A network-based strategy of price correlations for optimal cryptocurrency portfolios

Ruixue Jing[a], Luis E C Rocha[a,b]

[a]Department of Economics, Ghent University, Ghent, Belgium
[b]Department of Physics and Astronomy, Ghent University, Ghent, Belgium

Abstract

A cryptocurrency is a digital asset maintained by a decentralised system using cryptography. The complex correlations between the cryptocurrencys' prices may be exploited to understand the market dynamics and build efficient investment portfolios. We use network methods to select cryptocurrencies and the Markowitz Portfolio Theory to create portfolios agnostic to future market behaviour. The performance of our network-based portfolios is optimal with 46 cryptocurrencies and superior to benchmarks for short-term investments, reaching up to $1,066\%$ average expected returns within 1 day. Cryptocurrency portfolio investment may be competitive but calls for caution given the high variability of prices.

Keywords: Cryptocurrency, Network Model, Portfolio, Optimisation, Price Correlation, Financial Market

1. Introduction

Cryptocurrencies are digital currencies using cryptography to validate financial transactions. The lack of governmental regulations and the intangible nature of cryptocurrencies lead to price volatility and thus scepticism on such currencies as financial assets (Jacobs, 2018). Complex systems such as the crypto market are expected to contain underlying mechanisms driving the dynamics of prices that can be exploited to extract information to understand market behaviour (Hong and Yoon, 2022). This happens because cryptocurrencies' prices and investors' sentiments are fundamentally interconnected. To what extent cryptocurrency prices follow rational decisions or gambling is unclear. The short-term memory of cryptocurrency prices makes price predictions challenging, and consequently, standard strategies of portfolio-based investment unsuitable (Chen, 2023; Latif et al., 2023).

Research on the stock market proposes to map the correlations between asset prices into distance correlation networks. The minimum spanning tree (MST) of such networks contains the least correlated stocks in its periphery in contrast to highly correlated stocks in the centre of the tree (Onnela et al., 2003). Using the Markowitz or modern portfolio theory (MPT) (Danko and Šoltés, 2018), the least correlated assets can be selected to build optimal portfolios. MSTs were also used to develop a stock market filtering method to show that the whole market can be simulated by this reduced version (Esfahanipour and Zamanzadeh, 2015). Attempts to apply such techniques outside the traditional capital market are limited. Brauneis and Mestel (2019) reported that MPT portfolios could obtain higher Sharpe ratios than a single cryptocurrency. Their study focused on top cryptocurrency like Bitcoin and used an equally-weighted cryptocurrency asset class to Markowitz's mean-variance framework to derive the risk-adjusted out-performance. It remains unclear what optimal weights and the number of cryptocurrencies must be included to make efficient portfolios. Another study showed that the Foster-Hart risk optimisation and the generalised auto-regressive conditional heteroscedasticity (GARCH) model with only four major cryptocurrencies could produce a more profitable portfolio than the traditional equally weighted portfolio method (Kurosaki and Kim, 2022). Bitcoin has been one of the highest-performing assets for profit, and thus the focus of previous research. Portfolio research also usually considers just a few cryptocurrencies and ignored less popular currencies with potential high gains despite the associated risks. Xie et al. (2021) showed the economic benefits and costs of adding safe-haven cryptocurrencies (like Tether) and traditional assets (like gold) to portfolios to outperform naked portfolios (with only Bitcoin). Furthermore, cryptocurrency's prospect has been shown to undermine gold's hedging ability (Su et al., 2020).

Network-based methods considering similarity and technology indicators have been tested for price prediction of individual cryptocurrencies (Lucchini et al., 2020; Zhong et al., 2023). The advantage of network modelling is to simultaneously include all available assets and their relations in a single framework that can be globally optimised. The main limitation of previous research is that the highly fluctuating nature of cryptocurrencies' prices makes the prediction task and thus price-based portfolio design difficult (Zhong et al., 2023; Latif et al., 2023). A recent literature review also pointed out that portfolios containing only cryptocurrencies may be too risky but acknowledges that studied sample sizes are small and diversification strategies must be further analysed (Almeida and Gonçalves, 2023). We hypothesise instead that the interdependences between cryptocurrencies, measured using pair-wise price correlations, fluctuate less than individual prices, i.e. the correlation between currencies' prices are more persistent than individual prices and thus can be exploited to estimate currencies' performance. We thus propose a network-based methodology able to simultaneously map all pair-wise price correlations and minimise the global correlation in a computationally efficient and temporally agnostic way. Thus, past correlations can be used to estimate future returns in uncertain scenarios without the need to predict individual prices. The rationale is that maximising diversity, by minimising the total correlation between portfolio's assets, reduces risk and increases returns. We test our hypothesis by developing a method to find the least correlated cryptocurrencies within a large pool of currencies (more than 5,000), building optimal portfolios using a multi-step procedure that involves selecting the least correlated currencies using network methods and then optimising the portfolio using MPT, and comparing their performance with benchmarks.

2. Materials and Methods

2.1. Modern Portfolio Theory

A fundamental issue in building investment portfolios is optimally allocating funds (Sulistiawan and Rudiawarni, 2017). The number of assets and the amount of each asset are difficult to estimate. The Modern Portfolio Theory (MPT) (Fabozzi et al., 2008) provides a framework to estimate optimal amounts (i.e. weights) of each asset to construct diversified portfolios able to maximise returns for a given level of risk. It is based on the variance and correlation of the assets in the portfolio. Though cryptocurrencies are not financial instruments, price correlations are crucial for portfolio optimisation and diversification. The optimal weight for each cryptocurrency i ($i = 1, 2, ..., n$) in a portfolio with n currencies is obtained by solving the optimisation problem:

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}\Sigma\mathbf{w} \tag{1}$$

subject to:

$$\mathbf{m}^T\mathbf{w} \geq \mu_0 \quad ; \mathbf{e}^T\mathbf{w} = 1 \tag{2}$$

where $\mathbf{w} = (w_1, ..., w_i ..., w_n)$ is a vector of weights for cryptocurrency i, and Σ is the covariance matrix of the cryptocurrencies' returns. The set $\mu_i = E(r_i)$, $\mathbf{m} = (\mu_1, ..., \mu_i, ..., \mu_n)$ is a vector of expected returns for each cryptocurrency, $\mathbf{e}$ denotes the vector of ones, and μ_0 the acceptable expected return.

The expected return of the portfolios can then be calculated by:

$$r_{\text{portfolio}} = \sum_{i=1}^{n} w_i r_i \tag{3}$$

The portfolio volatility is calculated using the covariance matrix and the above-described portfolio weights.

$$\sigma_{\text{portfolio}} = \sqrt{\mathbf{w}^T \cdot \Sigma \cdot \mathbf{w}} \tag{4}$$

The software Python is used to estimate the optimal weights using the PyPortfolioOpt module.

2.2. *Network-based Portfolios*

Our network methodology aims to find the least correlated cryptocurrencies to be included in a portfolio to achieve diversification in a computationally efficient way. The first step maps all pair-wise correlations in a network and finds a globally optimal hierarchical structure of the relative correlations between currencies. The second step produces a ranking of the (globally) least correlated cryptocurrencies. Let $P_{i,t}$ be the price (in USD) of cryptocurrency i at time t, with $i = 1, 2, ...N$ and $t = 1, 2, ..., t_w$ in days. The log return $r_{i,t} = \ln(P_{i,t}/P_{i,t-1})$ is used to obtain a stationary time-series. The Pearson correlation coefficient between the log returns of i and j over the period Δt is:

$$\rho_{ij} = \frac{Cov(r_i, r_j)}{\sigma_i \sigma_j} \tag{5}$$

The correlation coefficient is transformed into a distance metric (Mantegna, 1999) using:

$$d_{ij} = 2\sqrt{(1 - \rho_{ij})} \tag{6}$$

A network $G(N, E)$ is defined as a set of N nodes $i = \{1, 2, 3, ..., N\}$ and a set of E links connecting those nodes (Costa et al., 2011). In our model, a node i represents one cryptocurrency, and the distance correlation d_{ij} corresponds to a weighted-link (i, j) connecting the respective nodes. The distance for all pairs of cryptocurrencies is mapped to a fully connected weighted network. The Minimum Spanning Tree (MST) is a filtering algorithm used to extract an optimal weighted acyclic structure with N connected nodes, such that the sum of the weights of all links is less than or equal to the total link weight of every other possible spanning tree (Kruskal, 1956; Tarjan, 1982). The MST thus returns a globally optimal hierarchical structure maximising the correlations between all pairs of cryptocurrencies (i.e. minimising the total distance correlation). We then calculate the centrality L_i of cryptocurrency i, that is the average distance between i and any other cryptocurrency j in the MST:

$$L_i = \frac{\sum_{i \neq j, j=1}^{N} l(i, j)}{N - 1} \tag{7}$$

where $l(i, j)$ is the shortest path (in terms of weighted-links) between the respective nodes. The centrality-based rank of a cryptocurrency indicates the level of decorrelation with other currencies, i.e. cryptocurrencies with higher L_i are (globally) least correlated to any other cryptocurrency (farther apart in the correlation-distance space) than those with lower L_i. The portfolio (MSTn) is then built by adding n cryptocurrencies according to their decreasing order of importance to guarantee maximum diversification, and the weights are calculated using the MPT (Section 2.1).

We developed four benchmark models: (i) an entirely agnostic model where n cryptocurrencies are chosen uniformly at random from a pool of N cryptocurrencies (RANDn); (ii) by selecting n currencies according to increasing order of network centrality L_i (BMmstn); (iii) the top five cryptocurrencies with the largest average market capitalisation throughout the study period, i.e. Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC), Tether (USDT) and XRP (TOP5); (iv) only Bitcoin (BTC), which is the market leader.

2.3. *Cryptocurrency Price Data*

The data set consists of the daily market price (at midnight) of 5,450 cryptocurrencies collected from several online sources (See SI). We extracted the top 1,000 currencies with the highest market cap on 22.02.2022 and then those active from 15.10.2017 to 15.10.2022 ($T_{\text{total}} = 1,827$ days). Currencies with missing prices for more than ten days are also excluded. The final data set contains time series of prices (in USD) for $N = 157$ cryptocurrencies (Table S1).

3. Results

3.1. *Distance Correlation Matrix*

Table 1 shows the descriptive statistics of the log-transformed return price ($r_{i,t}$) for the top 5 (BTC: Bitcoin; ETH: Ethereum; LTC: Litecoin; USDT: Tether; XRP) and bottom 5 (OCN: Odyssey; DLT: Agrello; ENG: Enigma; ETP:

	BTC	ETH	LTC	USDT	XRP	OCN	DLT	ENG	ETP	FUEL
n	1826	1826	1826	1826	1826	1826	1826	1826	1826	1826
Mean	0.001	0.001	0.000	0.000	0.000	-0.005	-0.003	-0.003	-0.003	-0.003
Std deviation	0.041	0.053	0.057	0.004	0.065	0.286	0.093	0.087	0.076	0.360
Minimum	-0.480	-0.570	-0.466	-0.057	-0.541	-9.954	-0.960	-0.769	-0.829	-2.943
Median	0.001	0.001	-0.001	0.000	0.000	-0.004	-0.002	-0.003	-0.003	-0.003
Maximum	0.203	0.246	0.426	0.053	0.618	6.099	0.809	0.889	0.591	2.920
Skewness	-0.884	-0.858	-0.089	0.049	0.901	-17.687	-0.298	0.420	-0.411	-0.130
Kurtosis	11.994	9.842	9.054	77.783	15.878	910.374	14.945	14.048	15.464	33.991
ADF test	-29.573	-29.102	-16.278	-10.841	-29.258	-74.698	-13.064	-46.035	-29.398	-8.376

Table 1: Descriptive statistics and augmented Dickey–Fuller (ADF) test of log-transformed returns ($r_{i,t}$) taken at n times (days) for the top 5 and bottom 5 cryptocurrencies according to their average market capitalisation.

Metaverse; FUEL: Etherparty) cryptocurrencies according to their average market capitalisation during the study period (See Table S2 for the correlation distance between the same pairs of cryptocurrencies).

To build the portfolios, the distance correlation $d_{ij}(\tau)$ is calculated in intervals of $\Delta t = 30$ days shifted by $\tau = 1$ day at a time (rolling window) to capture the fluctuations of the correlation networks (Rocha et al., 2017). Portfolio estimates are independent of τ, and Δt is chosen to capture the short-term heterogeneity of correlations; a larger Δt reduces the statistical power of the estimates (See SI). Figure 1C shows that the average distance correlation $\langle d \rangle = \sum_\tau d_{ij}(\tau)$ fluctuates between 1 and 2, indicating an overall positive correlation in the crypto market. At turbulent periods, for example, the Bitcoin crash in January 2018 and the covid-19 outbreak from February to March 2020, $\langle d \rangle$ drops significantly, indicating a strong correlation between cryptocurrencies. Figure 1B shows that pairwise correlations are highly heterogeneous (right-skewed distribution) during critical events in contrast to a normal distribution during quiet periods (Fig.1A).

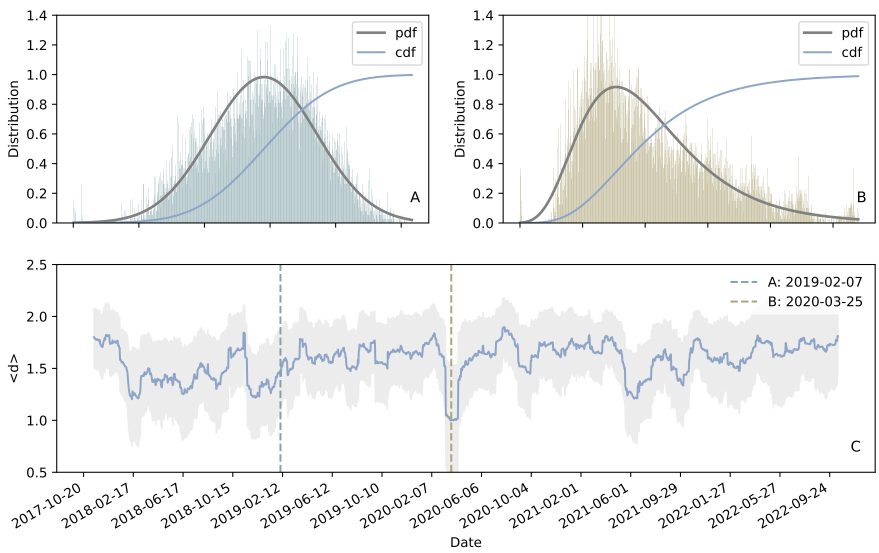

Figure 1: A) and B) show the distribution of distance correlations, respectively, during a quiet (fitted by the normal distribution, p-value < .01, $\mu = 1.45$, $\sigma = 0.41$) and a turbulent period (fitted by the log-normal distribution, p-value < .01, $\mu = -0.24$, $\sigma = 0.40$, $s = 1.19$) (See panel C) for the selected dates)). C) shows the evolution of the average distance correlation $\langle d \rangle$ during the study period and two selected dates. The shaded area shows the standard error.

3.2. *Performance Network-based Portfolio*

The first step when designing a portfolio is identifying the optimal number of assets. While diversification is essential, over-diversification increases management costs, may reduce significant gains, and will not eliminate unsystematic risk (Merker et al., 2019). To estimate the optimal number of cryptocurrencies, we calculate a single MST for a given reference time τ (see SI) and select n currencies for the portfolio, following our proposed ranking method. Once n is defined, we apply the MPT to estimate the optimal weights of each currency in the portfolio for that time τ. The training period is $[\tau - \Delta t, \tau]$. Figures 2A,B show respectively the average expected return $\langle r_{\mathrm{MST}} \rangle$ and average expected risk $\langle \sigma_{\mathrm{MST}} \rangle$, where the average is taken over all times τ to include quiet and turbulent periods in the performance assessment. This first model is only based on the training period as thus used as a reference portfolio. The portfolio achieves high levels of return with associated low risks; the return reaches a plateau, whereas the risk decreases with an increasing number of cryptocurrencies.

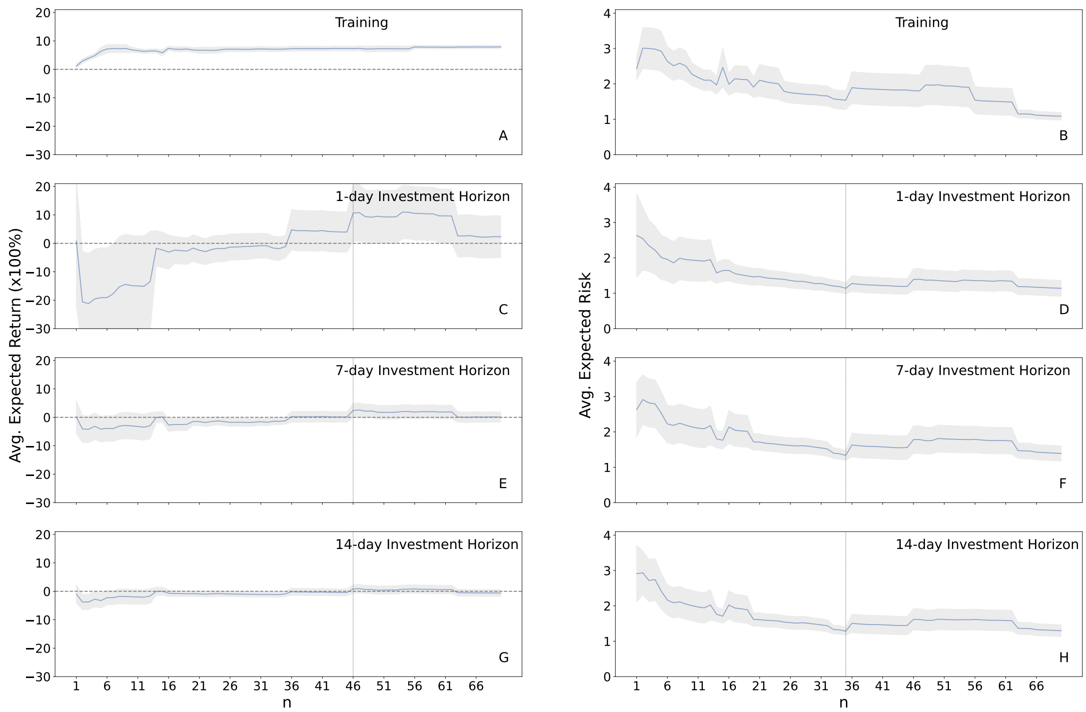

Figure 2: Average expected returns $\langle r_{\mathrm{MST}} \rangle$ and average expected risk $\langle \sigma_{\mathrm{MST}} \rangle$ for portfolios with n cryptocurrencies during the A,B) training and C-H) future periods. The shaded area shows the standard error. The dashed vertical lines highlight the optimal number of currencies for maximum return and minimum risk.

We now assess the performance of the same portfolio for future investments. We build a portfolio using data from the period $[\tau - \Delta t, \tau]$ and evaluate that same portfolio in the period $[\tau, \tau + t_{horizon}]$. In other words, we implicitly assume that prices and correlations between cryptocurrencies remain constant. Our optimal portfolio is agnostic to future events because we only use prices from the training period. Figure 2C shows that the portfolio performs poorly in the future (1 day) for less than $n = 35$ currencies but achieves the highest $\langle r_{\mathrm{MST}} \rangle$ with $n = 46$ currencies. Figure 2D shows that $\langle \sigma_{\mathrm{MST}} \rangle$ decreases by adding more currencies, and although it is not minimum for $n = 46$, it is comparatively lower than other cases. The performance decreases for longer investment horizons (see SI for all horizons), but the peaks of $\langle r_{\mathrm{MST}} \rangle$ were consistently at $n = 46$ (Figs. 2E,G). The lowest $\langle \sigma_{\mathrm{MST}} \rangle$ was achieved with $n = 35$ currencies (with slightly negative average returns), yet, not significantly lower than the case of $n = 46$ currencies, suggesting a good compromise return vs risk.

Figures 3A,B compares our network-based portfolio to the four benchmarks regarding average return and risk, respectively. The performance of our portfolio is always superior, achieving more than 1,000% average expected return for a 1-day horizon. The peak of $\langle r_{\mathrm{MST46}} \rangle$ at 6 days suggests weekly cycles in the crypto market requiring

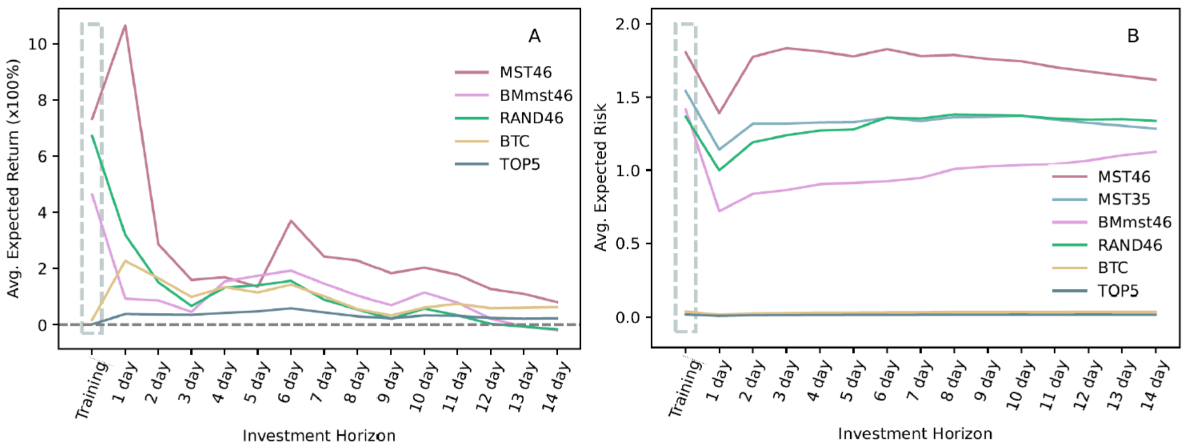

Figure 3: A) shows the average expected return and B) the average expected risk for our network-based (MST46) portfolio with $n = 46$ currencies and benchmark portfolios for various investment time horizons.

further research. The performance of the RAND46 portfolio (i.e. formed by 46 cryptocurrencies selected randomly, see Section 2.2), is lower and equivalent to a solo Bitcoin (BTC) investment. TOP5 gives the lowest expected return in the two weeks. Our portfolio with $n = 46$ currencies consistently shows the highest average risk but is not substantially higher than other benchmarks.

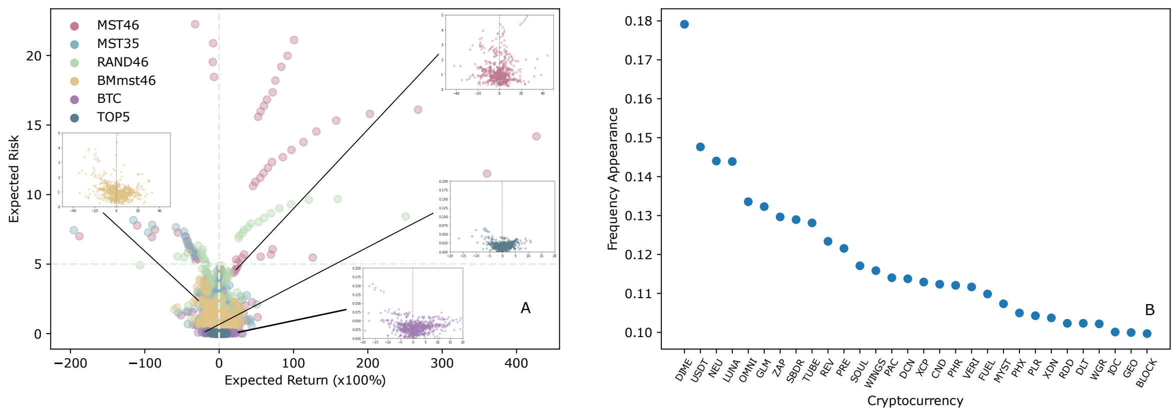

Figure 4: A) shows the expected return and risk for various portfolios during different training periods for each investment horizon (1 to 14 days, 840 points per portfolio model). Our network-based optimal portfolio (MST46) reaches the highest and second highest expected returns of 42,674% and 36,036% with the investment horizon of 1 day for two different training periods. The insets zoom in the performance of the portfolios for low expected return and risk. B) shows the ranking of the top 30 cryptocurrencies based on the frequency of appearance in the network-based optimal portfolios with $n = 46$ currencies (MST46).

The relative reduction in risk is disproportional to the decrease in the returns. An increase in the investment horizon from 1 day to 2 days (training period with the highest return) decreases the return of our portfolio by 37.28% and the risk by 13.51%. The best (1-day) and the worst (14 days) investment horizons show a relative decrease of 92.40% in returns, whereas the risk decrease by only 15.83%. In contrast, for the benchmark TOP5, the worst investment horizon decreases the return by 62.71%, whereas the risk increases by 5.88%. Conservative portfolios give positive returns with low risk, but aggressive investors may consider high returns associated with increased risks (Fig. 4). The benchmarks TOP5, BTC, and BMmst46 provide the lowest risks, generating positive returns, respectively, in 59.3%, 57.8%, and 55.5% of the agnostic periods (for all investment horizons); that compares to 53.2% in our portfolio (MST46).

The most frequent currencies in our optimal portfolios (MST46) appear in at least 9.97% of the times, with Dimecoin (DIME) included 17.91% of the times (Fig. 4B). Although DIME is not a top market-value cryptocurrency, it is useful for everyday micro-transactions due to its relatively fast block time (i.e. transactions can be confirmed quickly). Except for USDT (14.8%), the other cryptocurrencies included in our benchmarks are not in the top 30. BTC is used 3.3% of the times, whereas ETH, LTC, and XRP are used, respectively, 7.1%, 1.1%, and 1.8%. USDT is a cryptocurrency designed to maintain a stable (one-to-one) value relative to USD, and thus not recommended for speculative investment (see SI for all frequencies).

3.3. *Financial Analysis*

Although the crypto market is relatively young compared to the stock and exchange markets, it has played an increasingly important role in the financial system, calling the attention of central banks and often reflecting the bilateral relations between countries (Qin et al., 2022). Investors with different profiles have joined the crypto market, but speculative investment drives the highly fluctuating prices. The capital market efficiency hypothesis indicates that new information in the market is immediately reflected in asset prices. Neither technical nor fundamental analysis can generate excess returns. All stocks trade at their fair market value, which means only inside information can result in outsized risk-adjusted returns (Fama, 1970). Investors can only diversify the unsystematic risk associated with the corresponding cryptocurrency's risk (e.g., technology and credibility). Efficient disclosure policy can also lessen information asymmetry in the capital market and keep returns from being too volatile (Gharbi et al., 2014). The inherent (systematic) market risk cannot be eliminated (associated with regulatory and macroeconomic forces). It will affect all currencies (Jalan and Matkovskyy, 2023), calling for portfolio diversification as a risk-reducing mechanism. The performance could be further improved by using dynamic hedging techniques, as proposed for traditional markets (Antonakakis et al., 2020; Adekoya et al., 2022).

Investors must balance their portfolios based on investment goals and risk tolerance. The diversification theory states that efficient strategies require allocating mixed products to offset the loss of one asset (Chatterjee and Wernerfelt, 1991). While there is no one-size-fits-all, portfolios in traditional markets usually include more than ten stocks (Evans and Archer, 1968) but may contain hundreds of stocks (e.g., combined holdings of mutual funds), observing higher risk for short-term investment (Stewart, 2019). Our portfolio achieves the best performance in the short-term with $\sim 25\%$ of the total number of available cryptocurrencies ($N = 157$). This result may indicate a more general relationship where potentially an optimal number of cryptocurrencies exists. Increasing the number of assets beyond that point accumulates systematic risks to the detriment of a potential return increase. The high short-term return indicates that the market is partially uncorrelated. A dynamic weighting strategy could be used to rebalance portfolios regularly (Antonakakis et al., 2020). The crypto market is not immune to contagion risk (da Gama Silva et al., 2019), where correlations increase with exogenous shocks (Broadstock et al., 2012), but also depend on technological aspects (Lucchini et al., 2020). Such correlations could be exploited via the distance correlation network (Allen and Babus, 2009) to improve estimators using multivariate regression (Chamberlain, 1982) instead of univariate temporal models (Murray et al., 2023).

There is room for profit when trading in this volatile market (Balcilar et al., 2017) but not without risks. Most investors in the capital market are conservative and rational, aiming for profitable allocations. Due to uncertain information, investors follow herd behaviour (Lin et al., 2013) and mimic peer investors. Some investors also follow rational herding in which market participants react to information about the behaviour of other market agents or participants rather than the behaviour of the market per se (Devenow and Welch, 1996). Most crypto investors tend to buy "popular" currencies like Bitcoin and Ethereum, leading to higher investment barriers (fear, inequitable access, insufficient funds (Lieberman, 1987)) than observed in the traditional markets. Our results, however, show that the most "popular" currencies do not lead to the most profitable portfolios. Non-mainstream cryptocurrencies may have technological advantages such as faster transaction time or be designed for specific industries or niches, providing more attractiveness and thus better diversification than mainstream cryptocurrencies. On the other hand, they may have lower liquidity and trading volumes, potentially increasing their risk.

4. Conclusion

Our study found that although the crypto market shows highly fluctuating prices compared to the traditional financial market, optimising global price correlations can be exploited via network modelling to create efficient agnostic portfolios. Price cross-correlations are higher than auto-correlations (due to the absence of memory), which opens a methodological venue for portfolio modelling. Network-based portfolios can achieve high returns, with acceptable risks, within a short-term investment horizon but with degrading performance in the long term. The optimal number of currencies is found empirically. Future research should automate this detection by exploiting price patterns during model training. Reducing the risk associated with crypto portfolios, particularly for long-term investment, should be further investigated, e.g., through hedging techniques. Trade indicators and characteristics of cryptocurrencies, e.g. trade volume, liquidity, blockchain technology, and niche market, combined with cross-correlations, might help improve portfolio performance. Persistent gains may also be achieved by applying other network-based strategies to correlation networks to find the optimal currencies and reduce the associated risk. Our data set is limited to daily closing prices and misses intra-day trade; higher resolution data, e.g. prices per minute, might also improve estimates. According to our results, the price dynamics of cryptocurrencies discourage long-term investment, and portfolios must be reviewed nearly daily to obtain the best results. Given the non-negligible risks, cryptocurrency investment remains recommended to investors who can timely pool sufficient diversification in their crypto portfolios. Our best-performing portfolios barely included the most popular currencies regarding market capitalisation. We thus recommend that investors pay attention to non-mainstream cryptocurrencies that can achieve higher returns, albeit with stronger price fluctuations. Regulators must also monitor non-mainstream currencies to minimise speculation and destabilisation of the market. Diversification has been increasingly facilitated via the popularisation of trade platforms, but the pool size (i.e. the number of available cryptocurrencies to invest) per platform still needs to be improved. Trade platforms vary considerably in the number of available currencies, and thus individual investors should consider multiple platforms to access non-mainstream currencies, which may be costly. Regulators could encourage trade platforms to provide more currencies, ensure price transparency in real-time and guarantee liquidity to increase market efficiency and democratise investment power. Such tasks might be challenging in the crypto eco-system, which is highly decentralised by design but would make it easier to monitor trade and avoid excessive speculation.

Acknowledgements

The authors declare that they have no conflict of interest. R.J. is funded by the China Scholarship Council (CSC) from the Ministry of Education of P. R. China. L.E.C.R. is partially funded by the Bijzonder Onderzoeksfonds (BOF) from Ghent University, Belgium.

References

O. B. Adekoya, A. B. Akinseye, N. Antonakakis, I. Chatziantoniou, D. Gabauer, and J. Oliyide. Crude oil and islamic sectoral stocks: Asymmetric tvp-var connectedness and investment strategies. *Resources Policy*, 78:102877, 2022. ISSN 0301-4207.

F. Allen and A. Babus. Networks in finance. *The network challenge: Strategy, profit, and risk in an interlinked world*, 367, 2009.

J. Almeida and T. C. Gonçalves. Portfolio diversification, hedge and safe-haven properties in cryptocurrency investments and financial economics: A systematic literature review. *Journal of Risk and Financial Management*, 16(1):3, 2023.

N. Antonakakis, J. Cunado, G. Filis, D. Gabauer, and F. P. de Gracia. Oil and asset classes implied volatilities: Investment strategies and hedging effectiveness. *Energy Economics*, 91:104762, 2020.

M. Balcilar, E. Bouri, R. Gupta, and D. Roubaud. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling*, 64:74–81, 2017.

A. Brauneis and R. Mestel. Cryptocurrency-portfolios in a mean-variance framework. *Finance Research Letters*, 28:259–264, 2019.

D. C. Broadstock, H. Cao, and D. Zhang. Oil shocks and their impact on energy related stocks in China. *Energy Economics*, 34(6):1888–1895, 2012.

G. Chamberlain. Multivariate regression models for panel data. *Journal of Econometrics*, 18(1):5–46, 1982.

S. Chatterjee and B. Wernerfelt. The link between resources and type of diversification: Theory and evidence. *Strategic Management journal*, 12 (1):33–48, 1991.

J. Chen. Analysis of bitcoin price prediction using machine learning. *Journal of Risk and Financial Management*, 16(1):51, 2023.

L. d. F. Costa, O. N. O. J. Oliveira Jr., G. Travieso, F. A. Rodrigues, P. R. V. Boas, L. Antiqueira, M. P. Viana, and L. E. C. Rocha. Analyzing and modeling real-world phenomena with complex networks: A survey of applications. *Advances in Physics*, 60(3):329–412, 2011.

P. V. J. da Gama Silva, M. C. Klotzle, A. C. F. Pinto, and L. L. Gomes. Herding behavior and contagion in the cryptocurrency market. *Journal of Behavioral and Experimental Finance*, 22:41–50, 2019.

J. Danko and V. Šoltés. Portfolio creation using graph characteristics. *Investment Management & Financial Innovations*, 15(1):180, 2018.

A. Devenow and I. Welch. Rational herding in financial economics. *European Economic Review*, 40(3-5):603–615, 1996.

A. Esfahanipour and S. Zamanzadeh. A stock market filtering model based on minimum spanning tree in financial networks. *AUT Journal of Modeling and Simulation*, 45(1):67–75, 2015.

J. L. Evans and S. H. Archer. Diversification and the reduction of dispersion: An empirical analysis. *The Journal of Finance*, 23(5):761–767, 1968.

F. J. Fabozzi, H. M. Markowitz, and F. Gupta. Portfolio selection. *Handbook of Finance*, 2, 2008.

E. F. Fama. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417, 1970.

S. Gharbi, J.-M. Sahut, and F. Teulon. R&d investments and high-tech firms' stock return volatility. *Technological Forecasting and Social Change*, 88:306–312, 2014.

M. Y. Hong and J. W. Yoon. The impact of covid-19 on cryptocurrency markets: A network analysis based on mutual information. *Plos ONE*, 17 (2):e0259869, 2022.

G. Jacobs. Cryptocurrencies & the challenge of global governance. *Cadmus*, 3(4):109–123, 2018.

A. Jalan and R. Matkovskyy. Systemic risks in the cryptocurrency market: Evidence from the FTX collapse. *Finance Research Letters*, page 103670, 2023.

J. B. Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, 7(1):48–50, 1956.

T. Kurosaki and Y. S. Kim. Cryptocurrency portfolio optimization with multivariate normal tempered stable processes and Foster-Hart risk. *Finance Research Letters*, 45:102143, 2022.

N. Latif, J. D. Selvam, M. Kapse, V. Sharma, and V. Mahajan. Comparative performance of LSTM and ARIMA for the short-term prediction of bitcoin prices. *Australasian Accounting, Business and Finance Journal*, 17(1):256–276, 2023.

M. B. Lieberman. Excess capacity as a barrier to entry: An empirical appraisal. *The Journal of Industrial Economics*, pages 607–627, 1987.

W. T. Lin, S.-C. Tsai, and P.-Y. Lung. Investors' herd behavior: Rational or irrational? *Asia-Pacific Journal of Financial Studies*, 42(5):755–776, 2013.

L. Lucchini, L. Alessandretti, B. Lepri, A. Gallo, and A. Baronchelli. From code to market: Network of developers and correlated returns of cryptocurrencies. *Science Advances*, 6(51):eabd2204, 2020.

R. N. Mantegna. Hierarchical structure in financial markets. *The European Physical Journal B*, 11(1):193–197, 1999.

C. K. Merker, S. W. Peck, C. K. Merker, and S. W. Peck. Over-diversification. *The Trustee Governance Guide: The Five Imperatives of 21st Century Investing*, pages 87–94, 2019.

K. Murray, A. Rossi, D. Carraro, and A. Visentin. On forecasting cryptocurrency prices: A comparison of machine learning, deep learning, and ensembles. *Forecasting*, 5(1):196–209, 2023.

J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertesz, and A. Kanto. Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, 68(5):056110, 2003.

M. Qin, T. Wu, R. Tao, C.-W. Su, and S. Petru. The inevitable role of bilateral relation: A fresh insight into the Bitcoin market. *Economic Research-Ekonomska Istraživanja*, 35(1):4260–4279, 2022.

L. E. C. Rocha, N. Masuda, and P. Holme. Sampling of temporal networks: Methods and biases. *Physical Review E*, 96:052302, 2017.

R. Stewart. The benefits of uncorrelated assets and strategies. *FS Advice*, page 9, 2019.

C.-W. Su, M. Qin, R. Tao, and X. Zhang. Is the status of gold threatened by Bitcoin? *Economic Research-Ekonomska Istraživanja*, 33(1):420–437, 2020.

D. Sulistiawan and F. A. Rudiawarni. Do accrual minimise (maximise) stock risk (return)?: Evidence from Indonesia. *International Journal of Globalisation and Small Business*, 9(1):20–28, 2017.

R. E. Tarjan. Sensitivity analysis of minimum spanning trees and shortest path trees. *Information Processing Letters*, 14(1):30–33, 1982.

Y. Xie, S. B. Kang, and J. Zhao. Are stablecoins safe havens for traditional cryptocurrencies? An empirical study during the covid-19 pandemic. *Applied Finance Letters*, 10:2–9, 2021.

C. Zhong, W. Du, W. Xu, Q. Huang, Y. Zhao, and M. Wang. LSTM-ReGAT: a network-centric approach for cryptocurrency price trend prediction. *Decision Support Systems*, 169:113955, 2023.

Supplementary Information

4.1. Data source

The daily price data of the cryptocurrencies used in our study was obtained from the following websites: www.investing.com, coinmarketcap.com, www.coindesk.com, and www.marketwatch.com. Table S1 shows the code of all the cryptocurrencies used in our study and the respective names can be retrieved from those websites.

ADA	BLOCK	OCEANp	PPC	ETH	QTUM	STORJ	LBC	MTL	XCP
ADX	BNB	OCN	CND	ETP	RCN	STRAX	LINK	MYST	XDN
AE	BNT	OK	CVC	EVX	RDD	SWFTC	LRC	NAS	XEM
AION	BTC	OMG	DASH	FAIR	RDN	SYNX	LSK	TOA	XLM
AMB	BTM	OMNI	DCN	FCT	REV	SYS	LTC	TRX	XRP
ANT	BTS	ONION	DCR	FLASH	RLC	GBYTE	LUNA	TUBE	XST
AOAR	BTU	PAC	DENT	FLO	SALT	GEO	MAID	UBQ	XTZ
ARDR	NEBL	PART	DGB	FTC	SBDR	GLM	MANA	USDT	XVG
ARK	NEO	PASC	DIME	FUEL	SC	GRC	MDA	VAL	ZAP
AST	NEU	PAY	DLT	FUN	SCRT	ICX	MDT	VERI	ZCL
BAT	NLG	PHR	DNT	GAME	SMART	IGNIS	MHC	VET	ZEC
BCD	NMC	PHX	DOGE	GAS	SNC	IOC	IOTA	VIA	ZEN
BCH	NMR	PIVX	EDG	PPT	SNM	JNT	MKR	VIB	ZRX
BCN	NXS	PLR	EMC2	PRE	SNT	KEY	MLN	WAVES	
BITCNY	NXT	PND	ENG	PRO	SOUL	KMD	MONA	WGR	
BLK	OAX	POT	ETC	QRL	STEEM	KNC	MTH	WINGS	

Table S1: The list of the $N = 157$ cryptocurrencies used in our study.

4.2. Minimum Spanning Tree

Figure S1 shows an example of the crypto correlation distance network filtered by using the Minimum Spanning Tree method (MST) on April 15, 2018.

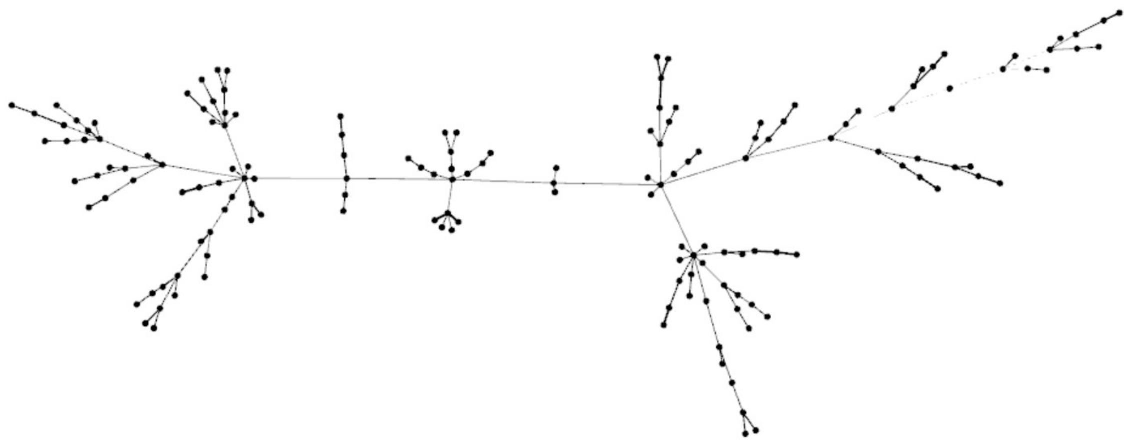

Figure S1: Example of the Minimum Spanning Tree of the crypto correlation distance network on April 15, 2018.

4.3. Performance for various investment horizons

Figure S2 shows the predicted average expected returns (with standard errors) and average expected volatility (risk, with standard errors) for the network-based portfolios with $n = 1$ to $n = 70$ cryptocurrencies.

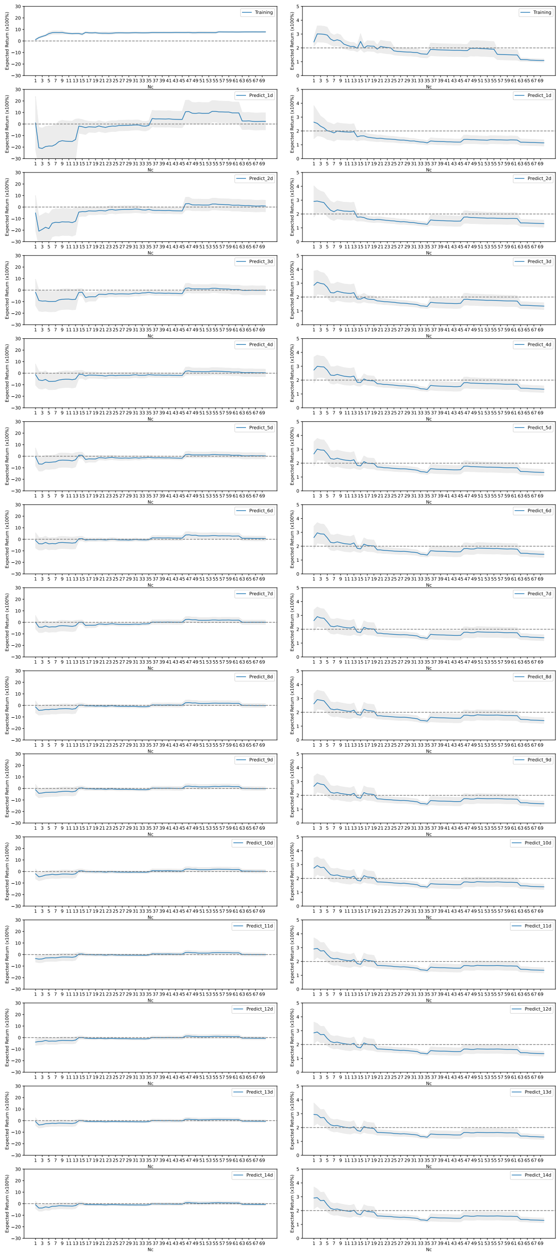

Figure S2: Predicted average expected return and average risk for investment horizons from 1 to 14 days. Grey areas show the respective standard errors.

4.4. Frequency of occurrence of cryptocurrencies in the optimal portfolios MST.

Figure S3 shows the frequency of occurrence of the cryptocurrencies used in the optimal network-based portfolios.

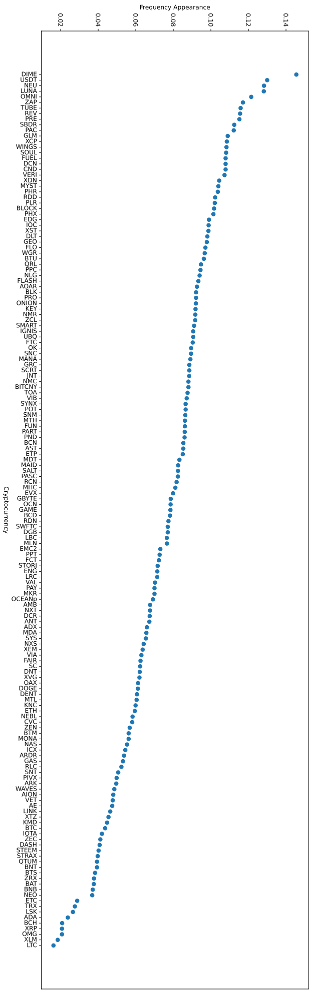

Figure S3: Ranking of the cryptocurrencies according to its frequency of use in the optimal network-based portfolios.

4.5. Distance correlation network for difference window sizes

Figure S4 shows the distributions and evolution of the distance correlation networks for different time windows Δt.

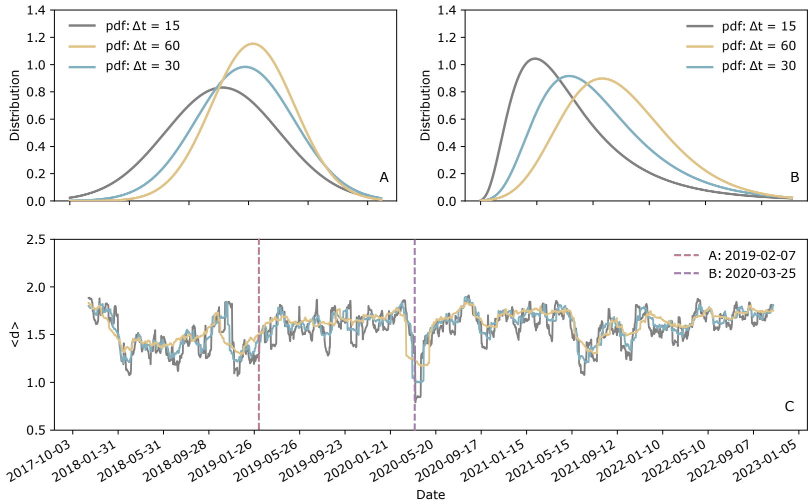

Figure S4: A) and B) show the distributions of distance correlations for Δt = 15 days, Δt = 30 days, and Δt = 60 days, respectively, during a quiet period (fitted by the normal distribution, Δt = 15 days:p-value < .01, μ = 1.28, σ = 0.48; Δt = 30 days: p-value < .01, μ = 1.45, σ = 0.41; Δt = 60 days: p-value < .01, μ = 1.52, σ = 0.35) and a turbulent period (fitted by the log-normal distribution, Δt = 15 days:p-value < .01, μ = −0.09, σ = 0.79, s = 0.57; Δt = 30 days: p-value < .01, μ = −0.24, σ = 0.40, s = 1.19; Δt = 60 days: p-value < .01, μ = −0.89, σ = 2.02, s = 0.23) (See panel C) for the selected dates). C) shows the evolution of the average distance correlation $\langle d \rangle$ during the study period and two selected dates.

4.6. Descriptive statistics

Table S2 shows the correlation analysis of log-transformed return for the top 5 (BTC, Bitcoin; ETH, Ethereum; LTC: Litecoin, USDT: Tether, XRP) and bottom 5 (OCN: Odyssey, DLT: Agrello, ENG: Enigma, ETP: Metaverse, FUEL: Etherparty) cryptocurrencies according to their average market capitalization during the study period.

	BTC	ETH	LTC	USDT	XRP	OCN	DLT	ENG	ETP	FUEL
BTC	1	0.522	0.763	0.026	0.563	0.103	0.268	0.452	0.519	0.115
ETH	0.522	1	0.537	0.030	0.454	0.073	0.329	0.360	0.386	0.114
LTC	0.763	0.537	1	0.063	0.646	0.117	0.288	0.443	0.509	0.109
USDT	0.026	0.030	0.063	1	-0.007	-0.039	-0.037	-0.013	-0.033	0.005
XRP	0.563	0.454	0.646	-0.007	1	0.088	0.251	0.357	0.440	0.114
OCN	0.103	0.073	0.117	-0.039	0.088	1	0.045	0.096	0.062	0.005
DLT	0.268	0.329	0.288	-0.037	0.251	0.045	1	0.235	0.292	0.064
ENG	0.452	0.360	0.443	-0.013	0.357	0.096	0.235	1	0.356	0.077
ETP	0.519	0.386	0.509	-0.033	0.440	0.062	0.292	0.356	1	0.080
FUEL	0.115	0.114	0.109	0.005	0.114	0.005	0.064	0.077	0.080	1

Table S2: Correlation matrix of log-transformed returns ($r_{i,t}$) for the top 5 and bottom 5 cryptocurrencies according to their average market capitalisation.