Nonlinear Dyn (2022) 107:4001–4017
https://doi.org/10.1007/s11071-021-07166-9

ORIGINAL PAPER

Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time

Nick James · **Max Menzies**

Received: 30 July 2021 / Accepted: 19 December 2021 / Published online: 3 January 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract This paper introduces new methods to study behaviours among the 52 largest cryptocurrencies between 01-01-2019 and 30-06-2021. First, we explore evolutionary correlation behaviours and apply a recently proposed turning point algorithm to identify regimes in market correlation. Next, we inspect the relationship between collective dynamics and the cryptocurrency market size—revealing an inverse relationship between the size of the market and the strength of collective dynamics. We then explore the time-varying consistency of the relationships between cryptocurrencies' size and their returns and volatility. There, we demonstrate that there is greater consistency between size and volatility than size and returns. Finally, we study the spread of volatility behaviours across the market changing with time by examining the structure of Wasserstein distances between probability density functions of rolling volatility. We demonstrate a new phenomenon of increased uniformity in volatility during market crashes, which we term *volatility dispersion*.

Keywords Cryptocurrency · Time series analysis · Nonlinear dynamics · Correlations · Collective behaviours

N. James (✉)
School of Mathematics and Statistics, University of Melbourne, Parkville, VIC 3010, Australia
e-mail: nick.james@unimelb.edu.au

M. Menzies
Beijing Institute of Mathematical Sciences and Applications, Tsinghua University, Beijing 101408, China

1 Introduction

Over the last several years, the cryptocurrency market has attracted substantial interest from both institutional and retail investors. The market has experienced significant growth in total assets, accompanied by commensurate levels of volatility. Following the COVID-19 and BitMEX market crash in early 2020, cryptocurrency prices had a strong rally and then a subsequent decline in prices across the market. Given the differing views on their long-term viability, understanding the changing nature of cryptocurrencies' returns and volatility as the market grows in size is a timely priority for investors and policymakers alike. Our paper meets this purpose by analysing the collective dynamics of cryptocurrencies over time regarding returns, volatility, and market size.

Our paper builds on a long literature studying the dynamics of financial markets. Across several fields, researchers have been interested in the time-varying nature of financial market behaviours for some time, particularly correlations [1–7]. The application of classical statistical techniques, such as ARCH and GARCH models [8–10] and other parametric models, can encounter difficulties due to the non-stationary nature of financial markets. Certain models for descriptive analysis or portfolio selection that perform well during extended bull market periods can suffer during sudden market crises. The literature has split into two approaches to this problem. Statistical researchers

have introduced explicit methodologies to model non-stationarity [11], while nonlinear dynamics researchers have taken a more descriptive approach to the time-varying dynamics.

In the nonlinear dynamics community, such market dynamics have been studied with a range of methodologies, including chaotic systems [12–14], clustering [15,16], sample entropy [17,18], and principal components analysis [19–21]. Various asset classes have attracted interest, including equities [22], fixed income [23], and foreign exchange [24]. Researchers have also explored extreme behaviours [25] and the propagation of structural breaks [26,27] in price and volatility time series. The evolutionary nature of volatility, often studied via the framework of volatility clustering and regimes, has been of interest for many years [8,10,28–36]. Overall, such financial research uses many of the same techniques from time series analysis that are used in other domains [37–45].

In recent years, substantial research has focused specifically on the unique dynamics of cryptocurrencies. Various topics of interest include Bitcoin and other cryptocurrencies' price dynamics [46–50], fractal patterns, [51–54], cross-correlation and scaling effects [55–60]. A great deal of this work has addressed how these dynamics have changed over time, in particular during market crises such as the COVID-19 market crash [61–73].

This paper prioritises understanding the changing nature of time-varying parameters that describe the market dynamics of cryptocurrencies. We are particularly interested in a descriptive analysis of these non-stationary dynamics and an identification of different patterns of behaviour, particularly around crises. Our primary parameters of interest are cryptocurrency returns $R(t)$, volatility $\Sigma(t)$, which have been well-studied and known to be highly non-stationary, as well as market size $M(t)$. This latter quantity is particularly relevant to a nascent market such as cryptocurrencies to determine if market dynamics change meaningfully as the market develops.

This work aims to study the changing dynamics of the cryptocurrency market, incorporating and building on much of the aforementioned literature. We take an interest both across our entire 2.5-year window of analysis as well as in several distinct periods, which are outlined in Sect. 2. In Sect. 3, we begin by building on the rich literature studying cryptocurrencies' significant correlations, proposing a new framework to analyse the correlation structure of the market. In Sect. 4, we investigate the relationship between the collective dynamics of the market and the time-varying total market capitalisation of all cryptocurrencies. From there in Sect. 5, we study the relationship between market size and returns and volatility, demonstrating greater quantitative consistency between market size and volatility than size and returns. Finally, Sect. 6 introduces a new technique for understanding the changing spread of volatility across the cryptocurrency market as a whole, terming this *volatility dispersion*. There we introduce a new quantity $\mathrm{Var}(\mathbf{p}(t))$ to describe the spread of volatility across a market that we also observe is highly non-stationary. Overall, we propose a new methodology to understand the evolution of correlation structures, reveal new insights about the relationship between market size and attributes such as collective dynamics, returns, and volatilities, and propose a new way to understanding the spread of volatility with time. Our insights are summarised in Sect. 7.

2 Data

In the proceeding sections, we analyse cryptocurrency data between 01-01-2019 and 30-06-2021. We study 52 cryptocurrencies that possess sufficient histories. In Sect. 3, we partition our analysis into five discrete periods to explore correlation behaviours at varying times. These periods are defined as follows:

1. Pre-COVID: 01-01-2019 to 28-02-2020;
2. Peak COVID: 01-03-2020 to 30-05-2020;
3. Post-COVID: 31-05-2020 to 31-08-2020;
4. Bull: 01-09-2020 to 14-04-2021;
5. Bear: 15-04-2021 to 30-06-2021.

Cryptocurrency data are sourced from https://coinmarketcap.com/. A full list of cryptocurrencies studied in this paper is available in Appendix A.

3 Temporal evolution of market correlation

Like many asset classes in financial markets, cryptocurrency returns are characterised by distributions that exhibit significant tail risk. In particular, their relative infancy and polarising views of the asset's long-term viability make their price behaviour more susceptible to extreme volatility and erratic behaviours [27]. Figure 1a displays the general volatility in log returns, and

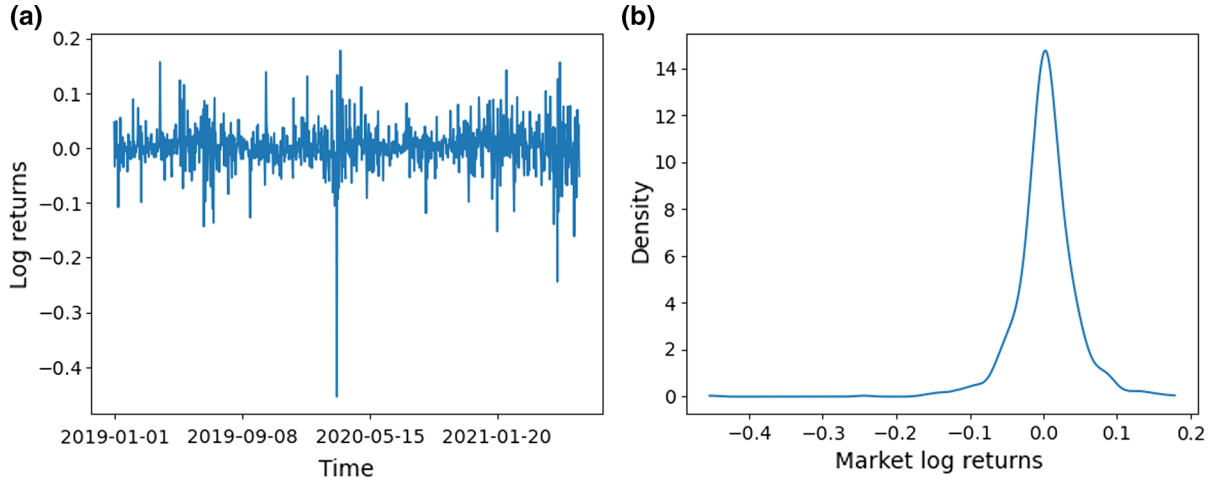

Fig. 1 Cryptocurrency market log returns, depicted (**a**) as a function of time and (**b**) as a distribution

Fig. 1b highlights the negative skew in the returns distribution throughout our analysis window.

Prior research has demonstrated that correlations are most strongly positive during market crises, where many investors engage in the systematic sale of assets [74,75]. This may be due to the growth of quantitatively driven asset managers and their use of algorithms that may induce simultaneous indiscriminate selling. That is, market systematic behaviours, particularly during bear markets, appear to be more pronounced due to an *algorithmic herd mentality*. Behaviour has been particularly volatile since 2020, where after the COVID-19 and BitMEX crash, the cryptocurrency market experienced an unprecedented rise (and subsequent fall) in total market capitalisation. With all this in mind, our objective is to study the temporal evolution of the market's correlation structure and contrast the market's collective similarity between different periods.

3.1 Analysis of collective correlation over time

Let our period of analysis 01-01-2019 to 30-06-2021 be indexed $t = 0, 1, ..., T$, where $T = 911$. Let $c_i(t), i = 1, ..., N, t = 0, ..., T$ be the multivariate time series of cryptocurrency daily closing prices. We first generate a multivariate time series of log returns, $R_i(t), t = 1, ..., T$, as follows:

$$R_i(t) = \log\left(\frac{c_i(t)}{c_i(t-1)}\right). \tag{1}$$

Our primary objects of study in this section are correlation matrices of log returns data across specified periods. These periods may roll forward with time or remain static. Let $a \leq t \leq b$ be such a period of analysis, an interval of $S = b - a + 1$ days. We standardise the cryptocurrency log returns over such a period by defining $\tilde{R}_i(t) = [R_i(t) - \langle R_i \rangle]/\sigma(R_i), a \leq t \leq b$, where $\langle . \rangle$ is an average and $\sigma(.)$ is a standard deviation operator, each computed over the same interval. The correlation matrix Ψ is then defined as follows: let $\tilde{R}$ be a $N \times S$ matrix defined by $\tilde{R}_{it} = \tilde{R}_i(t), i = 1, ..., N, t = a, ..., b$ and define

$$\Psi = \frac{1}{S}\tilde{R}\tilde{R}^T. \tag{2}$$

Explicitly, individual entries are defined by

$$\Psi_{ij} = \frac{\sum_{t=a}^{b}(R_i(t) - \langle R_i \rangle)(R_j(t) - \langle R_j \rangle))}{\left(\sum_{t=a}^{b}(R_i(t) - \langle R_i \rangle)^2 \sum_{t=a}^{b}(R_j(t) - \langle R_j \rangle)^2\right)^{1/2}}, \tag{3}$$

for $1 \leq i, j \leq N$. All entries Ψ_{ij} lie in $[-1, 1]$. If we wish to explicitly note the interval over which these are defined, we may denote this matrix $\Psi^{[a:b]}$. To quantify the total strength in correlation behaviours across the market, we compute an appropriately normalised L^1 norm of the matrix Ψ. That is, let

$$\|\Psi\|_1 = \frac{1}{N^2}\sum_{i,j=1}^{N}|\Psi_{ij}|. \tag{4}$$

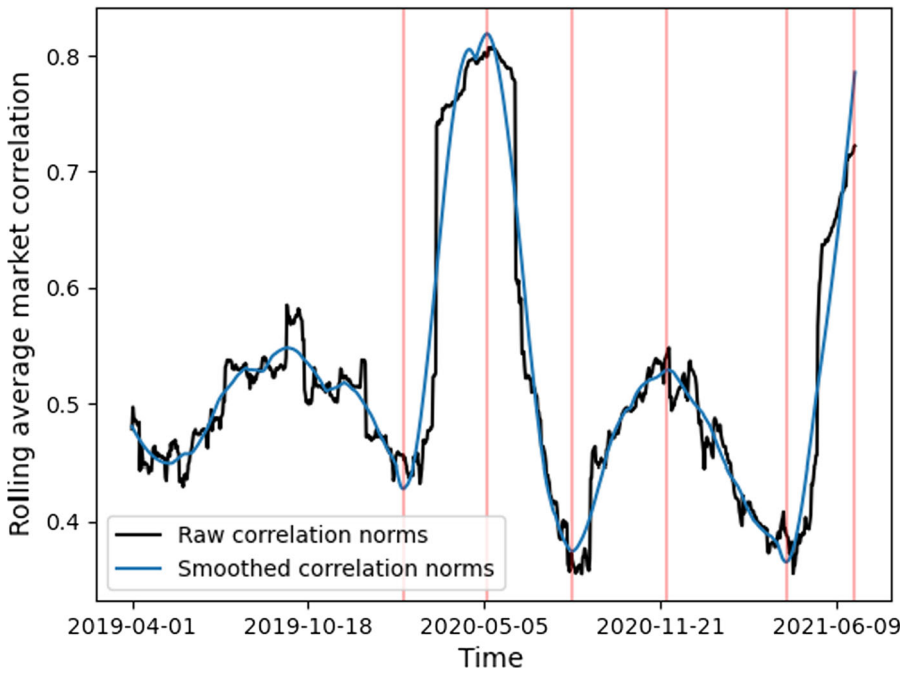

Fig. 2 Time-varying correlation matrix norms, $\nu^{\Psi}(t)$ and its smoothed counterpart $\nu_s^{\Psi}(t)$, as defined in Sect. 3.1. Six non-trivial local maxima and minima are annotated, detected at 03-02-2020, 08-05-2020, 13-08-2020, 29-11-2020, 13-04-2020 and 30-06-2020

This gives the average absolute correlation of all cryptocurrencies over the interval $a \leq t \leq b$. To explore the temporal evolution of collective strength in correlation behaviours, we examine the changes in matrix Ψ as our interval $[a, b]$ rolls forward. Specifically, we set $S = 90$ and compute the time-varying L^1 norm of a 90-day rolling window, $\nu^{\Psi}(t) = \|\Psi^{[t-S+1:t]}\|_1$. We also apply a *Savitzky–Golay* filter to produce a smoothed function $\nu^{\Psi}(t)$. We then apply a recently introduced turning point algorithm [76], detailed in Appendix B, to generate a set of non-trivial local maxima and minima in the total correlation behaviours. While some previous work explores the evolution of the first eigenvalue [21], our methodology is the first we know to study the L^1 norm of the correlation matrix and apply a bespoke turning point algorithm to study non-trivial peak and trough propagation. The norm function $\nu^{\Psi}(t)$, its smoothed analogue $\nu_s^{\Psi}(t)$, and the detected turning points are all displayed in Fig. 2.

Examining Fig. 2 reveals six non-trivial local maxima and minima in the overall magnitude of correlations. Of particular note are the local maximum identified on 08-05-2020 and the minimum on 13-08-2020. These dates reflect the COVID-19 market crash and the subsequent recovery in the cryptocurrency market, respectively. Clearly, correlations are most significant during the crash and weakest during the subsequent growth in the market. The local maximum on 30-06-2020 reflects a significant increase in correlation behaviours during the latter part of our analysis window. This is most likely indicative of the aggressive sell-off in cryptocurrency assets from approximately mid-April until the present. Again, we see a striking pattern where the general growth of the cryptocurrency market is inversely related to the collective strength of correlations.

This is broadly consistent with general economic intuition. There is a substantial amount of literature in the quantitative finance and applied mathematics communities that highlights that an increase in correlation among financial securities is observed during times of crisis [21]. This corresponds to spikes in the first eigenvalue in the correlation matrix—indicating increases in collective market behaviour. When money flows out of the cryptocurrency market, through losses and sale of assets, this would lead to increased correlations and a spike in the first eigenvalue. Thus, the inverse relationship we identify is consistent with what one would expect based on prior research [1].

However, our findings go further and may have fruitful implications for investment managers. Identifying such peaks and troughs in correlation behaviours may provide a decision support tool as to whether they are in a market environment where security selection is

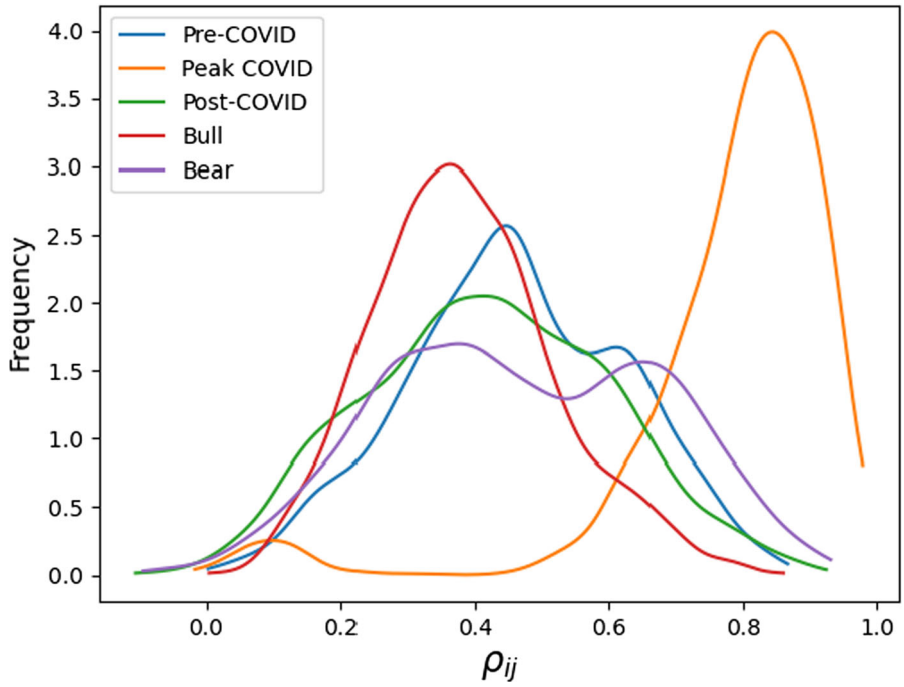

Fig. 3 Kernel density estimates of correlation matrix entries corresponding to our five time partitions: Pre-COVID, Peak COVID, Post-COVID, Bull, and Bear. The Peak COVID correlation matrix has the highest average correlation, indicating extraordinarily high correlations during this period

of greater or less importance. In market environments where correlations are strongly positive, it may be more difficult for managers to produce return streams that exhibit lower market beta. The approximate periodicity we observe in local maxima and minima may encourage cryptocurrency investors to engage in cyclical patterns of more bullish and bearish investing to avoid exposure in riskier periods.

3.2 Comparison of particular periods

To further elucidate the findings of the previous section, we partition our analysis window into five periods and analyse the correlations separately within each of these windows. Let Ψ^{Pre}, Ψ^{Peak}, Ψ^{Post}, Ψ^{Bull} be the correlation matrices obtained across the entire time intervals specified in Sect. 2. In the notation of Sect. 3.1, $\Psi^{\text{Pre}} = \Psi^{[1:59]}$, and so on.

Figure 3 displays kernel density estimates of the entries of our 5 correlation matrices. The findings are generally consistent with Fig. 2. The distribution of elements Ψ_{ij} in the four matrices Ψ^{Pre}, Ψ^{Post}, Ψ^{Bull} and Ψ^{Bear} are highly similar. All four distributions have means in the range 0.38 to 0.46 and exhibit similar variability around their means. By contrast, Ψ^{Peak} has a mean value of 0.78—highlighting the spike in correlations during the COVID-19 crisis. One finding of slight surprise is the Bear partition exhibiting correlations more like the Pre-COVID, Post-COVID, and Bull windows than the Peak COVID window. It is quite possible that if the dates of the Bear period were brought closer to the present day, including beyond 30-06-2021, this distribution would yield a significantly higher average correlation score. The means and standard deviations of the entries of all five matrices are recorded in Table 1.

The observed increase in correlations during market crises is consistent with prior research that has been published in the literature. However, one new observation is the stark contrast in correlation behaviours in periods directly adjacent to an acute period of crisis. One could imagine that the distribution of cryptocurrency correlations transitions gradually downwards following a market crisis. Instead, periods directly after such crises exhibit correlations that are more similar to relatively stable market conditions.

4 Collective dynamics and market size

In the previous section, we observed weaker correlation behaviours while the market grew and stronger correlation behaviours when the market was in crisis or contracting. In this section, we take a closer examination

Table 1 Mean and standard deviation of correlation matrix entries across the five periods, as analysed in Sect. 3.2

Period	Mean of entries Ψ_{ij}	Standard deviation
Pre-COVID	0.456	0.164
Peak COVID	0.784	0.166
Post-COVID	0.421	0.182
Bull	0.383	0.135
Bear	0.421	0.182

of the underlying market effect of collective dynamics and how this relates to the current market size.

First, we apply principal components analysis (PCA) to our time-varying correlation matrices. This procedure learns the linear map Ω such that our standardised returns matrix $\tilde{R}$ is transformed into a matrix of uncorrelated variables Z, that is, $Z = \Omega\tilde{R}$. The rows of Z represent the principal components (PCs) of the matrix $\tilde{R}$, while the rows of Ω consist of the principal component coefficients. The matrix is ordered such that the first row is along the axis of most variation in the data. Its corresponding eigenvalue λ_1 is thus of substantial practical importance, quantifying the greatest extent of variance in the data. It has been referred to as representative of the collective strength of the market [1]. All subsequent PCs, subject to the constraint that they are mutually orthogonal, maximise the variance along their respective axes. Continuing this procedure effectively diagonalises the correlation matrix, $\Psi = EDE^T$, where D is a diagonal matrix of eigenvalues and E is an orthogonal matrix. By (3), Ψ is a symmetric positive semi-definite matrix with all eigenvalues real and non-negative, so we may order them $\lambda_1 \geq ... \geq \lambda_N \geq 0$. Each λ_i quantifies the extent of variance along the ith principal component axis. Thus, we may normalise the eigenvalues by defining $\tilde{\lambda}_i = \frac{\lambda_i}{\sum_{j=1}^N \lambda_j}$ to determine the proportion of all variance accounted for by each step in the PCA.

This quantity is related to the norm of the correlation matrix defined in Sect. 3.1. Indeed, by the spectral theorem, the largest (absolute value) eigenvalue of a symmetric matrix coincides with the matrix' *operator norm* [77]. That is,

$$|\lambda_1| = \|\Psi\|_{op} = \max_{x\in\mathbb{R}^N-\{0\}} \frac{\|\Psi x\|}{\|x\|}. \tag{5}$$

Next, every diagonal entry of Ψ is equal to 1, so the trace of Ψ is equal to N, and thus $\sum_{j=1}^N \lambda_j = N$. Hence $\tilde{\lambda}_1 = \frac{1}{N}\|\Psi\|_{op}$. Thus, both $\tilde{\lambda}_1$ and $\|\Psi\|_1$, as defined in (4), are appropriately normalised norms, with values in [0, 1]. Succinctly put, $\tilde{\lambda}_1$ is a normalised operator norm while $\|\Psi\|_1$ is a normalised L^1 norm. For the remainder of this section, our central object of study is the changing value of $\tilde{\lambda}_1$, which represents both the first proportion of explanatory variance, as well as a normalised operator norm of the matrix.

Just like Sect. 3.1, we set $S = 90$ and analyse a 90-day rolling window. Let $\tilde{\lambda}_1(t)$ be the normalised first eigenvalue of the matrix $\Psi^{[t-S+1:t]}$. We plot this over our analysis window in Fig. 4. There are two primary findings of interest. First, consistent with earlier experiments, there is a spike in $\tilde{\lambda}_1(t)$ during early 2020. This reflects the highly correlated behaviours of cryptocurrencies and indiscriminate selling during the COVID-19 pandemic. Subsequently, $\tilde{\lambda}_1(t)$ declines until May 2021. This most evident decline corresponds to the cryptocurrency bull market, where total cryptocurrency assets grew by several orders of magnitude. Next, there is a significant rise in $\tilde{\lambda}_1(t)$ towards the end of our analysis window, corresponding to higher collective strength in the market. This increase occurred contemporaneously with the aggressive sell-off in cryptocurrency assets, suggesting that there may be a relationship between the size of the market and the strength of the underlying collective dynamics. Broadly, the trajectory of $\tilde{\lambda}_1(t)$ is highly similar to that of $\nu^{\Psi}(t)$ as depicted in Fig. 2.

To investigate further, we quantitatively incorporate the size of the cryptocurrency market changing over time. Let $M_i(t), i = 1, ..., N, t = 0, ..., T$ be the multivariate time series of cryptocurrency market sizes $M_i(t)$ over our analysis window. Due to the significant volatility exhibited by the market, we compute a rolling average of the entire market, defined by $\tilde{M}(t) = \frac{1}{S}\sum_{k=t-S+1}^{t}\sum_{i=1}^{N} M_i(k), t = S, ..., T$. We include the plot of this varying over time in Fig. 4.

While previous work [21] has studied the first eigenvalue in isolation and compared its properties between cryptocurrency and equity markets, this work is the first we know of to examine its relationship with market size over time. In particular, we reveal an inverse relationship between the size of the market and the first eigenvalue of the correlation matrix $\tilde{\lambda}_1(t)$. To quantify this observation, we compute the correlation between the rolling size of the cryptocurrency market, $\tilde{\lambda}_1(t)$, and the collective strength of the market, $\tilde{M}(t)$. The correlation between these two series is computed to be

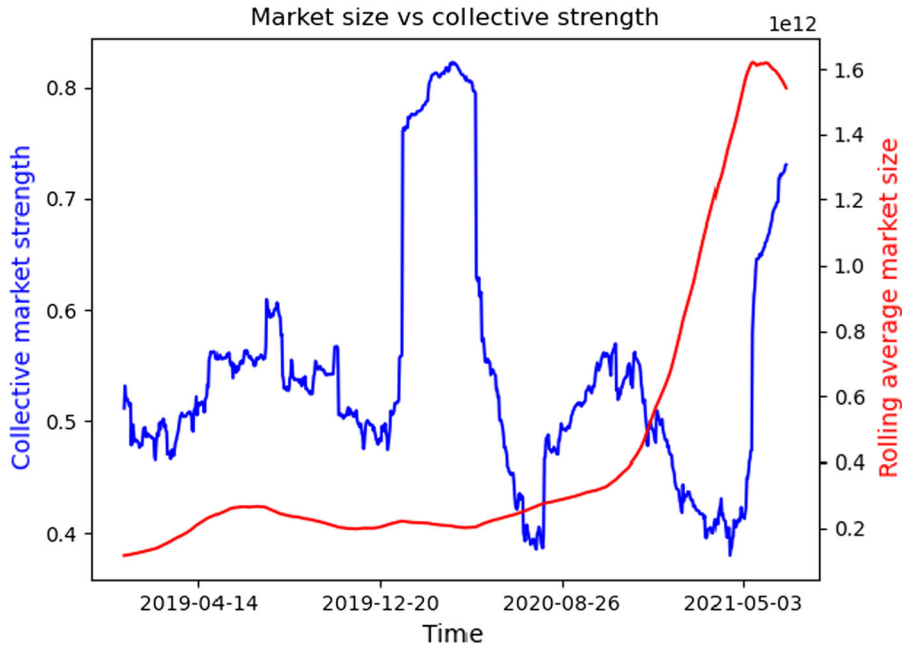

Fig. 4 $\tilde{\lambda}_1(t)$, which represents the collective strength of the market and $\tilde{M}(t)$, which represents the size of the cryptocurrency market, as defined in Sect. 4. An inverse relationship is observed with correlation $\rho^{\tilde{M},\tilde{\lambda}_1} = -0.122$

$\rho^{\tilde{M},\tilde{\lambda}_1} = -0.122$. While this cannot show causation, it suggests a possibility that as the market grows in size, the strength of collective dynamics may decline.

This would be a noteworthy finding with important implications for the future of the cryptocurrency market, especially given the divided views on the future of the market's viability. Suppose one of two contrived scenarios exist; a "bull case" where cryptocurrency prices recover and cryptocurrency becomes a systemically important asset class, and a "bear case" where prices continue to decline and cryptocurrencies lose the interest of institutional investors. Our findings indicate that in the bull case, behaviours may become increasingly fragmented and heterogeneous, and there will be opportunities for skilful security selection to generate portfolio alpha. In the bear case, where prices decline and the size of the market decreases, behaviours may become more homogeneous. This could mean fewer opportunities for alpha generation through security selection, as correlations will be strongly positive.

5 Inconsistency analysis between market size, returns, and volatility

We now extend our study of cryptocurrency market sizes to incorporate their relationship with returns and volatility behaviours individually. Specifically, we investigate the consistency among cryptocurrencies between the attributes of market size, returns and volatility, and how this changes over time. For this purpose, we define three distance matrices and appropriately normalise them.

Let our full period of analysis 01-01-2019 to 30-06-2021 be indexed $t = 0, 1, ..., T$, where $T = 911$. Let $M_i(t)$, $t = 0, ..., T$, be the multivariate time series of market size on each day and let $R_i(t)$, $t = 1, ..., T$, be the multivariate time series of log returns, as defined by (1). Let $S = 90$ days and let $\sigma_i(t)$, $t = S, ..., T$ be the multivariate time series of 90-day rolling volatility. At each t, this is defined as the standard deviation of the log returns of the prior 90 days. Now, we may construct distance matrices for each $t = S, ..., T$ as follows:

$$D^M_{ij}(t) = \frac{1}{S}\left|\sum_{k=t-S+1}^{t}[M_i(k) - M_j(k)]\right|; \tag{6}$$

$$D^R_{ij}(t) = \left|\sum_{k=t-S+1}^{t}[R_i(k) - R_j(k)]\right|; \tag{7}$$

$$D^\Sigma_{ij}(t) = \left|\sigma_i(t) - \sigma_j(t)\right|. \tag{8}$$

Thus, $D^M(t)$, $D^R(t)$ and $D^\Sigma(t)$, respectively, measure discrepancy between cryptocurrencies with respect to average market size, total returns, and rolling volatility, each over the 90-day period concluding on day t, just like our study of correlation norms and market size in

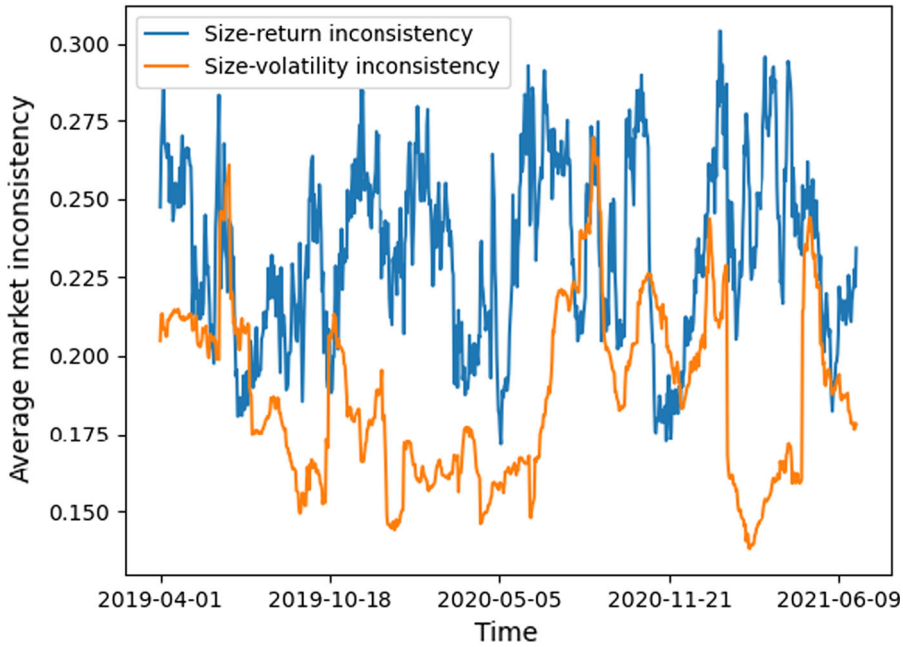

Fig. 5 Time-varying inconsistency matrix norms between market size and returns, $\nu^{INC}_{M,R}$, and size and volatility, $\nu^{INC}_{M,\Sigma}$, as defined in Sect. 5. The size–volatility inconsistency is essentially always lower, indicating more consistency in relationships between cryptocurrencies regarding their size and volatility than their size and returns

Sects. 3.1 and 4. We now convert these three distance matrices into affinity matrices whose elements lie in [0, 1]:

$$A^M_{ij}(t) = 1 - \frac{D^M_{ij}(t)}{\max_{kl}\{D^M_{kl}(t)\}}; \tag{9}$$

$$A^R_{ij}(t) = 1 - \frac{D^R_{ij}(t)}{\max_{kl}\{D^R_{kl}(t)\}}; \tag{10}$$

$$A^\Sigma_{ij}(t) = 1 - \frac{D^\Sigma_{ij}(t)}{\max_{kl}\{D^\Sigma_{kl}(t)\}}. \tag{11}$$

These affinity matrices are appropriately normalised and can be compared directly to study the consistency between cryptocurrency market size and returns or volatility. We generate two *inconsistency matrices* as follows:

$$\mathrm{INC}^{M,R}(t) = A^M(t) - A^R(t); \tag{12}$$

$$\mathrm{INC}^{M,\Sigma}(t) = A^M(t) - A^\Sigma(t). \tag{13}$$

Larger absolute values of the entries of $\mathrm{INC}^{M,R}$ indicate that the relationship between two cryptocurrencies regarding market size and returns is quite different, analogously for $\mathrm{INC}^{M,\Sigma}$. To study the degree of consistency between these attributes in totality across our collection, we compute the L^1 norm of the resulting inconsistency matrices and study how these norms evolve over time. That is, for $t = S, ..., T$, we compute an analogous quantity as defined in (4):

$$\nu^{INC}_{M,R}(t) = \|INC^{M,R}(t)\|; \tag{14}$$

$$\nu^{INC}_{M,\Sigma}(t) = \|INC^{M,\Sigma}(t)\|. \tag{15}$$

Figure 5 displays the time-varying inconsistency norms. It is clear that throughout our period of analysis, there is greater consistency between volatility and size than returns and size, as indicated by the smaller values of $\nu^{INC}_{M,\Sigma}(t)$. This is what one would expect in a more established asset class such as equities. For instance, large-cap equities typically exhibit lower volatility than small-cap equities, creating some consistency between market size and volatility. By contrast, returns and size are shown to be significantly more inconsistent, highlighting that the relationship between size and returns is less clear than that between size and volatility. Furthermore, it suggests that the size of cryptocurrencies is by no means a good representation of the future expectation of returns.

Previous work [27] has used inconsistency matrices to study the extent of inconsistency between returns and volatility. There, we identified the most anoma-

lous individual cryptocurrencies in these attributes. In this work, we take quite a different direction. First, our inconsistency matrices are time-varying. Next, rather than identifying individual cryptocurrencies, we study temporal trends in the collective extent of inconsistency by examining the inconsistency matrices' norms as a function of time. Further, we compare the inconsistency between returns and market size with volatility and market size, incorporating a new parameter. Our finding is also new and of interest, namely that size–returns is more inconsistent than size–volatility. Such a finding could be of great interest to risk managers looking to find factors or exposures their portfolios are most in need of diversifying away from.

6 Temporal changes and dispersion of volatility

Having identified greater consistency in volatility and market size behaviours, we now more closely examine the structure of collective volatility over time. The behaviour of volatility and the general identification of regimes in financial markets is a topic of great interest. Many parametric statistical models, such as regime-switching models, assume a fixed number of volatility regimes for a candidate modelling problem. Often the selection of this number is quite arbitrary. Like all parametric models, if assumptions are misspecified, the resulting estimates can be highly inaccurate.

We take a different approach to the analysis of collective volatility behaviours and detect a new phenomenon that we term *volatility dispersion*. To do so, set our window length as $S = 90$ days. For each $t = S, ..., T$, we may consider the 90-day rolling volatilities $\sigma_i(t)$, as discussed in the previous section. For a fixed t, we normalise the vector $(\sigma_1(t), ..., \sigma_N(t)) \in \mathbb{R}^N$ by its total sum to produce a probability vector $\mathbf{p}(t)$ that measures the concentration of volatility across a selection of cryptocurrencies. That is, let

$$p_i(t) = \frac{\sigma_i(t)}{\sum_{j=1}^{N} \sigma_j(t)}. \tag{16}$$

For example, if $\mathbf{p}(t) = \frac{1}{N}(1, 1, ..., 1) \in \mathbb{R}^N$, this indicates that the 52 cryptocurrencies have identical volatility measured over the past 90 days, while a value of $(1, 0, ..., 0) \in \mathbb{R}^N$ indicates that all volatility is observed in the first currency, with none in any of the others. Thus, $p_i(t)$ is a measure not of the absolute size of volatility but the proportional contribution of one cryptocurrency to the total volatility of the collection.

With these probability vectors $\mathbf{p}(t)$, $t = S, ..., T$, we define a $(T-S+1) \times (T-S+1)$ distance matrix using the Wasserstein distance between these distributions at all points in time. That is, let d^W be the L^1-Wasserstein metric [78], and let

$$D^{vol}(s, t) = d^W(\mathbf{p}(s), \mathbf{p}(t)) \quad \forall s, t, \in [S, ..., T]. \tag{17}$$

We then apply hierarchical clustering to our distance matrix, $D^{vol}(s, t)$, and study the resulting dendrogram (Fig. 6a). It is worthy to note that the Wasserstein distance, unlike the L^1 norm between vectors, does not distinguish probability vectors based on their order; it is essentially a distance between vectors as sets (possibly with repetition). We make this choice because we are not interested in distinguishing periods where a particular cryptocurrency is highly volatile, but whether the volatility is, broadly speaking, spread out or concentrated among the collection as a whole.

Figure 6a highlights several interesting findings. First, two volatility clusters are identified, one dominant cluster of volatility behaviours (with two subclusters) and a smaller cluster that is highly different to the rest of the collection. The latter cluster consists of probability vectors $\mathbf{p}(t)$ generated at times within the COVID-19 market crash. Interestingly, this cluster does not display overwhelming self-similarity; instead, it is relatively diffuse but exhibits significant differences with the concentrated subcluster of the majority cluster. The anomalous behaviours of the COVID-19 market crash are further demonstrated, with such profound distances to other (highly variable) periods in the market over the past several years.

To further investigate the nature of this split, we perform a closer analysis of the probability vectors $\mathbf{p}(t)$ over time. Considered as a distribution over [0, 1], we compute the time-varying *intra-volatility variance* $\mathrm{Var}(\mathbf{p}(t))$. We display this as a function of time in Fig. 6b. This supports the separation of behaviours seen in Fig. 6a. The COVID-19 market crash exhibits markedly lower variance among the individual rolling volatilities $p_i(t)$. Interestingly, this is also observed towards the end of our analysis window. A lower value of the variance of $\mathbf{p}(t)$ means that the contribution to the variance of each individual cryptocurrency is more uniform across the collection. That is, the proportion of the market's total volatility is more spread out among

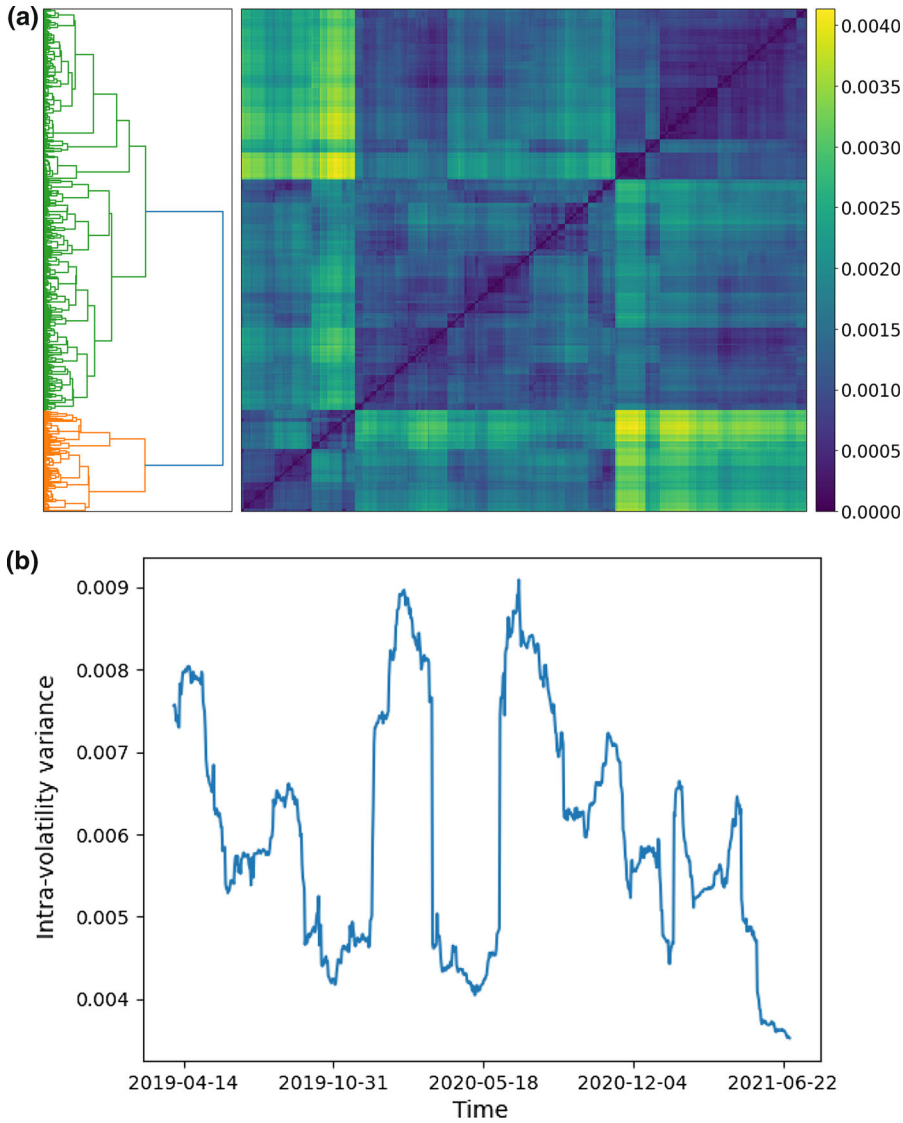

Fig. 6 In (**a**), we perform hierarchical clustering on the matrix D^{vol} between normalised vectors of rolling volatility. The dendrogram groups dates $s, t \in [S, ..., T]$ according to similarity between their corresponding vectors $\mathbf{p}(s)$ and $\mathbf{p}(t)$. Two clusters are observed, with the secondary cluster associated with times t observed during COVID-19. In (**b**), we elucidate this finding more closely, plotting the variance of the probability vector $\mathbf{p}(t)$ over time. The variance of this vector is lower during COVID-19, indicating vectors where all volatilities are closer together. That is, the total volatility of the market is spread out more over all the constituent cryptocurrencies. We term this increased volatility dispersion

all the cryptocurrencies during the COVID-19 market crisis. While it is predictable that the absolute volatility would spike during a crisis, it is an unexpected finding that there would be less deviation between the different volatilities of individual cryptocurrencies—essentially, everything is similarly volatile together.

In Appendix C, we present two theoretical results on the distances $D^{vol}(s, t)$ between normalised volatility vectors and intra-volatility variance $\mathrm{Var}(\mathbf{p}(t))$. These propositions identify the *uniform distribution of volatility* $\mathbf{p}_0 = \frac{1}{N}(1, 1, ..., 1) \in \mathbb{R}^N$ and the *one-shot distributions of volatility* $\mathbf{q}_k = (0, ..., 0, 1, 0, ..., 0)$ as the two extremal possible spreads of volatility. The uniform has the lowest intra-volatility variance, a one-shot has the greatest intra-volatility variance, and the greatest possible value of the Wasserstein distance is between the uniform and a one-shot distribution. These propositions demonstrate that, in a precise sense, our study of volatility dispersion investigates the extent that rolling volatility vectors sit between two extremes: the case where all the volatility of the market is uniformly distributed across every asset, and a case where all volatility is concentrated in a single asset.

Fund managers are often warned of the potential issues when over-fitting models to study parameters such as volatility. The volatility dispersion framework

we introduce in this section supports the "Occam's Razor" principle familiar to many in statistical learning. Rather than trying to capture the true complexity of the volatility process, using just two regimes to capture low and high volatility periods, may work best. However, the interpretations of volatility dispersion are new and different from those gained from traditional volatility clustering, indicating periods where volatility is relatively uniform and less avoidable regardless of asset selection.

7 Conclusion

In Sect. 3, we study the time-varying evolution of correlations among our collection of cryptocurrencies, and explicitly compare the distribution of correlation coefficients over five discrete time windows. The most notable findings were the spike in correlations during the COVID-19 market crash and the drop in correlations during the subsequent bull market run in late 2020 to early 2021. Broadly, both experiments in this section allude to a clear association between strong collective correlation among cryptocurrencies and periods of declining value in the market.

In Sect. 4, we investigate the aforementioned association more closely. First, we explore the time-varying explanatory variance of the correlation matrix's first eigenvalue $\tilde{\lambda}_1(t)$, where a noticeable spike is seen during the COVID-19 pandemic. Next, we directly compare $\tilde{\lambda}_1(t)$ with the rolling size of the cryptocurrency market, and show a negative correlation therein. This suggests that as cryptocurrency assets rise, the strength of collective dynamics may weaken. Thus, in a scenario where cryptocurrency assets rise significantly and the asset class gains further prominence, the strength of collective dynamics may decline—leading to more heterogeneous behaviours. This would place greater importance on high-quality security selection when investing in cryptocurrencies.

In Sect. 5, our experiments reveal greater consistency in volatility and market size behaviours, wherein cryptocurrencies similar in market size are more likely to exhibit commensurate levels of volatility. By contrast, there is less consistency between cryptocurrency size and returns—suggesting that the cryptocurrencies' size does not provide a good indication of future expected returns.

Finally, in Sect. 6, we study the structure of volatility behaviours over time, applying hierarchical clustering to the distances between distributions of rolling volatility at all points in time. Our technique suggests that there are two volatility patterns—times where the total volatility of the market is more dispersed across the entire collection, and times where it is more concentrated in fewer particular cryptocurrencies. We reveal that the COVID-19 market crash not only features higher volatilities in general, but that the total volatility is more evenly spread across all individual cryptocurrencies. This technique could be used as an accompanying tool to estimate the number of regimes in more traditional, parametric regime-switching models in the econometric and statistical modelling literature. Volatility dispersion also provides independent confirmation of and a new approach to studying increased heterogeneity of market dynamics during crises, complementing our study of collective correlations. It is effectively a more direct measure of assets being uniformly volatile together than the first eigenvalue of the correlation matrix.

Many possibilities exist for future work building upon the techniques and findings of this paper. First, one could study the associations between market size and the strength of collective dynamics in the cryptocurrency market with alternative methodologies. For example, suitable distances between normalised trajectories could replace the use of correlations. Second, one could compare studied relationships between the size of the cryptocurrency market and its underlying dynamics with those of more traditional asset classes. Third, given the cryptocurrency market's relative infancy, our findings may turn out to be transient; if the market continues to grow, perhaps the inverse relationship between total market size and collective dynamics will not hold in future crises.

Several of our new methodologies and findings may have particular promise in future research and applications. Our time-varying analysis of the total extent of inconsistency between parameters could reveal suitable predictors to incorporate in trading strategies either aimed at maximising returns or minimising volatility. Our volatility dispersion analysis is also promising. Unlike typical methods of volatility clustering or regime-switching models, we compare similarity between all windows of time, not just adjacent periods. In this paper, our analysis has been entirely descriptive, but future work could employ it in a predictive

context where we expect volatility to be rather uniform across the market. Such times could prompt an entire withdrawal of funds to safe-haven assets such as gold or cash, as a uniform spread of volatility could mean any investment in the cryptocurrency market would carry significant risk. Finally, volatility dispersion could be applied to other financial and economic securities beyond cryptocurrencies. For instance, one could identify clusters of macroeconomic behaviour using data such as interest rates, GDP, inflation, unemployment, and others. One could explore the dispersion of these factors individually, or identify clusters of economic behaviour with a higher dimensional distance measure where a variety of metrics are incorporated.

Overall, this paper reveals several key relationships between cryptocurrencies' collective dynamics, market size, returns, and volatility and analyses these behaviours over time. During COVID-19 and towards 30 June 2021, correlation behaviours are stronger, and volatility is more uniformly spread across the entire market. Individual cryptocurrencies' market sizes are shown to be more consistent with volatility than returns, while the total market size is inversely associated with the quantifiable strength of collective dynamics. Both the lack of consistency between market size and returns and the high correlations across the market during crises present significant challenges to investors aiming to select optimal portfolios of cryptocurrencies on either a long or short-term basis.

Acknowledgements The authors would like to thank Georg Gottwald for some helpful discussions.

Author contributions Each author is an equal first author, playing an equal role in every aspect of the manuscript.

Funding No specific funding was received for this manuscript.

Data availability and material All data are publicly available at https://coinmarketcap.com/.

Declarations

Conflict of interest The authors have no conflicts of interest to report.

Appendix A: Securities analysed

In Table 2, we list the 52 cryptocurrencies analysed in this paper, both their tickers and names.

Table 2 Cryptocurrency tickers and names

Ticker	Coin Name	Ticker	Coin Name
BTC	Bitcoin	WAVES	WAVES
ETH	Ethereum	CEL	Celsius
BNB	Binance Coin	DASH	Dash
ADA	Cardano	ZEC	Zcash
XRP	XRP	MANA	Decentraland
DOGE	Dogecoin	ENJ	Enjin Coin
BCH	Bitcoin Cash	HOT	Holo
LTC	Litecoin	QNT	Quant
LINK	Chainlink	KCS	KuCoin
ETC	Ethereum Classic	NEXO	Nexo
XLM	Stellar	BAT	Basic Attention Token
THETA	Theta	ZIL	Zilliqa
VET	VeChain	BTG	Bitcoin Gold
FIL	Filecoin	BNT	Bancor
TRX	Tron	ONT	Ontology
SMR	Monero	ZEN	Horizen
EOS	EOS	SC	Siacoin
CRO	Crypto.com	DGB	Digibyte
MKR	Maker	QTUM	QTUM
BSV	Bitcoin SV	CHSB	SwissBorg
NEO	NEO	ZRX	0x
XTZ	Tezos	RVN	Ravencoin
MIOTA	IOTA	OMG	OMG Network
DCR	Decred	NANO	Nano
HT	Huobi Token	ICX	ICON
XEM	NEM	FTM	Fantom

Appendix B: Turning point algorithm

In this section, we provide more details for the identification of turning points (local maxima and minima) of the smoothed function $\nu_s^{\Psi}(t)$, used in Sect. 3. We begin by linearly adjusting this function to ensure that the global minimum of the resulting function is 0. Specifically, let $\hat{\nu}(t) = \nu_s^{\Psi}(t) - \min_t\{\nu_s^{\Psi}(t)\}$, $t = S, ..., T$.

Following [76], we apply a two-step algorithm to the smoothed and minimum-adjusted $\hat{\nu}(t)$. The first step produces an alternating sequence of local minima (troughs) and local maxima (peaks), which may include some immaterial turning points. The second step refines this sequence according to chosen conditions and parameters. The most important conditions

to initially identify a peak or trough, respectively, are the following:

$$\hat{\nu}(t_0) = \max\{\hat{\nu}(t) : \max(0, t_0 - l) \le t \le \min(t_0 + l, T)\}, \tag{18}$$

$$\hat{\nu}(t_0) = \min\{\hat{\nu}(t) : \max(0, t_0 - l) \le t \le \min(t_0 + l, T)\}, \tag{19}$$

where l is a parameter to be chosen. We follow [76] in their choice of $l = 17$. Defining peaks and troughs according to this definition alone has some flaws, such as the potential for two consecutive peaks.

Instead, we implement an inductive procedure to choose an alternating sequence of peaks and troughs. Suppose t_0 is the last determined peak. We search in the period $t > t_0$ for the first of two cases: if we find a time $t_1 > t_0$ that satisfies (19) as well as a non-triviality condition $\hat{\nu}(t_1) < \hat{\nu}(t_0)$, we add t_1 to the set of troughs and proceed from there. If we find a time $t_1 > t_0$ that satisfies (18) and $\hat{\nu}(t_0) \ge \hat{\nu}(t_1)$, we ignore this lower peak as redundant; if we find a time $t_1 > t_0$ that satisfies (18) and $\hat{\nu}(t_1) > \hat{\nu}(t_0)$, we remove the peak t_0, replace it with t_1 and continue from t_1. A similar process applies from a trough at t_0.

At this point, the time series is assigned an alternating sequence of troughs and peaks. However, some turning points are immaterial and should be removed. Let $t_1 < t_3$ be two peaks, necessarily separated by a trough. We select a parameter $\delta = 0.2$, and if the *peak ratio*, defined as $\frac{\hat{\nu}(t_3)}{\hat{\nu}(t_1)} < \delta$, we remove the peak t_3. If two consecutive troughs t_2, t_4 remain, we remove t_2 if $\hat{\nu}(t_2) > \hat{\nu}(t_4)$, otherwise remove t_4. That is, if the second peak has size less than δ of the first peak, we remove it.

Finally, we use the same *log-gradient* function between times $t_1 < t_2$, defined as

$$\text{log-grad}(t_1, t_2) = \frac{\log \hat{\nu}(t_2) - \log \hat{\nu}(t_1)}{t_2 - t_1}. \tag{20}$$

The numerator equals $\log(\frac{\hat{\nu}(t_2)}{\hat{\nu}(t_1)})$, a "logarithmic rate of change". Unlike the standard rate of change given by $\frac{\hat{\nu}(t_2)}{\hat{\nu}(t_1)} - 1$, the logarithmic change is symmetrically between $(-\infty, \infty)$. Let t_1, t_2 be adjacent turning points (one a trough, one a peak). We choose a parameter $\epsilon = 0.01$; if

$$|\,\text{log-grad}(t_1, t_2)| < \epsilon, \tag{21}$$

that is, the average logarithmic change is less than 1%, we remove t_2 from our sets of peaks and troughs. If t_2 is not the final turning point, we also remove t_1. After these refinement steps, we are left with an alternating sequence of non-trivial peaks and troughs—these are marked in Fig. 2.

Appendix C: Theoretical properties of volatility dispersion

In the two propositions below, let $\mathbf{p}(t)$ be an arbitrary probability vector of volatilities. That is, $\mathbf{p}(t) = (a_1, ..., a_N) \in \mathbb{R}^N$ with $a_i \ge 0$ and $\sum a_i = 1$. Let the *uniform distribution of volatility* be the probability vector given by $\mathbf{p}_0 = \frac{1}{N}(1, 1, ..., 1) \in \mathbb{R}^N$. For any $j = 1, ..., N$, let the *one-shot distribution of volatility* on asset j be $\mathbf{q}_k = (0, ..., 0, 1, 0, ..., 0)$, where the k-th coordinate is a 1. That is, $\mathbf{p}_0$ signifies a uniform spread of volatility across all assets, while $\mathbf{q}_k$ represents a spread of volatility exhibited entirely by one asset, with all other assets having zero volatility.

We prove two properties concerning the distance $D^{vol}(s, t) = d^W(\mathbf{p}(s), \mathbf{p}(t))$ and the intra-volatility variance $\text{Var}(\mathbf{p}(t))$, as defined in Sect. 6.

Proposition 1 *Let $\mathbf{p}(t)$ be a volatility vector. The least possible value of Var$(\mathbf{p}(t))$ is zero, exhibited uniquely by the uniform $\mathbf{p}_0$. The greatest possible value is $1 - \frac{1}{N^2}$, exhibited only by the one-shot distributions $\mathbf{q}_k$.*

Proof The variance can be expressed as

$$\text{Var}(\mathbf{p}(t)) = \sum_{i=1}^{N} \left(a_i - \frac{1}{N}\right)^2, \tag{22}$$

which is clearly at most zero, equality if and only if all $a_i = \frac{1}{N}$, that is, for $\mathbf{p}(t) = \mathbf{p}_0$.

An alternative formulation of the variance is

$$\text{Var}(\mathbf{p}(t)) = \sum_{i=1}^{N} a_i^2 - \frac{1}{N^2} \tag{23}$$

$$\le \sum_{i=1}^{N} a_i^2 + 2\sum_{i<j} a_i a_j - \frac{1}{N^2} \tag{24}$$

$$= \left(\sum_{i=1}^{N} a_i\right)^2 - \frac{1}{N^2} \tag{25}$$

$$= 1 - \frac{1}{N^2}. \tag{26}$$

Equality is obtained if and only if all cross terms $a_i a_j = 0$ for $i \neq j$. This means that only one value of a_i may be nonzero, say $i = k$. By the condition that a_i sum to 1, this nonzero value must be 1, so $\mathbf{p}(t) = \mathbf{q}_k$. □

Proposition 2 *The greatest possible distance $D^{vol}(s,t) = d^W(\mathbf{p}(s), \mathbf{p}(t))$ occurs between the uniform distribution $\mathbf{p}_0$ and a one-shot distribution $\mathbf{q}_k$. Specifically,*

$$D^{vol}(s,t) \leq \frac{2}{N}\left(1 - \frac{1}{N}\right), \tag{27}$$

and equality occurs only if $\{\mathbf{p}(t), \mathbf{p}(s)\} = \{\mathbf{p}_0, \mathbf{q}_k\}$ as sets.

Proof Let $\mathbf{p}(t) = (a_1, ..., a_N)$ be an arbitrary probability vector of volatilities. As discussed in Sect. 6, our implementation of the Wasserstein distance is essentially between vectors treated as sets (possibly with repetition in their elements). Specifically, we associate to each probability vector $\mathbf{p}(t)$ a probability measure defined as a weighted sum of Dirac delta measures

$$\mu = \frac{1}{N}\sum_{i=1}^{N} \delta_{a_i}. \tag{28}$$

When μ and ν are probability measures on $\mathbb{R}$ that have cumulative distribution functions F and G, there is a simple representation of the Wasserstein distance as follows [78]:

$$d^W(\mu, \nu) = \int_0^1 |F^{-1} - G^{-1}| dx, \tag{29}$$

where F^{-1} and G^{-1} are the quantile functions associated with F and G, respectively [79]. In our scenario, we integrate μ in (28) to compute F. Let $a_{(1)}, a_{(2)}, ..., a_{(N)}$ be the unique reordering of $a_1, ..., a_N$ in non-decreasing order. That is, $a_{(1)} \leq a_{(2)} \leq ... \leq a_{(N)}$ and $\{a_{(1)}, a_{(2)}, ..., a_{(N)}\} = \{a_1, ..., a_N\}$. Integrating μ, we can see that F is a piecewise-constant increasing step function:

$$F = \sum_{i=1}^{N-1} \frac{i}{n} \mathbb{1}_{[a_{(i)}, a_{(i+1)})} + \mathbb{1}_{[a_{(N)}, \infty)}. \tag{30}$$

Thus, its associated quantile function is

$$F^{-1} = \sum_{i=1}^{N} a_{(i)} \mathbb{1}_{(\frac{i-1}{n}, \frac{i}{n}]}. \tag{31}$$

Now let $\mathbf{p}(s) = (b_1, ..., b_N)$ be an alternative probability vector of volatilities. Let ν, G and G^{-1} be its associated measure, cumulative distribution function and quantile function associated, respectively. In particular,

$$G^{-1} = \sum_{i=1}^{N} b_{(i)} \mathbb{1}_{(\frac{i-1}{n}, \frac{i}{n}]}. \tag{32}$$

Thus,

$$d^W(\mathbf{p}(s), \mathbf{p}(t)) = \int_0^1 |F^{-1} - G^{-1}| dx \tag{33}$$

$$= \int_0^1 \left| \sum_{i=1}^{N} (a_{(i)} - b_{(i)}) \mathbb{1}_{(\frac{i-1}{n}, \frac{i}{n}]} \right| dx \tag{34}$$

$$= \frac{1}{N}\sum_{i=1}^{N} |a_{(i)} - b_{(i)}| \tag{35}$$

$$= \frac{1}{N}\sum_{i=1}^{N-1} |a_{(i)} - b_{(i)}| + \frac{1}{N}|(a_{(N)} - \frac{1}{N}) - (b_{(N)} - \frac{1}{N})| \tag{36}$$

$$\leq \frac{1}{N}\sum_{i=1}^{N-1} (a_{(i)} + b_{(i)}) + \frac{1}{N}((a_{(N)} - \frac{1}{N}) + (b_{(N)} - \frac{1}{N})) \tag{37}$$

$$= \frac{1}{N}\sum_{i=1}^{N} a_{(i)} + \frac{1}{N}\sum_{i=1}^{N} b_{(i)} - \frac{2}{N^2} \tag{38}$$

$$= \frac{2}{N} - \frac{2}{N^2} \tag{39}$$

$$= \frac{2}{N}\left(1 - \frac{1}{N}\right). \tag{40}$$

These inequalities hold by the triangle inequality, and the non-negativity of $a_{(i)}$, $b_{(i)}$, $a_{(N)} - \frac{1}{N}$ and $b_{(N)} - \frac{1}{N}$. These last two are non-negative as the condition $\sum a_i = 1$ implies $a_{(N)} \geq \frac{1}{N}$.

Finally, suppose equality holds. Then for each $i = 1, ..., N-1$, we must have $|a_{(i)} - b_{(i)}| = a_{(i)} + b_{(i)}$. This means at least one of $a_{(i)}$ or $b_{(i)}$ is zero, for each i. Without loss of generality, suppose $a_{(N-1)} = 0$. As $0 \leq a_{(1)} \leq a_{(2)} \leq ... \leq a_{(N-1)} \leq a_{(N)}$, this immediately shows $a_{(i)} = 0$ for all $i = 1, ..., N-1$. So $a_{(N)} = 1$ and up to reordering, $\mathbf{p}(t)$ is a one-shot probability vector $(1, 0, ..., 0)$. Thus, $\mathbf{p}(t) = \mathbf{q}_k$ for some k.

In addition, equality forces $|(a_{(N)} - \frac{1}{N}) - (b_{(N)} - \frac{1}{N})| = (a_{(N)} - \frac{1}{N}) + (b_{(N)} - \frac{1}{N})$. As both terms are non-negative, this implies one of $a_{(N)} - \frac{1}{N}$, $b_{(N)} - \frac{1}{N}$ must be zero. But we've established $a_{(N)} - \frac{1}{N} = 1 - \frac{1}{N}$

is nonzero. Thus, $b_{(N)} - \frac{1}{N}$ must be zero. So $b_{(N)} = \frac{1}{N}$. And yet $b_{(N)}$ is the largest of the b_i, elements that sum to 1. This implies all the b_i must be equal to $\frac{1}{N}$. So $\mathbf{p}(s)$ must be the uniform distribution $\mathbf{p}_0$. It is then routine to demonstrate that $d^W(\mathbf{q}_k, \mathbf{p}_0)$ attains the theoretical upper bound of $\frac{2}{N}\left(1 - \frac{1}{N}\right)$. □

Remark 1 The above proof can be made more conceptual by noticing that the condition $\sum_{i=1}^{N} a_i = 1$ shows directly that

$$\int_0^1 F^{-1} dx = \frac{1}{N}. \tag{41}$$

Thus, there is a quick bound on $\int_0^1 |F^{-1} - G^{-1}| dx$ of $\frac{2}{N}$. To tighten the bound, one can notice that the quantile functions F^{-1} and G^{-1} necessarily have a "rectangle in common" in their area under the curve. This rectangle has area $\frac{1}{N^2}$, which provides the strengthening of the bound. If x is the coordinate on the domain of F^{-1} and G^{-1} and y is the coordinate on the codomain, this rectangle can be described as $1 - \frac{1}{N} \leq x \leq 1, 0 \leq y \leq \frac{1}{N}$ (so is in fact a square).

Viewed together, these propositions demonstrate that, in a precise sense, our study of volatility dispersion investigates the extent that real rolling volatility data sits between two extremes: the case where all the volatility of the market is uniformly distributed across every asset, and a case where all volatility is concentrated in a single asset.

References

1. Fenn, D.J., Porter, M.A., Williams, S., McDonald, M., Johnson, N.F., Jones, N.S.: Temporal evolution of financial-market correlations. Phys. Rev. E **84**(2), 026109 (2011). https://doi.org/10.1103/physreve.84.026109
2. Münnix, M.C., Shimada, T., Schäfer, R., Leyvraz, F., Seligman, T.H., Guhr, T., Stanley, H.E.: Identifying States of a Financial Market, Scientific Reports **2**(1), 1–6 (2012). https://doi.org/10.1038/srep00644
3. Vicente, R., de Toledo, C.M., Leite, V.B., Caticha, N.: Underlying dynamics of typical fluctuations of an emerging market price index: the Heston model from minutes to months. Phys. A: Stat. Mech. Appl. **361**(1), 272 (2006). https://doi.org/10.1016/j.physa.2005.06.095
4. Wang, G.J., Xie, C.: Cross-correlations between the CSI 300 spot and futures markets. Nonlinear Dyn. **73**(3), 1687 (2013). https://doi.org/10.1007/s11071-013-0895-7
5. Wang, G.J., Xie, C.: Cross-correlations between WTI crude oil market and U.S. stock market: a perspective from econophysics. Acta Phys. Pol. B **43**(10), 2021 (2012). https://doi.org/10.5506/aphyspolb.43.2021
6. Miśkiewicz, J.: Network analysis of cross-correlations on forex market during crises globalisation on forex market. Entropy **23**(3), 352 (2021). https://doi.org/10.3390/e23030352
7. Pan, R.K., Sinha, S.: Collective behavior of stock price movements in an emerging market. Phys. Rev. E **76**(4), 046116 (2007). https://doi.org/10.1103/physreve.76.046116
8. Lamoureux, C.G., Lastrapes, W.D.: Persistence in variance, structural change, and the GARCH model. J. Bus. Econ. Stat. **8**(2), 225 (1990). https://doi.org/10.2307/1391985
9. Chu, J., Chen, S., Nadarajah, S., Osterrieder, J.: GARCH modelling of cryptocurrencies. J. Risk Fin. Manag. **10**(4), 17 (2017)
10. Kumar, A.S., Anandarao, S.: Volatility spillover in crypto-currency markets: some evidences from GARCH and wavelet analysis. Phys. A: Stat. Mech. Appl. **524**, 448 (2019). https://doi.org/10.1016/j.physa.2019.04.154
11. Dahlhaus, R.: Fitting time series models to nonstationary processes. Annals Stat. **25**(1), 1 (1997). https://doi.org/10.1214/aos/1034276620
12. Cai, G., Yu, H., Li, Y.: Localization of compact invariant sets of a new nonlinear finance chaotic system. Nonlinear Dyn. **69**(4), 2269 (2012). https://doi.org/10.1007/s11071-012-0425-z
13. Tacha, O.I., Munoz-Pacheco, J.M., Zambrano-Serrano, E., Stouboulos, I.N., Pham, V.T.: Determining the chaotic behavior in a fractional-order finance system with negative parameters. Nonlinear Dyn. **94**(2), 1303 (2018). https://doi.org/10.1007/s11071-018-4425-5
14. Szumiński, W.: Integrability analysis of chaotic and hyperchaotic finance systems. Nonlinear Dyn. **94**(1), 443 (2018). https://doi.org/10.1007/s11071-018-4370-3
15. Heckens, A.J., Krause, S.M., Guhr, T.: Uncovering the dynamics of correlation structures relative to the collective market motion. J. Stat. Mech. Theory Exper. **2020**(10), 103402 (2020). https://doi.org/10.1088/1742-5468/abb6e2
16. James, N., Menzies, M.: Association between COVID-19 cases and international equity indices. Phys. D: Nonlinear Phenom. **417**, 132809 (2021). https://doi.org/10.1016/j.physd.2020.132809
17. Wu, Y., Shang, P., Xia, J.: Inverse sample entropy analysis for stock markets. Nonlinear Dyn. **103**(1), 741 (2021). https://doi.org/10.1007/s11071-020-06118-z
18. Chen, Y., Lin, A.: Weighted link entropy and multiscale weighted link entropy for complex time series. Nonlinear Dyn. **105**(1), 541 (2021). https://doi.org/10.1007/s11071-021-06599-6
19. Laloux, L., Cizeau, P., Bouchaud, J.P., Potters, M.: Noise dressing of financial correlation matrices. Phys. Rev. Lett. **83**(7), 1467 (1999). https://doi.org/10.1103/physrevlett.83.1467
20. Kim, D.H., Jeong, H.: Systematic analysis of group identification in stock markets. Phys. Rev. E **72**(4), 046133 (2005). https://doi.org/10.1103/physreve.72.046133
21. James, N.: Dynamics, behaviours, and anomaly persistence in cryptocurrencies and equities surrounding COVID-19.

Phys. A: Stat. Mech. Appl. **570**, 125831 (2021). https://doi.org/10.1016/j.physa.2021.125831
22. Wilcox, D., Gebbie, T.: An analysis of cross-correlations in an emerging market. Phys. A: Stat. Mech. Appl. **375**(2), 584 (2007). https://doi.org/10.1016/j.physa.2006.10.030
23. Driessen, J., Melenberg, B., Nijman, T.: Common factors in international bond returns. J. Int. Money Fin. **22**(5), 629 (2003). https://doi.org/10.1016/s0261-5606(03)00046-9
24. Ausloos, M.: Statistical physics in foreign exchange currency and stock markets. Phys. A: Stat. Mech. Appl. **285**(1–2), 48 (2000). https://doi.org/10.1016/s0378-4371(00)00271-5
25. Qi, D., Majda, A.J.: Using machine learning to predict extreme events in complex systems. Proc. Natl. Acad. Sci. **117**(1), 52 (2019). https://doi.org/10.1073/pnas.1917285117
26. Telli, Ş, Chen, H.: Structural breaks and trend awareness-based interaction in crypto markets. Phys. A: Stat. Mech. Appl. **558**, 124913 (2020). https://doi.org/10.1016/j.physa.2020.124913
27. James, N., Menzies, M., Chan, J.: Changes to the extreme and erratic behaviour of cryptocurrencies during COVID-19. Phys. A: Stat. Mech. Appl. **565**, 125581 (2021). https://doi.org/10.1016/j.physa.2020.125581
28. Shah, D., Isah, H., Zulkernine, F.: Stock Market Analysis: A Review and Taxonomy of Prediction Techniques. Int. J. Fin. Stud. **7**(2), 26 (2019). https://doi.org/10.3390/ijfs7020026
29. Kirchler, M., Huber, J.: Fat tails and volatility clustering in experimental asset markets. J. Econ. Dyn. Control **31**(6), 1844 (2007). https://doi.org/10.1016/j.jedc.2007.01.009
30. Baillie, R.T., Morana, C.: Modelling long memory and structural breaks in conditional variances: An adaptive FIGARCH approach. J. Econ. Dyn. Control **33**(8), 1577 (2009). https://doi.org/10.1016/j.jedc.2009.02.009
31. Hamilton, J.D.: A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica **57**(2), 357 (1989). https://doi.org/10.2307/1912559
32. Lavielle, M., Teyssière, G.: Adaptive Detection of Multiple Change-Points in Asset Price Volatility. In: *Long Memory in Economics* (Springer Berlin Heidelberg, 2007), pp. 129–156. https://doi.org/10.1007/978-3-540-34625-8_5
33. Prakash, A., James, N.,Menzies,M., Francis, G.: Structural clustering of volatility regimes for dynamic trading strategies. Appl. Math. Finance (2022) (in print). https://doi.org/10.1080/1350486X.2021.2007146
34. Guidolin, M., Timmermann, A.: Asset allocation under multivariate regime switching. J. Econ. Dyn. Control **31**(11), 3503 (2007). https://doi.org/10.1016/j.jedc.2006.12.004
35. Yang, Q.Q., Ching, W.K., Gu, J.W., Siu, T.K.: Market-making strategy with asymmetric information and regime-switching. J. Econ. Dyn. Control **90**, 408 (2018). https://doi.org/10.1016/j.jedc.2018.04.003
36. de Zeeuw, A., Zemel, A.: Regime shifts and uncertainty in pollution control. J. Econ. Dyn. Control **36**(7), 939 (2012). https://doi.org/10.1016/j.jedc.2012.01.006
37. Hethcote, H.W.: The mathematics of infectious diseases. SIAM Rev. **42**(4), 599 (2000). https://doi.org/10.1137/s0036144500371907
38. James, N., Menzies, M.: Trends in COVID-19 prevalence and mortality: a year in review. Phys. D: Nonlinear Phenom. **425**, 132968 (2021). https://doi.org/10.1016/j.physd.2021.132968
39. Vazquez, A.: Polynomial growth in branching processes with diverging reproductive number. Phys. Rev. Lett. **96**(3), 038702 (2006). https://doi.org/10.1103/physrevlett.96.038702
40. Mendes, C.F., Beims, M.W.: Distance correlation detecting Lyapunov instabilities, noise-induced escape times and mixing. Phys. A: Stat. Mech. Appl. **512**, 721 (2018). https://doi.org/10.1016/j.physa.2018.08.028
41. Mendes, C.F.O., da Silva, R.M., Beims, M.W.: Decay of the distance autocorrelation and Lyapunov exponents. Phys. Rev. E **99**(6) (2019). https://doi.org/10.1103/physreve.99.062206
42. Rizzi, R., Mahata, P., Mathieson, L., Moscato, P.: Hierarchical clustering using the arithmetic-harmonic cut: complexity and experiments. PLoS ONE **5**(12), e14067 (2010). https://doi.org/10.1371/journal.pone.0014067
43. Shang, K., Yang, B., Moore, J.M., Ji, Q., Small, M.: Growing networks with communities: A distributive link model. Chaos: Interdiscipl. J. Nonlinear Sci. **30**(4), 041101 (2020). https://doi.org/10.1063/5.0007422
44. James, N., Menzies, M., Radchenko, P.: COVID-19 second wave mortality in Europe and the United States. Chaos: Interdiscipl. J. Nonlinear Sci. **31**, 031105 (2021). https://doi.org/10.1063/5.0041569
45. Machado, J.A.T., Lopes, A.M.: Rare and extreme events: the case of COVID-19 pandemic. Nonlinear Dyn. (2020). https://doi.org/10.1007/s11071-020-05680-w
46. Chu, J., Nadarajah, S., Chan, S.: Statistical analysis of the exchange rate of bitcoin. PLOS ONE **10**(7), e0133678 (2015). https://doi.org/10.1371/journal.pone.0133678
47. Lahmiri, S., Bekiros, S.: Chaos, randomness and multifractality in Bitcoin market. Chaos Solitons Fractals **106**, 28 (2018). https://doi.org/10.1016/j.chaos.2017.11.005
48. Kondor, D., Pósfai, M., Csabai, I., Vattay, G.: Do the rich get richer? An empirical analysis of the Bitcoin transaction network. PLoS ONE **9**(2), e86197 (2014). https://doi.org/10.1371/journal.pone.0086197
49. Bariviera, A.F., Basgall, M.J., Hasperué, W., Naiouf, M.: Some stylized facts of the Bitcoin market. Phys. A: Stat. Mech. Appl. **484**, 82 (2017). https://doi.org/10.1016/j.physa.2017.04.159
50. Alvarez-Ramirez, J., Rodriguez, E., Ibarra-Valdez, C.: Long-range correlations and asymmetry in the Bitcoin market. Phys. A: Stat. Mech. Appl. **492**, 948 (2018). https://doi.org/10.1016/j.physa.2017.11.025
51. Stosic, D., Stosic, D., Ludermir, T.B., Stosic, T.: Multifractal behavior of price and volume changes in the cryptocurrency market. Phys. A: Stat. Mech. Appl. **520**, 54 (2019). https://doi.org/10.1016/j.physa.2018.12.038
52. Stosic, D., Stosic, D., Ludermir, T.B., Stosic, T.: Exploring disorder and complexity in the cryptocurrency space. Phys. A: Stat. Mech. Appl. **525**, 548 (2019). https://doi.org/10.1016/j.physa.2019.03.091
53. Manavi, S.A., Jafari, G., Rouhani, S., Ausloos, M.: Demythifying the belief in cryptocurrencies decentralized aspects. A study of cryptocurrencies time cross-correlations with common currencies, commodities and financial indices. Phys. A: Stat. Mech. Appl. **556**, 124759 (2020). https://doi.org/10.1016/j.physa.2020.124759

54. Ferreira, P., Kristoufek, L., de Area Leão Pereira, E.J.: DCCA and DMCA correlations of cryptocurrency markets. Phys. A: Stat. Mech. Appl. **545**, 123803 (2020). https://doi.org/10.1016/j.physa.2019.123803
55. Drożdż, S., Gębarowski, R., Minati, L., Oświęcimka, P., Wątorek, M.: Bitcoin market route to maturity? Evidence from return fluctuations, temporal correlations and multiscaling effects. Chaos Interdiscipl. J. Nonlinear Sci. **28**(7), 071101 (2018). https://doi.org/10.1063/1.5036517
56. Drożdż, S., Minati, L., Oświęcimka, P., Stanuszek, M., Wątorek, M.: Signatures of the crypto-currency market decoupling from the Forex. Future Internet **11**(7), 154 (2019). https://doi.org/10.3390/fi11070154
57. Drożdż, S., Minati, L., Oświęcimka, P., Stanuszek, M., Wątorek, M.: Competition of noise and collectivity in global cryptocurrency trading: Route to a self-contained market. Chaos Interdiscipl. J. Nonlinear Sci. **30**(2), 023122 (2020). https://doi.org/10.1063/1.5139634
58. Gębarowski, R., Oświęcimka, P., Wątorek, M., Drożdż, S.: Detecting correlations and triangular arbitrage opportunities in the Forex by means of multifractal detrended cross-correlations analysis. Nonlinear Dyn. **98**(3), 2349 (2019). https://doi.org/10.1007/s11071-019-05335-5
59. Drożdż, S., Kwapień, J., Oświęcimka, P.: Complexity in economic and social systems. Entropy **23**(2), 133 (2021). https://doi.org/10.3390/e23020133
60. Wątorek, M., Kwapień, J., Drożdż, S.: Financial return distributions: past, present, and COVID-19. Entropy **23**(7), 884 (2021). https://doi.org/10.3390/e23070884
61. Wątorek, M., Drożdż, S., Kwapień, J., Minati, L., Oświęcimka, P., Stanuszek, M.: Multiscale characteristics of the emerging global cryptocurrency market. Phys. Rep. (2020). https://doi.org/10.1016/j.physrep.2020.10.005
62. Corbet, S., Larkin, C., Lucey, B.: The contagion effects of the COVID-19 pandemic: evidence from gold and cryptocurrencies. Fin. Res. Lett. **35**, 101554 (2020). https://doi.org/10.1016/j.frl.2020.101554
63. Conlon, T., McGee, R.: Safe haven or risky hazard? Bitcoin during the Covid-19 bear market. Fin. Res. Lett. **35**, 101607 (2020). https://doi.org/10.1016/j.frl.2020.101607
64. Conlon, T., Corbet, S., McGee, R.J.: Are cryptocurrencies a safe haven for equity markets? An international perspective from the COVID-19 pandemic. Res. Int. Bus. Fin. **54**, 101248 (2020). https://doi.org/10.1016/j.ribaf.2020.101248
65. Ji, Q., Zhang, D., Zhao, Y.: Searching for safe-haven assets during the COVID-19 pandemic. Int. Rev. Fin. Anal. **71**, 101526 (2020). https://doi.org/10.1016/j.irfa.2020.101526
66. Lahmiri, S., Bekiros, S.: The impact of COVID-19 pandemic upon stability and sequential irregularity of equity and cryptocurrency markets. Chaos Solitons Fractals **138**, 109936 (2020). https://doi.org/10.1016/j.chaos.2020.109936
67. Zhang, D., Hu, M., Ji, Q.: Financial markets under the global pandemic of COVID-19. Fin. Res. Lett. **36**, 101528 (2020). https://doi.org/10.1016/j.frl.2020.101528
68. He, Q., Liu, J., Wang, S., Yu, J.: The impact of COVID-19 on stock markets. Econ. Polit. Stud. **8**, 1 (2020). https://doi.org/10.1080/20954816.2020.1757570
69. Zaremba, A., Kizys, R., Aharon, D.Y., Demir, E.: Infected markets: Novel coronavirus, government interventions, and stock return volatility around the globe. Fin. Res. Lett. **35**, 101597 (2020). https://doi.org/10.1016/j.frl.2020.101597
70. Akhtaruzzaman, M., Boubaker, S., Sensoy, A.: Financial contagion during COVID-19 crisis. Fin. Res. Lett. **38**, 101604 (2020). https://doi.org/10.1016/j.frl.2020.101604
71. Okorie, D.I., Lin, B.: Stock markets and the COVID-19 fractal contagion effects. Fin. Res. Lett. **38**, 101640 (2020). https://doi.org/10.1016/j.frl.2020.101640
72. Naeem, M.A., Bouri, E., Peng, Z., Shahzad, S.J.H., Vo, X.V.: Asymmetric efficiency of cryptocurrencies during COVID19. Phys. A: Stat. Mech. Appl. **565**, 125562 (2021). https://doi.org/10.1016/j.physa.2020.125562
73. Curto, J.D.: To keep faith with homoskedasticity or to go back to heteroskedasticity? The case of FATANG stocks. Nonlinear Dyn. **104**(4), 4117 (2021). https://doi.org/10.1007/s11071-021-06535-8
74. Sandoval, L., Franca, I.D.P.: Correlation of financial markets in times of crisis. Phys. A: Stat. Mech. Appl. **391**(1–2), 187 (2012). https://doi.org/10.1016/j.physa.2011.07.023
75. James, N., Menzies, M.: A new measure between sets of probability distributions with applications to erratic financial behavior, J. Stat. Mech. 123404 (2021). https://doi.org/10.1088/1742-5468/ac3d91
76. James, N., Menzies, M.: COVID-19 in the United States: trajectories and second surge behavior. Chaos Interdiscipl. J. Nonlinear Sci. **30**, 091102 (2020). https://doi.org/10.1063/5.0024204
77. Rudin, W.: *Functional Analysis* (McGraw-Hill Science, New York, 1991)
78. del Barrio, E., Giné, E., Matrán, C.: Central limit theorems for the Wasserstein distance between the empirical and the true distributions. Annals Probab. **27**(2), 1009 (1999). https://doi.org/10.1214/aop/1022677394
79. Gilchrist, W.: *Statistical Modelling with Quantile Functions* (Chapman and Hall/CRC, Boca Raton, 2000). https://doi.org/10.1201/9781420035919

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