



Cryptocurrency co-investment network: token returns reflect investment patterns

Luca Mungo^{1,2*} , Silvia Bartolucci³ and Laura Alessandretti⁴

*Correspondence:

luca.mungo@maths.ox.ac.uk

¹Mathematical Institute, University of Oxford, Oxford, United Kingdom

²Institute for New Economic Thinking, University of Oxford, Oxford, United Kingdom

Full list of author information is available at the end of the article

Abstract

Since the introduction of Bitcoin in 2009, the dramatic and unsteady evolution of the cryptocurrency market has also been driven by large investments by traditional and cryptocurrency-focused hedge funds. Notwithstanding their critical role, our understanding of the relationship between institutional investments and the evolution of the cryptocurrency market has remained limited, also due to the lack of comprehensive data describing investments over time. In this study, we present a quantitative study of cryptocurrency institutional investments based on a dataset collected for 1324 currencies in the period between 2014 and 2022 from Crunchbase, one of the largest platforms gathering business information. We show that the evolution of the cryptocurrency market capitalization is highly correlated with the size of institutional investments, thus confirming their important role. Further, we find that the market is dominated by the presence of a group of prominent investors who tend to specialise by focusing on particular technologies. Finally, studying the co-investment network of currencies that share common investors, we show that assets with shared investors tend to be characterized by similar market behaviour. Our work sheds light on the role played by institutional investors and provides a basis for further research on their influence in the cryptocurrency ecosystem.

1 Introduction

Since the introduction of Bitcoin in 2009 [1], the cryptocurrency market has experienced bewildering growth, surpassing an overall value of one trillion dollars in early 2021. Beyond private investors, the development of the market was fostered by cryptocurrency hedge funds and Venture Capital (VC) funds, with institutional investments in cryptocurrency-related projects reaching an estimated amount of 17 billion US dollars in 2021 [2, 3].

A growing number of traditional financial firms and investment funds in Europe and the U.S. are also exploring avenues for investments in cryptocurrency via different channels, including, but not limited to, including cryptocurrency into their portfolios, investing through tokenization in equity of blockchain companies, and exploiting more regulated tools such as crypto futures, options, and ETFs [3, 4]. Unfriendly regulations, high volatility, and lack of reliable valuation tools, amongst other issues, have so far hindered widespread adoption and institutionalisation of these assets [3, 5, 6]. Most cryptocurrency

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

platforms, for instance, lack regulatory and supervisory oversight concerning trading, disclosure, anti-money laundering, and consumer protection measures, forming what has also been described as a “shadow financial system” [7]. Nonetheless, recent challenging events affecting the economy and markets, i.e., the U.S. elections, Brexit in Europe, and the global pandemic, have gradually accelerated the uptake [3]. Despite these developments, the effects of institutional investments on the cryptocurrency market are still little understood, also due to the lack of comprehensive quantitative data.

Moreover, it has recently been flagged that the participation of institutional investors in both crypto and traditional markets might lead to potential spillovers and increased contagion risks between traditional finance and decentralised finance (DeFi)¹ [4]. Understanding the behaviour of institutional investors and its effect on the structure and evolution of the cryptocurrency markets is therefore of paramount importance to quantify the mutual impact between DeFi and traditional entrepreneurial finance [4, 8].

This paper aims to study the link between institutional investments and cryptocurrencies’ market trends systematically and quantitatively, exploiting a novel combination of data sources on a larger sample of cryptocurrencies. Our analysis exploits network science tools to study the structure and evolution of the co-investment network, i.e., constructed as an undirected network of cryptocurrencies (nodes) connected if they share a common investor. In particular, we aim to tackle the following two main research questions: (i) Do connections in the co-investment network reflect intrinsic similarities (e.g., in terms of technology or use cases) between cryptocurrencies? (ii) Is the co-investment network related to cryptocurrencies’ market dynamics? First, we investigate the connection between the co-investment network structure and various features of cryptocurrencies, such as their supported blockchain protocols and use cases. Then, we examine the relation between the co-investment network structure and the correlation between the market behaviour of pairs of tokens measured in terms of correlations of their returns (i.e., the percentage changes in their prices over time).

The article is organised as follows: in Sect. 2, we provide an overview of the relevant literature; in Sect. 3, we describe how the data was collected and integrated and the methodologies and algorithms employed for this study; in Sect. 4.1, we describe the co-investment network and study how the cryptocurrency features (e.g., type of blockchain protocol, use case) are related to the network structure; in Sect. 4.2 we study the connection between the structure of the co-investment network and market properties of different assets. In Sect. 5, we conclude.

2 Related work

Our work contributes to the literature on (i) characterising cryptocurrency market dynamics, (ii) constructing optimal portfolios of currencies, and (iii) quantifying and characterising institutional investments in cryptocurrency-related projects.

A growing body of literature has so far focused on the properties of the rapidly evolving crypto market ecosystem, shedding light on critical aspects such as assessing market efficiency and maturity [9, 10], detecting and characterising asset pricing bubbles due to endogenous and exogenous events [11, 12]. The dynamics of competition between currencies [13, 14], and the impact of collective attention [15] have also been closely analysed.

¹The term “decentralised finance” refers to financial services, such as lending or asset trading, provided through decentralized platforms, as opposed to traditional centralized financial institutions.

Given the digital and decentralised nature of crypto assets, a major focus has been to understand the drivers of price fluctuations and how to properly value these assets. Studies using empirical data have focused on understanding and predicting the price dynamics of cryptocurrencies using machine learning techniques with different input features [15–20]. Socio-economic signals, such as sentiment index gathered from social media platforms [21, 22], also appear to be strongly intertwined with the price dynamics [23, 24]. Research has also shown that movements in the market can be tied to macroeconomic indicators, media exposure, and public interest [25, 26], policies and regulations [27], and indeed the behaviour of other financial assets [28].

In the context of institutional investments, the recent growing interest in mixed portfolios of crypto and traditional assets [4] has paved the way to research looking at optimal portfolio allocation strategies. Studies have focused on the composition of mixed portfolios, i.e., including traditional (bonds, commodities, etc.) and crypto assets [29, 30], and crypto-only portfolios [31, 32] testing the performances of different allocation and rebalancing strategies. Specific strategies, e.g., introducing so-called stop-loss rules, have been tested specifically as they would make crypto portfolios more appealing to institutional investors due to lower risks associated with volatility [33].

Concerning characterising and quantifying institutional interest and investments in cryptocurrency projects, most of the research available is based on qualitative surveys by private companies of investors in Europe and the U.S., which aim to identify market trends and issues, e.g., barriers to adoption and current channels to exposure in cryptocurrencies [3, 4]. In Sun, 2021 [34], for instance, the authors surveyed 33 Asian firms to investigate whether price volatility lowers institutional investors' confidence and to quantify the role played by the familiarity of investors with the technology in the selection of crypto assets. In [35] the authors analysed the connection between investors' ESG preferences and crypto investments exposure using household-level portfolio data gathered from the Austrian Survey of Financial Literacy (ASFL). The analysis suggests that crypto investments are more strongly driven by social and ethical preferences compared to traditional investments (e.g., bonds). In [7], the authors analyse the drivers of crypto adoption, and assess institutional investors' crypto exposure via different channels (e.g., banks, exchanges, etc.). In [36], a comprehensive review of typical crypto investors' behaviour and their effect, including understanding drivers of investors' sentiment and attention and detecting herding behaviour. In [37] the authors provide a first quantitative exploration of the investor's network focusing on data for investments on ~ 300 ERC-20 tokens.² Their analysis shows that less central tokens in the investment network have also low market capitalization (i.e., the overall dollar value of all the tokens) and trading volume, poor liquidity, and high volatility. Our analysis builds directly on their approach, by considering an extended set of cryptoassets, as well as a novel combination of data, which also includes information on the technological features of the assets considered.

²An ERC-20 token is a type of digital asset that runs on the Ethereum blockchain, following a standardized set of rules so it can easily interact with other apps and tokens. Essentially, it is a special type of currency that can be used in a variety of online applications and services.

3 Data and methods

3.1 Data description

In this paper, we use three main data types, (i) cryptocurrency price time series data, (ii) cryptocurrency metadata describing projects' technological features and/or their use case and functionalities, and (iii) data capturing information on investment rounds in cryptocurrency projects.

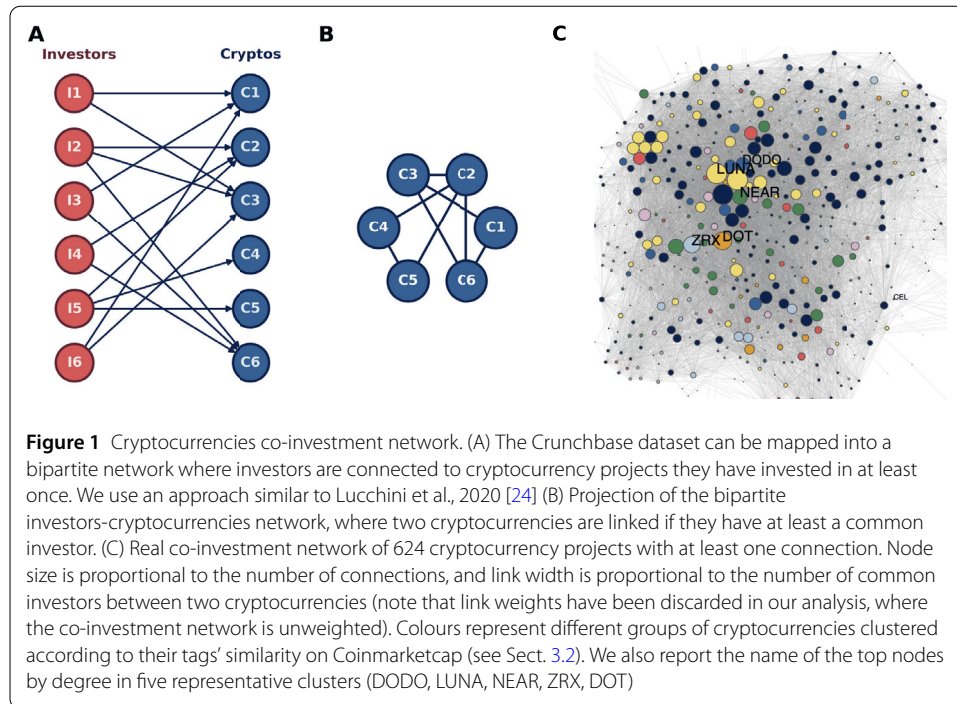
Market data (i) and cryptocurrency metadata (ii) were extracted from the website Coinmarketcap [38]. The data covers 1324 cryptocurrency projects over eight years, spanning from 2014 to 2022. It is important to note that the term 'cryptocurrency' here encompasses various types of blockchain-based digital assets. This includes traditional cryptocurrencies like Bitcoin and Litecoin, which are standalone digital currencies operating on their own blockchains, and blockchain-based tokens, such as the previously mentioned ERC-20 tokens on the Ethereum blockchain and analogous tokens on other platforms. These tokens have a range of applications, and they can represent various assets or functionalities within decentralized applications. A notable example within this group is stablecoins, which are typically designed to minimize price volatility by being pegged to more stable assets such as fiat currencies.

Market data consists of each cryptocurrency's opening price, closing price, and traded volume, sampled weekly.

Coinmarketcap also assigns tags describing the main features of the different cryptocurrencies. Metadata can be broadly classified into three categories. The first is *technology*-related specifications, which refer to the underlying blockchain technology that the cryptocurrency employs (e.g., Proof-of-Work vs. Proof-of-Stake algorithms—these are different methods used to validate transactions and create new blocks in the blockchain). The second is *ecosystem*-related information, indicating whether the cryptocurrency operates on an independent blockchain or as part of an existing one, as well as whether it is part of decentralized finance (DeFi) projects. The third category relates to the *use case*, or the specific purpose and utility of the cryptocurrency (e.g., it could be used for facilitating distributed storage, as a fan token for a particular brand or celebrity, or simply as a digital store of value, like digital gold). See Appendix A.5 for a list of available tags used to categorize these aspects and their respective frequency. The dataset contains 226 unique tags. Cryptocurrencies' tags might change over time as, for instance, the project pivots its scope or new categories are invented. Thus, the data we collected and used in the analysis should be understood as a snapshot of the cryptocurrency environment at the time they were gathered (August 2021).

Coinmarketcap also provides cryptocurrencies' webpage URLs, which are used to merge market-related data with investment data.

Finally, the investments' data (iii) is gathered from Crunchbase [39], a commercial database covering worldwide innovative companies and accessed by 75M users each year. The data is sourced through two main channels: an extensive investor network and community contributors. Investors commit to keeping their portfolios updated to get free access to the dataset. More than 600k executives, entrepreneurs, and investors update over 100k company, people, and investor profiles per month. Crunchbase processes the data with machine learning algorithms to ensure accuracy and scan for anomalies, ultimately verified by a team of data experts at Crunchbase. Due to its broad coverage, the data has been used in thousands of scholarly articles and technical reports [39, 40]. Information on



Crunchbase includes an overview of the company's activities, number of employees, and detailed information on funding rounds, including investors and—more rarely—amounts raised. We provide detailed information on the features contained in this dataset in Appendix A.4.

We merged the Crunchbase data on investment rounds with Coinmarketcap data via the companies' webpage URLs. After merging, the dataset includes 4395 investments made in 1458 rounds by 1767 investors to 1324 cryptocurrency projects appearing on Crunchbase. The total investments amount to \$13B US dollars in the period considered (2008–2022). When merging with the time series data, we can still track 624 cryptocurrency projects.

3.2 Methods

In this section, we review the methods used for our analyses. We first describe the co-investment network and the approach we used to cluster its nodes. Later, we explain our analysis of the interplay between the network structure and the market dynamics.

Co-investment network The main object considered in our study is the cryptocurrencies' co-investment network. Figure 1, A shows how the co-investment network is constructed as a monopartite projection of the bipartite network where investors are connected to cryptocurrency projects they have funded at least once. In the resulting co-investment network (Fig. 1, B)—which is unweighted and undirected—nodes represent different cryptocurrencies, and the presence of a link means that the two nodes share at least one common investor. Figure 1C, shows the real co-investment network composed of 624 cryptocurrency projects. The node sizes are proportional to their degree, and the link widths are proportional to the number of common investors between two cryptocurrencies. In the rest of this paper, the co-investment network will be characterised by a *binary* and symmetric adjacency matrix A , with entries $a_{ij} \in \{0, 1\}$, recording only the information on whether *at least one shared investor* exists between two cryptocurrencies.

Clustering algorithm We assign a vector \mathbf{x}_i to each cryptocurrency, where, for every tag j , $x_{i,j} = 1$ if the j -th tag (see Table 6) is assigned to the i -th cryptocurrency, and $x_{i,j} = 0$ otherwise. We used the Ward Aggregative Clustering [41] algorithm to divide the cryptocurrencies into different clusters based on the observations $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$. The algorithm uses a “bottom-up” approach: each observation is initially placed in its own clusters, and clusters are merged sequentially according to some criterion until the desired number of clusters is reached. Ward’s algorithm specifically prescribes to merge, at each iteration, the pair of clusters S_i, S_j that minimizes the distance $\Delta(S_i, S_j)$, defined as

$$\Delta(S_i, S_j) = \sum_{l \in S_i \cup S_j} \|\mathbf{x}_l - \boldsymbol{\mu}_{i+j}\|^2 - \sum_{l \in S_i} \|\mathbf{x}_l - \boldsymbol{\mu}_i\|^2 - \sum_{l \in S_j} \|\mathbf{x}_l - \boldsymbol{\mu}_j\|^2 = \frac{|S_i||S_j|}{|S_i| + |S_j|} \|\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\|^2, \quad (1)$$

where $|S_i|$ is the number of observations in cluster S_i , $\boldsymbol{\mu}_i$ is the mean of points in S_i , $\boldsymbol{\mu}_j$ is the mean of points in S_j , and $\boldsymbol{\mu}_{i+j}$ is the mean of points in $S_i \cup S_j$. The number of clusters k is an input of the clustering algorithm. Using the elbow method (see Appendix A.1) we set $k = 12$. We opted for Ward’s Agglomerative Clustering Algorithm over alternatives such as k -means and k -modes due to its propensity for generating more equal cluster sizes [42, 43]. Minimizing the total within-cluster variance, which often results in clusters that are similarly sized in terms of variance, Ward’s method provides a more regular partitioning of the data. Since our data is sparse (i.e., each cryptocurrency only has a handful of tags), other alternatives would put most of the cryptocurrencies in a single cluster. However, we show in Appendix A.1 that our conclusions are robust with respect to the clustering algorithm choice.

Clustering evaluation and benchmarks We investigate whether the clusters obtained via the previous procedure reflect the underlying network structure by studying the in-density and out-density of links according to the partitioning defined by the clusters. Given the $N \times N$ adjacency matrix A of our co-investment network and the clustering $S^* = \{S_1, \dots, S_k\}$, we define the *in-density* of a cluster S_i as

$$\rho_i^i = \frac{1}{|S_i|(|S_i| - 1)} \sum_{j, k \in S_i, j \neq k} A_{jk}, \quad (2)$$

and its *out-density* as

$$\rho_i^o = \frac{1}{|S_i|(N - |S_i|)} \sum_{j \in S_i, k \notin S_i} A_{jk}. \quad (3)$$

These metrics are used to study whether cryptocurrencies with similar characteristics—clustered according to the Coinmarket cap tags—are more strongly interconnected (higher in-cluster density) in the co-investment network among themselves rather than with groups of dissimilar cryptocurrencies. We, then, compare the *in-densities* and *out-densities* of the clusters identified by the clustering algorithm with those of random clusters. To generate the random clusters, we simply assign each cryptocurrency to one of the twelve possible clusters with equal probability. In Sect. A.3 of the Appendix, we repeat the analysis with several different node similarity metrics including the *Jaccard index*, the *cosine similarity* (also known as *Salton index*), the *Adamic–Adar index*, and the *resource allocation index*, showing that our findings are robust with respect to different metrics.

Time series processing The investigation of the co-investment network's relationship with the cryptocurrency market is conducted by computing cryptocurrencies' returns correlation. The primary objects of this analysis are cryptocurrencies' weekly closing price (i.e., the final price at which the cryptocurrency is traded during a specific trading week) time series $p_i(t), i = 1, \dots, N$. We compute their log returns as

$$r_i(t) = \log\left(\frac{p_i(t+1)}{p_i(t)}\right), \quad (4)$$

and use the *leave-one-out* rescaling described in [44] to define the rescaled returns,

$$\tilde{r}_i(t) = \frac{r_i(t) - \mathbb{E}_{t'}[r_i(t')]}{\sqrt{\mathbb{V}_{t' \neq t}[r_i(t')]}}, \quad (5)$$

where the average of the returns $\mathbb{E}_{t'}[r_i(t')]$ is computed over all times t' , but the variance $\mathbb{V}_{t' \neq t}[r_i(t')]$ is computed from the time series where the observation corresponding to $t' = t$ has been removed. The correlation matrix of the time series \tilde{r}_i is defined as

$$C_{ij} = \mathbb{E}_t[\tilde{r}_i(t)\tilde{r}_j(t)]. \quad (6)$$

Cryptocurrencies' prices usually move coherently, increasing or decreasing simultaneously [45–47]. This collective behaviour of the market makes returns strongly correlated and hides the more subtle effects we want to highlight. Therefore, we adopt the following strategy to remove the so-called *market component* from the correlation matrix characterising common price co-movements [48]. We first compute the set of eigenvalues $\lambda_1, \dots, \lambda_N$ of the correlation matrix, the corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_N$, and the modes $m_i(t)$, defined as

$$m_i(t) = \sum_{j=1}^N \mathbf{v}_{ij} \tilde{r}_j(t). \quad (7)$$

We call *market mode* the mode $m_1(t)$ associated with the largest eigenvalue λ_1 . The time series $\tilde{r}_i(t)$ can now be written as linear combinations of the modes $m_i(t)$,

$$\tilde{r}_i(t) = \sum_{j=1}^N c_{ij} m_j(t). \quad (8)$$

We can now define the *adjusted* time series $r'_i(t)$,

$$r'_i(t) = \sum_{j>1}^N c_{ij} m_j(t), \quad (9)$$

and the corresponding *adjusted* correlation matrix C' ,

$$C'_{ij} = \mathbb{E}_t[r'_i(t)r'_j(t)]. \quad (10)$$

Network correlation and random benchmarks We compute the average value of the raw and adjusted correlations C and C' (defined in Eq. (6), (10) respectively) restricted to the pairs of cryptocurrencies (i, j) that are linked (i.e., share an investor) in the co-investment network. For any (binary) adjacency matrix \mathbf{M} characterising the co-investment network, we define

$$C_M = \mathbb{E}_{ij}[C_{ij}M_{ij}|M_{ij} > 0], \tag{11}$$

and

$$C'_M = \mathbb{E}_{ij}[C'_{ij}M_{ij}|M_{ij} > 0], \tag{12}$$

where the average runs over all pairs (i, j) of connected nodes. The values of C_M and C'_M range from -1 to 1 , where -1 indicates a perfect inverse correlation, 0 indicates no correlation, and 1 indicates a perfect positive correlation between pairs of cryptocurrencies. High values (close to 1) suggest that the cryptocurrencies move in tandem, while a value around 0 would indicate a lack of any significant relationship in their returns.

We compute C_A and C'_A over the adjacency matrix A of the real co-investment network and compare them with the values obtained on three random network models: the *Erdős-Rényi* model [49], the *Stochastic Block Model* [50], and the *Configuration Model* [51]. Here—to mimic the properties of the real co-investment network—we have constructed undirected and unweighted random networks as benchmarks.

For every model, we sample $n = 1000$ network instances R_1, \dots, R_n at random, and compute the mean and standard deviation of the sets $\{C_{R_1}, \dots, C_{R_n}\}$ and $\{C'_{R_1}, \dots, C'_{R_n}\}$. All models are parametrized to match the empirical properties of the co-investment network. The probability of a link p in the Erdős-Rényi model is set to match the co-investment network's empirical density,

$$p = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j>i}^N A_{ij}.$$

Blocks in the Stochastic block model match the clusters found with the clustering algorithm and the densities within- and across- clusters are equal to the empirical values. Finally, the degree sequence in the configuration model matches the empirical degree sequence.

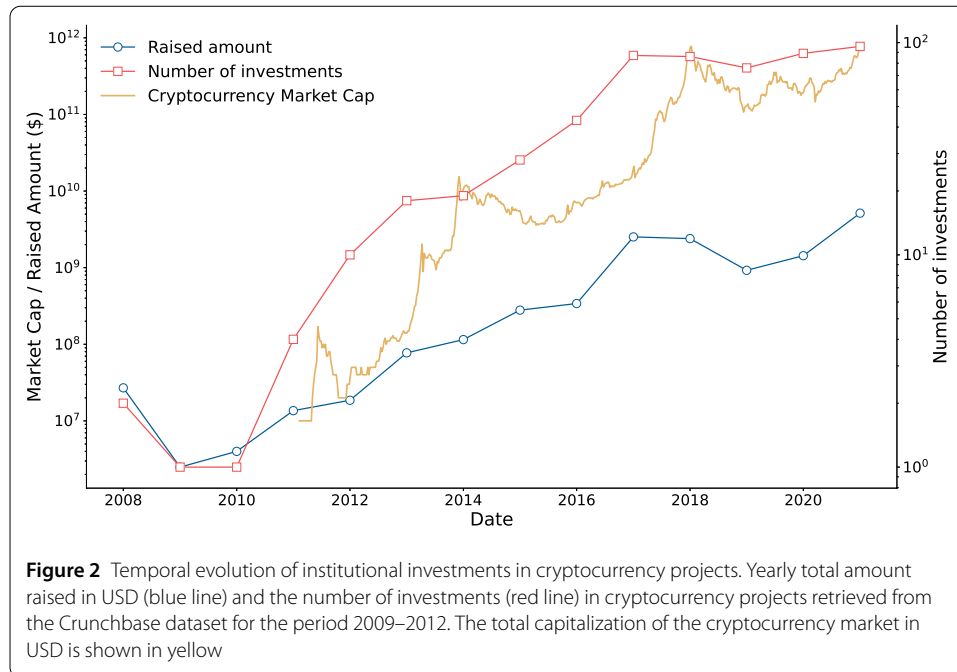
4 Results

4.1 Structure of the cryptocurrency co-investment network

In this section, we analyze the relationship between institutional investments and the properties of the cryptocurrency market.

We start by quantifying the joint evolution of the number and volume of investments together with the growth of the cryptocurrency market. In Fig. 2, we show the evolution of the total raised amount, number of investments, and market capitalization³ of the cryptocurrency ecosystem. Overall, we find that the number of investments, as well as the

³The market capitalization of a token is the total value of all its units in circulation, calculated by multiplying the current price per token by the total number of tokens available.

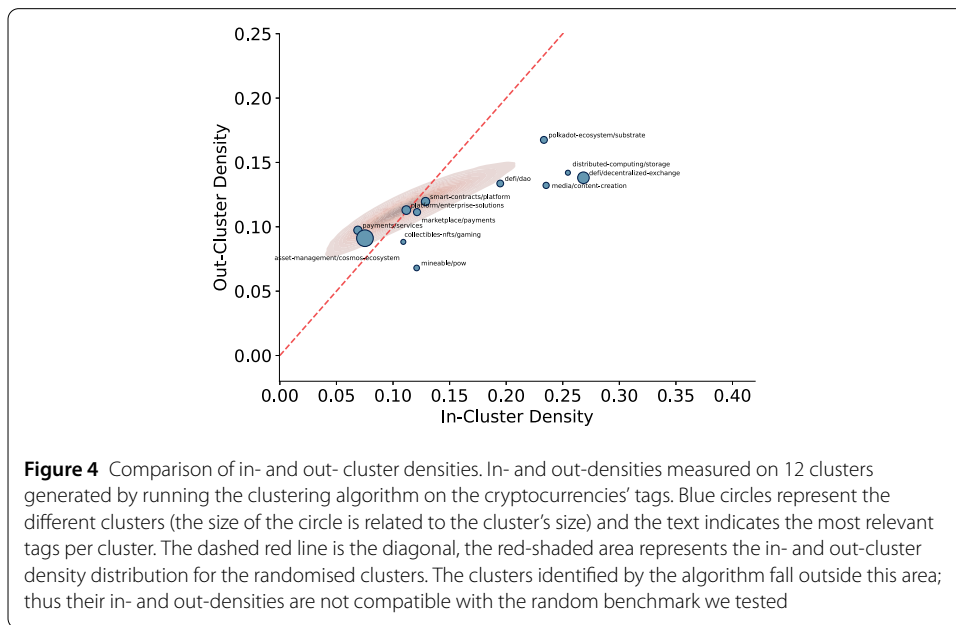
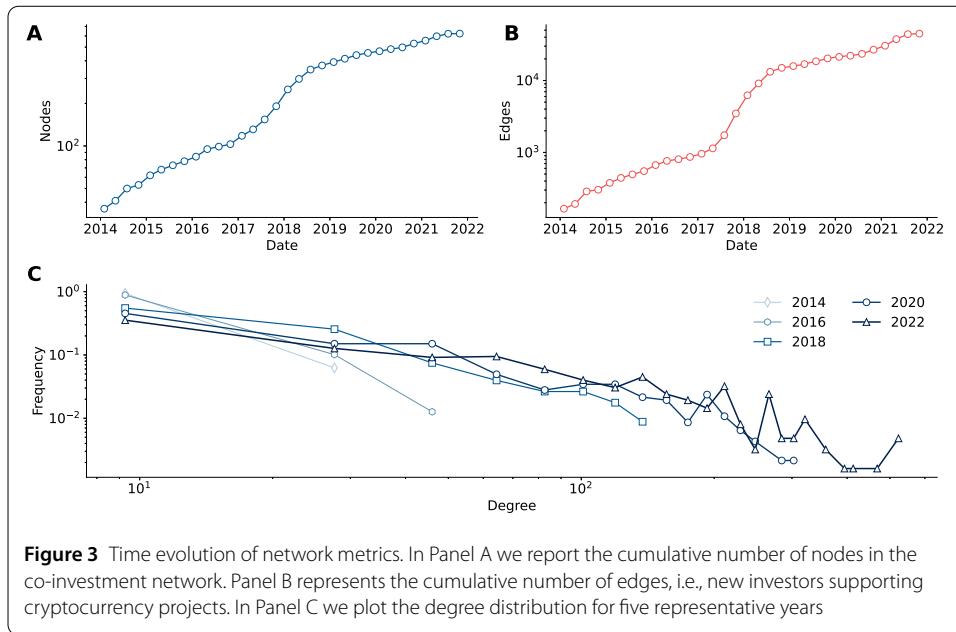


amount raised, has been steadily growing since 2012. Moreover, we find a positive correlation between the cryptocurrency market capitalization (MC) and both the total volume of investments/raised amount in dollars (VI) and the number of investments (NI). The Spearman correlation amounts respectively to $\rho_{MC-VI} = 0.79$ and $\rho_{MC-NI} = 0.81$, suggesting that the crypto market and the volume of investments have evolved hand in hand.

Next, we turn to studying the evolution of the co-investment network in time (see Fig. 3). We find that, since 2014, the network has grown steadily in terms of the cumulative number of nodes (panel A), i.e., cryptocurrency projects funded by institutional investors, and the cumulative number of edges (panel B), i.e., common investors between cryptocurrencies. Interestingly, the growth displays a steeper increase around 2017–2019, consistently with the rapid increase in demand for cryptocurrencies and the rise of Bitcoin’s valuation over those years [52]. Turning our attention to the number of connections per node, we observe that the degree distribution of the co-investment network is heavy-tailed, with most nodes having a single connection and only a few having hundreds of neighbours (see Fig. 1C). Interestingly, the shape of the distribution has been relatively stable over time (see Fig. 1C), in line with the findings discussed in Ref. [37], where the authors studied the co-investment network restricted to ERC-20 tokens only.

Which factors may explain the observed structure of the cryptocurrency co-investment network? In the following, we test the hypothesis that the structure of the co-investment network is partly determined by the properties characterising different cryptocurrency projects (e.g., their underlying technology or their purpose) because investors tend to specialize and invest in specific types of cryptocurrencies. More formally, we assess whether two cryptocurrencies with similar properties are also more likely to be connected in the co-investment network compared to any random pair of currencies.

To this end, we assign each cryptocurrency to a cluster, based on its properties (see Sect. 3.2 for more details). Then—for each cluster i —we calculate the in-cluster density ρ_i^i and the out-cluster density ρ_i^o , as defined in Eq. (2) and Eq. (3) respectively. We then



compare the in- and out-cluster densities: if ρ_i^i is significantly higher than ρ_i^o , then there is a higher density of links among cryptocurrencies with similar properties.

Indeed, we observe that the densities inside clusters of similar cryptocurrencies tend to be larger than those across clusters (see Fig. 4), which confirms our hypothesis. In practice, this implies that similar cryptocurrency projects (i.e., those that share a common set of tags), tend to share a larger number of investors compared to any two randomly chosen projects.

Importantly, we find that—when cryptocurrencies are assigned to random clusters—the relation between the in- and out-density is significantly different (see red shaded area in Fig. 4). Thus, our results reveal that there is a non-trivial connection between the topol-

ogy of the network and the intrinsic features of cryptocurrency projects. In particular, they hint at the presence of specialised investors who do not simply invest in the whole cryptocurrency ecosystem but rather focus on specific technologies and/or use cases.

4.2 Interplay between the co-investment network structure and returns correlations

In this section, we investigate the interplay between the structure of the co-investment network and the cryptocurrency market properties. More specifically, we test if the price returns of cryptocurrencies that share common investors are more correlated than one would expect by random chance.

To this end, we compute the average returns correlation C_A defined in Eq. (11) across pairs of cryptocurrencies sharing a link in the real co-investment network (described by its adjacency matrix A). We also compute average returns correlation of cryptocurrency pairs sharing a link on random network benchmarks including (i) an Erdős-Rényi network, (ii) a configuration model and (iii) a stochastic block model parametrized to reproduce some of the features of the real network (e.g., number of nodes, number of clusters, degree distribution—as detailed in Sect. 3).

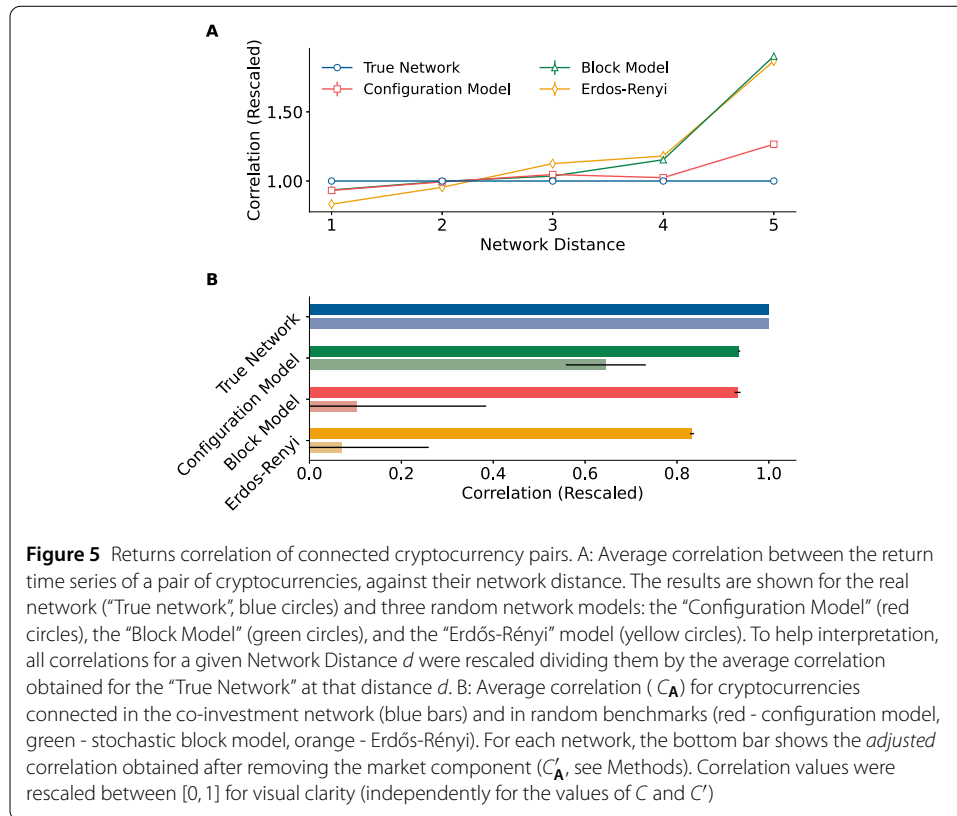
Figure 5 compares the values of the correlation for the real co-investment network and the benchmarks respectively. The correlation values displayed can be found in Table 1 and Table 2 of the Appendix. In Panel A of Fig. 5, the returns correlation between cryptocurrency pairs is plotted against their network distance, defined as the shortest path between the two nodes in the network. Our findings indicate that the average correlation decreases as the distance in the network increases. Cryptocurrencies that are “close” in the co-investment network are, on average, more correlated than the random benchmarks; conversely, pairs of cryptocurrencies that are distant in the network are less correlated than the benchmarks.

Figure 5, Panel B summarizes the average returns correlation for the real network (blue) and random networks (green, red, and orange). The lighter shades of colour display the values of the correlation C'_A for the adjusted time series, where the market component has been removed (see Sect. 3.2). Once again, the figure shows that the average correlation on the real network is significantly larger than on all the benchmarks tested, suggesting that the network’s structure may directly impact the cryptocurrencies’ market behaviour. Furthermore, the gap between real and random correlation widens significantly after removing the time series as discussed in Sect. 3.2.

Overall, our results reveal that the returns of cryptocurrencies that share a common investor have a stronger correlation than one would expect by random chance, revealing that assets with shared investors tend to be characterized by similar market dynamics.

5 Discussion

In this paper, we have analyzed an ecosystem of 1324 cryptocurrency projects that received 4395 investments from 1767 investors for a total amount of \$13B appearing on Crunchbase. We have built and analysed the co-investment network, where two cryptocurrencies are linked if they share an investor. We have also clustered cryptocurrency projects based on metadata and tags from the Coinmarketcap website and studied the community structure.



As hinted by previous research and surveys concerning institutional and individual crypto investors' preferences [3, 4, 37, 53], our results show that investors tend to specialise and focus on particular technologies, use cases, and features of the cryptocurrency projects they decide to include in their portfolio.

We have also analyzed the relationship between the co-investment network and the cryptocurrencies' market properties. We showed that the presence of a link in the co-investment network translates into a higher correlation in cryptocurrencies' returns. The marginal increase in the correlation of cryptocurrency returns decreases as the distance between the considered pairs of cryptocurrencies in the co-investment network increases.

Our work has limitations that, hopefully, can be turned into future avenues of research. As stated above, we also provide access to the co-investment network reconstructed from Crunchbase to ease further explorations and extensions of our work. Firstly, our data collection process stopped over the summer of 2021, before the second major cryptocurrency crash and the default of established players such as Terra, Celsius, and FTX. It is legit to wonder to what extent our results would hold in the new regime, where the general sentiment towards cryptocurrencies has pivoted.

Secondly, some prominent players in the cryptocurrencies' ecosystem are not associated with a company, but rather with different types of organizations including *Decentralized Autonomous Organizations* (DAOs), foundations, or even no legal entity at all. The nature of the investment may also vary substantially. For instance, instead of buying a share of the company, investors may, e.g., lend money to DeFi protocols in exchange for tokens as rewards (a practice known as *liquidity mining* [54]). These new organization types and forms of investment are scarcely represented in our dataset, therefore we can only offer a

partial view of the cryptocurrencies' investment ecosystem. Finally, most of our analysis was performed on a static network. However, how the network grows, what the different investment strategies adopted by an investor are, and how they depend on the market are also clearly worth analyzing.

In light of the recent crypto market crash events—from the stablecoin pair Terra—Luna to large exchanges [55–57]—understanding the crypto market connectedness at the investors level helps shed light on possible contagion channels posing threat to the ecosystem overall stability.

Appendix

A.1 Methods

We have tested four algorithms to cluster our set of cryptocurrencies based on the associated Coinmarketcap tags, namely Ward’s iterative clustering, *k-means* [58], *k-modes* [59], and an agglomerative clustering algorithm based on cosine distance between data points. We eventually settled for Ward’s algorithm due to its propensity to generate more equally-sized clusters [42, 43]. However, other algorithms resulted in similar, non-random partitions of cryptocurrencies into clusters as shown in Fig. 6.

However, the algorithm choice might be not optimal, and more sophisticated clustering algorithms could lead to more insightful partitions of our data. Specifically, it should be mentioned that Ward’s algorithm, as well as *k-means*, computes Euclidean distances to divide data points into clusters, which is, arguably, not the optimal way of computing distances when dealing with binary data.

To select the total number of cryptocurrencies’ clusters we employ the elbow method. For each possible partition $\mathbf{S} = S_1, \dots, S_k$ of the dataset, we define a loss function $L(S)$ as

$$L(\mathbf{S}) = \sum_{i=1}^k \sum_{j \in S_i} \|\mathbf{x}_j - \boldsymbol{\mu}_i\|^2, \tag{13}$$

where \mathbf{x}_j is the vector of tags observations for cryptocurrencies belonging to the partition S_i and $\boldsymbol{\mu}_i$ is its mean. We ran the clustering algorithm for several different values of k , and computed the value of the loss function for the set of optimal partitions $\{S_{k=1}^*, S_{k=2}^*, \dots, S_{k=N}^*\}$, where N is the total number of cryptocurrencies considered in our study.

The elbow method prescribes choosing the maximum number of clusters before the curve becomes flat. Intuitively, the method recommends picking a point where the marginal decrease in the loss function is not worth the additional cost of creating another cluster. Figure 7 shows that a value around $k = 12$ is compatible with the elbow method in our case.

A.2 Further analysis

The tables below report the results used to build Fig. 5. In particular, we show the mean correlation defined in Eq. (12) and its variance computed over 1000 realizations of the random networks and on the real co-investment network (Eq. (11)). In Table 1 we report the results as a function of the network distance, while in Table 2 computed over all pairs

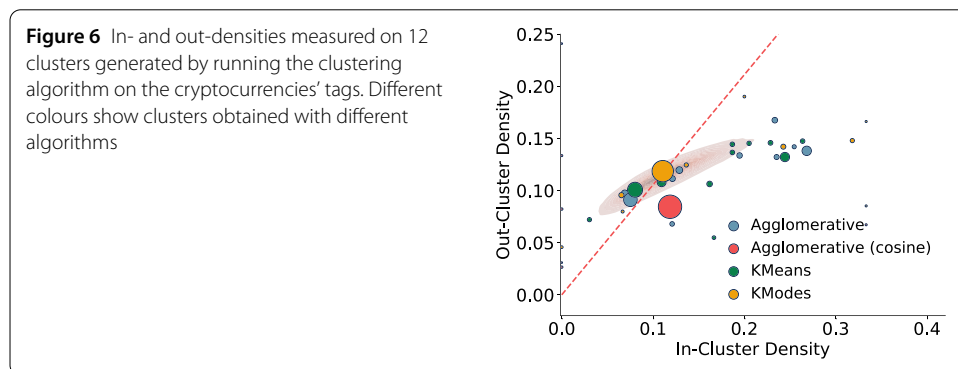


Figure 6 In- and out-densities measured on 12 clusters generated by running the clustering algorithm on the cryptocurrencies’ tags. Different colours show clusters obtained with different algorithms

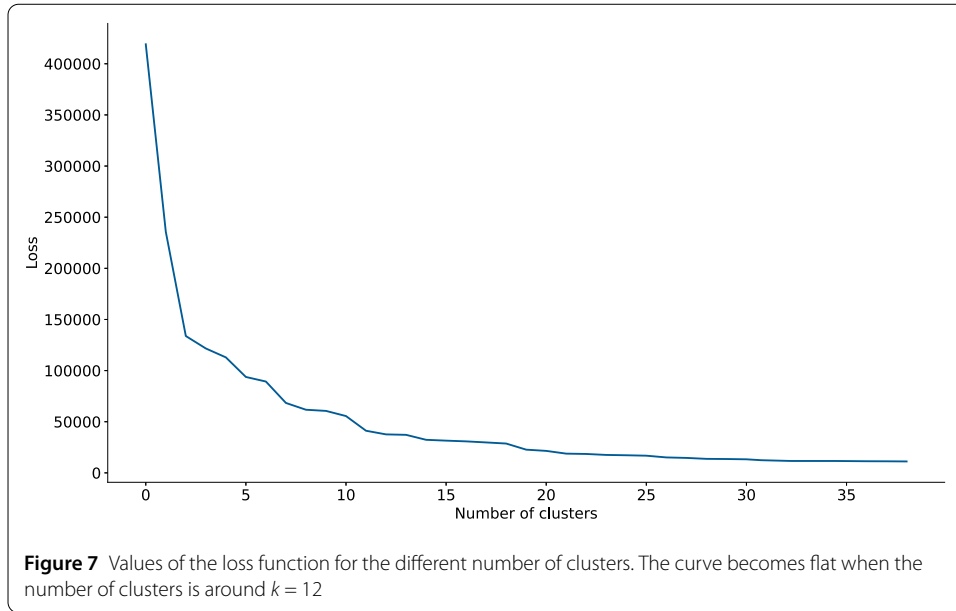


Table 1 Correlation values as a function of the distance for Fig. 5 A comparing results for the real co-investment network and the three random benchmarks (Configuration Model, Block Model and Erdős-Rényi)

Distance	True network	Configuration model	Block model	Erdős-Rényi
1	0.38	0.354±0.002	0.355±0.001	0.316±0.002
2	0.331	0.329±0.001	0.330±0.000	0.316±0.000
3	0.281	0.294±0.001	0.291±0.000	0.316±0.001
4	0.268	0.274±0.003	0.309±0.001	0.316±0.003
5	0.17	0.215±0.010	0.323±0.002	0.317±0.012

Table 2 Correlation values for the real co-investment network and the three random benchmarks (Configuration Model, Block Model and Erdős-Rényi) used in Fig. 5, B

Model	Correlation	Rescaled	Correlation, cleaned data	Rescaled, cleaned data
True Network	0.380	1.000	0.010	1.0
Block Model	0.355±0.001	0.935±0.002	6.64e-03±8.96e-04	0.645±0.087
Configuration Model	0.354±0.002	0.932±0.006	1.06e-03±2.90e-03	0.103±0.282
Erdős-Rényi	0.316±0.002	0.833±0.004	7.33e-04±1.94e-03	0.071±0.188

of cryptocurrencies, including the raw correlation values as well as correlations computed on ‘cleaned data’ obtained by removing the market mode (see Eq. (10) and rescaling the correlation to be in the range [0, 1] and included in the figure.

A.3 Clusters analysis

To better characterise the similarity between nodes belonging to the same clusters as defined in Sect. A.1, we compute four well-known similarity measures [60], the *Jaccard index*, the *cosine similarity* (also known as *Salton index*), the *Adamic-Adar index*, and the *resource allocation index*. The Jaccard index measures the similarity between two nodes’ sets of neighbours and is defined as the size of the intersection divided by the size of the union of the sets. The cosine similarity counts the number of common neighbours but penalizes nodes that have a higher degree. The Adamic-Adar index and the resource allocation index count the number of common neighbours, but they assign a lower weight

Table 3 Data entries in the organization Crunchbase bundle

Bundle	Columns name	Description
Organization	uuid	Organization unique Identifier
	name	Company's name
	permalink	
	cb_url	Company url on Crunchbase
	rank	Crunchbase rank
	created_at	Record creation date
	updated_at	Last record update
	legal_name	Company legal name
	roles	Company, Investor, or both
	domain	Company's website domain
	homepage_url	Company's homepage URL
	country_code	
	state_code	
	region	
	city	
	address	
	postal_code	
	status	
	short_description	
	category_list	Company classification (e.g., Enterprise Software, Financial Services, Social Media)
	category_groups_list	Company classification (e.g., Content and Publishing, Internet Services)
	num_funding_rounds	Number of funding rounds
	total_funding_usd	Total Funding raised in USD
	total_funding	Total funding raised
	total_funding_currency_code	Funding currency
	founded_on	
	last_funding_on	
	closed_on	
	employee_count	
	email	
	phone	
	facebook_url	
	linkedin_url	
twitter_url		
logo_url		
alias1	Other company's names	
alias2		
alias3		
primary_role	Either "company" or "investor"	
num_exits		

to neighbours that have a high degree. If we call $\Gamma(i)$ the set of neighbors of a node i , we can define these measures as

$$d_{ij}^{\text{jaccard}} = \frac{|\Gamma(i) \cap \Gamma(j)|}{|\Gamma(i) \cup \Gamma(j)|},$$

$$d_{ij}^{\text{cosine}} = \frac{|\Gamma(i) \cap \Gamma(j)|}{\sqrt{|\Gamma(i)| \times |\Gamma(j)|}},$$

$$d_{ij}^{\text{Adamic-Adar}} = \sum_{k \in \Gamma(i) \cap \Gamma(j)} \frac{1}{\log |\Gamma(k)|},$$

$$d_{ij}^{\text{RA}} = \sum_{k \in \Gamma(i) \cap \Gamma(j)} \frac{1}{|\Gamma(k)|}.$$

Table 4 Data entries in the Crunchbase funding rounds bundle

Bundle	Columns name	Description
Funding Rounds	uuid	Funding round unique identifier
	name	Funding round name (e.g., Angel Round—Facebook)
	permalink	
	cb_url	Crunchbase url
	rank	Crunchbase company rank
	created_at	Record creation date
	updated_at	Record last update
	country_code	
	state_code	
	region	
	city	
	investment_type	Investment type (e.g., angel, seed, series a)
	announced_on	
	raised_amount_usd	
	raised_amount	
	raised_amount_currency_code	
	post_money_valuation_usd	
	post_money_valuation	
	post_money_valuation_currency_code	
	investor_count	Number of investors
org_uuid	Investee unique identifier	
org_name	Investee name	
lead_investor_uuids	Lead investor's unique identifier.	

Table 5 Data entries in the Crunchbase investment bundle

Bundle	Columns name	Description
Investments	uuid	Investment unique identifier
	name	Investment's name (e.g., Accel investment in Series A—Facebook)
	permalink	
	cb_url	Crunchbase's investment url
	created_at	Record creation date
	updated_at	Record last update
	funding_round_uuid	
	funding_round_name	
	investor_uuid	
	investor_name	
	investor_type	Either "organization" or "person"
	is_lead_investor	

For each cluster S_k , we compute the average value of each metric within and outside the cluster. The average similarity inside the cluster is

$$d_k^{\text{in}} = \frac{1}{|S_k| \times (|S_k| - 1)} \sum_{i,j \in S_k, i \neq j} d_{ij},$$

and the average similarity outside the cluster is

$$d_k^{\text{out}} = \frac{1}{|S_k| \times (N - |S_k|)} \sum_{i \in S_k, j \notin S_k} d_{ij},$$

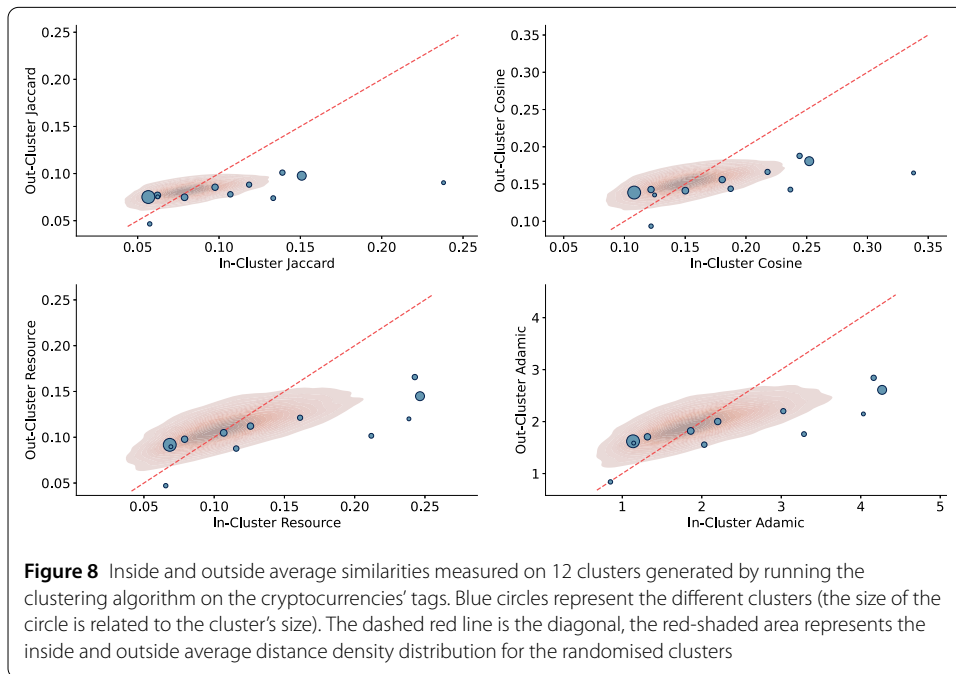
where d_{ij} represents one of the four metrics defined above. Figure 8 shows the values of the in- and out-average similarity metrics for the 12 cryptocurrency clusters described in Sect. 4 and compares them with those obtained for 1000 random clustering assignments.

Table 6 Coinmarketcap cryptocurrencies tags and their frequency characterising the cryptocurrencies present in the co-investment network

0	mineable: 465	defi: 333	platform: 188
1	collectibles-nfts: 139	yield-farming: 129	payments: 127
2	pow: 98	marketplace: 97	binance-smart-chain: 86
3	masternodes: 84	decentralized-exchange: 83	smart-contracts: 82
4	exnetwork-capital-portfolio: 72	hybrid-pow-pos: 72	medium-of-exchange: 65
5	polkadot-ecosystem: 53	governance: 53	script: 53
6	dao: 49	enterprise-solutions: 47	ethereum: 47
7	privacy: 42	gaming: 41	media: 40
8	pos: 38	asset-management: 37	kinetic-capital: 36
9	stablecoin: 32	centralized-exchange: 32	distributed-computing: 31
10	services: 28	ai-big-data: 28	content-creation: 27
11	cosmos-ecosystem: 26	staking: 26	iot: 26
12	pantera-capital-portfolio: 23	alameda-research-portfolio: 23	filesharing: 23
13	tokenized-stock: 22	sha-256: 22	substrate: 22
14	polkastarter: 20	amm: 20	memes: 19
15	sports: 18	gambling: 18	derivatives: 18
16	storage: 17	x11: 16	oracles: 16
17	rebase: 16	solana-ecosystem: 16	stablecoin-asset-backed: 16
18	entertainment: 15	store-of-value: 14	polkadot: 14
19	yield-aggregator: 14	wallet: 14	dao-maker: 14
20	coinbase-ventures-portfolio: 13	duckstarter: 13	binance-launchpad: 13
21	wrapped-tokens: 12	seigniorage: 12	interoperability: 12
22	lending-borrowing: 10	binance-chain: 10	cms-holdings-portfolio: 10
23	dapp: 10	insurance: 10	dcg-portfolio: 9
24	multicoins-capital-portfolio: 9	launchpad: 9	polychain-capital-portfolio: 9
25	hashkey-capital-portfolio: 9	fan-token: 9	synthetics: 8
26	poolz-finance: 8	binance-labs-portfolio: 8	three-arrows-capital-portfolio: 8
27	placeholder-ventures-portfolio: 7	blockchain-capital-portfolio: 6	scaling: 6
28	social-money: 6	fabric-ventures-portfolio: 6	crowdfunding: 6
29	dpos: 5	boostvc-portfolio: 5	arrington-xrp-capital: 5
30	framework-ventures: 4	defi-index: 4	trustswap-launchpad: 4
31	discount-token: 4	state-channels: 3	coinfund-portfolio: 3
32	logistics: 3	dex: 3	a16z-portfolio: 3
33	marketing: 3	e-commerce: 3	tourism: 3
34	health: 2	research: 2	loyalty: 2
35	dragonfly-capital-portfolio: 2	identity: 2	energy: 2
36	parafi-capital: 1	huobi-capital: 1	metaverse: 1
37	yearn-partnerships: 1	defiance-capital: 1	ledgerprime-portfolio: 1
38	data-provenance: 1	sharing-economy: 1	zero-knowledge-proofs: 1
39	paradigm-xzy-screener: 1	electric-capital-portfolio: 1	1confirmation-portfolio: 1
40	binance-launchpool: 1	video: 1	analytics: 1
41	music: 1	cybersecurity: 1	prediction-markets: 1
42	fenbushi-capital-portfolio: 1	options: 1	education: 1
43	real-estate: 1	x13: 1	aave-tokens: 1
44	avalanche-ecosystem: 1	mobile: 1	galaxy-digital-portfolio: 1
45	crowdsourcing: 1	hardware: 0	reputation: 0
46	usv-portfolio: 0	jobs: 0	stablecoin-algorithmically-stabilized: 0
47	quark: 0	multiple-algorithms: 0	equihash: 0
48	events: 0	winklevoss-capital: 0	art: 0
49	atomic-swaps: 0	cryptonight: 0	communications-social-media: 0
50	neoscrypt: 0	social-token: 0	dag: 0
51	heco: 0	retail: 0	eth-2-0-staking: 0
52	philanthropy: 0	commodities: 0	ringct: 0
53	transport: 0	sharding: 0	quantum-resistant: 0
54	ethash: 0	vr-ar: 0	hospitality: 0
55	asset-backed-coin: 0	layer-2: 0	blake2b: 0
56	hybrid-dpow-pow: 0	hacken-foundation: 0	adult: 0
57	manufacturing: 0	sha-256d: 0	search-engine: 0
58	ontology: 0	dagger-hashimoto: 0	poc: 0
59	pos-30: 0	blake256: 0	blake: 0
60	hybrid-pos-lpos: 0	geospatial-services: 0	m7-pow: 0

Table 6 (Continued)

61	fashion: 0	cryptonight-lite: 0	tron: 0
62	mimble-wimble: 0	lp-tokens: 0	poi: 0
63	lyra2rev2: 0	agriculture: 0	posign: 0
64	timestamping: 0	pop: 0	lpos: 0
65	sidechain: 0	platform-token: 0	eos: 0
66	hybrid-pow-npos: 0	lellantusmw: 0	groestl: 0
67	cosmos: 0	x11gost: 0	scrypt-n: 0
68	food-beverage: 0	tpos: 0	qubit: 0
69	x15: 0	sha-512: 0	data-availability-proof: 0
70	cuckoo-cycle: 0	escrow: 0	rollups: 0
71	hybrid-pos-pop: 0	yescript: 0	rpos: 0
72	x14: 0	post: 0	blake2s: 0
73	nist5: 0	bulletproofs: 0	sigma: 0
74	argon2: 0	lyra2re: 0	xevan: 0
75	waves: 0		



Nodes belonging to the same cluster tend to be more similar, in a way that is not compatible with a random benchmark.

A.4 Crunchbase dataset

Crunchbase provides information on worldwide innovative companies. The dataset covers several aspects of the companies, spanning from a basic description of the business description to their financial status, board composition, and even media exposition. The dataset is organized in different bundles that reflect this different information. The bundles are:

- *Company-related: organizations* (including information on parent companies, organization descriptions, and their division in categories) and *investment funds*.
- *Investment-related: funding rounds* (group of investments in a single company), *investments* (specific investor-to-company transaction), *investors*, *acquisitions*, *IPOs*.

- *People-related*: people covered in the dataset, the jobs they have, and the degrees they hold, with a focus on *investment partners*.
- *Event-related*: events description and *event appearances* of specific companies.

For the sake of this paper, the relevant bundles concern organization, funding rounds, and investments. We detail their content in Tables 3, 4, 5.

A.5 Coinmarketcap cryptocurrency tags

Table 6 contains together with their respective frequency gathered from Coinmarketcap for all the cryptocurrency projects analysed in this paper. Given the heterogeneity of the cryptocurrency market in terms of use case and/or supporting technology, the tags created by Coinmarketcap help label and distinguish the different types of cryptocurrencies based on 'intrinsic' features related to the nature of the project.

Acknowledgements

The authors acknowledge Crunchbase for easing data access.

Data availability

Aggregated data describing the bipartite network of investors and cryptocurrencies is available [here](#). All code for pre-processing and analysing the data is available [here](#).

Declarations

Competing interests

The authors declare that they have no competing interests.

Author contributions

LM collected the data and performed the analysis. All authors analysed the results and wrote the manuscript. All authors read and approved the final manuscript.

Author details

¹Mathematical Institute, University of Oxford, Oxford, United Kingdom. ²Institute for New Economic Thinking, University of Oxford, Oxford, United Kingdom. ³Dept. of Computer Science, University College London, London, United Kingdom. ⁴DTU Compute, Technical University of Denmark, Kongens Lyngby, Denmark.

Received: 30 May 2023 Accepted: 25 December 2023 Published online: 20 February 2024

References

1. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. <https://bitcoin.org/bitcoin.pdf>
2. Kochkodin B Venture Capital Makes a Record \$17 Billion Bet on Crypto World. Accessed: 2022-11-20 (2022). <https://www.bloomberg.com/news/articles/2021-06-18/venture-capital-makes-a-record-17-billion-bet-on-crypto-world?sref=3REHEaVI>
3. Neureuter J (2021) The institutional investor digital assets study. Technical report, Fidelity Digital Assets. <https://www.fidelitydigitalassets.com/sites/default/files/documents/2021-digital-asset-study.pdf>
4. Institutionalisation of crypto-assets and DeFi-TradFi interconnectedness. Accessed: 2022-11-20 (2022). <https://www.oecd-ilibrary.org/content/paper/5d9dddbbe-en>
5. Rauchs M, Blandin A, Bear K, McKeon SB (2019) 2nd global enterprise blockchain benchmarking study. Available at SSRN 3461765
6. Gurguc Z, Knottenbelt W Cryptocurrencies: overcoming barriers to trust and adoption. Retrieved from Imperial College London website: <https://www.imperial.ac.uk/media/imperial-college/research-centres-andgroups/ic3re/cryptocurrencies-overcoming-barriers-to-trust-and-adoption.pdf> (2018)
7. Auer R, Farag M, Lewrick U, Orazem L, Zoss M (2023) Banking in the shadow of bitcoin? The institutional adoption of cryptocurrencies
8. Shakhnov K, Zaccaria L (2020) (r) evolution in entrepreneurial finance? the relationship between cryptocurrency and venture capital markets. Technical report, Einaudi Institute for Economics and Finance (EIEF)
9. Sigaki HY, Perc M, Ribeiro HV (2019) Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market. *Sci Rep* 9(1):1–9
10. Vidal-Tomás D, Ibañez A (2018) Semi-strong efficiency of bitcoin. *Finance Res Lett* 27:259–265
11. Chen CY-H, Hafner CM (2019) Sentiment-induced bubbles in the cryptocurrency market. *J Financ Risk Manag* 12(2):53
12. Vidal-Tomás D, Bartolucci S (2023) Artificial intelligence and digital economy: divergent realities. Available at SSRN 4589333
13. Dowd K, Greenaway D (1993) Currency competition, network externalities and switching costs: towards an alternative view of optimum currency areas. *Econ J* 103(420):1180–1189

14. Luther WJ (2016) Cryptocurrencies, network effects, and switching costs. *Contemp Econ Policy* 34(3):553–571
15. ElBahrawy A, Alessandretti L, Baronchelli A (2019) Wikipedia and cryptocurrencies: interplay between collective attention and market performance. *Front Blockchain* 2:12
16. Alessandretti L, ElBahrawy A, Aiello LM, Baronchelli A (2018) Machine learning the cryptocurrency market. *Complexity* 2018
17. Walther T, Klein T, Bouri E (2019) Exogenous drivers of bitcoin and cryptocurrency volatility—a mixed data sampling approach to forecasting. University of St. Gallen. Research Paper (2018/19)
18. McNally S, Roche J, Caton S (2018) Predicting the price of bitcoin using machine learning. In: 2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP). IEEE Press, New York, pp 339–343. <https://ieeexplore.ieee.org/abstract/document/8374483>
19. Chen Z, Li C, Sun W (2020) Bitcoin price prediction using machine learning: an approach to sample dimension engineering. *J Comput Appl Math* 365:112395
20. Akyildirim E, Goncu A, Sensoy A (2020) Prediction of cryptocurrency returns using machine learning. *Ann Oper Res*: 1–34
21. García D, Tessone CJ, Mavrodiev P, Perony N (2014) The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. *J R Soc Interface* 11(99):20140623. <https://doi.org/10.1098/rsif.2014.0623>
22. Aste T (2018) Cryptocurrency market structure: connecting emotions and economics. special issue of digital finance on cryptocurrencies. *Digit Finance*
23. Ortu M, Uras N, Conversano C, Bartolucci S, Destefanis G (2022) On technical trading and social media indicators for cryptocurrency price classification through deep learning. *Expert Syst Appl* 198:116804
24. Lucchini L, Alessandretti L, Lepri B, Gallo A, Baronchelli A (2020) From code to market: network of developers and correlated returns of cryptocurrencies. *Sci Adv* 6(51). <https://doi.org/10.1126/sciadv.abd2204>
25. Lyócsa S, Molnár P, Plíhal T, Širaňová M (2020) Impact of macroeconomic news, regulation and hacking exchange markets on the volatility of bitcoin. *J Econ Dyn Control* 119:103980
26. Corbet S, Larkin C, Lucey BM, Meegan A, Yarovaya L (2020) The impact of macroeconomic news on bitcoin returns. *Eur J Finance* 26(14):1396–1416
27. Borri N, Shakhnov K (2020) Regulation spillovers across cryptocurrency markets. *Finance Res Lett* 36:101333
28. Nguyen KQ (2022) The correlation between the stock market and bitcoin during Covid-19 and other uncertainty periods. *Finance Res Lett* 46:102284
29. Koutsouri A, Poli F, Alfieri E, Petch M, Distaso W, Knottenbelt WJ (2020) Balancing cryptoassets and gold: a weighted-risk-contribution index for the alternative asset space. In: *Mathematical research for blockchain economy*. Springer, Berlin, pp 217–232
30. Platanakis E, Urquhart A (2020) Should investors include bitcoin in their portfolios? A portfolio theory approach. *Br Account Rev* 52(4):100837
31. Hu Y, Rachev ST, Fabozzi FJ (2019) Modelling crypto asset price dynamics, optimal crypto portfolio, and crypto option valuation. ArXiv preprint. [arXiv:1908.05419](https://arxiv.org/abs/1908.05419)
32. Ahelegbey DF, Giudici P, Mojtahedi F (2021) Crypto asset portfolio selection. Available at SSRN 3892999
33. Białkowski J (2020) Cryptocurrencies in institutional investors' portfolios: evidence from industry stop-loss rules. *Econ Lett* 191:108834
34. Sun W, Dedahanov AT, Shin HY, Li WP (2021) Factors affecting institutional investors to add crypto-currency to asset portfolios. *N Am J Econ Finance* 58:101499
35. Ciaian P, Cupak A, Fessler P, Kancs D (2022) Environmental-social-governance preferences and investments in crypto-assets. ArXiv preprint. [arXiv:2206.14548](https://arxiv.org/abs/2206.14548)
36. Almeida J, Gonçalves TC (2023) A systematic literature review of investor behavior in the cryptocurrency markets. *J Behav Exp Finance* 100785
37. Liu S-H, Liu XF (2021) Co-investment network of ERC-20 tokens: network structure versus market performance. *Front Phys* 9:55
38. Coinmarketcap. <https://coinmarketcap.com/>. Accessed: 2022-07-16
39. Dalle J-M, den Besten M, Menon C (2017) Using crunchbase for economic and managerial research. Technical report, OECD. <https://doi.org/10.1787/6c418d60-en>. <https://www.oecd-ilibrary.org/content/paper/6c418d60-en>
40. den Besten ML (2020) Crunchbase research: monitoring entrepreneurship research in the age of big data. Available at SSRN 3724395
41. Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58(301):236–244
42. Everitt BS, Landau S, Leese M, Stahl D (2011) *Cluster analysis*. Wiley, New York
43. Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *J Classif* 31:274–295
44. Bouchaud J-P, Potters M (2003) *Theory of financial risk and derivative pricing*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/cbo9780511753893>
45. Katsiampa P, Corbet S, Lucey B (2019) High frequency volatility co-movements in cryptocurrency markets. *J Int Financ Mark Inst Money* 62:35–52
46. Koutmos D (2018) Return and volatility spillovers among cryptocurrencies. *Econ Lett* 173:122–127
47. Stosic D, Stosic D, Luderer TB, Stosic T (2018) Collective behavior of cryptocurrency price changes. *Phys A, Stat Mech Appl* 507:499–509
48. Laloux L, Cizeau P, Bouchaud J-P, Potters M (1999) Noise dressing of financial correlation matrices. *Phys Rev Lett* 83(7):1467
49. Erdős P, Rényi A (1959) On random graphs I. *Publ Math (Debr)* 6:290–297
50. Karrer B, Newman MEJ (2011) Stochastic blockmodels and community structure in networks. *Phys Rev E* 83:016107. <https://doi.org/10.1103/PhysRevE.83.016107>
51. Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45(2):167–256. <https://doi.org/10.1137/S003614450342480>
52. Institutional investors will bet big on cryptocurrencies in 2018 (2018) Technical report, Cointelegraph. <https://cointelegraph.com/news/institutional-investors-will-bet-big-on-cryptocurrencies-in-2018>

53. Ciaian P, Rajcaniova M, Kancs D (2016) The economics of bitcoin price formation. *Appl Econ* 48(19):1799–1815
54. Fan S, Min T, Wu X, Wei C (2022) Towards understanding governance tokens in liquidity mining: a case study of decentralized exchanges. *World Wide Web*: 1–20
55. Briola A, Vidal-Tomás D, Wang Y, Aste T (2022) Anatomy of a stablecoin's failure: the terra-luna case. *Finance Res Lett* 103358
56. Hermans L, Ianiro A, Kochanska U, Törmälehto V-M, van der Kraaij A, Simón JMV et al (2022) Decrypting financial stability risks in crypto-asset markets. *Financ Stab Rev* 1
57. Chipolina S (2022) FT Cryptofinance: crypto's Lehman moment. Accessed: 2022-11-30. <https://www.ft.com/content/a4d31278-a5d9-4a02-a169-4996f4a8e8f8>
58. Lloyd S (1982) Least squares quantization in pcm. *IEEE Trans Inf Theory* 28(2):129–137. <https://doi.org/10.1109/TVT.1982.1056489>
59. Huang Z (1998) Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Min Knowl Discov* 2(3):283–304. <https://doi.org/10.1023/A:1009769707641>
60. Lü L, Zhou T (2011) Link prediction in complex networks: a survey. *Phys A, Stat Mech Appl* 390(6):1150–1170. <https://doi.org/10.1016/j.physa.2010.11.027>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
