Université Paris-Dauphine
TOBAM

Copula-Based Trading of Cointegrated Cryptocurrency Pairs

Crypto Trading Research Project

Élèves :
Paul-Antoine Charbit
Jérémie Vilpellet
Matthieu Monnot

TOBAM tutor :
Jean Czyhir

Supervisor :
Tatjana Puhan

11 janvier 2024

Table des matières

1 Introduction

This paper takes part of a research project in quantitative crypto trading conducted by three students of Université Paris-Dauphine and supervised by Tobam quant researchers. The implemented approach in this study has been mainly inspired by M. TADI and J. WITZANY's paper of 2023 [1].

Pair trading is as an algorithmic trading strategy that involves the simultaneous buying and selling of two closely related financial instruments. The strategy is based on the premise that, over time, certain pairs of assets exhibit a historical correlation in their price movements. The ultimate aim is to take advantage of temporary divergences from their established statistical relationship, buying the related undervalued asset and selling the other, as there is a high probability the two prices will converge again. In this way, a mispricing index is design to evaluate a divergence among a period of time and trigger an opening trade position for the pair.

A lot of work have been done on this topic to create the more profitable pair trading strategy, first with distance method and then using cointegration approach. Our study tends to replicate and challenge one of the most advanced algorithm that generalize these cases involving copulas functions. In fact, the sinews of war is to best fit the dependency relationship between two assets. In this way, the study conducted calibrate copulas distribution to stationary spread process of cointegrated pairs and compute a mispricing index that serve as a threshold to open trade positions when an abnormal price movement happened.

This paper propose something new in this approach because the commonly used cumulative mispricing indices (CMI) is computed on assets log-returns and is not necessary stationary. In fact, CMI is constructed in order to contain past information (sum of MI over a certain period) and thus tend to be a random walk and does not garantee the stationarity needed to ensure a mean-reverting process. Hence, returns are replaced with stationnary spread processes between cointegrated altcoins and bitcoin, so we can directly drawn uniform random variables from those spread (which already contains past information) and thus use MI (conditionnal probability) as trading signal without construct a possibly non mean reverting signal. Now and then a price abnormal divergence from its behaviour will present a safer opportunity.

The crypto market have been selected because it is highly volatile so there are more trading opportunities. Also, high frequency trading can be perform on this market then it is very appropriate for copulas pair trading strategies. We have chosen bitcoin as a reference coin for cointegration rather than any altcoins because of its dominance in overall transactions and its central place within crypto markets.

To find the best possible cryptocurrency pair we identify cointegrated coins with bitcoin during a given period. We suppose that all differentiated price is stationnary and so coins are I(1). We set up both linear and non linear cointegration test with EG and KSS test on the 19 spreads implicating bitcoin and other coins. Then we fit a copula distribution to these two selected coins and build trading signals when there is an extreme deviation of both spreads related to the other (extreme conditionnal probability occurence which is derivated from copula relation).

The first objective of this paper is to reimplement the whole strategy of the article mentionned and compare results. Then we will challenge our choices in order to try the robustness of the approach. We will analyse if other configurations have a significant influence on the retrun and results in general. Finally, we will conclude on the more profitable strategy.

2 Data presentation

2.1 Source

The data used in this paper is composed of 20 cryptocurrencies available on Binance Exchange, they are listed in the table 7. In this study, coins are quoted in USDT which is a stablecoin sticked onto the US Dollar. The crypto market being open twenty-four hours a day, seven days a week we decided to use the close prices hourly to produce our analysis. Then we based our analysis on prices of assets every hour during the entire period from 01-01-2021 to 01-19-2023.
As the trading strategy cycle last four weeks with three weeks of training and one of trading we can highlight 104 cycles in the studied period. For each formation period, we identify the best cryptocurrencies pair that will be trade during the following trading period (see figure 1).

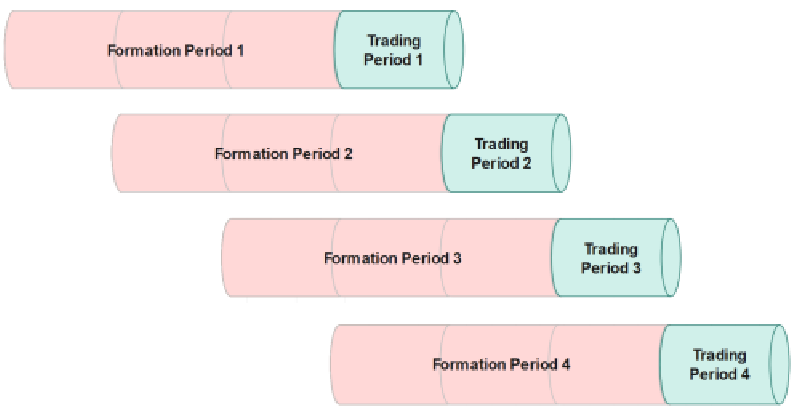


FIGURE 1 – Training and Trading Periods

2.2 Statistics

In order to give a statistical overview of assets traded but also to realized the order of magnitude of the cryptocurrency market, it is interesting to calculate some basic metrics often used in traditional markets such as mean, volatility and overall return. These are resume in the table 1.

Moreover, in the cryptocurrency market, a coin's dominance is the ratio of its market capitalization to the cumulative market capitalization of cryptocurrencies. This is a great way to see the weight of a coin in relation to the entire crypto market. It is calculated by dividing the market cap of a coin by the overall market cap of the other coins, then multiplying it by 100. The bitcoin dominance during the period studied has significatively decreased going from about 70% at the right beginning to under 40% at its lowest, meaning that the digital gold has lost his supremacy over the others and it has consequences on its price (see figure 2).

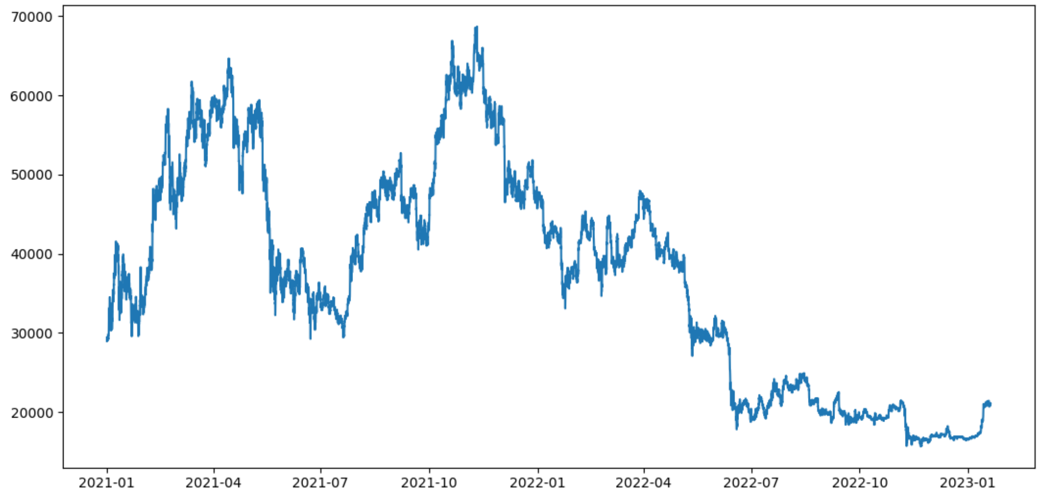


FIGURE 2 – BTCUSDT price

The trend we identify here can have a significant impact on the result of a trading method if we consider bitcoin as a reference for the benchmark. The trading strategy developed in this paper overcome this issue because trades are uncorrelated of market trends and the BTC position is always neutral as we buy and sell two different spreads, all implicating BTCUSDT. Thereby for a pair ETH-LTC, if ETH is overevaluated we sell ETH and buy BTC on one hand and sell BTC and buy LTC on the other hand. Overall we have buy and sell the same amount of BTC.

Coins	Mean	Volatility	Return
ADA	1.1	171%	87.9%
ATOM	18.9	208%	94.9%
BAT	0.65	192%	21.1%
BCH	394	160%	-64.7%
BNB	349	150%	690%
BTC	37306	107%	-27.7%
DASH	129	175%	-50.7%
EOS	3.1	172%	-62.4%
ETC	34.3	194%	266%
ETH	2355	139%	109%

Coins	Mean	Volatility	Return
IOTA	0.82	188%	-30.6%
LINK	18.2	178%	-41.7%
LTC	132	160%	-32.5%
ONT	0.68	186%	-55.9%
TRX	0.07	149%	125%
XLM	0.25	186%	-33.4%
XMR	205	157%	5.96%
XRP	0.68	176%	75.6%
XTZ	3.3	191%	-50.4%
ZEC	117	191%	-30.1%

Table 1 – Statistical overview of the entire period in USDT

3 Method

As mentioned in the introduction, the approach implemented in this paper directly comes from TADI and WITZANY article (2023 [1]). The method is the following.

First, we calculate spreads of each altcoin with BTCUSDT using a linear regression. Then in order to select the optimal cryptocurrency pair for the trading period we apply both Engle-Grenger (EG) and Kapetanios-Shin-Snell (KSS) cointegration test separatly. These tests provide the information whether a spread of two assets is stationary or in other words if two series are cointegrated. Tests perfomed needed two $I(1)$ series in input so the spread can be $I(0)$ if the pair is cointegrated.

The goal here is to find cointegrated coins with BTCUSDT then if there are two cointegrated coins we can build a pair of assets that have the same behaviour over time. As it can be several cryptocurrencies cointegrated with BTCUSDT we use another correlation metric to rank them from the more correlated with BTC to the least one. This metric is Kendall's tau which assesses the similarity in the orderings of data points between two variables. Concretely, it counts the number of simultaneous increase or decrease between two series and positively weight them. Conversely, it adds negative weights if simultaneous variations are not in the same direction. Finally, we are able to select the top two coins cointegrated with BTC and having the highest statistical correlation with it in terms of Kendall's tau.

After that, we work on the two spreads selected during the previous step. The aim here is to identify and fit the best marginal distribution of these two spreads. We have tried gaussian, student and cauchy distribution for which we determined optimal parameters using the maximum likelihood approach. We then chose the best distribution between these three based on the Akaike Information Criterion (AIC) which take into consideration the likelihood value and penalize it with the number of parameters. Then we transform the marginal distribution into uniform variables using the Probability Integral Transformation. Concretely, we apply the cumulative distribution function of the selected marginal distribution to the spread realisation itself so we create random uniform variables.

Now we introduce the concept of copula which is in two dimensions a function that maps the dependency relationship between two uniform variables into a variable in $[0, 1]$. The copula plays a crucial role in allowing dependence to be modeled independently of the marginal characteristics of the variables. In other words, it offers a flexible approach to representing dependency relationships between random variables. Sklar formulated a fundamental theorem, stating that there always exists a copula that can relate the marginal distributions of random variables to the joint distribution functions.

Then, with the same method than previously, we select the best fitting copula distribution among differents famillies reported in the table 12. The maximum likelihood and AIC allow us to find the copula distribution which fit most to the real dependance between assets.

Finally, we compute copula conditional probabilities which are partial derivatives of the copula distribution with respect to one of its component. From these probabilities, we can deduce if a spread is deviating from it's long term period (relatively to formation period) knowing the information contain in the other one. Then we define a threshold for these Mispricing index via back-testing to open or close position on the pair of coins (in fine a trading signal is a dummy variable wether the MI cross the threshold or not).

4 Statistic description of the trading strategy

First step is to select, during the training period, the best pair in terms of cointegration and correlation with BTCUSDT. This pair will be traded in the following trading week. Selected pairs for each trading week are reported in the table 11.

$$S_t^i = \text{BTCUSDT}_t - \hat{\beta}_i P_t^i, \quad i = 1, 2, \ldots, 19$$

We can analyse some divergences from the paper under study, mainly because cointegration tests are not perfectly robust and don't implement the same approach to verify the stationnarity of a process. For example, ADF test compute a stationary test on a single series on which it tests the presence of a trend, a constant or neither. Unlike that, the EG test takes 2 series in input and compute itself the regression to create the spread between BTC and the other altcoins. Then it process to an ADF test on centered and trend-less residuals.

The occurences of cryptocurrencies in pair trading are presented in table 3. The distribution of occurences raise an important point regarding the choices of altcoins but also on the reference coin compared to which we carry out cointegration and which is not traded (BTCUSDT). We find high occurence for ETH, BNB and LTC for instance but very low ones for about 10 out of 19 altcoins and only 1 for ZEC. This leads us to think that there are two categories of coins, with radically different influences on the market in general, on the one hand BTC, ETH, LTC and BNB that also correspond to the highest crypto market capitalisations and on the other hand the rest of the alctoins which follow the trends but with more erratic movement.

EG test Coin	Occurrence	Frequence
ETH	43	41.34%
BNB	34	32.69%
LTC	30	28.84%
EOS	14	13.46%
ETC	13	12.50%
DASH	10	9.61%
LINK	10	9.61%
BAT	9	8.65%
BCH	9	8.65%

KSS test Coin	Occurrence	Frequence
ETH	37	35.57%
LTC	34	32.69%
BNB	33	31.73%
BCH	21	20.19%
ETC	14	13.46%
DASH	11	10.57%
EOS	11	10.57%
XRP	9	8.65%
LINK	8	7.69%

FIGURE 3 – Coin occurrence and frequence in pair trading

Moreover, the result reveals around 80% of similarity with the reference article which is important but can be questionned for the last 20%. That is a point on which we found the approach not completely robust because the selected pair for each week really depends on the parametrisation of the cointegration test. After several test, we found that the results are best fitted for an OLS regression without constant to find the spread, and then to apply an ADF test with 6 lags and without neither linear trend or constant. This approach is problematic because first, it does not adhere to the standard procedure of an ADF test in terms of testing for the significance of a trend or a constant before testing for a possible unit root. Secondly, it seems surprising to fix the number of lags without optimizing them through criteria such as AIC or BIC. Just to understand the issue, on the article report, we can analyse about 30% of divergence between selected pair with ADF test or KSS one. This being said, when we analyze the differences on a week-by-week basis, we observe that the majority of discrepancies are largely due to the threshold setting. Most discrepancies occur when p-values are close to 10%. This results in two scenarios : either we accept an asset that is likely rejected by the authors, where Kendall's tau of this asset with BTC is higher than for the asset they accept, or we reject an asset that they choose. In almost all these cases, we can observe a p-value between 0.09 and 0.10.
Finally, to discriminate among the remaining candidates, authors suggest using Kendall's tau. Although the notation used in the paper might be confusing, it's important to note that the tau is not calculated on the spreads, but rather for each coin with the BTC. Of course, this calculation must be done on the return series, as we do not want to compute descriptive statistics on non-stationary series.

Now we have selected the best pair for each trading week, we fit marginal distributions of the two spreads on gaussian, student or cauchy distributions and apply the cumulative distribution function of the selected law on the spread to make it uniform. Then we calibrate different families of copulas during the training weeks again and also report there occurences over the whole period (table 2).

Finally, we can evaluate our copula-based strategy over the 104 trading periods as expected. This step needs the spreads process of the selected pair on the corresponding trading period. As a reminder, spreads are fitted on the first three weeks of the period and then we apply the fitted spread on the last week for trading. Spreads of the trading

Selected Copulas	EG test	KSS test
Gaussian	6.1%	6.7%
Student-t	13.3%	3.8%
Clayton	4.1%	7.7%
Frank	3.1%	3.8%
Gumbel	1.0%	3.8%
Joe	6.2%	3.8%
BB1	1.0%	2.9%
BB6	0%	2.9%
BB7	5,1%	11.5%
BB8	11.2%	4.8%
Tawn type 1	26.5%	25.0%
Tawn type 2	30.4%	23.1%

TABLE 2 – Copulas occurrence and frequence over weeks

period are then used to compute marginal and copula distributions of the selected laws and mispricing indeces which are the derivatives of the copula distribution with respect to its two component. Again, marginal and copula distributions are fitted on the training period and apply on the following trading week. Mispricing indeces are calculated using numerical derivatives of the copula realisations regarding its normal components. The detailed formula is presented below.

$$h_{1|2} := h(u_1|u_2) = P(U_1 \leq u_1|U_2 = u_2) = \frac{\partial C(u_1, u_2)}{\partial u_2}$$
$$h_{2|1} := h(u_2|u_1) = P(U_2 \leq u_2|U_1 = u_1) = \frac{\partial C(u_1, u_2)}{\partial u_2}$$

This step gives series of derivatives for each component for all trading weeks. After all we define trading rules to take advantages of an abnormal behaviour between two cointegrated coins. The aim here is to enter a trade when the fisrt coin is overvalued (resp. undervalued) compared to the other coin of the traded pair. Then we generate signals to open a short position (resp. long) for the first coin and a long position (resp. short) for the second coin. The same method is operate with the second coin as reference thus in order to find better opportunities we combine both information and open a position when the two mispricing indeces indicate a consistent signal. In other words, a trade position is open whether the first coin is overvalued compares to the other one (first mispricing index) and the second coin is undervalued in regard to the first one (second mispricing index). Trading rules can be resume like in the table 3 and 4. When $h_{1|2}$ is under a certain threshold the spread realisation u_1 is considered undervalued then we long S_1 which is BTCUSDT $- \beta_1 \times P_1$ so as long as we have a neutral position on BTCUSDT we short $\beta_1 \times P_1$ because the spread has high probability to increase.

Trading Rule	Signals
If $h_{1\|2} < \alpha_1$ and $h_{2\|1} > 1 - \alpha_1$	open long S_1 and short S_2
If $h_{1\|2} > 1 - \alpha_1$ and $h_{2\|1} < \alpha_1$	open short S_1 and long S_2
If $\|h_{1\|2} - 0.5\| < \alpha_2$ and $\|h_{2\|1} - 0.5\| < \alpha_2$	close both positions

TABLE 3 – Trading Rules in Terms S_1 and S_2

Trading Rule	Signals
If $h_{1\|2} < \alpha_1$ and $h_{2\|1} > 1 - \alpha_1$	open long $\beta_2 \times P_2$ and short $\beta_1 \times P_1$
If $h_{1\|2} > 1 - \alpha_1$ and $h_{2\|1} < \alpha_1$	open short $\beta_2 \times P_2$ and long $\beta_1 \times P_1$
If $\|h_{1\|2} - 0.5\| < \alpha_2$ and $\|h_{2\|1} - 0.5\| < \alpha_2$	close both positions

TABLE 4 – Trading Rules in Terms P_1 and P_2

The final step is to back-test the trading strategy on different values of α_1 which is the threshold to open positions. We consider the threshold to close positions at $\alpha_2 = 0.10$ as it is imposed in the related article. Thus we compute various metrics for tested strategies to evaluate their performance and at the end compare them to a classic buy & hold strategy. Results are presented in table 5.

We can highlight, first of all, that authors mentionned back-testing approach to best select the threshold value α_1 but they never tried to optimize this value on the historic data. They simply arbitrate between the total return and the overall risk as they introduced a risk-adjusted profitability. Thus they conclude that the strategy with threshold $\alpha_1 = 0.1$ maximize this profitability as, beyond this value, the risk increases while returns decrease.

Here, results aren't as simple to interpret because for EG test for instance, the pattern observed in the related article is confirm between $\alpha_1 = 0.1$ and $\alpha_1 = 0.15$ but reversed between $\alpha_1 = 0.15$ and $\alpha_1 = 0.2$. We still have consistent results concerning the number of transactions that rise with α_1 meaning there are more opportunities with a more permissive threshold. But considering returns and risk, we found the $\alpha_1 = 0.2$ strategy better that when it is at 0.15. To settle between first and last configuration we can refer to the theoretical aspect that a lower threshold is more robust but also on the non negligible 2% difference for transaction costs. Then regarding KSS test strategy, we retrieve the pattern of the article with higher returns and a lower risk for $\alpha_1 = 0.1$, so there is no ambiguity concerning the best threshold. Additionally, it is clearly identified here that a large part of the returns will go to the exchange for transaction fees and it becomes very worrying beyond 10% like using $\alpha_1 = 0.2$.

We finally choose the first threshold $\alpha_1 = 0.1$ as it, statistically and empirically, provides better trades than the second and third one. For both EG and KSS cointegration tests we found the first configuration more profitable and less risky.

To conclude this analysis, it is worth to notice that both strategies largely outperform the buy & hold one. In terms of returns and more especially the risk incurred, the buy & hold strategy is completely less rewarding : the Annualized Standard Deviation during the studied period is very high and do not garantee such good returns as a result thus the profitability is clearly limited.

Pairs Trading with EG Test

	$\alpha_1 = 0.10$	$\alpha_1 = 0.15$	$\alpha_1 = 0.20$
Total Return	49.1%	39.6%	51.1%
Annualized Return	20.2%	16.9%	20.9%
Annualized Standard Deviation	20.3%	21.6%	19.4%
Annualized Sharpe Ratio	0.99	0.78	1.07
Maximum Drawdown	-9.5%	-14.7%	-12.2%
Return over Maximum Drawdown	5.2	2.7	4.2
Number of Transactions	92	116	150
Transaction Costs over Gross P&L	3.8%	5.9%	5.9%

Pairs Trading with KSS Test

	$\alpha_1 = 0.10$	$\alpha_1 = 0.15$	$\alpha_1 = 0.20$
Total Return	55.1%	34.4%	25.4%
Annualized Return	22.2%	14.9%	11.4%
Annualized Standard Deviation	19.0%	20.4%	20.8%
Annualized Sharpe Ratio	1.17	0.73	0.55
Maximum Drawdown	-5.08%	-12.9%	-15.8%
Return over Maximum Drawdown	10.9	2.66	1.61
Number of Transactions	82	102	136
Transaction Costs over Gross P&L	3.0%	5.9%	10.7%

Buy & Hold Strategy

	Bitcoin Buy & Hold	Portfolio Buy & Hold
Total Return	-33.9%	28.8%
Annualized Return	-17.0%	14.4%
Annualized Standard Deviation	72.6%	98.6%
Annualized Sharpe Ratio	-0.23	0.15
Maximum Drawdown	-77.1%	-82.2%
Return over Maximum Drawdown	-0.44%	0.35

TABLE 5 – Performance Metrics for Different α_1 Values

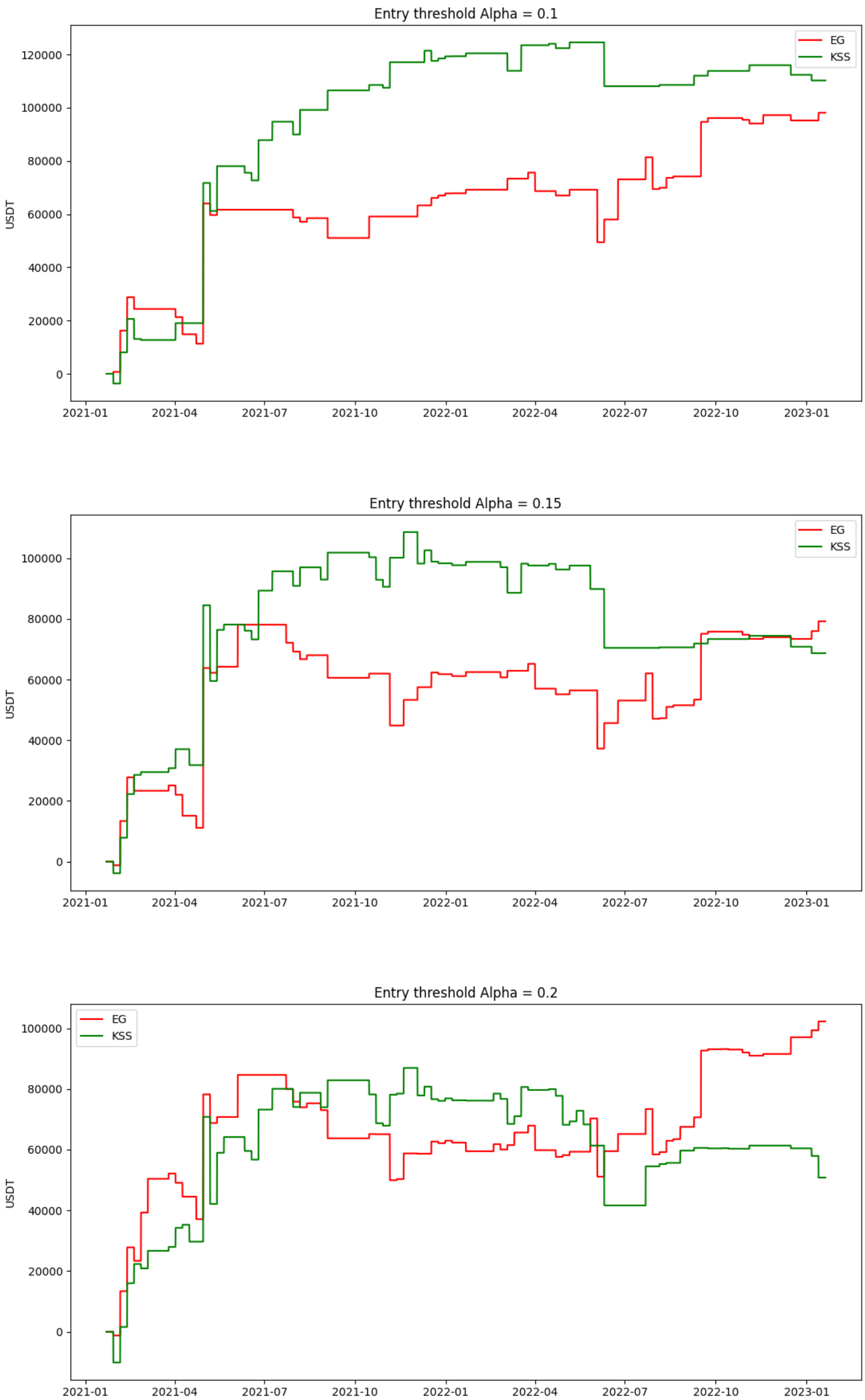


FIGURE 4 – Returns over time for both EG and KSS tests with different entry thresholds

5 Further contributions

After reimplementing the article mentionned, we were very interested in challenging the choices that have been made and the constraints imposed. In this way we found several aspects of the paper that can be questioned.

We undertook a research that deals with the asset of reference for pair trading. As we have seen before, the method involve stationary spread process rather than returns not to have to use the cumulative mispricing index which is not mean-reverting. Then, authors of the article choose bitcoin as a base to evaluate the cointegrated coins and select the traded pair. This choice is not argued in the article but as we have seen previously it could be questionned. Since the bitcoin dominance has suffered a significant drawdown during the studied period, it is relevant to challenge this choice. In order to conclude to the benefit of such a change in our strategy we develop the same approach than previously, simply changing the reference coin to ETH, BNB and LTC as they are the most represented in selected pairs but also because they concentrate a greater capitalisation and then drive the market. We use the KSS cointegration test and threshlod $\alpha_1 = 0.1$ to analyse performances. We found -40.8% overall return for ETH but 28.9% and 10.0% for respectively BNB and LTC. We conclude that BTC remain at a central place in the cryptocurrency market even if its dominance over capitalisation decrease. We also notice that a such pair trading strategy do not garantee positive results as long as we trade on spreads during a relative short period of time. Thus a trading signal can be incorrect and the two traded coins can continue their abnormal behaviour at least untill the end of the trading period producing a loss over the trading week. We finally highlight that number of transactions is significatively higher for these other reference coins (ETH : 166, BNB : 140 and LTC : 107 rather than 82 for BTC). We interpret this phenomenon with a higher volatility of these coins as they are in mean 1.5 times more volatile than BTC on the period. Volatility increase the number of trading signal, considering α_1 fixed, without impoving trade quality. Results are presented in table 6.

Pairs Trading with KSS Test and $\alpha_1 = 0.1$

	ETH	BNB	LTC
Total Return	-40.75%	28.9%	10.0%
Annualized Return	-26.4%	12.2%	20.9%
Annualized Standard Deviation	33.1%	18.9%	19.4%
Annualized Sharpe Ratio	-0.8	0.65	0.22
Maximum Drawdown	-42.0%	-16.7%	-13.1%
Return over Maximum Drawdown	0.97	1.65	0.8
Number of Transactions	166	140	116
Transaction Costs over Gross P&L	-8.1%	9.7 %	23.2%

TABLE 6 – Performance metrics for other reference coins

6 Robustness

The robustness of a trading strategy is crucial to analyse in order to replicate the apporach. That is why we perform some sanity checks throughout our study and confront results we found. The aim here is to ensure that the strategy profitability do not depends on arbitrary parametrization and constraints concerning all step of the implementation.

First of all, regarding cointegration test, we perform two different tests EG and KSS ones and our overall results remain quite equivalent for both of them. Furthermore, to closely replicate the paper's results, we deliberately overfit by selecting six lags for the ADF test. However, this approach can be contested due to its lack of robustness. Therefore, we also utilize the AIC criteria, choosing lags for each regression that minimize noise, aiming for as 'white' noise as possible. Ultimately, the results in terms of total returns are quite similar, with 49.1 versus 43.0 respectively.

Then, we also challenge our strategy on the reference coin on which we test altcoins cointegration. Originally, the method only focuses on BTC but we tried on other central coins as ETH, BNB and LTC to analyse returns resilience on this constraint. We have found really divergent results for ETH and less rewarding ones for BNB and LTC meaning the strategy can not be based on other reference coins even if they share some characteristic like dominance and market capitalisation.

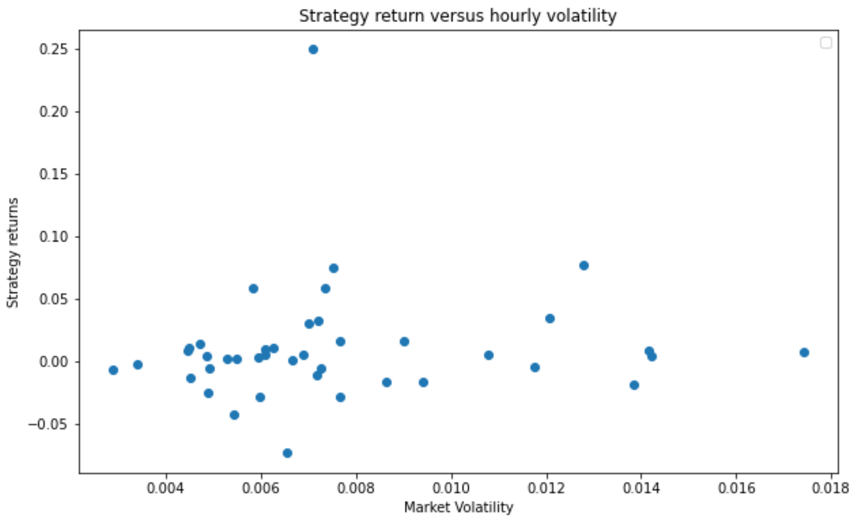
We tried to see how the strategy's performance behaved in relation to the crypto market. We defined a benchmark crypto index that corresponds to the sum weighted by the prices of each crypto.

$$\text{Poids}_{i,t} = \frac{\text{Prix}_{i,t}}{\sum_{j=1}^{20} \text{Prix}_{j,t}}$$

$$\text{Indice}_t = \sum_{i=1}^{20} \text{Prix}_{i,t} \times \text{Poids}_{i,t}$$

We looked at whether our strategy was sensitive to market performance and volatility. Regardless of the test used, the relationship between the performance of our strategy and market performance appears to be independent. In bullish and bearish periods, the strategy's performance is not affected. We therefore manage to generate an alpha.


Strategy return versus market returns
Strategy returns
Market returns
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3


Strategy return versus hourly volatility
Strategy returns
Market Volatility
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
0.004
0.006
0.008
0.010
0.012
0.014
0.016
0.018

7 Conclusion

This paper contributes to the field of quantitative crypto trading by developing a unique pairs trading framework, built on state-of-the-art technics and combining different methodologies to model dependency relationships. The approach integrates copula and cointegrated based strategies for twenty cryptocurrencies. The research focuses on identifying coins cointegrated with a reference asset (BTCUSDT) using both the EG two-step method and the KSS cointegration tests. After ranking the cointegrated coins based on Kendall's Tau correlation coefficient, we select the first two assets for trading, updated weekly.

The methodology implemented in this study is inspired by M. TADI and J. WITZANY's 2023 paper. Unlike traditional pair trading approaches, this paper introduces a novel aspect by incorporating stationary spread process in copula component thus the mispricing index contain historical information on relative price from one coin to the other. Cointegration also reinforce the mean-reverting signature of generated trading signals and confirm the real contribution of the paper to the domain.

The research presented aims to re-implement and compare the entire strategy of the referenced article but also to evaluate the robustness of the approach. Factors such as the training period lenght, the reference asset, and optimizing α_1 via back-testing are analyzed for their impact on returns and overall results. The study concludes by determining the most profitable strategy and configuration and its robustness concerning the different steps.

Findings highlight the significance of choosing an appropriate entry threshold in pairs trading, influencing both volatility and returns. The proposed pairs trading strategy exhibits a high risk-adjusted return and return over maximum drawdown compared to a buy-and-hold approach, suggesting its effective utilization for profitability. The study underscores the effectiveness of pairs trading in the cryptocurrency market and emphasizes the importance of meticulous coin pair and copula model selection.

We conclude this study with some further ideas that would certainly increase the profitability of the strategy. First, a real back-testing to optimize both α_1 and α_2 independently of market conditions (especially volatility) would be relevant. Then, to challenge the training and trading period lenghts as coins can be more cointegrated on different timescales. And finally, we were also curious to know if the adds of other marginal distribution to fit the spread between bitcoin and altcoins would have change the overall results

Références

[1] Tadi and Witzany 2023, *Copula-Based Trading of Cointegrated Cryptocurrency Pairs*, https://arxiv.org/abs/2305.06961

[2] Hansen Pei, *Copula for Pairs Trading : A Unified Overview of Common Strategies*, https://hudsonthames.org/copula-for-pairs-trading-overview-of-common-strategies

[3] Myriam Chabot 2013, *Concepts de dépendance et copules*, https://camus.espaceweb.usherbrooke.ca/revue/revue4/chabot.pdf

[4] Tsung et al. 2007, *Finite mixture modelling using the skew normal distribution*, https://www.jstor.org/stable/24307705?seq=3

[5] Hélène Hamisultane 2002, *Modèle à correction d'erreur (MCE) et applications*, https://shs.hal.science/cel-01261167/document

[6] Hendry and Juselius 1999, *Explaning cointegration analysis : part 1*

A Appendix

Symbol	Cryptocurrency
ADA	Cardano
ATOM	Cosmos
BAT	Basic Attention Token
BCH	Bitcoin Cash
BNB	Binance Coin
BTC	Bitcoin
DASH	Dash
EOS	Eos.io
ETC	Ethereum Classic
ETH	Ethereum
IOTA	Iota
LINK	Chainlink
LTC	Litecoin
ONT	Ontology
TRX	Tron
XLM	Stellar
XMR	Monero
XRP	Ripple
XTZ	Tezos
ZEC	Zcash

TABLE 7 – Cryptocurrencies and their symbols

Week	Coin1	Coin2	Pval_Coin1	Pval_Coin2
1	ETH	LTC	0.094	0.074
2	LTC	BCH	0.028	0.002
3	ETH	LTC	0.039	0.028
4	ETH	LTC	0.087	0.022
5	LTC	EOS	0.074	0.071
6	LINK	TRX	0.092	0.073
7	LINK	TRX	0.099	0.047
8	ETH	LTC	0.003	0.016
9	ETH	LTC	0.037	0.019
10	ETH	BCH	0.093	0.065
11	LTC	BCH	0.083	0.003
12	ATOM	ADA	0.013	0.008
13	0	0	0.0	0.0
14	ADA	EOS	0.035	0.008
15	EOS	LTC	0.045	0.021
16	BNB	BAT	0.049	0.018
17	BNB	TRX	0.052	0.019
18	ETH	BCH	0.096	0.068
19	ETH	TRX	0.046	0.054
20	LINK	ETC	0.083	0.044
21	TRX	LTC	0.007	0.013
22	0	0	0.0	0.0
23	0	0	0.0	0.0
24	ETH	BNB	0.056	0.069
25	LTC	EOS	0.02	0.03
26	ETH	LTC	0.086	0.004
27	BNB	DASH	0.05	0.028
28	ETH	LINK	0.079	0.069
29	BNB	LTC	0.052	0.008
30	BNB	EOS	0.032	0.026
31	ETH	LINK	0.025	0.013
32	ETH	ONT	0.054	0.05
33	LTC	XRP	0.063	0.023
34	EOS	XRP	0.05	0.046
35	ETH	EOS	0.061	0.036
36	XRP	LINK	0.097	0.07
37	ETH	BNB	0.051	0.032
38	ETH	LINK	0.008	0.031
39	DASH	XLM	0.065	0.0
40	LTC	BNB	0.087	0.088
41	BCH	ETC	0.095	0.087
42	BCH	LTC	0.007	0.028
43	ETH	ETC	0.082	0.002
44	ETH	ETC	0.038	0.006
45	EOS	ETC	0.029	0.004

Continued on next page

Table 8 – *Continued from previous page*

Week	Coin1	Coin2	Pval_Coin1	Pval_Coin2
46	EOS	ETC	0.03	0.001
47	ETC	XRP	0.056	0.097
48	LTC	DASH	0.079	0.01
49	ETH	ETC	0.037	0.007
50	ETH	ETC	0.037	0.002
51	ETH	ETC	0.071	0.009
52	LTC	EOS	0.004	0.03

Table 8 – Selected pairs and p-values for each trading week (Page 1)

Week	Coin1	Coin2	Pval_Coin1	Pval_Coin2
53	EOS	XRP	0.018	0.001
54	0	0	0.0	0.0
55	XLM	BAT	0.031	0.081
56	ETH	BNB	0.048	0.02
57	ETH	BNB	0.01	0.016
58	ETH	BNB	0.024	0.036
59	ETH	BNB	0.029	0.09
60	ETH	BNB	0.023	0.022
61	ETH	BNB	0.089	0.01
62	BNB	ONT	0.007	0.073
63	BNB	LTC	0.075	0.018
64	LINK	XLM	0.026	0.03
65	XRP	XLM	0.003	0.091
66	ETH	BNB	0.001	0.03
67	ETH	BNB	0.039	0.011
68	ETH	BNB	0.013	0.001
69	0	0	0.0	0.0
70	BNB	ETC	0.017	0.051
71	LINK	ETC	0.091	0.056
72	LINK	EOS	0.044	0.027
73	DASH	ETC	0.023	0.043
74	DASH	IOTA	0.076	0.036
75	ZEC	BCH	0.04	0.044
76	ETH	ZEC	0.002	0.089
77	ETH	BNB	0.071	0.06
78	ETH	BNB	0.013	0.003
79	BNB	LTC	0.013	0.005
80	BNB	LTC	0.013	0.058
81	ETH	LTC	0.008	0.008
82	LTC	ADA	0.012	0.002
83	LTC	DASH	0.002	0.041
84	ETH	LTC	0.032	0.001
85	ETH	IOTA	0.083	0.019
86	BAT	IOTA	0.0	0.003
87	BNB	ETC	0.02	0.069
88	BNB	TRX	0.014	0.07
89	DASH	TRX	0.003	0.031
90	ETH	DASH	0.002	0.013
91	ETH	BAT	0.066	0.019
92	ETH	BAT	0.002	0.082
93	BNB	LTC	0.017	0.028
94	BCH	DASH	0.008	0.049
95	ETH	BCH	0.049	0.016
96	ADA	DASH	0.005	0.026
97	ETH	IOTA	0.095	0.011

Continued on next page

Table 9 – *Continued from previous page*

Week	Coin1	Coin2	Pval_Coin1	Pval_Coin2
98	ADA	ETC	0.037	0.038
99	XTZ	BAT	0.001	0.012
100	ETH	ONT	0.008	0.036
101	0	0	0.0	0.0
102	ETH	BNB	0.075	0.044
103	BNB	LINK	0.007	0.024
104	LINK	DASH	0.068	0.094

TABLE 9 – Selected pairs and p-values for each trading week (Page 2)

Week	Coin1	Coin2	t-stat_Coin1	t-stat_Coin2
1	BCH	BNB	-2.526	-2.462
2	LTC	BCH	-3.513	-1.967
3	LTC	ETC	-3.604	-7.78
4	ETH	LTC	-2.454	-2.199
5	LTC	BCH	-2.958	-3.054
6	LTC	BCH	-2.831	-2.713
7	LTC	BCH	-2.841	-2.524
8	ETH	LTC	-2.809	-2.012
9	LTC	BCH	-3.619	-2.174
10	LTC	BCH	-2.939	-2.443
11	LTC	BCH	-2.592	-2.636
12	BCH	ATOM	-2.123	-2.781
13	ADA	XTZ	-3.244	-2.013
14	LTC	ADA	-2.067	-3.742
15	EOS	LTC	-2.386	-2.173
16	TRX	BAT	-2.661	-4.042
17	BNB	TRX	-3.007	-2.925
18	TRX	IOTA	-3.091	-2.589
19	BCH	TRX	-2.7	-2.577
20	BCH	XTZ	-2.923	-2.181
21	BCH	XRP	-2.508	-3.101
22	BNB	ATOM	-1.926	-2.028
23	BNB	LINK	-2.24	-3.528
24	ETH	BNB	-2.097	-2.722
25	LTC	BNB	-2.34	-2.327
26	LTC	XRP	-2.696	-4.406
27	XRP	DASH	-2.282	-2.646
28	ETH	BCH	-1.958	-2.617
29	BNB	LTC	-2.437	-2.658
30	BNB	EOS	-2.22	-3.786
31	LTC	LINK	-2.389	-3.086
32	ETH	LTC	-2.498	-2.674
33	LTC	BNB	-2.742	-2.029

Continued on next page

Table 10 – *Continued from previous page*

Week	Coin1	Coin2	t-stats_Coin1	t-stats_Coin2
34	EOS	LTC	-2.441	-2.049
35	EOS	BNB	-2.22	-4.912
36	ETC	XRP	-2.065	-2.095
37	ETH	BNB	-2.83	-2.11
38	ETH	BCH	-1.934	-2.464
39	ETH	XRP	-1.945	-2.318
40	DASH	BNB	-2.191	-2.094
41	BCH	ETC	-2.077	-1.997
42	BCH	LTC	-3.164	-2.07
43	ETC	EOS	-3.854	-2.448
44	ETH	ETC	-2.709	-3.158
45	EOS	ETC	-4.464	-3.067
46	EOS	ETC	-2.337	-3.556
47	ETC	XRP	-2.66	-2.688
48	LTC	DASH	-3.27	-4.347
49	ETH	ETC	-2.287	-2.125
50	ETH	ETC	-3.318	-3.029
51	ETH	ETC	-2.446	-3.421
52	LTC	EOS	-3.073	-2.641

Table 10 – Selected pairs and t-stats for each trading week (Page 1)

Week	Coin1	Coin2	t-stat_Coin1	t-stat_Coin2
53	EOS	XRP	-2.454	-3.526
54	LTC	BNB	-2.045	-2.064
55	EOS	LTC	-2.253	-2.114
56	ETH	LTC	-2.494	-2.83
57	ETH	BNB	-2.871	-3.164
58	ETH	BNB	-2.853	-3.401
59	ETH	BNB	-3.397	-2.808
60	ETH	BNB	-2.453	-2.991
61	BNB	ADA	-2.99	-2.246
62	ETH	BNB	-2.021	-2.781
63	BNB	LTC	-1.958	-3.517
64	XRP	LINK	-2.12	-2.412
65	XRP	XLM	-3.556	-2.654
66	ETH	BNB	-3.666	-3.012
67	ETH	BNB	-3.292	-2.827
68	ETH	BNB	-3.471	-3.404
69	BCH	XLM	-2.059	-2.072
70	BNB	ETC	-2.372	-2.805
71	ETC	BNB	-2.309	-2.43
72	LTC	BNB	-1.933	-4.402
73	EOS	DASH	-2.322	-3.288
74	ETH	BNB	-2.064	-2.537
75	ETH	IOTA	-3.209	-1.942
76	ETH	ZEC	-2.923	-3.0
77	ETH	BNB	-2.75	-2.475
78	ETH	BNB	-3.207	-2.362
79	BNB	LTC	-2.894	-3.309
80	BNB	ZEC	-3.016	-2.507
81	LTC	LINK	-3.306	-2.817
82	LTC	ADA	-2.93	-3.402
83	LTC	DASH	-4.118	-2.663
84	ETH	LTC	-2.853	-3.898
85	IOTA	LINK	-2.631	-2.092
86	BAT	IOTA	-3.739	-3.005
87	BAT	ETC	-2.487	-2.698
88	DASH	BAT	-2.399	-2.788
89	ETH	DASH	-2.62	-3.535
90	ETH	DASH	-2.641	-3.012
91	ETH	DASH	-2.879	-2.288
92	ETH	BAT	-3.041	-2.178
93	BNB	LTC	-2.373	-3.478
94	BCH	DASH	-4.599	-2.703
95	BCH	DASH	-7.665	-2.556
96	ETH	ADA	-2.355	-2.722
97	ETH	ADA	-2.55	-2.529

Continued on next page

Table 11 – *Continued from previous page*

Week	**Coin1**	**Coin2**	**t-stat_Coin1**	**t-stat_Coin2**
98	ETH	ADA	-2.116	-2.651
99	ETH	ATOM	-1.922	-2.044
100	ETH	BCH	-2.097	-2.021
101	LINK	BCH	-2.093	-2.234
102	ETH	LTC	-2.267	-2.133
103	BNB	LINK	-3.834	-2.789
104	LINK	ONT	-2.621	-2.656

Table 11 – Selected pairs and t-stats for each trading week (Page 2)

Copulas Type	**Distribution**
Elliptical	Gaussian Student-t
Archimedean	Clayton Frank Joe BB1 BB6 BB7 BB8
Extreme-Value	Tawn type 1 Tawn type 2
Archimax	Gumbel

Table 12 – Families of Copulas and Distribution used