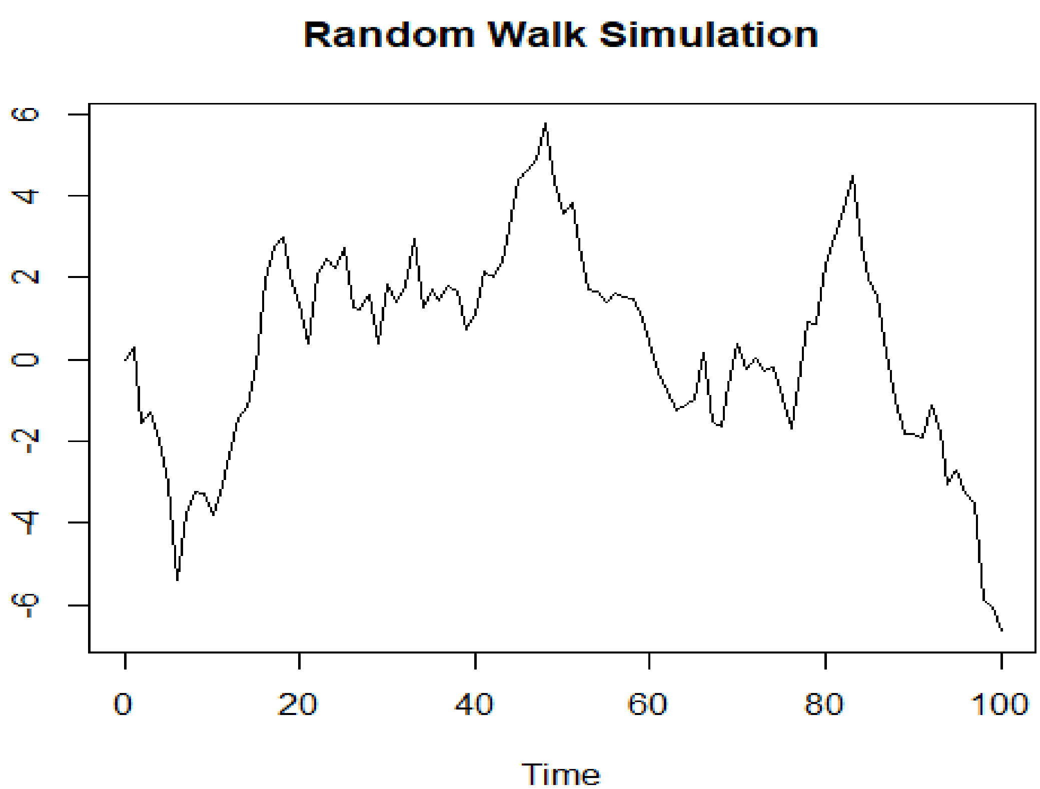


Figure 1: Random Walk Simulation

2.2.2 Stationarity

The fundamental idea behind stationarity is that the probability laws which govern the behavior of a stochastic process do not change over time. The statistical properties from observations of a stationary process are the same regardless of time in the process. There are two different kinds of stationarity - strict stationarity, and weak stationarity also referred to as covariance stationarity. A process $\{Y_t\}$ is said to be strictly stationary if the joint distribution of $Y_{t1}, Y_{t2}, \dots , Y_{tn}$ is the same as $Y_{t1-k}, Y_{t2-k}, \dots , Y_{tn-k}$ for all t time periods and all k lags (Cryer & Kung-Sik, 2009).

Strong stationarity is difficult to assess and the weak stationary process will be considered for the scope of this paper. A weak stationary process does not consider the joint distribution of the random variables.

For a weakly stationary process, it follows that $E(Y_t) = E(Y_{t-k})$ for every t and k. Hence, the mean function is constant over time (Cryer & Kung-Sik, 2009).

In addition,

$$Var\ (Y_t) = Var\ (Y_{t-k}), \tag{6}$$

for every t and k which makes the variance constant over time. The covariance is independent of time and only a function of the lag length (Cryer & Kung-Sik, 2009).

Therefore, the moments of a weakly stationary process are as follows:

Moment	Criteria	Formally
1st Mean	Mean is constant over time and independent of time	$\mu_t = \mu_{t-k}$
2st Mean	Variance is constant over time and independent of time	$\mu_{t,t} = \gamma_{0,0}$
3st Mean	Covariance is constant over time and independent of time	$\mu_{t,t-k} = \gamma_{0,k}$

Table 1: Moments of a Stationary Process

Shocks to stationary time series are temporary over time and the effects of the shocks will therefore dissipate. The time series will eventually revert to its long-time mean (Asterious & Hall, 2011).

2.2.3 White noise

The white noise process is a simple case of a probabilistic time series and the simplest case of a stationary time series. A white noise process is constructed by drawing an observation with a value from a normal distribution where the parameters are fixed and do not change over time at each time instance (Cryer & Kung-Sik, 2009). A white noise process is denoted as

$$Y_t = e_t \tag{7}$$

A simulation of a white noise process with 100 observations and a mean of 0 is found in Figure 2 below.

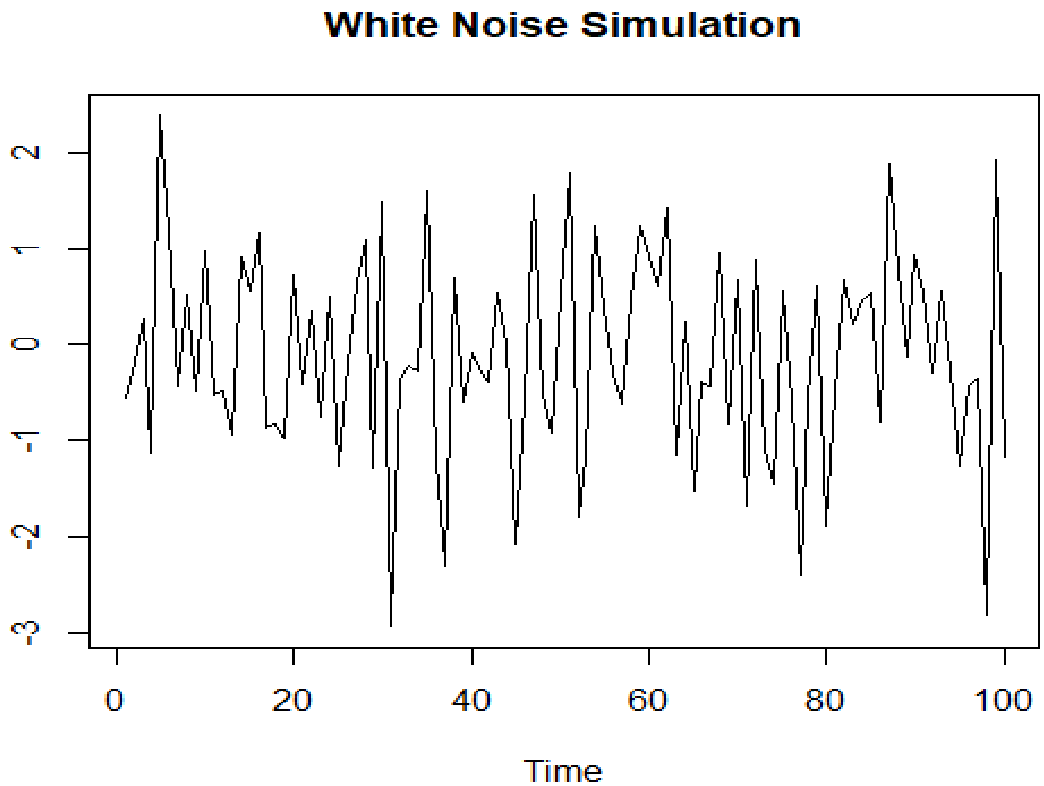


Figure 2: White Noise Simulation

2.2.4 The difference in time series and order of integration

The difference of a time series is the series of changes from one period to the next. For example, if the value of a time series is Y_t then the first difference of Y at period t is given by

$$\nabla Y_t = Y_t - Y_{t-1} \tag{8}$$

The order of integration is commonly denoted by

$$Y_t \sim I(d), \tag{9}$$

where d represents the least amount of differences in order to achieve a covariance stationary time series. An $I(0)$ is a covariance stationary process and the most common difference in order to achieve stationarity is the first difference that is that the time series is integrated of order 1; $I(1)$ (Neusser, 2016).

2.2.5 Unit root

An autoregressive process is a process that regresses on itself. The assumption of an $AR(1)$ model is that the time series of Y_t is mostly determined by the value in the prior period. Therefore, what occurs in time t is highly dependent on what happened in $t - 1$ and what will occur in time $t + 1$ will in turn be largely dependent on the series in the present time t (Asterious & Hall, 2011).

Consider the following $AR(1)$ model

$$Y_t = \phi Y_{t-1} + e_t, \quad (10)$$

where the residuals are white noise, there are in general three cases:

Case 1: If $|\phi| < 1$ then the series is stationary.

Case 2: If $|\phi| = 1$ then the series is non-stationary, that is has a unit root.

Case 3: If $|\phi| > 1$ then the series will explode.

Simulations of case 1, 2 and 3 for 1000 observations is found below in Figure 3.

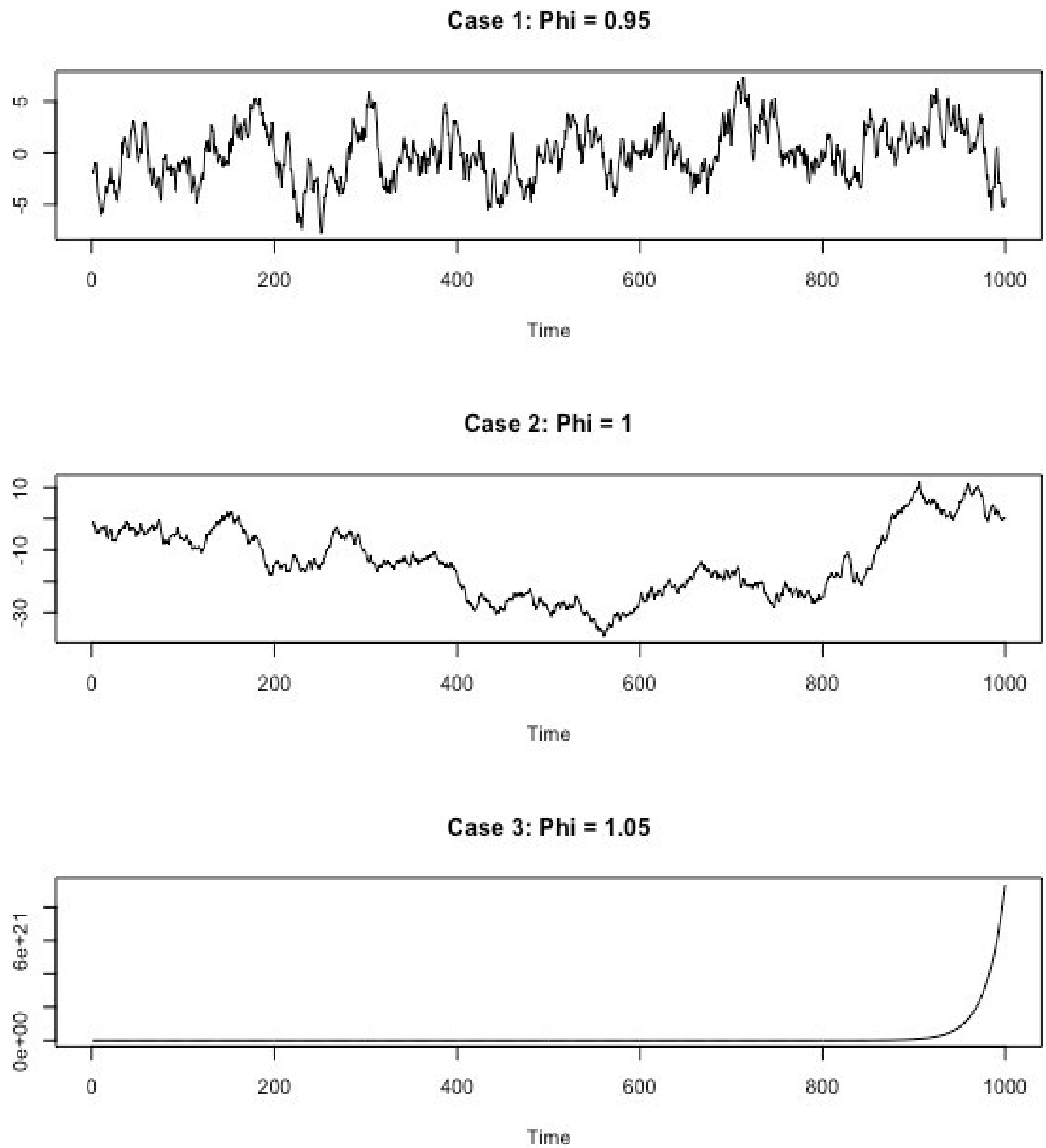


Figure 3: Different Phis of an AR(1) Process

To test the order of integration is to test the number of unit roots. The number of unit roots is therefore the difference required to obtain a stationary process for example the first or second difference (Asterious & Hall, 2011).

2.2.6 Augmented Dickey-Fuller test

The ADF test is a unit root test where lagged terms are added to the Y variable to remove possible autocorrelation. The number of lags is determined by the Akaike information criterion (AIC) or the Schwartz Bayesian criterion (SBC). The test has the following form

$$\Delta Y_t = a_0 + a_1 Y_{t-1} + a_2 t + \sum_{i=1}^{n} \beta_i \nabla Y_{t-1} + e_i, \tag{11}$$

where $a/$ is the intercept, $\sum_{i=1}^{n} \beta_i \nabla Y_{t-1}$ is the sum of the differentiated lagged Y's together with their coefficients. The null of the test is $a_0 = 0$ and the alternative hypothesis $a_0 < 1$. Rejecting the null will indicate that Y_t does not exhibit a unit root and therefore is stationary. This is obtained by comparing the ADF test statistic with a critical value at a given significance level (Asterious & Hall 2016). The test statistic of the ADF test is given by

$$ADF_{obs} = \frac{\widehat{\alpha_1}}{\hat{\sigma}_{\widehat{\alpha_1}}}. \tag{12}$$

2.2.7 Cointegration

Even though a group of variables is individually non-stationary, a linear combination of the series can form a stationary time series under the condition that they are individually integrated of the same order (Vidyamurthy 2004). That means that a linear combination of X_t and Y_t can form an $I(0)$ and a stationary process.

A linear combination of X_t and Y_t is obtained by regressing one of the time series on the other

$$Y_t = \hat{\beta}_1 + \hat{\beta}_1 X_t + \hat{e}_t \tag{13}$$

By taking the residuals we get

$$\hat{e}_t = Y_t + -\hat{\beta}_1 + \hat{\beta}_2 X_t \tag{14}$$

If $\hat{e}_t \sim I\ (0)$ and stationary, then X_t and Y_t are cointegrated (Asterious & Hall, 2011). $\hat{e}_t$ in the context of pairs trading will be the spread between assets in a pair.

2.2.8 The log-normal process

The most used model for modeling financial assets is the log-normal process, where the logarithm of the price of an asset is assumed to exhibit a random walk process. This implies that the price of the asset in the next period is approximately the price at the current period and is in probability theory referred to as a martingale. This means that the conditional expectation of a value in the next time point, given all prior values, is equal to the present value. As mentioned in section 2.2.1, taking the first difference of a random walk yields a stationary process, which can also be interpreted as the return of the asset or the increment of a random walk at a time point (Vidyamurthy 2004).

Likewise, the set of increments from a random walk obtained by taking the first difference is by definition drawings from a normal distribution. However, because of the martingale property of the random walk, the predicted increment of a random walk is zero, which is not handy when predicting asset prices with the goal of making money. The predicted value two steps further in time is still zero, but with increased variance. Nevertheless, because of the mean-reverting property of stationary time series, the researcher can predict the increment to the next value in a stationary process. Still, financial assets are modeled as random walks, which are not stationary, and the predicted value is equal to the value at present. However, due to cointegration, the researcher can find linear combinations of assets whose time series are combined stationary and therefore are predictable (Vidyamurthy 2004).

2.3 Cryptocurrency Market

A cryptocurrency is a digital asset whose primary purpose is to work as medium of exchange. Cryptocurrencies use cryptography and blockchain technology to ensure that all transactions are secured and everything new that appears in the blockchain is controlled by its digital infrastructure. The cryptocurrency market characterizes itself as a volatile space, especially when compared to traditional markets. While stocks or bond markets deal on average with no more than single percent digit price movements per day, cryptocurrencies can experience double- or triple-digit price movements, upwards and downwards. The volatile and current non-regulatory nature calls into question whether cryptocurrencies are a viable addition to one's asset portfolio. The advantage of this new developing asset class brings is potential portfolio diversification and some researchers question if cryptocurrencies can be a good hedging tool. Researchers are studying if people can store money in cryptocurrencies like they do with gold. This characteristic makes cryptocurrencies potentially less susceptible to shocks in the traditional market, making them a hedge option. Furthermore, cryptocurrencies are traded on digital exchanges that do not operate with a closing or opening time. Cryptocurrency exchanges are open 7 days a week, 365 days a year. Being able to trade at any time of the day or night brings a bigger window of opportunity to make trades.

2.3.1 Is Cryptocurrency Money?

Money is typically defined by economists as having three attributes: it functions as a medium of exchange, a unit of account, and a store of value. Even today, many researchers are trying to answer the question of whether if cryptocurrencies can satisfy those three points.

Gervais et al. (2014) analyzed if Bitcoin is a decentralized currency and, i.e., there is no They concluded that Bitcoin is not a "normal" currency, but it seems to be decentralized differently as it has no central entity or the administrator for the currency like a government or bank.

Yermack (2015) also tried to assess if Bitcoin is a real currency. The author concluded that Bitcoin fails to satisfy the criteria of fiat currencies and that Bitcoin appears to behave more like a speculative investment than a currency.

Bjerg (2016) tries to explain how Bitcoin is money. As we know that fiat money must be used as a store of value, a medium of exchange, and a unit of account, the author conclude that Bitcoin is commodity money without gold, fiat money without a state, and credit money without debt. According to the author, Bitcoin is something between money and a commodity, although closer to this last one.

2.3.2 Review of Developed Research

Cryptocurrencies besides being alternative money might also be a good investing commodity due to their liquidity. On the other hand, they are very volatile, so it is as easy to win as to lose money. In this section, the authors are checking if cryptocurrencies are worth investing in or not.

Feng et al. (2017) analyze informed trading in the Bitcoin market, from 2011 until 2017, with data taken from bitcoincharts.com. These authors construct their own order-size-based measure to detect informed trading and conclude that informed trading in the Bitcoin market suggests that people who get information before it is widely available, profit on their private information, at the cost of other market participants losses. Hence, the lack of clear regulatory laws and regulatory authorities are potential reasons for the existence of informed trading.

Brauneis and Mestel (2018) analyze cryptocurrencies' prices, using data from 2015 until 2017, provided by coinmarketcap.com. The authors concluded that cryptocurrencies become less predictable as liquidity increases and that the bid-ask spread shows the expected negative effect towards efficiency.

Corbet et al. (2018) explore dynamic relationships between cryptocurrencies and other financial assets, using data from CryproCompare.com for cryptocurrencies and Bloomberg for financial assets. They use the generalized variance decomposition methodology by Diebold and Yilmaz (2012) and conclude that cryptocurrencies may offer diversification benefits for investors with short investment horizons and that time variation in the linkages reflects external economic and financial shocks.

Gkillas and Katsiampa (2018) apply extreme value theory, a theory that tries to uncover the characteristics of the distribution tails of asset returns, to assess which cryptocurrency is the most and least risky. They used the five largest cryptocurrencies, each one from the earliest date available until 2017, taken from www.coindesk.com for Bitcoin and coinmarketcap.com for the remaining cryptocurrencies. They use the peaks-over-threshold, which is a method to extract extremes and find that Bitcoin Cash is the riskiest cryptocurrency, while Bitcoin and Litecoin are the least risky cryptocurrencies in terms of investing.

Ciaian et al. (2018) investigate the virtual relationships between the Bitcoin and the Altcoin markets, which are markets for other cryptocurrencies which prices are correlated with the Bitcoin price. They use data on virtual currency supply and demand data from 2013 until 2016 for Bitcoin, 6 major Altcoins, and 10 minor Altcoins, which they extract from quandl.com and coinmarketcap.com. Additionally, they use for commodities like oil and gold, including two exchange rates for the USD/EUR and the CNY/USD. The authors use the Autoregressive Distributive Lag (ADL) model and conclude that the Bitcoin and the Altcoin markets are interdependent, the Bitcoin-Altcoin price relationship is significantly stronger in the short-run than in the long-run, and finally, in the long-run, macro-financial indicators determine the Altcoin price formation to a slightly greater degree than Bitcoin does.

3 Cointegration Approaches

The applicability of the cointegration technique to asset allocation was pioneered by Lucas (1997) and Alexander (1999). Its key characteristics, i.e. mean-reverting tracking error, enhanced weights stability, and better use of the information comprised in the stock prices, allow a flexible design of various funded and self-financing trading strategies, from the index and enhanced index tracking to long-short market neutral and alpha transfer techniques (Caldeira & Moura, 2013).

3.1 Engle-Granger´s approach

As mentioned in the introduction of this paper, in 1987 Engle and Granger introduced a way to test for cointegrated relationships between different time series. To understand the approach, consider two given time series X_t and Y_t, where X_t is $I(0)$ and Y_t is $I(1)$. Thereby, any linear combination of the series

$$\theta_1 X_t + \theta_2 X_t \tag{15}$$

will always be $I(1)$, that is non-stationary. This is because the behavior of the non-stationary $I(1)$ series will dominate the behavior of the stationary series.

However, if X_t and Y are both $I(1)$, then a linear combination of the series in equation 15 is likely to be non-stationary $I(1)$ too. Although this is usually the case, there are cases where a linear combination of two non-stationary time series can result in a stationary process and the time series is then said to be cointegrated.

Estimating the parameters of the long-term relationship and investigating if the time series are cointegrated or not is difficult. Therefore, Engle and Granger introduced a method for estimating parameters of the relationship and checking for cointegration. The method is as follows:

First, test whether the time series are integrated of the same order. This is in this thesis is tested through the ADF test, in order to deduce the number of unit roots. The time series must be integrated of the same order and cannot be stationary.

Second, if the variables are integrated of the same order, the long-run relationship is estimated by regressing one variable on the other

$$Y_t \;=\; a_0 \;+\; \beta_1 X_t + e_t \tag{16}$$

which can be written as

$$e_t \;=\; Y_t - a_0 \;-\; \beta_1 X_t \tag{17}$$

If e_t is stationary $(I(0))$, then the variables are cointegrated. This is tested through either the ADF or the PP test, only this time on the residual time series. If the test is rejected it can be concluded that the variables have a cointegrated relationship (Asterious & Hall 2016). The procedure for JOE test will be outlined below in section 3.3.

3.2 Phillips-Perron test

As outlined in section 2.2.6 the ADF test is based on the assumption that the error terms have constant variance and are statistically independent. The PP test, however, which was developed as a generalization of the ADF test, has a milder assumption regarding the error terms.

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

$$Y_t \;=\; a_0 \;+\; a_1 Y_{t-1} + e_t, \tag{18}$$

where the null is that $a_1 = 1$ and the alternative that $a_1 < 1$. Rejecting the null will indicate that Y_t does not have a unit root and is therefore stationary.

Whereas the ADF test adds lagged differentiated terms to handle higher-order correlations, the PP test modifies the coefficient a_1 from the $AR(1)$ regression for the serial correlation in e_t. The derivation of the PP-test is beyond the scope of this thesis.

3.3 Johansen´s approach and test

Vector autoregression (VAR) is essential in order to understand Johansen's test. A vector autoregression is a matrix that contains two or more regressions, where each variable is regressed on n number of lags of the other variables and n number of lags of the variable itself. Each variable is also regressed on a constant. A VAR system can take the following form

$$Y_t \ = a \ + \ \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \ldots + \beta_n Y_{y-n} \ + \ e_t, \tag{19}$$

where Y_t is a vector, β_k act as an j by j matrix of the coefficients, where $k = 1,2,3, \ldots, n$, a represent a j by one matrix of the constants and e_t represent the error terms in the same matrix as a.

If a model contains three or more variables, there is a possibility that more than one cointegrated relationship exists. As a rule of thumb, for n number of variables there can at most be $(n - 1)$ cointegration. Johansen's approach can detect multiple cointegrated relationships due to the use of a VAR system. Compared to Engle and Granger's approach, which can just detect one cointegrated relationship.

Using the framework by Asterious and Hall (2016), the derivation of Johansen's approach to detecting cointegration for a vector of two-time series $X_t = [Y_t, Z_t]$, is as follows

$$\begin{cases} Y_t \\ Z_t \end{cases} = \begin{matrix} \pi_{11} Y_{t-1} + \pi_{12} Z_{t-1} + e_{1t} \\ \pi_{21} Y_{t-1} + \pi_{22} Z_{t-1} + e_{2t} \end{matrix}. \tag{20}$$

Now Y_t and Z_t are cointegrated, if

$$\begin{cases} \Delta Y_t \\ \Delta Z_t \end{cases} = \begin{matrix} \alpha(\beta_1 Y_{t-1} + \beta_2 Z_{t-1}) + e_{1t} \\ \alpha(\beta_1 Y_{t-1} + \beta_2 Z_{t-1}) + e_{2t} \end{matrix}, \tag{21}$$

where $\beta_1 Y_{t-1} + \beta_2 Z_{t-1}$ is a stationary process.

This can also be represented using matrices

$$\begin{pmatrix} \Delta Y_t \\ \Delta Z_t \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \cdot \begin{pmatrix} Y_{t-1} \\ Z_{t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}. \tag{22}$$

Then Y_t and Z_t are cointegrated, if

$$\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} = \begin{pmatrix} \alpha_1\beta_1 & \alpha_1\beta_2 \\ \alpha_2\beta_1 & \alpha_2\beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \cdot (\beta_1\ \beta_2), \tag{23}$$

where $\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix}$

Thereby Y_t and Z_t are cointegrated if the rank of Π is one. The rank of the matrix Π represents the maximum number of linearly independent rows of Π.

The rank of Π is estimated by two different likelihood ratio tests, both based on eigenvalues, that is the number of characteristic roots. The first method tests the null that $Rank(\Pi) = r$ against the alternative that $Rank(\Pi) = r + 1$. In other words, the null is that there are r cointegrated vectors and at most r cointegrated relationships. Meanwhile, the alternative suggests that there are $r + 1$ vectors. The method orders the eigenvalues in descending orders and tests if they are significantly different from zero. For example, consider n characteristic roots, $-\lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_n$. If there is no cointegration, then all roots will be equal to zero. Hence, $-T \ln(1 - \hat{\lambda}_{r+1})$ will also be zero. Nonetheless, if the rank is equal to one implying one cointegrated relationship, then $\lambda_1 > 0$ which leads to $-T \ln(1 - \hat{\lambda}_{r+1}) < 0$.

There are two methods to get the statistics used to test if the characteristic roots are different from zero. The first is as follow

$$\lambda_{max}(r, r + 1) = -T \ln(1 - \hat{\lambda}_{r+1}) \tag{24}$$

The second method is conducted by the likelihood ratio test for the trace of Π. The null, in this case, is that the number of cointegrated vectors is at most r (Asterious & Hall 2016). Where the test statistic is

$$\lambda_{trace} = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_{r+1}) \quad (25)$$

In this thesis, the first method will be used when following Johansen´s approach.

3.4 Drawdowns of the Engle-Granger´s approach, the ADF & the PP Test

Following Engle-Granger´s approach, one must regress one-time series on the other. As an example, consider the time series X_t and Y_t, the approach does not explain which time series to regress on the other and why. One can either regress X_t on Y_t or vice versa. This forces the researcher to choose between two different regressions often with different residuals. In asymptotic theory, when the sample sizes go to infinity, the residuals of the regressions are equivalent. However, the sample size for economic data is rarely large enough to result in equal series when time series is regressed upon each other (Asterious & Hall, 2016).

One drawdown of the ADF and the PP test is that when a process is stationary, yet close to having a unit root, the power is low (Brooks, 2002). Another drawdown of the ADF and PP test is that they over-reject the null when the moving average root of the process is negative (Schwert, 1989).

Likewise, as mentioned in sections 3.1 and 3.2, neither the ADF nor the PP test can test for more than one cointegrated relationship.

3.5 Drawdowns of Johansen´s approach and test

One of the assumptions of Johansen´s test is that the cointegrated vector is constant during the test period which is a strong assumption since long-run relationships of the underlying variables can vary, particularly if the test period is long. In addition, using the VAR method is of theoretical nature which can make the model hard to interpret (Brooks, 2002).

4 Methodology/Implementation of the Strategy

4.1 Data

The data used in this thesis has been imported from Yahoo Finance and consists of data from the 10 cryptocurrencies with the highest market capitalization on January 1st 2020. The data pans from January 1st 2020 to July 1st 2021. Since cryptocurrencies are traded 24 hours every day there are no closing days to consider. The prices in the study will therefore be the 24-hour price change and all prices in the study are denoted in United States dollars ($).

Choice of cryptocurrencies is the top 10th cryptocurrencies with the highest market capitalization between January 1st 2020 and July 1st 2021, excluding stable coins and cryptocurrencies that were not created after January 1st 2020. The selected cryptocurrencies and respective market capitalization between the specified data range can be found in Table 3 below.

	Name	**Symbol**	**Market Cap (01/01/2020)**	**Market Cap (01/07/2021)**
1	Bitcoin	BTC	$130,580M	$629,339M
2	Etherium	ETH	$14,271M	$246,278M
3	XRP	XRP	$8,349M	$30,511M
4	Bitcoin Cash	BCH	$3,719M	$9,355M
5	Litecoin	LTC	$2,679M	$9,183M
6	Cardano	ADA	$867M	$42,667M
7	Binance Coin	BNB	$2,129M	$44,222M
8	Polygon	MATIC	$38M	$6,791M
9	Dogecoin	DOGE	$249M	$31,852M
10	Chainlink	LINK	$632M	$7,958M

Table 2: Cryptocurrencies by Market Capitalization

4.2 Selection of Pairs

Pairs are formed through testing 10 cryptocurrencies whether cointegration exists and if each individual time series exhibit a unit root. There can at most be 45 cointegrated pairs when performing a cointegration test on 10 cryptocurrencies in a window. The maximum number of pairs is given by

$$\frac{n(n-1)}{2} \tag{26}$$

where n is the number of cryptocurrencies.

4.2.1 Identifying Pairs - ADF, PP, and JOE approach

As mentioned in section 3.1, it is redundant to test already stationary time series for cointegration. Therefore, all cryptocurrencies are first tested on whether their individual time series have a unit root by the ADF test. All currencies which have a unit root will then be tested for cointegration with other cryptocurrencies which have a unit root using the ADF, PP, and JOE test and all decisions will be made at the 5% significance level.

When testing for a unit root, each cryptocurrency will be tested on whether each individual time series is integrated of order 1. For the scope of this thesis, all cryptocurrencies will only be tested if they are integrated into order 1 and this is only tested through the ADF test although time series can potentially be integrated into other orders.

For the ADF and PP cointegration test, the second step is to regress the time series on each other and determine if the residual time series is stationary. A stationary residual time series between two cryptocurrencies would mean rejecting the null hypothesis that the time series exhibits a unit root.

The test regression for the ADF test is found in equation 11 and the test regression for PP test is found in equation 18.

For JOE cointegration test, a bivariate vector of two cryptocurrencies is set up. Thus, Y_t and Z_t are the time series of two cryptocurrencies that are individually integrated of order 1 (see equation 20).

Johansen's reduced rank regression is then used to estimate α and β and leads to the rank of Π which is derived in section 3.3 and becomes the number of cointegrated relationships which in this thesis must be one or zero. The method to test the null hypothesis will in this study be based on maximum eigenvalue and the test statistic can be found in equation 24.

4.3 Methodology

4.3.1 Testing and trading windows

The methodology for trading pairs will follow the following scheme:

1. Test if each individual cryptocurrency exhibits a unit root over a 3- or 6-month testing window through the ADF test. There will in total be 5 testing windows for the 3-month method and 2 for the 6-month method.

2. All cryptocurrencies which exhibit a unit root will be tested for cointegration through the ADF test, PP test, or JOE test over a 3- or 6-month testing window. The cointegrated cryptocurrencies will form pairs and a portfolio of pairs will be formed.

3. All cointegrated pairs will be traded over a trading window that spans over the last day of the testing window to 3 or 6 months further in time. There will in total be 5 trading windows for the 3-month method and 2 for the 6-month method.

The testing and trading windows for the 3-month and 6-month procedures are illustrated in Figure 4 below.

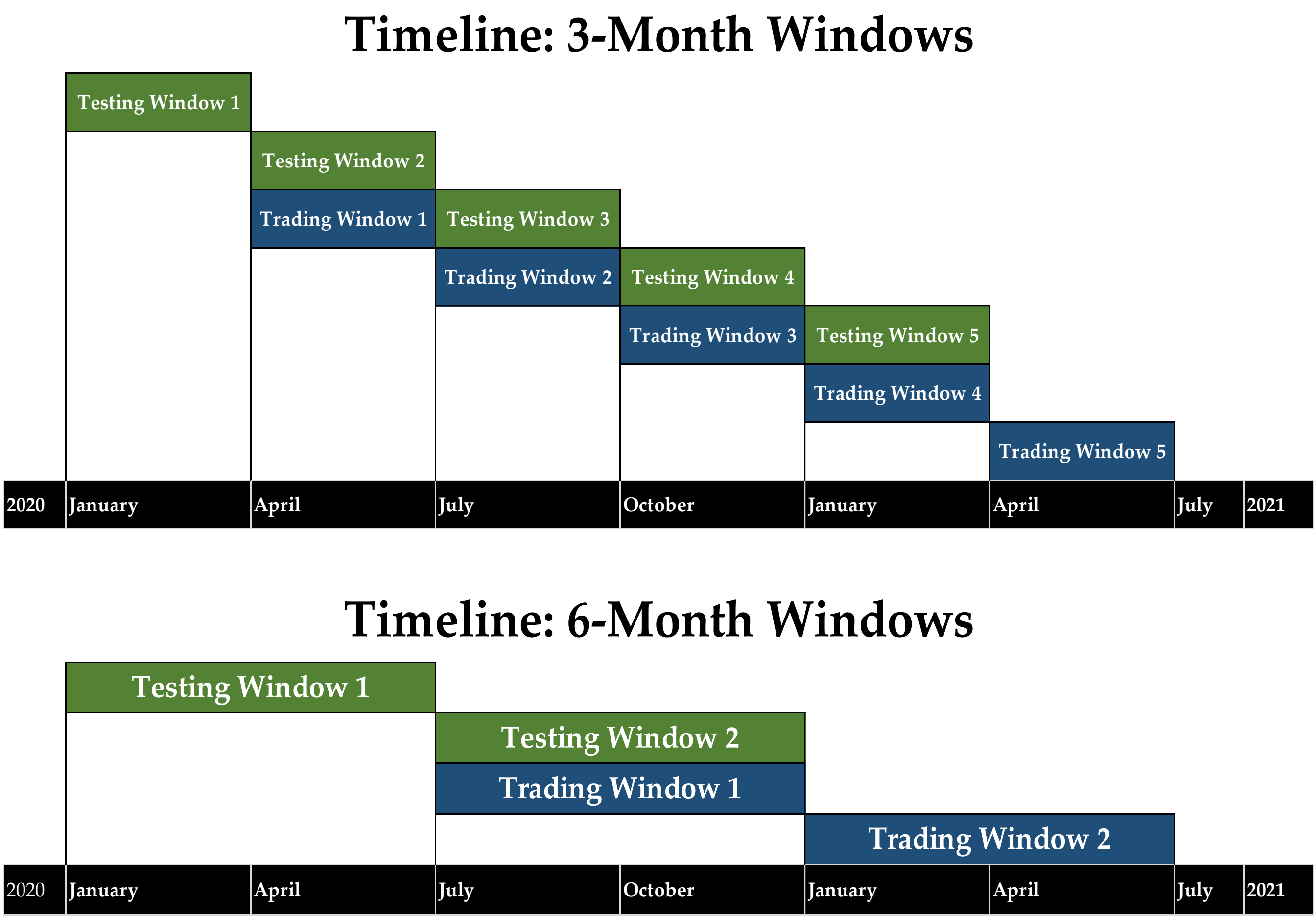


Figure 4: Testing and Trading Windows

4.3.2 Z-Score and Trading Strategy

Pairs trading strategies are based on capitalizing on the oscillations around the mean of the spread. It usually requires the trader to trade an equal amount in asset Y of price Y_t and asset X of the price βX_t by

$$Y_t \ = \beta X_t, \tag{27}$$

where β is the coefficient that makes the price of X and Y equal when a position is opened (Ting, 2017).

Sophisticated pairs trading strategies can involve different weightings instead of equal weightings but are considered out of scope for this paper.

Recall that $\hat{e}_t$ represents the spread between two assets at a given time and can be obtained by regressing one asset on the other. In order to easily generate trading signals, we compute the dimensionless z-score defined as

$$z_t = \frac{\hat{e}_t - \mu_{spread}}{\sigma_{spread}}, \quad (28)$$

Where z_t is the standard deviation away from 0 at time t, $\hat{e}_t$ is the value of the spread at a given time t, μ_{spread} is the mean of the spread and σ_{spread} is the standard deviation of the spread (Palomar, 2018). A mathematical representation of the spread and $\hat{e}_t$ is found in equation 14.

Positions are taken when z& drift from 0 and reach a certain value. These rules are referred to as thresholds in this paper and are pre-specified trading strategies on when to open and close positions. The spread is expected to be shorted when z& reaches a specific positive threshold and the trader is expected to long the spread if the spread reaches a certain negative threshold.

According to Ting (2017) thresholds should be set to maximize returns and minimize the number of transactions. Therefore, thresholds should be as far away from zero as possible while still capitalizing on the mean-reverting property of the spread in order to reduce transaction costs and increase the likelihood of high-quality trades.

The trading strategy chosen for all portfolios in this study is as follows:

Trading Strategy

Short the spread if $z_t > 1.75$	Buy βX shares and short-sell Y shares.
Long the spread if $z_t < -1.75$	Short-sell βX and buy Y shares.
Stop-loss if $z_t > 4$	Exit both positions
Stop-loss if $z_t < -4$	Exit both positions
Close positions if the spread reverts back to its mean $z_t = 0$	Exit both positions

Table 3: Trading Strategy
Note: z_t is the value of the normalized spread at a time *t*.

For portfolios where transaction costs are included, a 1% transaction cost will be subtracted from the cumulative return of every position when a position is opened and closed.

4.3.3 Buy and Hold index

The buy and hold index consists of the cumulative return of buying and holding each cryptocurrency in a trading window without a pairs trading strategy. The buy and hold return of every portfolio - Augmented Dickey-Fuller, Johansen's, or Phillips Perion portfolio is given by

$$Return\ Test_{Buy\ and\ Hold} = \frac{Return\ Cryptocurrency\ in\ window\ t}{Number\ of\ Cryptocurrencies\ in\ window\ t}, \tag{29}$$

where *window t* is the amount of trading windows. The return of each test is then combined by

$$Return\ Test_{Buy\ and\ Hold} = \frac{Return\ ADF_{Buy\ and\ hold} + Return\ JOE_{Buy\ and\ hold} + Return\ PP_{Buy\ and\ hold}}{3}, \tag{30}$$

The buy and hold index will be used to compare the return of a pairs trading strategy.

4.3.4 Sharpe Ratio

William F. Sharpe introduced the Sharpe Ratio as a way to measure mutual funds returns adjusted to risk exposure. The ratio aims to describe the difference in expected return in excess of the risk-free rate for one more unit of volatility (Sharpe, 1994). In general, investors prefer a portfolio with high a Sharpe ratio (SR) over a portfolio with a low Sharpe ratio, ceteris paribus. The Sharpe ratio is given by

$$S_p = \frac{E(r_p) - r_f}{\sigma_p}, \tag{31}$$

where $E(r_p)$ is the expected return of the portfolio, r_f is the risk-free rate and σ_p is the standard deviation of the portfolio.

5 Empirical Results

This section aims to evaluate the performance of the methodology of this paper. This is done by assessing the number of cointegrated cryptocurrencies whose individual time series exhibit a unit root and have a cointegrated relationship for each testing period by using the ADF, JOE, or PP test. These pairs that exhibit a cointegration relationship will form an individual portfolio which will be compared with other portfolios with the same testing and trading window procedure.

5.1 Cointegrated Pairs

Table 4 shows the number of cointegrated pairs for every testing method. Johansen´s approach detected the most pairs for the 3-month method and ADF's approach for the 6-month method. Furthermore, 21 pairs were identified through Johansen's approach following the 3-month procedure and 3 were formed with the 6-month procedure. The PP test detected only 5 pairs. 5 in the 3-month window and 6 for the 6-month window procedure. ADF test detected 12 pairs with the 3-month procedure and 10 pairs to the 6-month procedure. In total, 38 pairs were identified with the 3-month window method and 19 pairs were identified through the 6-month method.

The following Table 4 shows the number of cointegrated pairs for the 3-month and 6-month window.

Number of cointegrated Pairs; 3-month window			
	ADF	PP	JOE
Trading Window 1	1	0	7
Trading Window 2	3	1	4
Trading Window 3	2	0	3
Trading Window 4	5	2	1
Trading Window 5	1	2	6
Total:	12	5	21
Number of cointegrated Pairs; 6-month window			
Trading Window 1	9	3	2
Trading Window 2	1	3	1
Total:	10	6	3

Table 4: Number of Traded Cointegrated Pairs
Note: Number of pairs where each individual time series has a unit root and the residual series is stationary ($p < 0,05$).

5.2 Performance Evaluation

Table 5 shows the cumulative return, standard deviation, and Sharpe ratio of each portfolio without transaction costs and with transaction costs set at 1% for every position and each corresponding buy and hold benchmark.

The buy and hold benchmark had a return of 2508,03% for the 3-month window procedure and 3662,19% for the 6-month window procedure by the end of the trading period. All portfolios have in every window underperformed their buy and hold benchmark and all portfolios with the 6- month window procedure have outperformed their 3-month counterpart except for the portfolio formed through the JOE test where the 3-month portfolio slightly outperformed the 3-month portfolio by 116%.

All the portfolios formed presented a positive cumulative return at the end of the trading periods. Hence, Phillips Peron's portfolio following the 3-month methodology was the only portfolio with a noticeable lower return at the end of the trading period.

The highest cumulative return without transaction costs at a time point was found in all three statistical tests, on the 6-month window, trading window 1, with BTC/DOGE pair, accumulating a return of 98,29%.

The lowest cumulative return was from the ADF portfolio in the trading window 4 of the 3-month window method with the pair ETH/LTC netting a negative accumulative return of 18%.

It is to notice that the cumulative returns with transaction costs seem not to vary much the strategy overall returns.

The null returns, in which the testing window confirmed cointegration relationship, but in the trading window the orders did not find the requirements to open, were found only in the 3-month method in all three statistical tests. Concretely, this null result appears both in the ADF test and PP test with BNB/BCH pair (window 4), and in the JOE test with BTC/XRP (Window 1).

The following Table 5 shows the performance results of each trading window, for the ADF test, PP test, JOE test, and Buy & Hold (benchmark portfolio), with and without transaction costs.

Without transaction costs

	ADF		PP		JOE		B&H	
Windows	3-Month	6-Month	3-Month	6-Month	3-Month	6-Month	3-Month	6-Month
Return	259,45%	417,10%	42,20%	335,28%	393,62%	277,96%	2508,03%	3662,19%
StDev	0,025	0,023	0,021	0,030	0,019	0,036	3,207	9,388
SR	5,381	14,725	3,820	17,043	9,388	26,283	7,821	3,901

With transaction costs

	ADF		PP		JOE		B&H	
Windows	3-Month	6-Month	3-Month	6-Month	3-Month	6-Month	3-Month	6-Month
Return	256,43%	412,74%	41,64%	331,93%	389,61%	275,18%	2508,03%	3662,19%
StDev	0,025	0,023	0,021	0,030	0,019	0,035	3,207	9,388
SR	5,318	14,545	3,770	16,873	9,286	26,021	7,821	3,901

Table 5: Performance Results

Note: Performance statistics for pairs trading portfolios formed through the ADF test, Johansen's test, Phillips Peron's, and a buy and hold strategy; Risk-Free Risk = 0.

6 Conclusion

In this work the pairs trading strategy was applied to the cryptocurrency market. For this purpose, we pursued the cointegration methodology, where the objective was to implement and compare the Augmented Dickey-Fuller test, Johansen's test, and Phillips Peron's test. The overall performance of these tests was compared to a simple buy and hold benchmark portfolio. Additionally, we took into account the possible impact of the transaction costs.

Our results show that our pairs trading strategy did not outperform the Buy and Hold benchmark portfolio, meaning that, in our time window, holding the pairs resulted in considerable higher returns. This investment strategy is often described as market neutral, in the sense that mitigates the net exposure to the market, so the return provided should not be affected by market direction.

Although all portfolios underperformed their buy and hold benchmark, the 6-month trading and testing procedure yielded a higher return compared with the 3-month procedure. Therefore, we can conclude that cointegrated relationships between cryptocurrencies are more likely to hold over a longer period of time.

Of the three proposed approaches, Engle-Granger´s approach employing the ADF test was the best predictor of cointegration relationships with an excess mean return of 338%.

The transaction costs, set at 1%, did not have a meaningful impact on the return, given the low number of opened and closed positions, assuming a spread oscillating around 1.75 standard deviations between cryptocurrencies.

This research concludes that the pairs trading strategy based on cointegration relationships can successfully be applied to the cryptocurrency market and profits can be generated.

6.1 Limitations and Future Work

This work has limitations, the performance of a pairs trading portfolio is highly dependent on the trading rules of the pairs trading algorithm. For example, if the spread is mean reverting but the spread does not oscillate around 1.75 standard deviations, then the strategy will not capitalize on the mean-reverting property of the spread. This means that a portfolio formed through another statistical test could potentially perform better with other trading rules in the trading algorithm.

Also, all capital has been invested in every window regardless of the amount of traded pairs in one window. This means that all capital is invested in only one pair if one pair is detected in a testing window. Hence, to weigh invested capital differently is suggested for further studies. That could be, for example, to only allocate 1% to 10% of invested capital in one pair regardless of the number of pairs in a portfolio.

Another suggestion for further research is to hold positions until they should be closed according to the pairs trading strategy. In this thesis, every open position in this study is closed by the end of a trading window which is not ideal or realistic. Hence, evaluating the performance of portfolios where positions are closed when the spread reverts to the mean or reaches a stop-loss regardless of window and time is suggested. Also, analyzing why the different cointegration tests detect pairs differently is suggested for further research.

The temporal line of the data chosen could be misconstrued since it captures for the most part a bull phase, where the most gains occur in this period. Comparisons should be made with a bear phase.

Lastly, analyzing shorter time windows is suggested. That can for example be to follow the methodology of this paper with 1-month windows. However, shorter windows lead to higher transaction costs meaning that there will always be a trade-off between the length of windows and transaction costs.

References

ASTERIOUS, D. & Hall, G. (2016) *Applied Econometrics*. London: Palgrave MacMillan.BITCOIN. (2019). *What is Bitcoin*. Available at: https://bitcoin.org/en/faq#what-is-bitcoin [2019-04-20]

DICKEY, D. A.; FULLER, W. A. Distribution of the estimator for auto-regressive time series with a unit root. Journal of the American Statistical Association, v. 74, p. 427-31, 1979.

BODIE, K. & MARCUS, A. (2013) *Essentials of Investments.* New York: McGraw-HillEducation.

BROOKS, C. (2002). *Introductory Econometrics for Finance.* Cambridge: CambridgeUniversity Press.

CALDEIRA, J. & MOURA, G. "Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy." Available: SSRN 2196391 (2013).

COINMARKETCAP (2019) *Global Charts.* Available at: https://coinmarketcap.com/historical/20210701/ [2021-08-20].

CRYER, J. & KUNG-SIK, C. (2009) *Time Series Analysis with applications in R*. Berlin:Springer Verlag Gmbh.

DICKEY, D. A.; FULLER, W. A. Distribution of the estimator for auto-regressive time series with a unit root. Journal of the American Statistical Association, v. 74, p. 427-31, 1979.

ENGLE, R. F. & GRANGER, C. W. J. (1987) Co-integrationand Error Correction: Representation Estimation, and Testing, *Econometrica,* vol 55, No 2, pp 251-276

GATEV, E., GOETZMANN, N. & ROUWENHORST, G. (1999) Pairs Trading Performance of a relative value arbitrage rule. *Review of Financial Studies*. Vol 19, pp 797-827.

HARRIS, R. I. D. Cointegration analysis in econometric modelling. London: Prentice Hall, 1995. 176 p.

HILD, A. & OLSSON, M. (2019) Pairs Trading, Cryptocurrencies and Cointegration. *Department of Statistics: Uppsala University*

JOHANSEN, S. Likelihood-based inference in cointegrated vector auto-regressive models. New York: Oxford University Press, 1995. 267 p. (Advanced Texts in Econometrics).

KHANDANI, A. E., & LO, A. W. 2007. What happened to the quants in August 2007? Journal of Investment. Management, 5, 5–54.

MACLEOD, W. VAN VUREEN, G. (2015) Interpreting the Sharpe ratio when excessreturns are negative. *Investment Analysis Journal, vol 33*

MILUTINOVIC, M. (2018) *Cryptocurrency*. Ekonomika (Vol 64)

ROSIC, A. (2016) *What is cryptocurrency*. Available at: https://blockgeeks.com/guides/what-is-cryptocurrency/ [2021-08-20].

SCHWERT, W. (1989) Why does stock market volatility change over time?. *The Journal offinance,* vol 44.

SHARPE, W. (1994) The Sharpe Ratio. *Journal of Portfolio Management*, vol. 21, no. 1,pp. 49-58.

SYCZEWSKA, E. (2011) COINTEGRATION SINCE GRANGER: EVOLUTION A DEVELOPMENT. *Quantitative Methods In Economics, Vol. XII, No 1, 2011,* pp.147-164

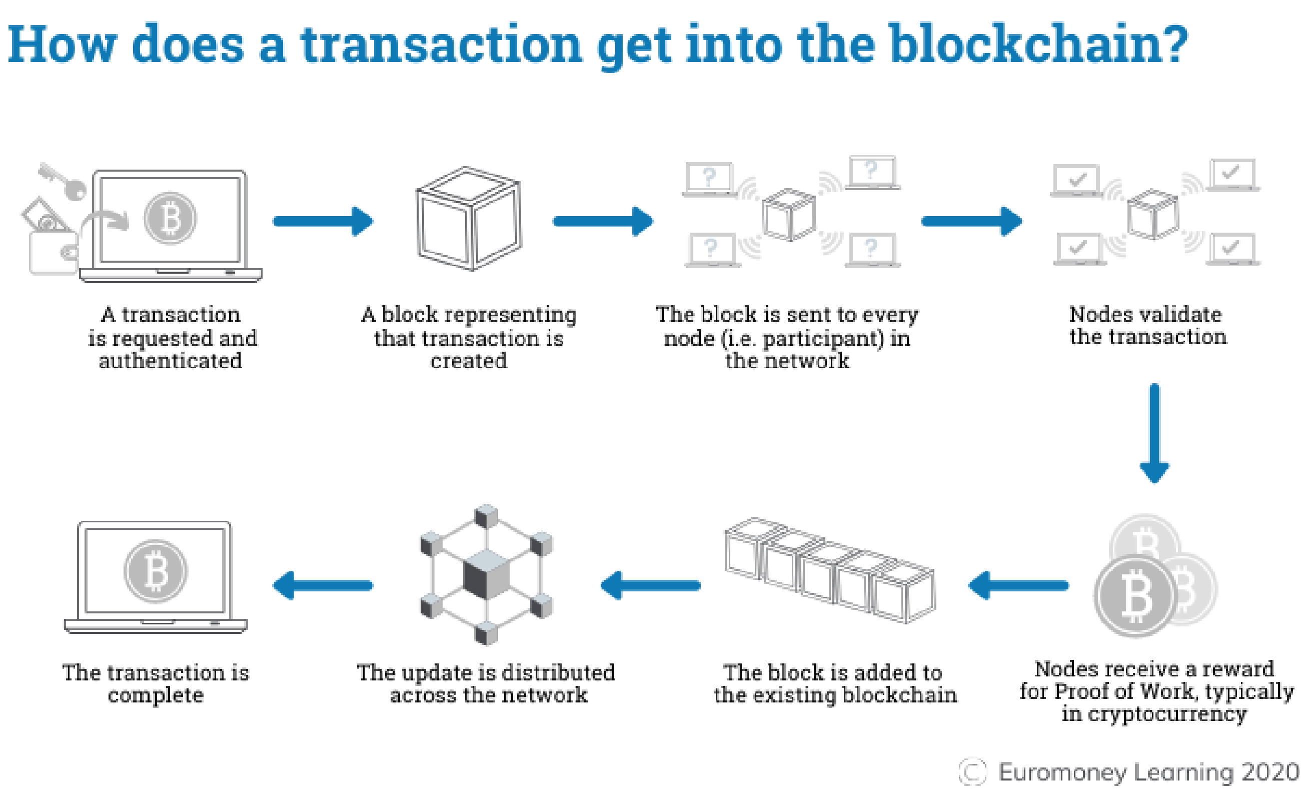
VIDYAMURTHY, G. (2004) *Pairs Trading, Quantitative Methods and Analysis*. NewJersey: Wiley Finance.

YOGESH, J. (2017) 5 impacts of Bitcoin on Economy, Banking & Finance.

XU & LUETKEPOHL (2012) The Role of the Log Transformation in Forecasting Economic Variables. *Emperical Economics,* vol 42, pp 619-638.

Appendix

How miners confirm create coins and confirm transactions:



Available at: https://www.euromoney.com/learning/blockchain-explained/how-transactions-get-into-the-blockchain , 2021)

List of Pairs ADF test 3-month window:

Testing Window	Asset 1	Asset 2	ADF p-value
1	BTC	DOGE	0,0287
2	BTC	DOGE	0,0043
2	BNB	LTC	0,0001
2	ADA	LINK	0,0043
3	BTC	ADA	0,0347
3	XRP	BCH	0,0256
4	BTC	BNB	0,0254
4	ETH	LTC	0,0489
4	ETH	MATIC	0,001
4	BNB	BCH	0,0095
4	BCH	MATIC	0,0213
5	BTC	BCH	0,0366

List of Pairs ADF test 6-month window:

Testing Window	Asset 1	Asset 2	ADF p-value
1	BTC	DOGE	0,0084
1	ETH	LTC	0,0446
1	ETH	LINK	0,0274
1	BNB	XRP	0,0023
1	BNB	BCH	0,017
1	BNB	LTC	0,0039
1	BNB	LINK	0,05
1	DOGE	LTC	0,0419
1	BCH	LTC	0,036
2	BTC	DOGE	0,001

List of Pairs PP test 3-month window:

Testing Window	Asset 1	Asset 2	PP p-value
2	DOGE	XRP	0,0284
4	BTC	DOGE	0,0138
4	BNB	BCH	0,0114
5	BTC	LTC	0,0354
5	ETH	XRP	0,0246

List of Pairs PP test 6-month window:

Testing Window	Asset 1	Asset 2	PP p-value
1	BTC	DOGE	0,0227
1	BCH	LTC	0,0102
1	BTC	DOGE	0,0001
2	BTC	LTC	0,0303
2	ETH	DOGE	0,0147
2	DOGE	LTC	0,0004

List of Pairs JOE test 3-month window:

Testing Window	Asset 1	Asset 2	JOE p-value
1	BTC	DOGE	0,0332
1	BTC	XRP	0,02808
1	BNB	XRP	0,01989
1	ADA	LTC	0,0104
1	DOGE	XRP	0,00183
1	DOGE	LTC	0,0092
1	XRP	LTC	0,01856
2	BNB	DOGE	0,01995
2	BCH	LTC	0,0258
2	BCH	LINK	0,02209
2	BCH	MATIC	0,02011
3	BTC	XRP	0,00985
3	DOGE	MATIC	0,01369
3	XRP	LINK	0,0455
4	BCH	LTC	0,00667
5	BTC	BNB	0,00425
5	ETH	DOGE	0,04228
5	ETH	XRP	0,00339
5	DOGE	XRP	0,00811
5	DOGE	LINK	0,03517
5	BCH	MATIC	0,030482

List of Pairs JOE test 6-month window:

Testing Window	Asset 1	Asset 2	JOE p-value
1	BTC	DOGE	0,01643
1	BNB	XRP	0,03821
2	BCH	LTC	0,0458

Testing and trading windows:

Testing window (3-months)		
	Start date	*End date*
Testing window 1	2020-01-01	2020-04-01
Testing window 2	2020-04-01	2020-07-01
Testing window 3	2020-07-01	2020-10-01
Testing window 4	2020-10-01	2021-01-01
Testing window 5	2021-01-01	2020-04-01
Trading window (3-months)		
	Start date	*End date*
Trading window 1	2018-02-01	2018-05-01
Trading window 2	2018-05-01	2018-08-01
Trading window 3	2018-08-01	2018-11-01
Trading window 4	2018-11-01	2019-02-01
Trading window 5	2019-02-01	2019-05-01
Testing window (6-months)		
	Start date	*End date*
Testing window 1	2020-01-01	2020-07-01
Testing window 2	2020-07-01	2021-01-01
Trading window (6-months)		
	Start date	*End date*
Trading window 1	2020-07-01	2021-01-01
Trading window 2	2021-01-01	2021-07-01