Erasmus University Rotterdam

Erasmus School of Economics

Econometrics and Management Science

Master Thesis: Quantitative Finance

High-Frequency Trading of Cryptocurrencies Through Short-Term Cointegration Pairs-Trading Strategies

Author

Remco de Bruijn

384741

Supervisor

Dick van Dijk

Blocktraders

Sander Adam & Maurits Hartman

May 24, 2019

Abstract

This paper focuses on estimating long-term relations of cryptocurrency exchange rates and trading based on short-term deviations from this long-run equilibrium. These short-term deviations are exploited by means of statistical arbitrage strategies incorporating cointegration methods.

Cointegrating relations in the cryptocurrency market are analyzed, sorted and chosen by means of the Engle-Granger two-step method (EG2SLS). Additionally, we perform a Johansen test for cointegration on the resulting EG2SLS cointegrating relations. Next, we form a pairs-trading portfolio based on the estimated cointegrating relations. A high-frequency trading bot is then configured to trade with this portfolio. We then compare the cointegration-based pairs-trading performance with and without the additional Johansen test step.

The weekly performance of both cointegration-based pairs-trading methods is historically simulated on 16 weeks of cryptocurrency data. We conclude that the Johansen-assisted market neutral pairs-trading strategy outperforms its EG2SLS counterpart. The Johansen and EG2SLS pairs-selection methods result in a weekly return of 6.81% and 5.97%, respectively, including representative transaction costs.

To decrease future computation time, we also consider a method to find a set of exchange rates with a higher probability of being cointegrated. We estimate lead/lag relations of exchange rates in the cryptocurrency market by means of Granger causality tests. We conclude that the lagging exchange rates are more likely to be cointegrated. This means that we can define a subset of all exchange rates, which are more probable of being cointegrated.

Index terms— Cointegration, VECM, cryptocurrencies, high-frequency trading, Granger causality

Contents

1 Introduction

The rise in popularity of cryptocurrencies and high-frequency trading reveals a new market to be subjected to econometric analysis. This market shows strong signs of co-movements between different cryptocurrency exchange rates. This paper focuses on estimating long-term relations of cryptocurrency exchange rates and trading based on short-term deviations from this long-run equilibrium. These short-term deviations are exploited by means of statistical arbitrage strategies incorporating cointegration methods. In collaboration with Blocktraders, a high-frequency cryptocurrency trading (HFT) company, we aim to implement a statistical arbitrage strategy that successfully trades based on short-term trend and long-term equilibrium predictions of cryptocurrency exchange rates. Furthermore, we distinguish several trading characteristics, such as level of diversification and trading frequency, to analyze what is most adequate for HFT bot behavior.

Many high-frequency traders are now joining the cryptocurrency market with continuously adapting trading algorithms. By observing the behavior of this increasingly popular market these past few months, we witnessed co-movements between exchange rates which could set the stage for successful pairs or leading indicator based trading strategies. The observation of co-movements gives us reason to believe that cointegrated currency exchange rates exist in cryptocurrency exchanges and, when properly distinguished, make for excellent short-term portfolios. Cointegration refers to an statistical phenomenon in which a linear combination of non-stationary time series describe a stationary process. When two cryptocurrency exchange rates are in a cointegrating relation, we can measure any short-term weakness in this relation and trade based on the expectation of mean-reversion.

A cointegration-based pairs-trading strategy consists of the selection of a universe, the identification of cointegrated exchange rates en lastly the development of a trading algorithm. In this paper we consider the largest cryptocurrency exchange in terms of volume, two distinct identification methods for cointegration and a trading bot written in Python. Cointegrating relations can be detected by means of cointegration tests, with which we can estimate the deviation of pairs in this relation from their long-term equilibrium. To clarify, if two exchange rates are cointegrated and have deviated from their long-term mean, then we expect this pair to mean-revert. Once the cointegrating relations have been selected, a set of trading rules are necessary to work with the resulting trading signals. The basic pairs-trading decisions are based on deviations from the long-term equilibrium, where the researcher invests in the opposite direction of the deviation (Caldeira and Moura (2013)). Therefore, each time a cointegrating relation is currently deviated from its long-term mean, we can establish a 'position' consisting of a long and short trade. Moreover, when the cointegrating returns to its long-term mean, the position is closed.

The mean, or long-term equilibrium, and any short-term deviations are estimated by means of the Engle Granger 2-step method (EG2SLS) (Engle and Granger (1987)) or the Johansen method (Johansen (1995), Dwyer (2015)). This is a cointegration-based approach to a pairs-trading strategy, which historically has already been proven to be successful (Caldeira and Moura (2013), Miao (2014), Huck and Afawubo (2015)) . The EG2SLS applies two OLS regressions to identify cointegrating relations between currency exchange rates and to estimate deviations from the associated long-term equilibrium. This results in a set of cointegrating relations between cryptocurrency exchange rates with which we can build a pairs-trading portfolio. In this paper, we extend the EG2SLS method by fitting a Vector Error Correction Model (VECM) and perform a Johansen test for cointegration on the identified cointegrating relations. The VECM, as opposed to the OLS regressions, includes short-term trends, which can give rise to different results than the EG2SLS. In other words, the Johansen test for cointegration is used as a 'check' after we establish a set of cointegrating relations with the EG2SLS method. Therefore, the addition of the Johansen test for cointegration will result in a subset of the original

EG2SLS cointegrating relations.

Performing many regressions over all combinations of cryptocurrency exchange rates can be computationally intensive, especially in HFT context. Therefore, we research a method to define a set of exchange rates which are more likely to be cointegrated. Pre-selection of securities, in this case exchange rates, in pairs-trading can significantly decrease computation time (Papadakis and Wisocky (2007), Huck and Afawubo (2015)). To approach this issue, we consider leading and lagging indicators in the cryptocurrency market. These indicators are identified by means of a Granger causality test. We hypothesize that lagging exchange rates are more likely to be cointegrated, given that these exchange rates jointly follow the leading exchange rate(s). These lagging exchange rates are then expected to have co-movements around a long-term equilibrium that is shaped by the leading exchange rate. If this is the case, we can first test for Granger causality once to distinguish a set of potentially cointegrated exchange rates. Next, we can perform the tests for cointegration on the resulting set.

To elaborate, in a market where cryptocurrency sentiment changes rapidly, the underlying exchange rate behavior changes as well. We expect that the popularity of a cryptocurrency exchange rate will affect its volume and, in turn, its leading/lagging behavior towards other exchange rates. In such a market, we could start the week by testing for Granger causality and identifying a set of lagging, or 'Granger caused', exchange rates. Then, throughout the week, we can perform the EG2SLS and Johansen methods on this set of exchange rates. This can save time when working with high frequency data (1 second or faster). In this paper we use a sample rate of 60 seconds due to time constraints, however the strategy should be just as viable in higher sample rates (Perlin (2009), Miao (2014)).

To test our lagging cryptocurrency exchange rate hypothesis, we must have an overview of the trades based on the EG2SLS and Johansen methods. This is done by recording the amount of pairs-trading positions that we open. Additionally, we estimate Granger causality relations between cryptocurrency exchange rates. The hypothesis can be confirmed if the results show that a significant proportion of trades are in Granger caused exchange rates. This would mean that we have a method to save computation time by identifying a set of potentially cointegrated exchange rates, before applying cointegration methods. The Granger causality analysis is done by measuring the percentage of time that exchange rates are significant leading indicators, according to the Granger causality test by means of the VECM. Therefore, it is possible to have several 'market leaders'. This means that the percentage we calculate defines the consistency of certain leading/lagging indicators over time.

Therefore, this paper researches whether or not high-frequency pairs-trading, by means of cointegration, produces significant, positive returns. In addition, we research whether Granger caused exchange rates have an increased probability of being cointegrated.

Both cointegration and Granger causality tests are applied on 16 weeks of intraday minute cryptocurrency data. For our analysis we use Binance data, because it is the largest cryptocurrency exchange, in terms of volume, at the moment of writing. The data consists of 17 exchange rates, all of which have Tether (USDT) as quote currency. USDT is currently the most traded stablecoin on Binance.

The cointegration-based pairs-trading methods are historically simulated, where the trades are performed using a HFT bot. We research different trading characteristics, such as moving window sizes and trading frequency settings. The moving window defines how many observations we keep into account when estimating a model at any moment in time. Trading frequency is affected by the threshold at which we consider a deviation to be (i) large enough to open a position and (ii) small enough to close an existing position. We integrate transaction costs in all trading decisions (the transaction fees are generally between 0.01% and 0.2% in cryptocurrency market exchanges), therefore we can see the consequences of different trading frequencies. This means that we would only open a position if a deviation at least exceeds the transaction costs.

The results, using a sample rate of 60 seconds, show that the Johansen method outperforms the EG2SLS with an average weekly return of 6.81%, compared to 5.97%. In addition, the Johansen case only performed around half of the original EG2SLS trades (77 and 152 positions per week, respectively). This means that a Johansen based pairs-trading algorithm results in 0.20% returns per position, whereas the EG2SLS strategy made 0.14% per position. In conclusion, our research shows that high-frequency pairs-trading in the cryptocurrency market is significantly profitable. More specifically, there exist cointegrating relations in the cryptocurrency market with which a profitable pairs-trading portfolio can be formed. Moreover, analysis on weekly returns for both methods show that the Johansen-based returns have a tighter confidence interval. Therefore, we conclude that the pairs-selection by means of the Johansen method outperforms the EG2SLS method in terms of returns and variance.

Apart from our research on cointegration-based pairs-trading strategies, we also examine the existence of leading/lagging indicators by means of Granger causality. The results over 16 weeks of cryptocurrency market data show that the larger (in terms of volume) coins, Bitcoin (BTC) and Ethereum (ETH), Granger cause most of the other, smaller, coins. Interestingly, the other leading indicators are Bitcoin Cash (BCH) and Ethereum Classic (ETC), which are both hard forks of the original BTC and ETH branches, respectively. The concept of hard forks is further discussed in section 5.3. In addition, the Granger causality tests show that, most of the time, the cryptocurrency market is led by EOS, which most likely is the result of a temporary popularity boost in the relatively small sample of data used in our paper. More on this issue can be found in section 3. Lastly, we identify several lagging cryptocurrency exchange rates in the market. These exchange rates are traded the most by our algorithm and are also part of the top cointegrating relations revealed in the EG2SLS and Johansen-based pairs-trading analysis. This supports our hypothesis that we can find a subset of potentially cointegrated exchange rates by performing Granger causality tests, adjusted for short-term trends.

In summary, we research the profitability of two possible approaches to pairs-trading, the EG2SLS and Johansen method, and we examine the connection between Granger caused exchange rates and cointegration. This paper consists of a literature review in section 2, a data section 3, a methodology on the execution and assessment of a pairs-trading strategy in section 4.1, the methodology of estimating Granger causality in section 4.2, the results in section 5, discussion in section 6 and lastly, the conclusion in section 7.

2 Review of literature

High-frequency traders from the stock and FX markets (King and Rime (2010)) are now joining the much more volatile, unregulated, cryptocurrency market. Therefore, many trading strategies that are implemented come from the FX market. There exist 234 cryptocurrency exchanges at the moment of writing [1], each with their own transaction cost functions (also known as fee brackets or fee schedules). Due to the relatively low amount of research in the cryptocurrency market at the moment, it is useful to apply the vast amount of FX and stock market research to this emerging digital currency market. This subsection will focus on relating FX and stock market research to the cryptocurrency market.

Pairs-trading is an investment strategy with which investors search for long-term relations of two securities and invest based on assumed mean-reversion. If a paired set of securities in a pairs-trading portfolio are deviated from their long-term equilibrium, the investor expects that these revert to their mean. There exist different approaches to estimating these long-term relations, such as minimum distance (Gatev et al. (2006), Perlin (2009)), cointegration (Caldeira and Moura (2013), Miao (2014)) or stochastic spread (Elliott et al. (2005)) to name a few. Another, widely used, method of finding statisti-

[1] https://coinmarketcap.com/

cal relations between securities is that of correlation. The effect of correlation is related to that of cointegration, however they emphasize different notions. A high correlation underlines co-movements in assets, however this does not necessarily depict long-term equilibria. As a result, hedging strategies based on correlation require frequent rebalancing. Cointegration measures long-term co-movements in prices, also in periods where correlation seems low (Alexander (1999), Miao (2014)).

Huck and Afawubo (2015) discuss the real-world applications of three pairs-trading algorithms, namely cointegration, stationarity and minimum distance. They conclude that the cointegration-based strategy, incorporating risk factors and transaction costs, is the most profitable and robust.

Cointegration-based analyses are usually long-term, however Perlin (2009) and Miao (2014) report that relatively high-frequency data (sample rate of 15 minutes, 1 day respectively) overperform compared to lower-frequency observations. This is also intuitive, due to the fact that there are simply more trading moments when more observations are regarded within the same time period. In our paper, we focus on implementing a cointegration-based pairs-trading strategy on a sample rate of 1 minute. More information on the data can be found in section 3.

As a safety measure, investors impose a stop-loss condition to prevent large losses in pairs-trading (Nath (2003), Miao (2014), Leung and Li (2015)). Given that a cointegrating relation is a mean-reverting process, many academic papers on cointegration do not implement a stop-loss condition (Elliott et al. (2005), Gatev et al. (2006), Perlin (2009)). Cryptocurrencies, in comparison with fiat currencies, are considered to have growth value. This valuation is affected by changes in the blockchain technology of these cryptocurrencies. These changes, called updates/forks [2], can be compared to a company/financial institution announcement in the stock/FX market. Shifts in expected growth of the cryptocurrencies bring a new layer of volatility to the table, which can have an effect on the

[2] A fork, or hard fork, occurs when a cryptocurrency splits into two cryptocurrencies.

durability of a cointegrating relation. This means that such an event hypothetically could affect an exchange rate such that they are no longer a pair. For this reason, we implement a stop-loss condition for the trading bot which is discussed in section 4.3.

The aforementioned cryptocurrency market risk gives rise to the need of diversification. Gatev et al. (2006) state that bankruptcy risk exists for the securities of a pairs-trading strategy as an idiosyncratic component, which can result in asymmetric returns from the long and short components. They note that there are diversification benefits from combining multiple pairs in a portfolio, this is shown by comparing the number of months with negative returns between a portfolio of the top 5 and top 20 pairs. In section 4.3 we propose a similar method to analyze the effect of diversification in the cryptocurrency market. The next section will focus on the implementation of a cointegration-based pairs-trading strategy.

3 Data

Trading in the cryptocurrency market is possible through many exchanges all over the world. In comparison to the FX market, the cryptocurrency market features advantages such as low transaction costs and higher volatility. On the other hand, the FX market is more stable, liquid and offers a wide variety of options for trade. Moreover, many cryptocurrency exchanges have suffered hacks resulting in large losses. The extreme highs and lows of this unregulated market show that it is still far from stability and safety. This high volatility is naturally of interest and can potentially be exploited by means of econometric analysis. The trading behavior amongst the cryptocurrency exchanges is quite different, aspects such as trading volume, average spread and even price tick-sizes differ. This is partially a result of the fact that the cryptocurrency market is unregulated.

The Blocktraders company supplies data from Binance[3], the largest (by volume) cryp-

[3] https://www.binance.com

Table 1 Data characteristics of cryptocurrency exchange rates on the Binance exchange between '2018-05-19' and '2018-09-03' (107 days) with quotation USDT.

Characteristic	Binance
Frequency	Per second
Exchange rates	332
Distinct exchange rates (with USDT)	17
Horizon	'2018-05-19 12:00:00' until '2018-09-03 23:59:59'
Trading volume	41 per second
Observations per exchange rate	9,288,000

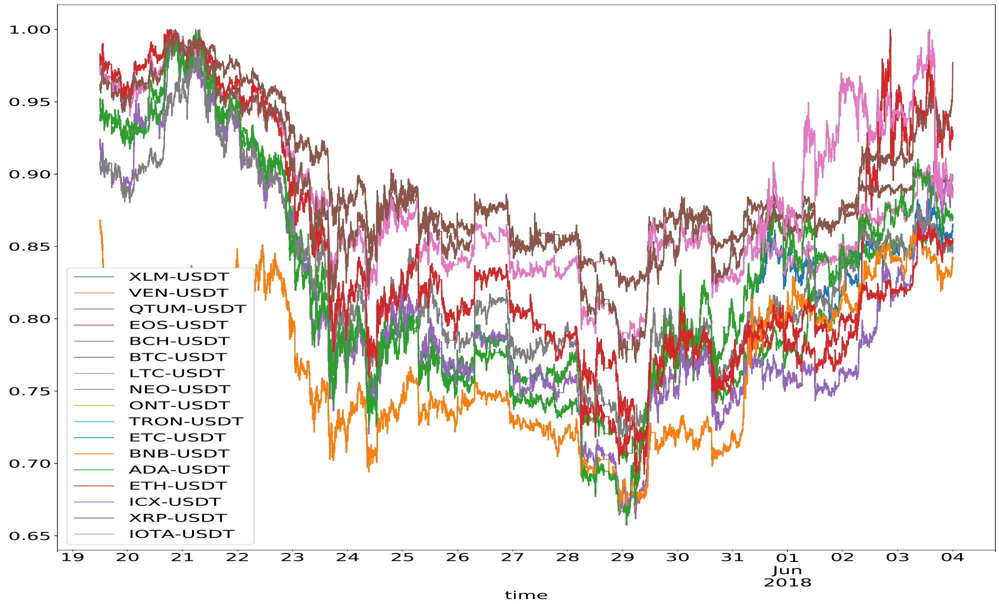
tocurrency exchange at the moment of writing. Information on Binance can be seen in table 1.

We opt for a high-volume exchange because limit orders, placed in low-volume exchange rates, will take significantly longer to fill. One could also consider market orders in this case, because these trigger instantly at the expense of increased transaction costs and the spread. Pairs-trading strategies, however, usually profit from small deviations, which makes it sensitive to increases in transaction costs.

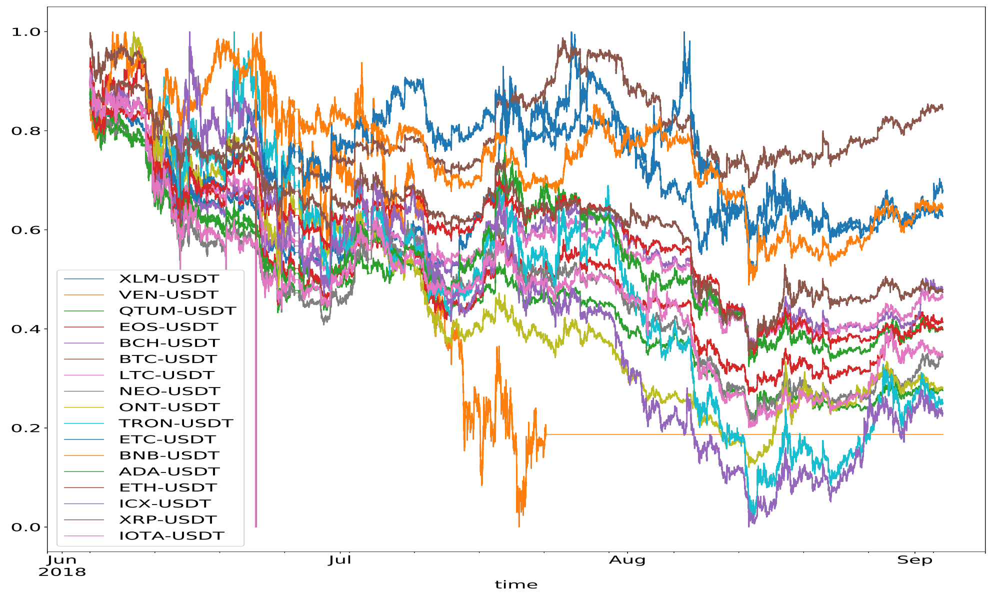
The analyses in this paper will be performed done using mid-price data of the cryptocurrency exchange rates extracted from the Binance order books. Each cryptocurrency quotation is in terms of quote Tether (USDT), which is a proxy for the USD. The USDT/USD exchange rate is therefore stationary around 1.

Considering a sample rate of 1 second for 17 exchange rates over 107 days, we nearly have 160 million observations in total. We have decided to resample the data to 60 seconds throughout this paper. This is done because (i) our analyses are computationally intensive and (ii) many exchange rates have little/no variance at a 1 second sample rate due to low volume. The low amount of variance in a 1 second sample lets us resample the data without the loss of too much information.

Our data is split into two sections, namely a training and testing set, to optimize our trading parameters. A small set of 15 days ($\approx 15\%$ of our sample) is used to perform a grid search on variables such as window size, optimized over final returns of the pairs-



(a) 15 days of normalized cryptocurrency prices in training period from '2018-05-19 12:00:00' until '2018-06-03 23:59:59'



(b) 92 days of normalized prices of each considered cryptocurrency prices in the subsequent period from '2018-06-04 00:00:00' until '2018-09-03 23:59:59'. Note the flatline of VEN-USDT from around the 23rd of July, which is the result of Binance changing the VEN ticker to VET due to a rebranding.

Figure 1: Plots of the training and remainder period cryptocurrency prices, extracted from Binance. The model parameters are estimated in the training period. Afterwards, the performance of the models is evaluated in the entire period in which the parameters are not adjusted.

trading algorithm. The grid search is performed on a small sample to gain some general insight on variable sensitivity, while not over-optimizing the results for the relatively small remaining time frame of our sample. The Binance data consists of intraday observations between '2018-05-19' and '2018-09-03' (16 weeks) with a frequency of 1 minute for 17 base currencies with quotation USDT. In the training period, from '2018-05-19 12:00:00' until '2018-06-03 23:59:59', we fit the models and choose the best performing variables based on a grid of possible values. Afterwards, we assess the model performance on the subsequent period from '2018-06-04 00:00:00' until '2018-09-03 23:59:59' (92 days). Figure 1 shows a division of two plots with the intervals in which the models are trained and tested. It also shows signs of co-movements of the normalized prices. Table 2 contains the summary statistics of the Binance data. We see that a few exchange rates have missing observations, such as ICX-USDT, ONT-USDT and worst case VEN-USDT. Figure 1b shows that VEN-USDT flatlines after the 23rd of July 2018 due to a ticker change from VEN to VET (VeChain). Despite the change in sample after 23rd of July, we choose to include this data in our sample to best represent the realistic situation of the unpredictability of the cryptocurrency market. If a estimation window contains more than 10% missing observations for an exchange rate, we do not include it in our EG2SLS regression. Otherwise the last known price is used to fill the missing observations. Furthermore, a position cannot be opened if there are missing observations for an exchange rate in a cointegrated pair.

Table 2 This table contains the Binance data summary statistics. All exchange rates are in quote currency USDT.

Statistic	XLM	VEN	QTUM	EOS	BCH	BTC	LTC	NEO	ONT	TRON	ETC	BNB	ADA	ETH	ICX	XRP	IOTA
count	825761	636039	928800	853117	928800	928800	928800	928800	757331	729970	724211	928800	928800	928800	714129	928800	825760
missing	103039	292761	0	75683	0	0	0	0	171469	198830	204589	0	0	0	214671	0	103040
mean	0.24	2.04	8.63	8.27	784.62	6964.50	85.86	34.66	3.75	0.03	15.23	13.17	0.15	450.56	1.26	0.47	0.99
std	0.04	0.33	3.33	2.65	188.05	651.14	22.42	12.49	1.69	0.01	1.85	1.89	0.04	110.19	0.47	0.11	0.35
min	0.17	1.57	3.56	4.18	472.26	5760.42	49.37	0.00	1.03	0.02	10.11	8.54	0.08	250.71	0.45	0.25	0.40
25%	0.21	1.83	5.88	6.47	650.94	6421.75	67.03	24.23	2.35	0.03	13.52	12.05	0.12	376.17	0.86	0.36	0.72
50%	0.23	1.83	8.15	8.01	761.54	6738.82	82.43	33.49	3.36	0.03	15.48	13.45	0.14	462.78	1.31	0.46	0.99
75%	0.27	2.26	10.67	9.74	879.59	7482.00	97.90	39.97	5.10	0.04	16.68	14.35	0.17	517.83	1.61	0.54	1.14
max	0.35	2.97	17.13	15.64	1311.79	8593.34	140.93	66.36	8.59	0.05	20.67	17.44	0.26	724.00	2.39	0.71	2.01

The models are estimated on very large amounts of observations in relatively short time periods, which means that the results are highly sensitive to economic circumstances over the regarded horizon. For example, section 5.3 on Granger causality shows that the EOS/USDT exchange rate is a leading indicator (as opposed to Bitcoin and Ethereum), which was only the case in our sample due to a relatively short hype. This example shows us that a small window in the cryptocurrency market will bias the results, regardless of the number of observations. As a precaution, we will consider the weekly performance of each model in terms of robustness to better compare their relative performance (note that the mentioned problem still persists).

4 Methodology

This section is divided into 4 parts. First we explain how the pairs-selection process of the currency exchange rates is performed, second, we explain the methodology of testing for Granger causality to find leading and lagging indicators, third, we approach the methods of pairs-trading and lastly we describe the performance assessment of our trading mechanisms.

4.1 Pairs-selection

Error correction models can be used to estimate the long-run equilibrium resulting from a cointegrating relation between two, or more, variables. While both methods identify the presence of cointegration in multivariate sets, the two-step Engle-Granger method regards cointegrating relations individually, whereas the vector error correction model (VECM) in collaboration with the Johansen test can identify the number of cointegrated relationships within the set. In this paper we want to research the difference in performance of the EG2SLS compared to a Johansen based approach. Since our entire analysis is bivariate, they differ only on the in/exclusion of short-term trends.

Initially, we perform the Engle-Granger two-step method for each combination of cryptocurrency exchange rates $M = \binom{N}{2}$. It is our expectation that many of these exchange rates will be part of cointegrating relations, for that reason we will perform an analysis on a subset of all cryptocurrency exchange rates containing the k strongest cointegrating relations. The parameter k can be regarded as a 'level of diversification' in the sense that the higher k is, the more distinct cointegrating relations are considered each window for trading. We can also estimate a bivariate VECM to gain additional, short-term, insight in the relation between two exchange rates. It is our expectation that the short-term trends in the VECM may have an effect on the conclusion of cointegration.

First, we will elaborate the steps required to perform the Engle-Granger 2 stage least squares (EG2SLS) procedure. Secondly, we will describe the VECM and the additional explanatory power of the Johansen method with respect to the Engle-Granger method.

4.1.1 Engle Granger 2 stage least squares

The error-correction models (ECMs) have a close relationship with cointegration. Engle and Granger (1987) explain the mechanics of the ECMs as that a proportion of the disequilibrium from one period is corrected in the next period. To estimate these disequilibria, we must first establish which exchange rates are cointegrated.

Engle and Granger (1987) propose a test for cointegration between time series, the Engle-Granger two-stage least squares method. The EG2SLS method essentially checks whether or not a linear combination of two non-stationary variables is stationary, which is the definition of cointegration. The steps are as follows, assume that x_t and y_t are two time series.

1. Test x_t and y_t for a unit root by means of the augmented Dickey-Fuller test. If the series are $I(1)$, we can proceed.

2. Perform an ordinary least squares (OLS) regression of y_t on x_t with an intercept: $y_t = \alpha + \beta x_t + \epsilon_t$

3. Test the estimated residuals $\hat{\epsilon}$ for a unit root with the augmented Dickey-Fuller test. Under H_0 we assume non-stationarity, or the presence of one or more unit roots. If the residuals are stationary, we can reject the null hypothesis of no cointegration.

Note that if a time series is stationary after differencing once, it is said to be $I(1)$ (a series that is $I(p)$ must be differenced p times before being stationary). Furthermore, the ADF test on the residuals in step 3) are performed with the MacKinnon (1991) critical values, adjusted for the number of variables.

The (non)-stationarity of the exchange rates from each currency $y_j, j \in \{1, \ldots, N\}$ to USDT is tested by means of the augmented Dickey-Fuller test, where we pick p from the p_{max} distinct regressions based on the Akaike information criterion (AIC). Here p denotes the number of lags included in the regression and N are the amount of distinct exchange rates, thus we regress:

$$\Delta y_{j,t+1} = \alpha + \beta t + \delta y_{j,t} + \sum_{k=0}^{p} \gamma_k \Delta y_{j,t-k} + \epsilon_{t+1} \tag{1}$$

and for $p \in \{1, 2, ..., p_{max}\}$, choose the lag with the highest value for the AIC. The ADF regression contains an intercept and a time trend. We test for a unit root under the null hypothesis $\delta = 0$, if the null is rejected then we assume the series is stationary. Furthermore, the lags are included to rid the effect of serial correlation, a relation which decreases the power of the ADF test. We use the same number of lags in both the first and third step of the EG2SLS method. The maximum number of lags p for the augmented Dickey-Fuller test was researched by Ng and Perron (1995) and can be determined by using a rule of thumb (Miao (2014)):

$$p_{max} = \lfloor 12 * (T/100)^{1/14} \rfloor \tag{2}$$

where T is the sample size. If both time series y_t and x_t are $I(1)$, by failing to reject the null-hypothesis of the ADF test, then it is possible to test for a cointegration relation. If there exists a linear combination of y_t and x_t such that the resulting vector is $I(0)$, then y_t and x_t are cointegrated. After we perform the ADF test on all time series of cryptocurrency exchange rate prices, and conclude that they are $I(1)$, we can continue with the EG2SLS. For each combination of currency exchange rates $i \in \{1, \dots, M\}$, perform an OLS regression of exchange rate $y_{i,t}$ on $x_{i,t}$ with an intercept:

$$y_{i,t} = \alpha + \beta x_{i,t} + \epsilon_{i,t} \tag{3}$$

where i denotes one of the cryptocurrency exchange rate combinations. The following step in the Engle-Granger two-step procedure is to test residual $\hat{\epsilon}_i$ for a unit root with an ADF test, without a constant or trend. Harris (1995) notes that the constant/trend deterministic components should either be added to the cointegration regression (eq. 3) or the ADF test, not both. If we reject the null hypothesis, the residual is stationary. Assuming we passed the first step of the EG2SLS, this means that there exists a linear combination of two non-stationary series, which is stationary. Therefore, from the rejection of the null we conclude y and x are cointegrated.

From the resulting cointegrated cryptocurrency exchange rates, we sort the top-k cointegrating relations based on the lowest p-values ensuing from the ADF unit root test of residuals $\epsilon_{i,t}$. This notion of "level of cointegration" is used to sort all cointegrating relations, which helps us prioritize investments.

4.1.2 Vector error correction model

The multivariate counterpart of the ADF test, mentioned above, was researched by Johansen (1995). The Johansen test by eigenvalues attempts to maximize canonical correlations based on eigenvalues of transformations of the data. The idea is to find the rank of Π of the vector autoregression (VAR) model of time series $\boldsymbol{x}$:

$$\Delta \boldsymbol{x}_t = \Pi \boldsymbol{x}_{t-1} + \sum_{i=1}^{l-1} \Pi_i \Delta \boldsymbol{x}_{t-i} + u_t \tag{4}$$

For $l = 1$ the second term on the right-hand side drops out. The matrix Π can be written as:

$$\Pi = \alpha\beta'$$

where β is a vector/matrix of cointegrating vectors, $\beta'\boldsymbol{x}_{t-1}$ denotes the cointegrating relationship and α is the adjustment coefficient. If $rank(\Pi) = 0$, there are no cointegrating relations. The rank of Π is equal to the amount of cointegrating vectors, therefore the Johansen test makes use of eigenvalues to compute $rank(\Pi)$. The test iteratively examines the ordered eigenvalues and tests for $H_0 : \lambda_i = 0$, where λ_i is the ith eigenvalue. If $H_0 : \lambda_i = 0$ is rejected we conclude $rank(\Pi) = i - 1$. In addition, if we have n variables and $rank(\Pi) = n$, we can conclude that the variables have no unit roots. This is because every variable can be written as a linear combination of the others.

The Johansen method will be used as a final check after the EG2SLS, meaning that we will already have a set of cointegrating relations. Therefore we only need a bivariate model to check each individual pair. Next, we write the bivariate VECM(p) of currencies y_i and x_i $i \in \{1, \ldots, M\}$ as:

$$\begin{bmatrix} \Delta y_{i,t} \\ \Delta x_{i,t} \end{bmatrix} = \mu_0 + \Pi_1 \begin{bmatrix} y_{i,t-1} \\ x_{i,t-1} \end{bmatrix} + \sum_{k=1}^{p} \Gamma_k \begin{bmatrix} \Delta y_{i,t-k} \\ \Delta x_{i,t-k} \end{bmatrix} + \begin{bmatrix} u_{i1,t} \\ u_{i2,t} \end{bmatrix} \tag{5}$$

where Π_1 should have rank 1, which can be confirmed by testing the residuals $\hat{\epsilon}_{i,t-1}$ from equation 3. The model contains a constant inside and outside of the cointegration relation, these are identified when there is a unique long-run equilibrium. The short term trends in the VECM can impact the conclusion of cointegration when compared to the EG2SLS. If $rank(\Pi_1) = 1$ we can rewrite equation 5 by factorizing $\Pi = \alpha\beta'$, as follows:

$$\begin{bmatrix} \Delta y_{i,t} \\ \Delta x_{i,t} \end{bmatrix} = \mu_0 + \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} y_{i,t-1} - \beta_0 - \beta_1 x_{i,t-1} \end{bmatrix} + \sum_{k=1}^{p} \Gamma_k \begin{bmatrix} \Delta y_{i,t-k} \\ \Delta x_{i,t-k} \end{bmatrix} + \begin{bmatrix} u_{i1,t} \\ u_{i2,t} \end{bmatrix} \tag{6}$$

where α are the loading coefficients which symbolize the speed of adjustment of the mean-reversion process and β contains the cointegrating vector. The error terms $u_{i\{1,2\},t}$ are white noise processes. The number of lags p are selected to maximize the AIC of the resulting regressions in equation 6. The deviations from the long-term equilibrium of the cointegrating relations are captured by the $\hat{\epsilon}_{i,t-1}$ term from equation 3.

The existence of a cointegrating relation between two exchange rates can be examined by means of the EG2SLS method or by the rank of Π in the VECM. Given that the latter incorporates short-term trends in the data, it can be that the conclusion of the existence of cointegration differs for these two procedures. Therefore, we will analyze the effect of the in/exclusion of these short-term trends in cointegration-based pairs-trading strategies. This means that when we trade using the Johansen method, a cointegrating relation will be discarded when $rank(\Pi) \neq 1$.

In summary, to perform a pairs-trading strategy based on the Johansen method we first obtain the top-k cointegrating relations by applying the EG2SLS method. Secondly, we estimate a VECM(p) and assess the rank of Π for each pair. If $rank(\Pi) \neq 1$ for a certain exchange rate in a window, then we exclude that exchange rate from any trades at that time.

4.2 Granger causality

In this section we try to find evidence of Granger causality in the cryptocurrency market. If this is the case, we can identify leading and lagging indicators. Once the lagging exchange rates have been identified, we can compare these with the most traded cointegrated pairs to find whether Granger caused exchange rates are more prone to being cointegrated. Granger causality is tested by means of a VECM(p) model. Let $\mathbf{Y}_t = \{y_1, \ldots, y_N\}'$ be a $(Nx1)$ vector of cryptocurrency exchange rate prices, then we can write the VECM(p) model as:

$$\boldsymbol{\Delta Y}_t = \boldsymbol{\mu} + \Pi_0 \boldsymbol{Y}_{t-1} + \sum_{k=1}^{p} \Pi_k \boldsymbol{\Delta Y}_{t-k} + \boldsymbol{\epsilon}_t \tag{7}$$

where $\boldsymbol{c}$ is a constant and $\boldsymbol{Y}_i$ Granger causes causes $\boldsymbol{Y}_j$ if and only if element $\boldsymbol{\Pi}_{(j,i)}$ is significant for $k = \{1, \ldots, p\}$. These tests are performed in section 5.3.

4.3 Pairs-trading

In this subsection we approach two different methods to trade with the deviations $\hat{\epsilon}_{i,t+h}$ from the model in equation 3. We consider when and how often to open/close positions, based on our desired trading frequency and the transaction costs.

A pairs-trading strategy is generally market-neutral, however, the cryptocurrency does not offer many shorting opportunities yet. Few exchanges offer shorting opportunities and, when they do, it usually holds for a small subset of exchange rates. In this paper we can theorize the situation when short contracts are properly implemented in the cryptocurrency market. In our simulation, we expect to profit more from many short-term trades at the cost of smaller deviations in comparison with a strict mean-reversion pairs-trading strategy. We distinguish these two methods by factoring in the holding time of a position Δ_t, which will be further explained in this section.

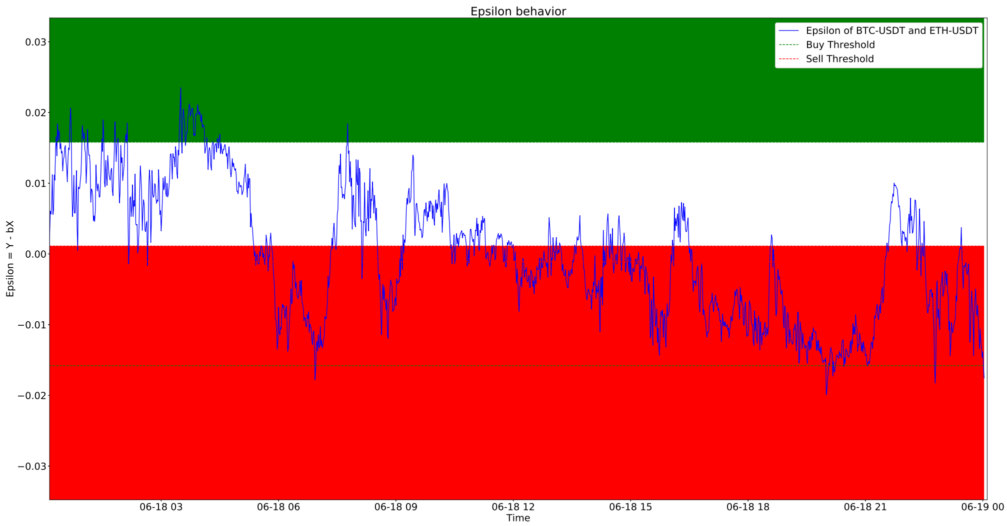


Figure 2: This graph shows a snapshot of $\hat{\epsilon}$ of the regression between BTC-USDT and ETH-USDT from equation 3. The trading thresholds c and s are based on opening/closing percentiles $p = 97$ and $q = 55$, respectively. When $\hat{\epsilon}$ hits the c threshold, in the green area, a short contract is opened for Bitcoin-USDT and a long contract is opened in ETH-USDT. The position is closed shortly afterwards, when $\hat{\epsilon}$ hits the s limit, in the red area. The example given in this figure would result in 7 closed positions.

The process of cointegration-based pairs-trading is clarified in figure 2, in which the trading decisions are optimized by minimizing holding time Δ_t and maximizing $\hat{\epsilon}_{i,t} - \hat{\epsilon}_{i,t+t_{close}}$, where $\hat{\epsilon}_{i,t}$ is the residual of regressing BTC-USDT on ETH-USDT. Assume BTC-USDT and ETH-USDT are cointegrated and we have 2 thresholds, one for entering the position c and one for exiting s. Each threshold is calculated by taking a percentile of $\hat{\epsilon}_i$. If $\hat{\epsilon}_{i,t}$ is large enough and exceeds the entering threshold, we assume that BTC-USDT/ETH-USDT will decrease/increase (green zone of figure 2). So at time t we open position by buying ETH-USDT and selling BTC-USDT. The position is closed at t_{close} when $\hat{\epsilon}_{i,t}$ mean-reverts to the closing threshold (red zone of figure 2).

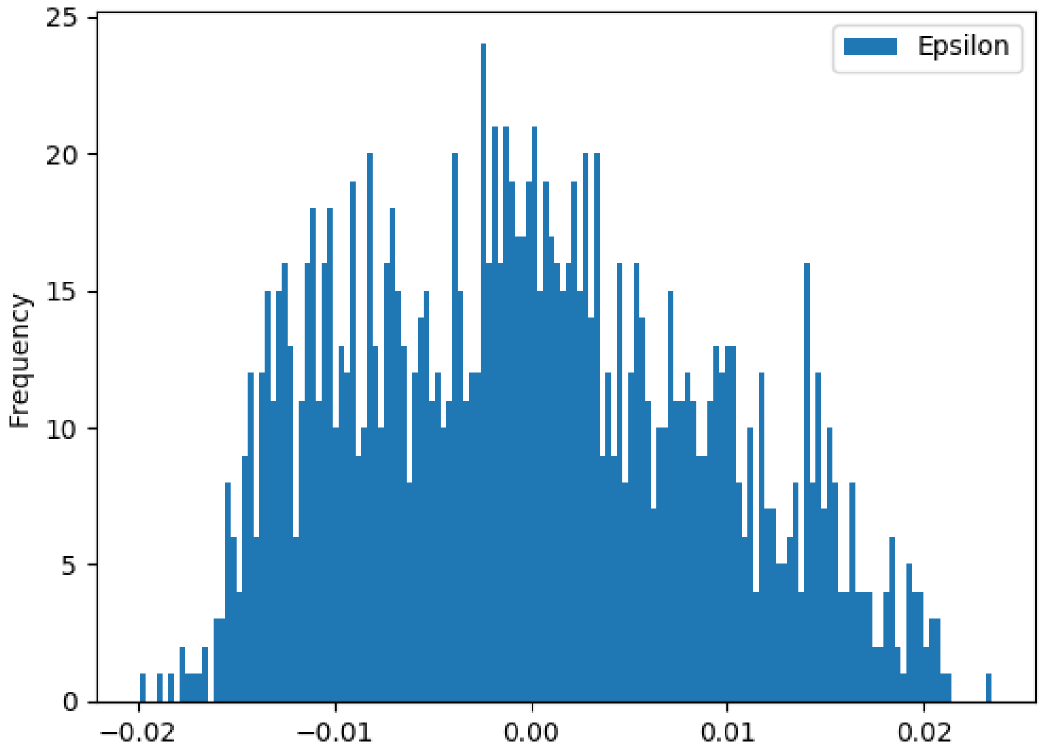


Figure 3: Histogram of a single window of $\hat{\epsilon}_{i,t:T}$

Due to the transaction costs in every cryptocurrency market, we require that the deviation from the long-term equilibrium should at least be larger than the transaction costs. The trading bot is instructed to place a limit order when deviation $|\epsilon_{i,t}| > c$, where $\epsilon_{i,t}$ indicates the deviation from the long-term equilibrium of cryptocurrency i at time t

and c denotes the trading threshold. Furthermore, we require a numerical value to attach to both thresholds. This is done by means of percentiles of $\hat{\epsilon}_i$ from equation 3. Certain asymmetries can cause $\hat{\epsilon}_i$ to be skewed, as shown in figure 3. In this case it is useful to adjust the percentiles, which we will elaborate further. To open a position, keeping into account transaction costs and asymmetric deviations, we define the following equation:

$$c_i = \max\left(c^t + s_i, \frac{\hat{\epsilon}_{i,t:T,\left(p\%\right)} + |\hat{\epsilon}_{i,t:T,\left(1-p\%\right)}|}{2}\right) \tag{8}$$

where c^t are the transaction costs, s_i is the selling threshold and $\hat{\epsilon}_{i,t:T,\left(p\right)}$ denotes the p^{th} percentile of $\hat{\epsilon}_{i,t:T}$ from equation 3. Normally the selling threshold $s_i = 0$, meaning that we await (complete) mean-reversion. However, as can be seen in figure 2, sometimes the deviations narrowly miss the red area. This means that the trade would remain open for a much longer time, which results in greater opportunity costs than the small loss of transaction costs. To look at this effect, we consider making the selling threshold s_i non-zero. To close a position, we define the selling threshold s_i in a similar fashion:

$$s_i = \frac{\hat{\epsilon}_{i,t:T,\left(q\%\right)} + |\hat{\epsilon}_{i,t:T,\left(1-q\%\right)}|}{2} \tag{9}$$

where q will be the 50% percentile in the basic pairs-trading case. By increasing q, we increase the number of possible trades ceteris paribus (at the cost of closing with smaller returns).

Using the fact that the trades are based on the behavior of epsilon, we can approximate the number of trades based on its order statistic. The maximum function is taken to make sure that the trade at least exceeds the transaction costs in expectation. Due to the skewness seen in figure 3, we take $p >= 50\%$ and take the average of the ϵ_i order statistic. Given that, most likely, the percentile $p < 50\%$ is negative, we must take the absolute value of $\hat{\epsilon}_{i,t:T,\left(1-p\%\right)}$. The left and right percentiles of the histogram of $\hat{\epsilon}_{i,t:T}$ should be equal in absolute value as $t \to \infty$, taking the average is a precaution for the event of strong shocks in a small window. By defining the trading threshold in this manner, we adjust the threshold dynamically for pairs with different levels of volatility.

Together, the thresholds c_i and s_i establish the areas for the basic behavior of the trading bot. These thresholds are a result of the estimation of $\hat{\epsilon}_{i,t:T}$ and opening/closing percentiles p and q. In case the size of the trading window is smaller than the transaction costs $c_i - s_i < c^t$, as a result of the percentiles, we adjust p and q with small increments to make the trading window larger $(c_i - s_i)$ until the second argument of equation 8 is larger than the first. For the next iterated pair, the percentiles p and q are reset to their original values.

The holding time of the algorithm is tested in 2 different manners, namely (1) with a direct approach, where place a limit order when $|\hat{\epsilon}_{ti}| > c_i$ and close the position when $|\hat{\epsilon}_{i,t+\Delta_t}| < s_i$ (from eq. 3) with sell limit $s_i \approx 0$, and (2) by means of a more risk-averse approach which discourages long holding times. Method (1) can potentially hold funds for a very long time, because mean-reversion is anticipated but may not be realized due to the strict selling threshold. Method 2 penalizes long holding times and sells at profitability, thus method 2 suffers decreased profits in exchange for a shorter holding period. The second method opens at the same moment as method 1 and closes when

$$|\hat{\epsilon}_{t+\Delta t,i}| < s_i(1 + \frac{\Delta t}{h}) \tag{10}$$

where Δt is the total holding time of the current trade. Given our expectation that the first method can be illiquid, we chose to compare the methods in this paper to better understand the trade-off between trading frequency and returns. Preferably, $|\hat{\epsilon}_{ti}|$ is maximized, however since we do not know price fluctuations a priori, we are bound to a deterministic trading threshold. Equation 10 is expected to result in a high amount of transactions, due to the relation often being true for small Δt. The method parameters are summarized in table 3. In this paper we train both methods with a cryptocurrency dataset and test their performance over all weeks in our data. By analyzing the weekly results of the methods, we can gain insight in both the robustness and accuracy of the methods.

Table 3 This table summarizes the trading threshold parameters c and s for methods 1 and 2.

Method	Opening Threshold c	Selling Threshold s
1	c_i	s_i
2	c_i	$s_i(1 + \frac{\Delta t}{h})$

As a safeguard we deem it wise to implement a 'stop'-condition for the trading bots in case the market, for instance, keeps deviating from the long-term equilibrium. Miao (2014) does this by applying a stop-loss condition when the regression residual $\hat{\epsilon}_t$ from equation 3 hits 2δ positive or negative standard deviations from the long-term mean μ. The problem with this stop condition is that the loss is 'accepted' no matter the behavior of the price; as soon as it surpasses the threshold, we close. Therefore we suggest an alternative, more statistically relevant, stopping condition, based on the stationarity of the price. The biggest problem with the trading algorithm is that, for example, ϵ_i from figure 2 would become non-stationary. This means that exchange rate i would no longer be part of a cointegrating relation. This can happen when a hard fork or blockchain update occurs for an individual coin. This, in turn, affects the exchange rates of the currency. Regardless of a two exchange rates being cointegrated, we believe that these events can affect the cointegrating relation between two exchange rates. We suggest a relatively small moving window h which performs an ADF test to confirm whether $\hat{\epsilon}$ is still stationary ($I(0)$), if this is not the case we impose a full-stop. Thus if the null of non-stationarity is failed to be rejected, we close our current opened trade(s). It is our belief that this method will better acknowledge deviations from the long-term mean and evaluate whether or not the deviation, $\epsilon_t > 2\delta$, has a decent probability of recovering towards the mean μ.

Lastly, we recalculate the Engle-Granger two-step procedure each $\dfrac{h}{4}$ (quarter window) seconds to see if there are any changes in the top ranked cointegrating relations. If it is the case that currency exchange rate a no longer belongs to the top-k cointegrating relations, we sell exchange rate a when method x ($x = 1, 2$) allows it and commence the subsequent trade in a exchange rate from the list of k openable positions. A simplified flowchart of the bots' mechanics is shown in figure 4. The positions cap (maximum number of simultaneous opened positions) is equal to the number of cointegrating relations k; note that this is not necessary, but this is done for simplicity.

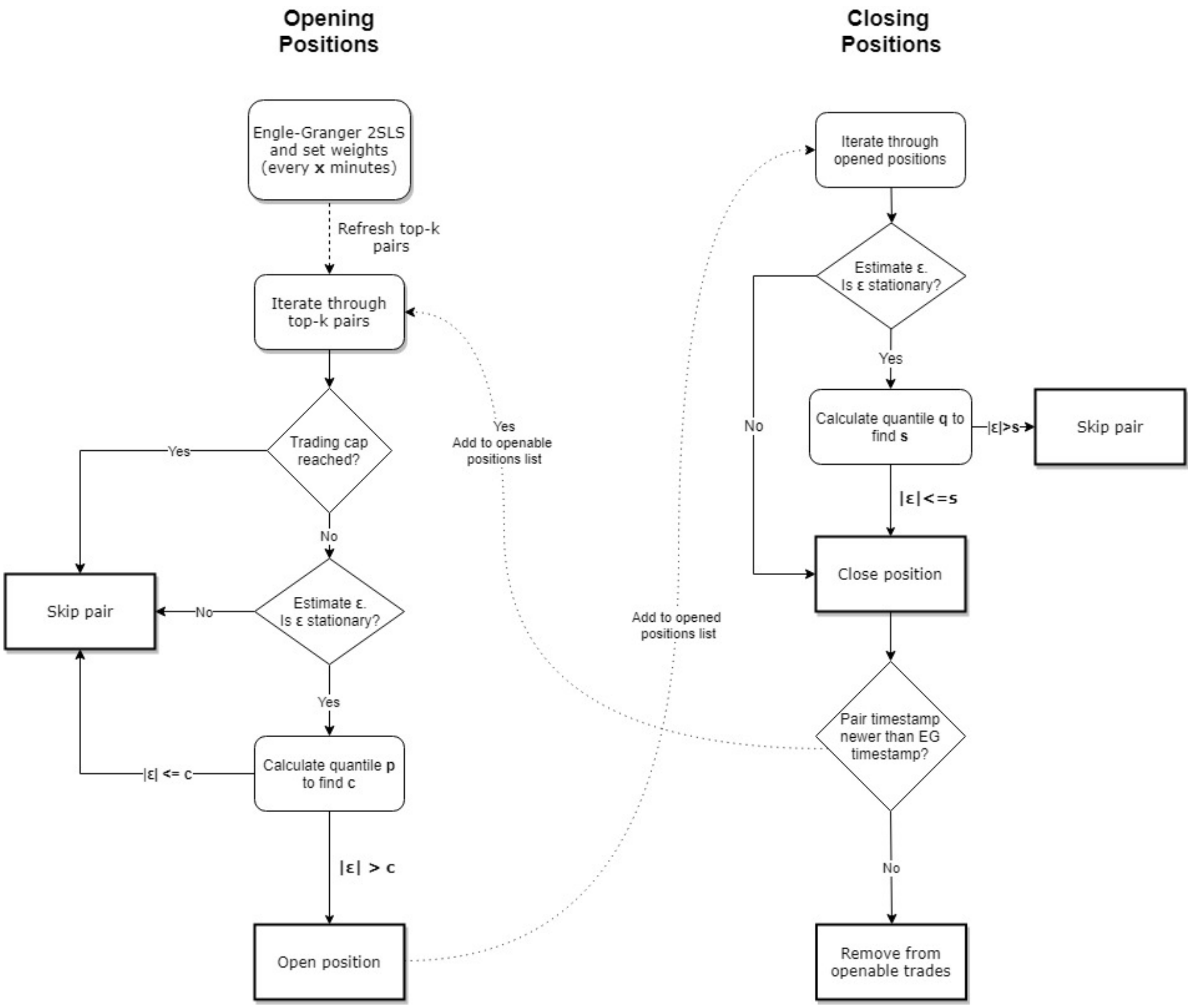


Figure 4: Flow chart of the trading decisions made by the multivariate trading algorithm. When opening a position, the bot iterates through the k cointegrating relations in order of lowest p-value of the ADF test of the Engle-Granger 2 step procedure (EG2SLS). The weights are either uniformly distributed or based on their corresponding p-value. Once a position is opened, the exchange rate is removed from the list of openable positions and added to the opened trades list.
When closing a position, the algorithm iterates through all opened trades. If $\hat{\epsilon}$ is non-stationary, the stop-loss condition kicks in, closing the position. Whenever a position is closed, the corresponding timestamp is compared to the most recent refresh of the top-k cointegrating relations list to see if the exchange rate is still assumed to be amongst this group. This is done in the case that there is still a trade open when the mentioned list is refreshed.

4.4 Evaluation

This subsection explains how we assess the performance of both pairs-trading methods. Deviations from the long-term mean are estimated as a moving window of h seconds. In the training set the model parameters are optimized based on the number of trades that are made and their final returns. The methods are afterwards examined based on their weekly performance over the full sample, in which no parameters are changed. Furthermore, each week within the data is separately regarded in terms of performance. The relative performance in each week, between the 2 trading methods, can be used to regard the robustness of our strategy.

The theoretical analysis of the cointegration-based pairs-trading strategy will be done as a market neutral portfolio, thus when a short trade is opened in Y, we use these funds to open a long in X. Short contracts are valued at a proportion of 1000 USDT at open, therefore profits are not reinvested over time. For the remainder of this paper, we use $ to denote USDT. We consider two ways of dividing the 1000$ capital when opening a position, namely a uniform distribution or a weighted distribution based on the ADF p-values. To elaborate, each time the Engle-Granger 2-step procedure (EG2SLS) is performed the weights for all k cointegrating relations are estimated based on uniform or ADF p-value distribution. This means that the 1000$ is redivided together with each reevaluation of the top-k cointegrating relations. Thus, in the uniform case, if $k = 5$ and we have just estimated the top-k cointegrating relations, we allocate a weight of 200$ per exchange rate even though it is uncertain that the full 1000$ (all k positions opened) will be invested. We expect that investing uniformly will mitigate more losses than when basing capital on the ADF p-values, due to diversification of the assets (especially for large k).

In case the exchange server is down, we close any outstanding trades and wait h seconds after the server is back online before we resume investing. Furthermore, highly illiquid exchange rates might have zero order books in an observation, in this case we do not trade this time frame. However, it is unlikely that such an illiquid exchange rate will be included in the top cointegrating relations. Lastly, we assume a transaction cost of $0,075\%$ per trade of the total amount traded of the purchased currency, which is the current transaction cost at Binance.

5 Results

The results are divided into a grid search, in which the optimal parameters p, q, h, k (opening percentile, closing percentile, window size and level of diversification, respectively) are estimated and evaluated, and a multivariate part where the performance of the Engle-Granger 2-step cointegration pairs-trading strategy is compared to a Johansen-based strategy.

The grid search is performed using the EG2SLS method, thus we do not regard the Johansen method. This is because the VECM takes significantly longer to estimate over all grid points. As a solution, we choose to generalize the optimal grid search parameters of the EG2SLS to the cointegration model including short-term trends. Furthermore, it is our belief that the general sensitivity of each regarded variable will not be affected by the in/exclusion of short-term trends. This is because our implementation of the Johansen method will only allow a subset of the original trades to open, compared to the base case. Since the grid search tests parameters ceteris paribus, we expect to gain sufficient insight on the underlying sensitivity of each variable.

A simple analysis was performed on the training set to determine a level of aggregation for which the data contains enough information, while not losing too many trading opportunities. To clarify,

As mentioned in section 3, certain exchange rates with a relatively low trading volume (in comparison to the liquidity of Bitcoin for example) will not be updated as frequently as once per second. Together with the notion of heavy computational intensity of our analysis, we use 60 second sample data. This will be done by resampling the observations and selecting the last value of every 60-second period, because this is the true price which would hold at the moment of a trade (as opposed to using the mean or first value of each 60 seconds).

5.1 Grid search

In this section we evaluate the grid search results to fix parameters p, q, h, k (opening percentile, closing percentile, window size and level of diversification) and the best method type to use for the rest of this paper. A grid search is performed to obtain insight on variable sensitivity, while limiting data mining by restricting the grid sizes. We perform the grid search on a small sample of weeks 20 until 22 (2018-05-19 12:00:00 until 2018-06-03 23:59:00) of intraday minute data. The results are evaluated by reviewing not only the returns of each grid, but also the sensitivity of the considered parameters.

The 144 iterations of the grid search can be found in table 13 of the appendix. The results show that the worst performing window size h is the smallest of 8 hours, apart from the top two results. By observing the fail rate, and overall performance, of the 6-hour window (360 seconds * 60 seconds) runs, it quickly becomes clear that such a short window produces more extreme results than those with larger windows. This is due to the fact that small windows cause $\hat{\epsilon}$ to be more sensitive to shocks. Despite using a percentile to define trading thresholds c and s, a higher sensitivity to shocks means that the volatility will increase. An increased volatility results in a more frequent breach of the trading threshold, increasing the probability of opening/closing a position. This effect can be seen in number of opened positions with respect to the window size, which can also be a result of our stop-loss condition based on the non-stationarity of $\hat{\epsilon}$. In comparison with the other two window sizes, the ADF test has much fewer degrees of freedom in the 6-hour window. The returns are strongly dependent of transaction costs, given that opening and closing a position results in 4 times the transaction fees. For example, if the average investment size of a position would be 500\$, then opening a position would cost $500\$ * 0.075\% * 2 = 0.75\$$ (long trade and a short trade), therefore the costs of a position are 1.50\$ and with 1000 positions opened this means that we require at least 1500\$ worth of profits just to break even.

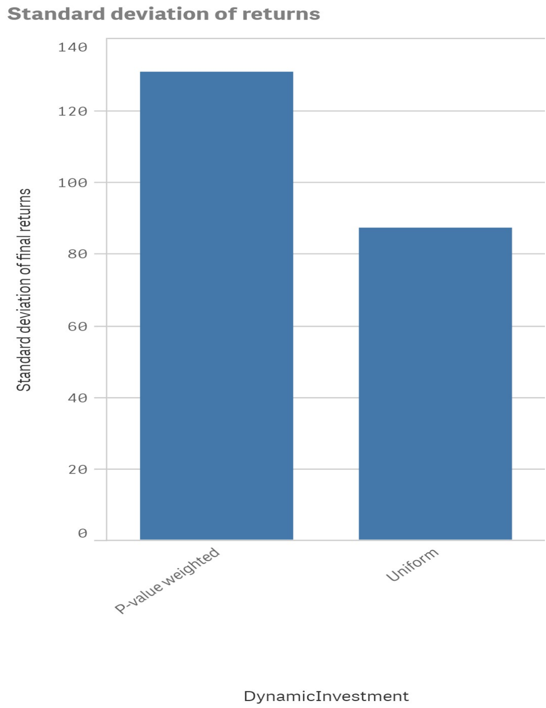


Figure 5: In this bar chart we plot the standard deviations of the final grid search returns per investment type (p-value weighted or uniform). The data of this plot can be found in table 13.

Another interesting detail is we see a clear increase in the number of positions when the threshold window (difference between p and q) shrinks. This is what we would expect, because a lower p means that positions require a smaller deviation from the long-term equilibrium to be opened. Similarly, a higher q leads to a decreased trading time, leaving more room for additional trades. There are, however, some cases with few trades which can be seen from the low investment averages. This results from both the asymmetric investments based on ADF test p-values and windows where percentile p is not reached, leading to that there are relatively more moments in time where few/no trade is opened. Note that because the returns are relative to the number of opened positions, we calculate the average investment over time to compute returns. This is because the size of a single position depends on diversification parameter k. Therefore it is also important to keep in mind the absolute returns from each iteration. There is a reasonable consistency in the relation between the percentage of failed trades and the final returns of each iteration.

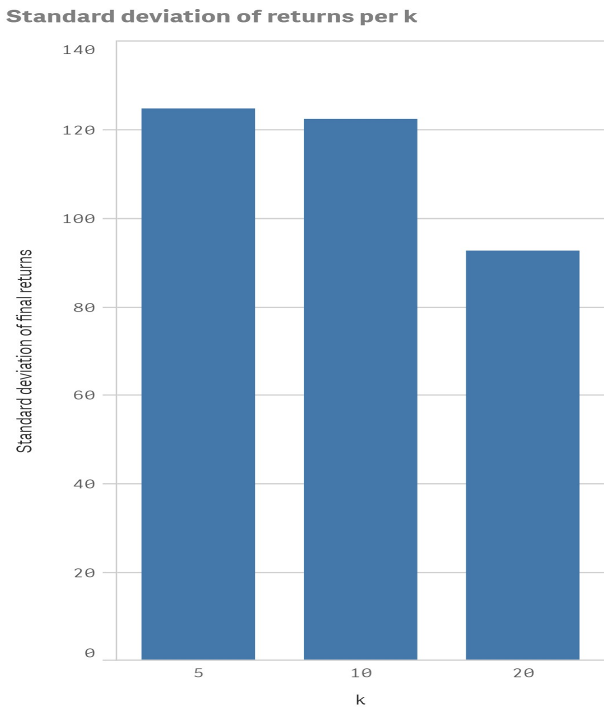


Figure 6: In this bar chart we plot the standard deviations of the final grid search returns per value of k.

In table 13 of the appendix we see that the ADF test p-value weighted investment method underperforms compared to a uniform distribution of capital. The more extreme values for best and worst trade in the ADF p-value weighted case backs the notion of loss in diversification, in contrast to the uniform case. This difference in variance is represented in figure 5. The results suggest that the first ranked cointegrating relation (by p-value) does not necessarily outperform the top-k pairs. The effect of k on the standard deviations of the final returns is shown in figure 6, where we see a negative correlation between k and the variance of returns. Furthermore, we do not see a significant difference in the fail rates of trades for both investment types. For the reason of diversification, we will only work with a uniform distribution of capital when opening positions from now on. In terms of robustness, it is preferred to have decreased risk as the cost of smaller returns at the close of a position.

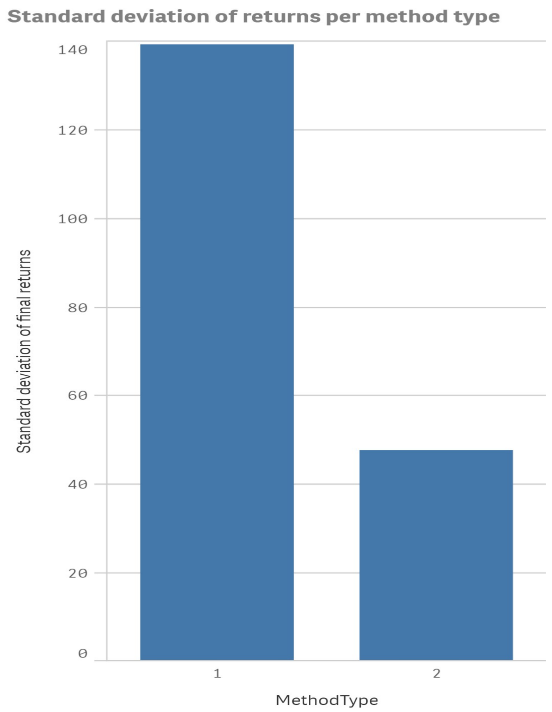


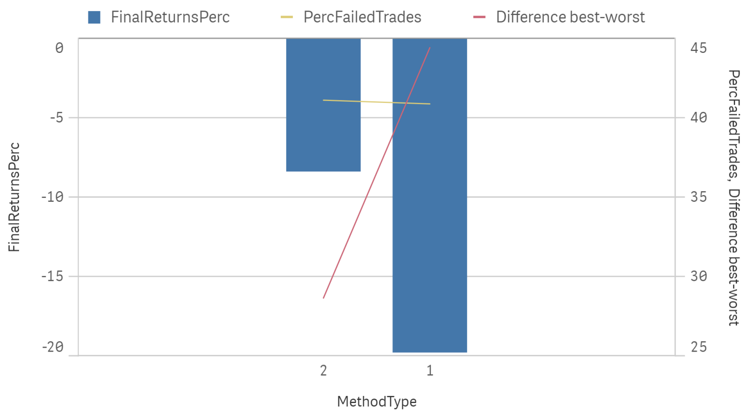
Figure 7: In this bar chart we plot the standard deviations of the final grid search returns per method.

The standard deviations of the final returns per method are plotted in figure 7, where method 2 is shown to be more robust. From the top 10 grid iteration results, we see that a window size h of 1 day and a uniform distribution of investment capital produce the best results. Interestingly, there are many poorly performing iterations with a 2-day window; even though we hypothesized that an optimal window size would be 1 day, because it, allegedly, has useful daily cyclical economic data. Therefore, a similar idea is expected to hold for a 2 day period. We reach the opposite conclusion in the grid results for a number of possible reasons. First of all, the sample of the entire grid search is quite small, especially in terms of the amount of information in the observations. Secondly, as more historical data is inserted into the estimation window, we increase the probability of mistaking a general economic trend with distinct co-movements between exchange rates, which is less desirable in high-frequency trading based on cointegration. To elaborate, when the window size becomes too large, a deviation from the general trend tends to be registered as a 'cointegrating relation' in disequilibrium, where the mean is the economic trend.

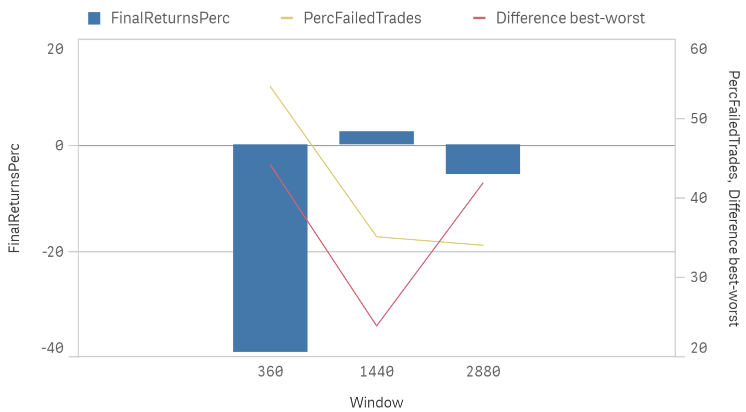
In figure 8 we plot the distinct method types and window sizes against the average returns, percentage of failed trades and differences between the best and worst trades. A large difference between the best and worst trade indicates a larger variance in the returns.

The relative performance of each method is visualized in figure 8a. The graph makes clear that both methods perform similarly in terms of fail rates, however method 2 has a much smaller difference between the best and worst trades, on average. The latter indicates that the second method has a lower variance in the returns, by settling for smaller increments (deviation difference between position open and close), resulting in overall better performance. The omission of the more extreme returns, both positive and negative, ensure a more robust trading method. This can also be seen in the different results for investment weights, namely in the p-value weighted iterations we see relatively much larger best and worst trades to the uniform case.

The optimal values for the variables used in the weekly historical simulation are based on the best performing iteration of the grid search. The results also show that method 2 clearly outperforms method 1 in terms of returns. Based on the arguments made in this section, we choose to set the optimal parameters equal to the iteration in bold from table 13. The subsequent analysis in section 5.2, will be performed using method 2 with



(a) This graph shows the average performance of each method type over all grid search iterations. The bars represent the average final returns and the yellow/red lines (secondary y-axis) constitute the percentage of failed trades and the difference between the best and worst trades, respectively.



(b) This graph shows the average performance of each regarded window size over all grid search iterations. The bars represent the average final returns and the yellow/red lines (secondary y-axis) constitute the percentage of failed trades and the difference between the best and worst trades, respectively.

Figure 8: Two figures on the sensitivity of methods and window sizes on the average returns.

parameters $h = 1440$, $k = 10$, $p = 99$, $q = 53$ and a uniform investment method.

5.2 Historical simulation

In this section we look at the weekly performance of method 2 using both the EG2SLS and Johansen methods. All analyses in this section are done using the optimized variables mentioned in the subsection 5.1. Table 4 contains the test results of the initial requirement for cointegration, which is non-stationarity. The results of the first step in the EG2SLS method show that each exchange rate is non-stationary, however to be sure we perform the ADF-test each window. In the case of a window of one day (used throughout this subsection), not a single exchange rate was found to be stationary over time [4].

Table 4 The augmented Dickey-Fuller test results where the null hypothesis of non-stationarity is failed to be rejected for all 17 cryptocurrency exchange rates. The value 0 in the second row represents that we fail to reject the null hypothesis. This is the first necessary check of the currency exchange rates, confirming that they are $I(1)$. In addition of this snapshot of the data, this test is performed every window of the analysis on $\hat{\epsilon}$ to confirm stationarity. All exchange rates are in quote currency USDT.

ADF test	XLM	VEN	QTUM	EOS	BCH	BTC	LTC	NEO	ONT	TRON	ETC	BNB	ADA	ETH	ICX	XRP	IOTA
Null Hypothesis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pval (MacKinnon (1991))	0.63	0.66	0.29	0.54	0.32	0.53	0.48	0.53	0.81	0.19	0.24	0.25	0.34	0.17	0.41	0.24	0.69
Critical Value	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41	-3.41

The analysis is run using method 2 with parameters $h = 1440$, $k = 10$, $p = 99$, $q = 53$ and a uniform investment method on 60 second intraday data between '2018-05-19 12:00:00' and '2018-09-03 23:59:00' with and without the inclusion of short-term trends in the pairs-selection process.

Both the Johansen and the EG2SLS-based pairs-trading strategies result in positive returns over the sample using the optimized parameters. The Johansen approach has a better performance overall, indicating that there is significant information in the short-term trends of cryptocurrency exchange rates in terms of cointegrating relations.

[4] Apart from VEN-USDT, due to its ticker change, as mentioned in section 3

The weekly results of the EG2SLS-based pairs-trading strategy can be found in table 5. From the table we see that, on average, the algorithm made around 27.73 USDT per week based on an average investment of 465.55 USDT (5.97% in 16 weeks). The worst performing weeks are 22 and 31, however, in figure 9a we see that week 31 had mostly positive returns. The graph also shows the negative increments at the close of the week, which is a result of closing all opened positions at that time. The summary statistics of the portfolios plotted in figure 9a can be found in table 11.

Table 5 This table contains the weekly pairs-trading strategy performance results of the EG2SLS method. The analysis was performed on aggregated data of 1 minute. NoPos stands for the total number of opened positions. RpP and RpD stand for returns per position and returns per day, respectively. % Failed is the percentage of positions closed with non-positive returns. AI is the average USDT capital invested in the opened positions over time. BestTrade, WorstTrade and FinalIncome are all in USDT

Week	Method	Window	k	Horizon	p	q	NoPos	NoLongs	% Failed	AI	BestTrade	WorstTrade	FinalIncome	RpP	RpD	FinalReturns
20	2	1440	10	2160	99	53	19	8	42.11	414.20	0.94	-4.23	-3.34	-0.18	-0.40	-0.81
21	2	1440	10	10080	99	53	158	96	29.11	427.83	3.42	-4.40	36.40	0.23	1.22	8.51
22	2	1440	10	10080	99	53	86	52	47.67	386.20	2.20	-6.10	-35.20	-0.41	-1.30	-9.11
23	2	1440	10	10080	99	53	200	113	32.00	471.08	4.15	-6.32	35.32	0.18	1.07	7.50
24	2	1440	10	10080	99	53	171	89	26.90	479.64	3.28	-6.85	47.04	0.28	1.40	9.81
25	2	1440	10	10080	99	53	205	113	35.12	528.41	3.85	-9.57	-1.68	-0.01	-0.05	-0.32
26	2	1440	10	10080	99	53	204	116	39.22	527.94	2.28	-9.40	1.27	0.01	0.03	0.24
27	2	1440	10	10080	99	53	102	62	25.49	371.26	4.50	-12.94	-3.70	-0.04	-0.14	-1.00
28	2	1440	10	10080	99	53	182	88	36.81	453.29	3.62	-5.48	48.88	0.27	1.54	10.78
29	2	1440	10	10080	99	53	166	82	30.12	494.57	4.09	-3.80	60.03	0.36	1.73	12.14
30	2	1440	10	10080	99	53	155	80	31.61	554.97	3.81	-10.99	-4.21	-0.03	-0.11	-0.76
31	2	1440	10	10080	99	53	132	68	40.15	473.07	6.56	-10.98	-19.66	-0.15	-0.59	-4.16
32	2	1440	10	10080	99	53	152	85	29.61	514.26	15.07	-6.73	72.40	0.48	2.01	14.08
33	2	1440	10	10080	99	53	176	79	24.43	416.38	5.89	-52.49	111.50	0.63	3.83	26.78
34	2	1440	10	10080	99	53	179	95	21.23	448.12	5.76	-15.02	88.41	0.49	2.82	19.73
35	2	1440	10	10080	99	53	155	69	21.94	487.51	2.95	-14.54	10.21	0.07	0.30	2.09
Avg	**2**	**1440**	**10**	**9585**	**99**	**53**	**152.63**	**80.94**	**32.10**	**465.55**	**4.52**	**-11.24**	**27.73**	**0.14**	**0.83**	**5.97**

With the addition of short term trends when estimating the existence of cointegration, the Johansen method proves to have 'hedging' potential. The weekly performance of the market neutral pairs-trading strategy with the inclusion of short term trends by means of the Johansen method can be found in tables 6 and 12. The first table contains the weekly performance and the latter contains summary statistics of the USDT portfolio value over time, which are also plotted in figure 9b. The summary statistics show a decreased standard deviation in all but one week and the upper/lower bounds are greatly reduced in variance and size. Together with the observation that the Johansen-based pairs-trading algorithm has higher returns by only performing half of the original trades, the results indicate that the inclusion of short-term trends greatly increase robustness of the algorithm's performance.

Table 6 This table contains the weekly pairs-trading strategy performance results with the Johansen method. The pairs are primarily selected with EG2SLS and afterwards reviewed by the Johansen method to obtain a subset of the original trades. The analysis was performed on aggregated data of 1 minute and with a uniform investment type. NoPos stands for the total number of opened positions. RpP and RpD stand for returns per position and returns per day, respectively. % Failed is the percentage of positions closed with non-positive returns. AI is the average USDT capital invested in the opened positions over time. BestTrade, WorstTrade and FinalIncome are all in USDT.

Week	Method	Window	k	Horizon	p	q	NoPos	NoLongs	% Failed	AI	BestTrade	WorstTrade	FinalIncome	RpP	RpD	FinalReturns
20	2	1440	10	2160	99	53	5	1	20.00	94.26	0.83	-0.56	1.03	0.21	0.55	1.09
21	2	1440	10	10080	99	53	65	47	29.23	222.03	3.42	-4.12	14.22	0.22	0.91	6.40
22	2	1440	10	10080	99	53	42	20	45.24	160.97	5.38	-4.27	-1.68	-0.04	-0.15	-1.05
23	2	1440	10	10080	99	53	106	53	29.25	257.06	5.27	-7.10	27.20	0.26	1.51	10.58
24	2	1440	10	10080	99	53	75	39	24.00	209.97	5.10	-5.16	37.79	0.50	2.57	18.00
25	2	1440	10	10080	99	53	116	61	33.62	263.45	3.85	-9.92	1.01	0.01	0.05	0.38
26	2	1440	10	10080	99	53	103	61	36.89	350.94	2.75	-4.67	-6.80	-0.07	-0.28	-1.94
27	2	1440	10	10080	99	53	53	28	22.64	189.84	3.28	-7.97	15.05	0.28	1.13	7.93
28	2	1440	10	10080	99	53	105	43	36.19	211.02	2.99	-5.48	23.29	0.22	1.58	11.04
29	2	1440	10	10080	99	53	99	55	28.28	278.72	4.43	-2.75	32.97	0.33	1.69	11.83
30	2	1440	10	10080	99	53	98	56	27.55	317.43	4.37	-5.37	26.27	0.27	1.18	8.28
31	2	1440	10	10080	99	53	46	28	39.13	134.72	4.66	-10.93	-9.87	-0.21	-1.05	-7.33
32	2	1440	10	10080	99	53	78	42	33.33	254.76	15.07	-4.99	30.47	0.39	1.71	11.96
33	2	1440	10	10080	99	53	94	44	28.72	263.94	5.89	-49.70	37.03	0.39	2.00	14.03
34	2	1440	10	10080	99	53	82	45	24.39	232.75	5.76	-17.61	40.87	0.50	2.51	17.56
35	2	1440	10	10080	99	53	71	34	19.72	259.89	3.09	-12.45	0.54	0.01	0.03	0.21
Avg	**2**	**1440**	**10**	**9585**	**99**	**53**	**77.38**	**41.06**	**29.89**	**231.36**	**4.76**	**-9.57**	**16.84**	**0.20**	**1.00**	**6.81**

The results in table 7 show the top- and bottom-10 cointegrating relations between the regarded USDT exchange rates. These are cointegrating relations which triggered the bot to trade in the weeks 20 until 36. Given our market neutral strategy, each cointegrating relation symbolizes two trades. Amongst the top cointegrating relations (and top traded exchange rates in table 8) we mainly see the presence of low-volume exchange rates. Furthermore, the dominant cointegrating relations over time consist of mainly lagging exchange rates, which is discussed in section 5.3.

Table 7 Top- and bottom-5 proportion of opened positions of cointegrating relations by means of the EG2SLS and the Johansen methods. The first row of the Johansen column means that QTUM-USDT and NEO-USDT where cointegrated and jointly traded 2% of the time in the weeks 20 until 36.

EG2SLS			**Johansen**		
Exchange rate 1	**Exchange rate 2**	**Perc**	**Exchange rate 1**	**Exchange rate 2**	**Perc**
BCH-USDT	NEO-USDT	1,84%	QTUM-USDT	NEO-USDT	2,16%
QTUM-USDT	NEO-USDT	1,84%	NEO-USDT	ADA-USDT	1,81%
QTUM-USDT	ADA-USDT	1,84%	QTUM-USDT	ADA-USDT	1,81%
BTC-USDT	ADA-USDT	1,76%	NEO-USDT	ONT-USDT	1,63%
NEO-USDT	ADA-USDT	1,64%	NEO-USDT	ICX-USDT	1,52%
⋮	⋮	⋮	⋮	⋮	⋮
VEN-USDT	IOTA-USDT	0,08%	VEN-USDT	BNB-USDT	0,12%
VEN-USDT	NEO-USDT	0,08%	VEN-USDT	QTUM-USDT	0,06%
BCH-USDT	TRON-USDT	0,04%	BCH-USDT	TRON-USDT	0,06%
VEN-USDT	QTUM-USDT	0,04%	VEN-USDT	ICX-USDT	0,06%
VEN-USDT	ICX-USDT	0,04%	VEN-USDT	IOTA-USDT	0,06%

QTUM-USDT, NEO-USDT and ADA-USDT are part of many cointegrating relations in the sample of weeks 20 until 36. This can also be seen in the number of individual trades of each exchange rate, which is shown in table 8. The number of trades in VEN-USDT is by far the lowest, however this is most likely also a result of the ticker change mentioned in section 3 which gives VEN-USDT less trading opportunities. The difference between the Johansen and EG2SLS methods is minimal, in terms of individually traded exchange rates. This is expected, given that the Johansen method depicts a subset of the original EG2SLS trades. A significant difference in traded exchange rates between the two could be a result of aberrant behavior in short-term trends of certain exchange rates.

In the 16 weeks of minute-data there are 130.320 trading moments (excluding the first window each week). The simulations from tables 5 and 6 had 4884 ($2442 * 2$) and 2476

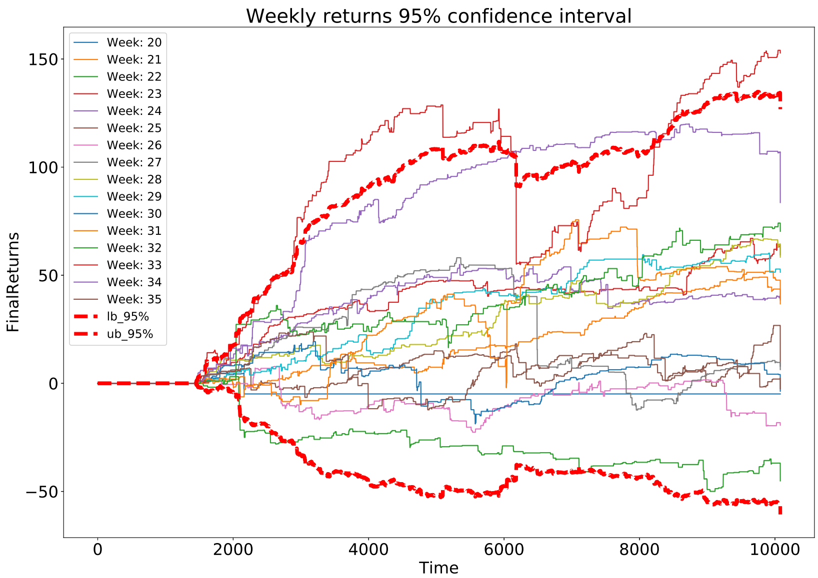
Table 8 Proportion of trades per exchange rate between weeks 20 and 36 as a result of the EG2SLS/Johansen cointegrating relations.

EG2SLS		**Johansen**	
Exchange rate	**Perc**	**Exchange rate**	**Perc**
QTUM-USDT	9,10%	QTUM-USDT	9,10%
ADA-USDT	9,10%	ADA-USDT	8,80%
NEO-USDT	8,10%	NEO-USDT	8,10%
BTC-USDT	7,90%	XRP-USDT	7,30%
XRP-USDT	7,20%	ONT-USDT	7,00%
BCH-USDT	7,10%	BTC-USDT	6,60%
ONT-USDT	6,40%	BCH-USDT	6,40%
ICX-USDT	6,00%	ICX-USDT	6,40%
IOTA-USDT	5,90%	IOTA-USDT	6,00%
LTC-USDT	5,50%	ETH-USDT	5,30%
EOS-USDT	5,30%	EOS-USDT	5,30%
ETH-USDT	5,30%	LTC-USDT	5,20%
XLM-USDT	4,60%	XLM-USDT	5,00%
TRON-USDT	4,00%	BNB-USDT	4,40%
BNB-USDT	3,90%	TRON-USDT	4,10%
ETC-USDT	3,40%	ETC-USDT	3,60%
VEN-USDT	1,20%	VEN-USDT	1,30%

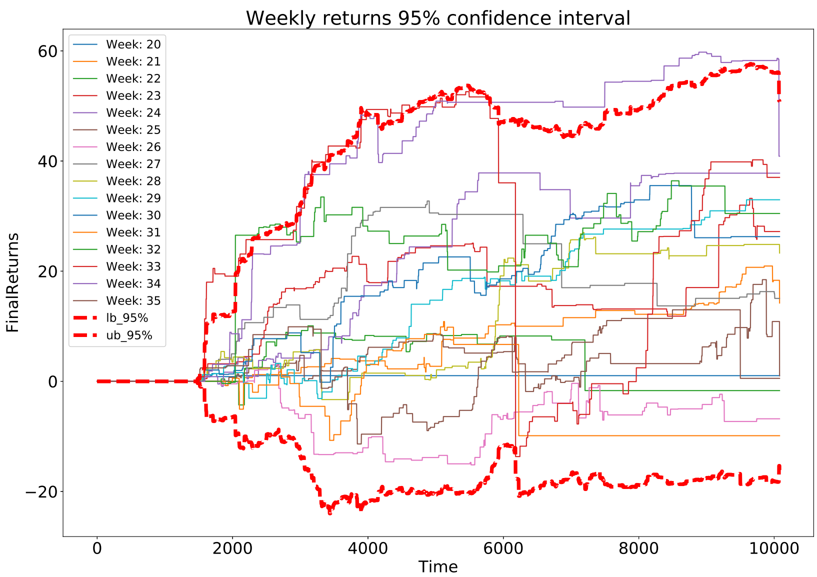
$(1238 * 2)$ trades in total, respectively. This means that the EG2SLS / Johansen methods traded 3.75% / 1.90% of the time and took on average 26.68 / 52.63 minutes per trade, respectively.

To understand the observed strong outliers in the weekly returns, we plot the prices during those weeks. These prices can be seen in figures 10. Graph 10a reveals a strong increase in the market on the 29th of May and graph 10b shows a similar upwards shock on the 30th of June, such shocks can result in large negative returns if a short position had already been opened by the trading algorithm before that time. This was the case in week 22 (figure 10a) with a BTC-USDT/LTC-USDT position, where our short position in LTC-USDT suffered a great loss due to its increase in price. In addition, graph 10c of the prices in week 32 shows diverging exchange rates from the general equilibrium which can result in heavy losses when closed prematurely, such as at the end of the week. The largest outlier occurred in week 33, where despite a position was closed with a loss of 52.49$, it closed with the highest returns of all weeks.

The graphs also show that there is a server error around the 26th of June. In such a case, the bot closes all opened positions at the time of the error, which can result in large amounts of missed profits because we don't await mean-reversion. Note that table 5 contains the best/worst trade which are the max/min increments (including transaction costs) of table 11, where the latter contains the returns in USDT over time. The average investment over all weeks, keeping in mind that $k = 10$ and a single position is opened for $1000/10 = 100$ USDT, shows us that on average there were 4 to 5 trades open at any given moment in time. In closing, the market neutral pairs-trading strategy without inclusion of short-term trends results in 6 out of 16 weeks with non-positive returns. A significant notice is that by running the analysis on each week independently, we add unnecessary risk at the end when all trades must be closed. Notwithstanding that the average worst closed position is nearly twice as bad as the average best position, the returns at the end of the simulation are positive.



(a) Weekly performance of the pairs-trading strategy with EG2SLS pairs-selection. The dotted red lines symbolize the 95% confidence interval.

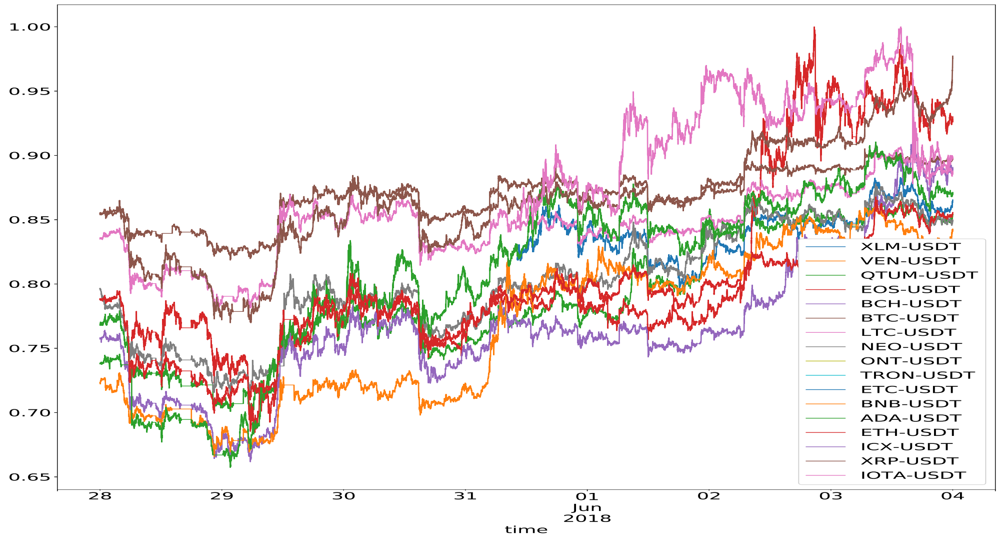


(b) Weekly performance of the pairs-trading strategy using the Johansen method. The dotted red lines symbolize the 95% confidence interval. Note the ticksize of the y-axis in comparison to graph (a).

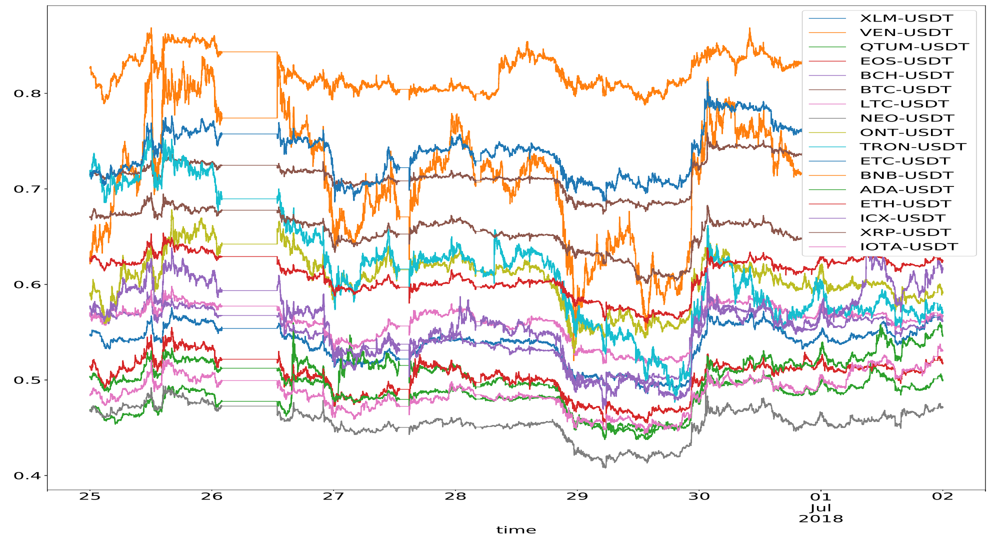
Figure 9: Two plots of the weekly performance of section 5.2, including a 95% confidence interval.

Additionally, we have performed the pairs-trading strategy with the EG2SLS method on all weeks with a decreased opening threshold percentile p in tables 14 and 15 of the appendix from which we see the effect in the increased number of trades. Albeit the average returns are higher, the summary statistics show that, as a result of lowered value for p, the variance increases significantly.

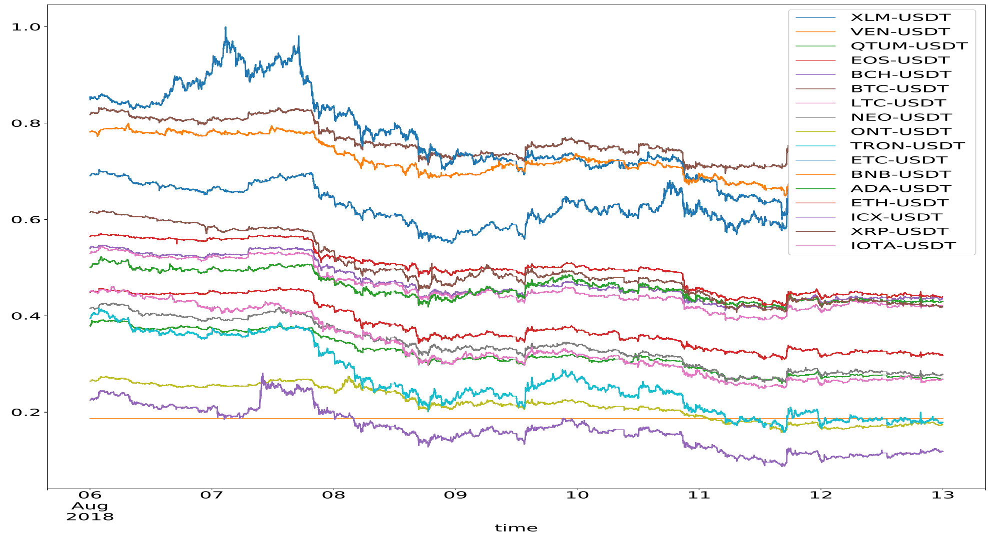
The plots of the weekly portfolio in figures 9 have the same general trends each week, because the Johansen-assisted trades are a subset of the Engle-Granger 2-step trades. The results show that the rank of Π is a more robust estimator for the existence of cointegration than EG2SLS. Despite the fact that not all extreme trades have been deselected in the Johansen case, such as the large negative return in week 33, the VECM-assisted case has only 3 of 16 weeks with negative returns. The average investment size of 231.36 elucidate that, on average, there are between 2 and 3 trades open at any point in time. This corresponds with the number of trades relative to table 5. The returns per trade is large in both cases, meaning that the strategy would still perform adequately on other markets/exchanges with larger transaction costs.



(a) These are the prices of the Binance cryptocurrencies in week 22 of 2018. There is a large shock (≈18% increase) visible on the 29th of May. Shocks in the market similar to these can have an adverse effect on the pairs-trading algorithm.



(b) These are the prices of the Binance cryptocurrencies in week 26 of 2018. Aside from the server error on the 26th, we see a 26% increase of the prices on the night of the 29th of June. In addition, the Stellar (XLM) exchange rate XLM-USDT increased by 83% during that day.



(c) These are the prices of the Binance cryptocurrencies in week 32 of 2018. We see some strong shocks as well as divergence of XLM-USDT near the end of the week. This could cause larger returns/losses at the closing of the week.

Figure 10: Three figures zooming in on the outliers in the results of section 5.2. The three outlined weeks are the movements of prices in the Binance exchange in a bullish, neutral and bearish market.

5.3 Granger causality

By means of Granger causality testing we can estimate the direction(s) of co-movements between two exchange rates. Once the causal relations in the cryptocurrency market have been identified, revealing the leading and lagging indicators, we can compare these results to those of section 5.2. To visualize these relations, we plot these directions for all available exchange rates, using a Python module called Graphviz [5]. This visualization summarizes Granger causality between our sample exchange rates over time by estimating the amount of time a exchange rate significantly Granger causes another exchange rate.

[5]https://www.graphviz.org

The Granger causality was tested using equation 7 as a moving window of 1 day by means of the Python Statsmodels [6] module. The results where afterwards aggregated over the number of weeks mentioned in each table and graph to obtain the percentage of time the causation was significant. The analyses in tables 9 and 16 are done with significance level $\alpha = 5\%$. Note that the figures only show edges with a weight above 70% to reduce complexity of the graphs, the details can all be found in tables 9 and 16. For each element (i, j) in the aforementioned tables, where $i = j$, we have denoted the causality as 0. Additionally, we omitted VEN-USDT in the Granger causality tests because there were too many missing values as can be seen in figure 1b, which would bias the aggregate results over all weeks.

Table 9 This table contains the aggregate Granger causality relations of each cointegrating relation between week 20 and 36 of 2018 with a level of significance of 95%. Causality is shown from row to column, meaning that Bitcoin (BTC) caused Cardano (ADA) 91.1% of the time, whereas Cardano caused Bitcoin in 25.7% of the time.

	ADA	BCH	BNB	BTC	EOS	ETC	ETH	ICX	IOTA	LTC	NEO	ONT	QTUM	TRON	XLM	XRP
XLM	0.156	0.268	0.300	0.213	0.097	0.242	0.237	0.167	0.075	0.242	0.157	0.139	0.222	0.212	0.000	0.318
QTUM	0.232	0.254	0.186	0.200	0.102	0.256	0.148	0.263	0.182	0.225	0.207	0.176	0.000	0.308	0.296	0.231
EOS	**0.894**	**0.805**	0.727	0.792	0.000	**0.839**	**0.914**	**0.805**	**0.889**	**0.882**	**0.860**	**0.927**	0.673	**0.971**	0.677	**0.905**
BCH	0.593	0.000	0.476	0.372	0.366	0.389	0.529	0.300	0.349	0.652	0.627	0.450	0.492	0.500	0.341	0.620
BTC	**0.911**	0.744	0.537	0.000	0.458	0.659	0.517	**0.867**	**0.804**	0.753	0.787	**0.837**	0.788	**0.839**	**0.809**	0.760
LTC	0.284	0.239	0.185	0.210	0.265	0.143	0.191	0.281	0.404	0.000	0.364	0.250	0.268	0.158	0.212	0.327
NEO	0.370	0.269	0.303	0.267	0.256	0.200	0.048	0.500	0.388	0.291	0.000	0.420	0.256	0.465	0.216	0.325
ONT	0.304	0.175	0.273	0.306	0.122	0.323	0.312	0.275	0.250	0.286	0.380	0.000	0.255	0.282	0.389	0.259
TRON	0.423	0.214	0.105	0.226	0.086	0.259	0.310	0.256	0.406	0.316	0.163	0.256	0.410	0.000	0.364	0.250
ETC	0.209	0.361	0.250	0.122	0.129	0.000	0.333	0.444	0.160	0.286	0.400	0.226	0.205	0.370	0.303	0.366
BNB	0.283	0.238	0.000	0.293	0.273	0.300	0.200	0.414	0.391	0.222	0.455	0.333	0.395	0.526	0.300	0.417
ADA	0.000	0.333	0.283	0.257	0.258	0.326	0.250	0.384	0.463	0.388	0.340	0.333	0.295	0.423	0.359	0.344
ETH	0.703	**0.882**	0.667	0.603	0.286	0.333	0.000	0.514	**0.867**	0.787	**0.825**	0.719	0.656	0.759	0.632	**0.870**
ICX	0.164	0.225	0.172	0.133	0.098	0.185	0.143	0.000	0.194	0.188	0.224	0.157	0.211	0.231	0.133	0.308
XRP	0.311	0.268	0.222	0.187	0.143	0.293	0.217	0.308	0.368	0.231	0.273	0.204	0.321	0.333	0.409	0.000
IOTA	0.278	0.233	0.043	0.118	0.167	0.240	0.200	0.306	0.000	0.277	0.143	0.159	0.182	0.062	0.200	0.193

The average direction of Granger causality between the weeks 20 and 36 of 2018 are shown in figure 11. A more detailed overview is displayed in table 9. Interestingly, the entire market is led by EOS. This is a result of a temporary increase in popularity in EOS during that period, this is reflected in the market capitalization in June/July, which is nearly double of that in August/September (from 10 billion \$ to 5 billion \$).

[6] https://www.statsmodels.org

As expected, the market is mostly lead by Ethereum (ETH) and Bitcoin (BTC). Apart from market cap, Ethereum and Bitcoin have a strong influence on the market through sentiment. This is largely due to the fact that most exchanges only quote in fiat currencies or Bitcoin/Ethereum (and artificial USDT). In addition, many currencies are issued on the Ethereum blockchain using the Ethereum Request for Comment (ERC-20) standard, which in turn is also a driver for co-movements. Figure 12 clearly shows the influence of Bitcoin and Ethereum in the second 'row' of coins. The other two currencies in the same row are Bitcoin Cash (BCH) and Ethereum Classic (ETC), which are, respectively, hard forks of Bitcoin and Ethereum. A hard fork is when a cryptocurrency splits into two cryptocurrencies. Regardless of having the same parent cryptocurrency, both figures 11 and 12 show that BCH and ETC have little causality towards the rest of the market (given the number of outgoing edges). This is also a result of their market capitalization, which is merely a fraction of their parents Bitcoin and Ethereum.

We see that the Binance coin (BNB), on average, does not have any co-movements with the other cryptocurrencies in the dataset. This is not surprising, as the Binance coin is seen as a stable cryptocurrency, isolated from many cryptomarket shocks. To clarify, the Binance coin is specific for the Binance exchange, as the name suggests, and is used for paying trading fees, transaction costs and other payments on the exchange. Furthermore, the BNB coin offers a type of dividend to the holders of the coin, based on the profits made by the Binance exchange. The disconnection of BNB-USDT from the other exchange rates is also reflected in the number of trades shown in table 8. Additionally, figures 11, 12 show that BNB is the only currency exchange rate which has no significant incoming/outgoing edges, meaning that its price movement is too erratic as compared to the rest of the sample.

Tables 7 and 8, show that the lagging currency exchange rates are traded most often. The nodes in the lower 'level' of graph 11, with the exclusion of BTC-USDT, make up the exchange rates in the top-10 cointegrating relations. This indicates that the lagging exchange rates are more likely to be cointegrated, given that they jointly follow the leading exchange rate(s). The inclusion of BTC-USDT in the top cointegrating relations is most likely due to the strength of Bitcoin in the cryptocurrency market. Not only in terms of volume over all exchanges, but arguably also due to its robustness, the Bitcoin dominates the cryptocurrency market. We do, however, see in table 8 that the fraction of trades in Bitcoin in the Johansen case is around 1.3% lower (largest difference between the methods). This difference is most likely caused by the short-term trends in Bitcoin. This indicates that several trades were discarded by the Johansen method which would normally trigger under the EG2SLS. Given that BTC-USD Granger causes the other most traded exchange rates, a shock in the BTC-USD price will affect its short-term trends differently than its lagging exchange rates. If we also note our level of aggregation (minute-data), it is possible that these short-term trends can be found in the lagging exchange rates, but not in the BTC-USD trends.

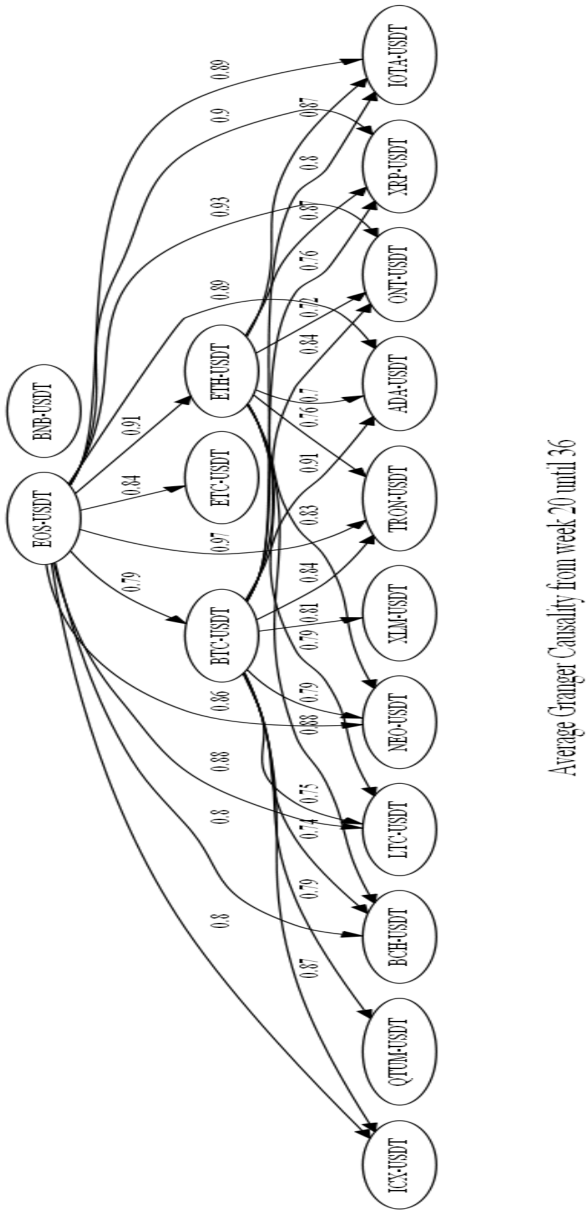


Figure 11: This is a visual representation of the data in table 9, which shows the average Granger causality between weeks 20 and 36 of 2018 with a level of significance of 95%. We only show edges with a weight larger than 70% to decrease complexity. The weights are based on the average causality over time (1 if X causes Y, otherwise 0) and rounded to two decimals. Thus on average, EOS-USDT significantly causes BTC-USDT in 79% of time between weeks 20 and 36, with a level of significance of 5%. All exchange rates are in quote currency USDT.

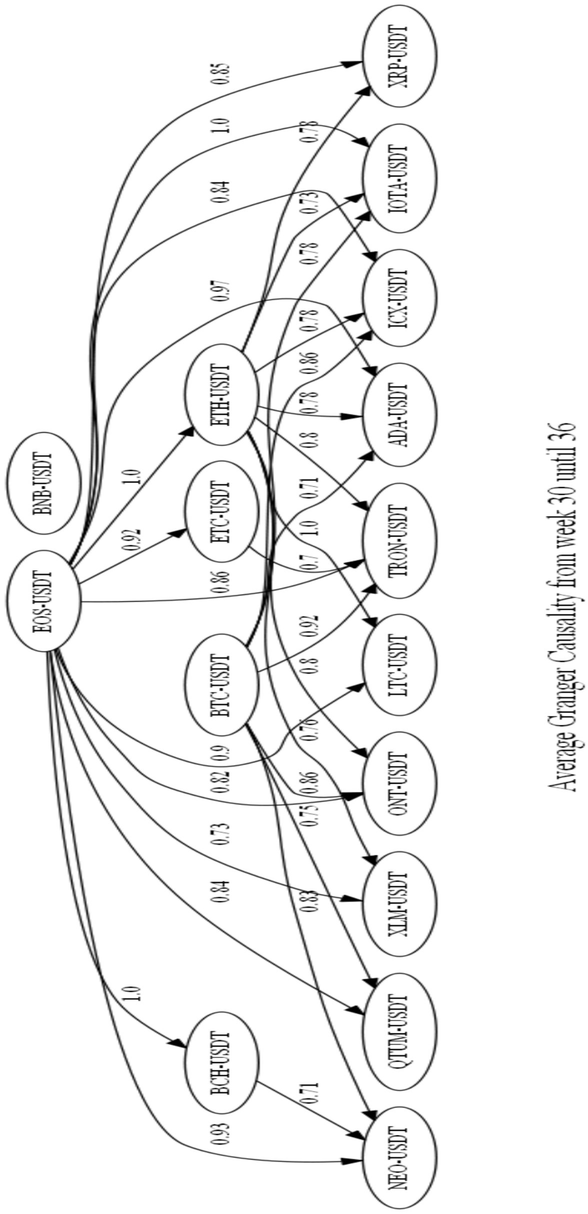


Figure 12: This is a visual representation of the data in table 16 found in the appendix, which shows the average Granger causality between weeks 30 and 36 of 2018 with a level of significance of 95%. We only show edges with a weight larger than 70% to decrease complexity. Again, all numbers are rounded to two decimals.

6 Discussions

In this section we discuss certain shortcomings of the analysis, especially with respect to a practical implementation. In addition, we consider the robustness of the results and run several Granger causality tests to better understand the relations between the exchange rates on Binance.

6.1 Limitations

Aside from the overall positive returns for both analyses, there are certain shortcomings which must be addressed. Firstly, it is currently not possible to short trade on such a variety of cryptocurrency exchange rates. Consequently, making a 'market neutral'-type (relative to the market) trading strategy in practice could be done by creating a division of funds over all regarded cryptocurrency exchange rates and increasing/decreasing their corresponding investment when going long/short, with the additional risk of holding the coins. For example, if we regard 10 exchange rates and a portfolio value of 1000 USD, we invest 100 USD in each cryptocurrency. If we need to long/short an exchange rate, we buy/sell the 100 USD worth of that currency. Alternatively, it is possible to invest in solely one direction of the deviation from the equilibrium, in which case we would not have a market neutral portfolio.

Secondly, in the simulation we assume that all limit orders can be fully realized, however, in practice it is possible to receive partially completed orders. Nevertheless, given that transaction costs on Binance are percentual, this should not have any severe consequences on the strategy performance. Note that if market orders would be used, we would additionally have to cross the spread.

Lastly, the sample which is used (from '2018-05-19 12:00:00' until '2018-09-03 23:59:00') reflects only a small window of the cryptocurrency market behavior, albeit containing many observations. The core issue is that there is little informative data available of this

Table 10 Table containing summarizing sensitivity statistics of the grid search results in table 13

p	q	NoPos	BestTrade	WorstTrade	Difference best and worst trade	AI	% Failed	Income
97	53	685.94	9.08	-28.88	37.96	491.45	40.50%	-$80.85
97	55	645.56	8.51	-29.43	37.94	469.90	40.85%	-$81.84
99	53	421.56	10.00	-24.97	34.97	377.80	41.02%	-$42.51
99	55	425.22	10.14	-24.62	34.77	362.99	40.94%	-$43.60

Window	NoPos	BestTrade	WorstTrade	Difference best and worst trade	AI	% Failed	Income
360	1,044.25	10.50	-33.48	43.98	322.03	53.80%	-$159.33
1440	381.00	9.04	-14.59	23.63	425.94	34.91%	$8.54
2880	208.46	8.76	-32.85	41.61	528.64	33.78%	-$35.80

k	NoPos	BestTrade	WorstTrade	Difference best and worst trade	AI	% Failed	Income
5	326.92	10.22	-32.26	42.49	491.43	42.82%	-$74.92
10	527.25	9.88	-26.47	36.35	431.09	39.74%	-$60.88
20	779.54	8.19	-22.19	30.38	354.09	39.93%	-$50.80

new market, after the bubble in December/January of 2017/2018. All data before 2017 represents an entirely different market, where either most cryptocurrencies were not available or their market capacity was an iota of what it is today. Therefore, it is difficult to predict the future merit of trading strategies in the cryptocurrency market due to the relatively sentiment-driven evolvement of this market. Regardless, the analyses performed in this paper still show the power of the inclusion of short-term trends in cointegration-based pairs-trading strategies.

6.2 Robustness analysis

In this section, we look at the weekly results of the market neutral pairs-trading strategy with and without the inclusion of short-term trends to gain insight on the algorithms' performance in terms of robustness. In addition, we examine the sensitivity of the variables in table 13.

From the grid search results in section 5.1 we see that window size h has a strong effect on performance. Figure 8b shows that, in the small sample of the grid search, the optimal return are for a window size of 1 day. The sensitivity of variables p, q, and h

are shown in table 10. Firstly, we see that a smaller value for p allows for more trades and thus results in a higher average investment. Additionally, we see that the sensitivity on the number of trades of q is smaller than that of p. This is due to the fact that the difference in time when hitting the $q = 53$ or $q = 55$ threshold is smaller and the occurrence is more probable than a equal difference in percentile p (figure 2). Secondly, the window size shows a negative correlation with the number of trades. As previously mentioned, a smaller window is more sensitive to shocks in the exchange rates which is reflected in the percentage of failed trades. This does not, however, explain why the difference between the best and worst trade behavior. This can simply be explained by the fact that the maximum and minimum increment size are not as good a metric as the percentage of non-positive trades (including transaction costs). It is important to note that the results in the sensitivity table are based on aggregate results of table 13 of the appendix and are therefore not ceteris paribus (they are on aggregate level over the entire table). The average investment increases with window size because the short window trades with a much shorter Δ_t, due to a higher sensitivity to shocks, whereas large windows have more overlapping trades requiring a larger amount of capital. Interestingly, we see a positive correlation of diversification and returns, as k increases. The average returns increase as we regard more cointegrating relations (and increase the maximum number of simultaneous opened positions, since this is equal to k), however this is due to the algorithm losing less as a result of a lower initial investment. The results in section 5.2 also show that there most likely never are more than ten simultaneous trades (averages are both below 5). This is also the reason why $k = 20$ is not part of the optimal iteration from section 5.1. The overall sensitivity to each variable could be mapped in a more robust way by increasing the grid and sample sizes, due to time management this was simplified although this can be done in future research.

As previously mentioned in the result section 5.2, the tighter confidence interval bounds of the Johansen method based strategy and the weekly performance lead to the

Table 11 Summary statistics of the weekly returns when using the pairs-trading strategy without the inclusion of short-term trends. The corresponding time series are plotted in figure 9a. Please note that the statistics are computed excluding the primary window, because there are no trades during that time. The last two columns symbolize the 95 % confidence interval lower and upper bounds.

Week	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	lb_95%	ub_95%
count	720	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640
mean	2.69	19.32	-23.32	25.38	39.99	5.94	-4.22	16.82	26.36	28.35	-1.66	8.92	43.55	65.46	72.53	19.36	-33.03	60.77
std	1.51	14.83	9.87	9.30	17.14	5.73	6.68	17.83	15.87	18.41	9.44	15.13	21.82	29.87	30.84	12.01	19.07	35.08
min	-3.34	-0.35	-35.20	0.00	-0.49	-6.59	-17.54	-4.65	0.00	0.00	-21.56	-20.26	-0.40	0.00	0.00	-1.07	-66.06	0.01
25%	1.04	6.65	-31.75	23.39	32.43	1.11	-9.29	0.47	12.19	9.47	-5.80	-0.25	29.05	36.29	55.89	8.28	-49.55	30.39
50%	3.57	17.05	-25.37	25.94	45.13	5.73	-5.63	10.97	18.62	29.62	-1.04	4.85	36.31	75.28	84.63	20.65	-33.03	60.77
75%	3.92	28.22	-18.04	33.73	52.84	11.39	1.69	37.10	37.01	41.51	3.69	21.36	63.22	90.61	99.19	30.45	-16.52	91.14
max	4.18	48.23	7.05	40.19	61.77	18.56	11.72	41.90	53.06	60.03	16.99	39.34	83.13	113.48	103.59	36.91	-0.01	121.52

conclusion that a short-term trend assisted pairs-trading strategy is more robust than one solely dependent on the EG2SLS.

Table 12 Summary statistics of the weekly returns when using the pairs-trading strategy with the addition of short-term trends (Johansen method). The corresponding time series are plotted in figure 9b. Please note that the statistics are computed excluding the primary window, because there are no trades during that time. The last two columns symbolize the 95 % confidence interval lower and upper bounds.

Week	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	lb_95%	ub_95%
count	720	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640
mean	0.81	8.09	3.86	17.19	24.46	4.17	-6.83	19.20	12.19	15.82	19.63	-3.32	25.70	27.46	43.87	3.72	-11.19	28.03
std	0.38	5.91	4.65	7.24	14.00	3.52	4.96	8.66	10.65	11.95	10.45	7.04	8.22	20.10	15.91	7.00	6.46	16.18
min	0.00	0.00	-4.29	0.00	-0.92	-1.85	-15.17	0.00	-3.31	-3.13	-0.16	-10.72	-0.39	-13.66	0.00	-11.37	-22.38	0.01
25%	0.83	2.64	-1.68	13.14	8.09	1.89	-12.28	13.70	2.35	2.63	13.41	-9.87	25.02	12.19	39.49	0.00	-16.78	14.02
50%	0.85	8.61	6.78	17.26	29.65	3.22	-6.10	17.00	5.62	18.29	20.27	-4.55	27.18	31.75	50.66	3.14	-11.19	28.03
75%	1.03	12.17	8.23	22.94	37.79	6.18	-3.95	30.31	23.16	27.65	26.27	3.05	30.47	42.72	54.47	9.90	-5.60	42.04
max	1.41	20.94	10.08	33.22	37.84	18.46	2.99	32.74	27.03	32.97	35.54	10.85	36.41	52.99	59.77	12.99	0.00	56.05

6.3 Extensions

Research in the cryptocurrency market is still in its infancy. Apart from implementing known trading strategies from the FX market, there are new methods emerging which can be tested in this high-frequency environment such as deep learning algorithms or Bayesian neural networks. Furthermore, the cryptocurrency market currently does not have many possibilities for shorting exchange rates. If shorting would be possible in the near future, it can be interesting to implement and evaluate a market neutral pairs-trading strategy in practice. Given the sentimental nature of the cryptocurrency market, one could perform a pairs-trading strategy using an asymmetric VECM.

The cryptocurrency market shows forms of non-stationary behavior until the transaction costs of the regarded exchange. In addition, the fact that traders can have different fee brackets and there is no Smart Order Routing between cryptocurrency exchanges (yet), means that there is a very complex threshold regarding this non-stationary behavior. To analyze this behavior, one could extend our research by implementing a (dynamic) threshold VECM to perform a cointegration-based pairs-trading strategy. Furthermore, Bekiros and Diks (2008) perform non-linear Granger causality tests, adjusting for cointegrating relations, to test for lead/lag relations between spot and future prices of commodities. This non-linear test can be used for our Granger causality-based pre-selection of exchange rates.

With regards to the development of the cryptocurrency market, it is useful to keep an eye on the effects of incipient blockchain technologies in modern cryptocurrencies. The reason is that the previous emerged foundation for cryptocurrencies, Ethereum, spawned a wave of new coins flooding the market. Given our Granger causality analysis from section 5.3, we expect similar co-movements as we saw from all the ERC-20 coins.

Furthermore, it can be advantageous to research the changes in expected growth of the cryptocurrencies. In the long-run, it is our belief that not many coins will survive, chiefly with regards to a practical implementation of cryptocurrencies in everyday transactions. If we manage to differentiate the more robust/trustworthy coins, especially in the cointegrating relations pool, we can reduce losses as a result of bankruptcy of altcoins in the long-run.

7 Conclusions

In this paper we research whether or not high-frequency pairs-trading, by means of cointegration, produces significant positive returns. Specifically, we research the difference in pairs-trading performance of two different pairs-selection methods, which differ on the inclusion of short-term trends. In addition, we performed Granger causality tests to examine the directionality of relations between cryptocurrency exchange rates. We then analyzed the presence of the Granger caused exchange rates in the EG2SLS/Johansen-based cointegrating relations.

The Granger causality tests revealed that the lagging exchange rates are traded the most in the pairs-trading simulation. In addition, these lagging exchange rates form the elements of the top cointegrating relations. The results show that Granger causality tests can be used to find a subset of time series to perform a cointegration-based pairs-trading algorithm on, namely the lagging exchange rates. There currently is no literature, that we know of, on the pre-selection of exchange rates for cointegration by means of Granger causality tests. Further research, in another market and/or over a longer time period, comparing performance with and without the use of Granger causality tests, will help support our observations. Therefore, we conclude that the connection between Granger caused exchange rates and cointegrating relations shows promise, but more research, over a longer time period, is required to truly test the validity of this hypothesis.

In both cases, the Johansen-assisted and the EG2SLS cointegration pairs-trading strategies had positive returns on average. In addition, basing the existence of cointegration on the rank of Π from the VECM increases the robustness of the strategy. This follows from the fact that the Johansen-based method trades in about half the cases of the EG2SLS, resulting in less sensitive results and a steady positive weekly return of approximately 6.80%.

In conclusion, a high-frequency pairs-trading algorithm by means of cointegration produces significant positive returns in the cryptocurrency market. Additionally, the short-term trends in the VECM contain significant information to reduce variance in cointegration-based pairs-trading returns.

References

Alexander, C. (1999). Optimal hedging using cointegration. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 357(1758):2039–2058.

Bekiros, S. D. and Diks, C. G. (2008). The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality. *Energy Economics*, 30(5):2673–2685.

Caldeira, J. and Moura, G. (2013). Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. *SSRN*.

Dwyer, G. P. (2015). The johansen tests for cointegration. *White Paper*.

Elliott, R. J., Hoek, J. V. D., and Malcolm, W. P. (2005). Pairs trading. *Quantitative Finance*, 5(3):271–276.

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

Gatev, E., Goetzmann, W. N., and Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, 19(3):797–827.

Harris, R. D. (1995). *Using cointegration analysis in econometric modelling*. Harvester Wheatsheaf, Prentice Hall.

Huck, N. and Afawubo, K. (2015). Pairs trading and selection methods: is cointegration superior? *Applied Economics*, 47(6):599–613.

Johansen, S. (1995). Identifying restrictions of linear equations with applications to simultaneous equations and cointegration. *Journal of econometrics*, 69(1):111–132.

King, M. R. and Rime, D. (2010). The $4 trillion question: what explains fx growth since the 2007 survey? *BIS Quarterly Review*.

Leung, T. and Li, X. (2015). Optimal mean reversion trading with transaction costs and stop-loss exit. *International Journal of Theoretical and Applied Finance*, 18(03):1550020.

MacKinnon, J. G. (1991). Critical values for cointegration tests. In *Eds.), Long-Run Economic Relationship: Readings in Cointegration*. Citeseer.

Miao, G. J. (2014). High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. *International Journal of Economics and Finance*, 6(3):96.

Nath, P. (2003). High frequency pairs trading with us treasury securities: Risks and rewards for hedge funds. *SSRN*.

Ng, S. and Perron, P. (1995). Unit root tests in arma models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, 90(429):268–281.

Papadakis, G. and Wisocky, P. (2007). Pairs trading and accounting information. *Working Paper, MIT Sloan School of Management, Cambridge, MA*.

Perlin, M. S. (2009). Evaluation of pairs-trading strategy at the brazilian financial market. *Journal of Derivatives & Hedge Funds*, 15(2):122–136.

8 Appendix

8.1 Grid search results

Table 13 This table shows the grid search results per method on window size h, top-k selectable cointegrating relations, percentiles p and q ($*100$) and investment type. The grid search was performed on 60 second aggregated data, therefore all window sizes h are based on observations of 60 seconds. Furthermore, this analysis is performed using the EG2SLS and not with the Johansen method. The table is sorted by FinalReturns. **Columns:** InvestType is the investment type, which is uniform or weighted by ADF test p-value of $\hat{\epsilon}$. NoPos is the total number of positions opened. NoLongs is the number of long positions opened in the Y exchange rate of the equation 3. % Failed is the number of failed positions over time, where 'failed' means that the returns where non-positive. AI is the average USDT capital invested in the opened positions over time. BestTrade, WorstTrade and FinalIncome are all in USDT. RpP stands for returns per position and equals the FinalReturns divided by the number of positions. FinalReturns is the percentage of final returns divided by the average investment.

Method	Window	k	p	q	InvestType	NoPos	NoLongs	% Failed	AI	BestTrade	WorstTrade	FinalIncome	RpP	FinalReturns
2	360	20	99	55	P-value weighted	95	37	62.11	42.53	31.72	-3.83	34.82	0.37	81.88
2	360	10	99	55	P-value weighted	66	32	57.58	44.86	23.69	-3.83	29.13	0.44	64.95
2	**1440**	**10**	**99**	**53**	**Uniform**	**143**	**78**	**26.57**	**243.26**	**7.77**	**-4.83**	**83.12**	**0.58**	**34.17**
2	1440	10	99	55	Uniform	134	80	29.85	214.79	9.40	-3.72	59.38	0.44	27.65
2	1440	5	99	55	Uniform	128	70	26.56	280.24	6.80	-6.35	74.14	0.58	26.46
2	2880	20	99	55	Uniform	226	126	26.99	327.12	10.27	-4.60	61.38	0.27	18.77
2	1440	5	99	53	Uniform	107	57	29.91	334.78	8.69	-6.15	59.60	0.56	17.80
2	2880	20	99	53	Uniform	221	122	27.15	328.17	10.69	-4.51	58.07	0.26	17.69
2	1440	5	97	53	Uniform	296	146	37.84	555.13	12.95	-9.67	67.31	0.23	12.12
1	1440	5	99	53	Uniform	234	126	34.62	523.83	7.23	-13.19	62.31	0.27	11.90
2	1440	5	97	55	Uniform	279	139	37.63	517.57	12.95	-11.57	56.81	0.20	10.98
1	1440	5	99	55	Uniform	235	127	34.47	523.22	7.23	-13.19	56.14	0.24	10.73
2	1440	20	99	53	Uniform	179	103	33.52	187.80	2.35	-3.62	14.23	0.08	7.58
1	1440	5	97	53	Uniform	313	156	36.10	611.18	6.30	-13.19	43.30	0.14	7.09
2	2880	10	99	55	Uniform	153	82	32.03	396.83	5.75	-11.26	27.11	0.18	6.83
2	1440	10	99	53	P-value weighted	143	78	26.57	324.58	40.23	-20.00	22.08	0.15	6.80
2	1440	10	97	55	Uniform	435	238	36.32	464.25	6.48	-7.13	31.30	0.07	6.74
1	1440	5	97	55	Uniform	314	158	37.58	609.86	6.30	-13.19	40.85	0.13	6.70
2	1440	20	99	55	Uniform	159	88	33.96	158.02	2.35	-3.51	9.88	0.06	6.25
2	1440	10	99	55	P-value weighted	134	80	29.85	254.08	40.23	-20.00	15.71	0.12	6.18
2	2880	20	99	55	P-value weighted	226	126	26.99	444.90	12.68	-18.61	26.96	0.12	6.06
2	2880	10	99	53	Uniform	158	85	32.91	420.12	5.75	-11.26	24.79	0.16	5.90
1	1440	10	99	55	Uniform	407	226	35.38	488.64	3.62	-6.46	28.16	0.07	5.76
2	1440	20	97	55	Uniform	607	326	36.24	367.94	3.27	-3.81	20.61	0.03	5.60
2	2880	20	99	53	P-value weighted	221	122	27.15	429.38	12.68	-18.19	23.83	0.11	5.55
2	1440	10	97	53	Uniform	475	257	37.89	505.42	6.48	-7.13	27.36	0.06	5.41
2	360	5	99	53	P-value weighted	87	44	60.92	69.63	12.45	-8.13	3.60	0.04	5.17

Table 13 continued from previous page

Method	Window	k	p	q	InvestType	NoPos	NoLongs	% Failed	AI	BestTrade	WorstTrade	FinalIncome	RpP	FinalReturns
1	1440	10	99	53	Uniform	398	219	35.43	493.69	3.62	-6.46	22.40	0.06	4.54
2	1440	5	99	53	P-value weighted	107	57	29.91	417.88	30.51	-22.58	18.38	0.17	4.40
2	1440	20	97	53	Uniform	662	352	36.10	392.28	3.27	-5.44	13.66	0.02	3.48
1	1440	10	97	55	Uniform	513	279	36.84	551.30	3.15	-7.13	17.51	0.03	3.18
1	1440	10	97	53	Uniform	499	270	36.87	557.18	3.15	-7.13	11.34	0.02	2.04
1	2880	20	97	53	Uniform	346	179	29.77	420.20	3.15	-5.05	7.20	0.02	1.71
2	360	10	99	53	P-value weighted	71	34	57.75	58.99	23.69	-25.45	0.86	0.01	1.45
2	2880	10	97	53	Uniform	239	124	30.13	521.78	5.75	-10.36	6.73	0.03	1.29
1	2880	20	97	55	Uniform	354	183	29.94	419.82	3.23	-5.06	4.42	0.01	1.05
2	2880	20	97	53	Uniform	350	177	30.57	418.19	3.23	-4.92	4.17	0.01	1.00
2	2880	20	97	55	Uniform	354	182	30.51	417.82	3.23	-4.99	2.48	0.01	0.59
1	2880	10	97	53	Uniform	237	124	30.38	524.76	5.75	-9.99	0.63	0.00	0.12
1	1440	5	99	53	P-value weighted	234	126	34.62	519.98	9.73	-21.25	0.50	0.00	0.10
2	2880	10	97	55	Uniform	239	123	30.96	517.23	5.75	-10.60	0.26	0.00	0.05
1	2880	10	97	55	Uniform	239	125	30.54	523.62	5.75	-10.03	-1.13	0.00	-0.22
1	1440	20	99	55	Uniform	595	318	37.31	397.82	3.11	-7.75	-1.21	0.00	-0.30
1	2880	20	99	53	Uniform	275	141	33.09	377.88	3.15	-4.84	-3.03	-0.01	-0.80
1	1440	5	99	55	P-value weighted	235	127	34.47	528.64	9.73	-22.11	-4.68	-0.02	-0.89
1	2880	20	99	55	Uniform	280	143	32.86	378.12	3.23	-4.86	-4.61	-0.02	-1.22
1	1440	20	97	53	Uniform	716	379	36.03	443.86	3.11	-7.84	-5.89	-0.01	-1.33
1	1440	20	99	53	P-value weighted	581	306	37.69	391.10	8.51	-19.49	-5.74	-0.01	-1.47
1	1440	20	97	55	Uniform	735	391	37.01	443.79	3.11	-7.84	-7.26	-0.01	-1.64
1	1440	20	99	53	Uniform	581	306	37.69	402.97	3.11	-7.75	-8.81	-0.02	-2.19
2	1440	5	99	55	P-value weighted	128	70	26.56	269.54	9.73	-25.05	-5.92	-0.05	-2.20
1	2880	5	99	53	P-value weighted	102	51	42.16	585.31	14.72	-43.85	-15.67	-0.15	-2.68
1	1440	20	99	55	P-value weighted	595	318	37.31	398.51	8.51	-20.27	-10.74	-0.02	-2.70
1	2880	5	99	55	P-value weighted	104	52	41.35	551.80	14.72	-45.34	-16.27	-0.16	-2.95
1	1440	5	97	55	P-value weighted	314	158	37.58	601.02	9.45	-21.84	-18.16	-0.06	-3.02
2	360	20	99	53	P-value weighted	111	50	61.26	33.83	3.62	-3.83	-1.16	-0.01	-3.43
1	1440	5	97	53	P-value weighted	313	156	36.10	603.20	9.45	-20.98	-22.27	-0.07	-3.69
1	1440	10	99	55	P-value weighted	407	226	35.38	498.49	8.70	-21.21	-20.12	-0.05	-4.04
1	2880	10	99	55	Uniform	187	97	34.22	476.28	5.75	-9.67	-19.80	-0.11	-4.16
2	1440	5	97	55	P-value weighted	279	139	37.63	491.09	10.79	-24.78	-20.79	-0.07	-4.23
1	1440	10	99	53	P-value weighted	398	219	35.43	477.31	8.70	-20.39	-20.38	-0.05	-4.27
2	2880	10	99	53	P-value weighted	158	85	32.91	448.29	8.25	-41.75	-20.27	-0.13	-4.52
2	1440	5	97	53	P-value weighted	296	146	37.84	519.44	9.45	-21.84	-24.23	-0.08	-4.66
1	2880	10	99	53	Uniform	183	96	34.97	476.14	3.95	-9.58	-22.91	-0.13	-4.81
2	2880	10	99	55	P-value weighted	153	82	32.03	435.15	8.25	-44.08	-23.46	-0.15	-5.39
2	1440	20	97	55	P-value weighted	607	326	36.24	392.49	10.75	-22.73	-21.19	-0.03	-5.40
1	1440	10	97	55	P-value weighted	513	279	36.84	541.50	8.70	-20.95	-30.77	-0.06	-5.68
2	2880	5	97	55	P-value weighted	139	69	36.69	687.57	14.72	-64.24	-39.50	-0.28	-5.74
2	1440	10	97	55	P-value weighted	435	238	36.32	441.94	10.75	-23.78	-25.48	-0.06	-5.77
2	2880	20	97	53	P-value weighted	350	177	30.57	547.89	15.48	-49.31	-33.87	-0.10	-6.18
1	1440	10	97	53	P-value weighted	499	270	36.87	528.40	8.70	-20.13	-34.03	-0.07	-6.44
2	2880	10	97	53	P-value weighted	239	124	30.13	641.42	15.68	-63.96	-41.55	-0.17	-6.48
1	1440	20	97	55	P-value weighted	735	391	37.01	417.59	7.87	-20.03	-27.74	-0.04	-6.64
2	1440	20	97	53	P-value weighted	662	352	36.10	389.74	9.33	-20.03	-25.95	-0.04	-6.66
1	2880	10	99	53	P-value weighted	183	96	34.97	537.30	8.43	-43.85	-36.12	-0.20	-6.72
2	2880	20	97	55	P-value weighted	354	182	30.51	555.57	15.48	-51.53	-38.55	-0.11	-6.94
1	1440	20	97	53	P-value weighted	716	379	36.03	417.83	7.87	-19.24	-29.95	-0.04	-7.17
2	1440	20	99	55	P-value weighted	159	88	33.96	133.36	3.93	-17.88	-9.69	-0.06	-7.27
1	2880	10	99	55	P-value weighted	187	97	34.22	502.83	8.43	-45.34	-37.00	-0.20	-7.36
2	1440	10	97	53	P-value weighted	475	257	37.89	466.49	9.33	-20.95	-35.93	-0.08	-7.70
1	2880	5	97	53	P-value weighted	141	71	37.59	722.14	14.72	-64.24	-59.42	-0.42	-8.23

Table 13 continued from previous page

Method	Window	k	p	q	InvestType	NoPos	NoLongs	% Failed	AI	BestTrade	WorstTrade	FinalIncome	RpP	FinalReturns
2	2880	5	97	53	P-value weighted	143	72	37.06	688.32	14.72	-64.24	-57.02	-0.40	-8.28
2	2880	10	97	55	P-value weighted	239	123	30.96	645.37	15.68	-63.96	-54.70	-0.23	-8.48
1	2880	5	97	55	P-value weighted	143	72	37.76	688.23	14.72	-64.24	-61.04	-0.43	-8.87
2	2880	5	99	53	Uniform	95	52	41.05	533.93	5.18	-24.93	-60.26	-0.63	-11.29
1	2880	20	97	53	P-value weighted	346	179	29.77	584.57	12.68	-46.91	-69.24	-0.20	-11.84
2	2880	5	99	55	Uniform	95	52	40.00	522.37	5.18	-24.93	-63.07	-0.66	-12.07
2	2880	5	99	53	P-value weighted	95	52	41.05	539.62	9.26	-56.26	-67.50	-0.71	-12.51
1	2880	10	97	53	P-value weighted	237	124	30.38	662.49	8.25	-63.96	-83.26	-0.35	-12.57
1	2880	20	97	55	P-value weighted	354	183	29.94	547.89	12.68	-49.31	-71.75	-0.20	-13.10
2	2880	5	97	53	Uniform	143	72	37.06	645.68	5.18	-31.48	-87.94	-0.61	-13.62
1	2880	10	97	55	P-value weighted	239	125	30.54	627.94	8.25	-63.96	-86.10	-0.36	-13.71
2	2880	5	99	55	P-value weighted	95	52	40.00	551.27	8.49	-56.26	-75.91	-0.80	-13.77
2	2880	5	97	55	Uniform	139	69	36.69	621.92	5.21	-31.45	-91.04	-0.65	-14.64
1	2880	5	97	55	Uniform	143	72	37.76	644.20	5.18	-31.78	-100.52	-0.70	-15.60
1	2880	5	97	53	Uniform	141	71	37.59	649.21	5.18	-31.10	-102.85	-0.73	-15.84
1	2880	20	99	53	P-value weighted	275	141	33.09	538.00	12.68	-46.91	-87.10	-0.32	-16.19
1	2880	20	99	55	P-value weighted	280	143	32.86	499.03	12.68	-49.31	-90.75	-0.32	-18.19
2	360	5	99	55	P-value weighted	75	37	61.33	70.45	12.45	-10.63	-13.07	-0.17	-18.55
2	1440	20	99	53	P-value weighted	179	103	33.52	151.88	7.31	-28.78	-29.39	-0.16	-19.35
1	2880	5	99	55	Uniform	104	52	41.35	573.52	5.30	-30.37	-120.35	-1.16	-20.99
1	2880	5	99	53	Uniform	102	51	42.16	577.75	5.18	-29.70	-123.09	-1.21	-21.31
2	360	5	97	53	Uniform	433	221	53.12	224.99	5.46	-6.51	-53.70	-0.12	-23.87
2	360	5	97	55	Uniform	275	140	53.45	134.83	2.99	-5.67	-37.54	-0.14	-27.84
2	360	10	97	53	Uniform	689	351	51.52	201.39	3.28	-6.14	-58.66	-0.09	-29.13
2	360	10	99	55	Uniform	66	32	57.58	32.74	2.45	-3.77	-9.86	-0.15	-30.10
2	360	5	99	55	Uniform	75	37	61.33	64.96	3.36	-5.17	-20.80	-0.28	-32.02
1	360	10	99	53	Uniform	1484	757	49.39	481.30	5.73	-11.90	-156.20	-0.11	-32.45
2	360	20	97	53	Uniform	1036	526	51.06	164.74	3.04	-3.43	-53.47	-0.05	-32.46
1	360	10	99	55	Uniform	1506	769	49.67	475.42	5.73	-12.05	-157.49	-0.10	-33.13
1	360	5	99	53	Uniform	898	467	51.11	536.70	9.91	-23.70	-179.99	-0.20	-33.54
1	360	20	99	55	Uniform	2299	1164	49.33	388.67	3.43	-6.14	-132.41	-0.06	-34.07
1	360	20	99	53	Uniform	2261	1143	48.83	391.63	3.43	-6.06	-133.80	-0.06	-34.16
2	360	5	99	53	Uniform	87	44	60.92	69.49	2.81	-4.40	-24.28	-0.28	-34.93
2	360	10	99	53	Uniform	71	34	57.75	43.27	2.69	-4.22	-15.24	-0.21	-35.23
1	360	5	99	55	Uniform	910	472	51.98	533.69	9.91	-24.00	-190.06	-0.21	-35.61
1	360	20	97	53	Uniform	2771	1395	51.39	442.98	3.19	-6.06	-183.00	-0.07	-41.31
2	360	20	99	53	Uniform	111	50	61.26	34.79	1.29	-2.25	-14.41	-0.13	-41.41
1	360	5	97	53	Uniform	1141	575	53.20	624.10	11.97	-23.70	-264.23	-0.23	-42.34
1	360	20	97	55	Uniform	2834	1432	52.05	440.45	3.19	-6.14	-187.26	-0.07	-42.52
2	360	20	97	55	Uniform	662	343	53.47	102.67	2.50	-3.43	-44.51	-0.07	-43.35
2	360	5	97	53	P-value weighted	433	221	53.12	246.42	14.96	-22.44	-107.76	-0.25	-43.73
2	360	10	97	55	Uniform	441	233	51.25	124.52	1.51	-6.14	-54.51	-0.12	-43.78
1	360	10	97	53	Uniform	1860	934	52.37	552.91	4.48	-11.90	-244.21	-0.13	-44.17
1	360	5	97	55	Uniform	1159	584	54.10	620.70	11.97	-24.00	-274.60	-0.24	-44.24
2	360	10	97	53	P-value weighted	689	351	51.52	226.60	14.95	-21.09	-101.77	-0.15	-44.91
1	360	10	97	55	Uniform	1898	956	53.00	546.88	4.48	-12.05	-247.51	-0.13	-45.26
2	360	20	97	53	P-value weighted	1036	526	51.06	189.45	14.95	-20.54	-92.26	-0.09	-48.70
1	360	5	99	53	P-value weighted	898	467	51.11	599.51	18.92	-101.07	-311.84	-0.35	-52.02
1	360	20	99	53	P-value weighted	2261	1143	48.83	455.18	18.85	-98.38	-252.98	-0.11	-55.58
1	360	5	99	55	P-value weighted	910	472	51.98	592.28	18.92	-102.35	-335.60	-0.37	-56.66
1	360	10	99	53	P-value weighted	1484	757	49.39	565.70	18.88	-99.38	-334.12	-0.23	-59.06
1	360	20	99	55	P-value weighted	2299	1164	49.33	448.49	18.85	-99.63	-274.34	-0.12	-61.17
1	360	10	99	55	P-value weighted	1506	769	49.67	545.84	18.88	-100.64	-340.18	-0.23	-62.32
2	360	20	99	55	Uniform	95	37	62.11	23.24	1.73	-2.25	-15.00	-0.16	-64.52

Table 13 continued from previous page

Method	Window	k	p	q	InvestType	NoPos	NoLongs	% Failed	AI	BestTrade	WorstTrade	FinalIncome	RpP	FinalReturns
1	360	20	97	53	P-value weighted	2771	1395	51.39	496.20	20.02	-98.38	-320.47	-0.12	-64.58
1	360	5	97	53	P-value weighted	1141	575	53.20	678.87	15.62	-101.07	-448.41	-0.39	-66.05
2	360	5	97	55	P-value weighted	275	140	53.45	165.14	8.42	-22.23	-111.37	-0.40	-67.44
1	360	20	97	55	P-value weighted	2834	1432	52.05	490.41	20.02	-99.63	-338.35	-0.12	-68.99
1	360	5	97	55	P-value weighted	1159	584	54.10	667.40	15.62	-102.35	-468.51	-0.40	-70.20
1	360	10	97	53	P-value weighted	1860	934	52.37	626.76	15.61	-99.38	-450.82	-0.24	-71.93
2	360	20	97	55	P-value weighted	662	343	53.47	129.54	8.21	-20.34	-94.00	-0.14	-72.56
1	360	10	97	55	P-value weighted	1898	956	53.00	605.99	15.61	-100.64	-453.03	-0.24	-74.76
2	360	10	97	55	P-value weighted	441	233	51.25	150.20	8.31	-20.89	-113.63	-0.26	-75.65

8.2 Historical simulation results

Market neutral pairs-trading algorithm results:

Table 14 This table contains the weekly pairs-trading strategy performance results with the EG2SLS pairs-selection method. The analysis was performed on aggregated data of 1 minute and with opening percentile $p = 97$. RpP and RpD stand for returns per position and returns per day, respectively.

Week	Method	Window	k	Horizon	p	q	InvestType	NoPos	NoLongs	% Failed	AI	BestTrade	WorstTrade	FinalIncome	RpP	RpD	FinalReturns
20	2	1440	10	2160	97	53	Uniform	24	9	45.83	493.00	1.52	-4.23	-4.91	-0.20	-0.50	-1.00
21	2	1440	10	10080	97	53	Uniform	227	132	32.16	501.86	3.03	-4.40	39.27	0.17	1.12	7.82
22	2	1440	10	10080	97	53	Uniform	183	104	43.17	590.45	3.11	-6.10	-45.12	-0.25	-1.09	-7.64
23	2	1440	10	10080	97	53	Uniform	309	176	30.42	516.81	4.15	-6.76	59.40	0.19	1.64	11.49
24	2	1440	10	10080	97	53	Uniform	264	122	29.17	577.60	3.28	-6.80	38.59	0.15	0.95	6.68
25	2	1440	10	10080	97	53	Uniform	292	160	34.93	589.15	3.85	-9.80	7.93	0.03	0.19	1.35
26	2	1440	10	10080	97	53	Uniform	274	152	45.99	596.44	2.28	-9.98	-19.38	-0.07	-0.46	-3.25
27	2	1440	10	10080	97	53	Uniform	238	128	34.45	593.99	4.28	-10.72	6.36	0.03	0.15	1.07
28	2	1440	10	10080	97	53	Uniform	276	133	34.42	571.35	3.52	-5.96	58.36	0.21	1.46	10.22
29	2	1440	10	10080	97	53	Uniform	222	109	32.88	546.42	3.61	-6.34	51.38	0.23	1.34	9.40
30	2	1440	10	10080	97	53	Uniform	231	117	35.50	653.01	3.86	-11.13	-3.65	-0.02	-0.08	-0.56
31	2	1440	10	10080	97	53	Uniform	244	115	36.89	625.02	18.85	-10.97	36.77	0.15	0.84	5.88
32	2	1440	10	10080	97	53	Uniform	212	117	34.91	594.38	14.03	-6.73	70.31	0.33	1.69	11.83
33	2	1440	10	10080	97	53	Uniform	283	121	21.91	482.45	5.89	-52.29	152.82	0.54	4.53	31.68
34	2	1440	10	10080	97	53	Uniform	260	145	30.77	540.46	5.25	-10.30	83.61	0.32	2.21	15.47
35	2	1440	10	10080	97	53	Uniform	268	122	27.61	635.07	3.06	-11.73	-2.51	-0.01	-0.06	-0.40
Avg:	**2**	**1440**	**10**	**9585**	**97**	**53**	**Uniform**	**237.94**	**122.63**	**34.44**	**569.22**	**5.22**	**-10.89**	**33.08**	**0.11**	**0.87**	**6.25**

Table 15 Summary statistics of the weekly returns when using the pairs-trading strategy with only the EG2SLS method and with opening percentile $p = 97$. Please note that the statistics are computed excluding the primary window, because there are no trades during that time. The last two columns symbolize the 95 % confidence interval lower and upper bounds.

Week	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	lb_95%	ub_95%
count	720	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640	8640
mean	2.98	20.57	-29.30	39.99	37.91	6.93	-6.99	21.48	30.32	35.10	4.30	33.13	38.99	94.32	86.42	7.04	-42.93	76.01
std	1.84	16.31	10.71	14.12	12.50	6.21	6.92	21.02	18.36	19.98	8.44	25.29	18.56	39.56	34.92	9.53	24.78	43.88
min	-4.91	-1.22	-49.97	0.00	0.00	-6.52	-22.62	-12.36	0.00	0.00	-18.74	-9.87	-0.29	0.00	0.00	-12.26	-85.85	0.02
25%	0.75	7.41	-36.76	34.20	34.73	1.47	-12.71	5.92	13.11	14.95	0.46	7.79	27.27	67.78	74.12	1.98	-64.39	38.01
50%	3.96	16.87	-29.54	43.37	39.93	6.82	-7.19	10.93	24.68	41.89	6.62	37.33	34.21	105.65	101.87	7.09	-42.93	76.01
75%	4.50	34.02	-23.97	46.27	47.69	12.27	-1.17	43.40	40.95	54.87	10.82	53.07	53.08	124.74	113.71	12.84	-21.47	114.01
max	5.43	51.79	3.07	67.05	55.00	26.77	7.79	58.08	66.82	60.21	19.29	75.66	74.13	154.04	120.02	23.74	-0.01	152.00

8.3 Granger causality

Table 16 This table contains the aggregate Granger causality relations of each cointegrating relation between week 30 and 36 of 2018 with a level of significance of 95%. Causality is shown from row to column, meaning that Bitcoin (BTC) caused Cardano (ADA) 70.8% of the time, whereas Cardano caused Bitcoin in 16.7% of the time. All exchange rates are in quote currency USDT.

	ADA	BCH	BNB	BTC	EOS	ETC	ETH	ICX	IOTA	LTC	NEO	ONT	QTUM	TRON	XLM	XRP
XLM	0.184	0.250	0.333	0.083	0.133	0.261	0.143	0.217	0.053	0.286	0.174	0.111	0.296	0.000	0.000	0.261
QTUM	0.150	0.000	0.154	0.125	0.000	0.154	0.091	0.212	0.059	0.200	0.107	0.238	0.000	0.000	0.185	0.381
EOS	0.970	1.000	0.846	0.625	0.000	0.917	1.000	0.842	1.000	0.900	0.929	0.818	0.842	0.857	0.733	0.846
BCH	0.583	0.000	0.143	0.250	0.429	0.444	0.615	0.400	0.545	0.625	0.708	0.562	0.273	0.500	0.667	0.600
BTC	0.708	0.438	0.400	0.000	0.125	0.333	0.385	0.857	0.778	0.455	0.833	0.857	0.750	0.917	0.667	0.625
LTC	0.250	0.625	0.500	0.182	0.500	0.286	0.308	0.500	0.636	0.000	0.462	0.308	0.200	0.417	0.429	0.200
NEO	0.409	0.250	0.167	0.222	0.143	0.333	0.250	0.583	0.368	0.462	0.000	0.464	0.357	0.478	0.174	0.136
ONT	0.343	0.062	0.316	0.214	0.364	0.182	0.400	0.323	0.304	0.077	0.464	0.000	0.381	0.304	0.389	0.190
TRON	0.447	0.071	0.211	0.083	0.214	0.400	0.467	0.516	0.467	0.333	0.261	0.478	0.526	0.000	0.333	0.250
ETC	0.286	0.333	0.500	0.556	0.250	0.000	0.545	0.412	0.250	0.286	0.571	0.273	0.308	0.700	0.304	0.385
BNB	0.276	0.143	0.000	0.200	0.154	0.300	0.222	0.250	0.000	0.333	0.417	0.158	0.231	0.368	0.250	0.182
ADA	0.000	0.333	0.276	0.167	0.182	0.357	0.333	0.460	0.552	0.300	0.500	0.314	0.350	0.579	0.474	0.378
ETH	0.778	0.692	0.556	0.615	0.308	0.455	0.000	0.778	0.733	1.000	0.650	0.800	0.773	0.800	0.762	0.778
ICX	0.160	0.150	0.150	0.095	0.105	0.353	0.222	0.000	0.174	0.357	0.278	0.161	0.182	0.258	0.217	0.115
XRP	0.243	0.267	0.182	0.125	0.308	0.154	0.111	0.500	0.200	0.400	0.273	0.333	0.381	0.438	0.348	0.000
IOTA	0.241	0.182	0.143	0.111	0.167	0.000	0.200	0.087	0.000	0.273	0.053	0.261	0.294	0.067	0.158	0.100