entropy

Special Issue Reprint

Signatures of Maturity in Cryptocurrency Market

Edited by
Stanisław Drożdż, Jarosław Kwapień and Marcin Wątorek

www.mdpi.com/journal/entropy

MDPI

Signatures of Maturity in Cryptocurrency Market

Signatures of Maturity in Cryptocurrency Market

Editors

Stanisław Drożdż
Jarosław Kwapień
Marcin Wątorek

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Editors
Stanisław Drożdż
Cracov University of Technology
Kraków, Poland

Jarosław Kwapień
Polish Academy of Sciences
Kraków, Poland

Marcin Wątorek
Cracow University of Technology
Kraków, Poland

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Entropy* (ISSN 1099-4300) (available at: https://www.mdpi.com/journal/entropy/special_issues/sig_cryp).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-0365-8574-1 (Hbk)
ISBN 978-3-0365-8575-8 (PDF)
doi.org/10.3390/books978-3-0365-8575-8

Contents

About the Editors

Stanisław Drożdż

Stanisław Drożdż, Professor of physics and head of the Complex Systems Theory Department at Institute of Nuclear Physics (INP), Polish Academy of Sciences, and professor of computer science at the Cracow University of Technology. He received M.S. (1978) and Ph.D. (1982) degrees in physics, both from Jagiellonian University in Kraków, and a DSc (1988) in theoretical physics from INP. In 1994, he received the Polish state title of professor. His long-term scientific studies abroad include a role at Forschungszentrum Juelich, Germany, from 1983 to 1986 as a post-doc and from 1989 to 1991 as a senior scientist, as well as at the University of Illinois at Urbana-Champaign, USA, from 1993 to 1994 and at Bonn University, Germany, from 2001 to 2002 as a visiting professor. His research interests and activities include the quantum many-body problem, nonlinear dynamics, general aspects of complexity, brain-research-related issues, dynamics of financial markets, and quantitative linguistics.

Jarosław Kwapień

Jarosław Kwapień , Assoc. Prof., graduated from Jagiellonian University in Kraków with a master's degree in physics and earned his Ph.D. degree in physics (2001) based on his work on topics at the interface between statistical physics and neuroscience at Institute of Nuclear Physics (IFJ PAN) in Kraków, Poland. As a post-doctoral fellow at Forschungszentrum Juelich, Germany, he started his research on topics in econophysics and data science. Currently he works at IFJ PAN and continues to publish works in econophysics, but also in complex networks, quantitative linguistics, and other interdisciplinary topics.

Marcin Wątorek

Marcin Wątorek, Ph.D., is an assistant professor at the Faculty of Computer Science and Telecommunications at Cracow University of Technology and postdoctoral research scientist at the Faculty of Physics, Astronomy and Applied Computer Science in Jagiellonian University. He received an M.S. in mathematics from Cracow University of Technology (2015) and a Ph.D. in physics from the Institute of Nuclear Physics Polish Academy of Sciences (2020). The title ofhis doctoral dissertation was "Quantitative characteristics of complexity in the world cryptocurrency market," with supervisor professor Stanisław Drożdż. His research interest include the dynamics of financial markets (especially cryptocurrencies) and patterns in the human brain.

entropy

MDPI

Article

What Is Mature and What Is Still Emerging in the Cryptocurrency Market?

Stanisław Drożdż [1,2,†], Jarosław Kwapień [2,†] and Marcin Wątorek [1,*,†]

[1] Faculty of Computer Science and Telecommunications, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland; stanislaw.drozdz@ifj.edu.pl
[2] Complex Systems Theory Department, Institute of Nuclear Physics, Polish Academy of Sciences, ul. Radzikowskiego 152, 31-342 Kraków, Poland; jaroslaw.kwapien@ifj.edu.pl
* Correspondence: marcin.watorek@pk.edu.pl
† These authors contributed equally to this work.

Abstract: In relation to the traditional financial markets, the cryptocurrency market is a recent invention and the trading dynamics of all its components are readily recorded and stored. This fact opens up a unique opportunity to follow the multidimensional trajectory of its development since inception up to the present time. Several main characteristics commonly recognized as financial stylized facts of mature markets were quantitatively studied here. In particular, it is shown that the return distributions, volatility clustering effects, and even temporal multifractal correlations for a few highest-capitalization cryptocurrencies largely follow those of the well-established financial markets. The smaller cryptocurrencies are somewhat deficient in this regard, however. They are also not as highly cross-correlated among themselves and with other financial markets as the large cryptocurrencies. Quite generally, the volume V impact on price changes R appears to be much stronger on the cryptocurrency market than in the mature stock markets, and scales as $R(V) \sim V^{\alpha}$ with $\alpha \gtrsim 1$.

Keywords: blockchain; cryptocurrencies; time series; fluctuations; correlations; multifractality; market maturity; market impact

Citation: Drożdż, S.; Kwapień, J.; Wątorek, M. What Is Mature and What Is Still Emerging in the Cryptocurrency Market? *Entropy* **2023**, *25*, 772. https://doi.org/10.3390/e25050772

Academic Editor: Marcel Ausloos

Received: 21 April 2023
Revised: 4 May 2023
Accepted: 6 May 2023
Published: 9 May 2023

1. Introduction

Studying the world cryptocurrency market is welcome for many reasons. Up to now, it constitutes the most spectacular and influential application of the distributed ledger technology called the blockchain, which, in the underlying peer-to-peer network, allows for the same access to information for all participants [1,2]. Research on blockchain technology is also unique because all related data are publicly available in the form of the history of every operation performed on the network. Furthermore, the tick-by-tick data for each transaction made on the cryptocurrency exchange are freely available using the application programming interfaces (APIs) of a given exchange.

As far as the financial, economic, and, in general terms, social aspects of cryptocurrencies are concerned, a basic related question that arises is whether such digital products can be considered as a commonly accepted means of exchange [3–5]. This is a complex issue involving many social, economical, and technological factors, such as trust, perceived risk, peer opinions, transaction security, network size effect, supply elasticity, and so on. However, also from a dynamical perspective, for this to apply, a certain level of maturity expressed in terms of market efficiency, liquidity, stability, size, and other characteristics is required [6,7]. Moreover, the developed markets show several statistical properties that newly established emerging markets often lack. Among such properties, one can list the so-called financial stylized facts: heavy tails of the probability distribution functions of fixed-time returns, long-term memory of volatility, a hierarchical structure of the asset cross-correlations, multifractality, and a stable (or meta-stable) price impact function [8–11].

There is growing quantitative evidence that the cryptocurrency market continuously advances on a route to maturity understood as sharing its statistical properties with the traditional financial markets. For instance, the most popular and oldest cryptocurrency, bitcoin (BTC), has passed through two stages of the shaping of its probability distribution function (pdf). It started as an extremely volatile asset with pdf tails that used to decline according to a power law, with the exponent reaching almost the Lévy-stable regime (the Lévy parameter $\alpha \approx 2$) on short time scales over the years 2012–2013, but then, already in the years 2014-2015, the tails of its pdf became thinner and reached the inverse cubic behavior that is observed universally in the traditional financial markets [12]. From that moment on, BTC has maintained this property over the subsequent years [6,13,14]. The difference between BTC and traditional assets is that the inverse cubic behavior of the BTC pdf tails was reported to be preserved up to much longer sampling intervals due to their less frequent trading [12]. Similar effects were seen for other major crypto currencies, such as ETH [12,15]. Since BTC and the other cryptocurrencies are traded on many independent platforms that differ in trading frequency, the pdf properties of the same cryptocurrency can be different on different platforms [6]. This is quite a unique trait of the cryptocurrencies not observed, for example, in the stock markets and Forex. Heavy pdf tails were also found in time series of volume traded in time units [16,17], even in the case of cryptocurrencies [18,19]. These two quantities—the log-returns and volume—are related to each other, because the size of a trade can have a profound impact on price variation: large trades lead to large price jumps on average (although this relation might be more subtle [20–22]). Some authors argue that price impact assumes a functional form with a square-root dependence of the log-returns on volume [23–25] but others are cautious [21,22,26].

The long-term memory of volatility fluctuations is responsible for the effect of volatility clustering, i.e., periods of a volatile market with large-amplitude fluctuations are interwoven with periods of relatively tranquil dynamics. In addition, the volatility autocorrelation is of a power-law form [27]. This property has been seen in all financial markets and has also been found in cryptocurrency dynamics [14]. The range of memory is comparable in this case with the range for the stock and Forex markets [28,29]. The scale-free form of the autocorrelation function is connected to fractality, which also requires long-term or long-range correlations to be self-similar. The log-return fluctuations for all the traditional financial markets studied so far show multiscaling together with some other quantities, such as inter-transaction times [30–32]. Consistently, multifractal properties have been observed in the cryptocurrency market returns and inter-transaction times for different assets [6,18,33–39]. Apart from univariate multiscaling, its bivariate version has also been reported between log-returns for different cryptocurrencies: BTC and ETH [40].

Apart from correlations in time, asset–asset cross-correlations play an important role in the shaping of the financial market structure as they lead to the emergence of the hierarchical organization of the markets as well as coupling between different markets [41–44]. While the hierarchical cross-correlations among the assets traded on the same market are a clear indicator of market maturity, the role of potential couplings between different markets must be interpreted with care. This is because either the independent dynamics of a market or the profound coupling of a market with the world's leading markets, being the two opposite cases, can potentially be interpreted in favor of market maturity. The former because independence can be viewed as strength and as a possibility for using the assets traded on such a market as a safe haven in hedging strategies [45,46], and the latter because it suggests that such a market is a well-rooted part of global financial markets. However, intuitively, neither of these extremes seems to represent the notion of maturity well enough. It is more justified to view market maturity as the ability to switch its dynamics between independence and compliance because such a behavior can better reflect the complexity that one may expect to be the property characterizing a developed market. This is why neither the effect of the cryptocurrency market decoupling from Forex reported in [29] nor the effects of the cryptocurrency market independence [47–51] and strong coupling between the cryptocurrencies and traditional financial markets reported in [52–55], respectively, can alone be a signature of maturity. It is

rather the opposite: only such flexible dynamics swinging between idiosyncrasy and a strong subjugation of the market to an actual global trend can be a manifestation of market maturity.

In this work, stress was put on the investigation of current statistical properties of cryptocurrency log-returns and volume from the perspective of how these properties differ from their counterparts in the traditional financial markets: the stock markets, Forex, and commodity markets. One has to be aware, however, that the statistical approach constitutes only a segment of the issues related to market maturity.

2. Methods and Results

2.1. Empirical Dataset

The data set studied contains 1 min quotations of 70 cryptocurrencies that were among the most actively traded on the Binance exchange [56], which had the largest market share in 2022 [57], over the period from 1 January 2020 to 31 December 2022 (3 years). The quotes are expressed in USD Tether (USDT), a stablecoin linked to the US dollar, and its value is close to USD 1 by design [58]. Basic time series statistics corresponding to these 70 cryptocurrencies are collected in Table 1. For a time series of price quotations $Q(t_i)$, $i = 1, \ldots, T$, the equally spaced logarithmic returns $R_{\Delta t}(t_i) = \log Q(t_i) - \log Q(t_{i-1})$, where $t_i - t_{i-1} = \Delta t$, are derived. Figure 1 shows the evolution of the cumulative log-returns $\hat{R}_{\Delta t}(t_i) = \sum_{i=1}^{i} R_{\Delta t}(t_i)$ during the whole period covered by the data. In accordance with the actual cryptocurrency price quotes, in 2021, the whole market experienced a transition from the bull phase to the bear phase.

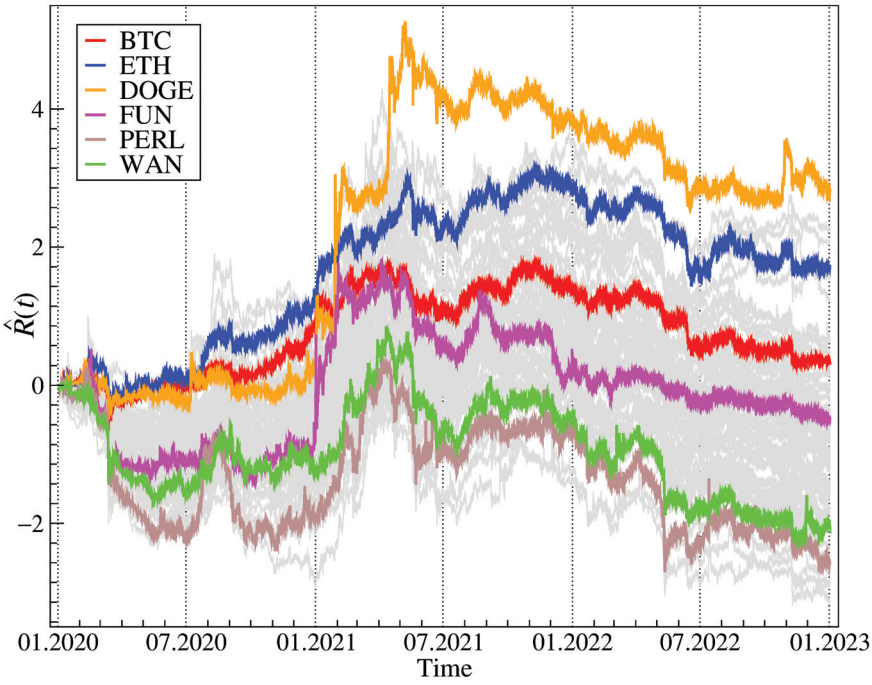


Figure 1. Evolution of the cumulative log-returns $\hat{R}(t)$ of the 70 cryptocurrencies over the time period from 1 January 2020 to 31 December 2022. The colors of two of the most liquid cryptocurrencies and a few other distinguished ones are indicated explicitly. The bulk of the cryptocurrencies is shown in the background (grey lines).

Table 1. Basic statistics of the cryptocurrencies considered in this study: the average inter-transaction time δt, the fraction of zero returns in time series %0, the average volume value traded per minute W, and market capitalization C on 1 January 2023. For the cryptocurrency name list, see Table A1 in Appendix A.

Ticker	δt [s]	%0	W [USDT]	C [$\times 10^6$ USD]	Ticker	δt [s]	%0	W [USDT]	C [$\times 10^6$ USD]
BTC	0.04	0.003	1,683,710	320,025	LINK	0.41	0.095	84,423	2856
ADA	0.24	0.121	172,891	8621	LTC	0.41	0.142	80,441	5096
ALGO	0.78	0.117	24,320	1267	MATIC	0.32	0.166	100,100	6638
ANKR	1.84	0.195	10,762	151	MFT	5.01	0.425	2436	54
ARPA	2.75	0.165	6082	33	MTL	3.16	0.400	5122	46
ATOM	0.58	0.109	42,048	2710	NEO	1.45	0.194	18,893	451
BAND	2.13	0.175	8285	49	NKN	2.99	0.425	5807	56
BAT	1.53	0.162	10,543	251	NULS	4.44	0.442	2845	12
BCH	0.70	0.140	48,288	1869	OMG	0.83	0.178	24,235	146
BEAM	5.30	0.433	2089	14	ONE	0.97	0.227	21,983	133
BNB	0.17	0.095	276,261	39,052	ONG	5.53	0.482	2297	71
CELR	1.77	0.292	10,843	68	ONT	1.28	0.149	16,136	134
CHZ	0.59	0.232	51,827	672	PERL	5.00	0.431	2406	7
COS	2.63	0.455	3575	18	QTUM	1.58	0.179	14,178	196
CTXC	3.42	0.464	3942	33	REN	2.72	0.207	6232	62
DASH	1.44	0.206	14,543	468	RLC	2.80	0.293	6090	95
DENT	1.24	0.353	16,417	68	RVN	1.82	0.202	9699	232
DOCK	5.39	0.455	2135	12	STX	4.42	0.416	3847	288
DOGE	0.20	0.173	247,343	9317	TFUEL	2.09	0.353	10,411	189
DUSK	2.97	0.441	3994	34	THETA	0.64	0.173	35,023	733
ENJ	1.17	0.225	21,114	243	TOMO	3.84	0.316	3581	24
EOS	0.53	0.147	59,616	948	TROY	3.20	0.381	3347	23
ETC	0.58	0.099	63,736	2188	TRX	0.46	0.142	71,306	5041
ETH	0.10	0.010	853,284	146,967	VET	0.52	0.093	55,362	1163
FET	2.65	0.255	7,909	75	VITE	4.22	0.469	3078	18
FTM	0.50	0.174	63,723	556	WAN	7.24	0.303	1609	34
FUN	3.91	0.538	2911	66	WAVES	1.19	0.177	19,265	144
HBAR	1.57	0.268	11,765	957	WIN	1.01	0.283	26,244	72
HOT	0.96	0.237	22,543	250	XLM	0.78	0.165	33,309	1894
ICX	2.64	0.306	6951	135	XMR	1.62	0.184	14,164	2707
IOST	1.40	0.199	14,551	129	XRP	0.21	0.071	229,976	17,055
IOTA	1.53	0.168	12,077	478	XTZ	1.08	0.137	19,407	663
IOTX	1.52	0.266	11,894	203	ZEC	1.15	0.240	20,010	597
KAVA	1.57	0.155	12,888	198	ZIL	1.03	0.145	20,195	258
KEY	2.83	0.358	4310	15	ZRX	3.04	0.214	5674	128

2.2. Cumulative Distribution Functions of Returns and Volume

The cumulative distribution function (cdf) $P(X > r_{\Delta t})$ can be calculated from the normalized returns $r_{\Delta t}(t_i) = (R_{\Delta t}(i) - \mu)/\sigma$, with μ and σ denoting sample mean and standard deviation, respectively. A form of this distribution varies among the markets and assets, but some interesting properties can be observed. There are generally three factors that shape it: the first one is liquidity, the second one is trading speed, and the third one is the overall market volatility [59]. If one focuses on a specific market, the most liquid assets show a faster decline in $P(X > r_{\Delta t})$ with $r_{\Delta t}$ than the less liquid ones for a given Δt [60]. However, most of the assets traded on mature markets reveal a power-law dependence of $P(X > r_{\Delta t})$ for some range of Δt [23,27,60–62]:

$$P(X > r_{\Delta t}) \sim |r_{\Delta t}|^{-\gamma}, \tag{1}$$

with $\gamma \approx 3$. It is observed for short sampling intervals and it is persistent for a range of Δt due to the existence of strong inter-asset correlations. This inverse cubic power-law dependence breaks for sufficiently long Δt and the cdf tails converge to the expected normal distribution. The speed of information processing on a given market also has influence on the crossover Δt. Since this speed increases with time as new technologies enter the service, we observe a gradual decrease in the crossover Δt across decades. The speed of market trading allows for a larger transaction number in time units, so this factor accelerates the market time even more [60]. The emerging markets, where investment strategies require the accommodation of significant risk, are thus highly volatile. The cdfs of the asset returns in this case often show heavy tails with the scaling exponent $\gamma \ll 3$, sometimes even in the Lévy-stable regime. In such markets, the inverse cubic behavior of $P(X > r_{\Delta t})$ may occur for some assets only, whereas, for the other assets, it cannot be found at all. This is why such extreme tails are often considered to be an indicator of market immaturity [14].

Based on the average inter-transaction time δt, we categorized the considered cryptocurrencies into three groups: I, the most frequently traded cryptocurrencies ($\delta t < 1s$); II, the cryptocurrencies with the average trading frequency ($1s \leq \delta t < 2s$); and III, the least frequently traded cryptocurrencies ($\delta t \geq 2s$). Then, we calculated the average cdfs for the cryptocurrencies belonging to each group. We show these cdfs in Figure 2 (left panel, dotted lines) together with the cdfs for a few selected individual cryptocurrencies (solid lines). Their form can be compared with the inverse cubic power-law model denoted by a dashed line. It can be seen that the average distributions have their tail close to a power law, with the exponent γ being close to 3. The most liquid cryptocurrencies—BTC and ETH—develop tails that show a cross-over from the power-law regime to a CLT-like regime for relatively small values of $|r_{\Delta t}|$ compared to both the average cdfs and to less frequently traded individual cryptocurrencies such as FUN, PERL, and WAN. The case of Dogecoin, which has the smallest slope in the middle of the distribution and, at the same time, does not have the thickest tail, is special. On the one hand, it can be included among the main cryptocurrencies due to the high frequency of transactions and capitalization, and, on the other, it was the subject of possible price manipulation through Elon Musk's tweets [63,64].

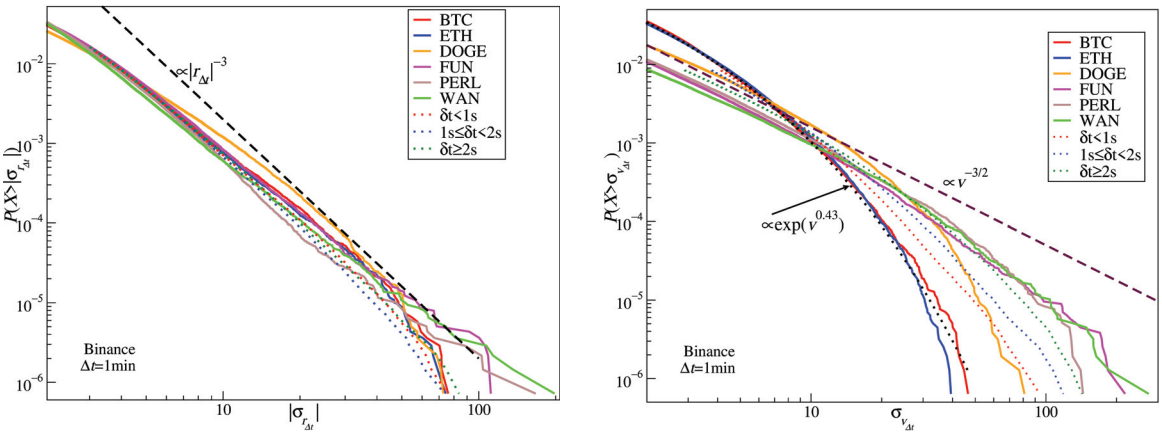


Figure 2. Cumulative distribution functions of the absolute normalized log-returns $r_{\Delta t}$ (**left**) and the normalized volume traded $v_{\Delta t}$ (**right**) for $\Delta t = 1$ min in units of the respective standard deviations σ for the selected cryptocurrencies with the highest liquidity (BTC and ETH) or the heaviest tails (DOGE, FUN, PERL, and WAN). The average cumulative distribution functions for the cryptocurrencies with the average inter-transaction time fulfilling the relations $\delta t < 1s$ (Group I, dotted red), $1s \leq \delta t < 2s$ (Group II, dotted blue), and $\delta t \geq 2s$ (Group III, dotted green) are also shown. Power laws with the scaling exponents γ and β assuming values typical for the financial markets—$\gamma = 3$ and $\beta = 3/2$—are denoted by dashed lines. There is also a stretched exponential function fitted to the $v_{\Delta t}$ distributions for BTC and ETH on the right (black dotted line).

Another quantity that is frequently observed to be power-law-distributed is normalized volume traded in time unit $v_{\Delta t}(t_i) = (V_{\Delta t}(i) - \mu)/\sigma$ [16,23]:

$$P(X > v_{\Delta t}) \sim v_{\Delta t}^{-\beta}. \tag{2}$$

In this case, the exponent is much lower than for the absolute returns and corresponds to the Lévy-stable regime: $\beta < 2$. It was argued that there exists a simple relation between both the exponents: $\beta = \gamma/2$ [23]. Figure 2 (right panel) shows the cumulative distribution functions for $v_{\Delta t}$ for the same individual cryptocurrencies and their Groups I-III as in Figure 2 (left panel). Now, the cdfs for BTC and ETH do not develop power-law tails. A model that best fits them is the stretched exponential function $P(X > v_{\Delta t}) \sim \exp \sigma_v^{-\eta}$ with $\eta = 0.43$. However, in the case of less frequently traded cryptocurrencies, which belong to Group III, one can observe the power-law relation. What makes the results obtained here different from their counterparts for, for instance, the stock markets, is that one does not find any cryptocurrency with its cdf being a power law with the exponent 3/2; the cdf tails decrease considerably faster here.

2.3. Price Impact

At this point, it is worthwhile to consider a possible causal relation between the returns and the volume despite the fact that no clear relation can be seen between their cdfs. It revokes the empirically well-documented observation that volume can influence price changes (both on the level of the order book and the level of the aggregated transaction volume), which is known in the literature as the price impact [21,23,65–68]. In order to investigate this issue, for each cryptocurrency, two parallel time series corresponding to $|R_{\Delta t}(t)|$ and $V_{\Delta t}(t)$ were input into the q-dependent detrended cross-correlation coefficient ρ_q measuring how correlated two detrended residual signals are across different scales [69]. The definition of the coefficient ρ_q, which allows one to quantify cross-correlations between two nonstationary signals, is based on the multifractal detrended cross-correlation analysis (MFCCA), whose algorithm can be sketched as follows [70].

In this particular case, there are two time series of length T and sampling intervals Δt: $|R_{\Delta t}(t_i)|$ and $V_{\Delta t}(t_i)$ with $i = 1, \ldots, T$. One starts the procedure by dividing each time series into $M_s = 2\lfloor T/s \rfloor$ non-overlapping segments of length s (called *scale*) going from both ends ($\lfloor \cdot \rfloor$ denotes the floor value). In each segment labelled by ν, both signals are integrated and polynomial trends $P^{(m)}_{\cdot,s,\nu}$ of degree m are removed:

$$\hat{R}_{\Delta t}(t_j, s, \nu) = \sum_{k=1}^{j} |R_{\Delta t}(t_{s(\nu-1)+k})| - P^{(m)}_{R,s,\nu}(t_j), \tag{3}$$

$$\hat{V}_{\Delta t}(t_j, s, \nu) = \sum_{k=1}^{j} V_{\Delta t}(t_{s(\nu-1)+k}) - P^{(m)}_{V,s,\nu}(t_j), \tag{4}$$

where $j = 1, \ldots, s$ and $\nu = 1, \ldots, M_s$. The detrended covariance is derived as

$$f^2_{|R|V}(s, \nu) = \frac{1}{s} \sum_{j=1}^{s} \left[\hat{R}_{\Delta t}(t_j, s, \nu) - \langle \hat{R}_{\Delta t}(t_j, s, \nu) \rangle_j\right] \left[\hat{V}_{\Delta t}(t_j, s, \nu) - \langle \hat{V}_{\Delta t}(t_j, s, \nu) \rangle_j\right], \tag{5}$$

where $\langle \cdot \rangle_j$ denotes the averaging over j. The detrended covariances calculated for all the segments ν are then used to determine the bivariate fluctuation function [70]:

$$F^{|R|V}_q(s) = \left\{ \frac{1}{M_s} \sum_{\nu=1}^{M_s} \mathrm{sgn}[f^2_{|R|V}(s, \nu)] |f^2_{|R|V}(s, \nu)|^{q/2} \right\}^{1/q}. \tag{6}$$

Apart from the bivariate form given by the formula above, the univariate fluctuation functions $F^{|R||R|}_q(s)$ and $F^{VV}_q(s)$ can also be calculated but, in this case, the covariance functions become variances and do not need to be factorized into the sign and modulus parts as no negative value can occur.

The above elements of the formalism allow one to introduce the q-dependent detrended cross-correlation coefficient $\rho_q(s)$ defined as [69]:

$$\rho^{|R|V}_q(s) = \frac{F^{|R|V}_q(s)}{\sqrt{F^{|R||R|}_q(s) F^{VV}_q(s)}}. \tag{7}$$

By manipulating the value of the parameter q, one can focus on the correlations between fluctuations in different size: the large fluctuations $q > 2$ or the small fluctuations $q < 1$. For $q = 2$, all the fluctuations in time series are considered with the same weights. For positive q, values of ρ_q are restricted to the interval $[-1, 1]$, with their interpretation being similar to the interpretation of the classic Pearson coefficient C: $\rho_q = 1$ means a perfect correlation, $\rho_q = 0$ means independence, and $\rho_q = -1$ means a perfect anticorrelation. For negative q, the interpretation of the coefficient is more delicate and requires some experience [69]. Figure 3 presents the coefficient $\rho_q(s)$ calculated in a broad range of scales s for the selected individual cryptocurrencies (BTC, ETH, DOGE, FUN, PERL, and WAN) and the average $\rho_q(s)$ for Groups I-III. While different data sets are characterized by different strength of the detrended cross-correlations with Group I cross-correlated the strongest and Group 3 the weakest, there is an explicit division of scales into the short-scale range ($s < 1000$ min), where the correlations monotonously increase with increasing s, and the long-scale range ($s > 1000$ min), where one observes a kind of saturation-like behavior. In the latter, the correlations are characterized by $0.75 \leq \rho_q(s) \leq 0.95$, which means that the cryptocurrency market does not differ from other financial markets and its volatility $|R_{\Delta t}|$ and volume traded are strongly correlated. The two distinguished scale ranges are related to the information-processing speed of the market: it requires some amount of time for the investors to fully react to the incoming information and to build up the cross-correlations. One might view this result as a counterpart of the Epps effect for the detrended volatility–volume data [6,28,71–73]. The main difference between this market

and the regular financial markets is the relatively long cross-over scale ($s \approx 1000$ min), which can be associated with its worse liquidity.

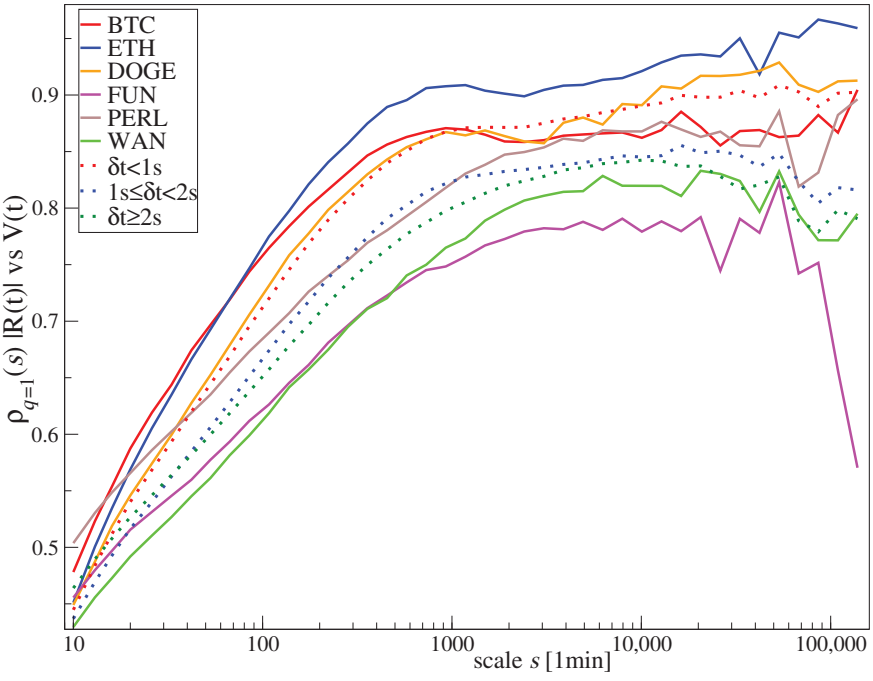


Figure 3. The q-dependent detrended cross-correlation coefficient $\rho_q(s)$ of order $q = 1$ calculated for volatility $|R_{\Delta t}(t)|$ and volume $V_{\Delta t}(t)$ (with $\Delta t = 1$ min) for the selected individual cryptocurrencies—BTC, ETH, DOGE, FUN, PERL, and WAN—where the cryptocurrency Groups I-III are characterized by a specific range of the average inter-transaction time: $\delta t < 1s$ (Group I, dotted red), $1s \leq \delta t < 2s$ (Group II, dotted blue), $\delta t \geq 2s$ (Group III, dotted green). The coefficient $\rho_q(s)$ has been averaged over all the cryptocurrencies belonging to a given group.

The next question to be asked is if there exists any functional relationship between $|R_{\Delta t}|$ and $V_{\Delta t}$. In order to address this question, $R_{\Delta t}$ vs. $V\Delta t$ scatter plots for six selected cryptocurrencies were created; see Figure 4. In general, the cross-correlations identified by means of $\rho_q(s)$ can also be confirmed visually on these plots: the larger the volume, the larger the volatility can be. However, no specific functional form of $R_{\Delta t}(V_{\Delta t})$ can be inferred from this picture. Therefore, it is instructive to change the presentation to the conditional probability plots of the form $\mathbb{E}[f(|r_{\Delta t}|)|v_{\Delta t}]$, where the expectation value $\mathbb{E}[\cdot]$ can be approximated by the mean $\langle\cdot\rangle$. From the perspective of a market with substantially limited liquidity, small price changes correspond to small transaction volumes and constitute market noise. Thus, one may expect that the most interesting relation between volatility and volume can be seen for large returns: $|r_{\Delta t}(t)| \gg 1$.

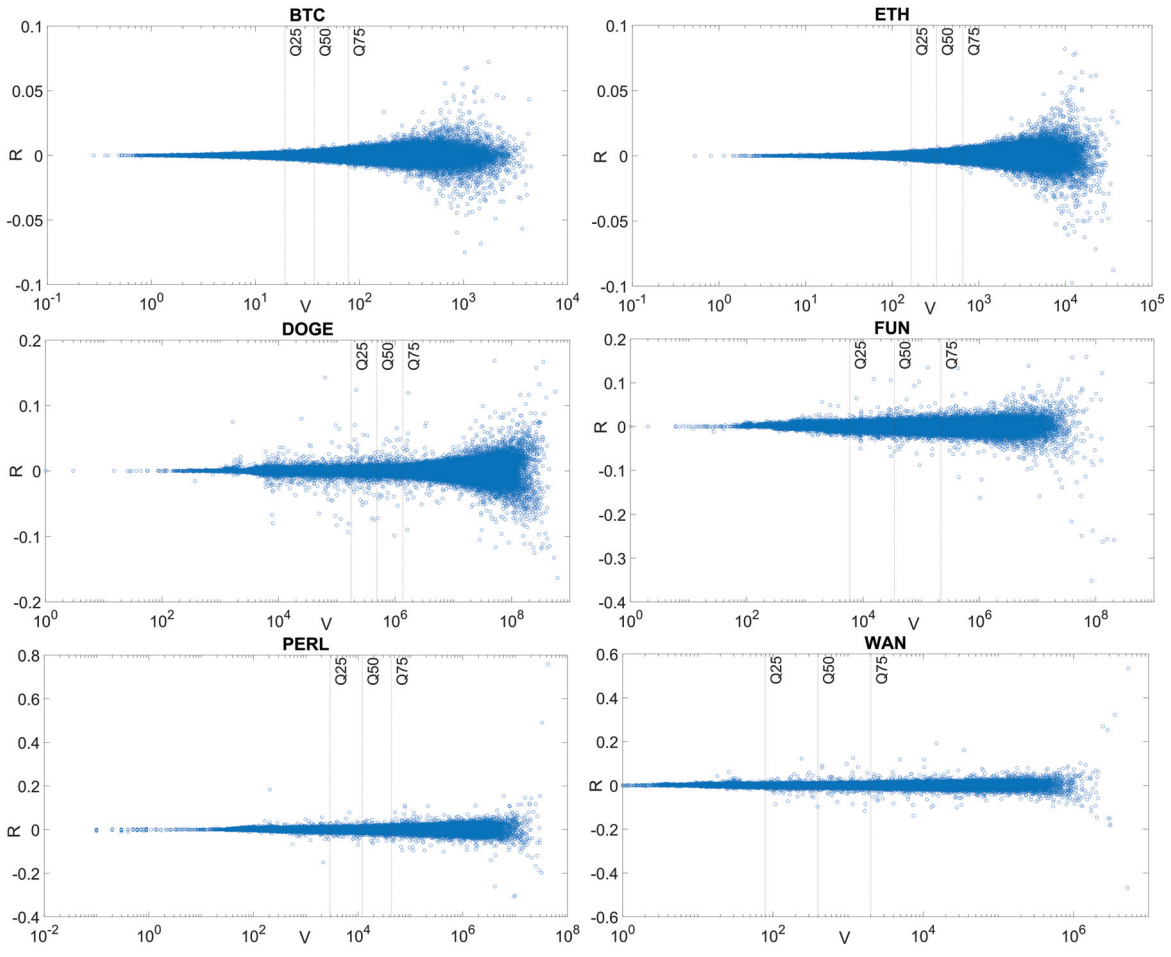


Figure 4. Scatter plots of the returns $R_{\Delta t}(t)$ and volume traded $V_{\Delta t}(t)$ for a few selected cryptocurrencies (BTC, ETH, DOGE, FUN, PERL, and WAN). Each point corresponds to a specific 1 min long interval in the whole 3-year-long period of interest. The vertical dashed lines in each panel denote the 25th, 50th, and 75th quantile of the volume probability distribution function for a given cryptocurrency. Note the logarithmic horizontal axis and the varying axis ranges among the panels.

The values of the normalized volume traded $v_{\Delta t}(t)$ were compartmentalized and, in each cell v_i, a fixed fraction $p \ll 1$ of the respective largest conditional volatility values was preserved for further analysis. A power-law function with the exponent κ is assumed to model a relation between the two quantities:

$$v_{\Delta t} \sim |r_{\Delta t}|^{\kappa}, \quad |r_{\Delta t}| \sim v_{\Delta t}^{\alpha}. \tag{8}$$

Figure 5 tests whether any of the relations of the form $\mathbb{E}[|r_{\Delta t}|^{\kappa}|v_{\Delta t}] \sim v_{\Delta t}$ hold for BTC if the following exponent values are selected: $\kappa = 0.2$, $\kappa = 0.5$, $\kappa = 1$, and $\kappa = 2$. The threshold value was chosen to be $p = 0.1$ because, for larger values, the relation becomes blurred and difficult to identify, whereas, for smaller values, too few data points can be considered, which amplifies the uncertainty. Looking at the graphs, one can reject the hypothesis that volatility and volume are related via $v_{\Delta t} \sim |r_{\Delta t}|^2$ (i.e., $\alpha = 0.5$) for all the sampling frequencies considered. In the case of the highest sampling frequency ($\Delta t = 1$ min), the data are approximated the best for $\kappa = 1$ and, secondarily, for $\kappa = 0.5$ and $\kappa = 0.2$, over the

broad volume range $1 < v_{\Delta t} < 16$. For $\Delta t \geq 10$ min, none of the values considered for κ work well, whereas, for $\Delta t = 5$ min, two cases cannot be excluded: $\kappa = 0.5$ and $\kappa = 0.2$. This means that the likely functional form of the price impact cannot be inferred based on the available data. Figure 6 presents the analogous results for ETH. The square-root form of the price impact (corresponding to $\kappa = 2$) can also be rejected in this case. However, it cannot be decided which of the remaining models ($\kappa \leq 1$) is the most likely.

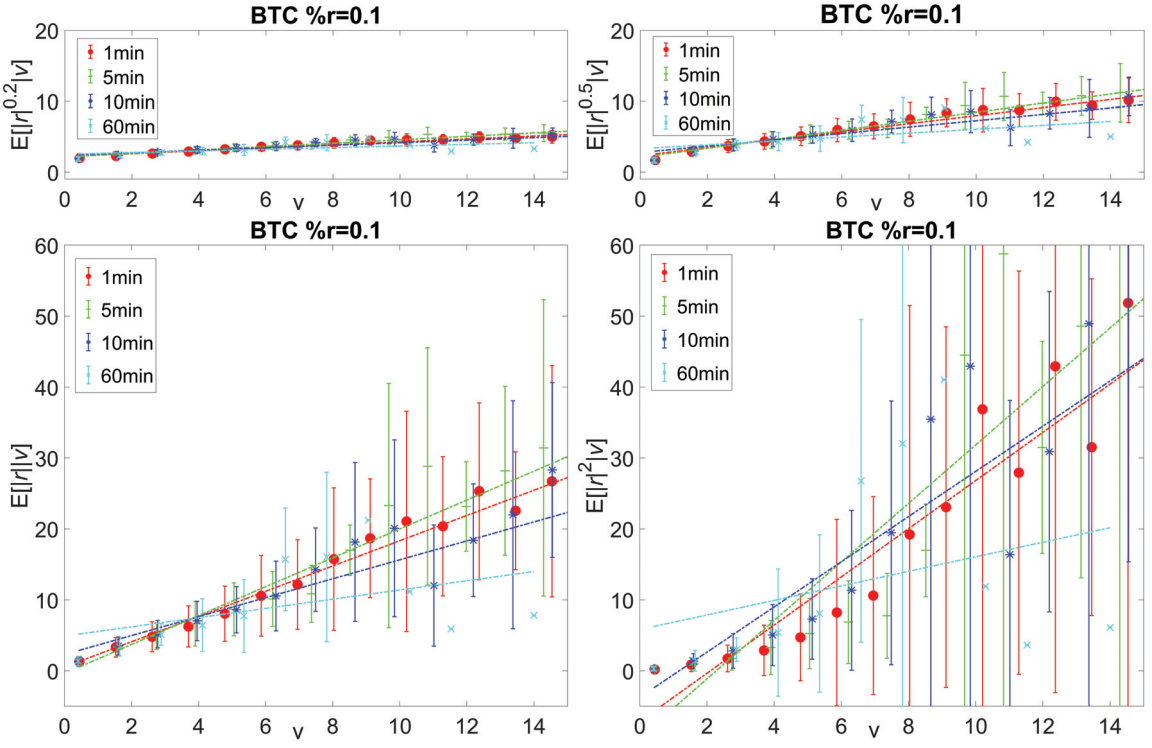


Figure 5. Conditional expectation $\mathbb{E}[|r_{\Delta t}|^{\kappa}|v_{\Delta t}]$ for BTC if only a p-fraction of the largest normalized returns $r_{\Delta t}$ is preserved for each value of the normalized volume $v_{\Delta t}$. Each panel shows the results for a specific value of κ together with a corresponding fitted power-law model. Four cases of the sampling interval are presented: $\Delta t = 1$ min, 5 min, 10 min, and 60 min. The error bars show the conditional standard deviation $\sigma[|r_{\Delta t}|^{\kappa}|v_{\Delta t}]$.

The fact that $\kappa \neq 2$ ($\alpha \neq 0.5$) and likely $\kappa \leq 1$ ($\alpha \geq 1$) for short sampling intervals is interesting because it makes the price impact function linear or superlinear ($\alpha \geq 1$): a result that differs from some earlier claims made for the regular financial markets, where the function was concave, at least for short and moderate sampling intervals [21,23]. There is also a discrepancy for the long sampling intervals because, in this case, the behavior reported for the regular markets was effectively linear, whereas here it remains undefined. It is noteworthy in this context that the superlinear ($\alpha > 1$) price impact for large Δt in Equation (8) could open the space for market manipulation [21], which, on the cryptocurrency trading platforms, can take the form of wash trading [18,74]. According to that, one can view the presented results as being in favor of the conclusion that full maturity is still ahead of the cryptocurrency market.

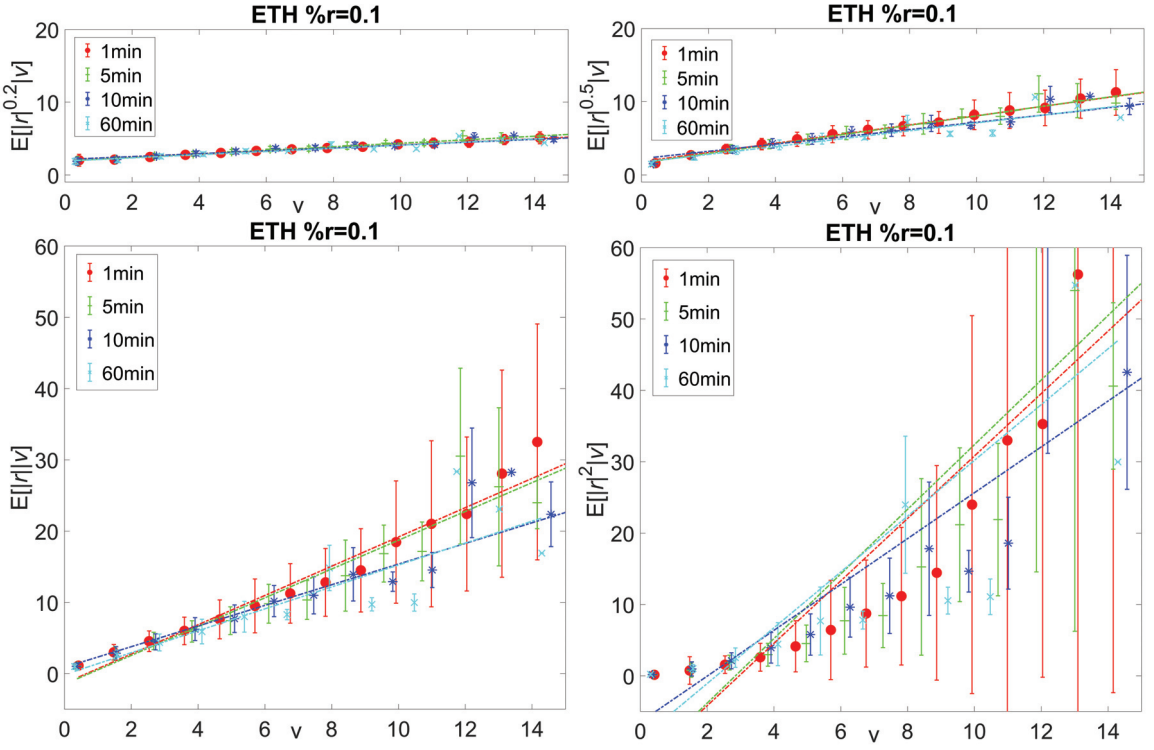


Figure 6. The same quantities as in Figure 5 for ETH.

2.4. Volatility Clustering and Long Memory

It takes some time for a market to completely absorb pieces of information that arrive there. This is a source of temporal market correlations that can be most easily observed in the price fluctuation amplitudes. Correlations are responsible for the phenomenon of volatility clustering, i.e., the existence of prolonged periods of fluctuations with elevated amplitude that are separated by quiet periods with more tamed fluctuations [75]. Volatility clustering is observed on all markets and can be quantified in terms of the autocorrelation function:

$$C(\tau) = \langle r_{\Delta t}(t) r_{\Delta t}(t-\tau) \rangle_t, \tag{9}$$

where τ is the lag time. The autocorrelation functions calculated from the absolute log-returns for several individual cryptocurrencies and the average autocorrelation functions calculated for Groups I–III are presented in Figure 7 on a double-logarithmic scale. In each case, one can identify at least one range of lags over which $C(\tau)$ shows power-law decay. For BTC, ETH, and FUN, there is only one such range corresponding to 10 min $\leq \tau \leq$ 500 min with a relatively small upper end. The same refers to WAN but, in this case, the upper end exceeds $\tau \approx$ 20,000 min (ca. two weeks). On the other hand, DOGE, PERL, and the average plots show two scaling regimes: the short-τ regime up to $\tau \approx$ 500–1000 min (less than a day) and the long-τ regime for 1000 min $< \tau <$ 20,000 min. In each case, $C(\tau)$ falls to 0 around $\tau \approx$ 100,000 min (more than 2 months). Compared to a more distant past, the scaling regions for BTC and ETH are shorter now (e.g., in the years 2016–2018, it reached $\tau =$ 1000 min [29]), which is consistent with the market time acceleration caused by an increased trading frequency. This overall picture for the cryptocurrency market does not depart much from the one observed in other financial markets. A power-law decaying autocorrelation function expressing the long memory of volatility is a common property that is a manifestation of the way that the market processes

information [27,76]. The time lag at which $C(\tau)$ reaches a statistically insignificant level is equal to the average length of a volatility cluster [76]. Due to the alternating character of market dynamics, where the high-volatility periods are interwoven with low-volatility periods, for larger time lags, the autocorrelation function becomes negative. Note that, due to the fact that volatility time series are unsigned, the long-range autocorrelations cannot be exploited for the related investment strategies.

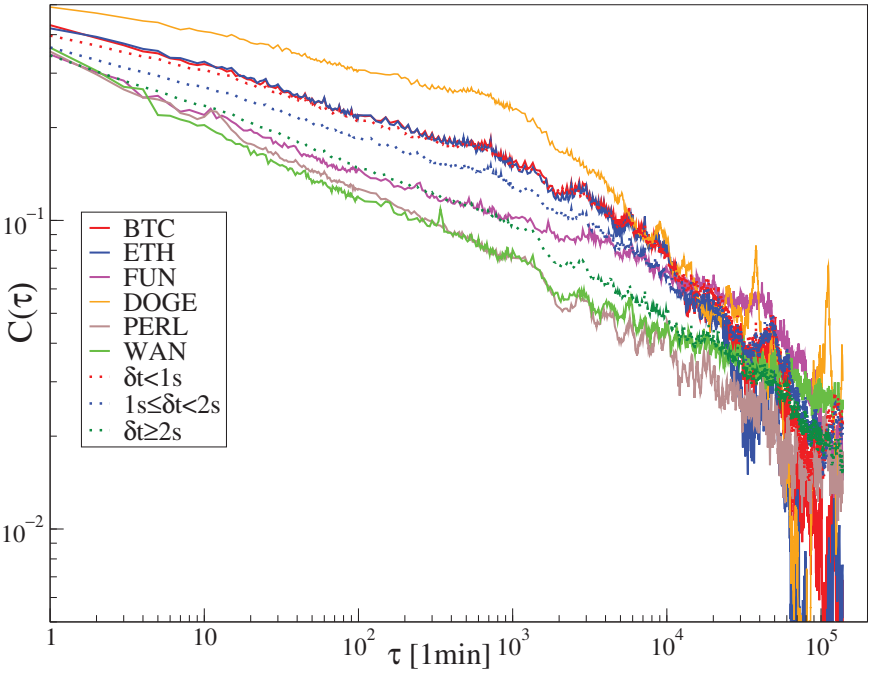


Figure 7. Autocorrelation function $C_{|r_{\Delta t}|}(\tau)$ of the absolute normalized log-returns $|r_{\Delta t}(t)|$ (volatility) calculated for the selected individual cryptocurrencies—BTC, ETH, DOGE, FUN, PERL, and WAN—as well as for the cryptocurrency Groups I-III characterized by specific range of the average inter-transaction time: $\delta t < 1s$ (Group I, dotted red), $1s \leq \delta t < 2s$ (Group II, dotted blue), $\delta t \geq 2s$ (Group III, dotted green). $C_{|r_{\Delta t}|}(\tau)$ has been averaged for each value of τ over all the cryptocurrencies belonging to a given group. Note the double-logarithmic scale.

2.5. Multiscaling of Returns

If the bivariate or univariate fluctuation functions can be approximated by a power-law relation

$$F_q^{AB}(s) \sim s^{h(q)}, \tag{10}$$

where $h(q)$ is a non-increasing function of q called the generalized Hurst exponent [77] and A and B stand for either R or V, the time series under study reveal a fractal structure. If $h(q) = \text{const} = H$, it means that this structure is monofractal, with H equal to the Hurst exponent, which is a measure of persistence; otherwise, it is multifractal [77]. Multifractal signals are governed by processes with long-range autocorrelations, which is why both properties are often observed together [78–81]. It is the case, for example, in financial data. If the relation (10) exists, it can be seen in double-logarithmic plots of $F_q^{AB}(s)$. Figure 8 displays $F_q^{RR}(s)$ for six cryptocurrencies, with $-4 \leq q \leq 4$ and $10 \leq s \leq 25{,}000$. Out of these, four cryptocurrencies show unquestionable power-law functions—BTC, ETH, DOGE, and FUN—for all used values of q and for at least a decade-long range of scales,

whereas PERL and WAN do not. The same result can be expressed in a different way by calculating the singularity spectra $f(\alpha)$ from $h(q)$ according to the following relations:

$$\alpha = h(q) + q\frac{dh(q)}{dq}, \quad f(\alpha = q(\alpha - h(q)) + 1. \tag{11}$$

The Hölder exponents α quantify the singularity strength and $f(\alpha_0)$ expresses the fractal dimension of a subset of singularities with strength $\alpha = \alpha_0$. While many theoretical singularity spectra are symmetric, in a practical situation, one observes spectra that are asymmetric [14,28,31,82–85]. The insets in Figure 8 show $f(\alpha)$ calculated from $F_q^{RR}(s)$ in the scaling regions of s. All the presented spectra are left-side asymmetric (their left shoulder, corresponding to positive qs, is longer). The origin of such a behavior can be twofold: the signals can develop heavy tails of the probability distribution functions that are unstable in the sense of Lévy yet their convergence to the normal distribution is slow [76], and the signals can be mixtures of processes that have different fractal properties: large fluctuations can be associated with a multifractal process (e.g., a multiplicative cascade), whereas small fluctuations can be monofractal. It often happens that the small fluctuations in financial time series are noise whereas the medium and large fluctuations carry meaningful information.

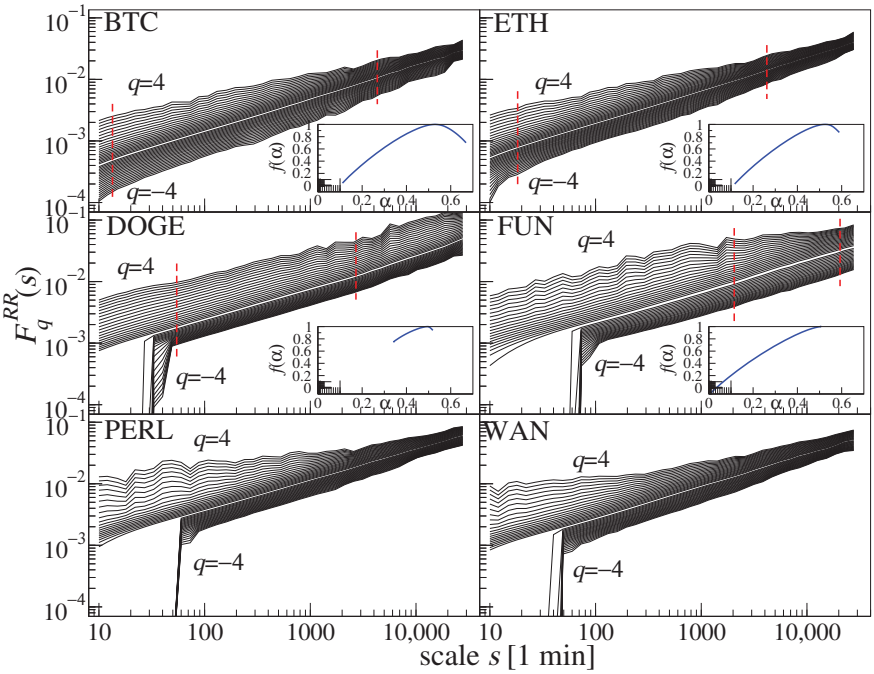


Figure 8. (Main plots) Univariate fluctuation functions $F_q^{RR}(s)$ calculated from the log-returns $R_{\Delta t}(t)$ with $\Delta t = 1$ min for BTC, ETH, DOGE, FUN, PERL, and WAN. The breakdown of scaling for small scales and negative values of q in some plots is an artifact related to long sequences of zero returns in time series. (Insets) Singularity spectra $f(\alpha)$ calculated from the corresponding fluctuation functions in the range denoted by dashed red lines (if possible).

It was reported in the literature that cryptocurrencies can also show such asymmetric $f(\alpha)$ spectra [6,14]. From the perspective of this study, it is interesting to note that the spectra for BTC calculated for different historical periods show an elongation of the right shoulder of $f(\alpha)$ that corresponds to small fluctuations. It can be interpreted as a gradual building of a multifractal structure in BTC price fluctuations that started from large returns only in the early stages of BTC trading and were imposed on the smaller returns as the

cryptocurrency market goes toward maturity. If one looks at Figure 8, BTC, ETH, and, to a lesser degree, DOGE—that is, the cryptocurrencies that are among the most capitalized ones—have noticeable right wings of $f(\alpha)$, whereas the more exotic cryptocurrencies, such as FUN, PERL, and WAN, do not develop the right wing at all. In agreement with what has been said before, despite various cryptoassets being traded on the same platforms, different ones can be found at different stages of the maturing process due to the different trading frequencies. This difference can also be observed in the possible scaling range of the fluctuation functions in Figure 8. In the case of the two most liquid cryptocurrencies, BTC and ETH, the $F_q^{RR}(s)$ scaling can be observed almost from the beginning of the scale range, whereas, in the case of less liquid cryptocurrencies, the range of satisfactory scaling is significantly shorter and $F_q^{RR}(s)$ even becomes singular on short scales due to the number of consecutive 1 min bins with zero returns. This is typical behavior in the case of less liquid financial instruments [14].

2.6. Cross-Correlations among Cryptocurrencies

Information that flows into the market may have the same impact on certain assets that, for example, share similar characteristics, such as the market sector, the main shareholders, or, in the case of cryptocurrency, the type or consensus mechanism [86]. This can lead to the emergence of cross-correlation between such assets and to a certain hierarchy of cross-correlations (e.g., sector, subsector, and bilateral ones) [87]. The correlation structure is a dynamical property that can change suddenly and substantially over time as the market reacts to perturbations [88]. In quiet times, it is well-shaped, elastic, and hierarchical, whereas, during periods of turmoil, it becomes centralized and rigid. This dual behavior is characteristic for the developed markets, while a lack of cross-correlations or a persistent centralization may be attributed to immaturity.

The market cross-correlation structure can be concisely characterized by the matrix approach. For a set of N time series of log-returns representing different cryptocurrencies $N(N-1)/2$, the q-dependent detrended cross-correlation coefficients $\rho_q^{ij}(s)$ can be calculated, where $i, j = 1, \ldots, N$ and $\rho_q^{ij} = \rho_q^{ji}$, which form a q-dependent detrended cross-correlation matrix $\mathbf{C}_q(s)$. Due to the fact that the cross-correlation strength increases typically with scale for all the asset pairs, the differences in $\rho_q^{ij}(s)$ are, on average, minimal for the shortest studied scale of $s = 10$ min. However, even in this case, it is easy to observe that different cryptocurrency pairs reveal strong differences. Figure 9 presents the complete matrix $\mathbf{C}_q(s)$ with the cryptocurrencies ordered according to the average inter-transaction time $\langle \delta t \rangle_t$. The ordering allows one to find even by eye a significant cross-correlation between $\langle \delta t \rangle_t$ and ρ_q^{ij}: the shorter this time is, the stronger the cross-correlations are. In full analogy to other markets, information needs time to propagate over the whole cryptocurrency market and the propagation speed is crucially dependent on the cryptocurrency liquidity, which can roughly be approximated by the transaction number per time unit. Based on the exact values of $\rho_q^{ij}(s)$, one can notice that even the least frequently traded cryptocurrencies from the considered basket develop statistically significant dependencies among themselves. This, however, might not be true for even less capitalized tokens, which can have idiosyncratic dynamics.

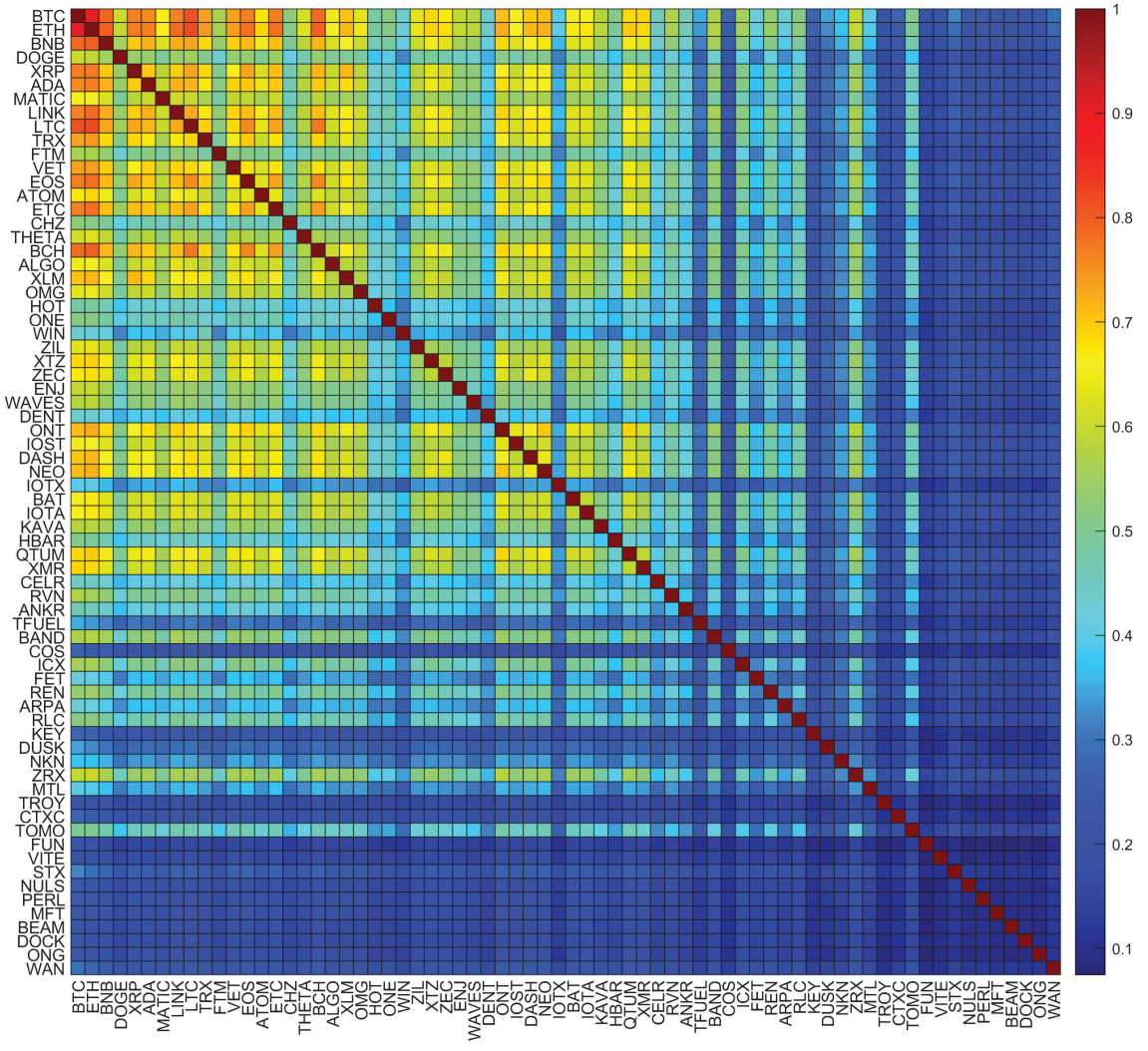


Figure 9. The q-dependent detrended cross-correlation matrix entries $\rho_q^{ij}(s)$ calculated from time series of log-returns representing 70 cryptocurrencies with $q = 1$ and $s = 10$ min. Cryptocurrencies have been sorted according to the average inter-transaction time δt in increasing order (top to bottom). The color-coding scheme is shown on the right.

The correlation matrix $\mathbf{C}_q(s)$ can be transformed into a distance matrix $\mathbf{D}_q(s)$ with the entries

$$d_q^{ij}(s) = \sqrt{2(1 - \rho_q^{ij}(s))}, \tag{12}$$

which differs from the former in that its entries $d_q^{(ij)}$ are metric. $\mathbf{D}_q(s)$ can be used for constructing a weighted graph with nodes representing cryptocurrencies and edges representing distances $d_q^{(ij)}(s)$. Next, by using the Prim algorithm [89], one can construct the corresponding q-dependent detrended minimal spanning tree (MST), which can be considered as a connected minimum-weight subset of the graph containing all N nodes and $N - 1$ edges. The MST topology depends strongly on the cross-correlation structure of a market. A centralized market corresponds to a star-like MST, whereas a market contain-

ing idiosyncratic assets shows an MST with elongated branches and no dominant hubs. Figure 10 exhibits two MSTs created from all 70 cryptocurrencies for $q = 1$ (left) and $q = 4$ (right). The former involves cross-correlations between the fluctuations in all magnitudes, whereas the latter involves only the large fluctuations. For $q = 1$, the structure is dual-star with BTC and ETH as its central hubs. This is not surprising as both cryptocurrencies are distinguished by their fame and large capitalization, which makes them a kind of reference for the remaining cryptocurrencies. On the other hand, for $q = 4$, the structure is more distributed, with a primary hub, BTC, and a few secondary ones: LTC, XMR, and ONT. This means that the relatively large fluctuations are not collectively correlated, unlike the majority of fluctuations, and more subtle dependencies are present. This is in parallel with the conclusions based on the multifractal analysis, which were large fluctuations that carried clearly multifractal characteristics and long-term correlations, whereas the small fluctuations were much more noisy. It is worth mentioning that a similar behavior can be observed in the stock market, where the cross-correlation structure carried by the large fluctuations is much richer than that carried by the medium and small fluctuations [90]. However, in the stock market, the heterogeneous cross-correlation structure is more pronounced even in the latter case [86,90]. Since there is no clear division into market sectors [91], the cryptocurrency market appears to be less developed from this particular perspective.

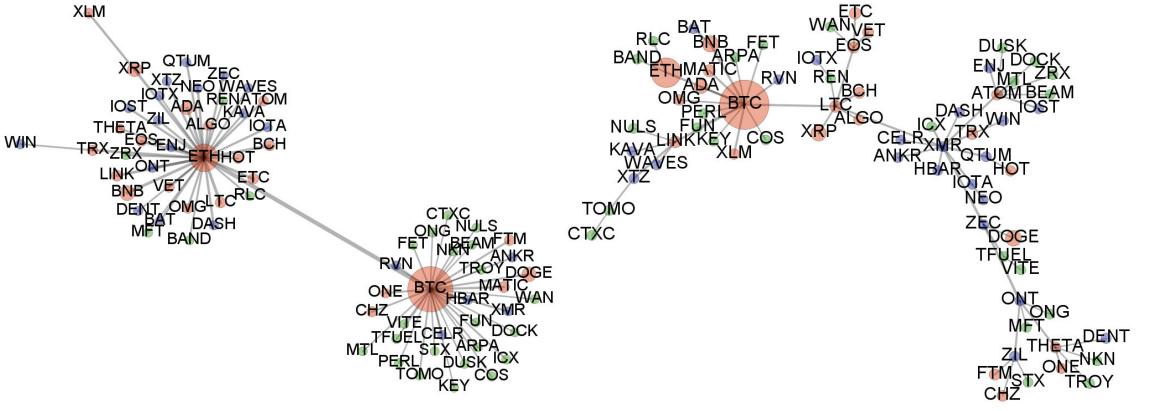


Figure 10. Minimal spanning trees calculated from a distance matrix $\mathbf{D}_q(s)$ based on $\rho_q(s)$ for $s = 10$ and for $q = 1$ (**left**) and $q = 4$ (**right**). Within each tree, the size of the vertex is proportional to the average value of the volume $W_{\Delta t}$ for $\Delta t = 1$ min, while the width of the edge is proportional to $1 - d_q^{ij}(s)$. The vertex sizes cannot be directly compared across the trees, however. Colors represent Groups I-III in terms of the trading frequency: $\delta t < 1s$ (Group I, red), $1s \leq \delta t < 2s$ (Group II, blue), and $\delta t \geq 2s$ (Group III, green).

2.7. Cross-Correlations between Cryptocurrencies and Other Markets

Recently, BTC and ETH have been found to be significantly coupled to the traditional financial markets during the period covering the COVID-19 pandemic and the bear market of 2022 [55]. This result has essential practical implications in risk management as cryptocurrencies cannot serve as hedging assets [92]. It differs from earlier findings that, before 2020, the cryptocurrency market was detached from the traditional markets [47,52,93,94], but, at the same time, it remains in agreement with the observations that COVID-19 changed the safe-haven paradigm and contributed to the correlation of major cryptocurrencies with traditional risk assets [53,95–98]. So far, only the most capitalized cryptocurrencies have been studied [55], and this is why cryptocurrencies with smaller capitalization were also studied here.

The time series of log-returns of 70 cryptocurrencies and 22 traditional financial instruments were collected from Dukascopy platform [99]. Among the latter, there are

contracts for difference (CFDs) representing the returns of 12 fiat currencies (AUD, CAD, CHF, CNH, EUR, GBP, JPY, MXN, NOK, NZD, PLN, and ZAR), 4 commodities (WTI crude oil (CL), high-grade copper (HG), silver (XAG), and gold (XAU)), 4 US stock market indices (Nasdaq 100 (NQ100), S&P500, Down Jones Industrial Average (DJI), and Russell 2000 (RUSSEL)), the main German stock index DAX 40 (DAX), and the Japanese Nikkei 225 (NIKKEI). All these instruments except for the non-US stock indices were expressed in USD. Their quotes cover a period from 1 January 2020 to 30 December 2022. The quotes were recorded over the trading hours, i.e., from Sunday 22:00 to Friday 20:15 UTC, with a break between 20:15 and 22:00 UTC each trading day. In order to assess the cross-correlations, the cryptocurrency time series were synchronized with those from Dukascopy. Cross-correlations were quantified by $\rho_q^{RR}(s)$.

Figure 11 shows the q-dependent detrended cross-correlation matrix $\mathbf{C}_q(s)$ entries for the inter-market pairs consisting of a cryptocurrency and a traditional asset. The first observation is that the maximum available values of the matrix entries do not exceed $\rho_q^{RR}(s) = 0.25$, which makes them much smaller than in the case of the inner cross-correlation among the cryptocurrencies. This is an expected effect because markets are typically more tightly coupled inside than outside. Among the strongest cross-correlations, one can point out the coupling of BTC and ETH with the American stock market indices ($\rho_q^{RR}(s) > 0.2$ and with NIKKEI and DAX ($0.15 < \rho_q^{RR}(s) < 0.2$). Considerably weaker yet still prominent are the cross-correlations between several other cryptocurrencies, such as XRP, ADA, LTC, LINK, VET, ETC, EOS, ATOM, and BCH on one side and the American indices ($0.15 < \rho_q^{RR}(s) < 0.2$). The relations between cryptocurrencies and fiat currencies remain moderate, with the AUD, CAD, and NZD being the relatively strongest ($0.1 < \rho_q^{RR}(s) < 0.15$). Contrary to this, the cryptocurrencies are the most decoupled from JPY, CHF, gold (XAU), and crude oil (CL). A general observation is that the less liquid a cryptocurrency is, the weaker its cross-correlation with traditional instruments. Here again, DOGE is somewhat of an exception and has a weaker cross-correlation than its trading frequency and capitalization would imply. However, it should be noted that the values collected in Figure 11 correspond to the shortest available scale of $s = 10$ min. How these values refer to the maximum cross-correlations for longer scales is documented in Figure 12. Here, the cross-correlation between the selected cryptocurrencies and their sets grouped according to the average inter-transaction time (Groups I–III) and NASDAQ 100 is presented. This particular choice of the traditional index was motivated by the fact that the cryptocurrency market is strongly cross-correlated with it [55]. Indeed, for much longer s, the values of $\rho_q^{RR}(s)$ grow significantly and even reach some saturation level resembling the Epps effect for $s > 500$ min, with the average values of $\rho_q^{RR}(s)$ in Groups I-III oscillating around 0.4 (for a given scale, $\rho_q^{RR}(s)$ decreases systematically with an increasing δt). The cryptocurrencies that are the most cross-correlated with NASDAQ 100, i.e., BTC and ETH, have maximum values of $\rho_q^{RR}(s) > 0.5$.

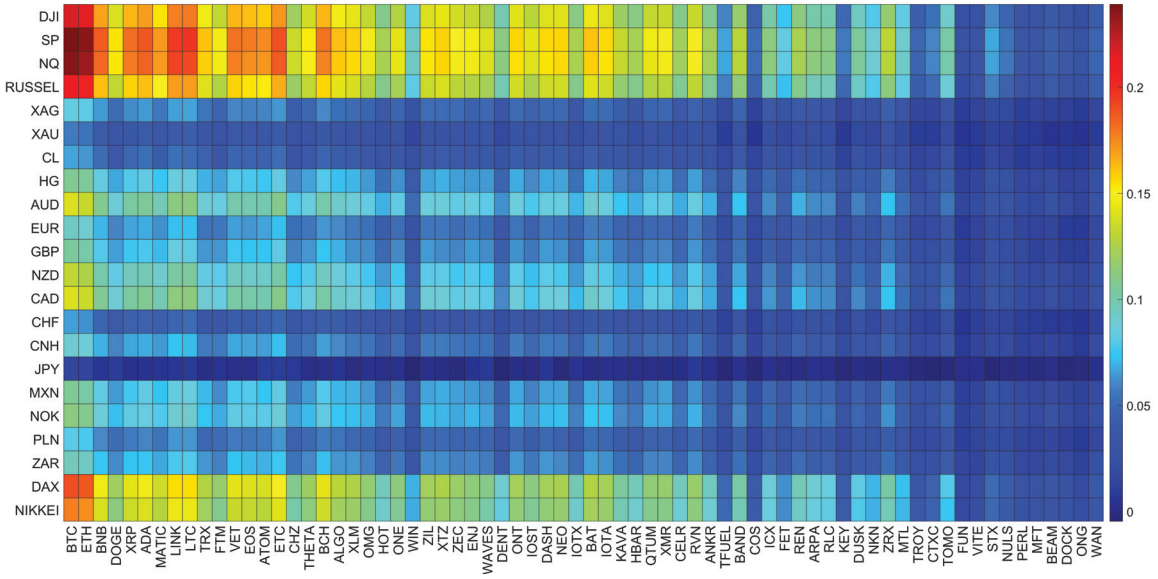


Figure 11. The q-dependent detrended cross-correlation matrix entries $\rho_q^{ij}(s)$ calculated from time series of log-returns representing selected cryptocurrencies and selected traditional financial instruments with $q = 1$ and $s = 10$ min. Cryptocurrencies have been sorted according to the average inter-transaction time δt in increasing order (top to bottom). The color coding scheme, which differs from the one in Figure 9, is shown on the right.

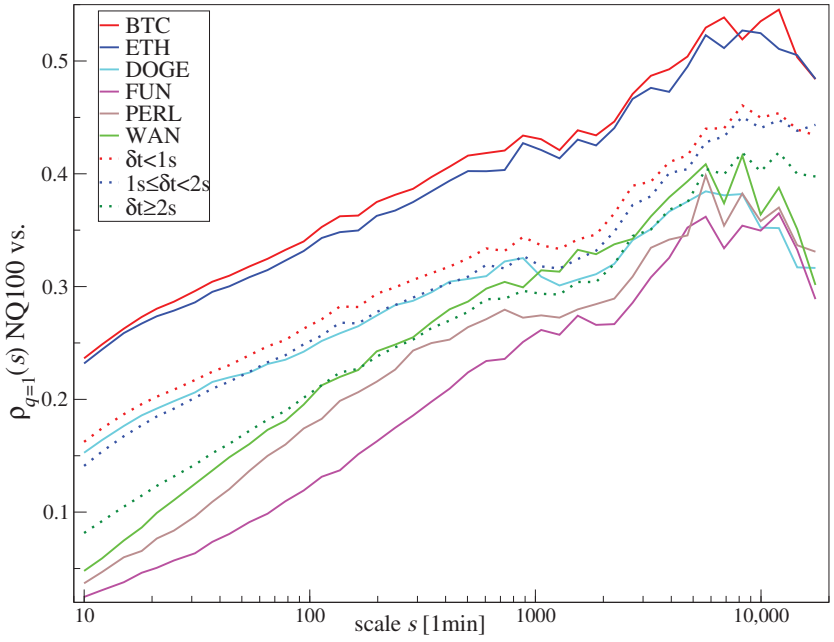


Figure 12. The q-dependent detrended cross-correlation coefficient $\rho_q^{RR}(s)$ calculated for the pairs of log-return time series consisting of NASDAQ 100 and a cryptocurrency (BTC, ETH, DOGE, FUN, PERL, or WAN) or a group of cryptocurrencies characterized by average inter-transaction time from a specific range: $\delta t < 1s$ (Group I, red), $1s \leq \delta t < 2s$ (Group II, blue), and $\delta t \geq 2s$ (Group III, green).

3. Conclusions

The statistical properties of price log-returns and the volume of the cryptocurrencies were the central points of the present study. The existence of the so-called financial stylized facts in the cryptocurrency market during the last 3 years was investigated and compared with the stylized facts observed in the traditional financial markets. Several characteristics were of particular interest: a tail behavior of the probability distribution functions for the log-returns and volume traded, the functional form of price impact, volatility autocorrelations, multiscaling, cross-correlations among the cryptocurrencies, and cross-correlations between the principal cryptocurrencies and selected traditional market assets. Almost all the analyzed characteristics of the cryptocurrency market were found to be in qualitative agreement with their counterparts from the traditional markets. It allows one to conclude that, from this particular perspective, the cryptocurrency market does not differ from the mature markets.

Despite such a positive conclusion, one still has to be cautious. First, the level of the maturity of the cryptocurrencies depends on their trading frequency. The most liquid ones, such as BTC and ETH, to a greater extent, have characteristics corresponding to mature financial markets, and the least liquid ones do not. Second, the price impact function, while also of a power-law form, results in being substantially different from its counterparts reported in the traditional markets (linear or convex here vs. concave there [21]). Third, while the statistical properties are important from a practical point of view as they can be exploited in various investment strategies, there are nevertheless many other important indicators of market maturity that were not investigated here. For example, the number of cryptocurrencies traded on the largest platforms, such as Binance, is so large that it already matches the world's largest markets, such as the New York Stock Exchange and NASDAQ. On the other hand, even the most recognized cryptocurrencies, such as BTC and ETH, show extreme volatility, which means that the market is still rather illiquid, and this property can question its maturity. There is another problem associated with the fact that the cryptocurrencies are often viewed as speculation toys rather than full-scale investment instruments. There are also numerous issues related to the limited reliability of the cryptocurrencies, their weak supply elasticity, etc. These problems, while important, were beyond the scope of this analysis, which one has to keep in mind when thinking about the given conclusions. Repeating this kind of analysis in future in order to follow how the cryptocurrency market changes seems to be a straightforward direction of potential future studies.

Author Contributions: Individual contributions of the authors are as follows. Conceptualization, S.D. and M.W.; methodology, S.D., J.K. and M.W.; software, M.W.; validation, S.D., J.K. and M.W.; formal analysis, S.D. and M.W.; investigation, S.D., J.K. and M.W.; resources, M.W.; data curation, M.W.; writing—original draft preparation, J.K.; writing—review and editing, S.D., J.K. and M.W.; visualization, M.W.; supervision, S.D. and J.K.; project administration, S.D. and M.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used in the article are freely available on Binance and Dukascopy platforms.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. List of cryptocurrencies from Binance.

Ticker	Name	Ticker	Name	Ticker	Name
ADA	cardano	FET	fetch	QTUM	qtum
ALGO	algorand	FTM	fantom	REN	ren
ANKR	ankr	FUN	funtoken	RLC	iexec
ARPA	arpa chain	HBAR	hedera	RVN	ravencoin
ATOM	cosmos	HOT	holo	STX	stacks
BAND	band protocol	ICX	icon	TFUEL	theta fuel
BAT	basic atention token	IOST	iost	THETA	theta
BCH	bitcoin cash	IOTA	miota	TOMO	tomochain
BEAM	beam	IOTX	iotex	TROY	troy
BNB	binance coin	KAVA	kava	TRX	tron
BTC	bitcoin	KEY	key	VET	vechain
CELR	celer network	LINK	chainlink	VITE	vite
CHZ	chiliz	LTC	litecoin	WAN	wanchain
COS	contentos	MATIC	polygon	WAVES	waves
CTXC	cortex	MFT	hifi finance	WIN	winklink
DASH	dash	MTL	metal	XLM	stellar
DENT	dent	NEO	neo	XMR	monero
DOCK	dock	NKN	nkn	XRP	ripple
DOGE	dogecoin	NULS	nuls	XTZ	tezos
DUSK	dusk network	OMG	omg network	ZEC	zcash
ENJ	enj coin	ONE	harmony	ZIL	zilliqa
EOS	eos	ONG	ontology gas	ZRX	0x
ETC	ethereum classic	ONT	ontology		
ETH	ethereum	PERL	perl		

References

1. Wattenhofer, R. *The Science of the Blockchain*; CreateSpace Independent Publishing Platform: Scotts Valley, CA, USA, 2016.
2. Lantz, L.; Cawrey, D. *Mastering Blockchain*; O'Reilly Media: Sebastopol, CA, USA 2020.
3. Corbet, S.; Lucey, B.; Urquhart, A.; Yarovaya, L. Cryptocurrencies as a financial asset: A systematic analysis. *Int. Rev. Financ. Anal.* **2019**, *62*, 182–199. [CrossRef]
4. Gil-Cordero, E.; Cabrera-Sánchez, J.P.; Arrás-Cortés, M.J. Cryptocurrencies as a financial tool: Acceptance factors. *Mathematics* **2020**, *8*, 1974. [CrossRef]
5. Cachanosky, N. Can cryptocurrencies become a commonly accepted means of exchange? In *The Economics of Blockchain and CryptocurrencyAIER Sound Money Project*; Working Paper No. 2020-14; Edward Elgar Publishing: Cheltenham, UK, 2020.
6. Wątorek, M.; Drożdż, S.; Kwapień, J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* **2021**, *901*, 1–82. [CrossRef]
7. James, N.; Menzies, M. Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time. *Nonlinear Dyn.* **2022**, *107*, 4001–4017. [CrossRef] [PubMed]
8. Ausloos, M. Statistical physics in foreign exchange currency and stock markets. *Physica A* **2000**, *285*, 48–65. [CrossRef]
9. Cont, R. Empirical properties of asset returns: Stylized facts and statistical issues. *Quant. Financ.* **2001**, *1*, 223–236. [CrossRef]
10. LeBaron, B. Agent-based financial markets: Matching stylized facts with style. In *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*; Cambridge University Press: Cambridge, UK, 2006; pp. 221–236.
11. Morone, A. Financial markets in the laboratory: An experimental analysis of some stylized facts. *Quant. Financ.* **2008**, *8*, 513–532. [CrossRef]

12. Wątorek, M.; Kwapień, J.; Drożdż, S. Financial return distributions: Past, present, and COVID-19. *Entropy* **2021**, *23*, 884. [CrossRef]
13. Begušić, S.; Kostanjčar, Z.; Eugene Stanley, H.; Podobnik, B. Scaling properties of extreme price fluctuations in Bitcoin markets. *Physica A* **2018**, *510*, 400–406. [CrossRef]
14. 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* **2018**, *28*, 071101. [CrossRef]
15. Pessa, A.A.; Perc, M.; Ribeiro, H.V. Age and market capitalization drive large price variations of cryptocurrencies. *Sci. Rep.* **2023**, *13*, 3351. [CrossRef] [PubMed]
16. Gopikrishnan, P.; Plerou, V.; Gabaix, X.; Stanley, H.E. Statistical properties of share volume traded in financial markets. *Phys. Rev. E* **2000**, *62*, R4493. [CrossRef] [PubMed]
17. Plerou, V.; Stanley, H.E.; Gabaix, X.; Gopikrishnan, P. On the origin of power-law fluctuations in stock prices. *Quant. Financ.* **2004**, *4*, 11–15. [CrossRef]
18. Kwapień, J.; Wątorek, M.; Bezbradica, M.; Crane, M.; Tan Mai, T.; Drożdż, S. Analysis of inter-transaction time fluctuations in the cryptocurrency market. *Chaos* **2022**, *32*, 083142. [CrossRef]
19. Navarro, R.M.; Leyvraz, F.; Larralde, H. Statistical properties of volume in the Bitcoin/USD market. *arXiv Prepr.* **2023**, arXiv:2304.01907.
20. Gillemot, L.; Farmer, J.D.; Lillo, F. There's more to volatility than volume. *Quant. Financ.* **2006**, *6*, 371–384. [CrossRef]
21. Bouchaud, J.P. Price impact. In *Encyclopedia of Quantitative Finance*; Cambridge University Press: Cambridge, UK, 2010; pp. 1–6.
22. Tóth, B.; Lempérière, Y.; Deremble, C.; Lataillade, J.D.; Kockelkoren, J.; Bouchaud, J.P. Anomalous price impact and the critical nature of liquidity in financial markets. *Phys. Rev. X* **2011**, *1*, 021006.
23. Gabaix, X.; Gopikrishnan, P.; Plerou, V.; Stanley, H.E. A theory of power-law distributions in financial market fluctuations. *Nature* **2003**, *423*, 267–270. [CrossRef]
24. Rak, R.; Drożdż, S.; Kwapień, J.; Oświęcimka, P. Stock returns versus trading volume: Is the correspondence more general? *Acta Phys. Pol. B* **2013**, *44*, 2035–2050. [CrossRef]
25. Bucci, F.; Benzaquen, M.; Lillo, F.; Bouchaud, J.P. Crossover from linear to square-root market impact. *Phys. Rev. Lett.* **2019**, *122*, 108302. [CrossRef]
26. Zarinelli, E.; Treccani, M.; Farmer, J.D.; Lillo, F. Beyond the square root: Evidence for logarithmic dependence of market impact on size and participation rate. *Mark. Microstruct. Liq.* **2015**, *1*, 1550004. [CrossRef]
27. Gopikrishnan, P.; Plerou, V.; Nunes Amaral, L.A.; Meyer, M.; Stanley, H.E. Scaling of the distribution of fluctuations of financial market indices. *Phys. Rev. E* **1999**, *60*, 5305–5316. [CrossRef] [PubMed]
28. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Rak, R. The foreign exchange market: Return distributions, multifractality, anomalous multifractality and the Epps effect. *New J. Phys.* **2010**, *12*, 105003. [CrossRef]
29. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Signatures of the crypto-currency market decoupling from the Forex. *Future Internet* **2019**, *11*, 154. [CrossRef]
30. Matia, K.; Ashkenazy, Y.; Stanley, H.E. Multifractal properties of price fluctuations of stocks and commodities. *Europhys. Lett.* **2003**, *61*, 422–428. [CrossRef]
31. Kwapień, J.; Oświęcimka, P.; Drożdż, S. Components of multifractality in high-frequency stock returns. *Physica A* **2005**, *350*, 466–474. [CrossRef]
32. Oświęcimka, P.; Kwapień, J.; Drożdż, S. Multifractality in the stock market: Price increments versus waiting times. *Physica A* **2005**, *347*, 626–638. [CrossRef]
33. Takaishi, T. Statistical properties and multifractality of Bitcoin. *Physica A* **2018**, *506*, 507–519. [CrossRef]
34. Kristjanpoller, W.; Bouri, E. Asymmetric multifractal cross-correlations between the main world currencies and the main cryptocurrencies. *Physica A* **2019**, *523*, 1057–1071. [CrossRef]
35. Han, Q.; Wu, J.; Zheng, Z. Long-range dependence, multi-fractality and volume-return causality of ether market. *Chaos: Interdiscip. J. Nonlinear Sci.* **2020**, *30*, 011101. [CrossRef]
36. Takaishi, T.; Adachi, T. Market efficiency, liquidity, and multifractality of Bitcoin: A dynamic study. *Asia-Pac. Financ. Mark.* **2020**, *27*, 145–154. [CrossRef]
37. Bariviera, A.F. One model is not enough: Heterogeneity in cryptocurrencies' multifractal profiles. *Financ. Res. Lett.* **2021**, *39*, 101649. [CrossRef]
38. Takaishi, T. Time-varying properties of asymmetric volatility and multifractality in Bitcoin. *PLoS ONE* **2021**, *16*, e0246209. [CrossRef] [PubMed]
39. Kakinaka, S.; Umeno, K. Asymmetric volatility dynamics in cryptocurrency markets on multi-time scales. *Res. Int. Bus. Financ.* **2022**, *62*, 101754. [CrossRef]
40. Wątorek, M.; Kwapień, J.; Drożdż, S. Multifractal cross-correlations of bitcoin and ether trading characteristics in the post-COVID-19 time. *Future Internet* **2022**, *14*, 215. [CrossRef]
41. Drożdż, S.; Gruemmer, F.; Ruf, F.; Speth, J. Towards identifying the world stock market cross-correlations: DAX versus Dow Jones. *Physica A* **2001**, *294*, 226–234. [CrossRef]
42. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L.A.N.; Guhr, T.; Stanley, H.E. Random matrix approach to cross correlations in financial data. *Phys. Rev. E* **2002**, *65*, 066126. [CrossRef]

43. Maslov, S. Measures of globalization based on cross-correlations of world financial indices. *Physica A* **2001**, *301*, 397–406. [CrossRef]
44. Nguyen, A.P.N.; Mai, T.T.; Bezbradica, M.; Crane, M. The Cryptocurrency Market in Transition before and after COVID-19: An Opportunity for Investors? *Entropy* **2022**, *24*, 1317. [CrossRef]
45. James, N.; Menzies, M.; Chin, K. Economic state classification and portfolio optimisation with application to stagflationary environments. *Chaos Solitons Fractals* **2022**, *164*, 112664. [CrossRef]
46. James, N.; Menzies, M.; Chan, J. Semi-metric portfolio optimization: A new algorithm reducing simultaneous asset shocks. *Econometrics* **2023**, *11*, 8. [CrossRef]
47. Corbet, S.; Meegan, A.; Larkin, C.; Lucey, B.; Yarovaya, L. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Econ. Lett.* **2018**, *165*, 28–34. [CrossRef]
48. Wang, P.; Zhang, W.; Li, X.; Shen, D. Is cryptocurrency a hedge or a safe haven for international indices? A comprehensive and dynamic perspective. *Financ. Res. Lett.* **2019**, *31*, 1–18. [CrossRef]
49. Shahzad, S.J.H.; Bouri, E.; Roubaud, D.; Kristoufek, L.; Lucey, B. Is Bitcoin a better safe-haven investment than gold and commodities? *Int. Rev. Financ. Anal.* **2019**, *63*, 322–330. [CrossRef]
50. Shahzad, S.J.H.; Bouri, E.; Roubaud, D.; Kristoufek, L. Safe haven, hedge and diversification for G7 stock markets: Gold versus bitcoin. *Econ. Model.* **2020**, *87*, 212–224. [CrossRef]
51. Bouri, E.; Shahzad, S.J.H.; Roubaud, D.; Kristoufek, L.; Lucey, B. Bitcoin, gold, and commodities as safe havens for stocks:New insight through wavelet analysis. *Q. Rev. Econ. Financ.* **2020**, *77*, 156–164. [CrossRef]
52. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Stanisz, T.; Wątorek, M. Complexity in economic and social systems: Cryptocurrency market at around COVID-19. *Entropy* **2020**, *22*, 1043. [CrossRef]
53. James, N. Dynamics, behaviours, and anomaly persistence in cryptocurrencies and equities surrounding COVID-19. *Physica A* **2021**, *570*, 125831. [CrossRef]
54. James, N.; Menzies, M.; Chan, J. Changes to the extreme and erratic behaviour of cryptocurrencies during COVID-19. *Physica A* **2021**, *565*, 125581. [CrossRef]
55. Wątorek, M.; Kwapień, J.; Drożdż, S. Cryptocurrencies are becoming part of the world global financial market. *Entropy* **2023**, *25*, 377. [CrossRef]
56. Binance. Available online: https://www.binance.com/ (accessed on 1 January 2023).
57. Marketshare. Available online: https://www.coindesk.com/markets/2023/01/04/binance-led-market-share-in-2022-despite-overall-decline-in-cex-volumes/ (accessed on 1 April 2023).
58. Tether. Available online: https://tether.to/ (accessed on 1 January 2023).
59. Farmer, J.D.; Gillemot, L.; Lillo, F.; Mike, S.; Sen, A. What really causes large price changes? *Quant. Financ.* **2004**, *4*, 383–397. [CrossRef]
60. Drożdż, S.; Forczek, M.; Kwapień, J.; Oświęcimka, P.; Rak, R. Stock market return distributions: From past to present. *Physica A* **2007**, *383*, 59–64. [CrossRef]
61. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Nunes Amaral, L.A.; Stanley, H.E. Universal and nonuniversal properties of cross-correlations in financial time series. *Phys. Rev. Lett.* **1999**, *83*, 1471–1474. [CrossRef]
62. Drożdż, S.; Kwapień, J.; Gümmer, F.; Ruf, F.; Speth, J. Are the contemporary financial fluctuations sooner converging to normal? *Acta Phys. Pol. B* **2002**, *34*, 4293–4306.
63. Nani, A. The doge worth 88 billion dollars: A case study of Dogecoin. *Convergence* **2022**, *28*, 1719–1736. [CrossRef]
64. Shahzad, S.J.H.; Anas, M.; Bouri, E. Price explosiveness in cryptocurrencies and Elon Musk's tweets. *Financ. Res. Lett.* **2022**, *47*, 102695. [CrossRef]
65. Dufour, A.; Engle, R.F. Time and the price impact of a trade. *J. Financ.* **2000**, *45*, 2467–2498. [CrossRef]
66. Weber, P.; Rosenow, B. Order book approach to price impact. *Quant. Financ.* **2005**, *5*, 357–364. [CrossRef]
67. Wilinski, M.; Cui, W.; Brabazon, A. An analysis of price impact functions of individual trades on the London Stock Exchange. *Quant. Financ.* **2014**, *15*, 1727–1735. [CrossRef]
68. Cont, R.; Kukanov, A.; Stoikov, S. The price impact of order book events. *J. Financ. Econom.* **2014**, *12*, 47–88. [CrossRef]
69. Kwapień, J.; Oświęcimka, P.; Drożdż, S. Detrended fluctuation analysis made flexible to detect range of cross-correlated fluctuations. *Phys. Rev. E* **2015**, *92*, 052815. [CrossRef] [PubMed]
70. Oświęcimka, P.; Drożdż, S.; Forczek, M.; Jadach, S.; Kwapień, J. Detrended cross-correlation analysis consistently extended to multifractality. *Phys. Rev. E* **2014**, *89*, 023305. [CrossRef] [PubMed]
71. Epps, T.W. Comovements in stock prices in the very short run. *J. Am. Stat. Assoc.* **1979**, *74*, 291–298.
72. Kwapień, J.; Drożdż, S.; Speth, J. Time scales involved in emergent market coherence. *Physica A* **2004**, *337*, 231–242. [CrossRef]
73. Toth, B.; Kertesz, J. The Epps effect revisited. *Quant. Financ.* **2009**, *9*, 793–802. [CrossRef]
74. Chen, J.; Lin, D.; Wu, J. Do cryptocurrency exchanges fake trading volumes? An empirical analysis of wash trading based on data mining. *Physica A* **2022**, *586*, 126405. [CrossRef]
75. Rak, R.; Drozdz, S.; Kwapień, J. Nonextensive statistical features of the Polish stock market fluctuations. *Physica A* **2007**, *374*, 315–324. [CrossRef]
76. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Rak, R. Quantitative features of multifractal subtleties in time series. *EPL* **2009**, *88*, 60003. [CrossRef]

77. Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Havlin, S.; Bunde, A.; Stanley, H.E. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A* **2002**, *316*, 87–114. [CrossRef]
78. Zhou, W.X. Finite-size effect and the components of multifractality in financial volatility. *Chaos, Solitons Fractals* **2012**, *45*, 147–155. [CrossRef]
79. Klamut, J.; Kutner, R.; Gubiec, T.; Struzik, Z.R. Multibranch multifractality and the phase transitions in time series of mean interevent times. *Phys. Rev. E* **2020**, *101*, 063303. [CrossRef] [PubMed]
80. Garcin, M. Fractal analysis of the multifractality of foreign exchange rates. *Math. Methods Econ. Financ.* **2020**, *13–14*, 49–73.
81. Kwapień, J.; Blasiak, P.; Drożdż, S.; Oświęcimka, P. Genuine multifractality in time series is due to temporal correlations. *Phys. Rev. E* **2023**, *107*, 034139. [CrossRef] [PubMed]
82. Oświęcimka, P.; Kwapień, J.; Drożdż, S.; Górski, A.; Rak, R. Multifractal Model of Asset Returns versus real stock market dynamics. *Acta Phys. Pol. B* **2006**, *37*, 3083–3096.
83. Oh, G.; Eom, C.; Havlin, S.; Jung, W.S.; Wang, F.; Stanley, H.E.; Kim, S. A multifractal analysis of Asian foreign exchange markets. *Eur. Phys. J. B* **2012**, *85*, 1–6. [CrossRef]
84. Wątorek, M.; Drożdż, S.; Oświęcimka, P.; Stanuszek, M. Multifractal cross-correlations between the world oil and other financial markets in 2012–2017. *Energy Econ.* **2019**, *81*, 874–885. [CrossRef]
85. Jiang, Z.Q.; Xie, W.J.; Zhou, W.X.; Sornette, D. Multifractal analysis of financial markets: A review. *Rep. Prog. Phys.* **2019**, *82*, 125901. [CrossRef]
86. James, N.; Menzies, M.; Gottwald, G.A. On financial market correlation structures and diversification benefits across and within equity sectors. *Physica A* **2022**, *604*, 127682. [CrossRef]
87. Kwapień, J.; Drożdż, S. Physical approach to complex systems. *Phys. Rep.* **2012**, *515*, 115–226. [CrossRef]
88. James, N. Evolutionary correlation, regime switching, spectral dynamics and optimal trading strategies for cryptocurrencies and equities. *Phys. D* **2022**, *434*, 133262. [CrossRef]
89. Prim, R.C. Shortest connection networks and some generalizations. *Bell Syst. Tech. J.* **1957**, *36*, 1389–1401. [CrossRef]
90. Kwapień, J.; Oświęcimka, P.; Forczek, M.; Drożdż, S. Minimum spanning tree filtering of correlations for varying time scales and size of fluctuations. *Phys. Rev. E* **2017**, *95*, 052313. [CrossRef] [PubMed]
91. Chaudhari, H.; Crane, M. Cross-correlation dynamics and community structures of cryptocurrencies. *J. Comput. Sci.* **2020**, *44*, 101130. [CrossRef]
92. James, N.; Menzies, M. Collective dynamics, diversification and optimal portfolio construction for cryptocurrencies. *arXiv Prepr.* **2023**, arXiv:2304.08902.
93. Urquhart, A.; Zhang, H. Is Bitcoin a hedge or safe haven for currencies? An intraday analysis. *Int. Rev. Financ. Anal.* **2019**, *63*, 49–57. [CrossRef]
94. 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. *Physica A* **2020**, *556*, 124759. [CrossRef]
95. Kristoufek, L. Grandpa, Grandpa, Tell Me the One About Bitcoin Being a Safe Haven: New Evidence From the COVID-19 Pandemic. *Front. Phys.* **2020**, *8*, 296. [CrossRef]
96. Yarovaya, L.; Matkovskyy, R.; Jalan, A. The COVID-19 black swan crisis: Reaction and recovery of various financial markets. *Res. Int. Bus. Financ.* **2022**, *59*, 101521. [CrossRef]
97. Wang, P.; Liu, X.; Wu, S. Dynamic linkage between Bitcoin and traditional financial assets: A comparative analysis of different time frequencies. *Entropy* **2022**, *24*, 1565. [CrossRef]
98. Zitis, P.I.; Kakinaka, S.; Umeno, K.; Hanias, M.P.; Stavrinides, S.G.; Potirakis, S.M. Investigating dynamical complexity and fractal characteristics of Bitcoin/US Dollar and Euro/US Dollar exchange rates around the COVID-19 outbreak. *Entropy* **2023**, *25*, 214. [CrossRef]
99. Dukascopy. Available online: https://www.dukascopy.com/swiss/pl/cfd/range-of-markets/ (accessed on 1 January 2023).

entropy

MDPI

Article

The Cryptocurrency Market in Transition before and after COVID-19: An Opportunity for Investors?

An Pham Ngoc Nguyen [1,2,*], Tai Tan Mai [1,3], Marija Bezbradica [1,3] and Martin Crane [1,3]

[1] School of Computing, Dublin City University, Collins Ave Ext, Whitehall, D09 Y074 Dublin, Ireland
[2] SFI Centre for Research Training in Artificial Intelligence, D02 FX65 Dublin, Ireland
[3] ADAPT Center for Digital Content Technology, D02 PN40 Dublin, Ireland
* Correspondence: ngocan.nguyenpham6@mail.dcu.ie

Abstract: We analyze the correlation between different assets in the cryptocurrency market throughout different phases, specifically bearish and bullish periods. Taking advantage of a fine-grained dataset comprising 34 historical cryptocurrency price time series collected tick-by-tick on the *HitBTC* exchange, we observe the changes in interactions among these cryptocurrencies from two aspects: time and level of granularity. Moreover, the investment decisions of investors during turbulent times caused by the COVID-19 pandemic are assessed by looking at the cryptocurrency community structure using various community detection algorithms. We found that finer-grain time series describes clearer the correlations between cryptocurrencies. Notably, a noise and trend removal scheme is applied to the original correlations thanks to the theory of random matrices and the concept of Market Component, which has never been considered in existing studies in quantitative finance. To this end, we recognized that investment decisions of cryptocurrency traders vary between bearish and bullish markets. The results of our work can help scholars, especially investors, better understand the operation of the cryptocurrency market, thereby building up an appropriate investment strategy suitable to the prevailing certain economic situation.

Keywords: cryptocurrencies; noise and trend effects; tick-by-tick data; network structure; community detection; COVID-19

Citation: Nguyen, A.P.N.; Mai, T.T.; Bezbradica, M.; Crane, M. The Cryptocurrency Market in Transition before and after COVID-19: An Opportunity for Investors? *Entropy* **2022**, *24*, 1317. https://doi.org/10.3390/e24091317

Academic Editors: Stanisław Drożdż, Jarosław Kwapień and Marcin Wątorek

Received: 19 August 2022
Accepted: 14 September 2022
Published: 19 September 2022

1. Introduction

The cryptocurrency market has become an attractive target for many financial investors in recent years due to its potential for rapid gains. One research topic being explored in this market is the correlation between different cryptocurrencies. Understanding how different assets interact with each other can help in portfolio optimization [1], predicting the future volatility or downturn [2] and also in observing the risk spillover that benefits portfolio diversification [3], to mention only a few.

Thanks to a network-based methodology, cryptocurrencies' cross-relationships can be learned and observed visually [4]. The idea of this method is that it builds up a network of different objects such that the distance between two objects depends on how similar they are: the shorter the distance, the more similar the two objects are. Eventually, we can see the interaction between objects by looking at their network's structure and analyzing characteristics of the network. Different network construction approaches have been explored in the literature, from Minimum Spanning Tree (MST) [5], k-Nearest neighbors (kNN) [6], planar maximally filtered graph (PMFG) [2] to Threshold Weighted-Minimum Dominating Set (TW-MDS) [7], to name but a few. In financial markets, normally, the similarity between two assets is measured by comparing the evolution of two corresponding price time series, one typical method to do this is Pearson correlation metric [8]. The study on correlation of traditional asset classes such as stocks, bonds, national fiat currencies and commodities has been developed a long time ago, with varying approaches invented to learn the correlation

between different entities in the same market but also between different asset classes, ranging from statistical [9,10] to AI-based methods [11].

Generally, there are two common shortcomings with correlation-related studies. Firstly, one mainly uses a low-frequency dataset such as daily or monthly, and this might cause a loss of important information from each time series, hence failing to reflect their true nature [12]. This appears to be a major concern in the cryptocurrency market, since it is well-known for its high fluctuations in terms of price movement. For example, in [13], the authors show that the losses of cryptocurrencies can reach 70% within one day. Recently, in 2020, by comparing the volatility in the returns between cryptocurrency and stock markets, the authors of [14] revealed that major cryptocurrencies such as BTC and ETH have volatilities of 5.68 and 7.10, respectively, which is two-fold higher than that of *S*&*P*500 and Euro Stoxx 50 indices. Notably, Dirk et al. calculated the daily price volatility of Bitcoin from 2001 until 2021 and found that there are extremely volatile days when the volatility can hit 120% [15]. Thus, using a high frequency means that we are ignoring valuable information (e.g., the intraday fluctuations of a time series) on purpose. As a result, this can adversely affect the correlation extracted from the dataset, potentially leading to inaccurate correlation-using experiments (e.g., portfolio optimization). Secondly, researchers tend to analyze the inter-relation between different time series by using trading price values reported on a website (e.g., Coinmarket (https://coinmarketcap.com/), Yahoo Finance (https://finance.yahoo.com/)). However, this practice deliberately ignores the effects of noise and trends in financial time series, which we will describe clearly in Section 4.

Another important factor to consider is the recent *COVID-19* pandemic which forced all countries to close off borders and restrict movements for residents as well as businesses [16]. This had a strong effect on the global downturn which occurred in March 2020 as a response to governments' efforts to control the disease spreading [17]. These historical events have been shown to disturb and devalue different financial asset classes such as stocks, bonds and also cryptocurrencies [18,19]. Instead of looking at the changes in time-series elements such as volumes, prices, returns and volatilities during the COVID-19 pandemic, in this study, we will investigate the impact of the pandemic by looking at the changes in *network structures* over time. Furthermore, based on these network's structures, we show how we can observe the corresponding community structures via community detection methods. The results from our experiment can be used to learn behaviours of investors in different periods of time, especially during downturn times in the financial market.

From the shortcomings of existing studies and utilizing the advantage of network-based analysis, this study aims to investigate the network structure of cryptocurrencies without noise and trend effects and how this structure changes under the impact of the COVID-19 pandemic. Specifically, the research target is to answer these research questions:

- RQ1. Is there evidence of the existence of noise and trend effects in the cryptocurrency market? If yes, how do noise and trend effects influence the interactions between cryptocurrencies? What does the network structure of these cryptocurrencies look like after removing noise and trend effects?
- RQ2. Does the network structure change when the level of granularity changes? If this is the case, what level of granularity should we use to obtain the true network structure?
- RQ3. Is there evidence that historical events such as the COVID-19 pandemic and the global downturn in 2020 changed the overall cryptocurrency network structure? If this is the case, how did they change it? Moreover, is there any possibility that this change was caused by a change in investors' investment strategy? In other words, does the way investors react to a downturn change the interactions between cryptocurrencies?

It should be noted that we are not new to the subject of time-varying cryptocurrency network structure, we merely build on work by the team of Drozdz, Watorek, Kwapien [20,21] as well as, more recently, Nie [22]. However, our work expands the existing studies since we consider the investment decisions of investors based on the observed network structure and we acknowledge the negative effect of not only trend but also noise

presenting in cryptocurrencies. As suggested by Miceli [23], the trend and noise removal results in a filtered MST that better explains investment strategy and also potentially uncovers endogenous or exogenous factors that drive the price of cryptocurrencies

To solve these research questions, we use a tick-by-tick dataset which consists of 34 price time series corresponding to 34 cryptocurrencies traded on the HitBTC exchange during the period between 13 February 2019 and 6 April 2021. When it comes to network formation, we calculate the correlation between cryptocurrencies by adopting the linear similarity measurement named Pearson and then construct a Minimum Spanning Tree (MST) based on these correlation coefficients. The noise and trend removal is carried out by applying Random Matrix Theory (RMT). Community structure is found by using community detection methods. In addition, different metrics are used to analyze the network structures and support our findings.

The remainder of the article is organized as follows: Section 2 presents an overview of the relevant literature. Section 3 provides a description of the dataset. Section 4 describes terminologies, methods and preprocessing procedures. Section 5 discusses the experimental results followed by implications and hypotheses. Finally, the conclusion of this study is given in Section 7.

2. Related Works

2.1. Correlation-Based Analysis in the Financial Markets

The topic of correlation analysis has a long history in connection with stock markets throughout various historical economic crises using different correlation-calculating metrics. In [24], the authors estimated the correlation between 116 *S&P*500 stocks between 1982 and 2000 using Pearson coefficient. They further used MST to build up a correlation-based network in order to observe time-varying correlations based on three network measuring metrics including *normalized tree length*, *survival ratio* and *mean occupation layer*. As a result, they pointed out a large change in the network structure during Black Monday. More recently, [6] came up with a Neural Network approach to construct a graph and found a dramatic difference in the network structure during the downturns in 2008, 2011 and 2020. In [1], a Pearson correlation matrix of 200 and 400 stocks from the CSI 300 and *S&P*500 index, respectively, was used to find an optimized portfolio following the Markowitz optimization scheme. Instead of using Pearson method, Liu et al's paper used an interesting alternative method Mutual Information to generate a distance metric to take account of non-linear effects in intra-day S&P stock data [25]. Other methods to estimate the correlation coefficients (i.e., Wavelet coherence, Fast Fourier Transform) and construct correlation-based networks (i.e., PMFG, threshold method) were introduced in several studies [2,11,26].

Different existing approaches to study the correlations in the stock market have been applied to digital coins. Some common conclusions from existing articles are that the cryptocurrency network changes over time but Ethereum tends to act as a central node in the whole network, i.e., it is a densely connected node [5,27,28]. A few works remedy the problem of dataset shortages that have been concerned in the traditional markets, i.e ones tended to use low-frequency data to implement their studies such as daily or weekly. However, they only account for a small portion of the existing literature. For example, Antonio et al. [29] used small frequency resolutions such as one hour and four hours and also consider daily data of 25 large market capitalized entities traded on the FTX exchange to discover the evolution of cryptocurrency network structures between different time frequencies. By using Pearson correlation-based MST, they found an increase in the complexity of networks' shape for coarser time resolutions. In other words, cryptocurrencies converge into a bigger group as resolution increases. On the contrary, the authors in [20] using multiple timescales starting at 10 min to 360 min proposed an opposite statement that low timescales cause the network to be centralized while it is distributed and more correlated at high timescales. They used the liquidity and capitalization differences among the assets to explain this result, since cryptocurrencies

with low capitalization are traded less frequently than those with large capitalization, it takes more time for a piece of market information to spread over such cryptocurrencies. Thus, they are more inclined to use longer scales. Notably, this is one of the very few studies that remove the trend effect from the original dataset. Interestingly, instead of using return time series like other researchers, a research using hourly realized volatility values was carried out to observe the risk spillover between 7 high-capitalized cryptocurrencies [3].

Different methods have been introduced to detect communities given a correlation matrix. The authors in [4] applied Louvain method on the MST of 119 cryptocurrencies to cluster potential communities. The time-varying dynamics from the community structures found suggests collective behaviour among these communities. With the communities found by the same method, the authors in [30] went one step further by using Principal Component Analysis (PCA) to find an optimal portfolio out of 200 cryptocurrencies in circulation. Another community detection method that is worth taking into consideration is Girvan–Newman, which has been adopted widely for multiple purposes such as link prediction, portfolio diversification, etc. [31,32]. A few other methods are also being used to grouping similar entities but are less popular such as Clauset algorithm, Stochastic block model (SBM), Latent Dirichlet Allocation (LDA) and Markov random field (MRF) [33]. One obstacle from existing studies is that some used a specific community detection algorithm only, raising a doubt about the robustness of the community structure. To this end, we first use the Louvain method to detect communities in our dataset and then adopt Girvan–Newman method to examine the robustness of the communities found earlier.

2.2. How the COVID-19 Pandemic Intervened on the Economy Worldwide

At the beginning of 2020, the economy of China started to be influenced by COVID-19, earlier than other countries. Moreover, as the world's hub for global manufacturing and trade, immediate adverse effects on the Chinese economy resulted in global impacts [16]. Different regulations have been applied to handle the disease, such as closing national borders as well as stopping business activities across the world, strongly influencing the global economy [16]. Eventually, the global financial panic in March 2020 took place. In [18], the authors pointed out that the similarity calculated by ACC and ADCC models between the US and Chinese markets increased dramatically during the pandemic. Regarding the stock prices, when the pandemic occurred, the prices of the US and Chinese stocks decreased but started to recover again since July 2020. This trend is also true for other emerging and developed stock markets in different countries from different continents such as Japan, Germany, Australia and Canada [34]. Likewise, even less risky assets such as gold were adversely affected [35]. The increase in the correlation between different financial markets in the presence of good and bad news has been observed for some decades. In [36], the authors stated that stocks are more affected by the presence of bad news, compared to good news. Moreover, bad news has a stronger correlation in traditional markets. These results align with what happened during the COVID-19 pandemic. Although the world continued facing different COVID-19 waves afterwards, its impact on different asset classes lessened significantly [37], stock prices increased and volatilities decreased again to their original values before the pandemic [38]. Furthermore, the connectedness between different assets also experienced a major decline [39].

In [19], the authors investigated the impact of the COVID-19 pandemic on the cryptocurrency market by using daily prices of 45 well-known cryptocurrencies between September 2019 and April 2020—the majority of which are also used in our present study. In particular, they measured the stability of cryptocurrency time series using Largest Lyapunov Exponent and Approximate Entropy. All time series are divided into two parts: the first part spans September to December 2019, considered normal time, while the second spans January to April 2020, considered a pandemic period. They revealed that the pandemic increases in cryptocurrency market uncertainty as prices fluctuated significantly. Moreover, the same experiment has also been carried out on the stock market, results indicating a lower level of price fluctuations in the stock compared to digital currencies.

Also on the same topic, Drozdz et al. [21] compared the Pearson correlation between the cryptocurrency market and different asset classes including stocks, fiat currencies and commodities, revealing that these conventional markets easily influence the cryptocurrency market when they are in turbulent times, while there is no significant correlation between digital currencies and other markets in normal times, given the time resolutions they used are 10 and 360 mins.

Reactions of the general public to the COVID-19 outbreak were also observed to examine its relationship with cryptocurrencies' returns. For example, authors in [40] measured the fear of people by the frequency of occurrence of keywords *COVID-19* and *coronavirus* on Google Trends (https://www.thinkwithgoogle.com/, accessed on 4 August 2022). Thanks to the vector autoregressive (VAR) models, they compared the evolution of this fear with the stock market's expectation of volatility VIX index (the VIX index is a measure of constant, 30-day expected volatility of the US stock market, derived from real-time, mid-quote prices of *S&P*500. Normally, it is calculated using the Black–Scholes formula) as well as the Bitcoin returns. They found that increases of fear can lead to Bitcoin crashes, as the correlation coefficient is -0.9. Furthermore, negative sentiment generated by coronavirus news is associated with market volatility, which is in line with other findings such as in [41]. Interestingly, some studies on the relationship between news-based sentiment and cryptocurrencies showed that, although both bad and good news cause the change in the returns and volatilities of cryptocurrencies, positive news has more effect on the volatilities and returns of cryptocurrencies in comparison with negative news [42–44].

Recently, network analysis in the cryptocurrency market during the COVID-19 pandemic has been carried out, with the common result being that the pandemic, as well as the global downturn, actually caused a change in the network structure of the cryptocurrency market. Specifically, cryptocurrencies tend to form bigger groups during the downtime, i.e., the number of potential clusters found in the network decreases during the downtime, with a few cryptocurrencies acting as central nodes. This topic has only been explored in a few studies to date [21,22,45,46]. Moreover, there are some gaps: (1) the lack of deep investigation of the network structure as they only consider MSTs; (2) the noise and trend effects are not removed; (3) data limitation issues.

We will address these shortcomings by doing deeper experiments on the network structure of the cryptocurrency market before, during and after the COVID-19 pandemic via a longer dataset with the effect of noise and trend removed. In addition, we will look at the way cryptocurrencies form a group during turbulent times by considering their rankings (identified by its market capitalization, the larger its maket capitalization, the higher its rank). We believe that this research can propose a better understanding of interconnections between digital currencies during standard and unstable periods. Furthermore, we also aim at understanding the investment decision of investors in different market states based on the results of community detection.

3. Data Description

All experiments in this study have been carried out based on a tick-by-tick price dataset (tick data are the highest resolution intraday data and consist of the sequence of each executed trade or bid/ask quote aggregated from an exchange) that was collected from the hitBTC exchange (a platform for digital asset and currency exchange to quickly and securely trade cryptocurrencies—website address: https://hitbtc.com/) from 13 February 2019 to 6 April 2021. The dataset comprises 34 cryptocurrencies with a hybrid of high and low rankings. Specifically, the highest rank is 1 (Bitcoin) while the lowest rank is 260 (FunToken), according to the price-checking website *Coinmarketcap* (https://coinmarketcap.com, accessed on 4 August 2022) in April 2021; full list in Table 1.

Table 1. A list of 34 cryptocurrencies used in this study. Abbreviations are put in parentheses.

Cryptocurrencies					
Argur (REP)	Bitcoin SV (BSV)	Ethereum Classic (ETC)	MaidSafeCoin (MAID)	Ontology (ONT)	Tron (TRX)
Bancor (BNT)	Cardano (ADA)	FunToken (FUN)	Maker (MKR)	Ox (ZRX)	Verge (XVG)
Basic Attention Token(BAT)	Decentraland (MANA)	ICON (ICX)	Monero (XMR)	QTUM	Zcash (ZEC)
Bitcoin (BTC)	Dogecoin (DOGE)	IOST	Nem (XEM)	Ripple (XRP)	Zilliqa (ZIL)
Bitcoin Cash (BCH)	EOS	Lisk (LSK)	NEO	Stellar (XLM)	
Bitcoin Gold (BTG)	Ethereum (ETH)	Litecoin (LTC)	OMG Network (OMG)	Tezos (XTZ)	

3.1. A Note on Data Sampling and Missing Data

Since price values are collected tick-by-tick, there is no fixed timescale for all cryptocurrencies leading to an inconsistency between the time series. For this reason, we re-sample the dataset by using data points at a specific timescale. In particular, we choose four different timescales, namely 30 min, 6 h, 12 h and 24 h. Each data point of a dataset is taken to be the price of the last transaction of 34 cryptocurrencies within the considered timescale. Eventually, we have four datasets corresponding to four different timescales. Table 2 shows the description of each re-sampled dataset.

Table 2. Characteristics of four re-sampled datasets at four different levels of granularity.

Level of Granularity	# Data Points	# Missing Values
30 min	37,632	289 (0.8%)
6 h	3136	24 (0.8%)
12 h	1568	12 (0.8%)
24 h	784	0 (0%)

Three out of four datasets have missing values with the same percentage of 0.8%. Note that a data point of a dataset is considered missing if at least one cryptocurrency does not have the price value at this data point. For each time series, instead of simply removing missing values from the time series and values from other time series from the same time, we replace missing values with the average value of the corresponding time series. This technique has been adopted in different research topics with good performance [47–49]. Furthermore, we notice that this does not change the statistical properties of the correlation between time series but, instead, helps to keep more information and thus the results found from conducting the experiments are more reliable and accurate.

3.2. Aggregational Gaussianity

Aggregational Gaussianity is considered a stylized fact in traditional financial markets. In [50], the authors observed the evolution of distributions of the IBM stock returns by looking at different levels of granularity, e.g., 30 min, one day, one week and one month, finding evidence of Aggregational Gaussianity. Another study on this topic drawing the same conclusion is described in [51]. However, these authors used different stocks and a higher set of timescales from one day to one year, showing that this stylized fact is also true for stocks at coarser time resolutions.

We investigate whether Aggregational Gaussianity exists in our log-return time series using a set of four timescales: 30 min, 6 h, 12 h and 1 day. We observe this statistical aspect by implementing three experiments: Firstly, we construct the histogram as well as kernel density estimation (KDE) for each cryptocurrency time series. Secondly, we generate the

Q-Q plot, which is a popular approach to test normality for a time series [52]. Lastly, we use the Lilliefors hypothesis test for normality [53]. We obtained the following findings: firstly, although the distributions of these cryptocurrency time series have a bell curve shape at all timescales considered, they are not (from the Q-Q plot and Lilliefors test) normally distributed; secondly, however, there appears to be evidence to say that Aggregational Gaussianity exists in all cryptocurrencies used in this present study from the Q-Q plots. This result is in line with existing findings in the cryptocurrency market such as [54,55].

4. Research Methodology

4.1. Correlation Matrix Based on Pearson Coefficients and Random Matrix Theory

Given x_i is the price time series of cryptocurrency i, we use its return values to find the correlation between cryptocurrencies. This is because Return values are represented as a percentage, making them scale-free and especially, stationary, which is an important requirement for many statistical tools, such as *Normalization*. Thus, we first calculate the corresponding return time series r_i as follows [56]: $r_i = log\left(x_i^t / x_i^{t-1}\right)$, where x_i^t is the price value of the cryptocurrency i at timestamp t.

Each of these return time series can be normalized as follows [57]: $\hat{r_i} = (r_i - \mu_i)/\sigma_i$, where μ_i and σ_i are the average value and standard deviation of time series i, respectively.

We form a $m \times n$ matrix $\mathbf{G}$ such that each column represents a normalized return time series of a cryptocurrency and each row represents a timestamp. The corresponding correlation matrix $\mathbf{C}$ can be expressed as follows [56]: $\mathbf{C} = \frac{1}{m}\mathbf{G}\mathbf{G}^\intercal$. In other words, each element C_{ij} of $\mathbf{C}$ shows the correlation strength between cryptocurrencies i and j by calculating the dot product of the two normalized return time series, $C_{ij} =< \hat{r_i}, \hat{r_j} >$. Such a correlation matrix is called *Pearson correlation matrix*.

It should be noted that Pearson correlation has some limitations as described in [58]. In particular, its sensitivity to outliers and inability to capture non-linear relationships both have the potential to cause misleading results. However, we believe that this correlation metric is appropriate to use in our study for the following reasons:

- Firstly, we make use of cryptocurrency returns in order to retain the statistical nature of the associated time series. While some authors have proposed addressing the non-linearity problem (e.g., Spearman [59] and Kendall [53]), these have the disadvantage of converting rational numbers into integer rankings, with the potential to lose out on critical information from financial time series [60]. Moreover, it has been shown that rank correlation metrics also suffer from the nonlinearity issue in some cases [58].
- Secondly, Pearson has been widely applied in the existing literature, not only in the cryptocurrency market [21,22,32] but also in markets for more traditional asset classes [2,6,24]. This strongly reinforces our belief in the applicability of this method of correlation calculation for our problem.
- Thirdly, rank-based correlation metrics require independent observations. This is a known weakness of non-linear correlation methods such as Spearman and Kendall [60]. On the other hand, Pearson works well for time series with duplicate observations (because there is no requirement for independent observations), as is the case in financial time series. For example, the price of a cryptocurrency can be unchanged for a period of time.

One issue raised from this type of matrix is the question of how reliable these correlations are, in other words, whether the correlation matrix shows genuine and authentic relationships between the considered time series. Thanks to the RMT [61], this hypothesis can be examined. Particularly, given a $m \times n$ random matrix $\mathbf{N}$ whose elements are distributed randomly with zero mean and unit variance, the eigenvalue distribution of the correlation matrix $\mathbf{C_N} = \frac{1}{m}\mathbf{N}\mathbf{N}^\intercal$ follows the Marchenko–Pastur probability density function [62] if the Quality Factor $Q = \frac{m}{n} \geq 1$ holds when the number of timestamps $m \to \infty$ and the number of features $n \to \infty$: $P(\lambda) = \frac{Q}{2\pi}\frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}$, where P is the Marchenko–Pastur

probability density function, λ is an eigenvalue of $\mathbf{C_N}$, $\lambda_\pm = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}$ are upper and lower limits, respectively.

From RMT, eigenvalues falling outside of $[\lambda_-, \lambda_+]$ are assumed to deviate from its expected predictions [63,64]. Hence, we can use this theory to test the reliability of the relationships in our empirical data [65]. That is, if an empirical correlation matrix actually has real valuable information, it must have eigenvalues that are outside the bounds of $[\lambda_-, \lambda_+]$. Otherwise, the empirical correlation matrix can be taken to contain mainly random noise. In this study, RMT has been used to test our correlation matrices. The results show that all correlation matrices are not random and contain valuable information.

4.2. Cleaning Trend and Noise Effects in the Cryptocurrency Market

4.2.1. Noise and Trend

The cryptocurrency market is known to have a higher percentage of noise than other traditional financial markets. According to [66], the average daily signal-to-noise ratio of the cryptocurrency market is 36%, which is extremely low compared to well-established US stock exchanges such as NYSE and NASDAQ, with an average daily signal-to-noise ratio of 90%, given the considered period between March 2017 and November 2017. The noise in the cryptocurrency market might come from different sources. For instance, there is no fixed volume for a transaction to be executed at a time, so investors can freely choose the amount that they want to trade; however, this issue causes one problem, in that investors can reduce the transaction costs by splitting their budget into smaller pieces and then buy one cryptocurrency many times with different amounts of volume and price, a practice which can trigger unforseen price movements, see [67]. Furthermore, cryptocurrencies' prices are vulnerable to "pump and dump" schemes [68], which have become pervasive recently, and also regulatory news enacted by national authorities [69]. All of these factors might intervene in the price movements of digital assets. Consequently, the correlation matrix between cryptocurrencies cannot explain their real connections as it is highly influenced by these noise factors.

On the other hand, the trend effect found in other correlated systems [70] might be found in the cryptocurrency market. Briefly speaking, a trend among cryptocurrencies means that they tend to move together in terms of price values. We notice that the majority of cryptocurrencies are created based on the protocol of leading cryptocurrencies such as Bitcoin and Ethereum (e.g., MKR, BNT, ICX, ETC and LTC) [71]. Moreover, cryptocurrencies' prices readily fluctuate with mass media [72], causing a herding behavior [72]. Similar characteristics contribute to creating a trend in cryptocurrencies.

Generally, these phenomena might be reasons for a high-value correlation matrix of cryptocurrencies from our dataset. Thus, it is important to remove of the existing noise and trend before moving on to further analysis.

4.2.2. Cleaning Method

In recent studies, different approaches have been proposed to remove the noise from a correlation matrix through modification of the corresponding eigenspectrum, e.g., Linear shrinkage [73], Eigenvector clipping [74], Non-linear shrinkage [75] and Rotationally invariant, optimal shrinkage [76]. One common obstacle for most of the existing cleaning methods is that they have parameters needing definition. This raises an obvious question: how do we choose these? It is acknowledged that a lot of effort has been made to obtain the right parameter values, i.e., the noise is removed completely without the loss of data information [77,78]. However, these optimization approaches have one issue, which is that they use the Frobenius norm in their formula, so they fail to work with outlier-containing data, a downside of the Frobenius metric [79]. On the other hand, Eigenvector Clipping distinguishes itself from others [74] as it does not require any training parameters, making its outcome robust and more reliable. Furthermore, this cleaning method is straightforward to implement, with the guaranteed efficiency as it keeps the information part, i.e., after the cleaning process, the trace of the correlation matrix remains unchanged [80]. This

method has shown good performance in different studies and has been applied widely to different topics such as programming education, portfolio optimization and signal processing [70,81,82]. The outstanding performance of the Eigenvector clipping encourages us to choose this method for our cleaning scheme.

Given eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_n$ and corresponding eigenvectors $v_1, v_2, \ldots, v_n$ of our empirical correlation matrix $\mathbf{C}$, we can identify $k \leq n$ such that $\lambda_k > \lambda_+$ and $\lambda_{k+1} \leq \lambda_+$. The Eigenvector clipping defines the denoised correlation matrix $\mathbf{C}_{denoised}$ by [83]:

$$\mathbf{C}_{denoised} = \Sigma_{i=1}^{n} \lambda_i^* v_i v_i^\intercal, \lambda_i^* = \begin{cases} \frac{\lambda_{k+1}+\lambda_{k+2}+\ldots+\lambda_n}{n-k}, \forall i \geq k+1 \\ \lambda_i, \forall i \leq k \end{cases} \quad (1)$$

Equation (1) uses the same eigenvectors as $\mathbf{C}$ but modifies their corresponding eigenvalues such that those greater than λ_+ remain unchanged while the rest will be replaced by their average value. Notably, although small eigenvalues are replaced, the trace of the denoised correlation matrix is equal to its origin.

Regarding the trend effect, it is explained by the first eigenvalue and eigenvector, referred to as "market component" [83]. The market component is proved to influence the outcome of the correlation matrix. In particular, it is involved in all interactions observed from the correlation matrix due to its enormous amount of information, consequently, lessening the performance of clustering algorithms [84]. Thus, removing this component is a necessary step to clean the trend effect so that a greater portion of the correlation can be explained by components that affect specific subsets of the cryptocurrencies and, hence, facilitate clustering algorithms to find dissimilarities across clusters. A cleaned correlation matrix $\mathbf{C}_{cleaned}$ is obtained by subtracting the market component from the denoised correlation matrix:

$$\mathbf{C}_{cleaned} = \mathbf{C}_{denoised} - \lambda_1 v_1 v_1^\intercal \quad (2)$$

We found that the connections between cryptocurrencies decrease greatly without noise and trend effects: large cryptocurrencies such as Bitcoin, Ethereum and Ripple do not see to affect the cryptocurrency market as they did before the cleaning process, since there is no strong connection between them and other cryptocurrencies. This result is in line with [70], where the Eigenvalue Clipping method was also used to clean the education-related correlation matrix.

4.3. Distance Matrix and Its Minimum Spanning Tree

Although the correlation coefficient can explain some aspects of the relationships between cryptocurrencies, it is not a metric [85]. Thus, the connections learned from the correlation matrix lack topological characteristics because they are not placed in a metric space [85]. To tackle this issue, a concept named *Distance Matrix* has been introduced to replace the correlation matrix.

Let $\mathbf{D}$ be a distance matrix deriving from $\mathbf{C}_{cleaned}$, then:

$$d_{ij} = \sqrt{2 * (1 - c_{ij})} \quad (3)$$

where $d_{ij} \in [0, 2]$ is an element of $\mathbf{D}$, with 0 indicates the complete similarity between 2 nodes while 2 indicates the complete difference between 2 nodes. From the Equation (3), we can prove that: (1) $d_{ij} \geq 0$, (2) $d_{ij} = 0$ if $i = j$ and (3) $d_{ij} = d_{ji}$, i.e., the requirements of a metric are satisfied [85]. By using the distance matrix, we can derive a network (graph) of cryptocurrencies (nodes) with a specific topology, where similar cryptocurrencies are close to each other and cryptocurrencies with different behaviors are far away from each other, the link (edge) between each pair of cryptocurrencies is their distance value. Thanks to this topology, different communities of cryptocurrencies can be observed.

One problem with this type of network is that it is dense. That is, for a set of N nodes, the corresponding graph deriving from $\mathbf{D}$ has $\frac{N\times(N-1)}{2}$ edges such that each vertex connects to all other vertices. To reduce the complexity of the network, we use a Minimum Spanning Tree (MST) [86], which refers to a special tree from the graph that links all vertices together in which its length is minimal. Particularly, it reduces the amount of redundant information since it only keeps the $N-1$ most important edges, i.e., $N-1$ shortest edges that are well connected. MST stems from graph theory and is applied widely to different fields [4,87,88], especially in financial markets [89–91]. To exploit the useability of MST, the dynamics of community structures in the stock market are observed by Huang et al. [92] with the dataset split into consecutive smaller periods and a MST constructed at each of them. Thus, the characteristics of a financial network can be captured by observing the evolution of MSTs. More recently, the cryptocurrency market was introduced and attracted a number of investors, and the demand for exploring the correlation between cryptocurrencies thereby emerged. However, this topic is rather new and needs more studies to be implemented [4,93].

There are two famous algorithms to find the MST, namely Prim [93] and Kruskal [94]. While both methods show good performance, Kruskal seems to be better in terms of time complexity. A comparison between the two from [95] shows that the prior works well with a big network, while the latter is dominant when the network is small, which is appropriate for this study as we have only 34 cryptocurrencies. Moreover, Kruskal is used more often in finance-related topics compared to other approaches [96–98], which strengthens the reliability of the algorithm. With these advantages, we choose Kruskal for this study.

4.4. Community Detection in the Cryptocurrency Market

Given a MST from the distance matrix $\mathbf{D}$, different communities are formed and can be recognized clearly, i.e., cryptocurrencies belonging to one community have short distance edges among them and the distance between two others in two different communities is longer than any edges of these two communities. However, there are less common cases in which some nodes are scattered between communities, or it is not visible from the graph how close the two communities are. This issue motivates us to further analyze the MST to optimize the clustering result using several community detection methods which have been developed [99–103]. Of these, the Louvain method is applicable across a wide range of domains [104–107]. Thus, we apply this method to our MST in order to obtain optimal communities. Theoretically, Louvain is an optimization problem that uses *Modularity* to measure the density of links inside communities compared to links between communities. The target of Louvain is to minimize the Modularity measure, which means that different authentic communities are clustered very tightly [108].

However, it is not convincing just to show results from one method only, as the community structure of a network might be just random. To overcome this issue, we also adopt another commonly used method named Girvan–Newman, which removes edges from the original graph one-by-one such that the edge having the highest number of shortest paths between nodes passing through it is removed first. Eventually, the graph breaks down into smaller pieces, so-called communities [109].

If the results proposed by these two community detection methods are similar, it implies that the relationship of the cryptocurrencies as well as their corresponding community structure are reliable and reflects their genuine characteristics. The results after applying these methods are shown in Section 5.

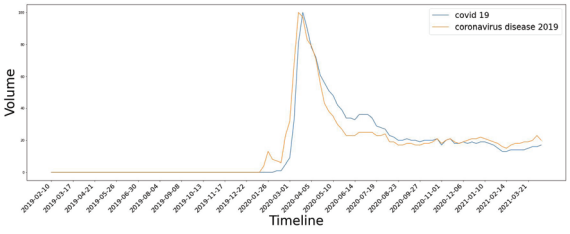
4.5. Time Window Division

Given the dataset described earlier, one important question about constructing a network structure in these cryptocurrencies is how to split the dataset into different consecutive periods. This is because a network structure corresponding to each period of time should be able to explain what has happened to the cryptocurrencies throughout that time, i.e., there must be a reason behind this topological structure. If we divided the dataset

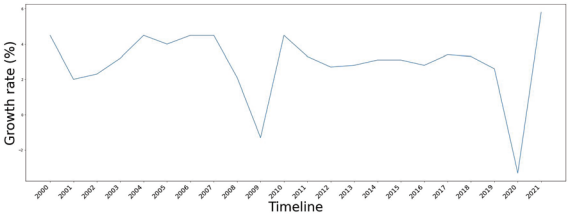
randomly, we could not capture important historical events at a specific period. As a result, the topology we found would be meaningless in the corresponding time window. To this end, we must select time windows rationally. We note that our dataset contains the period of the COVID-19 pandemic as well as the global downturn 2020. From the literature in Section 2, we see these historical events actually adversely influenced the financial markets. Thus, we postulate that the COVID-19 pandemic is a reasonable milestone to separate our dataset.

To verify the pandemic's impact on the global economy and thereby choose the right time windows for the dataset, we consider the movements of four different factors. Firstly, the attention to the COVID-19 pandemic, as measured by the frequency of COVID-19-related keywords searched on Google Trends. For this factor, we use two keywords including *COVID-19* and *coronavirus disease 19* . Secondly, we use the VIX index to observe fluctuations of the stock market, this index starts at 0 for no upper bound and a higher value implies that the stock market has stronger fluctuation. Thirdly, we also observe the prices of the *S*&*P*500 index, representing the US economy. Lastly, the growth rate of the world's GDP is used as a proxy for the development of the global economy in general.

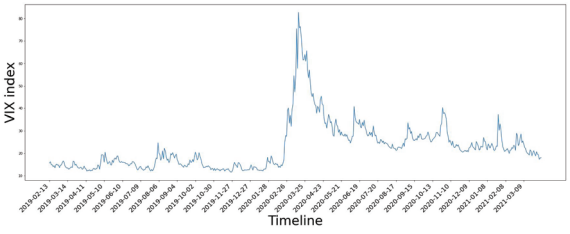
Figure 1 visualizes these aforementioned factors. From Figure 1a, people started to worry about this disease in January 2020. However, it was not until March 2020 that the COVID-19 pandemic actually caught the attention of people worldwide, as the volume of searches for COVID-19-related terms quickly peaked. This remained a topic of interest until July 2020. Furthermore, March 2020 was the month in which a pandemic-induced economic recession first occurred, seriously affecting the economy of nations worldwide. This effect is shown in Figure 1b–d. In particular, the GDP's growth rate decreased by 3.3% in 2020, which is the highest decrease ever, even worse than the Great Recession in 2007–2009 [110]. Simultaneously, the stock market fluctuated dramatically, which can be seen via the VIX index and the *S*&*P*500 index, both of which experienced a significant fall during March 2020. However, the economy started to recover afterward, the stock market became less fluctuated and the *S*&*P*500 index regained its original pre-pandemic value in July 2020.



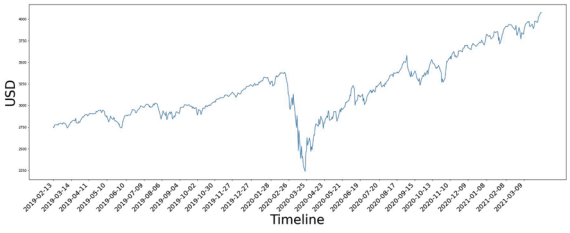
(**a**) Public attention



(**b**) GDP growth rate



(**c**) VIX index



(**d**) *S*&*P*5000 index

Figure 1. The reaction of general public and global economy to the COVID-19 pandemic. Four factors are considered: (**a**) Worldwide attention to the pandemic, (**b**) Global GDP growth, (**c**) VIX index, (**d**) *S*&*P*500 index.

Consequently, we split the 784 days from 13 February 2019 to 6 April 2021 into three time windows which correspond to three different stages, including normal time, downturn time and recovery time. The details for these time windows are shown in Table 3.

Table 3. Three time windows used in this work (time windows split to take into consideration the COVID-19 pandemic).

Time Window	Stage	Time Span	# Days
1	Normal time	13 February 2019–31 December 2019	322 days
2	Downturn time	1 January 2020–30 June 2020	182 days
3	Recovery time	1 July 2020–6 April 2021	280 days

5. Experimental Results and Discussion

This section sets out our three research questions. We will first examine the impact of noise and trend effects on the correlation between cryptocurrencies as well as their corresponding topological structure. Then, we observe the evolution of the structure according to the levels of granularity. Finally, the results from these two experiments will be used to construct the right network structure. Consequently, the corresponding community structure is identified, which is used to learn the investment decisions of crypto investors during the COVID-19 pandemic.

We note that all calculations in our study are implemented using *Python* programming language (version 3.7.14, designed by Guido van Rossum, Centrum Wiskunde & Informatica (CWI), The Netherlands). Regarding network-related calculations (e.g., network construction and network-involved metrics), we utilize the *networkx* (https://networkx.org/) package incorporated into Python.

5.1. The Response of Network Structures to Noise and Trend Effects

Given the fact that there are noise and trends in the cryptocurrency market, we examine whether these factors affect the cryptocurrency network structure. Since we have four datasets corresponding to four timescales (e.g., 30 min, 6 h, 12 h and 24 h), we use both metric-related methods and visualization for all available datasets to discover the discrepancy between original and cleaned (after removing noise and trends) datasets.

To show the difference between two network structures, we choose two such metrics to measure the connection strength in a network of cryptocurrencies:

- Residuality Coefficient [93]: This compares the relative strength of the connections above and below a threshold distance value. In this experiment, we use the highest distance value ensuring connectivity of the MST as the threshold, denoted L:

$$R = \frac{\Sigma_{[d_{ij}>L]} d_{ij}^{-1}}{\Sigma_{[d_{ij}\leq L]} d_{ij}^{-1}} \tag{4}$$

- MST-based mean distance [111]: this calculates the average distance of the MST:

$$M = \frac{1}{N-1} \Sigma_{d_{ij} \in MST} d_{ij} \tag{5}$$

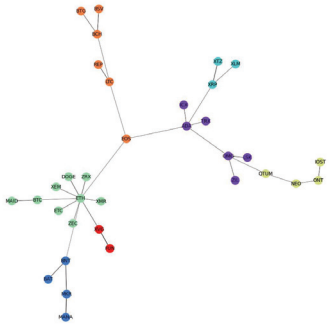
An increase in these means that cryptocurrencies are further from each other. By contrast, cryptocurrencies are closer to each other if these metrics decrease. Note that although both metrics are used to examine the connection strength of cryptocurrencies, the Residuality coefficient is known to be more vulnerable to the links between cryptocurrencies in different groups, i.e., if the connection strength between cryptocurrencies in different groups increases, the Residuality coefficient will decrease dramatically, and vice versa; the connections between cryptocurrencies within one group do not affect the Residuality coefficient much [112]. On the other hand, Mean distance is more vulnerable to the

links between cryptocurrencies belonging to one group, as it mainly uses the connections within a group to find the average value and ignores the connections between different groups [111].

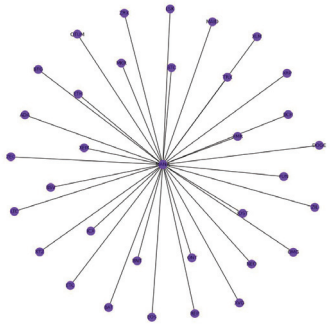
Table 4 shows the results of the two metrics using different levels of granularity. It is clear that both Residuality coefficients and Mean distance values increase significantly when the effects of noise and trend are dismissed. This phenomenon remains unchanged in different timescales, implying that this is a genuine characteristic of the cryptocurrency market. Furthermore, a visualization of network structures before and after cleaning is shown in Figure 2 to reinforce our finding. As can be seen, the topological structure changes after the noise and trends are removed. Moreover, what happens in each time window is that the number of communities decreases after removing these effects. From these figures and illustrations, we can conclude that the connections between cryptocurrencies are caused mainly by the noise and trend effects. That is, these factors result in different cryptocurrencies becoming closer to each other and forming a group. This phenomenon can be explained by low values for Residuality coefficients and Mean distance values in the original data compared to the cleaned data. A value less than unity of the prior metric means that there are few connections greater than the threshold L. Moreover, a small value of the latter metric means that cryptocurrencies within a group are closer to each other. In summary, each group of the network is compact with strong links inside, which helps the community detection algorithm to easily cluster them. In other words, the difference between different groups is clear because the links between different groups are weak, i.e., the ones greater than L. However, after cleaning the correlation matrix, cryptocurrencies that are closely related to each other through noise and trend become further away, i.e., the strong links between some cryptocurrencies are broken. This causes our metrics to increase dramatically, which means that the network structure starts to expand, forming a sparse network. For example, the Residuality coefficient of the second time window in the 30 min original data is 0.28, while it is 20 times higher after cleaning the effects of noise and trends. This fact is also true for the rest of our dataset. The result is in line with [20]; these authors did not consider the noise effect but, with the removal of trends, they found that the correlation between the 80 most liquid cryptocurrencies from 1 January 2020 to 1 October 2021 decreased.

Table 4. Cryptocurrency network connection strength through three time windows measured by Residuality Coefficient and Mean Distance. Four different granularity levels are considered, each with datasets, including original and cleaned dataset after removing noise and trend effects.

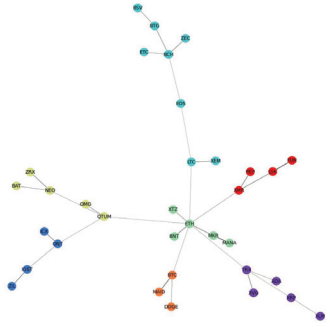
Metric	Data Type	Time Window	Granularity			
			30 min	6 h	12 h	24 h
Residuality Coefficient	Original Data	1	0.41	0.11	0.16	0.08
		2	0.28	0.111	0.06	0.05
		3	0.14	0.05	0.07	0.34
	Cleaned data	1	1.69	6.66	14.82	14.40
		2	5.98	8.90	14.41	15.34
		3	2.32	2.99	1.88	1.05
Mean distance	Original Data	1	1.08	0.82	0.80	0.76
		2	0.99	0.71	0.65	0.56
		3	0.98	0.57	0.46	0.45
	Cleaned data	1	1.29	1.38	1.42	1.42
		2	1.40	1.42	1.42	1.42
		3	1.29	1.12	1.01	1.22



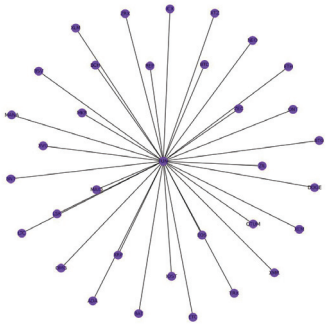
(**a**) Time window 1, original



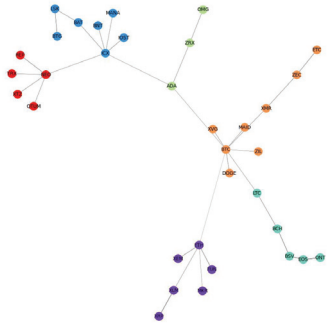
(**b**) Time window 1, cleaned



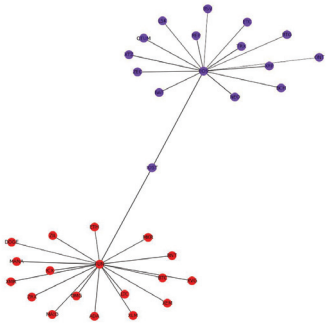
(**c**) Time window 2, original



(**d**) Time window 2, cleaned



(**e**) Time window 3, original



(**f**) Time window 3, cleaned

Figure 2. Cryptocurrency network structures using daily data. For each time window, Louvain method is applied to both original and cleaned data to detect existing communities. The illustrations on the left and right hand side are for the original and cleaned data, respectively, for 3 time windows referring to normal, downturn and recovery times, respectively.

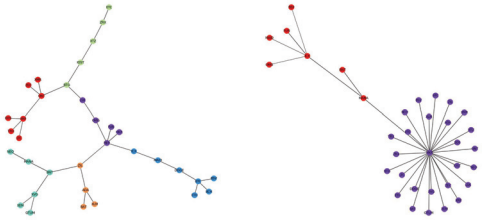
5.2. Real Network Structures in Different Levels of Granularity: An Experiment on Cleaned Data

In this section, we will construct the network structure of 34 cryptocurrencies removing the effects of noise and trends. By doing this, we can look at the evolution of network

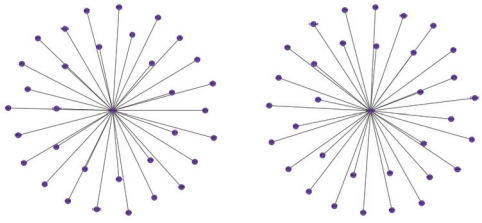
structures at each timescale over time and, of greater interest, the differences in the network structures between different timescales. Note that community detection results found by using Louvain algorithm are also included in these networks. The results of this experiment can shed light on the influence of timescales on cryptocurrencies' connections and what timescale should be used for cryptocurrency-related analysis.

5.2.1. The Evolution of the Cryptocurrency Network According to Timescales

Figures 3–5 show the results of network structures along with detected communities using the Louvain method with each figure representing a different time window. For each window, four network structures corresponding to four different levels of granularity are displayed. One obvious statement that can be made from the illustrations is that the community structures at each level of granularity change over time. Additionally, if we consider different levels of granularity at the same time, the number of detected communities tends to decrease when the timescale becomes more coarse-grained. For large timescales, such as 24 h, cryptocurrencies build up big groups with few cryptocurrencies acting as central nodes that link directly to the remainder. For example, in Figure 3d, MANA acts as a central node that links all other cryptocurrencies together. This explains why community detection techniques cannot distinguish several subsets as the network in this case is naturally one group. Figure 4d shows a similar pattern, while in Figure 5d there are two central nodes that create two big groups with relatively similar sizes. To this end, with low-frequency data, we expect we can predict the long-term trend of cryptocurrencies in the future by looking at the central nodes from their corresponding community structures. If this is the case, it will be very beneficial for investors who choose a long-term investment. However, this behaviour requires deeper investigation and will be the subject of further research.



(**a**) Time window 1, 30 min (**b**) Time window 1, 6 h

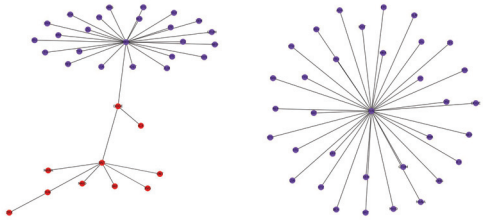


(**c**) Time window 1, 12 h (**d**) Time window 1, 24 h

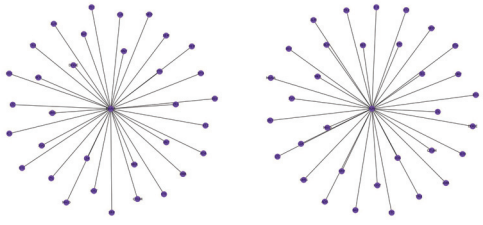
Figure 3. Network structure for the first time window, community detection is applied using Louvain method. Four different timescales are used, e.g., (**a**) 30 min, (**b**) 6 h, (**c**) 12 h, (**d**) 24 h.

We notice that the difficulty of detecting communities in this market increases with the timescale length. In other words, cryptocurrencies are more likely to belong to the same community if we just look at their price values at a high level of granularity such as daily. Thankfully, it can be explained based on the nature of the cryptocurrency market. In particular, the cryptocurrency market is well-known for its high volatility compared to

other traditional asset classes such as stocks, bonds and commodities [113–116]. In [117], the authors used 5 min data of Bitcoin prices traded on three different exchanges, Kraken, Bitstamp and Btcbox, during the period between 2017 and 2021 to calculate the realized volatility (the assessment of variation in returns for an asset by analyzing its historical returns within a defined time period) of this most stable and popular cryptocurrency. The results showed that although Bitcoin is the most valuable and trustworthy cryptocurrency, its volatility fluctuates from 4.8 to 7.5. By contrast, with the same level of granularity, the stock market seems to be more stable, as the realized volatility stood at roughly 0.58 during normal times [118] and increased to just around 1.0 during the COVID-19 pandemic [119]. These facts suggest that the cryptocurrency price fluctuations are dramatic even within a 5 min period. Consequently, using a low-frequency dataset such as 12 h or 24 h appears to cause a loss of important information that influences the results of analysis. This problem has also been described in earlier studies such as [12]. However, existing studies mainly focused on daily data to detect communities in the cryptocurrency market.



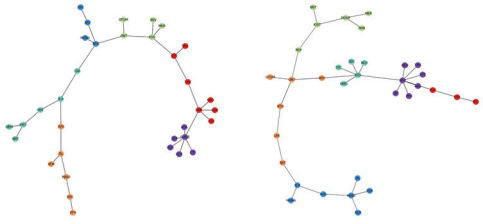
(**a**) Time window 2, 30 min (**b**) Time window 2, 6 h



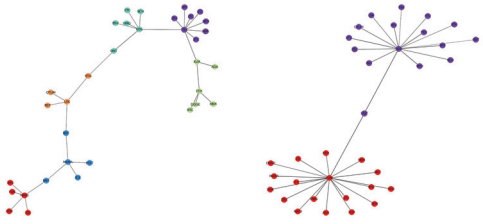
(**c**) Time window 2, 12 h (**d**) Time window 2, 24 h

Figure 4. Network structure for the second time window, community detection is applied using Louvain method. Four different timescales are used, e.g., (**a**) 30 min, (**b**) 6 h, (**c**) 12 h, (**d**) 24 h.

In this study, the loss of information by using large timescales including 6 h, 12 h and 24 h makes judging the correlation between different cryptocurrencies unclear. As a result, it affects the corresponding MST which can be seen in Figures 3–5. Ideally, we would like to use a dataset that is as fine-grained as possible. Unfortunately, our experiments show that for frequencies lower than 30 min, there are a huge amount of missing values as some cryptocurrencies are not traded frequently [120], thus requiring their removal or imputing a value. This adversely affects the correlation between time series and impacts our analysis. Finally, we choose a 30 min dataset for further experiments.



(**a**) Time window 3, 30 min (**b**) Time window 3, 6 h



(**c**) Time window 3, 12 h (**d**) Time window 3, 24 h

Figure 5. Network structure for the third time window, community detection is applied using Louvain method. Four different timescales are used, e.g., (**a**) 30 min, (**b**) 6 h, (**c**) 12 h, (**d**) 24 h.

5.2.2. Louvain vs. Girvan–Newman for Community Structure Detection

The Louvain method is our main technique for detecting communities but we also use the Girvan–Newman method to double-check the communities found. The *v-measure* gives the similarity between these two methods [121], shown in Table 5. This metric ranges from 0 to 1 such that 0 indicates a complete dissimilarity between two graphs while 1 indicates a complete similarity. We found that the *v-measure* in all cases is high with the lowest value of 0.82 from the third time window in the 6 h dataset in Table 5. That is, the Louvain method proposes similar results as Girvan–Newman. Thus, the communities found by Louvain are reliable for use in further analysis.

Table 5. *v-Measure* between Louvain and Girvan–Newman methods.

	Granularity			
	30 min	**6 h**	**12 h**	**24 h**
Time window 1	0.88	1.00	1.00	1.00
Time window 2	1.00	1.00	1.00	1.00
Time window 3	0.87	0.82	0.91	1.00

5.3. Analysis of Investors' Investment Decisions Based on the Time-Varying Network Structure

5.3.1. The Changes in Crypto Network Structure during Times of Crisis

To observe the growth of the network structure over time, we use *Degree Assortativity Coefficient* [122] and *Average Betweenness Centrality* [3]. However, these metrics fail to tell us the similarity between two networks. Thus, to statistically compare the topological change between two networks, we use three more metrics, including *v-measure*, *Degree centrality* [26] and *Eigenvalue method* [123,124].

Table 6 shows results of *Betweenness Centrality* and *Degree Assortativity*. Immediately, we can see that there is a huge change occurring in time window 2, which corresponds to the turbulent time caused by the pandemic on both metrics.

Regarding the *Betweenness Centrality*, this metric decreases from 0.15 in time window 1 to 0.05 in the next period before going back to its original value prior to the pandemic outbreak (time window 1). A reasonable explanation for this movement is that the network structure of the cryptocurency market during normal times appears to have a dispersive tendency with the whole network divided into multiple small-size groups such that each group share common characteristics. However, during COVID-19, these synchronize to form a big group. Thus, the number of groups decreases while the size of each group increases. This might be a consequence of an increase in the connectedness of cryptocurrencies during the pandemic, as shown in many research papers [11,27,45]. In the recovery time, however, the network started to divide into smaller parts again, perhaps because the cryptocurrency market overcame the most connected period and started to go back to its normal behavior.

The *Degree Assortativity* results strongly support those of the *Betweenness Centrality*. In particular, a negative value shows that high-degree nodes are more likely to link to low-degree nodes, which means that each group in the network has one node acting as a central node connecting to the rest. While the values in time window 1 and 3 are approximately the same, time window 2 shows a decline by nearly 50 percent. This indicates that the number of connections between high-degree nodes and low-degree nodes increases, i.e., the network forms big groups with a large number of leaf nodes in each group.

We notice that this time-varying structure is similar to what have been shown in works of Drozdz et al. [20,21], who stated that the market has a distributed-network topology or a hierarchical-network topology in which no node dominates the network during normal times. However, it becomes more centralized during the pandemic and started to distribute as this turbulent time is gone. More recently, another work proposed by Nie also confirmed the same result [22].

Table 7 shows results of the three similarity metrics for different time periods: *normal time* (time window 1), *downtime* (time window 2) and *recovery time* (time window 3). Each values shows the similarity between two time windows. For *v-measure*, the higher the value is, the more alike two networks are. On the other hand, for the remaining values, a lower value indicates that two networks are more similar.

The differences between time window 2 and the other two time windows are very clear. In particular, the *v-measure* between time window 1 and 3 is 0.32, meaning that communities found in time window 3 hold roughly one third of characteristics from time window 1's communities. By contrast, *v-measure* values between time window 1 and 2 as well as between time window 2 and 3 are negligible, standing at 0.04 and 0.02, respectively. Additionally, for the topological structures of MSTs, the other two metrics also show the same principle since time window 1 and 3 share common characteristics and the similarity degree of other cases are nearly zero. Remarkably, the *Eigenvalue method* shows a significant divergence of time window 2 with others, as shown in Table 7.

The severe pandemic and the global downturn of March 2020 together seem to have actually changed the way cryptocurrencies interact with each other. The changes of these interactions have created new communities and broken down old ones, i.e., some cryptocurrencies become closer to each other while others moved further away from each other due to the COVID-19 pandemic and the economic recession. Eventually, the topological structure during this turbulent time shows completely different patterns compared to the periods when the global market is stable. Furthermore, we noticed that the community structure started to recover back to its pre-COVID-19 levels after June 2020, which coincides with the time the global economy recovered and the COVID-19 pandemic had less impact. During this time, some characteristics of the network structure reappeared that are similar to the structure during the normal time (it is obvious that these structures are not fully similar because they change over time, as proven in previous sections and, in addition, after June 2020, the global economy started to recover, but not as well as in the past, and the pandemic still had an impact on the economy worldwide to some extent). This is why the *v-measure* between time window 1 and 3 increased significantly and the corre-

sponding differences measured by *Degree centrality* and *Eigenvalue method* are very small. The community structures for the three time windows are shown in Figures 3a, 4a and 5a.

Table 6. The growth of network structures over time measured by Betweenness Centrality and Degree Assortativity.

Metrics	Time Window 1	Time Window 2	Time Window 3
Betweenness centrality	0.15	0.05	0.16
Degree Assortativity	−0.49	−0.72	−0.51

Table 7. Similarity in network structures between different phases of the cryptocurrency market measured by three metrics. A higher value of *v-measure* indicates a greater similarity between two structures, whereas, higher values of *degree centrality* and *eigenvalue method* indicate more dissimilarity between two structures.

Time Window / Metrics	1 vs. 2	1 vs. 3	2 vs. 3
Degree centrality	0.5	0.09	0.42
Eigenvalue method	844.45	4.59	759.16
v-measure	0.04	0.32	0.02

5.3.2. Learning the Investment Decision of Crypto Traders Based on Ranking Distribution

The ranking of a cryptocurrency is measured by its market capitalization (a multiplication between the number of coins in circulation and the current market price of a single coin). We obtain cryptocurrencies' ranking on the https://coinmarketcap.com website (accessed on 15 August 2022).

We use this characteristic of cryptocurrencies to examine how they are distributed in each community of the cryptocurrency network. More importantly, we will have a look at the way cryptocurrencies form groups during different phases of the global economy by observing the distribution of ranking in each group between different periods of time.

Table 8 summarizes the results of community detection using the Louvain method. For each period of time, the found communities are listed with a set of cryptocurrencies and corresponding rankings belonging to each of them. We found that during the normal time, there is a mix between high-ranking and low-ranking cryptocurrencies in each community. For example, group 6 has a size of 7 including top-ranking cryptocurrencies such as BTC, ETH and BCH, while also having very low-ranking ones such as MAID and ICX. We pay more attention to communities found in the downturn time. At this phase, we recognized that the community formation of these cryptocurrencies seems to be dramatically different from the previous period. In particular, there are only two communities found during this period, while the other has six. More importantly, there seems to be a separation between high-ranking and low-ranking cryptocurrencies, because the majority of top-ranking cryptocurrencies belong to one group while the majority of low-ranking cryptocurrencies are in the other. Additionally, by comparing these results with the period of recovery, we noticed that this period shares common characteristics with both normal time and downturn time. Specifically, after the downturn time, cryptocurrencies started to separate from each other; this can be seen by looking at the number of communities during this time. There was an increase from 2 to 6, which is equal to the normal time case. While the majority of communities show a mix between high- and low-ranking cryptocurrencies, there are two communities that are similar to the downturn time: group 4 with all high-ranking cryptocurrencies and group 5 with all low-ranking cryptocurrencies.

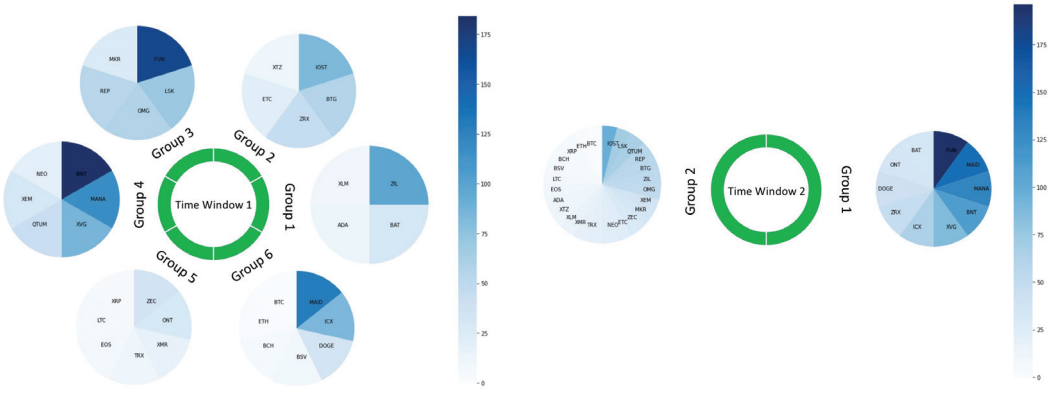
Table 8. Distributions of rankings in each community during different phases of the financial market: normal time, downturn time and recovery time. The rankings are sorted in ascending order. Bold values are minimum and maximum ranks in each period.

	Group	Cryptocurrencies	Rankings
Normal time	1	ADA, XLM, BAT, ZIL	10, 13, 32, 99
	2	BTG, IOST, XTZ, ZRX, ETC	12, 21, 45, 57, 83
	3	LSK, OMG, REP, FUN, MKR	26,54, 58, 70, 168
	4	NEO, MANA, BNT, XVG, XEM, QTUM	19, 31, 41, 86, 117, **184**
	5	ONT, ZEC, XMR, XRP, EOS, TRX, LTC	3, 6, 7, 11, 16, 29, 35
	6	ICX, MAID, DOGE, BTC, BSV, ETH, BCH	**1**, 2, 5, 9, 34, 84, 130
Downturn time	1	DOGE, ICX, BNT, MANA, ZRX, FUN, MAID, BAT, XVG, ONT	32, 33, 40, 45, 60, 81, 105, 124, 139, **196**
	2	ADA, BCH, BSV, BTC, BTG, EOS, ETH, ETC, IOST, LSK, LTC, MKR, NEO, OMG, QTUM, REP, TRX, XEM, XLM, XMR, XRP, XTZ, ZEC, ZIL	**1**, 2, 4, 5, 6, 7, 9, 11, 12, 15, 17, 18, 21, 22, 27, 30, 34, 48, 51, 53, 54, 62, 65, 91
Recovery time	1	BTG, MANA, BAT, ZEC	56, 62, 67, 107
	2	ONT, QTUM, EOS, BSV, MKR	24, 31, 53, 75, 88
	3	XVG ,ZIL, XEM, MAID, BTC, ETH	**1**, 2, 38, 48, 109, 136
	4	ADA, DOGE, XRP, BCH, XLM, LTC	6, 7, 9, 15, 16, 20
	5	OMG, BNT, IOST, REP, ICX, LSK	68, 78, 85, 100, 101, **140**
	6	ETC, ZRX, TRX, NEO, XMR, FUN, XTZ	17, 27, 33, 35, 64, 76, 129

Figure 6 shows the distribution of cryptocurrencies' rankings in three different phases of time. We use this visualization to show readers the changes of ranking distributions in a clearer and easier manner. Each community is represented by a circular shape while the rankings of cryptocurrencies are represented by the intensity of the blue color, i.e., the darker the blue, the lower the cryptocurrency's rank. Figure 6b shows that the circular shape of group 1 is clearly darker than that of group 2. On the other hand, there is a combination of both bright and dark blue in the majority of cases in two remaining sub-figures. Notably, Groups 3 and 5 in Figure 6c show a clear difference from the rest.

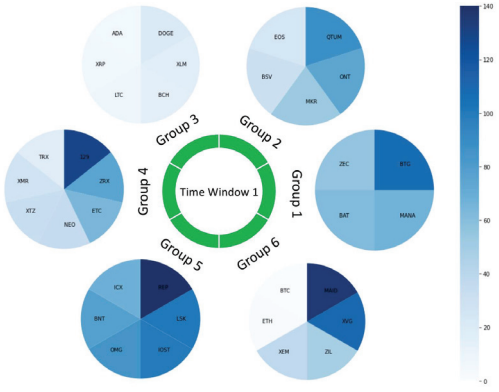
When it comes to these results, investors' investment decisions can be considered as potential explanations for the time-varying community structure. During normal times, i.e., when the financial market is stable and there is no major event occurring that impacts wider society, investors show a non-herding behaviour. That is, their decision for investing in a cryptocurrency is based on their own market analysis and is not influenced by other investors' choice. This might push up the vibrancy of the cryptocurrency market where a large number of coins with both high and low rankings are traded. As a result, there is a diversification in terms of rankings in each community. Empirically, it is found that there was no herding behavior before the pandemic. In particular, Larisa et al. in [125] used hourly price time series of multiple exchanges such as Binance, Bitbay, BitFinex, Coinbase and major cryptocurrencies including BTC, LTC and ETH to find the existence of herding before the start of COVID-19. Based on the *Cross Sectional Absolute Deviation* model, they found that the herding behavior was free during this time. By contrast, during a turbulent time, investors are panicked by the fluctuations of cryptocurrencies' price as well as being bombarded by bad news that strongly affect their investment. Different studies have been carried out to investigate the investors' behavior since the onset of the COVID-19 outbreak. Generally speaking, these reached the same conclusions: that the pandemic actually increased herding behavior in the cryptocurrency market. In [126], the authors used 43 cryptocurrencies with large market capitalization between 2013 and 2020; they

found that investors in the cryptocurrency market follow the consensus and the impact of coronavirus media coverage is significant on the herding behavior. In particular, news related to the coronavirus increases fear and affects the behavior of investors reducing appetite for risk. Consequently, investors disregard their private information and follow others' investment decisions. However, the impact of media is reduced when the market returns to a normal phase. This is in line with different studies that use different datasets and time periods [125,127,128]. More importantly, the way investors show herding behavior is that they tended to invest in the major and most-tradable cryptocurrencies [27]. This can be explained by the fact that high-ranking cryptocurrencies are more mature so they are more stable than the rest and are more likely to retain value under the uncertainty of the global financial market, causing a bias from investors [129]. Consequently, major cryptocurrencies were seen to increase in terms of trading volume and act as a store of value during the turbulence times [130]. In other words, there was a risk aversion occurring after the pandemic outburst as described in [72]. Eventually, all high-ranking cryptocurrencies belonged to one group.



(**a**) Time window 1

(**b**) Time window 2



(**c**) Time window 3

Figure 6. Cryptocurrency's rankings distributions in three different phases of time. Each community is represented by a circular shape while the rankings of cryptocurrencies in this community are given by the blue color intensity, i.e., the darker the blue, the lower the cryptocurrency's rank.

When it comes to low-ranking cryptocurrencies, we notice that cryptocurrencies with the lowest rankings in our dataset belong to another group. This might be because they receive the same treatment from investors during the downturn time, so they have the

same trend. One possible reason for this is that low-ranking cryptocurrencies are less likely to be used as an investment option during the downturn time because they have negligible value and bring more risk to investors. Instead, they are mainly used for other purposes, such as paying transaction fees, as currency for a smart contract or simply a token on a cryptocurrency platform used to access applications [27]. This seems reasonable as the pandemic stopped in-person interactions. Hence, they had to complete all work remotely. In this case, cryptocurrencies and blockchain technology are extremely useful since they thrive under the proposed online environment to resume working activities worldwide and also bring benefits to users. Being used for the same purpose causes a similarity between these cryptocurrencies.

All findings that we have shown earlier help us to explain the community structure in time window 3, which corresponds to the recovery period. During this time, the concerns about the pandemic started to decrease, meaning that not only cryptocurrency but also other traditional assets recovered with investors' newfound positive attitude bringing them back to normal trading. Crypto traders started to diversify their portfolio by investing in different low and high market-capitalized assets and making their own decisions [126]. However, one remarkable phenomenon that is worth taking into consideration is] risk-taking behavior. A piece of research implemented by Christoph et al. [131] used 100 return time series of risky stocks to conduct a survey related to the investment behavior of professional market traders. The responses of more than 800 participants revealed that a number of investors underestimate risk after prolonged exposure to high risk, as they become accustomed to the uncertainty of the economy. Thus, they go back to investing in risky assets or even engage in more risk-taking to gain more profits. This tendency explains the similarity in the community structures between time windows 1 and 3. However, as we can see, there exists one group with high-ranking cryptocurrencies and one group with low-ranking ones as a result of the risk aversion of a portion of investors after the great shock caused by the pandemic.

6. Limitations and Future Works

6.1. Limitations

Although the tick-by-tick dataset used in this study is large, which strengthens the results of the experiments, the number of cryptocurrencies should ideally be higher so that we can draw firmer conclusions regarding the cryptocurrency market (e.g., whether the results generalise for large-cap and small-cap crypto assets). This will be the subject of future work. Secondly, while 30 min granularity has been found to suffice for our calculations, it would be better if we could use a lower level, say 15 min or even finer. Unfortunately, some cryptocurrencies are not traded regularly causing a lot of missing values at these timescales. This will also be the subject of future work.

There is also a concern with respect to the use of Pearson correlation for clustering problems. In particular, although this correlation metric has been applied widely in the existing literature and proposed various findings in the financial markets [2,21,22], it is sensitive to outliers [58] and cannot capture non-linear relationships that might cause misleading results [25]. Consequently, this adversely affects the clustering results. Indeed, these issues are also observed in other correlation metrics such as Spearman and Kendall [58]. Furthermore, we noticed that the results of clustering vary significantly by using different correlation measuring methods. Thus, it is necessary to deeply investigate different methods for a specific research task and analyze the results from each of these methods. Indeed, the creation of new approaches for calculating correlation coefficients that overcome the current limitations needs to receive more attention.

6.2. Future Works

Understanding how cryptocurrencies are correlated with each other sheds light on portfolio optimization. Based on the outcome of this study, we can take one step further by constructing and comparing the portfolio optimizations at different phases of the market,

i.e., during bear and bull market periods. Therefore, the unique characteristics of an optimized portfolio at different market phases can, in theory, be learned and analysed. Secondly, we have noted that different network structures can be observed for a number of exchanges. Thereby, a comparison between them can be made. Another future plan which is worth taking into consideration is to observe the correlation using different techniques. For instance, we are aiming to use mutual information, which is successfully applied in [25], to estimate the correlation between two cryptocurrencies. This method can overcome obstacles from popular linear and non-linear methods since it can measure the non-linear correlation while allowing the existence of non-monotonic relationships. Lastly, we have noticed that the network structure of low-frequency data behaves differently to that of high-frequency data. We remark that we can expect to learn the long-term characteristics of cryptocurrencies based on this structure which could be potentially beneficial for investors who choose to make a long-term investment decision.

7. Conclusions

This research aims at answering three questions related to cross-correlations in the cryptocurrency market: Firstly, how do noise and trends in cryptocurrencies influence their cross-correlations and then the corresponding network structure? Secondly, what level of granularity should we use? Lastly, is the dramatic change in the cryptocurrency network structure during the pandemic caused by investors' investment strategy? We firstly analyze the effect of noise and trend in cryptocurrencies on their cross-correlations and then remove these factors thanks to Random Matrix Theory and Market Component. Four sub-datasets with different levels of granularity including 30 min, 6 h, 12 h and 1 day are created from the original tick-by-tick data to examine the importance of choosing the right frequency resolution. Then, we use MST to construct a correlation-based network and detect different potential communities by using Louvain and Girvan–Newman algorithms. We found that the correlations between cryptocurrencies are mainly caused by noise and trend effects, which might lead to a big problem for the traders' investment strategy because investors might be fooled by looking at the counterfeit relationship. It is necessary to analyze and explore real interactions between cryptocurrencies so that the evolution of the cryptocurrency market can be learned properly and thus investors can choose a good strategy for their investment. Moreover, the frequency resolution of our data plays an important role in the performance of correlation matrix and also community detection. Specifically, the finer the data, the more precise the community structure. Thus, we use a 30 min dataset, which is the finest available timescale in this study. The dramatic change in the community structures between bearish and bullish markets reveals a change in the investment decisions of investors. In particular, investors makes their own investment decisions based on their personal market analysis and experience during normal times. Eventually, this causes a diversification in the cryptocurrencies chosen to invest in, since not only high- but also low-ranking cryptocurrencies are added in the portfolios. On the other hand, investors tend to only trade cryptocurrencies with high market capitalization during turbulent times, while smaller cryptocurrencies are mainly used for other purposes, such as transaction fees, smart contracts tokens or simply used to run a digital platform.

Author Contributions: A.P.N.N.: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing—original draft, Visualization. T.T.M.: Methodology, Writing—review and editing, Supervision. M.B.: Conceptualization, Methodology, Data curation, Writing—review and editing, Supervision. M.C.: Conceptualization, Methodology, Data curation, Writing—review and editing, Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Science Foundation Ireland Centre for Research Training in Artificial Intelligence under grant number 18/CRT/6223.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors Martin Crane, Marija Bezbradica, Tai Tan Mai, wish to acknowledge the support, in part, from the Science Foundation Ireland under Grant Agreement No. 13/RC/2106_P2 at the ADAPT SFI Research Centre at DCU. ADAPT, the SFI Research Centre for AI-Driven Digital Content Technology, is funded by the Science Foundation Ireland through the SFI Research Centres Programme. The author An P. N. Nguyen wishes to acknowledge the support from Dublin City University's Research Committee and research grants from Science Foundation Ireland Centre for Research Training in Artificial Intelligence under grant number 18/CRT/6223.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, Y.; Jiang, X.F.; Tian, Y.; Li, S.P.; Zheng, B. Portfolio optimization based on network topology. *Physica A* **2019**, *515*, 671–681. [CrossRef]
2. Zhao, L.; Li, W.; Cai, X. Structure and dynamics of stock market in times of crisis. *Phys. Lett. A* **2016**, *380*, 654–666. [CrossRef]
3. Balcilar, M.; Ozdemir, H.; Agan, B. Effects of COVID-19 on cryptocurrency and emerging market connectedness: Empirical evidence from quantile, frequency, and lasso networks. *Physica A* **2022**, *604*, 127885. [CrossRef]
4. Chaudhari, H.; Crane, M. Cross-correlation dynamics and community structures of cryptocurrencies. *J. Comput. Sci.* **2020**, *44*, 101130. [CrossRef]
5. Wątorek, M.; Drożdż, S.; Kwapień, J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* **2021**, *901*, 1–82. [CrossRef]
6. Nie, C.X. A network-based method for detecting critical events of correlation dynamics in financial markets. *Europhys. Lett.* **2020**, *131*, 50001. [CrossRef]
7. Papadimitriou, T.; Gogas, P.; Gkatzoglou, F. The evolution of the cryptocurrencies market: A complex networks approach. *J. Comput. Appl. Math.* **2020**, *376*, 112831. [CrossRef]
8. Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–4.
9. Conlon, T.; Ruskin, H.J.; Crane, M. Cross-correlation dynamics in financial time series. *Physica A* **2009**, *388*, 705–714. [CrossRef]
10. Matos, J.; Gama, S.; Ruskin, H.; Sharkasi, A.; Crane, M. Correlation of worldwide markets' entropies. *Centro* **2006**. Available online: https://citeseerx.ist.psu.edu/messages/downloadsexceeded.html (accessed on 18 August 2022)
11. Goodell, J.W.; Goutte, S. Diversifying equity with cryptocurrencies during COVID-19. *Int. Rev. Financ. Anal.* **2021**, *76*, 101781. [CrossRef]
12. Mallinger-Dogan, M.; Szigety, M.C. Higher-Frequency Analysis of Low-Frequency Data. *J. Portf. Manag.* **2014**, *41*, 121–138. [CrossRef]
13. Brauneis, A.; Mestel, R. Price discovery of cryptocurrencies: Bitcoin and beyond. *Econ. Lett.* **2018**, *165*, 58–61. [CrossRef]
14. Caferra, R.; Vidal-Tomás, D. Who raised from the abyss? A comparison between cryptocurrency and stock market dynamics during the COVID-19 pandemic. *Financ. Res. Lett.* **2021**, *43*, 101954. [CrossRef]
15. Baur, D.G.; Dimpfl, T. The volatility of Bitcoin and its role as a medium of exchange and a store of value. *Empir. Econ.* **2021**, *61*, 2663–2683. [CrossRef]
16. Barua, S. Understanding Coronanomics: The Economic Implications of the Coronavirus (COVID-19) Pandemic. 2020. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3566477 (accessed on 17 July 2022).
17. Padhan, R.; Prabheesh, K. The economics of COVID-19 pandemic: A survey. *Econ. Anal. Policy* **2021**, *70*, 220–237. [CrossRef]
18. Yousfi, M.; Zaied, Y.B.; Cheikh, N.B.; Lahouel, B.B.; Bouzgarrou, H. Effects of the COVID-19 pandemic on the US stock market and uncertainty: A comparative assessment between the first and second waves. *Technol. Forecast. Soc. Chang.* **2021**, *167*, 120710. [CrossRef]
19. Lahmiri, S.; Bekiros, S. The impact of COVID-19 pandemic upon stability and sequential irregularity of equity and cryptocurrency markets. *Chaos Solit. Fractals* **2020**, *138*, 109936. [CrossRef]
20. Kwapień, J.; Wątorek, M.; Drożdż, S. Cryptocurrency Market Consolidation in 2020 & 2021. *Entropy* **2021**, *23*, 1674. [CrossRef]
21. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Stanisz, T.; Wątorek, M. Complexity in economic and social systems: Cryptocurrency market at around COVID-19. *Entropy* **2020**, *22*, 1043. [CrossRef]
22. Nie, C.X. Analysis of critical events in the correlation dynamics of cryptocurrency market. *Phys. A Stat. Mech. Its Appl.* **2022**, *586*, 126462. [CrossRef]
23. Miceli, M.A.; Susinno, G. Ultrametricity in fund of funds diversification. *Physica A* **2004**, *344*, 95–99. [CrossRef]
24. Onnela, J.P.; Chakraborti, A.; Kaski, K.; Kertesz, J. Dynamic asset trees and Black Monday. *Physica A* **2003**, *324*, 247–252. [CrossRef]
25. Liu, H.; Zou, J.; Ravishanker, N. Clustering high-frequency financial time series based on information theory. *Appl. Stoch. Model. Bus. Ind.* **2022**, *38*, 4–26. [CrossRef]
26. Durcheva, M.; Tsankov, P. Analysis of similarities between stock and cryptocurrency series by using graphs and spanning trees. In *AIP Conference Proceedings*; AIP Publishing LLC: Melville, NY, USA, 2019; Volume 2172, p. 090004.
27. Katsiampa, P.; Yarovaya, L.; Zięba, D. High-Frequency connectedness between bitcoin and other top-traded crypto assets during the COVID-19 crisis. *J. Int. Financ. Mark. Inst Money* **2022**, *79*, 101578. [CrossRef]

28. Francés, C.J.; Carles, P.; Arellano, D.J. The cryptocurrency market: A network analysis. *Esic Mark. Econ. Bus. J.* **2018**, *49*, 569–583. [CrossRef]
29. Briola, A.; Aste, T. Dependency structures in cryptocurrency market from high to low frequency. *arXiv* **2022**, arXiv:2206.03386.
30. Gavin, J.; Crane, M. Community Detection in Cryptocurrencies with Potential Applications to Portfolio Diversification. *arXiv* **2021**, arXiv:2108.09756.
31. Yu, A.; Bünz, B. *Community Detection and Analysis in the Bitcoin Network, CS 224W Final Report*; Stanford University: Stanford, CA, USA, 2015.
32. Atiya, H.R.; Nawaf, H.N. Prediction of Link Weight of bitcoin Network by Leveraging the Community Structure. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristiol, UK, 2020; Volume 928, p. 032080.
33. Jin, D.; Yu, Z.; Jiao, P.; Pan, S.; He, D.; Wu, J.; Yu, P.; Zhang, W. A survey of community detection approaches: From statistical modeling to deep learning. *IEEE Trans. Knowl. Data. Eng* **2021**. [CrossRef]
34. Cao, K.H.; Li, Q.; Liu, Y.; Woo, C.K. COVID-19's adverse effects on a stock market index. *Appl. Econ. Lett.* **2021**, *28*, 1157–1161. [CrossRef]
35. Ali, M.; Alam, N.; Rizvi, S.A.R. Coronavirus (COVID-19)—An epidemic or pandemic for financial markets. *J. Behav. Exp. Financ.* **2020**, *27*, 100341. [CrossRef]
36. De Luca, G.; Loperfido, N. A skew-in-mean GARCH model. In *Skew-Elliptical Distributions and Their Applications: A Journey beyond Normality*; Genton, M.G., Ed.; Chapman and Hall/CRC: London, UK, 2004; pp. 233–252.
37. Karamti, C.; Belhassine, O. COVID-19 pandemic waves and global financial markets: Evidence from wavelet coherence analysis. *Financ. Res. Lett.* **2022**, *45*, 102136. [CrossRef] [PubMed]
38. Zhang, W.; Hamori, S. Crude oil market and stock markets during the COVID-19 pandemic: Evidence from the US, Japan, and Germany. *Int. Rev. Financ. Anal.* **2021**, *74*, 101702. [CrossRef]
39. Adekoya, O.B.; Oliyide, J.A. How COVID-19 drives connectedness among commodity and financial markets: Evidence from TVP-VAR and causality-in-quantiles techniques. *Resour. Policy* **2021**, *70*, 101898. [CrossRef] [PubMed]
40. Chen, C.; Liu, L.; Zhao, N. Fear sentiment, uncertainty, and bitcoin price dynamics: The case of COVID-19. *Emerg. Mark. Financ. Trade* **2020**, *56*, 2298–2309. [CrossRef]
41. Baig, A.S.; Butt, H.A.; Haroon, O.; Rizvi, S.A.R. Deaths, panic, lockdowns and US equity markets: The case of COVID-19 pandemic. *Financ. Res. Lett* **2021**, *38*, 101701. [CrossRef]
42. Sapkota, N. News-based sentiment and bitcoin volatility. *Int. Rev. Financ. Anal.* **2022**, *82*, 102183. [CrossRef]
43. Bianchi, D. Cryptocurrencies as an asset class? An empirical assessment. *J. Altern. Investments* **2020**, *23*, 162–179. [CrossRef]
44. Anamika, A.; Subramaniam, S. Do news headlines matter in the cryptocurrency market? *Appl. Econ.* **2022**, 1–17. [CrossRef]
45. Vidal-Tomás, D. Transitions in the cryptocurrency market during the COVID-19 pandemic: A network analysis. *Financ. Res. Lett.* **2021**, *43*, 101981. [CrossRef]
46. Assaf, A.; Bhandari, A.; Charif, H.; Demir, E. Multivariate long memory structure in the cryptocurrency market: The impact of COVID-19. *Int. Rev. Financ. Anal.* **2022**, *82*, 102132. [CrossRef]
47. Masconi, K.L.; Matsha, T.E.; Erasmus, R.T.; Kengne, A.P. Effects of different missing data imputation techniques on the performance of undiagnosed diabetes risk prediction models in a mixed-ancestry population of South Africa. *PLoS ONE* **2015**, *10*, e0139210. [CrossRef] [PubMed]
48. Van der Heijden, G.J.; Donders, A.R.T.; Stijnen, T.; Moons, K.G. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: A clinical example. *J. Clin. Epidemiol.* **2006**, *59*, 1102–1109. [CrossRef] [PubMed]
49. Genolini, C.; Jacqmin-Gadda, H. Copy mean: A new method to impute intermittent missing values in longitudinal studies. *Open J. Stat.* **2013**, *3*, 26. [CrossRef]
50. Rydberg, T.H. Realistic statistical modelling of financial data. *Int. Stat. Rev.* **2000**, *68*, 233–258. [CrossRef]
51. Amien, I.; Rajaratnam, K.; Kruger, R. Inference of aggregational gaussianity in asset returns exhibiting a paretian-distribution. *Procedia Econ. Financ.* **2015**, *25*, 400–407. [CrossRef]
52. Kratz, M.; Resnick, S.I. The QQ-estimator and heavy tails. *Stoch. Model.* **1996**, *12*, 699–724. [CrossRef]
53. Abdi, H.; Molin, P. Lilliefors/Van Soest's test of normality. In *Encyclopedia of Measurement and Statistics*; SAGE Publications: Thousand Oaks, OK, USA, 2007; pp. 540–544.
54. Takaishi, T. Time-varying properties of asymmetric volatility and multifractality in Bitcoin. *PLoS ONE* **2021**, *16*, e0246209. [CrossRef]
55. Kulyatin, I. *Stylized Facts for Cryptocurrencies: A Sectoral Analysis*; Medium: New York, NY, USA, 2015.
56. Tsay, R.S. *Analysis of Financial Time Series*; John Wiley & Sons: Hoboken, NJ, USA, 2005.
57. Schäfer, R.; Guhr, T. Local normalization: Uncovering correlations in non-stationary financial time series. *Physica A* **2010**, *389*, 3856–3865. [CrossRef]
58. Tjøstheim, D.; Otneim, H.; Støve, B. Statistical Dependence: Beyond Pearson's ρ. *Stat. Sci.* **2022**, *37*, 90–109. [CrossRef]
59. Wissler, C. The Spearman correlation formula. *Science* **1905**, *22*, 309–311. [CrossRef]
60. Lieberson, S. Limitations in the application of non-parametric coefficients of correlation. *Am. Sociol. Rev.* **1964**, *29*, 744–746. [CrossRef]

61. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L.N.; Stanley, H.E. A random matrix theory approach to financial cross-correlations. *Physica A* **2000**, *287*, 374–382. [CrossRef]
62. Cacciapuoti, C.; Maltsev, A.; Schlein, B. Local Marchenko-Pastur law at the hard edge of sample covariance matrices. *J. Math. Phys.* **2013**, *54*, 043302. [CrossRef]
63. Laloux, L.; Cizeau, P.; Potters, M.; Bouchaud, J.P. Random matrix theory and financial correlations. *Int. J. Theor. Appl. Financ.* **2000**, *3*, 391–397. [CrossRef]
64. Mai, T.; Martin, C.; Marija, B. Student behaviours in using learning resources in higher education: How do behaviours reflect success in programming education? In Proceedings of the 7th International Conference on Higher Education Advances, Torino, Italy, 22–23 June 2021; pp. 47–55.
65. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L.A.N.; Guhr, T.; Stanley, H.E. Random matrix approach to cross correlations in financial data. *Phys. Rev. E* **2002**, *65*, 066126. [CrossRef]
66. Dimpfl, T.; Peter, F.J. Nothing but noise? Price discovery across cryptocurrency exchanges. *J. Financ. Mark.* **2021**, *54*, 100584. [CrossRef]
67. Sarr, A.; Lybek, T. Measuring Liquidity in Financial Markets. 2002. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=880932 (accessed on 27 July 2022).
68. Li, T.; Shin, D.; Wang, B. Cryptocurrency Pump-and-Dump Schemes. 2021. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3267041 (accessed on 5 August 2022).
69. Auer, R.; Claessens, S. Regulating cryptocurrencies: Assessing market reactions. *BIS Q. Rev. Sept.* **2018**. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3288097 (accessed on 8 August 2022).
70. Mai, T.T.; Bezbradica, M.; Crane, M. Learning behaviours data in programming education: Community analysis and outcome prediction with cleaned data. *Future Gener. Comput. Syst.* **2022**, *127*, 42–55. [CrossRef]
71. Zhang, S.; Lee, J.H. Analysis of the main consensus protocols of blockchain. *ICT Express* **2020**, *6*, 93–97. [CrossRef]
72. da Gama Silva, P.V.J.; Klotzle, M.C.; Pinto, A.C.F.; Gomes, L.L. Herding behavior and contagion in the cryptocurrency market. *J. Behav. Exp. Financ.* **2019**, *22*, 41–50. [CrossRef]
73. Burda, Z.; Jarosz, A. Cleaning large-dimensional covariance matrices for correlated samples. *Phys. Rev. E* **2022**, *105*, 034136. [CrossRef] [PubMed]
74. Bouchaud, J.P.; Potters, M. Financial applications of random matrix theory: A short review. *arXiv* **2009**, arXiv:0910.1205.
75. Ledoit, O.; Wolf, M. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Ann. Stat.* **2012**, *40*, 1024–1060. [CrossRef]
76. Bun, J.; Allez, R.; Bouchaud, J.P.; Potters, M. Rotational invariant estimator for general noisy matrices. *IEEE Trans. Inf. Theory* **2016**, *62*, 7475–7490. [CrossRef]
77. Ledoit, O.; Wolf, M. A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.* **2004**, *88*, 365–411. [CrossRef]
78. Bun, J.; Knowles, A. An optimal rotational invariant estimator for general covariance matrices: The outliers. *Preprint* **2018**. Available online: https://www.researchgate.net/profile/Joel-Bun/publication/323255675_An_Optimal_Rotational_Invariant_Estimator_for_General_Covariance_Matrices_the_outliers/links/5a89e7fba6fdcc6b1a424f88/An-Optimal-Rotational-Invariant-Estimator-for-General-Covariance-Matrices-the-outliers.pdf (accessed on 8 August 2022).
79. Shcherbakov, M.V.; Brebels, A.; Shcherbakova, N.L.; Tyukov, A.P.; Janovsky, T.A.; Kamaev, V.A. A survey of forecast error measures. *World Appl. Sci. J.* **2013**, *24*, 171–176.
80. Markowitz, R. Cleaning Correlation Matrices. 2016. Available online: https://www.cfm.fr/assets/ResearchPapers/2016-Cleaning-Correlation-Matrices.pdf (accessed on 8 August 2022).
81. Conlon, T.; Ruskin, H.J.; Crane, M. Random matrix theory and fund of funds portfolio optimisation. *Physica A* **2007**, *382*, 565–576. [CrossRef]
82. Yang, L.; McKay, M.R.; Couillet, R. High-dimensional MVDR beamforming: Optimized solutions based on spiked random matrix models. *IEEE Trans. Signal Process* **2018**, *66*, 1933–1947. [CrossRef]
83. Heimo, T.; Kaski, K.; Saramäki, J. Maximal spanning trees, asset graphs and random matrix denoising in the analysis of dynamics of financial networks. *Physica A* **2009**, *388*, 145–156. [CrossRef]
84. de Prado, M.M.L. *Machine Learning for Asset Managers*; Cambridge University Press: Cambridge, UK, 2020.
85. O'Searcoid, M. *Metric Spaces*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006.
86. Graham, R.L.; Hell, P. On the history of the minimum spanning tree problem. *IEEE Ann. Hist. Comput.* **1985**, *7*, 43–57. [CrossRef]
87. Ghosh, B.; Motagh, M.; Haghighi, M.H.; Vassileva, M.S.; Walter, T.R.; Maghsudi, S. Automatic detection of volcanic unrest using blind source separation with a minimum spanning tree based stability analysis. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7771–7787. [CrossRef]
88. Denkowska, A.; Wanat, S. Linkages and systemic risk in the European insurance sector. New evidence based on Minimum Spanning Trees. *Risk Manag.* **2022**, *24*, 123–136. [CrossRef]
89. Coelho, R.; Gilmore, C.G.; Lucey, B.; Richmond, P.; Hutzler, S. The evolution of interdependence in world equity markets—Evidence from minimum spanning trees. *Physica A* **2007**, *376*, 455–466. [CrossRef]
90. Kazemilari, M.; Mohamadi, A.; Mardani, A.; Streimikis, J. Network topology of renewable energy companies: Minimal spanning tree and sub-dominant ultrametric for the American stock. *Technol. Econ. Dev. Econ.* **2019**, *25*, 168–187. [CrossRef]

91. Nguyen, Q.; Nguyen, N.; Nguyen, L. Dynamic topology and allometric scaling behavior on the Vietnamese stock market. *Physica A* **2019**, *514*, 235–243. [CrossRef]
92. Huang, C.; Zhao, X.; Su, R.; Yang, X.; Yang, X. Dynamic network topology and market performance: A case of the Chinese stock market. *Int. J. Financ. Econ.* **2022**, *27*, 1962–1978. [CrossRef]
93. Giudici, P.; Pagnottoni, P.; Polinesi, G. Network models to enhance automated cryptocurrency portfolio management. *Front. Artif. Intell.* **2020**, *3*, 22. [CrossRef]
94. Kruskal, J.B. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Am. Math. Soc.* **1956**, *7*, 48–50. [CrossRef]
95. Huang, F.; Gao, P.; Wang, Y. Comparison of Prim and Kruskal on Shanghai and Shenzhen 300 Index hierarchical structure tree. In Proceedings of the 2009 International Conference on Web Information Systems and Mining, Shanghai, China, 7–8 November 2009; pp. 237–241.
96. Mantegna, R.N. Hierarchical structure in financial markets. *Eur. Phys. J. B* **1999**, *11*, 193–197. [CrossRef]
97. Zhao, B.; Yang, W.; Wen, J.; Zhang, W. The Financial Market in China under the COVID-19. *Emerg. Mark. Financ. Trade* **2022**, 1–13. [CrossRef]
98. Brida, J.G.; Risso, W.A. Hierarchical structure of the German stock market. *Expert Syst. Appl.* **2010**, *37*, 3846–3852. [CrossRef]
99. MacMahon, M.; Garlaschelli, D. Community detection for correlation matrices. *arXiv* **2013**, arXiv:1311.1924.
100. Chakrabarti, P.; Jawed, M.S.; Sarkhel, M. COVID-19 pandemic and global financial market interlinkages: A dynamic temporal network analysis. *Appl. Econ.* **2021**, *53*, 2930–2945. [CrossRef]
101. Bazzi, M.; Porter, M.A.; Williams, S.; McDonald, M.; Fenn, D.J.; Howison, S.D. Community detection in temporal multilayer networks, with an application to correlation networks. *Multiscale Model. Simul.* **2016**, *14*, 1–41. [CrossRef]
102. Anghinoni, L.; Vega-Oliveros, D.A.; Silva, T.C.; Zhao, L. Time series pattern identification by hierarchical community detection. *Eur. Phys. J. Spec. Top.* **2021**, *230*, 2775–2782. [CrossRef]
103. Raghavan, U.N.; Albert, R.; Kumara, S. Near linear time algorithm to detect community structures in large-scale networks. *Phys. Rev. E* **2007**, *76*, 036106. [CrossRef]
104. Pujol, J.M.; Erramilli, V.; Rodriguez, P. Divide and conquer: Partitioning online social networks. *arXiv* **2009**, arXiv:0905.4918.
105. Roma, G.; Herrera, P. Community structure in audio clip sharing. In Proceedings of the 2010 INCoS, Thessaloniki, Greece, 24–26 November 2010; pp. 200–205.
106. Wang, C.; Wang, F. GIS-automated delineation of hospital service areas in Florida: From Dartmouth method to network community detection methods. *Ann. GIS* **2022**, *28*, 93–109. [CrossRef]
107. Raeder, T.; Chawla, N.V. Market basket analysis with networks. *Soc. Netw. Anal. Min.* **2011**, *1*, 97–113. [CrossRef]
108. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008. [CrossRef]
109. Newman, M.E.; Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **2004**, *69*, 026113. [CrossRef] [PubMed]
110. Boyle, M.; Bellucco-Chatham, A. *2008 Recession: What the Great Recession Was and What Caused It*; Investopedia: New York, NY, USA, 2022.
111. Liang, J.; Li, L.; Zeng, D.; Zhao, Y. Correlation-based dynamics and systemic risk measures in the cryptocurrency market. In Proceedings of the IEEE 2018 ISI, Miami, FL, USA, 9–11 November 2018; pp. 43–48.
112. Spelta, A.; Araújo, T. The topology of cross-border exposures: Beyond the minimal spanning tree approach. *Physica A* **2012**, *391*, 5572–5583. [CrossRef]
113. Feng, W.; Wang, Y.; Zhang, Z. Can cryptocurrencies be a safe haven: A tail risk perspective analysis. *Appl. Econ.* **2018**, *50*, 4745–4762. [CrossRef]
114. Corbet, S.; Lucey, B.; Urquhart, A.; Yarovaya, L. Cryptocurrencies as a financial asset: A systematic analysis. *Int. Rev. Financ. Anal* **2019**, *62*, 182–199. [CrossRef]
115. Ammous, S. Can cryptocurrencies fulfil the functions of money? *Q. Rev. Econ. Financ.* **2018**, *70*, 38–51. [CrossRef]
116. Klein, T.; Thu, H.P.; Walther, T. Bitcoin is not the New Gold–A comparison of volatility, correlation, and portfolio performance. *Int. Rev. Financ. Anal.* **2018**, *59*, 105–116. [CrossRef]
117. Su, F.; Wang, X.; Yuan, Y. The intraday dynamics and intraday price discovery of bitcoin. *Res. Int. Bus. Financ.* **2022**, *60*, 101625. [CrossRef]
118. Megaritis, A.; Vlastakis, N.; Triantafyllou, A. Stock market volatility and jumps in times of uncertainty. *J. Int. Money Financ.* **2021**, *113*, 102355. [CrossRef]
119. Aslam, F.; Ferreira, P.; Mughal, K.S.; Bashir, B. Intraday volatility spillovers among European financial markets during COVID-19. *Int. J. Financ. Stud.* **2021**, *9*, 5. [CrossRef]
120. Jarosław, K.; Marcin, W.; Marija, B.; Martin, C.; Mai, T.; Stanisław, D. Analysis of inter-transaction time fluctuations in the cryptocurrency market. *Chaos Interdiscip. J. Nonlinear Sci.* **2022**, *32*, 083142.
121. Rosenberg, A.; Hirschberg, J. V-measure: A conditional entropy-based external cluster evaluation measure. In Proceedings of the 2007 EMNLP-CoNLL, Prague, Czech Republic, 28–30 June 2007; pp. 410–420.
122. Liang, J.; Li, L.; Zeng, D. Evolutionary dynamics of cryptocurrency transaction networks: An empirical study. *PLoS ONE* **2018**, *13*, e0202202.

123. Koutra, D.; Parikh, A.; Ramdas, A.; Xiang, J. Algorithms for Graph Similarity and Subgraph Matching. 2011. Volume 17. Available online: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.377.4291&rep=rep1&type=pdf (accessed on 8 August 2022).
124. Gera, R.; Alonso, L.; Crawford, B.; House, J.; Mendez-Bermudez, J.; Knuth, T.; Miller, R. Identifying network structure similarity using spectral graph theory. *Appl. Netw. Sci.* **2018**, *3*, 2. [CrossRef]
125. Yarovaya, L.; Matkovskyy, R.; Jalan, A. The effects of a'Black Swan'event (COVID-19) on herding behavior in cryptocurrency markets: Evidence from cryptocurrency USD, EUR, JPY and KRW Markets. *J. Int. Financ. Mark. Inst. Money* **2021**, *75*, 101321. [CrossRef]
126. Youssef, M.; Waked, S.S. Herding behavior in the cryptocurrency market during COVID-19 pandemic: The role of media coverage. *N. Am. J. Econ. Financ.* **2022**, *62*, 101752. [CrossRef]
127. Susana, D.; Kavisanmathi, J.; Sreejith, S. Does herding behaviour among traders increase during COVID 19 pandemic? Evidence from the cryptocurrency market. In Proceedings of the International Working Conference on Transfer and Diffusion of IT, Tiruchirappalli, India, 18–19 December 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 178–189.
128. Kaiser, L.; Stöckl, S. Cryptocurrencies: Herding and the transfer currency. *Financ. Res. Lett* **2020**, *33*, 101214. [CrossRef]
129. Mnif, E.; Salhi, B.; Mouakha, K.; Jarboui, A. Investor behavior and cryptocurrency market bubbles during the COVID-19 pandemic. *Rev. Behav. Financ.* 2022, *ahead-of-print*. [CrossRef]
130. Corbet, S.; Larkin, C.; Lucey, B.; Yarovaya, L. KODAKCoin: A blockchain revolution or exploiting a potential cryptocurrency bubble? *Appl. Econ. Lett* **2020**, *27*, 518–524. [CrossRef]
131. Huber, C.; Huber, J.; Kirchler, M. Market shocks and professionals' investment behavior—Evidence from the COVID-19 crash. *J. Bank. Financ.* **2021**, *133*, 106247. [CrossRef]

entropy

MDPI

Article

Collective Dynamics, Diversification and Optimal Portfolio Construction for Cryptocurrencies

Nick James [1,*,†] and Max Menzies [2,†]

[1] School of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia
[2] Beijing Institute of Mathematical Sciences and Applications, Tsinghua University, Beijing 101408, China; max.menzies@alumni.harvard.edu
* Correspondence: nick.james@unimelb.edu.au
† These authors contributed equally to this work.

Abstract: Since its conception, the cryptocurrency market has been frequently described as an immature market, characterized by significant swings in volatility and occasionally described as lacking rhyme or reason. There has been great speculation as to what role it plays in a diversified portfolio. For instance, is cryptocurrency exposure an inflationary hedge or a speculative investment that follows broad market sentiment with amplified beta? We have recently explored similar questions with a clear focus on the equity market. There, our research revealed several noteworthy dynamics such as an increase in the market's collective strength and uniformity during crises, greater diversification benefits across equity sectors (rather than within them), and the existence of a "best value" portfolio of equities. In essence, we can now contrast any potential signatures of maturity we identify in the cryptocurrency market and contrast these with the substantially larger, older and better-established equity market. This paper aims to investigate whether the cryptocurrency market has recently exhibited similar mathematical properties as the equity market. Instead of relying on traditional portfolio theory, which is grounded in the financial dynamics of equity securities, we adjust our experimental focus to capture the presumed behavioral purchasing patterns of retail cryptocurrency investors. Our focus is on collective dynamics and portfolio diversification in the cryptocurrency market, and examining whether previously established results in the equity market hold in the cryptocurrency market and to what extent. The results reveal nuanced signatures of maturity related to the equity market, including the fact that correlations collectively spike around exchange collapses, and identify an ideal portfolio size and spread across different groups of cryptocurrencies.

Keywords: cryptocurrencies; collective dynamics; time series analysis; portfolio optimization

Citation: James, N.; Menzies, M. Collective Dynamics, Diversification and Optimal Portfolio Construction for Cryptocurrencies. *Entropy* **2023**, *25*, 931. https://doi.org/10.3390/e25060931

Academic Editors: Stanisław Drożdż, Jarosław Kwapień, José F. F. Mendes and Marcin Wątorek

Received: 18 April 2023
Revised: 7 June 2023
Accepted: 12 June 2023
Published: 13 June 2023

1. Introduction

One of the topics at the heart of complex systems analysis is the study of financial markets. Financial markets have a diverse range of participants ranging from extremely sophisticated investors leveraging a technological and information advantage to retail investors who may purchase securities based on other fundamental intuitions. One asset class that has seen a significant degree of variance in the sophistication of the investor base is the cryptocurrency market. Over the last few years, the cryptocurrency market has gathered meaningful interest from institutional and retail investors alike. Despite exhibiting tumultuous changes in aggregate assets under management, the overall market has produced substantial growth in total assets since its inception. Given the relative immaturity of the cryptocurrency market, it is important to study the underlying dynamics of the market and contrast optimal trading and portfolio management strategies with that of more traditional asset classes such as the equity market. The main motivation of this paper is to investigate the next stage of the cryptocurrency market's evolution. Although the cryptocurrency market is young, we feel that it may be coming of age and exhibiting

signs of maturity, becoming more like the equity market. To assess this, we tactically assess whether certain phenomena such as collective movement, uniformity and diversification benefits are similar to that of the equity market.

It is worth commenting more broadly on financial market dynamics and the wealth of work that has been conducted on that topic before we focus on the cryptocurrency market most specifically. There are a variety of academic communities that have studied financial market dynamics and evolutionary changes in structural dynamics such as those in applied mathematics, complex systems and econometrics [1–3]. A wide range of data scientific methodologies has been used to study evolutionary dynamics in financial assets such as linear algebraic-inspired techniques [2,4–6], spectral methods such as random matrix theory [1,7–10], a variety of unsupervised learning methodologies [11,12], change point detection [13,14] and a litany of statistical modeling techniques [15].

Another topic of substantial interest to the financial markets community is that of nonstationarity, regime switching and the time-varying nature of model parameters for phenomena such as volatility. Such research dates back to autoregressive conditionally heteroskedastic (ARCH) models [16], generalized ARCH (GARCH) [17] and stochastic adaptations such as stochastic volatility models [18–20]. Recently, many researchers have explored adaptions to these fundamental models explicitly capturing dynamics exhibited by various time series. Some of these models include exponential general autoregressive conditional heteroskedastic models [21], Glosten–Jagannathan–Runkle GARCH [22], Threshold GARCH [23] and T-SV [24], Markov switching GARCH [25–27] and MS-SV [28]. Many financial mathematicians have also adopted Bayesian estimation methodologies [29–32], generally citing the need for uncertainty quantification in estimating model parameters. These modeling techniques have been widely applied to the study of several asset classes including equities, cryptocurrencies and fixed income [33–38]. Finally, we would be remiss not to mention the wide range of techniques in time series analysis that have been used to study financial problems [39–48], including cryptocurrencies [49–58] and diverse fields in socio- and econophysics [59–77].

Another topic of great interest across asset classes is the topic of portfolio optimization, and more generally, the essence of portfolio construction. The quantitative finance and econometrics communities have studied core issues related to portfolio diversification, where portfolios are optimized with respect to different objective functions [78–85]. More broadly, financial market dynamics are universally difficult to model. The seminal work of Markowitz in 1952 [78] proposed the concept of diversification as a superior framework for investing in multiple securities at a time. The principle underpinning diversification is built upon disassociating the risk of an individual and particular financial asset into a market (systematic) risk component and an asset-specific risk, called unsystematic risk. Diversification essentially equates to smoothing (or averaging over) unsystematic risk by investing in an appropriately large number of individual assets, which leads to candidate investment portfolios' only exposure being inherently due to market risk.

In recent work, the authors of this work and collaborators [86] perform a thorough inspection of diversification properties from the perspective of a pure equity portfolio. Precisely, they explore the changing diversification benefit of various portfolios spread across a range of industry sectors. While in more recent years investor composition has broadened to include the likes of quantitative and high-frequency investors, active investment management has historically been dominated by fundamental investors who make investment decisions based on the future potential of companies relative to market valuations (most commonly, the earnings the company produces relative to its share price). The authors hypothesized that there is more substantial diversification benefit investing across sectors, rather than within them. Indeed, different sectors exhibit varying performance during distinct market periods: some sectors may outperform in buoyant equity markets (such as information technology and often energy), while other sectors outperform in declining equity markets (such as healthcare, consumer staples and utilities).

The authors confirmed this hypothesis, producing four primary findings. First, they use time-varying PCA to highlight that the collective behavior of equities spikes during market crises, rendering diversification far less effective. Second, they demonstrate that various community detection algorithms such as modularity are unable to distinguish between heterogeneous equity sector dynamics during times of crisis. By contrast, during more buoyant equity market periods, equity sector behaviors are more easily distinguished. Third, they introduce a new metric to quantify the uniformity of market impact across equity sectors. There, they show substantial variance across the uniformity of market impact across independent equity sectors. Finally, they demonstrate that a best value equity portfolio exists with respect to evolutionary diversification benefits. They show that a portfolio of size 36, where 4 equities are sampled randomly from 9 different equity sectors, provides comparable diversification benefit to a portfolio of size 81, where 9 equities are randomly sampled from 9 equity sectors. Our critical focus is exploring diversification benefits for cost-conscious retail investors. These are investors who are intelligent, and may be financially interested but lack the resources to trade frequently in an efficient manner.

With respect to the signature of maturity, the cryptocurrency market is very much in its infancy when compared to the equity market. Cryptocurrency sectors are not well defined, and it is often hard to differentiate behaviors between cryptocurrency sectors [14]. If one explores candidate cryptocurrency sector themes online, categories such as wallet, web3, yield farming, play to earn, energy, decentralized finance, distributed computing and cybersecurity can be found. However, these categories frequently overlap or differ from source to source, and it is not necessarily clear how the behaviors of these cryptocurrency sectors relate to the underlying economy. In fact, it is unclear just how often cryptocurrencies are purchased with the underlying coin sector or thematic within the digital ecosystem in mind. We suspect that this phenomenon is especially pronounced among less sophisticated retail investors—where coins may be bought and sold based on factors such as their recent price and volume movements, and overall macroeconomic trends. Accordingly, in this work, we turn to the cryptocurrency market and adapt our experiments to test for alternative diversification strategies among retail cryptocurrency investors. Rather than testing diversification effectiveness among equity sectors, we use tranches of cryptocurrency market capitalizations to proxy sectors. We suspect that many cryptocurrency investors buy securities from platforms where they simply scan a list of assets that are ordered by market capitalization, and are unaware of many coins' association with a deeper role in the digital economy. We feel that this is an original and suitable measure of different "classes" of cryptocurrencies. Here, we apply the same fundamental methodologies to the cryptocurrency market as a means of testing the levels of maturity and sophistication in the cryptocurrency market.

This paper is structured as follows. In Section 2, we outline the data that we use in this paper. In Section 3, we study the evolution of the collective dynamics of the cryptocurrency market. We compare these findings to what has been observed over 20 years in the equity market and draw conclusions regarding the cryptocurrency market's signatures of maturity. In Section 4, we turn to the theme of optimal portfolio construction among cryptocurrency portfolios. There we study marginal diversification benefits as additional cryptocurrency sector deciles, and cryptocurrencies within deciles are sequentially added to a portfolio. In Section 6, we conclude and summarize our findings regarding recent signatures of maturity in the cryptocurrency market.

2. Data

Our data are chosen as follows. Our window of analysis ranges from 1 July 2019 to 14 February 2023. As of the final day of our analysis window, we drew the top 75 cryptocurrencies by market capitalization from Yahoo Finance [87], and restricted these to those with a price history dating back to 1 July 2019. This left 42 cryptocurrencies. We then discarded the two smallest, leaving $N = 40$ cryptocurrencies and their closing price data over $T = 1325$ days. The window of analysis includes periods of major disruption

for cryptocurrencies, such as the COVID-19 market crash in 2020, the BitMEX exchange market crash [88], and the collapse of the FTX exchange in late 2022 [89]. We divide the 40 remaining cryptocurrencies into 10 deciles each with four cryptocurrencies based on market capitalization as of the end of the analysis window. We list all cryptocurrencies analyzed in this paper in Table 1.

Table 1. Cryptocurrencies, their tickers and decile allocations

Cryptocurrency	Ticker	Decile
Bitcoin	BTC	1
Ethereum	ETH	1
Tether	USDT	1
Binance Coin	BNB	1
USD Coin	USDC	2
XRP	XRP	2
Cardano	ADA	2
Polygon	MATIC	2
Dogecoin	DOGE	3
Litecoin	LTC	3
TRON	TRX	3
Wrapped Bitcoin	WBTC	3
Chainlink	LINK	4
Cosmos	ATOM	4
UNUS SED LEO	LEO	4
OKB	OKB	4
Ethereum Classic	ETC	5
Filecoin	FIL	5
Monero	XMR	5
Bitcoin Cash	BCH	5
Stellar	XLM	6
VeChain	VET	6
Crypto.com Coin	CRO	6
Algorand	ALGO	6
Quant	QNT	7
Fantom	FTM	7
Tezos	XTZ	7
Decentraland	MANA	7
EOS	EOS	8
Bitcoin BEP2	BTCB	8
Theta Network	THETA	8
TrueUSD	TUSD	8
Rocket Pool	RPL	9
Chiliz	CHZ	9
USDP Stablecoin	USDP	9
Huobi Token	HT	9
KuCoin Token	KCS	10
Bitcoin SV	BSV	10
Dash	DASH	10
Zcash	ZEC	10

3. Collective Dynamics and Uniformity

Let $c_i(t)$, $i = 1, \ldots, N$, $t = 0, \ldots, T$ denote the multivariate time series of daily closing prices among our collection of N cryptocurrencies. Let $r_i(t)$ be the multivariate time series of log returns $i = 1, \ldots, N$, $t = 1, \ldots, T$, defined as

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

In this section, we analyze correlation matrices of log returns across rolling time windows of length τ; in this paper, we choose $\tau = 90$ days. We standardize the log returns

over such a window $[t-\tau+1, t]$ by defining $R_i(s) = [r_i(s) - \langle r_i \rangle]/\sigma(r_i)$ where $\langle . \rangle$ denotes the average over the time window $[t-\tau+1, t]$ and σ the associated standard deviation. Let $\mathbf{R}$ be the $N \times \tau$ matrix defined by $R_{is} = R_i(s)$ with $i = 1, \ldots, N$ and $s = t-\tau+1, \ldots, t$, and then the correlation matrix $\mathbf{\Psi}$ is then defined as

$$\mathbf{\Psi}(t) = \frac{1}{\tau}\mathbf{R}\mathbf{R}^T. \tag{2}$$

Individual entries describing the correlation between cryptocurrencies i and j can be written as

$$\Psi_{ij}(t) = \frac{1}{\tau}\frac{\sum_{s=t-\tau+1}^{t}(r_i(s) - \langle r_i \rangle)(r_j(s) - \langle r_j \rangle)}{\left(\sum_{s=t-\tau+1}^{t}(r_i(s) - \langle r_i \rangle)^2\right)^{1/2}\left(\sum_{s=t-\tau+1}^{t}(r_j(s) - \langle r_j \rangle)^2\right)^{1/2}}, \tag{3}$$

for $1 \leq i, j \leq N$. We may analogously define the cross-correlation matrices for each individual decile by restricting i and j to be chosen from a set of indices corresponding to that decile.

All entries Ψ_{ij} lie in the interval $[-1, 1]$. $\mathbf{\Psi}$ is a symmetric and positive semi-definite matrix with real and non-negative eigenvalues $\lambda_i(t)$, so we order them as $\lambda_1 \geq \cdots \geq \lambda_N \geq 0$. All the diagonal entries of $\mathbf{\Psi}$ are equal to 1, so the trace of $\mathbf{\Psi}$ is equal to N. Thus, we may normalize the eigenvalues by defining by N, to wit, $\tilde{\lambda}_i = \frac{\lambda_i}{\sum_{j=1}^{N}\lambda_j} = \frac{\lambda_i}{N}$. We display the rolling normalized first eigenvalue $\tilde{\lambda}_1(t)$ for both the entire collection of cryptocurrencies and the 10 deciles in Figure 1.

In Figure 1a, we see particular periods of heightened collective correlation between cryptocurrencies. In particular, we see extended periods of high correlation in early 2020, coinciding with COVID-19 and the BitMEX crash, and toward the end of 2022, reflecting the tumultuous period around the collapse of FTX. These are perhaps the most significant moments of collective crisis in the cryptocurrency market in the last three years. These broad trends are reflected on a decile-by-decile basis in Figure 1b, where each individual decile exhibits a rise in collective correlations in these two periods.

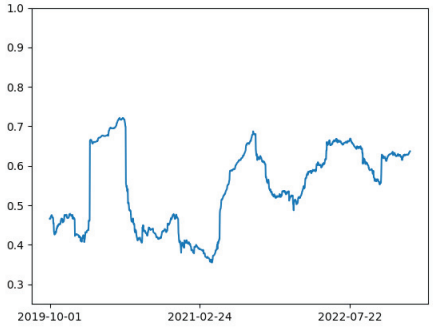
To a nuanced extent, this is a signature of growing maturity in the cryptocurrency market. Specifically, crisis periods are observed; there is a fairly robust time differential between crises; collective correlations rise during crises and fall outside these periods; such effect is seen rather uniformly among different sectors of the market. However, we must remark that the extent of maturity does not coincide with more established markets such as the equity market. The time differential between peaks in collective correlations is still notably shorter than it is for equities; for example, the large time differential between the Dot-com bubble, the global financial crisis and the COVID-19 crash. Moreover, the strength of collective correlations between deciles varies significantly, despite their sharing temporal patterns. Some deciles, such as the third, exhibit significantly higher collective behaviors than others such as the second, fourth and ninth, whereas these behaviors are much more uniform for equity indices.

Next, we turn to an analysis of the leading *eigenvector* $\mathbf{v}_1$ to complement our study of the leading eigenvalue. We analyze its *uniformity* via the following computation:

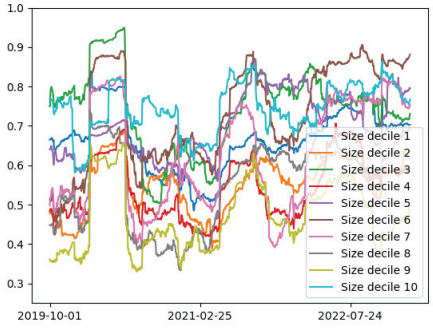
$$h(t) = \frac{|\langle \mathbf{v}_1, \mathbf{1} \rangle|}{\|\mathbf{v}_1\|\|\mathbf{1}\|}, \tag{4}$$

where $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbb{R}^N$ is a uniform vector of 1's. We may compute this for both the entire collection of cryptocurrencies and individual deciles, analogously to the leading eigenvalue. We observe that $h(t) \leq 1$ with $h(t) = 1$ if and only if $\mathbf{v}_1 = \mathbf{1}$ (up to scalar multiplication). In this case, all cryptocurrencies carry the same amount of variance in the "market effect" summarized by $\tilde{\lambda}_1(t)$. This can be used to quantify the potential benefit of diversification: increased values of $h(t)$ indicate increased interchangeability of

cryptocurrencies in the "market", and hence less opportunity for diversification or judicious selection of individual cryptocurrencies.



(**a**)

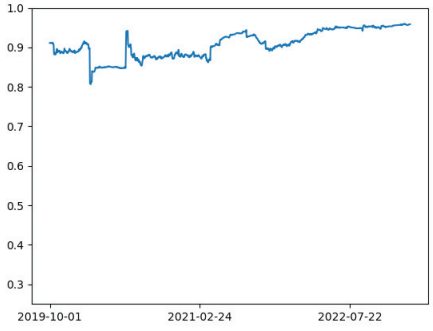


(**b**)

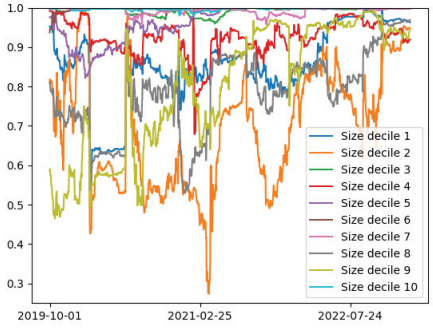
Figure 1. Normalized leading eigenvalue $\tilde{\lambda}_1(t)$ of the cross-correlation matrix as a function of time, for (**a**) the entire collection of cryptocurrencies and (**b**) the ten deciles. Like the equity market, collective correlations spike during market crises, such as COVID-19, and the collapse of exchanges BitMEX and FTX.

We display the rolling uniformity of the first eigenvector $h(t)$ for both the entire collection of cryptocurrencies and the 10 deciles in Figure 2. Unlike Figure 1, we observe a substantial difference compared to the equity market. In the case of the equity market, the uniformity for each sector and the entire market are consistent with the degree of collectivity. The degree of uniformity spikes during market crises such as the dot-com bubble GFC and COVID-19. This spike during market crises occurs for sectors (when studied independently) as well as across the entire market. The cryptocurrency market produces dramatically different findings to that of the equity market. Most notably, we see that there are substantial differences between the uniformity of independent sectors of the cryptocurrency market with that of the equity market. The cryptocurrency market clearly exhibits less uniformity during crises (which we see during the COVID-19 market crisis), and significantly higher variance between sectors of securities throughout our window of analysis. This is opposite to the finding of the equity market, where industry sectors exhibited more uniformity during crises. Another point to note is the stark contrast in how low the $h(t)$ values reach when comparing the two asset classes. In the case of equities, there is a clear lower bound around the value of 0.75, while for cryptocurrencies we see two groups of cryptocurrencies reach values below 0.5 (with one reaching less than 0.3)

during our analysis window. Our analysis, therefore, suggests that we see less persistent and amplified uniformity among cryptocurrencies when compared to equities.



(**a**)



(**b**)

Figure 2. Uniformity $h(t)$ of the leading eigenvector $\mathbf{v}_1$ of the cross-correlation matrix as a function of time, for (**a**) the entire collection of cryptocurrencies and (**b**) the ten deciles. The results are dramatically different compared to the equity market, with numerous deciles exhibiting strikingly low uniformity scores over time.

4. Portfolio Sampling

In this section, we perform an extensive sampling procedure to explore how diversification benefits depend on the number of cryptocurrencies held in a portfolio and on the number of decile sectors from which to choose those cryptocurrencies. Motivated by Section 3, we choose the normalized leading eigenvalue $\tilde{\lambda}_1(t)$ to quantify the potential diversification benefit. We investigate the diversification benefits of portfolios that consist of mn cryptocurrencies such that n cryptocurrencies are drawn from m separate deciles. Both the individual cryptocurrencies and the sector deciles are drawn randomly and independently with uniform probability. We draw $D = 500$ portfolios for each ordered pair (m, n).

To quantify the potential diversification benefit for a portfolio consisting of mn cryptocurrencies, we determine the $mn \times mn$ correlation matrix $\mathbf{\Psi}$ for each draw and calculate the associated normalized first eigenvalue $\tilde{\lambda}_{m,n}(t)$. Again, we use a rolling time window of length $\tau = 90$ days when determining the cross-correlation matrix. In particular, we record the 50th percentile (median) of the D values, which we denote $\tilde{\lambda}^{0.50}_{m,n}(t)$.

We analyze this quantity in two experiments. First, we compute the temporal mean of the median of the normalized first eigenvalue

$$\mu_{m,n} = \frac{1}{T-\tau+1} \sum_{t=\tau}^{T} \tilde{\lambda}_{m,n}^{0.50}(t) \tag{5}$$

as a measure of the diversification benefit of a portfolio with n cryptocurrencies in each of m decile sectors. Table 2 records these means $\mu_{m,n}$ for cryptocurrency portfolios across values of (m, n) for $1 \leq m \leq 10$ and $1 \leq n \leq 4$.

Table 2 shows the mean $\mu_{m,n}$ of the median normalized eigenvalue $\tilde{\lambda}_{m,n}^{0.50}(t)$ for various combinations of cryptocurrency sectors, and randomly sampled cryptocurrencies within each sector. We also mark in red a "greedy path" to decrease the overall average $\mu_{m,n}$ (that is, increase the overall diversification benefit of a portfolio) by greedily increasing either m or n at each stage. There are several key findings from this analysis. First, the overall structural finding with respect to optimal portfolio construction strongly resembles that of the equity market in [86]. We see incrementally greater benefit in diversifying across sectors rather than within them, and we see a significant reduction in marginal diversification benefit once a portfolio reaches a critical mass of securities (sampled from different sectors). This leads to the existence of a "best value" cryptocurrency portfolio, such as that seen in the equity market. This finding is slightly surprising and may support our hypothesis that retail cryptocurrency investors diversify across cryptocurrency market capitalization levels. Indeed, this may occur in the absence of clearly defined sector themes, which may exhibit different performances during different parts of the business cycle. As investors come to better understand cryptocurrencies, and cryptocurrencies related to separate aspects of the digital economy begin to perform differently during various market conditions, this diversification benefit may slightly alter and amplify. That is, rather than cryptocurrency market capitalization being a primary discriminator in diversified performance we may see a closer resemblance to the equity market with cryptocurrency sector themes more closely resembling equity dynamics.

Table 2. Average $\mu_{m,n}$ of the median normalized eigenvalue $\tilde{\lambda}_{m,n}^{0.50}(t)$ for different pairs of m sectors and n cryptocurrencies per sector. In red we display a greedy path that aims to increase the total diversification benefit (by decreasing $\mu_{m,n}$) at each step. We identify a best value cryptocurrency portfolio consisting of 4 sectors and 4 cryptocurrencies per sector. This (4,4) portfolio has nearly the same diversification benefit as the largest possible (10,4) portfolio, as we will also show in the next experiment.

	Number of Cryptocurrencies per Sector			
Number of Sectors	**1**	**2**	**3**	**4**
1	1	0.759	0.668	0.645
2	0.774	0.651	0.598	0.587
3	0.681	0.605	0.581	0.576
4	0.641	0.587	0.572	0.565
5	0.613	0.583	0.565	0.559
6	0.607	0.57	0.565	0.557
7	0.593	0.565	0.559	0.555
8	0.582	0.564	0.557	0.552
9	0.552	0.565	0.557	0.553
10	0.581	0.560	0.554	0.552

In the second experiment, we investigate which portfolio combinations (m, n) share the most similar evolution in their collective dynamics. For this purpose, we perform hierarchical clustering on the distance metric

$$d((m,n),(m',n')) = \frac{1}{T-\tau+1}\sum_{t=\tau}^{T}|\tilde{\lambda}^{0.50}_{m,n}(t) - \tilde{\lambda}^{0.50}_{m',n'}(t)|, \tag{6}$$

which computes the average absolute difference between the median eigenvalues of two portfolios (m, n) and (m', n'). This results in a 40×40 distance matrix for $1 \leq m \leq 10$, $1 \leq n \leq 4$. Hierarchical clustering is a convenient and easily visualizable tool to reveal the proximity between different elements of a collection. Here, we perform agglomerative hierarchical clustering using the average-linkage criterion [90]. The algorithm works in a bottom-up manner, where each ordered pair (m, n) starts in its own cluster, and pairs of clusters are successively merged as one traverses up the hierarchy.

The results of hierarchical clustering are displayed in Figure 3. The resulting structure is interesting. The dendrogram consists of four predominant groups of clusters. There is an easily identified outlier cluster, consisting of the smallest portfolios that provide the least diversification benefit. This cluster, located at the bottom of the dendrogram, includes portfolio combinations such as (1,1), (1,2) and (2,1). The second least diversified cluster is located at the top of the dendrogram and includes portfolio combinations such as (1,3), (1,4) and (4,1). Below this, is a significantly larger cluster consisting of portfolio combinations such as (8,1), (3,3) and (4,2). Finally, the largest, most well-diversified fourth cluster consists of portfolio combinations ranging from (4,3) through to (10,4). The size and range of portfolio combinations within this cluster have interesting implications for risk management in cryptocurrency portfolio diversification. The fact that combination (4,3) is in the same cluster as portfolio (10,4) suggests that comparable risk mitigation can be realized with a portfolio of size 12 when compared to a portfolio of size 40. This finding is not dissimilar to that proposed in [86], where a "best value" portfolio (9,4) is shown to provide comparable diversification benefit to a (9,9) portfolio. Furthermore, the sheer size of this cluster indicates that one may require a lower cardinality portfolio in cryptocurrency portfolio management than in equities when trying to attain a "best value" portfolio.

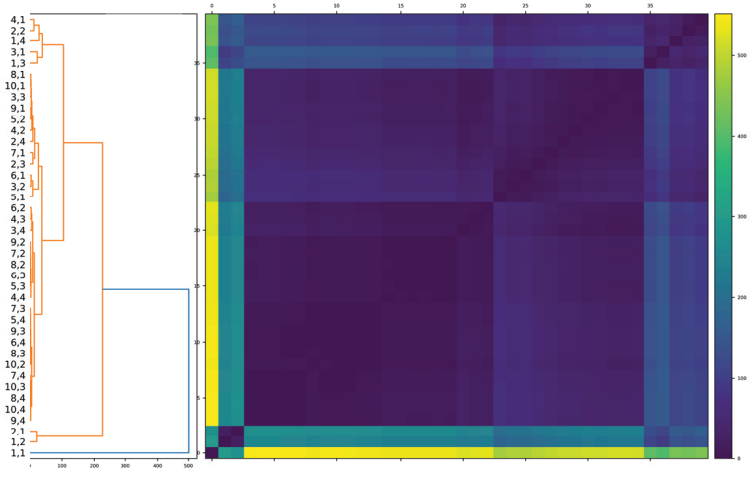


Figure 3. Results of hierarchical clustering applied to (6) between ordered pairs (m, n). A large majority cluster confirms the finding of Table 2 that the (4,4) portfolio is closely similar to the full (10,4) portfolio in its diversification benefit over time.

5. Discussion

This paper has investigated the cryptocurrency market with a focus on collective correlation dynamics and portfolio diversification benefits across different market capitalization size deciles. We choose to investigate deciles as an analog of industry sectors in the equity market motivated by the vagueness of existing cryptocurrency sectors and the hypothesis that many retail investors use size as a primary means of diversifying across highly different cryptocurrencies. Throughout the paper, we consistently observed signatures of maturity in the cryptocurrency market with nuanced differences relative to established patterns in the (much) more mature equity market.

In Section 3, we analyzed the collective dynamics of the cryptocurrency market, focusing on collective correlation strength and market uniformity summarized in the leading eigenvalue and eigenvector of the correlation matrix, examining both the full market and individual deciles. Our first finding is that collective correlations spike during market crises connected to cryptocurrency exchange crashes; this occurs in every decile and closely resembles analogous behavior in the more mature equity market. Other findings of this section portray a more nuanced picture of the differences between the cryptocurrency and equity markets. While collective correlations spike across every decile during a crisis, it is not the case that correlations within each decile sector are uniformly higher than collective correlations across the whole market, as we previously observed for the equity market. In addition, the uniformity $h(t)$ of different deciles over time exhibited a finding highly dissimilar from the equity market. This was the most significant difference relative to the equity market observed in this paper. While the uniformity (measuring the uniformity of different assets contributing toward the first principle component) was close to maximal 1 for every sector in the equity markets, that finding was not at all observed for the deciles of the cryptocurrency market. Curiously, it was observed for just two deciles consistently over time, but not the others. In addition, uniformity within deciles dropped during market crises, the opposite finding for the equity market.

These findings have significant implications for the alpha generation in the cryptocurrency market. The fact that collective correlations are so pronounced during market crises implies that alpha-generating opportunities based on systematic market movements would be more predictable and successful if performed on the short side. During market crises, correlations translate upward and cryptocurrencies of all sizes tend to decline. This would suggest that fundamentally-driven investment strategies may be more successful when implemented during buoyant equity markets, where there is less correlation among underlying securities. By contrast, during market crises (which are typically coincident among the equity and cryptocurrency asset classes) the collective strength of the market is so strong that the weight of underlying security investments driven by bottom-up analysis may be washed away by the sheer weight of money.

In Section 4, our portfolio sampling experiment investigated the diversification benefit of portfolios of total size mn spread evenly across m separate deciles. In a greedy experiment, we demonstrated that greater diversification benefit is generally obtained by increasing the number of decile sectors rather than the number of cryptocurrencies per decile, a result analogous to that observed for the cryptocurrency market. We followed this up with a careful experiment clustering different temporal trajectories of median normalized eigenvalue functions $\tilde{\lambda}_{m,n}(t)$. A large majority cluster showed a similar result as observed for the equity market, that a portfolio of spread (4,4) had near-identical diversification benefit as our maximal size (10,4) cryptocurrency collection.

Our findings in this section may drive decision-making for optimal portfolio construction for cryptocurrency investors. First, the emergence of a low-cardinality highly diversified portfolio implies that retail investors may gain exposure to high-quality diversification at low-cost. When contrasting this analysis with that of the equity market, if we were to assume equivalent transaction costs and equivalent periodicity of portfolio rebalancing, the cryptocurrency may be a more retail-friendly market for easy access and portfolio diversification. Of course, given that the equity market is so sophisticated, there

is a large number of index-tracking and factor-based investment strategies that may benefit retail investors. This could be an opportunity for asset managers and large investment firms who are looking to create cryptocurrency investment products and is certainly a signature of the market's maturity. Finally, our analysis supports the notion that the cryptocurrency market may be a suitable environment for skillful stock pickers. We have highlighted that a portfolio of just 16 stocks produces low correlation and significant diversification benefits. This would indicate that an investment portfolio built upon a smaller number of high-conviction ideas could thrive in the cryptocurrency market.

There are several insights contained within concerning the cryptocurrency market's levels of maturity. First, the overall structure of the aforementioned hierarchical clustering is highly similar to that of the equity market. We have identified heterogeneous clusters of diversification benefit, and highlight the existence of a "best value" cryptocurrency portfolio where comparable diversification benefit is attained with relatively fewer securities held in a portfolio. Second, a crucial corollary of this finding is that retail investors with limited ability to hold complex portfolios of many cryptocurrencies may be sufficiently diversified with a relatively small portfolio across just 16 cryptocurrencies. However, there are some key differences in the equity market. First, the link between underlying cryptocurrencies' business functions (at least those coins that have a business function) and various business cycles is far less clear than in the equity market. Perhaps when the market becomes more sophisticated and such technological understanding becomes more mainstream knowledge, this could change the landscape of cryptocurrency investing. This could lead to the development of better-understood and widely disseminated cryptocurrency investment principles, which may drive more predictable investment patterns during different market cycles. Such dynamics are likely to drive further differentiation in cryptocurrency price patterns in varying market cycles and may lead to further diversification benefits as the market approaches greater levels of maturity.

No analysis is without its limitations. There are several limitations in our work. First, we have only looked at a collection of 40 cryptocurrencies. This could be extended, and include a much wider variety of cryptocurrency securities. The difficulty here is that many smaller coins do not have sufficient time windows for us to analyze. However, as time goes on, doing such analysis on a larger number of coins will become easier and may provide more robust insights. Furthermore, we could extend our portfolio sampling analysis to explicitly study diversification benefits during various market conditions. In the near future, we may be able to compare a short and intense market crisis such as the COVID-19 market crash or the Bitmex crash with that of the Russian financial crisis—or something more protracted and systemic. At present, the data is most likely insufficient.

6. Conclusions

Overall, we have uncovered nuanced similarities and differences between the cryptocurrency and equity markets. These mathematical properties signal increased signatures of maturity in the collective dynamics and diversification benefit of different portfolio spreads and provide concrete suggestions to retail investors seeking a relatively low-complexity exposure to the cryptocurrency market. Cryptocurrency decile sectors have been shown to share several properties, but not all, with industry sectors of the equity markets, and the most relevant findings for small-scale investors interested in limited-size portfolios are shared.

There are a variety of opportunities for future research building upon the methodologies we have developed and the insights highlighted in this paper. First, it would be interesting to test how the cryptocurrency market compares with other more mature and better-established asset classes with respect to various statistical properties. A thorough inspection of metrics such as drawdowns, peak-to-trough analysis, changepoint propagation, intra and inter-asset correlations, etc., could reveal information as to how cryptocurrencies can complement a multi-asset investment portfolio. An additional avenue of future research could be studying the data at a higher sampling rate than daily data. Given the significant

composition of day traders in the cryptocurrency market, we may see patterns that deviate from what we see intra-day within the equity market. In a somewhat-related manner, studying these securities in a longer time horizon may highlight regime shifts in dynamics or optimal portfolio construction. One of the key assumptions in this work is our separating cryptocurrencies into size deciles. Further work could look into alternative bucketing criteria such as sector allocation, volatility or other measures of risk. Finally, given the number of quantitative investment firms with burgeoning high-frequency cryptocurrency trading operations, one could examine the effectiveness of frequency-based trading strategies to see if there is greater "power" with certain trading windows. This could reveal typical holding periods for investment firms that trade in the cryptocurrency market.

Author Contributions: Both authors contributed equally to every aspect of the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data analyzed in this article are publicly available at [87]. A cached copy of the data and code are available at https://github.com/nickjamesgithub/Cryptocurrency_Entropy.

Conflicts of Interest: The authors declare no conflict of interest.

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* **2011**, *84*, 026109. [CrossRef] [PubMed]
2. Laloux, L.; Cizeau, P.; Bouchaud, J.P.; Potters, M. Noise Dressing of Financial Correlation Matrices. *Phys. Rev. Lett.* **1999**, *83*, 1467–1470. [CrossRef]
3. Münnix, M.C.; Shimada, T.; Schäfer, R.; Leyvraz, F.; Seligman, T.H.; Guhr, T.; Stanley, H.E. Identifying States of a Financial Market. *Sci. Rep.* **2012**, *2*, 644. [CrossRef] [PubMed]
4. Kim, D.H.; Jeong, H. Systematic analysis of group identification in stock markets. *Phys. Rev. E* **2005**, *72*, 046133. [CrossRef] [PubMed]
5. Pan, R.K.; Sinha, S. Collective behavior of stock price movements in an emerging market. *Phys. Rev. E* **2007**, *76*, 046116. [CrossRef]
6. Wilcox, D.; Gebbie, T. An analysis of cross-correlations in an emerging market. *Phys. Stat. Mech. Its Appl.* **2007**, *375*, 584–598. [CrossRef]
7. Conlon, T.; Ruskin, H.; Crane, M. Random matrix theory and fund of funds portfolio optimisation. *Phys. Stat. Mech. Its Appl.* **2007**, *382*, 565–576. [CrossRef]
8. Bouchaud, J.P.; Potters, M. *Theory of Financial Risk and Derivative Pricing*; Cambridge University Press: Cambridge, UK, 2003. [CrossRef]
9. Burda, Z.; Görlich, A.; Jarosz, A.; Jurkiewicz, J. Signal and noise in correlation matrix. *Phys. Stat. Mech. Its Appl.* **2004**, *343*, 295–310. [CrossRef]
10. Sharifi, S.; Crane, M.; Shamaie, A.; Ruskin, H. Random matrix theory for portfolio optimization: A stability approach. *Phys. Stat. Mech. Its Appl.* **2004**, *335*, 629–643. [CrossRef]
11. Heckens, A.J.; Krause, S.M.; Guhr, T. Uncovering the dynamics of correlation structures relative to the collective market motion. *J. Stat. Mech. Theory Exp.* **2020**, *2020*, 103402. [CrossRef]
12. James, N.; Menzies, M. Association between COVID-19 cases and international equity indices. *Phys. Nonlinear Phenom.* **2021**, *417*, 132809. [CrossRef]
13. James, N.; Menzies, M.; Chan, J. Changes to the extreme and erratic behaviour of cryptocurrencies during COVID-19. *Phys. Stat. Mech. Its Appl.* **2021**, *565*, 125581. [CrossRef]
14. James, N. Dynamics, behaviours, and anomaly persistence in cryptocurrencies and equities surrounding COVID-19. *Phys. Stat. Mech. Its Appl.* **2021**, *570*, 125831. [CrossRef]
15. Ferreira, P.; Kristoufek, L.; de Area Leão Pereira, E.J. DCCA and DMCA correlations of cryptocurrency markets. *Phys. Stat. Mech. Its Appl.* **2020**, *545*, 123803. [CrossRef]
16. Engle, R. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* **1982**, *50*, 987–1007. [CrossRef]
17. Bollerslev, T. Generalized Autoregressive Conditional Heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [CrossRef]
18. Taylor, S. Financial Returns Modelled by the product of two stochastic processes, a study of daily sugar prices 1961–1979. In *Time Series Analysis: Theory and Practice 1*; North-Holland: Amsterdam, The Netherlands, 1982; pp. 203–226.

19. Taylor, S. *Modelling Financial Time Series*; Wiley: New York, NY, USA, 1986.
20. Taylor, S.J. Modelling Stochastic Volatility: A review and comparative study. *Math. Financ.* **1994**, *4*, 183–204. [CrossRef]
21. Nelson, D. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* **1991**, *59*, 347–370. [CrossRef]
22. Glosten, L.R.; Jagannathan, R.; Runkle, D.E. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *J. Financ.* **1993**, *48*, 1779–1801. [CrossRef]
23. Zakoian, J. Threshold heteroskedastic models. *J. Econ. Dyn. Control.* **1994**, *18*, 931–955. [CrossRef]
24. So, M.; Li, W.; Lam, K. On a Threshold Stochastic Volatility Model. *J. Forecast.* **2002**, *22*, 473–500. [CrossRef]
25. Cai, J. A Markov Model of Switching-Regime ARCH. *J. Bus. Econ. Stat.* **1994**, *12*, 309–316.
26. Hamilton, J.; Susmel, R. Autoregressive conditional heteroskedasticity and changes in regime. *J. Econom.* **1994**, *64*, 307–333. [CrossRef]
27. Gray, S. Modelling the conditional distribution of interest rates as a regime-switching process. *J. Financ. Econom.* **1996**, *2*, 211–250.
28. So, M.; Lam, K.; Li, W. A stochastic volatility model with markov switching. *J. Bus. Econ. Stat.* **1998**, *16*, 244–253.
29. Andersen, T.; Bollerslev, T.; Diebold, F.; Labys, P. Modeling and forecasting realized volatility. *Econometrica* **2003**, *71*, 579–625. [CrossRef]
30. Hwang, S.; Satchell, S. GARCH Model with Cross-sectional Volatility: GARCHX Models. *Appl. Financ. Econ.* **2007**, *15*, 203–216. [CrossRef]
31. Hansen, P.R.; Huang, Z.; Shek, H.H. Realized GARCH: A joint model for returns and realized measures of volatility. *J. Appl. Econom.* **2011**, *27*, 877–906. [CrossRef]
32. James, N.; Menzies, M. Optimally adaptive Bayesian spectral density estimation for stationary and nonstationary processes. *Stat. Comput.* **2022**, *32*, 45. [CrossRef]
33. Cerqueti, R.; Giacalone, M.; Mattera, R. Skewed non-Gaussian GARCH models for cryptocurrencies volatility modelling. *Inf. Sci.* **2020**, *527*, 1–26. [CrossRef]
34. Wan, Y.; Si, Y.W. A formal approach to chart patterns classification in financial time series. *Inf. Sci.* **2017**, *411*, 151–175. [CrossRef]
35. Stehlík, M.; Helperstorfer, C.; Hermann, P.; Šupina, J.; Grilo, L.; Maidana, J.; Fuders, F.; Stehlíková, S. Financial and risk modelling with semicontinuous covariances. *Inf. Sci.* **2017**, *394–395*, 246–272. [CrossRef]
36. Chu, C.S.J.; Santoni, G.J.; Liu, T. Stock market volatility and regime shifts in returns. *Inf. Sci.* **1996**, *94*, 179–190. [CrossRef]
37. yong Chen, G.; Gan, M.; long Chen, G. Generalized exponential autoregressive models for nonlinear time series: Stationarity, estimation and applications. *Inf. Sci.* **2018**, *438*, 46–57. [CrossRef]
38. Cerqueti, R.; Giacalone, M.; Panarello, D. A Generalized Error Distribution Copula-based method for portfolios risk assessment. *Phys. Stat. Mech. Its Appl.* **2019**, *524*, 687–695. [CrossRef]
39. Drożdż, S.; Kwapień, J.; Oświęcimka, P. Complexity in Economic and Social Systems. *Entropy* **2021**, *23*, 133. [CrossRef]
40. Liu, Y.; Cizeau, P.; Meyer, M.; Peng, C.K.; Stanley, H.E. Correlations in economic time series. *Phys. Stat. Mech. Its Appl.* **1997**, *245*, 437–440. [CrossRef]
41. James, N.; Menzies, M.; Chan, J. Semi-metric portfolio optimization: A new algorithm reducing simultaneous asset shocks. *Econometrics* **2023**, *11*, 8. [CrossRef]
42. Basalto, N.; Bellotti, R.; Carlo, F.D.; Facchi, P.; Pantaleo, E.; Pascazio, S. Hausdorff clustering of financial time series. *Phys. Stat. Mech. Its Appl.* **2007**, *379*, 635–644. [CrossRef]
43. Wątorek, M.; Kwapień, J.; Drożdż, S. Financial Return Distributions: Past, Present, and COVID-19. *Entropy* **2021**, *23*, 884. [CrossRef]
44. Prakash, A.; James, N.; Menzies, M.; Francis, G. Structural Clustering of Volatility Regimes for Dynamic Trading Strategies. *Appl. Math. Financ.* **2021**, *28*, 236–274. [CrossRef]
45. Drożdż, S.; Grümmer, F.; Ruf, F.; Speth, J. Towards identifying the world stock market cross-correlations: DAX versus Dow Jones. *Phys. Stat. Mech. Its Appl.* **2001**, *294*, 226–234. [CrossRef]
46. James, N.; Menzies, M.; Chin, K. Economic state classification and portfolio optimisation with application to stagflationary environments. *Chaos Solitons Fractals* **2022**, *164*, 112664. [CrossRef]
47. 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.* **2019**, *98*, 2349–2364. [CrossRef]
48. James, N.; Menzies, M. A new measure between sets of probability distributions with applications to erratic financial behavior. *J. Stat. Mech. Theory Exp.* **2021**, *2021*, 123404. [CrossRef]
49. Sigaki, H.Y.D.; Perc, M.; Ribeiro, H.V. Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market. *Sci. Rep.* **2019**, *9*, 1440. [CrossRef]
50. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Stanisz, T.; Wątorek, M. Complexity in Economic and Social Systems: Cryptocurrency Market at around COVID-19. *Entropy* **2020**, *22*, 1043. [CrossRef]
51. James, N.; Menzies, M. Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time. *Nonlinear Dyn.* **2022**, *107*, 4001–4017. [CrossRef]
52. 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 Interdiscip. J. Nonlinear Sci.* **2020**, *30*, 023122. [CrossRef]
53. Wątorek, M.; Drożdż, S.; Kwapień, J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* **2021**, *901*, 1–82. [CrossRef]

54. 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 Interdiscip. J. Nonlinear Sci.* **2018**, *28*, 071101. [CrossRef]
55. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Signatures of the Crypto-Currency Market Decoupling from the Forex. *Future Internet* **2019**, *11*, 154. [CrossRef]
56. Kwapień, J.; Wątorek, M.; Bezbradica, M.; Crane, M.; Mai, T.T.; Drożdż, S. Analysis of inter-transaction time fluctuations in the cryptocurrency market. *Chaos Interdiscip. J. Nonlinear Sci.* **2022**, *32*, 083142. [CrossRef]
57. Wątorek, M.; Kwapień, J.; Drożdż, S. Multifractal Cross-Correlations of Bitcoin and Ether Trading Characteristics in the Post-COVID-19 Time. *Future Internet* **2022**, *14*, 215. [CrossRef]
58. Wątorek, M.; Kwapień, J.; Drożdż, S. Cryptocurrencies Are Becoming Part of the World Global Financial Market. *Entropy* **2023**, *25*, 377. [CrossRef]
59. James, N.; Menzies, M. Spatio-temporal trends in the propagation and capacity of low-carbon hydrogen projects. *Int. J. Hydrog. Energy* **2022**, *47*, 16775–16784. [CrossRef]
60. Perc, M.; Miksić, N.G.; Slavinec, M.; Stožer, A. Forecasting COVID-19. *Front. Phys.* **2020**, *8*, 127. [CrossRef]
61. Machado, J.A.T.; Lopes, A.M. Rare and extreme events: The case of COVID-19 pandemic. *Nonlinear Dyn.* **2020**, *100*, 2953–2972. [CrossRef]
62. James, N.; Menzies, M. Distributional trends in the generation and end-use sector of low-carbon hydrogen plants. *Hydrogen* **2023**, *4*, 174–189. [CrossRef]
63. Merritt, S.; Clauset, A. Scoring dynamics across professional team sports: Tempo, balance and predictability. *EPJ Data Sci.* **2014**, *3*, 4. [CrossRef]
64. James, N.; Menzies, M. Equivalence relations and L^p distances between time series with application to the Black Summer Australian bushfires. *Phys. Nonlinear Phenom.* **2023**, *448*, 133693. [CrossRef]
65. Khan, M.K.; Khan, M.I.; Rehan, M. The relationship between energy consumption, economic growth and carbon dioxide emissions in Pakistan. *Financ. Innov.* **2020**, *6*, 1. [CrossRef]
66. James, N.; Menzies, M. Global and regional changes in carbon dioxide emissions: 1970–2019. *Phys. A Stat. Mech. Its Appl.* **2022**, *608*, 128302. [CrossRef]
67. Manchein, C.; Brugnago, E.L.; da Silva, R.M.; Mendes, C.F.O.; Beims, M.W. Strong correlations between power-law growth of COVID-19 in four continents and the inefficiency of soft quarantine strategies. *Chaos Interdiscip. J. Nonlinear Sci.* **2020**, *30*, 041102. [CrossRef] [PubMed]
68. Ribeiro, H.V.; Mukherjee, S.; Zeng, X.H.T. Anomalous diffusion and long-range correlations in the score evolution of the game of cricket. *Phys. Rev. E* **2012**, *86*, 022102. [CrossRef]
69. James, N.; Menzies, M. Estimating a continuously varying offset between multivariate time series with application to COVID-19 in the United States. *Eur. Phys. J. Spec. Top.* **2022**, *231*, 3419–3426. [CrossRef]
70. Li, H.J.; Xu, W.; Song, S.; Wang, W.X.; Perc, M. The dynamics of epidemic spreading on signed networks. *Chaos Solitons Fractals* **2021**, *151*, 111294. [CrossRef]
71. Blasius, B. Power-law distribution in the number of confirmed COVID-19 cases. *Chaos Interdiscip. J. Nonlinear Sci.* **2020**, *30*, 093123. [CrossRef]
72. James, N.; Menzies, M.; Chok, J.; Milner, A.; Milner, C. Geometric persistence and distributional trends in worldwide terrorism. *Chaos Solitons Fractals* **2023**, *169*, 113277. [CrossRef]
73. Clauset, A.; Kogan, M.; Redner, S. Safe leads and lead changes in competitive team sports. *Phys. Rev. E* **2015**, *91*, 062815. [CrossRef]
74. James, N.; Menzies, M. Dual-domain analysis of gun violence incidents in the United States. *Chaos Interdiscip. J. Nonlinear Sci.* **2022**, *32*, 111101. [CrossRef]
75. Perc, M.; Donnay, K.; Helbing, D. Understanding Recurrent Crime as System-Immanent Collective Behavior. *PLoS ONE* **2013**, *8*, e76063. [CrossRef]
76. Singh, P.; Watorek, M.; Ceglarek, A.; Fąfrowicz, M.; Lewandowska, K.; Marek, T.; Sikora-Wachowicz, B.; Oswiecimka, P. Analysis of fMRI Signals from Working Memory Tasks and Resting-State of Brain: Neutrosophic-Entropy-Based Clustering Algorithm. *Int. J. Neural Syst.* **2022**, *32*, 2250012. [CrossRef]
77. James, N.; Menzies, M.; Bondell, H. In search of peak human athletic potential: A mathematical investigation. *Chaos Interdiscip. J. Nonlinear Sci.* **2022**, *32*, 023110. [CrossRef]
78. Markowitz, H. Portfolio Selection. *J. Financ.* **1952**, *7*, 77. [CrossRef]
79. Sharpe, W.F. Mutual Fund Performance. *J. Bus.* **1966**, *39*, 119–138. [CrossRef]
80. Almahdi, S.; Yang, S.Y. An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown. *Expert Syst. Appl.* **2017**, *87*, 267–279. [CrossRef]
81. Calvo, C.; Ivorra, C.; Liern, V. Fuzzy portfolio selection with non-financial goals: Exploring the efficient frontier. *Ann. Oper. Res.* **2014**, *245*, 31–46. [CrossRef]
82. Soleimani, H.; Golmakani, H.R.; Salimi, M.H. Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm. *Expert Syst. Appl.* **2009**, *36*, 5058–5063. [CrossRef]
83. Vercher, E.; Bermúdez, J.D.; Segura, J.V. Fuzzy portfolio optimization under downside risk measures. *Fuzzy Sets Syst.* **2007**, *158*, 769–782. [CrossRef]

84. Bhansali, V. Putting Economics (Back) into Quantitative Models. *J. Portf. Manag.* **2007**, *33*, 63–76. [CrossRef]
85. Moody, J.; Saffell, M. Learning to trade via direct reinforcement. *IEEE Trans. Neural Netw.* **2001**, *12*, 875–889. [CrossRef] [PubMed]
86. James, N.; Menzies, M.; Gottwald, G.A. On financial market correlation structures and diversification benefits across and within equity sectors. *Phys. A Stat. Mech. Its Appl.* **2022**, *604*, 127682. [CrossRef]
87. Cryptocurrencies. Yahoo Finance. 2023. Available online: https://finance.yahoo.com/crypto/ (accessed on 14 February 2023)
88. Bambrough, B. Here's What Caused Bitcoin's 'Extreme' Price Plunge. Forbes. 2020. Available online: https://www.forbes.com/sites/billybambrough/2020/03/19/major-bitcoin-exchange-bitmex-has-a-serious-problem/ (accessed on 19 March 2020).
89. Huang, K. Why Did FTX Collapse? Here's What to Know. The New York Times. 2022. Available online: https://www.nytimes.com/2022/11/10/technology/ftx-binance-crypto-explained.html (accessed on 10 November 2022).
90. Müllner, D. Fastcluster: Fast Hierarchical, Agglomerative Clustering Routines forRandPython. *J. Stat. Softw.* **2013**, *53*, 1–18. [CrossRef]

entropy

MDPI

Article

Cryptocurrencies Are Becoming Part of the World Global Financial Market

Marcin Wątorek [1,*], Jarosław Kwapień [2] and Stanisław Drożdż [1,2,*]

[1] Faculty of Computer Science and Telecommunications, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland
[2] Complex Systems Theory Department, Institute of Nuclear Physics, Polish Academy of Sciences, Radzikowskiego 152, 31-342 Kraków, Poland
* Correspondence: marcin.watorek@pk.edu.pl (M.W.); stanislaw.drozdz@ifj.edu.pl (S.D.)

Abstract: In this study the cross-correlations between the cryptocurrency market represented by the two most liquid and highest-capitalized cryptocurrencies: bitcoin and ethereum, on the one side, and the instruments representing the traditional financial markets: stock indices, Forex, commodities, on the other side, are measured in the period: January 2020–October 2022. Our purpose is to address the question whether the cryptocurrency market still preserves its autonomy with respect to the traditional financial markets or it has already aligned with them in expense of its independence. We are motivated by the fact that some previous related studies gave mixed results. By calculating the q-dependent detrended cross-correlation coefficient based on the high frequency 10 s data in the rolling window, the dependence on various time scales, different fluctuation magnitudes, and different market periods are examined. There is a strong indication that the dynamics of the bitcoin and ethereum price changes since the March 2020 COVID-19 panic is no longer independent. Instead, it is related to the dynamics of the traditional financial markets, which is especially evident now in 2022, when the bitcoin and ethereum coupling to the US tech stocks is observed during the market bear phase. It is also worth emphasizing that the cryptocurrencies have begun to react to the economic data such as the Consumer Price Index readings in a similar way as traditional instruments. Such a spontaneous coupling of the so far independent degrees of freedom can be interpreted as a kind of phase transition that resembles the collective phenomena typical for the complex systems. Our results indicate that the cryptocurrencies cannot be considered as a safe haven for the financial investments.

Keywords: complexity; financial markets; cryptocurrencies; cross-correlations; multiscale; hedge

Citation: Wątorek, M.; Kwapień, J.; Drożdż, S. Cryptocurrencies Are Becoming Part of the World Global Financial Market. *Entropy* **2023**, *25*, 377. https://doi.org/10.3390/e25020377

Academic Editors: Panos Argyrakis and José F. F. Mendes

Received: 9 November 2022
Revised: 5 December 2022
Accepted: 16 February 2023
Published: 18 February 2023

1. Introduction

From the physics point of view, the financial markets are considered as one of the most complex systems we observe in our world [1]. Not only they are characterised by all the properties such systems can typically show, but there is also an important intelligent component involved that is decisively responsible for their enormous complexity. Among the well-known features of the financial markets is their flexibility in the transition between the disordered and ordered phases. Such a transition is the key feature associated with the market crashes but it is also often observed on the level of whole markets when some so-far independent markets start to have their dynamics substantially coupled (or *vice-versa*). Exactly this kind of phenomenon has recently been experienced by the cryptocurrency market, which has lost its relative dynamical autonomy and become significantly tied to the traditional financial instruments. In this work, we present quantitative arguments in support of this statement.

Since the inception of Bitcoin in 2009, the cryptocurrency market has experienced a rapid surge. Although it used to be a niche and traded unofficially in its early years [2], trading takes place now 24/7 on more than 500 exchanges [3]. The current (October 2022) capitalization of the cryptocurrency market is approximately 1 trillion USD [3], which can

be compared to the largest US tech stocks. During these 12 years of Bitcoin history, there were bubbles and crashes [4–6]. In particular, the foundation of Ethereum in 2015, which allowed for a new application of the blockchain technology in the form of smart contracts, and the subsequent Initial Coin Offer bubble [7] in 2018 reshaped the cryptocurrency market and made it appear in the public eye. A recent bubble in 2021, which was related to the adoption of DeFi (Decentralized Finance) and DEX (Decentralized Exchanges) trading [8], ended with a peak in November 2021, when the total market capitalization was close to 3 trillion dollars. Although there are more than 10,000 cryptocurrencies [3], Bitcoin and Ethereum are currently the most recognizable, and their share in the capitalization of the entire market changed from over 80% in early 2021 to 60% in October 2022 [3].

During these 12 years of development, the characteristics of the cryptocurrency market have changed significantly [9,10]. The properties of the cryptocurrency price return time series are now close to those observed in mature financial markets such as Forex [11]. However, it has long been believed that the cryptocurrency market, which itself is strongly correlated [12–18], especially during the COVID-19 period [19–23], has dynamics that is separate from the traditional financial markets [24–28] and that bitcoin can even serve as a hedge or safe haven [29,30] with respect to the stock market, Forex or the commodity market. The hedging potential of bitcoin was even compared to gold [31–35]. However, the results of many recent studies have changed this paradigm [36–43]. They reported that during the COVID-19 pandemic outburst and the related crash in March 2020 [44,45] the cryptocurrency market and, in particular, bitcoin was highly correlated with the falling stock markets [46–52]. Some studies even noted that this connection still occurred in the market recovery phase in the second half of 2020 [15,53].

The studies referenced above brought therefore rather mixed results and have led to uncertainty as to whether cryptocurrencies can be used for hedging the financial investments. This uncertainty opens space for further research on this topic and our study proceeds exactly in this direction. Our aim is to clarify whether the loss of the cryptocurrency market independence was temporary and primarily caused by the pandemic turmoil or it was only a part of a more general trend towards coupling of this market with the traditional financial markets. We intend to determine how strongly the cryptocurrency price changes are associated with the price changes in the traditional financial markets. To achieve that, the detrended multiscale correlation of the two principal cryptocurrencies: bitcoin (BTC) and ethereum (ETH) versus the traditional financial instruments: stock indices, commodities and currency exchange rates are studied based on high frequency data covering the period from January 2020 to October 2022, which is an extension of the period before 2020 that was analyzed in our earlier study [47]. The year 2022 is particularly interesting, because since the BTC price peak in November 2021, there is a joint bear market on the US tech stocks and the cryptocurrencies for the first time in the existence of the latter. On the basis of this observation, it is likely that there will be some detectable correlation between both markets. The year 2022 is also unique in the history of cryptocurrencies due to the presence of high inflation in the world against which Bitcoin was intended to protect [54–57].

Compared to the other articles dealing with correlations between the cryptocurrency market and traditional financial markets, in our research, the main task is to measure these correlations quantitatively on various time scales and for the fluctuations of various size. It can broaden the market practitioners' perspective on the investing and hedging possibilities by incorporating the fluctuation size as an additional dimension that might be considered while making investment decisions.

2. Data and Methodology

2.1. Data Sources and Preprocessing

In the present study, a data set of 24 financial time series representing contracts for difference (CFDs) from the Dukascopy trading platform [58] is considered. Unlike many other trading platforms, Dukascopy offers freely the high-frequency recordings of many

financial instruments, which is the main reason it has been chosen as the data source. CFDs are characterised by the price movements that are close to the price movements of the original instruments, so we consider them as reliable proxies. Apart from the two highest capitalized cryptocurrencies, BTC and ETH, it includes the most important traditional financial instruments: 12 fiat currencies (Australian dollar—AUD, Canadian dollar—CAD, Swiss franc—CHF, Chinese yuan—CNH, euro—EUR, British pound—GBP, Japanese yen—JPY, Mexican peso—MXN, Norwegian krone—NOK, New Zealand dollar—NZD, Polish zloty—PLN, and South African rand—ZAR), 4 commodities (WTI crude oil—CL, high grade copper—HG, silver—XAG, and gold—XAU), 4 US stock market indices (Nasdaq 100—NQ100, S&P500, Dow Jones Industrial Average—DJI, and Russell 2000—RUSSEL), German stock index—DAX 40—DAX, and the Japanese stock index—Nikkei 225—NIKKEI. All these instruments except for the non-US stock indices are expressed in USD (thus there is no USD in the data set) and their quotes cover a period from 1 January 2020 to 28 October 2022. Each week the quotes were recorded over the whole trading hours, i.e., from Sunday 22:00 to Friday 20:15 with a break between 20:15 and 22:00 each trading day (UTC) [58].

Cumulative log-returns of all the instruments considered are plotted in Figure 1 against time. The original price changes, sampled every $\Delta t = 10$ s, were transformed into logarithmic returns: $r(t_m) = \ln P_i(t_{m+1}) - \ln P_i(t_m)$, where $P_i(t_m)$ is a price quote recorded at time t_m ($m = 1, \ldots, T$) and i stands for a particular financial instrument. We use this particular time interval Δt, because such a data set was available from the source. However, it is satisfactory because it allows us to avoid excessive null returns, which lower reliability of the detrended analysis (see below). In order to obtain the indicative relationships among all the time series, the Pearson correlation coefficient C [59] was calculated for the log-returns $r(t_m)$ from January to October 2022, when the joint bear market mentioned above was observed. A correlation matrix obtained for 24 financial instruments is shown in Figure 2. While the Pearson coefficient is one of the most widely applied measures of time series dependencies (and this is why we also exploited it in our study), the results obtained with it have to be taken with some reserve in our context. This is because the statistical tests that we carried out, i.e., the Jacque-Bera test for normality and the ARCH test for no heteroskedasticity, both rejected the respective null hypotheses with high confidence (p-value < 0.00001), which means that the data under study was both heavy-tailed and heteroskedastic. Obviously, such a result is not surprising, because fat tails of the return distributions and volatility clustering are well-known effects observed in the financial time series [60–62]. Nevertheless, the very long time series that were analysed here and the statistical significance of the obtained results convinced us that the Pearson coefficient can still serve as an effective measure of the time series correlations even in such circumstances. Taking all this into account, a standard naming convention: small ($0.1 \leq C < 0.3$), medium ($0.3 \leq C < 0.$), and large ($0.5 \leq C \leq 1.0$) correlation was used to describe the coefficient values. The strongest cross-correlations (large, $C > 0.6$) can be seen for the stock indices, for BTC and ETH, for AUD, NZD and CAD, for XAU and XAG, and for EUR and GBP. If we consider the cross-correlations between BTC and the traditional instruments, the strongest ones can be seen for NQ100 and S&P500 (medium, $C \approx 0.32$), DJI and RUSSEL (medium, $C \approx 0.29$), DAX (small, $C \approx 0.24$) and NIKKEI, (small $C \approx 0.23$). The Pearson coefficient above 0.1 (small), is observed for BTC on the one side and HG, GBP and EUR, as well as the so-called commodity currencies: AUD, CAD, NZD, MXN, NOK, and ZAR, on the other side. The cross-correlations between ETH and the other instruments are even higher: $C \approx 0.38$ (medium) for SP500 and NQ100, $C \approx 0.35$ (medium) for DJI and RUSSEL, $C \approx 0.29$ (medium) for DAX, and $C \approx 0.27$ (small) for NIKKEI. The same is true for the cross-correlations between ETH and the commodity currencies: $C \approx 0.22$ (small) for AUD, CAD, NZD, $C \approx 0.17$ (small) for MXN, $C \approx 0.13$ (small) for NOK, and $C \approx 0.12$ (small) for ZAR. Among the commodities analyzed here, ETH is the most correlated with HG ($C \approx 0.15$, small). The statistical significance of the coefficient values presented in Figure 2 has been checked by calculating the range: $\bar{C} \pm \sigma_C$, where $\bar{C}$ denotes mean and σ_C denotes standard deviation of C, from 100 independent realisations of the shuffled time series. The

statistically insignificant correlation region is very close to 0 (the third decimal place), all the presented values, except DAX vs. JPY, are thus significant.

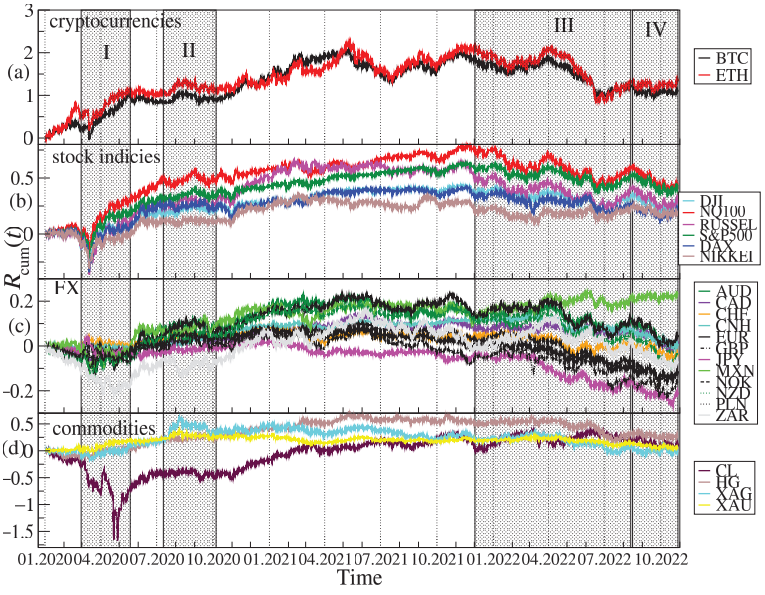


Figure 1. Evolution of the cumulative log-returns of the cryptocurrencies R_{cum} (**a**), the stock market indices (**b**), the fiat currencies (**c**), and the commodities (**d**) over a period from 1 January 2020 to 28 October 2022. Periods for which significant correlations between the cryptocurrencies and the US stock indices are distinguished by grey vertical strips. The most characteristic periods are denoted by Roman numerals: a COVID-19-related crash in March 2020 and a quick bounce in April–May 2020 (period I), new all-time highs of NQ100 and S&P500 and a September 2021 correction (period II), a bear phase in the cryptocurrency and stock markets since November 2021 (period III), and another downward wave of US stock indices after holiday upward correction along with the appreciating USD and inflation fears (period IV).

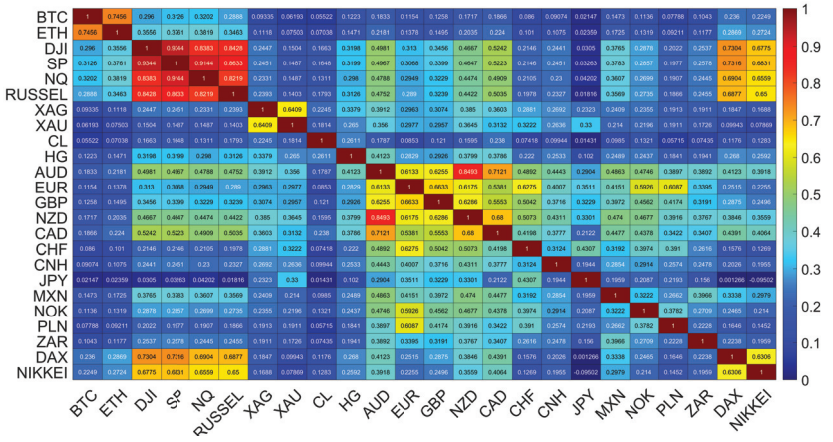


Figure 2. Correlation matrix of Pearson coefficients calculated for all possible pairs of the time series considered in this study (January to October 2022). All the values are statistically significant with p-value < 0.00001, except DAX vs. JPY, where $p = 0.1$.

2.2. The q-Dependent Detrended Correlation Coefficient

Since the Pearson correlation coefficient as a measure has its drawbacks in the case of heavy tails, heteroskedasticity, and multi-scale nonstationarity (which are observed in the cryptocurrency market [9]) the cross-correlations will henceforth be determined using an alternative, better tailored method: the q-dependent detrended cross-correlation coefficient $\rho_q(s)$ [63]. The detrended fluctuation analysis (DFA), which forms a basis for defining $\rho_q(s)$, was developed with the intention to allow for detecting the long-range power-law auto- and cross-correlations that produce trends on different time horizons [64]. Therefore, unlike more traditional methods of trend removal, both DFA and its derivative measures like the coefficient $\rho_q(s)$ can successfully deal with nonstationarity on all scales [65]. $\rho_q(s)$ enables, thus, considering the cross-correlation strength on different time scales and, if used in parallel with the multiscale DFA itself, is able to detect scale-free correlations. Moreover, owing to the parameter q, the correlation analysis can be focused on a specific range of the fluctuation amplitudes.

The steps to calculate $\rho_q(s)$ are as follows. Two possibly nonstationary time series $\{x(i)\}_{i=1,\ldots,T}$ and $\{y(i)\}_{i=1,\ldots,T}$ of length T are divided into M_s boxes of length s starting from its opposite ends and integrated. In each box, the polynomial trend is removed:

$$X_\nu(s,i) = \sum_{j=1}^{i} x(\nu s + j) - P^{(m)}_{X,s,\nu}(i), Y_\nu(s,i) = \sum_{j=1}^{i} y(\nu s + j) - P^{(m)}_{Y,s,\nu}(i), \tag{1}$$

where the polynomials $P^{(m)}$ of order m are applied. In this study $m = 2$ has been selected, which performs well for the financial time series [66]. After this step $2M_s$ boxes are obtained in total with detrended signals. The next step is to calculate the variance and covariance for each of the boxes ν:

$$f^2_{\mathrm{ZZ}}(s,\nu) = \frac{1}{s}\sum_{i=1}^{s}(Z_\nu(s,i))^2, \tag{2}$$

$$f^2_{\mathrm{XY}}(s,\nu) = \frac{1}{s}\sum_{i=1}^{s} X_\nu(s,i) \times Y_\nu(s,i), \tag{3}$$

where Z means X or Y. These quantities are used to calculate a family of the fluctuation functions of order q:

$$F^{(q)}_{\mathrm{ZZ}}(s) = \frac{1}{2M_s}\sum_{\nu=0}^{2M_s-1}\left[f^2_{\mathrm{ZZ}}(s,\nu)\right]^{q/2} \tag{4}$$

$$F^{(q)}_{\mathrm{XY}}(s) = \frac{1}{2M_s}\sum_{\nu=0}^{2M_s-1}\mathrm{sign}\left[f^2_{\mathrm{XY}}(s,\nu)\right]|f^2_{\mathrm{XY}}(s,\nu)|^{q/2}, \tag{5}$$

where a sign function $\mathrm{sign}\left[f^2_{\mathrm{XY}}(s,\nu)\right]$ is preserved in order to secure consistency of results for different qs.

The formula for the q-dependent detrended correlation coefficient is given as follows:

$$\rho^{\mathrm{XY}}_q(s) = \frac{F^{(q)}_{\mathrm{XY}}(s)}{\sqrt{F^{(q)}_{\mathrm{XX}}(s)F^{(q)}_{\mathrm{YY}}(s)}}. \tag{6}$$

For $q = 2$ the above definition can be viewed as a detrended counterpart of the Pearson cross-correlation coefficient C [67]. The parameter q plays the role of a filter suppressing $q < 2$ or amplifying ($q > 2$) the fluctuation variance/covariance calculated in the boxes ν (see Equations (4) and (5)). For $q < 2$ boxes with small fluctuations contribute more to $\rho_q(s)$, while for $q > 2$ the boxes with large fluctuations contribute more. Therefore, by using this measure, it is possible to distinguish the fluctuation size range that is a source of

the observed correlations. In the numerical calculations below, we used our own software in which we implemented the algorithm described above.

3. Results and Discussion

The aforementioned ability of $\rho_q(s)$ to quantify cross-correlation for various time scales (s-dependence) and fluctuation size (q-dependence) is documented in Figures 3 and 4, where the values of $\rho_q(s)$ calculated for BTC and ETH versus the traditional instruments (the same as Figure 1) calculated for the log-returns $r(t_m)$ from January to October 2022 is shown. One can immediately notice two properties: (i) the correlation strength increases with scale s for most financial instruments, and (2) the correlation strength is lower for $q = 4$ (i.e., for large fluctuations Figure 4). These properties, observed here for BTC and ETH versus the other instruments, are typical for the financial markets in general [68,69].

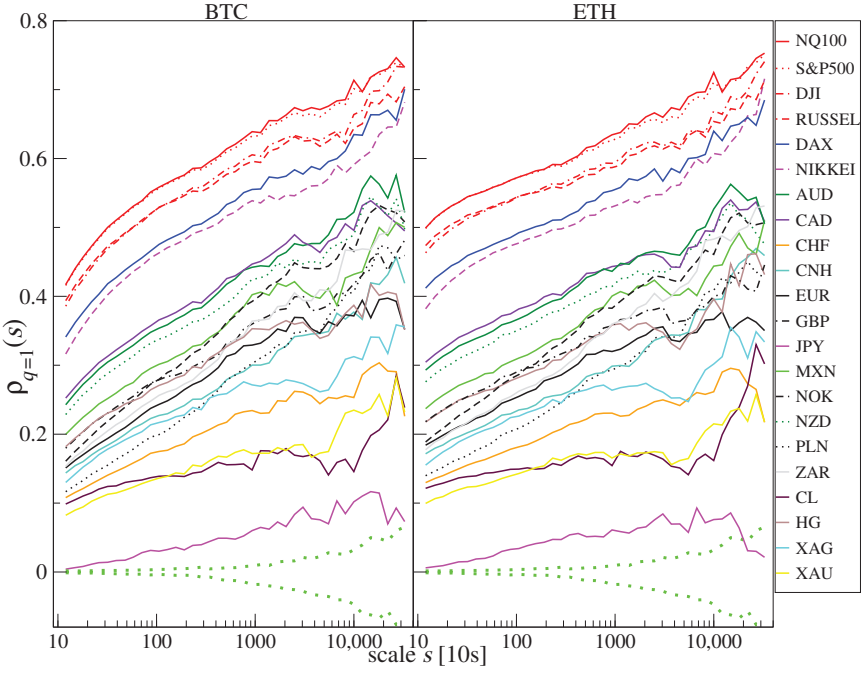


Figure 3. The q-dependent detrended cross-correlation coefficient $\rho_q(s)$ between BTC/USD (**right**) and ETH/USD (**left**) versus selected traditional financial instruments for $q = 1$, which does not favor any specific amplitude range. $\rho_q(s)$ for a range of time scales from $s = 12$ (2 min) to $s = 32,000$ ($\sim$4 trading days) is presented based on data from January to October 2022. The statistically insignificant correlation region (dotted green line) is given as $\pm$ standard deviation of $\rho_q(s)$ calculated from 100 independent realizations of the shuffled time series.

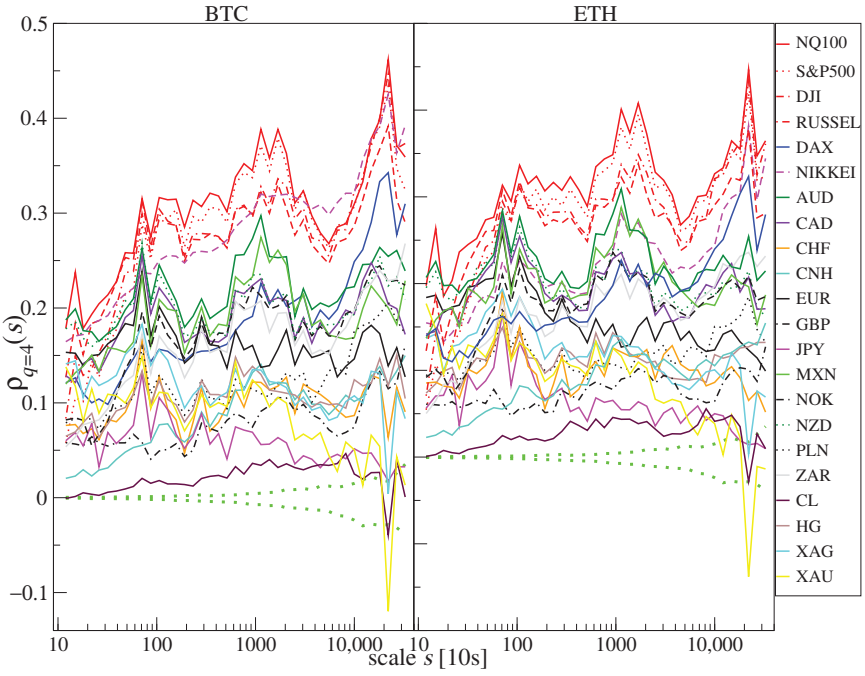


Figure 4. The q-dependent detrended cross-correlation coefficient $\rho_q(s)$ between BTC/USD (**right**) and ETH/USD (**left**) versus selected traditional financial instruments for $q = 4$, which amplifies the large return contributions. $\rho_q(s)$ for a range of time scales from $s = 12$ (2 min) to $s = 32,000$ ($\sim$4 trading days) is presented based on data from January to October 2022. The statistically insignificant correlation region (dotted green line) is given as $\pm$ standard deviation of $\rho_q(s)$ calculated from 100 independent realizations of the shuffled time series.

As in the case of the Pearson coefficient, the strongest cross-correlations measured by $\rho_q(s)$ for $q = 1$ (Figure 3) are BTC and ETH versus the stock indices NQ100 and S&P500. It is different for DJI, RUSSEL, DAX, and NIKKEI, which are less cross-correlated with the cryptocurrencies. What is interesting is that these correlations were stronger for ETH than for BTC, particularly on short time scales. For the shortest scale considered ($s = 12 = 2$ min), they started from $\rho_q(s) \approx 0.5$ in the case of ETH vs. NQ100 and S&P500 and from $\rho_q(s) \approx 0.4$ in the case of BTC vs. NQ100 and S&P500. For the longest scale considered ($s = 32,000 \approx 4$ trading days), the coefficient $\rho_q(s) \approx 0.75$ for BTC and ETH vs. NQ100 and S&P500. The lowest correlations and the weakest scale dependence are observed for JPY, where $\rho_q(s) \approx 0$. XAU and CL are slightly more correlated: $\rho_q(s) \approx 0.1$ and 0.2 for the longest scale s. Above them are XAG and CHF for which the correlations increase with s from 0.1 to 0.3. The cross-correlations for remaining fiat currencies and HG start from $\rho_q(s) \approx 0.15 \div 0.25$ for $s = 12$ and end at $\rho_q(s) \approx 0.35 \div 0.55$ for $s = 32,000$. If we focus on the large fluctuations and apply $q = 4$ (Figure 4), the cross-correlation levels are lower and approximately the same for BTC and ETH. Again, the most significant correlations are observed for the BTC and ETH vs. the US stock indices, but the correlations between BTC and NIKKEI are higher by ~ 0.05 than for ETH and NIKKEI. In the range of scales $4000 \leq s \leq 10,000$ the correlations between BTC and NIKKEI are the strongest. Unlike for $q = 1$, the cross-correlations for BTC and ETH vs. DAX are on the same level as vs. AUD, CAD, MXN, NZD, and NOK. Only for $s \approx 20,000$ the negative values of $\rho_q(s)$ can be found for BTC and ETH versus XAU. The statistical significance of $\rho_q(s)$ in each case was determined by calculating the standard deviation of $\rho_q(s)$ for 100 independent realizations of shuffled time series. This quantity is plotted in Figures 3 and 4 by green dotted lines. It

shows that the detrended cross-correlations are significant for all the instruments in the case of $q = 1$, except for the longest considered scales for JPY, while in the case of $q = 4$, the results for CL and XAU lack statistical significance for the longest considered scales.

Now, a time-dependent analysis of the cross-correlations measured by $\rho_q(s)$ for BTC and ETH versus the traditional financial instruments: AUD, CAD, CHF, CL, DAX, EUR, HG, JPY, MXN, NIKKEI, NQ100, S&P500, XAG and XAU will be presented. A 5-day rolling window with a 1-day step was applied on two time scales: $s = 12$ (2 min) and $s = 360$ (60 min) in order to calculate $\rho_q(s)$. A window of this length corresponds to a trading week. Figures 5 and 6 shows the results obtained for $q = 1$ and Figures 7 and 8 shows the results obtained for $q = 4$. The results for some assets presented in Figures 3 and 4 are omitted here because they are similar to the ones already shown. Our previous study [47] reported that before 2020 the cross-correlations for BTC and ETH versus the traditional instruments were close to 0. In this study, the period starting in 2020 is considered, thus. During these 2.5 unstable years, several important events that affected price changes in the financial markets could be observed.

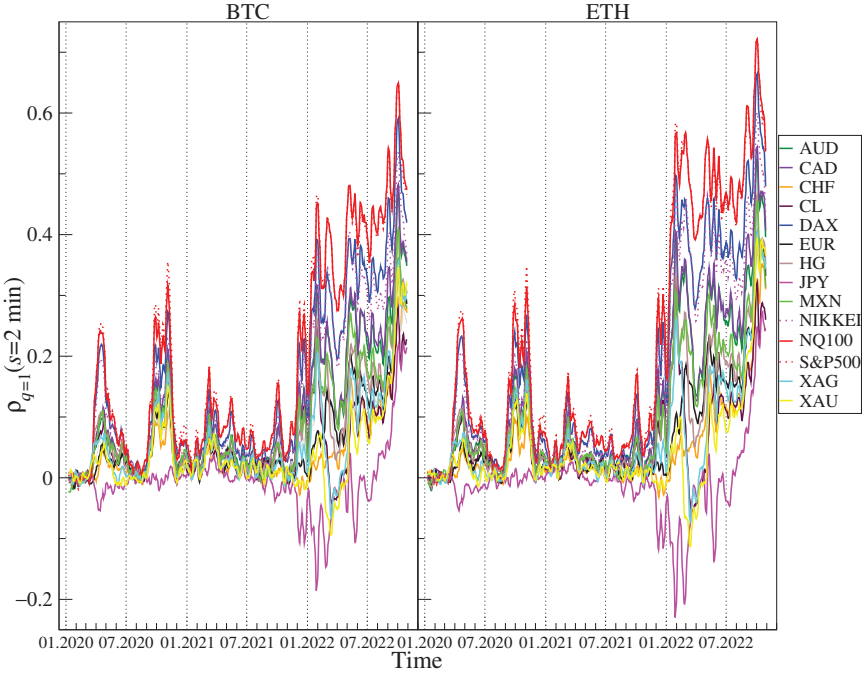


Figure 5. Evolution of the q-dependent detrended cross-correlation coefficient $\rho_q(s)$ with $q = 1$ and $s = 2$ min calculated in a 5-day rolling window with a 1-day step between 1 January 2020 and 28 October 2022 for the price returns of BTC/USD (**left**) and ETH/USD (**right**) versus the selected traditional assets: AUD, CAD, CHF, CL, DAX, EUR, HG, JPY, MXN, NIKKEI, NQ100, S&P500, XAG, and XAU. The statistically insignificant correlations are in the region $\rho_q(s) = 0 \pm 0.001$.

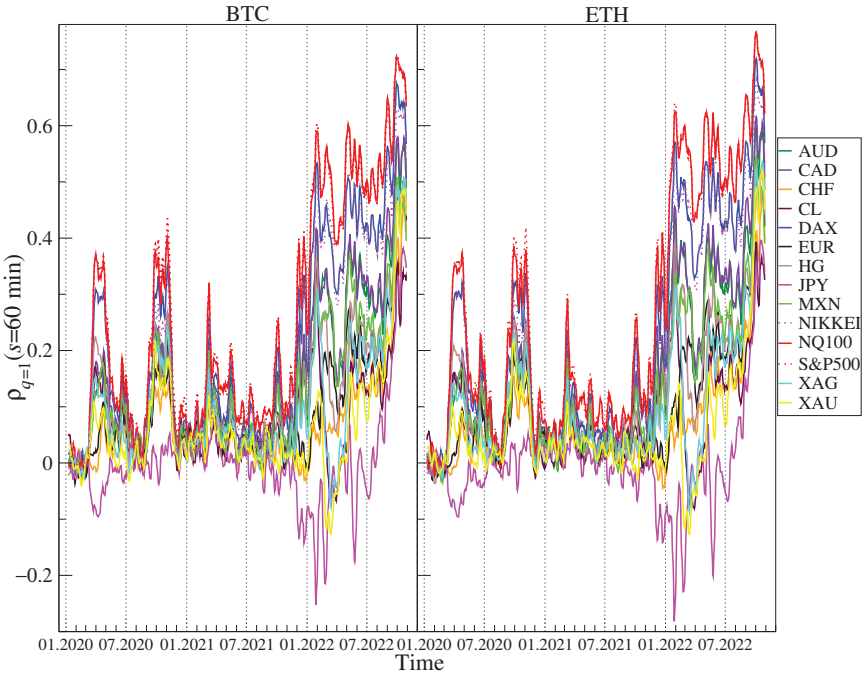


Figure 6. Evolution of the q-dependent detrended cross-correlation coefficient $\rho_q(s)$ with $q = 1$ and $s = 60$ min calculated in a 5-day rolling window with a 1-day step between 1 January 2020 and 28 October 2022 for the price returns of BTC/USD (**left**) and ETH/USD (**right**) versus the selected traditional assets: AUD, CAD, CHF, CL, DAX, EUR, HG, JPY, MXN, NIKKEI, NQ100, S&P500, XAG, and XAU. The statistically insignificant correlations are in the region $\rho_q(s) = 0 \pm 0.01$.

The first event was the outburst of the COVID-19 pandemic that caused a crash in March 2020 on almost all the financial instruments expressed in USD. Only JPY and CHF gained in early March 2020, but later on they also lost value against the US dollar. This price behavior during period I (see Figure 1) resulted in the appearance of a significant positive cross-correlation for BTC and ETH versus the risky assets such as the stock indices, CL, HG, and the commodity currencies (AUD, NZD, CAD, MXN, NOK), which can be seen in Figures 5 and 6. The largest values of $\rho_q(s)$ for BTC and ETH versus the stock indices are observed. In the case of $q = 1$ and $s = 2$ min, $\rho_q(s) \approx 0.2$ and in the case of $q = 1$ and $s = 60$ min, $\rho_q(s) \approx 0.4$. Such strong cross-correlations observed during the general meltdown are not that surprising, but still the joint behavior of the cryptocurrencies and, particularly, the stock indices is noteworthy because it has changed the view that the cryptocurrency market is independent. What is more interesting is the appearance of the even stronger positive cross-correlations for BTC and ETH versus almost all the other instruments except for JPY in the second half of 2020. The strongest cross-correlations are observed again for the stock indices, but very close were also those for CL, HG, XAG, XAU, and the commodity currencies. The highest values, $\rho_q(s) > 0.5$ for $q = 1$ and $s = 60$ min, were observed at the turn of September and October 2020 after the stock and cryptocurrency markets peaked and turned down at the beginning of September 2020. The third period of the significant cross-correlations for BTC and ETH versus the other instruments starts at the beginning of December 2021 after the November 2021's all-time highs on both the cryptocurrency and the US stock markets occurred. $\rho_q(s)$ grew above 0.5 for $q = 1$ and $s = 2$ min and above 0.6 for $q = 1$ and $s = 60$ min in January 2022, when both markets experienced strong declines. BTC and ETH dropped 50% from their peak price down to 33,000 USD and 2300 USD, respectively, S&P500 dropped 8% down to 4200 USD and

NQ100 dropped 18% to 13,700 USD that were their local lows on 22 January 2022. At that time, there were also significant negative cross-correlations for BTC and ETH versus JPY, which is typically considered as a safe currency during the market meltdowns. After local peak of cross-correlations at the beginning of May 2022, when S&P500, NQ100, BTC, and ETH broke into new lows below 4150, 13,000, 35,000, and 2200 levels, respectively, the cross-correlations for BTC and ETH vs. the remaining instruments were significant at approximately the same levels until mid-August 2022, when the holiday upward correction in the US stock indices ended. From that moment on, one can distinguish period IV, when another downward wave of US exchange indices took place, which lasted until mid-October. This was accompanied by a strengthening of the USD, and the EUR/USD exchange rate fell below 1. During this period, the cross-correlation of BTC and ETH with all instruments denominated in USD has started to increase. They were even significantly positive in the case of the least correlated JPY at a level above 0.2 for $s = 2$ min and 0.4 for $s = 60$ min. The cross-correlations peaked in the last week of September, when for NQ100 and S&P500 they first exceeded the level of 0.6 and in the case of $s = 60$ min, they were close to 0.8. They were again slightly higher for ETH.

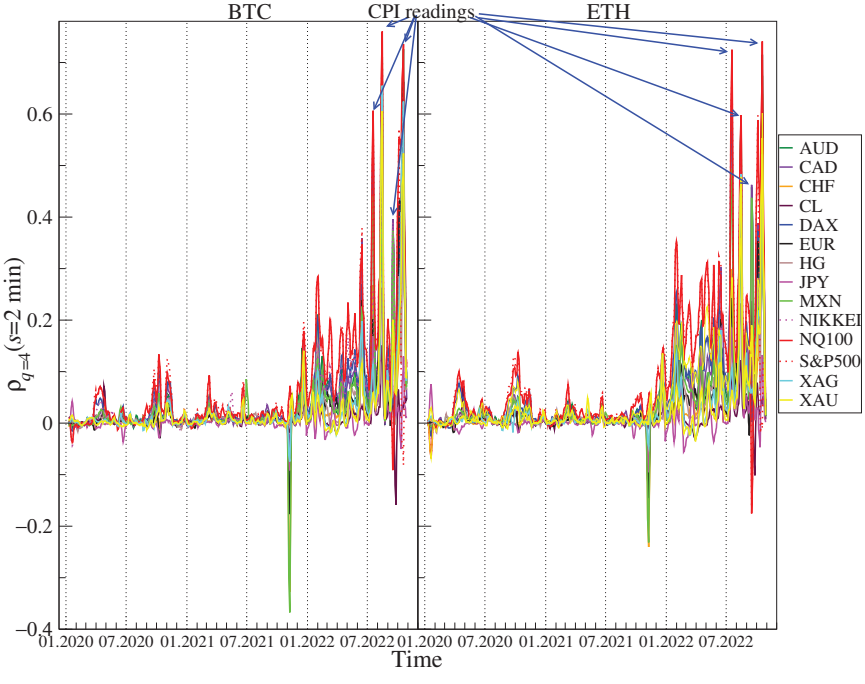


Figure 7. Evolution of the q-dependent detrended cross-correlation coefficient $\rho_q(s)$ with $q = 4$ and $s = 2$ min calculated in a 5-day rolling window with a 1-day step between 1 January 2020 and 28 October 2022 for the price returns of BTC/USD (**left**) and ETH/USD (**right**) versus the selected traditional assets: AUD, CAD, CHF, CL, DAX, EUR, HG, JPY, MXN, NIKKEI, NQ100, S&P500, XAG, and XAU. Higher levels of cross-correlations, associated with the Consumer Price Index (CPI) readings, are marked (see the event description in Figure 9). The statistically insignificant correlations are in the region $\rho_q(s) = 0 \pm 0.001$.

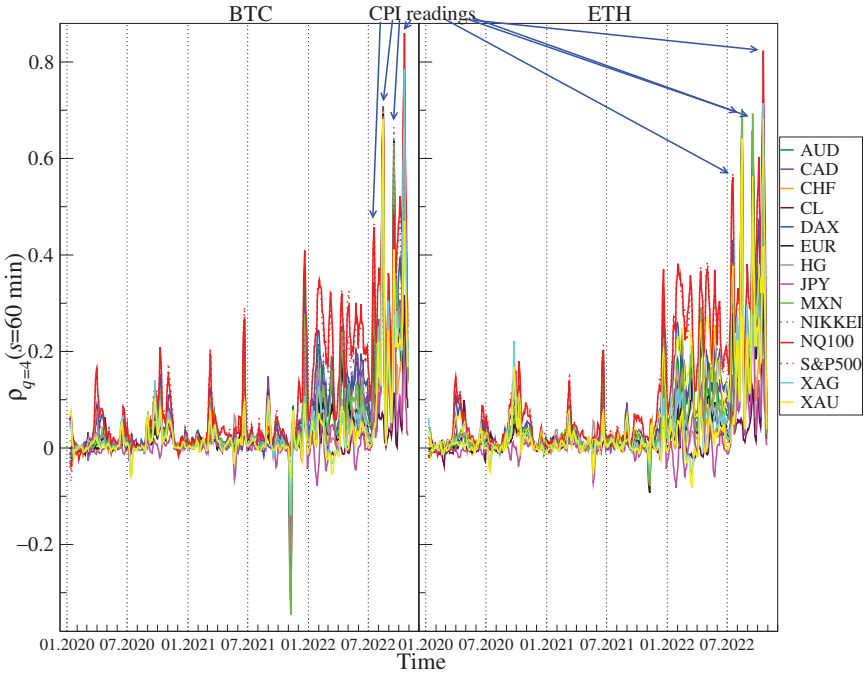


Figure 8. Evolution of the q-dependent detrended cross-correlation coefficient $\rho_q(s)$ with $q = 4$ and $s = 60$ min calculated in a 5-day rolling window with a 1-day step between 1 January 2020 and 28 October 2022 for the price returns of BTC/USD (**left**) and ETH/USD (**right**) versus the selected traditional assets: AUD, CAD, CHF, CL, DAX, EUR, HG, JPY, MXN, NIKKEI, NQ100, S&P500, XAG, and XAU. Higher levels of cross-correlations, associated with the Consumer Price Index (CPI) readings, are marked (see the event description in Figure 9. The statistically insignificant correlations are in the region $\rho_q(s) = 0 \pm 0.01$.

If large returns are considered ($q = 4$, Figures 7 and 8) the detrended cross-correlations for $s = 2$ min remain close to 0 and are statistically insignificant for most of the considered instruments until November 2021, when $\rho_q(s)$ for BTC and ETH versus most currencies, especially MXN, CHF, and, to a lesser extent, for AUD, NZD, EUR and CNH turn negative for short periods of time. As in the case of $q = 1$, the cross-correlations versus the US indices became significantly positive starting from December 2021. What is most interesting is that from July 2022, the cross-correlation levels in some weekly windows exceed those obtained for $q = 1$. For $s = 60$ min they are even higher than 0.8 in the case of NQ100. There are also high correlations of BTC and ETH vs. precious metals: gold and silver. Unlike average fluctuations ($q = 1$), here BTC is slightly more strongly correlated with traditional financial instruments.

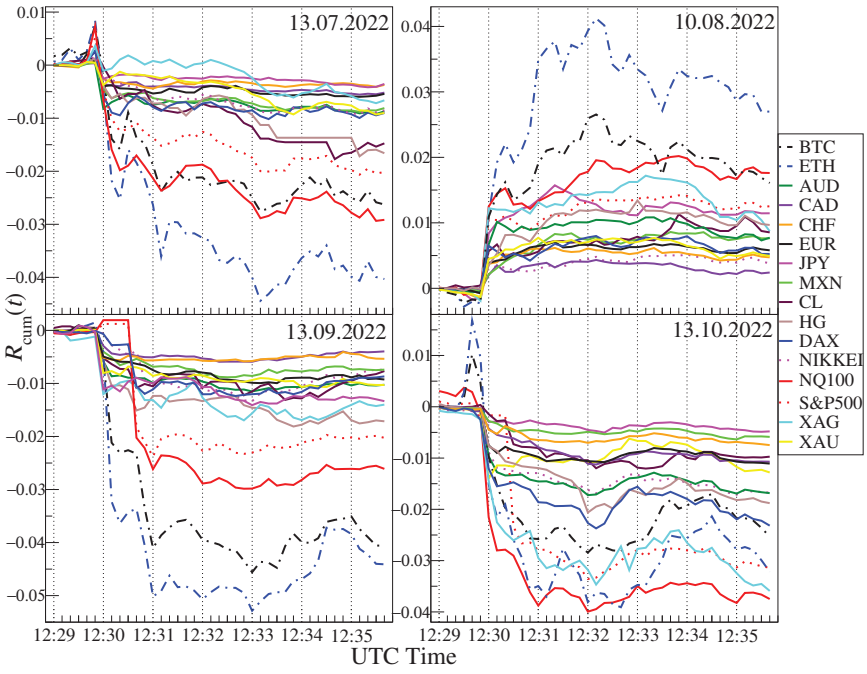


Figure 9. Evolution of the cumulative logarithmic returns R_{cum} of selected financial instruments: BTC, ETH, AUD, CAD, CHF, CL, DAX, EUR, HG, JPY, MXN, NIKKEI, NQ100, S&P500, XAG, and XAU on specific dates, around the publication time of the Consumer Price Index (CPI) report in the US.

After careful checking of the exact start and end dates of the sliding window with increased correlations for $q = 4$ and the time of large fluctuations, it turned out that the correlation of BTC and ETH with traditional financial instruments is (directly or indirectly via other markets) influenced by the CPI inflation data published every month at 12:30 UTC. Cumulative price changes in days during the CPI publication date 13 July 2022, 10 August 2022, 13 September 2022, 10 October 2022, from 12:29 to 12:35 are presented in the Figure 9. It can be clearly seen that in all four cases US tech stocks and cryptocurrencies price changes behave in the same way just after 12:30 UTC. It happened regardless of whether the surprise with the CPI data was positive or negative. In three cases, inflation data was higher than expected and surprised markets negatively, leading to declines. This is especially well visible in the case of 10 October 2022, when apart from US indices, XAG also follows the same trajectory. In the roling windows containing this day, the correlations were the strongest: 0.6 for S&P500, 0.79 for XAG and 0.86 for NQ100 vs. BTC and 0.6 for S&P500, 0.72 for XAG and 0.82 for NQ100 vs. ETH for $s = 60$ min. In one case, 10 August 2022, the inflation was lower than expectations, which resulted in an increase in all instruments. This price behavior means that cryptocurrencies have started to respond to readings from the economy, just like traditional financial instruments. Despite the fact that our analysis of the cross-correlations was carried out by means of the measures, which were unable to detect the direction of influence, it seems natural to infer that these were the economical data releases that had direct or indirect impact on the cryptocurrency market rather than the opposite. That is why we concluded about the direction above.

4. Conclusions

Based on the multiscale cross-correlation analysis performed for the data covering almost the last three years, it can be concluded that the cryptocurrency market dynamics

is substantially tied to the traditional financial markets. Consistently, the most liquid cryptocurrencies, BTC and ETH, cannot serve as a hedge or safe haven for the stock market investments, not only during the turbulent periods like the COVID-19 panic, where this effect is particularly strong, but also during the recent bear market period on tech stocks, which has been accompanied by the parallel bear market on cryptocurrencies. Many observations show that the COVID-19 pandemic may have changed the paradigm that the cryptocurrency market is a largely autonomous market. The recent market developments and the strong US dollar have additionally increased the strength of the cross-correlations between BTC and ETH on the one side and the US tech stocks on the other side. These observations support some earlier findings on the same subject (e.g., [22,70]). In contrast, as the cryptocurrency market was weakly correlated with other markets during 2021, our results cannot support directly a recent hypothesis that the quantitative easing could actually be responsible for these correlations [22]. The existence of links between the global economy and the cryptocurrency market are further strengthened by the reaction of the price changes of BTC and ETH to economic data, such as CPI inflation, in a similar way to traditional financial instruments. These results are able to remove or, at least, to suppress the uncertainty that recent literature on this topic has brought to the cryptocurrency investors. Now it is more clear that the cryptocurrencies can no longer serve as a convenient hedging target for the investors whose purpose is to diversify the risk.

Our study brings a strong indication that the cryptocurrency market has finally become a connected part of the global financial markets after 12 years of the maturation process. Whether such a direction of this market evolution remains in agreement with the early vision of the cryptocurrency creators can be disputed, however. We also face a related question: does the fact that we have got "just another part of the global financial market" deserve devoting so huge amounts of energy to it? Sooner or later this question must be addressed by the policy makers. Nevertheless, what becomes evident now is that it allows the market participants to broaden the spectrum of their investment possibilities.

Among the limitations that might have influenced our study and, subsequently, our conclusions, we have to mention that only two principal cryptocurrencies were studied. Although they are the most influential, the most frequently traded, and widely discussed cryptocurrencies, they by no means define the entire market. It is possible that an analysis that included some less important cryptocurrencies would bring different outcomes. This is especially likely for the marginal cryptocurrencies without any thinkable "fundamental" value, whose dynamics is driven predominantly by extreme speculation. However, as the cryptocurrency market is looked at by the most investors through the lens of BTC and ETH (as their capitalization indicates), this particular limitation does not seem discouraging to us. Currently, these two assets shape the whole cryptocurrency market and we expect them to continue doing it in the nearest future. Another limitation of our analysis is the particular selection of the traditional financial instruments. Indeed, they constitute only a small fraction of the available ones. We are convinced, though, that they are among the most observed and the most influential ones in the context of the global economy, which fully justifies our choice.

A more general observation that the cryptocurrency market has spontaneously coupled to the technological sector of the stock markets by reacting to some trigger provided by the external data inflow resembles analogous effects of the spontaneous emergence of order among the so-far independent degrees of freedom in the various complex systems. However, as complexity allows for flexible reacting of a system to both the external perturbations and internal processes, such effects of ordering in the financial markets have to be eventually counterbalanced by the opposite processes of disordering. Therefore, the market participants must be aware that the inter-market couplings may not last forever and they can significantly be weakened or even removed completely at some point in future. This is why the in-depth studies of the cross-market dependencies have to remain among the principal directions of the cryptocurrency research. Our future work will also deal with energy consumption of the cryptocurrency market.

Author Contributions: Conceptualisation, S.D., J.K. and M.W.; methodology, S.D., J.K. and M.W.; software, M.W.; validation, S.D., J.K. and M.W.; formal analysis, M.W.; investigation, S.D., J.K. and M.W.; resources, M.W.; data curation, M.W.; writing—original draft preparation, M.W.; writing—review and editing, S.D., J.K. and M.W.; visualisation, M.W.; supervision, S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data available freely from Dukascopy [58].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kwapień, J.; Drożdż, S. Physical approach to complex systems. *Phys. Rep.* **2012**, *515*, 115–226.
2. Bitcoin Pizzaday. Available online: https://www.investopedia.com/news/bitcoin-pizza-day-celebrating-20-million-pizza-order/ (accessed on 1 November 2022).
3. CoinMarketCap. Available online: https://coinmarketcap.com (accessed on 1 November 2022).
4. Gerlach, J.C.; Demos, G.; Sornette, D. Dissection of Bitcoin's multiscale bubble history from January 2012 to February 2018. *R. Soc. Open Sci.* **2019**, *6*, 180643.
5. Bellon, C.; Figuerola-Ferretti, I. Bubbles in Ethereum. *Financ. Res. Lett.* **2022**, *46*, 102387. [CrossRef]
6. Zitis, P.I.; Contoyiannis, Y.; Potirakis, S.M. Critical dynamics related to a recent Bitcoin crash. *Int. Rev. Financ. Anal.* **2022**, *84*, 102368. [CrossRef]
7. Aste, T. Cryptocurrency market structure: Connecting emotions and economics. *Digit. Financ.* **2019**, *1*, 5–21. [CrossRef]
8. Maouchi, Y.; Charfeddine, L.; Montasser, G.E. Understanding digital bubbles amidst the COVID-19 pandemic: Evidence from DeFi and NFTs. *Financ. Res. Lett.* **2022**, *47*, 102584. [CrossRef]
9. Wątorek, M.; Drożdż, S.; Kwapień, J.J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* **2021**, *901*, 1–82. [CrossRef]
10. Lahmiri, S.; Bekiros, S.; Bezzina, F. Complexity analysis and forecasting of variations in cryptocurrency trading volume with support vector regression tuned by Bayesian optimization under different kernels: An empirical comparison from a large dataset. *Expert Syst. Appl.* **2022**, *209*, 118349. [CrossRef]
11. 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* **2018**, *28*, 071101. [CrossRef]
12. 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* **2020**, *30*, 023122. [CrossRef]
13. Kaiser, L.; Stöckl, S. Cryptocurrencies: Herding and the transfer currency. *Financ. Res. Lett.* **2020**, *33*, 101214. [CrossRef]
14. Aslanidis, N.; Bariviera, A.F.; Perez-Laborda, A. Are cryptocurrencies becoming more interconnected? *Econ. Lett.* **2021**, *199*, 109725. [CrossRef]
15. Kwapień, J.; Wątorek, M.; Drożdż, S. Cryptocurrency Market Consolidation in 2020-2021. *Entropy* **2021**, *23*, 1674. [CrossRef]
16. Bae, G.; Kim, J.H. Observing Cryptocurrencies through Robust Anomaly Scores. *Entropy* **2022**, *24*, 1643. [CrossRef]
17. James, N.; Menzies, M. Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time. *Nonlinear Dyn.* **2022**, *107*, 4001–4017. [CrossRef]
18. Kakinaka, S.; Umeno, K. Asymmetric volatility dynamics in cryptocurrency markets on multi-time scales. *Res. Int. Bus. Financ.* **2022**, *62*, 101754. [CrossRef]
19. Arouxet, M.B.; Bariviera, A.F.; Pastor, V.E.; Vampa, V. COVID-19 impact on cryptocurrencies: Evidence from a wavelet-based Hurst exponent. *Phys. A* **2022**, *596*, 127170. [CrossRef]
20. Corbet, S.; Hou, Y.G.; Hu, Y.; Larkin, C.; Lucey, B.; Oxley, L. Cryptocurrency liquidity and volatility interrelationships during the COVID-19 pandemic. *Financ. Res. Lett.* **2022**, *45*, 102137. [CrossRef]
21. Nguyen, A.P.N.; Mai, T.T.; Bezbradica, M.; Crane, M. The Cryptocurrency Market in Transition before and after COVID-19: An Opportunity for Investors? *Entropy* **2022**, *24*, 1317. [CrossRef]
22. Kumar, A.; Iqbal, N.; Mitra, S.K.; Kristoufek, L.; Bouri, E. Connectedness among major cryptocurrencies in standard times and during the COVID-19 outbreak. *J. Int. Financ. Mark. Institutions Money* **2022**, *77*, 101523. [CrossRef]
23. Wątorek, M.; Kwapień, J.; Drożdż, S. Multifractal cross-Correlations of bitcoin and ether trading Characteristics in the Post-COVID-19 Time. *Future Internet* **2022**, *14*, 215. [CrossRef]
24. Corbet, S.; Meegan, A.; Larkin, C.; Lucey, B.; Yarovaya, L. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Econ. Lett.* **2018**, *165*, 28–34. [CrossRef]
25. Ji, Q.; Bouri, E.; Gupta, R.; Roubaud, D. Network causality structures among Bitcoin and other financial assets: A directed acyclic graph approach. *Q. Rev. Econ. Financ.* **2018**, *70*, 203–213. [CrossRef]
26. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Signatures of the Crypto-Currency Market Decoupling from the Forex. *Future Internet* **2019**, *11*, 154. [CrossRef]

27. Gil-Alana, L.A.; Abakah, E.J.A.; Rojo, M.F.R. Cryptocurrencies and stock market indices. Are they related? *Res. Int. Bus. Financ.* **2020**, *51*, 101063. [CrossRef]
28. 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* **2020**, *556*, 124759. [CrossRef]
29. Urquhart, A.; Zhang, H. Is Bitcoin a hedge or safe haven for currencies? An intraday analysis. *Int. Rev. Financ. Anal.* **2019**, *63*, 49–57. [CrossRef]
30. Wang, P.; Zhang, W.; Li, X.; Shen, D. Is cryptocurrency a hedge or a safe haven for international indices? A comprehensive and dynamic perspective. *Financ. Res. Lett.* **2019**, *31*, 1–18. [CrossRef]
31. Shahzad, S.J.H.; Bouri, E.; Roubaud, D.; Kristoufek, L.; Lucey, B. Is Bitcoin a better safe-haven investment than gold and commodities? *Int. Rev. Financ. Anal.* **2019**, *63*, 322–330. [CrossRef]
32. Shahzad, S.J.H.; Bouri, E.; Roubaud, D.; Kristoufek, L. Safe haven, hedge and diversification for G7 stock markets: Gold versus bitcoin. *Econ. Model.* **2020**, *87*, 212–224.
33. Bouri, E.; Shahzad, S.J.H.; Roubaud, D.; Kristoufek, L.; Lucey, B. Bitcoin, gold, and commodities as safe havens for stocks:New insight through wavelet analysis. *Q. Rev. Econ. Financ.* **2020**, *77*, 156–164. [CrossRef]
34. Thampanya, N.; Nasir, M.A.; Huynh, T.L.D. Asymmetric correlation and hedging effectiveness of gold & cryptocurrencies: From pre-industrial to the 4th industrial revolution. *Technol. Forecast. Soc. Chang.* **2020**, *159*, 120195.
35. Zhu, X.; Niu, Z.; Zhang, H.; Huang, J.; Zuo, X. Can gold and bitcoin hedge against the COVID-19 related news sentiment risk? New evidence from a NARDL approach. *Resour. Policy* **2022**, *79*, 103098. [CrossRef]
36. Conlon, T.; McGee, R. Safe haven or risky hazard? Bitcoin during the COVID-19 bear market. *Financ. Res. Lett.* **2020**, *35*, 101607. [CrossRef]
37. Kristoufek, L. Grandpa, Grandpa, Tell Me the One About Bitcoin Being a Safe Haven: New Evidence From the COVID-19 Pandemic. *Front. Phys.* **2020**, *8*, 296. [CrossRef]
38. Grobys, K. When Bitcoin has the flu: On Bitcoin's performance to hedge equity risk in the early wake of the COVID-19 outbreak. *Appl. Econ. Lett.* **2021**, *28*, 860–865. [CrossRef]
39. James, N.; Menzies, M.; Chan, J. Changes to the extreme and erratic behaviour of cryptocurrencies during COVID-19. *Phys. A* **2021**, *565*, 125581. [CrossRef]
40. Barbu, T.C.; Boitan, I.A.; Cepoi, C.. Are cryptocurrencies safe havens during the COVID-19 pandemic? A threshold regression perspective with pandemic-related benchmarks. *Econ. Bus. Rev.* **2022**, *8*, 22–49. [CrossRef]
41. Jareno, F.; Gonzàlez, M.D.L.O.; Belmonte, P. Asymmetric interdependencies between cryptocurrency and commodity markets: The COVID-19 pandemic impact. *Quant. Financ. Econ.* **2022**, *6*, 83–112. [CrossRef]
42. Kakinaka, S.; Umeno, K. Cryptocurrency market efficiency in short- and long-term horizons during COVID-19: An asymmetric multifractal analysis approach. *Financ. Res. Lett.* **2022**, *46*, 102319. [CrossRef]
43. Evrim Mandaci, P.; Cagli, E.C. Herding intensity and volatility in cryptocurrency markets during the COVID-19. *Financ. Res. Lett.* **2022**, *46*, 102382. [CrossRef]
44. Baker, S.R.; Bloom, N.; Davis, S.J.; Kost, K.; Sammon, M.; Viratyosin, T. The Unprecedented Stock Market Reaction to COVID-19. *Rev. Asset Pricing Stud.* **2020**, *10*, 742–758. [CrossRef]
45. Yarovaya, L.; Matkovskyy, R.; Jalan, A. The COVID-19 black swan crisis: Reaction and recovery of various financial markets. *Res. Int. Bus. Financ.* **2022**, *59*, 101521. [CrossRef]
46. 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. Financ.* **2020**, *54*, 101248. [CrossRef]
47. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Stanisz, T.; Wątorek, M. Complexity in economic and social systems: Cryptocurrency market at around COVID-19. *Entropy* **2020**, *22*, 1043. [CrossRef]
48. Caferra, R.; Vidal-Tomas, D. Who raised from the abyss? A comparison between cryptocurrency and stock market dynamics during the COVID-19 pandemic. *Financ. Res. Lett.* **2021**, *43*, 101954. [CrossRef]
49. Jiang, Y.; Lie, J.; Wang, J.; Mu, J. Revisiting the roles of cryptocurrencies in stock markets: A quantile coherency perspective. *Econ. Model.* **2021**, *95*, 21–34. [CrossRef]
50. James, N. Dynamics, behaviours, and anomaly persistence in cryptocurrencies and equities surrounding COVID-19. *Phys. A* **2021**, *570*, 125831. [CrossRef]
51. Elmelki, A.; Chaâbane, N.; Benammar, R. Exploring the relationship between cryptocurrency and S&P500: evidence from wavelet coherence analysis. *Int. J. Blockchains Cryptocurrencies* **2022**, *3*, 256–268.
52. Wang, P.; Liu, X.; Wu, S. Dynamic Linkage between Bitcoin and Traditional Financial Assets: A Comparative Analysis of Different Time Frequencies. *Entropy* **2022**, *24*, 1565. [CrossRef]
53. Balcilar, M.; Ozdemir, H.; Agan, B. Effects of COVID-19 on cryptocurrency and emerging market connectedness: Empirical evidence from quantile, frequency, and lasso networks. *Phys. A* **2022**, *604*, 127885. [CrossRef]
54. Conlon, T.; Corbet, S.; McGee, R.J. Inflation and cryptocurrencies revisited: A time-scale analysis. *Econ. Lett.* **2021**, *206*, 109996. [CrossRef]
55. Choi, S.; Shin, J. Bitcoin: An inflation hedge but not a safe haven. *Financ. Res. Lett.* **2022**, *46*, 102379. [CrossRef]

56. James, N.; Menzies, M.; Chin, K. Economic state classification and portfolio optimisation with application to stagflationary environments. *Chaos, Solitons Fractals* **2022**, *164*, 112664. [CrossRef]
57. Phochanachan, P.; Pirabun, N.; Leurcharusmee, S.; Yamaka, W. Do Bitcoin and Traditional Financial Assets Act as an Inflation Hedge during Stable and Turbulent Markets? Evidence from High Cryptocurrency Adoption Countries. *Axioms* **2022**, *11*, 339. [CrossRef]
58. Dukascopy. Available online: https://www.dukascopy.com/swiss/pl/cfd/range-of-markets/ (accessed on 1 November 2022).
59. Pearson, K. Note on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* **1895**, *58*, 240–242.
60. Gopikrishnan, P.; Plerou, V.; Nunes Amaral, L.A.; Meyer, M.; Stanley, H.E. Scaling of the distribution of fluctuations of financial market indices. *Phys. Rev. E* **1999**, *60*, 5305–5316. [CrossRef]
61. Cont, R. Empirical properties of asset returns: Stylized facts and statistical issues. *Quant. Financ.* **2001**, *1*, 223–236. [CrossRef]
62. Gabaix, X.; Gopikrishnan, P.; Plerou, V.; Stanley, H.E. A theory of power-law distributions in financial market fluctuations. *Nature* **2003**, *423*, 267–270. [CrossRef]
63. Kwapień, J.; Oświęcimka, P.; Drożdż, S. Detrended fluctuation analysis made flexible to detect range of cross-correlated fluctuations. *Phys. Rev. E* **2015**, *92*, 052815. [CrossRef]
64. Peng, C.K.; Buldyrev, S.V.; Havlin, S.; Simons, M.; Stanley, H.E.; Goldberger, A.L. Mosaic organization of DNA nucleotides. *Phys. Rev. E* **1994**, *49*, 1685–1689. [CrossRef]
65. Jiang, Z.Q.; Xie, W.J.; Zhou, W.X.; Sornette, D. Multifractal analysis of financial markets: A review. *Rep. Prog. Phys.* **2019**, *82*, 125901. [CrossRef]
66. Oświęcimka, P.; Drożdż, S.; Kwapień, J.; Górski, A. Effect of detrending on multifractal characteristics. *Acta Phys. Pol. A* **2013**, *123*, 597–603. [CrossRef]
67. Zebende, G. DCCA cross-correlation coefficient: Quantifying level of cross-correlation. *Phys. A* **2011**, *390*, 614–618. [CrossRef]
68. 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.* **2019**, *98*, 2349–2364. [CrossRef]
69. Wątorek, M.; Drożdż, S.; Oświęcimka, P.; Stanuszek, M. Multifractal cross-correlations between the world oil and other financial markets in 2012–2017. *Energy Econ.* **2019**, *81*, 874–885. [CrossRef]
70. Zhang, Y.J.; Bouri, E.; Gupta, R.; Ma, S.J. Risk spillover between Bitcoin and conventional financial markets: An expectile-based approach. *North Am. J. Econ. Financ.* **2021**, *55*, 101296. [CrossRef]

entropy

MDPI

Article

Investigating Dynamical Complexity and Fractal Characteristics of Bitcoin/US Dollar and Euro/US Dollar Exchange Rates around the COVID-19 Outbreak

Pavlos I. Zitis [1], Shinji Kakinaka [2], Ken Umeno [2], Michael P. Hanias [3], Stavros G. Stavrinides [3] and Stelios M. Potirakis [1,4,*]

[1] Department of Electrical and Electronics Engineering, University of West Attica, Ancient Olive Grove Campus, GR-12241 Aigaleo, Greece
[2] Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Sakyo, Kyoto 606-8501, Japan
[3] Department of Physics, International Hellenic University, GR-65404 Kavala, Greece
[4] Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, Metaxa and Vasileos Pavlou, GR-15236 Penteli, Greece
* Correspondence: spoti@uniwa.gr

Abstract: This article investigates the dynamical complexity and fractal characteristics changes of the Bitcoin/US dollar (BTC/USD) and Euro/US dollar (EUR/USD) returns in the period before and after the outbreak of the COVID-19 pandemic. More specifically, we applied the asymmetric multifractal detrended fluctuation analysis (A-MF-DFA) method to investigate the temporal evolution of the asymmetric multifractal spectrum parameters. In addition, we examined the temporal evolution of Fuzzy entropy, non-extensive Tsallis entropy, Shannon entropy, and Fisher information. Our research was motivated to contribute to the comprehension of the pandemic's impact and the possible changes it caused in two currencies that play a key role in the modern financial system. Our results revealed that for the overall trend both before and after the outbreak of the pandemic, the BTC/USD returns exhibited persistent behavior while the EUR/USD returns exhibited anti-persistent behavior. Additionally, after the outbreak of COVID-19, there was an increase in the degree of multifractality, a dominance of large fluctuations, as well as a sharp decrease of the complexity (i.e., increase of the order and information content and decrease of randomness) of both BTC/USD and EUR/USD returns. The World Health Organization (*WHO*) announcement, in which COVID-19 was declared a global pandemic, appears to have had a significant impact on the sudden change in complexity. Our findings can help both investors and risk managers, as well as policymakers, to formulate a comprehensive response to the occurrence of such external events.

Keywords: COVID-19; bitcoin; cryptocurrencies; forex market; complexity; entropy; multifractal analysis; complex systems; financial crisis; econophysics

Citation: Zitis, P.I.; Kakinaka, S.; Umeno, K.; Hanias, M.P.; Stavrinides, S.G.; Potirakis, S.M. Investigating Dynamical Complexity and Fractal Characteristics of Bitcoin/US Dollar and Euro/US Dollar Exchange Rates around the COVID-19 Outbreak. *Entropy* **2023**, *25*, 214. https://doi.org/10.3390/e25020214

Academic Editors: Stanisław Drożdż, Jarosław Kwapień and Marcin Wątorek

Received: 8 December 2022
Revised: 15 January 2023
Accepted: 18 January 2023
Published: 22 January 2023

1. Introduction

Financial markets are widely recognized as typical examples of complex dynamical systems [1]. Asset prices are created by a large number of nonlinear interactions between heterogeneous agents and complex events occurring in the external environment [2,3]. The properties observed in financial time series such as nonlinearity, long-range dependence [4,5], volatility clustering [6], fat tails [7,8], asymmetry [9], chaos [10,11], fractals and multifractals [12,13], and self-similarity [14] have attracted the interest of many scientists from different fields. In the last three decades, physicists have studied and developed models to understand the behaviors and interactions in financial systems, establishing an interdisciplinary research field known as Econophysics [15–17]. This term first appeared in the published article by Stanley et al. [18] when analyzing the Dow Jones index; they found

that stock returns followed a power law distribution. Since then, significant progress has been made in the field of Econophysics [19].

The dynamics of financial markets are difficult to understand not only because of the complexity of their internal elements but also because of the many intractable external factors acting on them. A recent example of an external factor causing disruptions in global financial markets is the outbreak of the COVID-19 pandemic. At its roots, the COVID-19 crisis is not a financial or economic crisis, it is a health crisis. Nevertheless, through its effects on supply and demand conditions, it evolved rapidly to a large-scale financial and economic crisis. In March 2020, the US stock market hit the circuit breaker mechanism four times in a period of ten days. Since its inception in 1987, the breaker has only ever been triggered once, in 1997. At the same time as the US crash, stock markets in Asia and Europe plunged also. More specifically, Japan's stock market fell by more than 20%, while the UK's main index, FTSE, fell by about 10.87% on 12 March 2020. Additionally, during the pandemic period, most economies experienced exchange rate volatility and currency depreciation due to capital outflows and market sentiments. Typical examples are the Australian dollar hitting a 17-year low of AUD 0.59215 and the New Zealand dollar hitting an 11-year low of NZD 0.5850. Furthermore, the price of gold dropped about 3.53%. It is worth noting that although gold is considered a strong safe haven for most developed markets during financial crises, there are findings showing that during the pandemic it was a weak safe haven for investors in the stock market [20]. The impact of COVID-19 affected even the newer asset classes such as cryptocurrencies. The declines in value of the three leading cryptocurrencies (Bitcoin, Ethereum, and Litecoin) exceeded 50% during the pandemic period.

The exchange rate is crucial for maintaining an economy's external stability. As exchange rate directly associates with foreign debt, capital flows, trade balance, and export competitiveness, maintaining a stable exchange rate is one of the policymakers' major concerns. On the other hand, several researchers argue that specific characteristics of cryptocurrencies, including the independence from monetary policy and the non-correlation with traditional assets, increase their resilience during crisis periods such as the recent pandemic crisis [21–23]. However, there is also the opposing view which argues that monetary policy has a significant impact on the price of cryptocurrencies as well as that the cryptocurrencies do not have zero correlation with other asset classes. For example, Chaoqun Ma et al. [24] found a strong response of Bitcoin prices to unexpected monetary policy actions, while Khanh Quoc Nguyen [25] found that S&P 500 returns significantly affected Bitcoin returns during the pandemic period. Therefore, it is concluded that understanding the pandemic's impact and the possible changes it caused in the cryptocurrency and foreign exchange markets is crucial for both investors and risk managers as well as policymakers.

Particularly useful conclusions about the effects of COVID-19 on financial markets can be obtained by studying changes in the multifractality and complexity of financial time series during the period around the COVID-19 outbreak. In the field of Econophysics, extensive research has been conducted on these topics. For example, Mnif et al. [26] utilized the multifractal detrended fluctuation analysis (MF-DFA) approach to investigate the degree of cryptocurrency efficiency before and after the COVID-19 outbreak using a limited time period, until 19 May 2020. Their results indicated that the pandemic outbreak positively affected the efficiency of the five cryptocurrencies that they studied. Naeem et al. [9] examined the asymmetric efficiency of the cryptocurrencies Bitcoin, Ethereum, Litecoin, and Ripple, using 1-h data. In their analysis, the authors utilized the A-MF-DFA and their results showed that the price of cryptocurrencies exhibited significant asymmetric multifractality. Moreover, they found that uptrends showed stronger multifractality than downtrends. Additionally, applying the time-varying deficiency measure, they found that the pandemic outbreak had a negative impact on the efficiency of the cryptocurrencies that they analyzed. Kakinaka and Umeno [27], applying the A-MF-DFA, examined the asymmetric multifractality along with the market efficiency of two main

cryptocurrencies (Bitcoin, Ethereum) during the pandemic period, taking into consideration different investment horizons. Their empirical results showed that the outbreak of COVID-19 affected the efficiency of the two cryptocurrencies differently in the short- and long-term horizons. More specifically, after the outbreak of COVID-19, Bitcoin and Ethereum in the short term exhibited stronger multifractality, while in the long term exhibited weaker multifractality. In addition, they studied the asymmetric market patterns between small and large price fluctuations and between upward and downward trends. These results confirmed that the outbreak caused a significant change in the level of asymmetry in cryptocurrency markets. Aslam et al. [28] applied the MF-DFA to study the efficiency of foreign exchange markets during the initial period of the COVID-19 pandemic. In their analysis, they used high-frequency data of six major currencies, during the period from 1 October 2019 to 31 March 2020. Before calculating the MF-DFA, they examined the inner dynamics of multifractality through seasonal and trend decomposition using loess. Their results indicated that efficiency of foreign exchange markets during the COVID-19 outbreak declined. Mensi et al. [29] examined the effect caused by the COVID-19 crisis on the pricing efficiency and asymmetric multifractality of major asset classes (US Treasury bond, US dollar index, S&P500, Brent oil, Gold, and Bitcoin). In their article, they applied the permutation entropy on intraday data from 30 April 2019 to 13 May 2020. Their results indicated that after the outbreak of COVID-19, the efficiency of all asset classes that they studied was deteriorated, and in most cases this deterioration was significant. In addition, using the A-MF-DFA, they found evidence of asymmetric multifractality in all markets. Drożdż et al. [30] studied the complexity of the cryptocurrency market in the period around the COVID-19 outbreak from three different perspectives. Their findings showed that throughout the time period analyzed, the returns of exchange rates were multifractal with intermittent signatures of bifractality that can be associated with the periods where the market was more volatile.

Lahmiri and Bekiros [31] investigated the time-varying characteristics of the informational efficiency in sixteen international stock markets and forty-five cryptocurrency markets before and during the pandemic period using the approximate entropy and Largest Lyapunov Exponent. Their results indicated that cryptocurrencies exhibited more irregularity and more instability during the pandemic period compared to international stock markets. Additionally, Lahmiri and Bekiros [32], applying Rényi entropy, analyzed the multiscale entropy function in the return time series of S&P500, Brent, WTI, Gas, Silver, Gold, Bitcoin, and VIX. Additionally, they analyzed the information sharing between these markets by estimating mutual information. Their results from Rényi entropy showed that for all market indices, disorder and randomness were more concentrated in less probable events. In addition, their results from the mutual information indicated that the information sharing network between markets has changed during the pandemic period. Wang J. and Wang X. [33] investigated the market efficiency of the S&P 500 Index, Gold, Bitcoin, and US Dollar Index during the extreme event of the COVID-19 pandemic using a multiscale entropy-based method. Their results indicated that, at all scales, the four markets' efficiency decreased abruptly and persistently during the period from February to March 2020. Market efficiency decreased the most in the S&P 500 Index and the least in the Bitcoin market. Additionally, their results showed that Bitcoin market efficiency was more resilient than the others during the extreme event. Fernandes et al. [34] investigated the informational efficiency and price disorder of five main cryptocurrencies (Ethereum, Bitcoin, Cardano, XRP, and BNB) before and during the pandemic period. In their article, the authors applied the permutation entropy and Fisher information measure to construct the Shannon–Fisher causality plane in order to map the cryptocurrencies and their respective locations in a two-dimensional plane. Their results indicated that all cryptocurrencies exhibited high but slightly varying informational efficiency during both periods. Additionally, their results showed that Cardano was the most efficient cryptocurrency. Kim and Lee [35] investigated the evolution of the complexity of the cryptocurrency market and analyzed the properties from the previous upward trend market in 2017 against the COVID-19 pandemic. In their article, the authors used three popular measures of complexity based on the nonlinear

analysis: sample entropy, approximate entropy, and Lempel–Ziv complexity. They studied the market complexity/unpredictability for forty-three cryptocurrency prices. They found that sample entropy, approximate entropy, and Lempel–Ziv complexity metrics of all markets could not generalize the COVID-19 effect of the complexity due to different patterns. Nevertheless, market unpredictability increased by the ongoing health crisis. Olbryś and Majewska [36] applied sample entropy to evaluate changes in the regularity of returns of thirty-six U.S. and European stock market indices during periods of uncertainty. Specifically, the authors studied the period of the Global Financial Crisis as well as the period of the COVID-19 pandemic. Their results showed that entropy decreased during the periods of turbulence, indicating that the regularity and predictability of stock market indices returns increased during these periods. In the field of Econophysics, the study of the complexity and multifractality of financial time series during the pandemic is a challenging topic. However, to the best of our knowledge, until now there has not been an in-depth comparative analysis of the effects of the pandemic on the complexity and fractal characteristics of the returns of two completely different currencies, such as BTC/USD and EUR/USD, that play a key role in the modern financial system.

In this article, we present a study of the temporal evolution of the multifractality and complexity of BTC/USD and EUR/USD returns for the period before and after the *WHO* announcement that declared COVID-19 a global pandemic (i.e., 11 March 2020). We chose to analyze and compare the effects of the pandemic on the two most representative currencies from the cryptocurrency and forex markets, respectively. Although these two markets are completely different from each other, they play a significant role in the modern financial system. More specifically, we applied the A-MF-DFA to investigate the temporal evolution of the asymmetric multifractal spectrum parameters $(\alpha_0,\ \Delta\alpha,\ A)$ before and after the outbreak of the pandemic. Although there are numerous studies that have followed a similar approach for the study of financial time series (e.g., [37–39]), as far as we are able to know, this is the first time that the temporal evolution of the specific parameters has been applied to BTC/USD and EUR/USD returns to study the period before and during COVID-19. At this point, it is important to mention that the analysis of the multifractal properties of financial time series has a wide contribution to the field of finance. For example, multifractality can be used to obtain better forecasts of tail risk as demonstrated by Batten et al. [40]. In addition, we examined the temporal evolution of four popular complexity measures. Although approximate and sample entropies are quite common for financial time series analysis [31,35,41], we chose to use Fuzzy entropy as it is considered as an upgraded alternative of approximate and sample entropy for evaluating the complexity, specifically for short time series contaminated by noise [42]. In combination, we chose to use the Shannon entropy as the standard information measure and Tsallis entropy as its non-extensive generalization, very closely related to multifractality. Additionally, we used another complexity measure, Fisher information. In financial data analysis, the application of Fisher information is very widespread for the construction of the Shannon–Fisher causality plane [34,43]. In the present article, we chose to investigate the temporal evolution of Fisher information as we believe that it can reveal useful elements for the evolution of the complexity of the dynamical system, providing a "mirror image" of the evolution of entropies, but also presenting the key difference of its so-called "locality" property (see Section 2.4). Our study attempts to provide a complete picture of the pandemic's impact in terms of the dynamical change of the complexity and the fractal characteristics of the two currencies. Additionally, our results provide useful conclusions about the behavior of two very different currencies during uncertainty periods. At the same time, interesting conclusions are drawn about the impact of *WHO* announcements and the reaction of investors to external events such as the pandemic. Our findings can help both investors and risk managers, as well as policymakers, to formulate a comprehensive response to the occurrence of such external events.

2. Materials and Methods

This section briefly presents the asymmetric multifractal detrended fluctuation analysis approach (Section 2.1), as well as key concepts of multifractal spectrum parameters (Section 2.1.1). Additionally, we present key notions and formulae related to Fuzzy entropy (Section 2.2), Tsallis entropy (Section 2.3), and Fisher information measure (Section 2.4).

2.1. Asymmetric Multifractal Detrended Fluctuation Analysis (A-MF-DFA)

The A-MF-DFA extends the MF-DFA method by considering positive and negative market trends [44,45]. First, the profile time series of each return time series $\{x_j : j = 1, \ldots, N\}$ are calculated as $X(t) = \sum_{j=1}^{t}(x_j - \overline{x})$ for $t = 1, \ldots, N$, where $\overline{x}$ is the average of the entire return time series. Then, the profile time series and the return time series are both divided into $N_n = \lfloor N/n \rfloor$ non-overlapping segments of length n. In case N is not a multiple of n, we repeat the division initially from the other end of the time series to take into account all the available data, making a total of $2N_n$ segments for both the profile and the return time series.

Next, the local trend of the profile series $\widetilde{X}_v(i)$, $i = 1, \ldots, n$ is calculated for each segment $v = 1, \ldots, 2N_n$, by fitting a least-square polynomial of degree 2 in order to detrend the corresponding profile $X_v(i)$, $i = 1, \ldots, n$. For the return time series, the local linear trend for each segment is also calculated to determine whether the return time series show an uptrend or downtrend. The different trends depend on the sign of each local slope $b_{n,v} \neq 0$, where $b_{n,v}$ represents the coefficient of the linear trend for segment v at scale n [27]. If $b_{n,v} > 0$ $(b_{n,v} < 0)$, the return time series have an upward (downward) trend within the vth segment.

Then, we define the residual variance as follows:

$$F^2(n, v) = \frac{1}{n}\sum_{i=1}^{n}\Big(X_v(i) - \widetilde{X}_v(i)\Big)^2. \tag{1}$$

By taking the average over corresponding segments, we can obtain the asymmetric qth order average fluctuation functions, which are then calculated by taking the average over the corresponding segments:

$$F_q^+(n) = \left\{\frac{1}{M^+}\sum_{v=1}^{2N_n}\frac{1+\mathrm{sgn}(b_{n,v})}{2}\Big[F^2(n, v)\Big]^{\frac{q}{2}}\right\}^{\frac{1}{q}}, \tag{2}$$

$$F_q^-(n) = \left\{\frac{1}{M^-}\sum_{v=1}^{2N_n}\frac{1-\mathrm{sgn}(b_{n,v})}{2}\Big[F^2(n, v)\Big]^{\frac{q}{2}}\right\}^{\frac{1}{q}}, \tag{3}$$

where $M^+ = \sum_{v=1}^{2N_n}(1+\mathrm{sgn}(b_{n,v}))/2$ and $M^- = \sum_{v=1}^{2N_n}(1-\mathrm{sgn}(b_{n,v}))/2$ are the number of total segments with directional trends. Note that for all $v = 1, \ldots, 2N_n$, $M^+ + M^- = 2N_n$ holds. Therefore, the qth order average fluctuation functions for the overall trend is written as:

$$F_q(n) = \left\{\frac{1}{2N_n}\sum_{v=1}^{2N_n}\Big[F^2(n, v)\Big]^{\frac{q}{2}}\right\}^{1/q}. \tag{4}$$

The calculation is repeated to find the fluctuation function for all box sizes n. If long-range power-law correlations are present, the function will increase with n as a power-law $F_q(n) \sim n^{h(q)}$. The scaling exponent $h(q)$, namely, the generalized Hurst exponent, is calculated by estimating the slope of the linear regression of $\log\big(F_q(n)\big)$ versus $\log(n)$. The asymmetric generalized exponents $h^+(q)$ and $h^-(q)$ are calculated in a similar way from the relationship $F_q^+(n) \sim n^{h^+(q)}$ and $F_q^-(n) \sim n^{h^-(q)}$. In this study, we consider n ranging from 8 to $N/4$ for the log-log linear regression to estimate the asymmetric generalized Hurst exponents.

2.1.1. Asymmetric Multifractal Spectrum Parameters

The multifractal characteristics of time series can be described not only by the generalized Hurst exponent $H(q)$ but also by the multifractal scaling exponent $\tau(q)$, and their relationship can be expressed as $\tau(q) = qH(q) - 1$. In the case that $\tau(q)$ and q are linearly related, the analyzed time series is monofractal. In the case that $\tau(q)$ and q have a nonlinear relationship, the analyzed time series is multifractal. Additionally, it is significant to note that the stronger their nonlinear relationship is, the stronger the multifractal characteristics are [46].

Moreover, using the multifractal (singularity) spectrum $f(\alpha)$ can also describe multifractional characteristics of time series. The multifractal spectrum is obtained by applying the first-order Legendre transform [39,46]:

$$\alpha = d\tau(q)/dq, \tag{5}$$

$$f(\alpha) = q\alpha - \tau(q), \tag{6}$$

where α is the singularity strength (also known as the Hölder exponent) that characterizes singularities in the time series. The interpretation of α is as follows: If $\alpha = 1$, then the distribution of the time series data is uniform. If $\alpha < 1$, then the singularity degree is larger. On the other hand, if $\alpha > 1$, then the singularity degree is smaller. The multifractal spectrum $f(\alpha)$ denotes the singularity content [46,47].

To analyze and make a solid understanding of the multifractal characteristics of a time series, a set of the asymmetric multifractal spectrum parameters $(\alpha_0,\ \Delta\alpha,\ A)$ has been suggested. More specifically, the maximum of the multifractal spectrum $f(\alpha)$ is used to detect the correlation behavior in terms of persistence and anti-persistence. The spectrum α_0 gives the maximum $f(\alpha)$, i.e., $f(\alpha_0) = 1$. At this spectrum, the measure provides information about the central tendency of the multifractal spectrum. If $\alpha_0 < 0.5$, then the correlations in the time series exhibit anti-persistent behavior (i.e., an increase is very likely to be followed by a decrease), if $\alpha_0 > 0.5$, then the correlations in the time series exhibit persistent behavior (i.e., an increase is very likely to be followed by an increase, and a decrease is very likely to be followed by a decrease), whereas if $\alpha_0 = 0.5$, then the time series displays characteristics of a standard non-correlated sequence [39,47,48]. By looking into the spectrum width, one can quantitatively detect the time series multifractality. Specifically, the width of the spectrum is estimated by the equation $\Delta\alpha = \alpha_{max} - \alpha_{min}$, and it reflects the degree of multifractality of the time series. The larger values of $\Delta\alpha$ are, the stronger the degree is and the more severe the fluctuations in the time series are. On the contrary, the smaller the values of $\Delta\alpha$, the more the time series is close to a monofractal behavior, indicating less significant fluctuations in the time series. The spectrum width should be equal to zero for a completely monofractal time series [39,49,50]. The dominance of small or large fluctuations is also an interesting characteristic of time series. This information can be extracted from the skew asymmetry of the multifractal spectrum, which is defined by the equation [51] $A = \frac{L-R}{R+L} = \frac{-\Delta S}{W}$, where $R = \alpha_{max} - \alpha_0$, $L = \alpha_0 - \alpha_{min}$, $\Delta S = R - L$, and $W = R + L = \Delta\alpha = \alpha_{max} - \alpha_{min}$. If $A > 0$ $(L > R)$, the spectrum is left-skewed, which means that the scaling behavior of large fluctuations dominates the multifractal behavior. On the contrary, if $A < 0$ $(L < R)$, then the spectrum is right-skewed, where the scaling behavior of small fluctuations dominates. The case of $A = 0$ indicates that the shape of multifractal spectra is symmetric [46,51].

Another multifractal spectrum asymmetry metric is the so-called truncation, defined as $\Delta f(a) = f(\alpha_{min}) - f(\alpha_{max})$ [49,52]. If $\Delta f(a) < 0$, the multifractal spectrum is right-truncated, i.e., it has a long left tail, indicating that the multifractal structure in the time series is insensitive to the local fluctuations with small magnitudes. In other words, the time series is less multifractal, closer to monofractal, for the small fluctuations than for the large fluctuations. If $\Delta f(a) > 0$, the multifractal spectrum is left-truncated, i.e., it has a long right tail, indicating that the multifractal structure is then insensitive to the local fluctuations with large magnitudes. It has to be noted that, very often, truncation and skew

asymmetries are directly related so that a left-skewed spectrum is also right-truncated, and a right-skewed is left-truncated. The absolute value of truncation, also known as "C-value", $\mathrm{C}-\mathrm{value} = |\Delta f(a)| = |f(\alpha_{min}) - f(\alpha_{max})|$ [49,52], indicates the degree of the truncation asymmetry, which also provides interesting information as C-values are known to illustrate the systems' underlying undulation or instability. The degree of undulation or instability becomes minimum when the C-value presents the smallest value (≈ 0) [49,52].

2.2. Fuzzy Entropy (FuzzyEn)

Expanding upon the concepts already established with approximate entropy (*ApEn*) and sample entropy (*SampEn*), Chen et al. [53,54] combined elements from Fuzzy sets and information theory to develop a fuzzy version of the *SampEn*. Fuzzy entropy (*FuzzyEn*) like its ancestors, *ApEn* and *SampEn* [54], is a "regularity statistic" that quantifies the (un)predictability of fluctuations in a time series. For the estimation of *FuzzyEn*, the similarity between vectors is defined based on fuzzy membership functions and the vectors' shapes. The gradual and continuous boundaries of the fuzzy membership functions lead to a series of advantages, such as the continuity as well as the validity of *FuzzyEn* at small values, higher accuracy, stronger relative consistency, and even less dependence on the data length. *FuzzyEn* can be considered as an upgraded alternative of *SampEn* (and *ApEn*) for the evaluation of complexity, especially for short time series contaminated by noise [55].

Similar to *SampEn*, *FuzzyEn* excludes self-matches. Nevertheless, it applies a slightly different definition for the employed first $N-m$ vectors of a length of m, by removing a baseline, $\bar{s}_i$:

$$\bar{s}_i = m^{-1} \sum_{j=0}^{m-1} s_{i+j}, \tag{7}$$

i.e., for the *FuzzyEn* estimations, we use the first $N-m$ of the vectors:

$$\mathbf{X}_i^m = \{s_i,\ s_{i+1}, \ldots, s_{i+m-1}\} - \bar{s}_i,\ \ i = 1,\ 2, \ldots,\ N-m+1, \tag{8}$$

Then, the similarity degree, D_{ij}^m, between each pair of vectors, $\mathbf{X}_j^m$ and $\mathbf{X}_i^m$, being within a distance, r, from each other is defined by a fuzzy membership function:

$$D_{ij}^m = \mu\left(d_{ij}^m, r\right), \tag{9}$$

where d_{ij}^m is, as in the case of *ApEn* and *SampEn*, the supremum norm difference between $\mathbf{X}_i^m$ and $\mathbf{X}_j^m$. For each vector, $\mathbf{X}_i^m$, we estimate the average similarity degrees with respect to all other vectors, $\mathbf{X}_j^m$, $j = 1, 2, \ldots, N-m+1$, and $j \neq i$ (i.e., excluding itself):

$$\phi_i^m(r) = (N-m-1)^{-1} \sum_{i=1, j\neq i}^{N-m} D_{ij}^m. \tag{10}$$

Then, we evaluate

$$\varphi^m(r) = (N-m)^{-1} \sum_{i=1}^{N-m} \phi_i^m(r), \tag{11}$$

and

$$\varphi^{m+1}(r) = (N-m)^{-1} \sum_{i=1}^{N-m} \phi_i^{m+1}(r). \tag{12}$$

The $FuzzyEn(m, r)$ is then defined as

$$FuzzyEn(m, r) = \lim_{N\to\infty} \left[\ln \varphi^m(r) - \ln \varphi^{m+1}(r)\right], \tag{13}$$

which, for finite time series, can be calculated by the statistic

$$FuzzyEn(m, r, N) = \ln \varphi^m(r) - \ln \varphi^{m+1}(r). \quad (14)$$

As mentioned above, *FuzzyEn* is a measure of estimation of the complexity. More specifically, lower *FuzzyEn* values demonstrate a larger chance that a set of data will be followed by similar data (regularity). Hence, lower values demonstrate larger regularity. Conversely, a greater value of *FuzzyEn* indicates a smaller chance of similar data being repeated (irregularity). Thus, greater values convey more randomness, disorder, and system complexity. Consequently, a low (high) value of *FuzzyEn* reflects a high (low) degree of regularity [42].

2.3. Tsallis Entropy

In a vast variety of systems that exhibit long-range interactions or long-term memory or being of a multifractal nature, they have been found to be better described by a generalized statistical-mechanical formalism proposed by Tsallis [56,57]. Tsallis, inspired by multifractal concepts, introduced an entropic expression characterized by an index, q_{TS}, which leads to non-extensive statistics [56,57]:

$$S_{q_{TS}} = k\frac{1}{q_{TS}-1}\left(1 - \sum_{i=1}^{W} p_i^{q_{TS}}\right), \quad (15)$$

where q_{TS} is a real number, k is the Boltzmann's constant from statistical thermodynamics, p_i are probabilities associated with the microscopic configurations, and W is their total number. It is important to note that there is a remarkable conceptual similarity between Tsallis' entropy definition and the notion of Rényi entropies.

The entropic index, q_{TS}, describes the deviation of Tsallis entropy from the standard Boltzmann–Gibbs entropy. Indeed, using $p_i^{(q_{TS}-1)} = e^{(q_{TS}-1)\ln(p_i)} \sim 1 + (q_{TS}-1)\ln(p_i)$ in the limit $q_{TS} \to 1$, we recover the Boltzmann–Gibbs entropy

$$S_1 = -k\sum_{i=1}^{W} p_i \ln(p_i), \quad (16)$$

as the thermodynamic analog of the information-theoretic Shannon entropy. From this point and for the rest of this article, we will refer to the entropy calculated by Equation (16) as the Shannon entropy.

For $q_{TS} \neq 1$, the entropic index, q_{TS}, characterizes the degree of non-extensivity reflected in the following pseudo-additivity rule:

$$\frac{S_{q_{TS}}(A+B)}{k} = \frac{S_{q_{TS}}(A)}{k} + \frac{S_{q_{TS}}(B)}{k} + (q_{TS}-1)\frac{S_{q_{TS}}(A)}{k}\frac{S_{q_{TS}}(B)}{k}, \quad (17)$$

where A and B are two subsystems. In case these subsystems have special probability correlations, extensivity does not hold for $q_{TS} = 1$ $(S_1 \neq S_1(A) + S_1(B))$, but may occur for $S_{q_{TS}}$, with a particular value of the index, $q_{TS} \neq 1$. Such systems are called non-extensive [56]. The cases $q_{TS} > 1$ and $q_{TS} < 1$ correspond to sub-additivity or super-additivity, respectively. As in the case of Rényi entropies, we may think of q_{TS} as a bias parameter: $q_{TS} < 1$ privileges rare events, while $q_{TS} > 1$ highlights prominent events [58].

It is noted that the parameter, q_{TS}, itself is not a measure of the complexity of the system but measures the degree of the non-extensivity of the system. In turn, the temporal variations of the Tsallis entropy, $S_{q_{TS}}$, for some q_{TS}, quantify the dynamical changes of the complexity of the system. In particular, lower $S_{q_{TS}}$ values characterize the portions of the signal with lower complexity [55].

2.4. Fisher Information Measure (FIM)

In the last decades, Fisher information has been increasingly gaining the interest of scientists of different scientific fields. It was first introduced by Fisher [59] as a representation of the amount of information in the results of experimental measurements of an unknown parameter of a stochastic system, or simply the amount of information that can be extracted from a set of measurements (or the "quality" of the measurements) [60]. Fisher information is a useful method to study non-stationary and complex time series [61]. It is used as a measure of the level of disorder of a system, behaving inversely to entropy, i.e., when the disorder increases, the entropy increases, while the Fisher information decreases. Fisher information has been successfully applied to many different systems, revealing its ability in describing the complexity of them [62–64]. Additionally, its use has been suggested to identify reliable precursors of critical events [65–67]. Moreover, Fisher information presents the so-called "locality" property in contrast to the "globality" of entropy, referring to the sensitivity of Fisher information in changes in the shape of the probability distribution corresponding to the measured variable, not presented by entropy [68,69]. The Fisher information measure can be expressed as

$$I_x = \sum_{n=1}^{N-1} \frac{[p(x_{n+1}) - p(x_n)]^2}{p(x_n)}. \tag{18}$$

The discrete probability distribution $p(x_n)$ corresponds to the specific values of the unknown underlying probability density function at the center values of the intervals $\{x_n\}$, which are not necessarily of equal length. The probability density function is usually approximated by a histogram, or by the kernel density estimator technique, employing different kernel functions such as the Gaussian kernel or Epanechnikov kernel [60].

3. Data and Results

The cryptocurrency market is a relatively new and emerging market, meaning that the trading mechanism is unique and makes it very different from traditional markets. More than 21,800 different cryptocurrencies are currently traded around the world with an estimated total market capitalization of over USD 843 billion (see, e.g., https://coinmarketcap.com/ (accessed on 7 December 2022)). On the other hand, the foreign exchange market is the largest financial market worldwide, with transactions amounting to trillions of US dollars daily [70]. In this article, we focused on the analysis of the two most representative currencies of these two markets, i.e., the BTC/USD and EUR/USD. Our analyses were applied to the daily logarithmic returns ($r_t = \ln p_t - \ln p_{t-1}$, where p_t denotes the price at time t) of the BTC/USD and EUR/USD during the period from 1 May 2019 to 20 January 2021. In an announcement by the *WHO* on 11 March 2020, the outbreak of COVID-19 was declared a global pandemic. Therefore, we considered the period from 1 May 2019 to 11 March 2020 as the pre-announcement period, and the period from 12 March 2020 to 20 January 2021 as the post-announcement period. All financial time series were taken from Yahoo Finance (http://finance.yahoo.com/ (accessed on 7 December 2022)).

In our study, we investigated the temporal evolution of complexity and fractal characteristics by using overlapping sliding windows (with window length equal to 512 samples and slide step equal to 1 sample). First, we investigated the temporal evolution of the multifractal spectrum parameters (α_0, $\Delta\alpha$, A) before and after the outbreak of the pandemic. Then, for the same time period, we extended our analysis by examining the temporal evolution of Fuzzy entropy, Tsallis entropy, Shannon entropy, and Fisher information.

At this point, we should mention that for the calculation of Tsallis entropy we have chosen to use the value $q_{TS} = 1.8$ for the non-extensive parameter, q_{TS}. On one hand, for financial time series q_{TS} has been found to take values $q_{TS} \sim 1.6 - 1.8$ [3], which has been discussed within the framework of the similarities in scaling properties and universality related to observables of extreme events from different disciplines (e.g., financial crisis, earthquake, epileptic seizure, magnetic storm, solar flare) [3,55,60,61]. On the other hand,

from the time series analysis point of view, the selection of the q_{TS} value for the calculation of the temporal variation of Tsallis entropy practically only affects the "separation" between the lower and higher complexity parts of the analyzed time series (e.g., min to max entropy values ratio, in direct analogy to the signal to noise ratio), while for the herein analyzed time series it was found that any q_{TS} value in the range $\sim 1.6 - 1.8$ leads to approximately the same results.

Figure 1c,d, depict the temporal evolution of α_0 values under different market trends of the BTC/USD and EUR/USD returns, respectively. By analyzing the overall trend of the BTC/USD returns, it is observed that the values of the α_0 fluctuate around 0.6 (Figure 1c). These results indicate that the returns time series is characterized by long-range correlations, both before and after the onset of COVID-19. By analyzing the downtrend markets of the BTC/USD returns, it is observed that the values of the α_0 fluctuate over 0.6 both before and after the outbreak of the pandemic, indicating persistent behavior. In the uptrend markets of the BTC/USD returns, the values of α_0 fluctuate between 0.5 and 0.65 for almost throughout the analysis period. An exception is a short period of time after the *WHO* announcement, where α_0 values fell below 0.5.

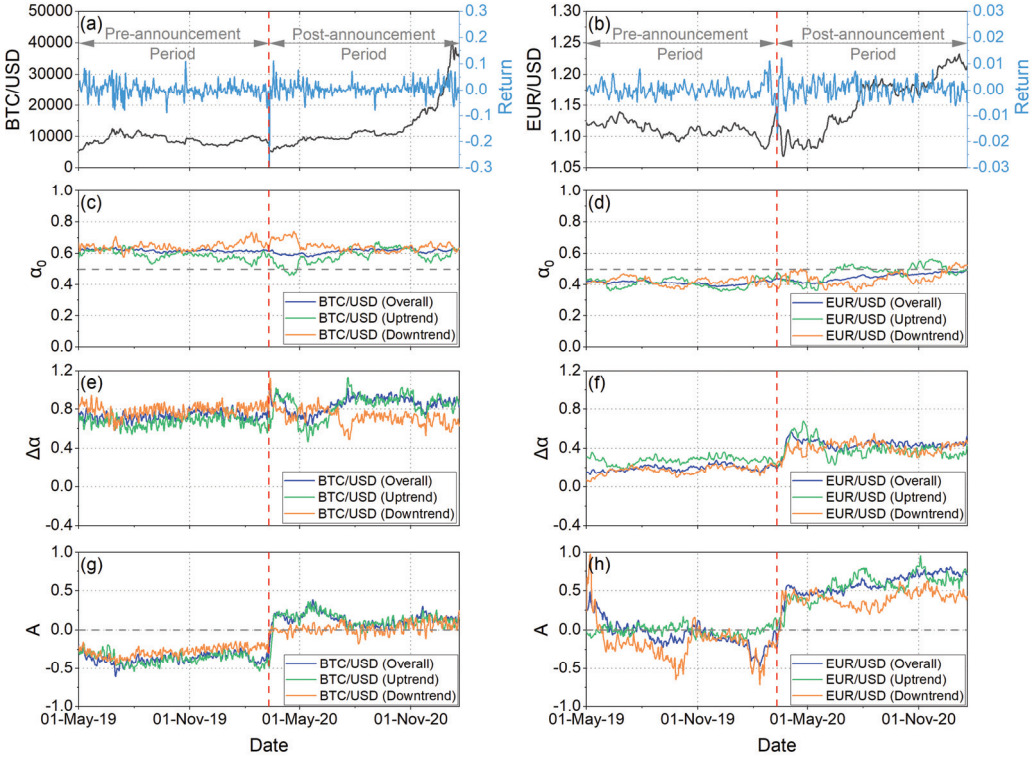


Figure 1. Comparative asymmetric multifractal analysis of BTC/USD (left panels) and EUR/USD (right panels) under different market trends. (**a**,**b**): Exchange rates and Returns. (**c**,**d**): Temporal evolution of α_0 parameter. (**e**,**f**): Temporal evolution of width of the multifractal spectrum. (**g**,**h**): Temporal evolution of the asymmetry parameter A values. The red vertical dash line corresponds to the date of the *WHO* announcement in which COVID-19 was declared a global pandemic (i.e., 11 March 2020). The period from 1 May 2019 to 11 March 2020 corresponds to the pre-announcement period, while the period from 12 March 2020 to 20 January 2021 corresponds to the post-announcement period.

Figure 1d depicts the temporal evolution of the α_0 values of the EUR/USD returns under different market trends. In this case, we observe that for the overall trend the α_0 values fluctuate around 0.4 during the whole pre-announcement period, while after the announcement they present a progressive increase approaching very close to $\alpha_0 = 0.5$ at the end of the considered analysis period. This suggests that the time series exhibit a "different" power-law correlation, such that large and small time series are more likely to alternate (anti-persistent behavior). It is worth mentioning that the downtrend α_0 values remain at the anti-persistent side except for the very last part of the analyzed period, while the uptrend α_0 values, although < 0.5 for the whole pre-announcement period, present an alternating behavior after the *WHO* announcement, taking values $\alpha_0 > 0.5$ for two two-month-long periods.

Figure 1e,f illustrate the width of the multifractal spectrum under different market trends of the BTC/USD and EUR/USD returns, respectively. As already mentioned in Section 2.1.1, the width of the multifractal spectrum $\Delta\alpha$ is a measure of the degree of multifractality. If a time series presents a smaller width of the multifractal spectrum, this indicates that the time series has lower heterogeneity, i.e., lower fluctuations and lower market risk [58]. The results show that throughout the period analyzed, the width of the multifractal spectrum receives higher values for the BTC/USD returns compared to the EUR/USD returns. Therefore, it can be concluded that EUR/USD is relatively more stable than BTC/USD. In addition, we observe that after the outbreak of the pandemic, the width of the multifractal spectrum increased for both BTC/USD returns and the EUR/USD returns for the overall trend. This suggests that after the outbreak of the pandemic, both currencies reacted similarly in terms of multifractality when observed from an overall trend point of view. The degree of multifractality increased, and, therefore, the fluctuations became more intense and the market risk increased. However, in terms of asymmetric multifractality, this is not always the case. When focused on downside markets of BTC/USD, the degree of multifractality decreased after the outbreak. More interestingly, downtrend multifractality was higher than the uptrend multifractality during the period before COVID-19, but during the pandemic the uptrend multifractality became higher. These findings reveal that the incremental multifractality in BTC/USD is due to intense fluctuations and higher heterogeneity during price increases, but not during price declines. In EUR/USD, it appears that the downside markets play a more important role in increasing multifractality. Nevertheless, both market trends may have had some impact in the post-announcement period increase in multifractality. It is important to note that the increase in multifractality in BTC/USD returns during COVID-19 is consistent with the existing literature as other studies have reached the same conclusion (e.g., [26,27]).

Figure 1g,h depict the temporal evolution of the asymmetry parameter A values under different market trends of the BTC/USD and EUR/USD returns, respectively. In the time period before the onset of the pandemic, the asymmetry parameter A of BTC/USD returns appears to have been consistently below 0, indicating relative dominance of the small fluctuations. Immediately after the date of the *WHO* announcement, there was a sharp change in the values of A in all market trends. Specifically, for both overall trend and uptrend markets, the values of A of the BTC/USD returns remain above 0 for the entire period after the outbreak of the pandemic. This sharp change shows a transition from a period where small fluctuations predominate (before the pandemic) to a period where large fluctuations predominate (during the pandemic). In the downtrend markets, the values of parameter A are almost at 0 for the entire period after the outbreak of the pandemic, indicating that the spectrum became practically symmetrical (Figure 1g). On the other hand, the values of the asymmetry parameter A of the EUR/USD returns for the uptrend markets are almost equal to 0 for the whole period before the outbreak of the pandemic. This fact indicates that the spectrum is practically symmetrical. On the contrary, analyzing the overall and downward trends of the market, we observe that the values of the asymmetry parameter A are below 0, almost for the entire period before the outbreak of the pandemic. Therefore, it is concluded that in the overall and downward trends of

the markets, they are dominated by the small fluctuations in EUR/USD returns before the outbreak of the pandemic. Immediately after the announcement date, the values of the asymmetry parameter A of the EUR/USD returns exceeded 0 in all market trends. This result shows that EUR/USD returns after the outbreak of the pandemic are dominated by large fluctuations (Figure 1h).

Figure 2c,d indicate that the effect of the announcement was, for all cases (for both BTC/USD and EUR/USD returns and for all three considered market trends), a sharp change towards right-truncation, which means that after the *WHO* announcement the multifractal structure in the time series became more insensitive to the local fluctuations with small magnitudes. On the other hand, Figure 2e,f show that the behavior of BTC/USD and EUR/USD returns was different concerning the degree of truncation asymmetry, indicated by the so-called C − value. Specifically, EUR/USD returns present C − values quite close to 0 before the *WHO* announcement, which means that the underlying system then presented the lowest possible undulation or instability. After the *WHO* announcement, the picture changed and for all market trends an increase in the undulation or instability of the underlying system is observed. In contrast, BTC/USD returns present a general trend (although with notable fluctuations for the overall and uptrend markets) towards a decrease in the C − values after the *WHO* announcement, which means that a trend for the decrease in the undulation or instability of the underlying system is observed. It is noted that the downtrend market after the *WHO* announcement presents C − values closer to 0, as compared with the uptrend and overall markets, indicating lower undulation or instability.

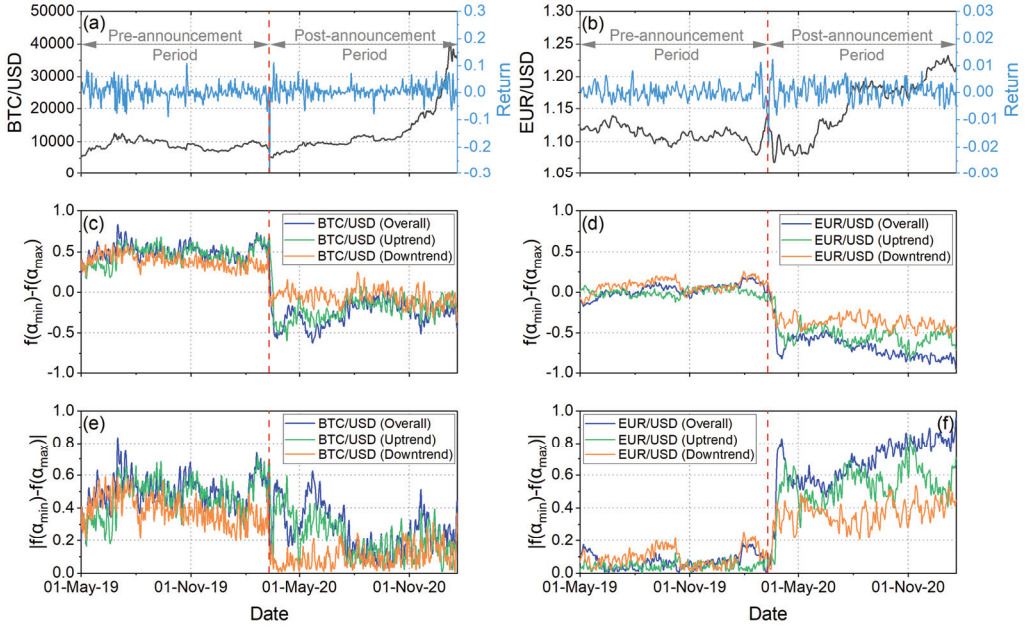


Figure 2. Comparative asymmetric multifractal analysis of BTC/USD (left panels) and EUR/USD (right panels) under different market trends. (**a**,**b**): Exchange rates and Returns. (**c**,**d**): Temporal evolution of truncation $\Delta f(a) = f(\alpha_{min}) - f(\alpha_{max})$. (**e**,**f**): Temporal evolution of the degree of truncation asymmetry, known as C − value $= |\Delta f(a)| = |f(\alpha_{min}) - f(\alpha_{max})|$. The red vertical dash line corresponds to the date of the *WHO* announcement in which COVID-19 was declared a global pandemic (i.e., 11 March 2020). The period from 1 May 2019 to 11 March 2020 corresponds to the pre-announcement period, while the period from 12 March 2020 to 20 January 2021 corresponds to the post-announcement period.

Moreover, we analyzed the temporal evolution of some complexity measures. Figure 3c,d illustrate the temporal evolution of the Fuzzy entropy of the BTC/USD and EUR/USD returns, respectively. As it has already been mentioned in Section 2.2., smaller values of Fuzzy entropy indicate a greater chance that a set of data will be followed by similar data (regularity). Conversely, larger values of Fuzzy entropy point to a lower chance of similar data being repeated (irregularity). As we observe in Figure 3c,d, the values of Fuzzy entropy dropped sharply in both BTC/USD and EUR/USD returns immediately after the *WHO* announcement. This fact indicates that in the pre-announcement period, both BTC/USD and EUR/USD returns were characterized by a higher degree of disorder and randomness, i.e., by higher complexity. In contrast, in the period during the pandemic, the values of Fuzzy entropy decreased, suggesting that the returns were characterized by a higher degree of order and lower complexity. Therefore, it is concluded that the pandemic led investors to behave in an "organized" (similar) way that thereby reduced the complexity of the two markets.

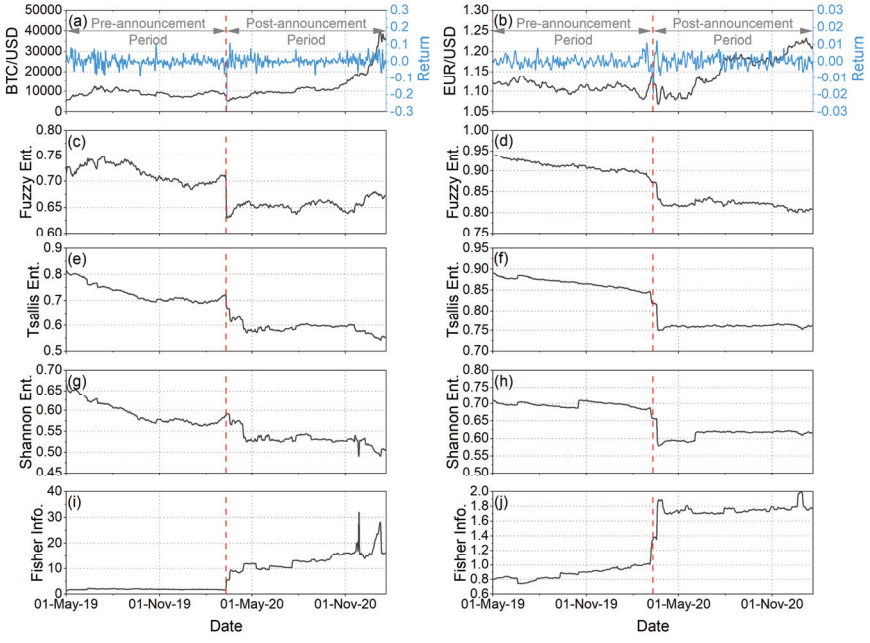


Figure 3. Comparative analysis of BTC/USD (left panels) and EUR/USD (right panels). (**a**,**b**): Exchange rates and Returns. (**c**,**d**): Temporal evolution of Fuzzy entropy. (**e**,**f**): Temporal evolution of Tsallis entropy. (**g**,**h**): Temporal evolution of Shannon entropy. (**i**,**j**): Temporal evolution of Fisher information. The red vertical dash line corresponds to the date of the *WHO* announcement in which COVID-19 was declared a global pandemic (i.e., 11 March 2020). The period from 1 May 2019 to 11 March 2020 corresponds to the pre-announcement period, while the period from 12 March 2020 to 20 January 2021 corresponds to the post-announcement period.

Corresponding results are obtained by also studying two quite popular complexity measures, i.e., the Shannon entropy (Figure 3g,h) and Tsallis entropy (Figure 3e,f). More specifically, the time variations of the Shannon entropy as well as the Tsallis entropy (for a given q_{TS}) quantify the dynamical changes of the information content and the complexity of the system. Smaller values characterize time series with lower complexity and randomness, as well as higher information content and order. Conversely, larger values characterize time series with higher complexity, disorder and randomness, as well as lower information content. As we observe in Figure 3e–h, during COVID-19, the

values of Tsallis and Shannon entropies were reduced in both BTC/USD and EUR/USD returns, indicating that the complexity of the two markets was reduced and the information content was increased. It is important to note that all the entropy measures we applied quickly adapted to market conditions, showing a sharp decrease immediately after the *WHO* announcement, with Shannon entropy being the exception in the case of BTC/USD. Additionally, it is of particular interest that the entropy values remained at low levels throughout the pandemic period we analyzed, showing that the effects of the pandemic were not short-term. Additionally, concerning the study of Lahmiri and Bekiros [32], although not the main finding of their analyses, it is nevertheless important to note that their results showed a decrease in Rényi entropy (and consequently a decrease in randomness) for the BTC/USD market during the pandemic compared to before.

In addition, we applied one more complexity measure, the Fisher information. Fisher information is a useful method to study non-stationary and complex time series. Fisher information is used as a measure of the degree of order of a system, behaving inversely to entropy, i.e., when the order increases, the entropy decreases, while the Fisher information increases. Moreover, unlike entropy, it is sensitive to changes in the shape of the probability distribution corresponding to the measured variable. Figure 3i,j illustrate the temporal evolution of the Fisher information of the BTC/USD and EUR/USD returns, respectively. We observe that immediately after the *WHO* announcement, the values of Fisher information increased in both BTC/USD and EUR/USD returns, indicating an increase in the order of the two markets.

At this point, it has to be mentioned that the observed decrease in randomness after the *WHO* announcement, indicated by all the applied complexity measures, is fully compatible with the corresponding increase of multifractality. Specifically, the more random a time series is, the more unifractal its scaling is, which means that a more multifractal time series can be considered as being farther away from "randomness" [71].

From the interpretation of our results in financial terms, useful conclusions are revealed. More specifically, in analyzing the values of α_0 for overall trend, as we have already mentioned, we observe that the BTC/USD returns show persistent behavior, while the EUR/USD returns exhibit anti-persistent behavior almost throughout the time period we studied them (Figure 1c,d). A persistent or anti-persistent market return series is characterized by a long memory effect. Therefore, what happens today, theoretically, will impact the future in a nonlinear fashion. For example, if a persistent market return change has been up (down) in the last period, then the changes will continue to be positive (negative) in the next period. On the other hand, anti-persistent markets are "mean-reverting." If the market return was up (down) in the previous period, it is more likely to be down (up) in the next period [72]. The long-memory characteristic in asset return is a fascinating topic for investors, risk managers, and scholars since appropriate return modeling is crucial for asset allocation and risk control. For example, existence of long memory in asset returns indicates that historical returns changes could be predictors of future returns changes [73]. Then, analyzing the $\Delta\alpha$ and A parameters, we observe that in the post-announcement period, mainly in the case of the EUR/USD, the degree of multifractal returns increased, and, therefore, fluctuations became more intense and market risk increased (Figure 1e,f). At the same time, we observe that in the post-announcement period, returns were dominated by large fluctuations (Figure 1g,h). Therefore, it is concluded that in the post-announcement period, EUR/USD returns experienced intense and large fluctuations. Similar behavior is observed for the overall trend and uptrend markets of the BTC/USD returns. Regarding the downtrend markets of the BTC/USD returns, it appears that during the pandemic period there were less intense fluctuations compared to the pre-pandemic period without small or large fluctuations in returns dominating. The analysis of the truncation asymmetry degree (Figure 2e,f), moreover, revealed that the *WHO* announcement had different impacts on BTC/USD and EUR/USD returns concerning the undulation or instability of the underlying system. For EUR/USD returns, the post-announcement period was characterized by an increase in the undulation or instability of the underlying system, whereas for BTC/USD

returns, the opposite behavior was generally observed. Analyzing the complexity measures (Fuzzy entropy, Tsallis entropy, Shannon entropy, and Fisher information) (Figure 3c–j), we observe a sharp decline in complexity (i.e., increase in the order and information content, and decrease in randomness) in the returns of both BTC/USD and EUR/USD in the post-announcement period. This fact, in financial terms, suggests that the pandemic led investors to behave in an "organized" (similar) way that thereby reduced the complexity of the two markets. In other words, after the outbreak of the pandemic, it seems that investors behaved like a herd. Therefore, it is concluded that although the fluctuations were larger and more intense after the outbreak of the pandemic, this was not carried out in a random way as investors seem to have behaved in an "organized" way; however, this behavior for BTC/USD returns was generally followed by a decrease in undulation or instability of the underlying system, while the opposite happened for EUR/USD returns.

Additionally, it is worth noting that the majority of the measures that we studied showed a strong change for both BTC/USD and EUR/USD returns immediately after the *WHO* announcement (11 March 2020), in which COVID-19 was mentioned for the first time as a pandemic. This fact indicates that the behavior of the system changed immediately after the *WHO*'*s* announcement, although the discussions about COVID-19 being a public health emergency of international concern had begun weeks before. Therefore, although many researchers accept the date of 2 January 2020 as the beginning of the COVID-19 pandemic crisis (e.g., [74–76]), we consider the most suitable start date of the pandemic to be 11 March 2020.

4. Conclusions

The detection of dynamical complexity in time series originated from various complex systems, including the disciplines of physics, finance, and medicine, and is one of the foremost problems in science. The measurement of complexity includes nonlinear statistics methods to extract hidden patterns as well as exploring multifractality, randomness, and information flows. Hence, complexity provides important information regarding the order or disorder states of a system under scrutiny. In the field of finance, the detection of the dissimilarity of complexity between order and disorder states (e.g., before and after the occurrence of extreme events) could shed light on the mechanisms associated with investor reaction to these events.

In this article, we studied the temporal evolution of the multifractality and complexity of BTC/USD and EUR/USD returns for the period before and after the *WHO* announcement that declared COVID-19 a global pandemic. In our study, we first examined the asymmetric multifractality through the analysis of the multifractal spectrum parameters as obtained by the A-MF-DFA method. Then, we extended our analysis by applying Fuzzy, Tsallis, and Shannon entropies as well as the Fisher information measure. Our results can be summarized as follows: (i) For the entire time period that we studied (i.e., before and during the pandemic), the behavior of BTC/USD returns was persistent in all trends of the market. On one hand, in the period before the outbreak of the pandemic, the behavior of EUR/USD returns was anti-persistent in all trends of the market. On the other hand, in almost the entire period after the outbreak of the pandemic, the returns of the EUR/USD exhibited anti-persistent behavior in the overall trend and downtrend markets, while the uptrend market presented an alternating behavior, including short periods of persistent dynamics. (ii) Throughout the period analyzed, the width of the multifractal spectrum received higher values for the BTC/USD returns compared to the EUR/USD returns. This implies that the multifractality of the BTC/USD returns was higher than the multifractality of the EUR/USD returns. In addition, after the outbreak of the pandemic, in the overall trend and uptrend markets, the width of the multifractal spectrum increased for both BTC/USD returns and EUR/USD returns. In the case of BTC/USD, the downtrend multifractality was higher in the pre-announcement period. In EUR/USD, it appears that the downtrend markets played an important role in increasing multifractality. Nevertheless, both market trends may have had some impact on the post-announcement period increase

in multifractality. (iii) In the pre-announcement period, small fluctuations in BTC/USD returns for all market trends dominated. In contrast, in the post-announcement period, large fluctuations in BTC/USD returns for overall trend and uptrend markets dominated, while in downtrend markets the spectrum became practically symmetrical. On the other hand, although in the uptrend markets the spectrum of EUR/USD returns was almost symmetrical, the returns in the overall trend and downtrend markets were dominated by small fluctuations for almost the entire pre-announcement period. During the pandemic period, the returns of the EUR/USD were dominated by large fluctuations in all market trends. (iv) For both BTC/USD and EUR/USD returns and all market trends, a sharp change towards becoming more insensitive to the local fluctuations with small magnitudes was observed after the *WHO* announcement. Nevertheless, the *WHO* announcement had different impacts on BTC/USD and EUR/USD returns concerning the undulation or instability of the underlying system. For EUR/USD returns, the post-announcement period was characterized by an increase in the undulation or instability of the underlying system, whereas for BTC/USD returns, the opposite behavior was generally observed. (v) Fuzzy entropy, non-extended Tsallis entropy, Shannon entropy, and Fisher information showed a sharp decrease in the degree of complexity immediately after the *WHO* announcement for both BTC/USD and EUR/USD. This fact shows that in the post-announcement period, the order and the information content of the systems increased, i.e., the randomness and complexity in the returns of the two currencies decreased. Therefore, in financial terms, we conclude that investors seem to have behaved in an "organized" way, as a herd. In addition, our analyses show that the date of the *WHO* announcement (11 March 2020) could be considered as the most appropriate date for the start of the pandemic. This element could be useful in future research studies.

The main finding that is revealed from our study is that immediately after the *WHO* announcement, the returns of both BTC/USD and EUR/USD presented a decrease in complexity and corresponding increase in multifractality, both indicating that they became less random compared to the pre-announcement period. Hence, it seems that although they are two such different currencies, which both play a key role in the modern financial system, they reacted in a similar way in response to the pandemic.

Author Contributions: Conceptualization, P.I.Z. and S.M.P.; methodology, P.I.Z., S.K., K.U., M.P.H., S.G.S. and S.M.P.; software, P.I.Z., S.K. and S.M.P.; validation, P.I.Z., M.P.H., S.G.S. and S.K.; formal analysis, P.I.Z. and S.K.; investigation, P.I.Z., S.K. and S.M.P.; data curation, P.I.Z.; writing—original draft preparation, P.I.Z., S.K. and S.M.P.; writing—review and editing, K.U., M.P.H., S.G.S. and S.M.P.; visualization, P.I.Z. and S.K.; supervision, S.M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: All financial time series used in this article are publicly available from Yahoo Finance (http://finance.yahoo.com/, accessed on 7 December 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zitis, P.I.; Contoyiannis, Y.; Potirakis, S.M. Critical Dynamics Related to a Recent Bitcoin Crash. *Int. Rev. Financ. Anal.* **2022**, *84*, 102368. [CrossRef]
2. Gopikrishnan, P.; Plerou, V.; Nunes Amaral, L.A.; Meyer, M.; Stanley, H.E. Scaling of the Distribution of Fluctuations of Financial Market Indices. *Phys. Rev. E* **1999**, *60*, 5305–5316. [CrossRef] [PubMed]
3. Potirakis, S.M.; Zitis, P.I.; Eftaxias, K. Dynamical Analogy between Economical Crisis and Earthquake Dynamics within the Nonextensive Statistical Mechanics Framework. *Phys. A Stat. Mech. Appl.* **2013**, *392*, 2940–2954. [CrossRef]
4. Mandelbrot, B. *The Fractal Geometry of Nature*; W.H. Freeman and Co. Ltd.: New York, NY, USA, 1982.
5. Yang, Y.H.; Shao, Y.H.; Shao, H.L.; Stanley, H.E. Revisiting the Weak-Form Efficiency of the EUR/CHF Exchange Rate Market: Evidence from Episodes of Different Swiss Franc Regimes. *Phys. A Stat. Mech. Appl.* **2019**, *523*, 734–746. [CrossRef]

6. Oh, G.; Kim, S.; Eom, C. Long-Term Memory and Volatility Clustering in High-Frequency Price Changes. *Phys. A Stat. Mech. Appl.* **2008**, *387*, 1247–1254. [CrossRef]
7. Aloui, C.; Mabrouk, S. Value-At-Risk Estimations of Energy Commodities via Long-Memory, Asymmetry and Fat-Tailed GARCH Models. *Energy Policy* **2010**, *38*, 2326–2339. [CrossRef]
8. Herrera, R.; Rodriguez, A.; Pino, G. Modeling and Forecasting Extreme Commodity Prices: A Markov-Switching Based Extreme Value Model. *Energy Econ.* **2017**, *63*, 129–143. [CrossRef]
9. 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.* **2021**, *565*, 125562. [CrossRef] [PubMed]
10. Adrangi, B.; Chatrath, A.; Dhanda, K.K.; Raffiee, K. Chaos in Oil Prices? Evidence from Futures Markets. *Energy Econ.* **2001**, *23*, 405–425. [CrossRef]
11. Lahmiri, S.; Bekiros, S. Chaos, Randomness and Multi-Fractality in Bitcoin Market. *Chaos Solitons Fractals* **2018**, *106*, 28–34. [CrossRef]
12. Mandelbrot, B. The Variation of Some Other Speculative Prices. *J. Bus.* **1967**, *40*, 393. [CrossRef]
13. He, L.Y.; Fan, Y.; Wei, Y.M. The Empirical Analysis for Fractal Features and Long-Run Memory Mechanism in Petroleum Pricing Systems. *Int. J. Glob. Energy Issues* **2007**, *27*, 492. [CrossRef]
14. He, L.Y.; Chen, S.P. Multifractal Detrended Cross-Correlation Analysis of Agricultural Futures Markets. *Chaos Solitons Fractals* **2011**, *44*, 355–361. [CrossRef]
15. Jaroonchokanan, N.; Termsaithong, T.; Suwanna, S. Dynamics of Hierarchical Clustering in Stocks Market during Financial Crises. *Phys. A Stat. Mech. Appl.* **2022**, *607*, 128183. [CrossRef]
16. Mantegna, R.N.; Stanley, H.E. *An Introduction to Econophysics: Correlations and Complexity in Finance*; Cambridge University Press: Cambridge, UK, 2004.
17. Kutner, R.; Ausloos, M.; Grech, D.; Di Matteo, T.; Schinckus, C.; Stanley, H.E. Econophysics and Sociophysics: Their Milestones & Challenges. *Phys. A Stat. Mech. Appl.* **2019**, *516*, 240–253. [CrossRef]
18. Stanley, H.E.; Afanasyev, V.; Amaral, L.A.N.; Buldyrev, S.V.; Goldberger, A.L.; Havlin, S.; Leschhorn, H.; Maass, P.; Mantegna, R.N.; Peng, C.K.; et al. Anomalous Fluctuations in the Dynamics of Complex Systems: From DNA and Physiology to Econophysics. *Phys. A Stat. Mech. Appl.* **1996**, *224*, 302–321. [CrossRef]
19. Kutner, R.; Schinckus, C.; Stanley, H.E. Three Risky Decades: A Time for Econophysics? *Entropy* **2022**, *24*, 627. [CrossRef]
20. Choudhury, T.; Kinateder, H.; Neupane, B. Gold, Bonds, and Epidemics: A Safe Haven Study. *Financ. Res. Lett.* **2022**, 102978. [CrossRef]
21. Urquhart, A.; Zhang, H. Is Bitcoin a Hedge or Safe Haven for Currencies? An Intraday Analysis. *Int. Rev. Financ. Anal.* **2019**, *63*, 49–57. [CrossRef]
22. Baur, D.G.; Hong, K.; Lee, A.D. Bitcoin: Medium of Exchange or Speculative Assets? *J. Int. Financ. Mark. Inst. Money* **2018**, *54*, 177–189. [CrossRef]
23. Melki, A.; Nefzi, N. Tracking Safe Haven Properties of Cryptocurrencies during the COVID-19 Pandemic: A Smooth Transition Approach. *Financ. Res. Lett.* **2021**, *46*, 102243. [CrossRef] [PubMed]
24. Ma, C.; Tian, Y.; Hsiao, S.; Deng, L. Monetary Policy Shocks and Bitcoin Prices. *Res. Int. Bus. Financ.* **2022**, *62*, 101711. [CrossRef]
25. Nguyen, K.Q. The Correlation between the Stock Market and Bitcoin during COVID-19 and Other Uncertainty Periods. *Financ. Res. Lett.* **2021**, *46*, 102284. [CrossRef]
26. Mnif, E.; Jarboui, A.; Mouakhar, K. How the cryptocurrency market has performed during COVID 19? A multifractal analysis. *Financ. Res. Lett.* **2020**, *36*, 101647. [CrossRef] [PubMed]
27. Kakinaka, S.; Umeno, K. Cryptocurrency Market Efficiency in Short- and Long-Term Horizons during COVID-19: An Asymmetric Multifractal Analysis Approach. *Financ. Res. Lett.* **2022**, *46*, 102319. [CrossRef]
28. Aslam, F.; Aziz, S.; Nguyen, D.K.; Mughal, K.S.; Khan, M. On the Efficiency of Foreign Exchange Markets in Times of the COVID-19 Pandemic. *Technol. Forecast. Soc. Chang.* **2020**, *161*, 120261. [CrossRef]
29. Mensi, W.; Sensoy, A.; Vo, X.V.; Kang, S.H. Pricing Efficiency and Asymmetric Multifractality of Major Asset Classes before and during COVID-19 Crisis. *N. Am. J. Econ. Financ.* **2022**, *62*, 101773. [CrossRef]
30. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Stanisz, T.; Wątorek, M. Complexity in Economic and Social Systems: Cryptocurrency Market at around COVID-19. *Entropy* **2020**, *22*, 1043. [CrossRef]
31. Lahmiri, S.; Bekiros, S. The Impact of COVID-19 Pandemic upon Stability and Sequential Irregularity of Equity and Cryptocurrency Markets. *Chaos Solitons Fractals* **2020**, *138*, 109936. [CrossRef]
32. Lahmiri, S.; Bekiros, S. Renyi Entropy and Mutual Information Measurement of Market Expectations and Investor Fear during the COVID-19 Pandemic. *Chaos Solitons Fractals* **2020**, *139*, 110084. [CrossRef]
33. Wang, J.; Wang, X. COVID-19 and Financial Market Efficiency: Evidence from an Entropy-Based Analysis. *Financ. Res. Lett.* **2021**, *42*, 101888. [CrossRef] [PubMed]
34. Fernandes, L.H.S.; Bouri, E.; Silva, J.W.L.; Bejan, L.; de Araujo, F.H.A. The Resilience of Cryptocurrency Market Efficiency to COVID-19 Shock. *Phys. A Stat. Mech. Appl.* **2022**, *607*, 128218. [CrossRef] [PubMed]
35. Kim, K.; Lee, M. The Impact of the COVID-19 Pandemic on the Unpredictable Dynamics of the Cryptocurrency Market. *Entropy* **2021**, *23*, 1234. [CrossRef] [PubMed]

36. Olbryś, J.; Majewska, E. Regularity in Stock Market Indices within Turbulence Periods: The Sample Entropy Approach. *Entropy* **2022**, *24*, 921. [CrossRef]
37. 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.* **2019**, *520*, 54–61. [CrossRef]
38. Telli, Ş.; Chen, H. Multifractal Behavior in Return and Volatility Series of Bitcoin and Gold in Comparison. *Chaos Solitons Fractals* **2020**, *139*, 109994. [CrossRef]
39. Telli, Ş.; Chen, H. Multifractal Behavior Relationship between Crypto Markets and Wikipedia-Reddit Online Platforms. *Chaos Solitons Fractals* **2021**, *152*, 111331. [CrossRef]
40. Batten, J.A.; Kinateder, H.; Wagner, N. Multifractality and Value-At-Risk Forecasting of Exchange Rates. *Phys. A Stat. Mech. Appl.* **2014**, *401*, 71–81. [CrossRef]
41. Olbrys, J.; Majewska, E. Approximate Entropy and Sample Entropy Algorithms in Financial Time Series Analyses. *Procedia Comput. Sci.* **2022**, *207*, 255–264. [CrossRef]
42. Mastrogiannis, D.; Potirakis, S.M. Experimental Study of the Dynamic Evolution of Cumulative Energy Release during LiF Fracture under Uniaxial Compression. *Int. J. Solids Struct.* **2018**, *132*, 59–65. [CrossRef]
43. Fernandes, L.H.S.; Araujo, F.H.A.; Silva, M.A.R.; Acioli-Santos, B. Predictability of COVID-19 Worldwide Lethality Using Permutation-Information Theory Quantifiers. *Results Phys.* **2021**, *26*, 104306. [CrossRef]
44. Cao, G.; Cao, J.; Xu, L. Asymmetric Multifractal Scaling Behavior in the Chinese Stock Market: Based on Asymmetric MF-DFA. *Phys. A Stat. Mech. Appl.* **2013**, *392*, 797–807. [CrossRef]
45. Mensi, W.; Lee, Y.J.; Al-Yahyaee, K.H.; Sensoy, A.; Yoon, S.-M. Intraday Downward/Upward Multifractality and Long Memory in Bitcoin and Ethereum Markets: An Asymmetric Multifractal Detrended Fluctuation Analysis. *Financ. Res. Lett.* **2019**, *31*, 19–25. [CrossRef]
46. Zhang, X.; Zhang, G.; Qiu, L.; Zhang, B.; Sun, Y.; Gui, Z.; Zhang, Q. A Modified Multifractal Detrended Fluctuation Analysis (MFDFA) Approach for Multifractal Analysis of Precipitation in Dongting Lake Basin, China. *Water* **2019**, *11*, 891. [CrossRef]
47. Telli, Ş.; Chen, H.; Zhao, X. Detecting Multifractality and Exposing Distributions of Local Fluctuations: Detrended Fluctuation Analysis with Descriptive Statistics Pooling. *Chaos Solitons Fractals* **2022**, *155*, 111678. [CrossRef]
48. Stošić, D.; Stošić, D.; Stošić, T.; Eugene Stanley, H. Multifractal Properties of Price Change and Volume Change of Stock Market Indices. *Phys. A Stat. Mech. Appl.* **2015**, *428*, 46–51. [CrossRef]
49. Boulassel, A.; Zaourar, N.; Gaci, S.; Boudella, A. A New Multifractal Analysis-Based for Identifying the Reservoir Fluid Nature. *J. Appl. Geophys.* **2021**, *185*, 104185. [CrossRef]
50. Cheng, Q.; Liu, X.; Zhu, X. Cryptocurrency Momentum Effect: DFA and MF-DFA Analysis. *Phys. A Stat. Mech. Appl.* **2019**, *526*, 120847. [CrossRef]
51. Drożdż, S.; Oświęcimka, P. Detecting and Interpreting Distortions in Hierarchical Organization of Complex Time Series. *Phys. Rev. E* **2015**, *91*, 030902. [CrossRef]
52. Chakraborty, B.; Haris, K.; Latha, G.; Maslov, N.; Menezes, A. Multifractal Approach for Seafloor Characterization. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 54–58. [CrossRef]
53. Chen, W.; Wang, Z.; Xie, H.; Yu, W. Characterization of surface EMG signal based on fuzzy entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2007**, *15*, 266–272. [CrossRef] [PubMed]
54. Chen, W.; Zhuang, J.; Yu, W.; Wang, Z. Measuring complexity using FuzzyEn, ApEn, and SampEn. *Med. Eng. Phys.* **2009**, *31*, 61–68. [CrossRef] [PubMed]
55. Balasis, G.; Donner, R.V.; Potirakis, S.M.; Runge, J.; Papadimitriou, C.; Daglis, I.A.; Eftaxias, K.; Kurths, J. Statistical Mechanics and Information-Theoretic Perspectives on Complexity in the Earth System. *Entropy* **2013**, *15*, 4844–4888. [CrossRef]
56. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
57. Tsallis, C. Generalized entropy-based criterion for consistent testing. *Phys. Rev. E* **1998**, *58*, 1442–1445. [CrossRef]
58. Zunino, L.; Pérez, D.G.; Kowalski, A.; Martín, M.T.; Garavaglia, M.; Plastino, A.; Rosso, O.A. Fractional Brownian Motion, Fractional Gaussian Noise, and Tsallis Permutation Entropy. *Phys. A Stat. Mech. Appl.* **2008**, *387*, 6057–6068. [CrossRef]
59. Fisher, R.A. Theory of statistical estimation. In *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1925; pp. 700–725.
60. Potirakis, S.M.; Minadakis, G.; Eftaxias, K. Analysis of Electromagnetic Pre-Seismic Emissions Using Fisher Information and Tsallis Entropy. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 300–306. [CrossRef]
61. Minadakis, G.; Potirakis, S.M.; Stonham, J.; Nomicos, C.; Eftaxias, K. The Role of Propagating Stress Waves on a Geophysical Scale: Evidence in Terms of Nonextensivity. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 5648–5657. [CrossRef]
62. Telesca, L.; Caggiano, R.; Lapenna, V.; Lovallo, M.; Trippetta, S.; Macchiato, M. The Fisher Information Measure and Shannon Entropy for Particulate Matter Measurements. *Phys. A Stat. Mech. Appl.* **2008**, *387*, 4387–4392. [CrossRef]
63. Telesca, L.; Caggiano, R.; Lapenna, V.; Lovallo, M.; Trippetta, S.; Macchiato, M. Analysis of Dynamics in Cd, Fe, and Pb in Particulate Matter by Using the Fisher–Shannon Method. *Water Air Soil Pollut.* **2008**, *201*, 33–41. [CrossRef]
64. Telesca, L.; Lovallo, M.; Kiszely, M.M.; Toth, L. Discriminating Quarry Blasts from Earthquakes in Vértes Hills (Hungary) by Using the Fisher-Shannon Method. *Acta Geophys.* **2011**, *59*, 858–871. [CrossRef]

65. Humeau, A.; Trzepizur, W.; Rousseau, D.; Chapeau-Blondeau, F.; Abraham, P. Fisher Information and Shannon Entropy for On-Line Detection of Transient Signal High-Values in Laser Doppler Flowmetry Signals of Healthy Subjects. *Phys. Med. Biol.* **2008**, *53*, 5061–5076. [CrossRef] [PubMed]
66. Telesca, L.; Lovallo, M.; Ramirez-Rojas, A.; Angulo-Brown, F. A Nonlinear Strategy to Reveal Seismic Precursory Signatures in Earthquake-Related Self-Potential Signals. *Phys. A Stat. Mech. Appl.* **2009**, *388*, 2036–2040. [CrossRef]
67. Telesca, L.; Lovallo, M.; Carniel, R. Time-Dependent Fisher Information Measure of Volcanic Tremor before the 5 April 2003 Paroxysm at Stromboli Volcano, Italy. *J. Volcanol. Geotherm. Res.* **2010**, *195*, 78–82. [CrossRef]
68. Frieden, B.R.; Binder, P.M. Physics from Fisher Information: A Unification. *Am. J. Phys.* **2000**, *68*, 1064–1065. [CrossRef]
69. Fath, B.D.; Cabezas, H. Exergy and Fisher Information as Ecological Indices. *Ecol. Model.* **2004**, *174*, 25–35. [CrossRef]
70. Mohamad, A.; Stavroyiannis, S. Do Birds of a Feather Flock Together? Evidence from Time-Varying Herding Behaviour of Bitcoin and Foreign Exchange Majors during COVID-19. *J. Int. Financ. Mark. Inst. Money* **2022**, *80*, 101646. [CrossRef]
71. Antoniades, I.P.; Karakatsanis, L.P.; Pavlos, E.G. Dynamical Characteristics of Global Stock Markets Based on Time Dependent Tsallis Non-Extensive Statistics and Generalized Hurst Exponents. *Phys. A Stat. Mech. Appl.* **2021**, *578*, 126121. [CrossRef]
72. Kyaw, N.A.; Los, C.A.; Zong, S. Persistence Characteristics of Latin American Financial Markets. *J. Multinatl. Financ. Manag.* **2006**, *16*, 269–290. [CrossRef]
73. Lahmiri, S.; Bekiros, S. The Effect of COVID-19 on Long Memory in Returns and Volatility of Cryptocurrency and Stock Markets. *Chaos Solitons Fractals* **2021**, *151*, 111221. [CrossRef]
74. Kocaarslan, B.; Soytas, U. The Asymmetric Impact of Funding Liquidity Risk on the Volatility of Stock Portfolios during the COVID-19 Crisis. *Sustainability* **2021**, *13*, 2286. [CrossRef]
75. Goodell, J.W.; Goutte, S. Co-Movement of COVID-19 and Bitcoin: Evidence from Wavelet Coherence Analysis. *Financ. Res. Lett.* **2020**, *38*, 101625. [CrossRef] [PubMed]
76. Erdem, O. Freedom and Stock Market Performance during COVID-19 Outbreak. *Financ. Res. Lett.* **2020**, *36*, 101671. [CrossRef] [PubMed]

entropy

MDPI

Article

Oracles in Decentralized Finance: Attack Costs, Profits and Mitigation Measures

Ayana T. Aspembitova * and Michael A. Bentley

Euler Labs, London EC1V 2NX, UK
* Correspondence: ayana.aspembitova@euler.xyz

Abstract: Decentralized finance (DeFi) is by far the most popular application of blockchain technology. Despite the wide acceptance of new financial instruments and services, there are still many unexplored areas in the field. We dedicate this research to the understanding of one of the most crucial limitations of decentralized finance—oracles. DeFi protocols, as well as other blockchain applications, function in a closed environment and regularly need to fetch real-world information (e.g., assets' prices)—the tool used for this purpose is called an oracle. We review the existing oracle types in DeFi applications and focus our research on the least explored one: when another protocol, typically a decentralized exchange, serves as a price oracle. After explaining the mechanisms behind the decentralized exchanges, we introduce an algorithmic model that allows one to safely design a decentralized oracle and adjust crucial parameters. We believe that understanding and implementing the logic presented in the model can help to reduce the chances of price manipulations attacks, which are the most frequent incident types in DeFi.

Keywords: DeFi; oracle; automated market makers; decentralized exchange; lending protocol

Citation: Aspembitova, A.T.; Bentley, M.A. Oracles in Decentralized Finance: Attack Costs, Profits and Mitigation Measures. *Entropy* **2023**, *25*, 60. https://doi.org/10.3390/e25010060

Academic Editors: Stanisław Drożdż, Jarosław Kwapień and Marcin Wątorek

Received: 6 November 2022
Revised: 13 December 2022
Accepted: 22 December 2022
Published: 28 December 2022

1. Introduction

Blockchain-based smart contracts have been successfully growing, and their use cases are quite innovative and have attracted lots of interest valued in the billions of dollars. However, there is a fundamental limitation of decentralized applications—they execute in a closed environment and a bridge service (oracle) is needed when obtaining information outside of the blockchain. As decentralized applications evolve and mature, oracles play an increasingly prominent role in ensuring the safety across smart contracts. Despite the critical role that oracles play in decentralized applications, the research is still in its infancy. In [1], the authors performed a bibliometric analysis and demonstrated the alarming scarcity of the research dedicated to blockchain oracles. Moreover, in the recent study of DeFi incidents [2], the authors empirically showed that oracle manipulation attacks are the most frequent incident types in DeFi. Although there are tools that can detect the price manipulation attacks [3,4], and identify new vulnerabilities in real time, there is still a need for prevention measures. The lack of understanding of oracles mechanics and functions concerns not only academic research but more so the real users of decentralized applications.

Decentralized finance (DeFi) uses blockchain technology to provide financial instruments without intermediaries in a trustless and transparent manner [5]. DeFi covers a wide range of financial products, offering innovative alternatives to traditional financial products, such as stablecoins, exchanges, lending protocols, insurance and yield farming protocols.

Here, we provide an overview on why DeFi rests heavily on the use of oracles and how information from the outside world can be retrieved. Generally speaking, there is some *ground truth* information that resides outside of smart contracts, and smart contracts need it for the proper performance. To obtain such ground truth, smart contracts need reliable *data sources*—any entity that stores the ground truth information (databases, sensors or other smart contracts). Then, *data feeders* report off-chain data to an on-chain system.

The systematic explanation on the existing type of oracles in a blockchain is provided in [6]. As for the decentralized financial applications, the ground truth needed is the price of the assets listed in a smart contract. Although there are many types of oracles with different functions and characteristics, the oracles currently used in DeFi can be broadly divided into two main categories—decentralized trust-based oracles and decentralized exchanges used as oracles.

Decentralized trust-based oracles function as a smart contract and do not rely on a single source of information. Instead, they query multiple sources and aggregate the obtained information into a single output. The papers [7–9] provided a detailed review on the architecture, workflow and weak points of various decentralized oracles, such as Chainlink [10,11], Provable [12], Oraclize, etc. Some DeFi applications are fetching the price information directly from the decentralized exchanges by either getting the spot price or aggregating the prices over a certain window size. Using the spot price can be very dangerous because the price can be easily manipulated [13–15]. Therefore, more and more DeFi applications started using the TWAP (time-weighted average price) instead—the output price is calculated as a weighted average over a certain time period and, therefore, the cost of price manipulation of the TWAP oracle increases linearly with the length of the TWAP averaging window, reducing the chance of an oracle hack.

In this paper, we focus on the decentralized exchanges (DEXs) used as oracles for DeFi protocols. While trust-based oracles have attracted some attention from the researchers, using DEXs directly as oracles is still not well understood. In [16], the authors analyzed the cost of TWAP manipulation when an arithmetic mean is used for the aggregation and also considered the possibility of an MMEV attack. Decentralized exchanges utilize the concept of automated market makers (to be explained in detail in Section 2). Our main contributions consist of the following: we systematize the existing knowledge about using automated market maker (AMM)-based decentralized exchanges as oracles, we derive attack costs for the most popular cost functions used in DEXs, then we derive the relations between protocol-specific parameters and oracle-specific parameters that impact the safety of using the DEX-based oracle and, finally, we develop the algorithmic model that allows to assess the risks of using oracles in a given protocol. Overall, knowing the mechanics behind the oracles' work would give a comprehensive understanding on how attacks can be performed. Implementing the logic presented in the model below would give the quantitative estimate on the cost a potential attacker needs for a successful attack. Knowing the mechanism behind the price oracle and being able to precisely estimate the cost of a potential attack provides an additional layer of security to the protocols using DEX-based oracles.

The paper is structured as follows. First, we review the most popular AMM-based decentralized exchanges, demonstrate their logic and the cost functions used for asset pricing. In the appendices, we derive the cost of the attacks for each type of AMM pricing function discussed in Section 2. Then, in Section 3, we discuss various aggregation methods that can be used in DEX-based oracles and show how they can be impacted by the price manipulation attack. Section 4 aggregates all the information obtained above and provides a step-by-step algorithm on how to mitigate attacks related to the DEX-based oracles on the example of a lending protocol. We simulate various attack scenarios to the lending protocol on two types of AMM cost functions—a constant product and stableswap. Finally, we conclude all the findings and discuss the future directions of this research in the last section.

2. Automated Market Makers

To understand the safety of using the DEX-based oracle, we need to first be familiar with how DEXs work and understand the mathematics behind it—the cost function utilized by the DEX. In this section, we first explain the mechanism of the decentralized exchange protocol and then review the popular AMM cost functions and demonstrate how they are used to price assets in a DEX.

An AMM-based decentralized exchange consists of pools of different assets (liquidity pools). Liquidity to these pools is provided by people who wish to gain income from the

transaction fees (liquidity providers). Each pool can have few assets (currently most of the pools have two assets) and users who want to exchange assets (traders) interact directly with the given pool to swap asset x to asset y.

In centralized exchanges, the price discovery happens by matching the sell and buy orders from various counterparties. In contrast to it, decentralized exchanges are based on the automated market-making mechanism (AMM). The AMM utilizes the cost function that discovers the price algorithmically—this function only allows counterparties to exchange the assets for the prices along the trajectory determined by the AMM formula and quantities of the available assets. Although the implementation of AMM functions to price assets in decentralized exchanges is quite novel, the idea of agents automatically placing bets and following prescribed rules is not new and has been implemented in many areas to aggregate the information—the prediction of building openings [17], sport matches [18], etc. Overall, the idea of automated market making is to define algorithmic rules for agents within the system to place their bets on a certain subject, aggregate them and derive a single function (*conservation function*) from the outcome. In DEXs, this process goes a little different. First, the conservation function is defined and then agents (traders and liquidity providers) match their trades, and whenever there is a trade that goes beyond the expectations of the cost function, it is punished by the algorithm and, therefore, it discourages agents to behave (trade) differently than prescribed by the conservation function. Although in DEXs agents are not algorithmic bots but real people, they do act in a way as was expected algorithmic bots to act to preserve the cost function.

In [19], Othman introduced five desideratas (desirable properties) for cost functions—monotonicity, convexity, bounded loss, translation invariance and positive homogeneity. As it was proven by [20], it is impossible for the cost function to satisfy all five properties. Therefore, all the cost functions utilized by AMMs satisfy only a few properties, while others are relaxed.

2.1. Logarithmic Market Scoring Rule

The first automated market maker for prediction markets was introduced by Hanson [21,22]. It has been quite popular due to its simple analytical form and satisfying the main desirable properties for cost functions (convexity, bounded loss and translation invariance).

The Logarithmic Market Scoring Rule (LMSR) conservation function for n assets is defined as:

$$C(x) = b \log(\sum_{i=1}^{n} exp(x_i/b) \tag{1}$$

where $b > 0$ is the liquidity parameter, it is strictly positive, constant and it is defined before the pricing of assets. b parameter controls the liquidity in the market—the higher the b, the less the price is shifted when assets are added. Moreover, it translates into the bigger maximum loss because the market maker's worst-case loss is the function of b which is b log n.

The derivative of the cost function $C(x)$ is the price function in the LMSR:

$$p_i(x) = \frac{exp(\frac{x_i}{b})}{\sum_j exp(\frac{x_j}{b})} \tag{2}$$

The LMSR is used in many settings, such as auctions, prediction markets, rating markets, etc. In decentralized finance, the LMSR has not been widely used for a few reasons: first, the LMSR does not satisfy the liquidity sensitivity property; second, it is quite easy and cheap to compromise the price of an asset when the LMSR is used as a cost function.

By allowing the parameter b to be the function of the outstanding quantities instead of being constant, the LMSR becomes liquidity sensitive—the Liquidity-Sensitive Logarithmic Market Scoring Rule (LS-LMSR), introduced by Othman [23]:

$$C(x) = b(x) \log(\sum_{i=1}^{n} exp(x_i/b(x)) \tag{3}$$

where function b is as follows:

$$b(x) = \alpha \sum_{i} x_i \tag{4}$$

where α is the parameter that is strictly positive and set before the pricing of assets. The possible maximum commission (also called vigorish v) depends on the α parameter and does not exceed v when α is set as follows [24]:

$$\alpha = \frac{v}{n \log n} \tag{5}$$

where n is the number of outcomes (assets in the pool for AMM-based DEX). Depending on what is the desired maximum commission v, the optimal parameter α can be easily calculated.

In decentralized finance, the LS-LMSR is used in applications such as Augur [25] and Gnosis [26].

2.2. Constant Product Market

Constant product AMM (CPAMM) cost function for n assets is defined as follows:

$$C(x) = \prod_{i=1}^{n} x_i \tag{6}$$

where $C(x)$ set as a constant.

Constant product AMM has many advantages that makes it suitable to be used in DEXs—it is simple to code into the smart contract, it is a convex function which meets the principles of supply and demand and it is also liquidity sensitive. Although it has been shown in [27] that prices in such a DEX can be inaccurate during volatile markets, this cost function still remains the most popular and being utilized by large DEXs, such as Uniswap [28,29].

Decentralized exchange pools consist of two tokens and Equation (6) becomes the following:

$$x \times y = k \tag{7}$$

where k is the constant, x is the amount of the first token and y is the amount of the second token.

The price for each token in a pool can be calculated by simply dividing the number of tokens in one reserve to the number of tokens in another. A more detailed review of constant product markets is given in [30,31].

2.3. Combination of Constant Sum and Constant Product Markets

In DeFi, there are assets that have the same value, for example, a different version of USD (USDC, USDT, etc.). Because the ratio between asset x and asset y in such pools is stable and close to 1, they are called *stableswap* pools. The pricing formula for stableswap pools was developed by the Curve team [32]. Essentially, this is a combination of the constant product market pricing formula $xy = k$ and the linear invariant $x + y = C$. The rationale behind adding the linear invariant term to the constant product formula is to achieve the closer peg 1:1 and allow lower slippage for stableswap pools. When using only a linear invariant formula, tokens are always traded at 1:1 with zero slippage; however, this might lead to the depleting of the pool's one token. Using only a constant product formula leads to larger slippage and a less stable peg. Therefore, the combination of these two curves allows to keep the pool balanced while providing a more stable peg.

The final stableswap curve formula looks as in Equation (8). For the full explanation and derivation, refer to the paper [33].

$$2^2A(x+y)+D=2^2AD+\frac{D^3}{2^2xy} \tag{8}$$

where A is the amplification factor for the linear invariant curve—the larger the A, the closer the curve to the linear. D is the total amount of tokens in the pool.

To calculate the price for the token, one needs to express the curve for y from Equation (8), and the derivative of that expression stands for the price. Stableswap AMM is widely used in many DEX pools that have the same price for both tokens.

3. Aggregation Methods

In this paper, we focus on oracles for DeFi applications that get price information directly from the decentralized exchanges by aggregating the output prices over a certain time period. Every time there is a new swap (trade) in the DEX, the price is updated in oracle and then the time-weighted average is calculated. To mitigate the possible effect of the price manipulation within one or a few blocks, one would prefer to use the time-weighted and/or liquidity-weighted average price.

In this section, we discuss various aggregation methods and show the impact of a price manipulation attack on each of them.

3.1. Arithmetic Mean Time-Weighted Average Price

The arithmetic mean TWAP over n price updates is calculated as follows:

$$TWAP=\frac{\sum_{i=1}^{n}t_ip_i}{\sum_{i=1}^{n}t_i} \tag{9}$$

where t_i is the time elapsed between the price update i and next price update $i+1$, and p_i is the price during that period. n is the averaging window.

We estimate the effect of manipulation on the TWAP price when the attacker wants to consistently manipulate the spot price for m times of price updates within the averaging window n.

$$TWAP_m=\frac{\sum_{i=1}^{n-m}t_ip_i+\sum_{j=n-m+1}^{n}t_jp_j}{\sum_{i=1}^{n}t_i} \tag{10}$$

assuming that attack would happen in the last m blocks. From here, we would like to estimate the p_j—how big should the manipulated price be that the attacker should target in order to achieve the desired effect on $TWAP_m$.

$$\sum_{j=0}^{m}t_jp_j=TWAP_m\times\sum_{i=1}^{n}t_i-\sum_{i=1}^{n-m}t_ip_i \tag{11}$$

In the case when the attacker does not want to be exposed to arbitrageurs and wants to manipulate the price within one block $m=1$, the manipulated price will be as follows:

$$p_j=\frac{TWAP_m\times\sum_{i=1}^{n}t_i-\sum_{i=1}^{n-m}t_ip_i}{t_j} \tag{12}$$

Although using the TWAP instead of a spot price is safer in terms of avoiding the malicious price manipulations, the output from the averaging might not be accurate.

3.2. Geometric Mean Time-Weighted Average Price

The geometric mean TWAP over n blocks can be calculated as the nth root of the product of the spot price on each block:

$$TWAP = (\prod_{i=1}^{n} p_i)^{\frac{1}{n}} \tag{13}$$

If the attacker wants to manipulate the geometric mean TWAP by manipulating the price over m blocks, then the target TWAP will be calculated as follows:

$$TWAP_m = (p^{n-m} \times q^m)^{\frac{1}{n}} \tag{14}$$

An attacker wanting to manipulate the $TWAP$ to some particular oracle price $TWAP_m$ over m blocks will need to know what spot price q they need to move the normal spot price p to in each of those blocks. It can be calculated by rearranging Equation (14):

$$q = \sqrt[m]{\frac{TWAP^n}{p^{n-m}}} \tag{15}$$

This equation shows that it is surprisingly difficult to move the geometric mean TWAP from the wider market spot price when manipulated blocks are few in number relative to unmanipulated blocks. That is, the spot price must be moved a significant distance from its wider market price in order to have even a modest impact on the geometric mean TWAP.

3.3. Median Time-Weighted Average Price

Using median time-weighted average prices as oracles has been discussed in [34], although they have not been widely implemented in practice yet. Theoretically, because the median is unaffected by the effect of outliers, it could be a solution to avoid single-block manipulation attacks, especially in pools with small liquidity. For an attacker to influence the oracle's final output price, they would need to control the last m block prices for at least half of the period of the window size.

From the economic point of view, storing price time series over a certain period to calculate the median could be very expensive in terms of the gas cost in the Ethereum blockchain. In alternative blockchains with a different technical design and cheaper gas cost, this could be possible if the pros of using the median TWAP outweigh the cons. In DeFi, median TWAP oracles have been implemented in the Euler Finance protocol as an alternative price source to the geometric mean TWAP [34].

4. Algorithmic Model to Estimate the Safety of TWAP Oracle

We have reviewed AMM cost functions in decentralized exchanges that are often used as price oracles. Moreover, we looked at the most popular methods for price aggregation in oracles. Now, we would like to systematize everything into the algorithmic model that allows to estimate the safety of any DEX-based TWAP oracle. Inputs in the model are parameters of the protocol using TWAP oracle and parameters of AMM that is being used as oracle. Outputs of the model are the attack cost AC and capital C needed to provide to the protocol under attack to be able to profit from it. Overall, the model output only tells how much funds a potential attacker needs to gain the profit from price manipulation. To assess the economic feasibility of such an attack, one would need to perform additional independent analysis, for instance, if model output tells that attack would cost 100 USD, then it is economically feasible for many people and, therefore, it is not safe. If model estimates AC and C to be big so only a few people can technically perform an attack, then one can conclude that chances of oracle attack are low. Outputs from the model can serve as a starting point to decide on the safety of TWAP-based oracle.

We first introduce the general algorithmic model that allows to estimate the feasibility of price manipulation attack. Then, we explain how the model can be implemented based on the example of lending protocol—this is the most popular use case of TWAP oracles. We provide a brief explanation of how lending protocols work, demonstrate how model can be used and simulate various attack scenarios for constant product and stableswap AMM-based TWAP oracles.

4.1. Algorithmic Model

Algorithm 1 shows step by step how to estimate the safety of TWAP oracle. Model outputs are attack cost AC and minimum collateral C an attacker needs to profit from their manipulation. Knowing these parameters allows us to estimate the economic feasibility of such an attack and, therefore, to decide whether it is safe to implement the given DEX as an oracle. Input parameters needed for the model can be divided as protocol specific, oracle specific (averaging window size WS) and DEX specific (liquidity L). These parameters can be easily found in protocol's web-page and in DEXs page (L). To find values of Δx (number of tokens needed to move the price to the target value) and Δy (number of tokens to receive in return after the swap), one needs to know the type of AMM that DEX uses to price assets. In Appendices A and B, we have derived equation for Δx and Δy for constant product and stableswap AMMs. Overall, all these parameters can be precisely found and no assumptions need to be made. There is one more parameter that is crucial to take into account—the number of blocks m during which attacker will try to manipulate the price. This can not be known upfront but its value affects the cost of an attack significantly. One would suppose that the value of m should be small because any deviation from the spot price would be noticed by arbitrageurs and set back to the real value, not allowing an attacker to manipulate for many blocks. However, in pools with infrequent trading activity, the arbitrage opportunity might go unnoticed for longer times. Moreover, there is a chance of multi-block MEV-style attack, where an attacker could cooperate with the miner to mine a few blocks in a row. This style of attack combined with the oracle price manipulation makes the oracle attack cost cheaper. In the examples and simulations below, we assume no MMEV attack and frequent trading activity in a pool.

Algorithm 1: Model to estimate the safety of DEX-based oracle

Input: Protocol risk parameters, WS, m, Δx, Δy, L
Output: AC, C

1 Calculate minimum deviation from the spot price needed to profit based on protocol-specific parameters.
2 Calculate target manipulation price from oracle $TWAP_m$.

$$TWAP_m = TWAP_s + TWAP_s \times \epsilon$$

3 Calculate how big should be the manipulation price p_m to achieve needed $TWAP_m$. For this:
(I) Estimate the oracle window size WS.
(II) Decide on number of blocks m to manipulate the price.
(III) Depending on aggregation method, calculate p_m using Equation (12) or Equation (14).
4 Calculate attack cost AC based on AMM-specific parameters (liquidity L) and p_m found above. Find Δx and Δy according to the AMM type, as shown in Appendices A and B.

$$AC = \Delta x - \Delta y$$

5 Calculate the minimum capital needed to obtain profit $Profit > 0$ depending on protocol's risk parameters.
6 Estimate the economic feasibility of an attack based on the attack cost AC and collateral C values.

4.2. Lending Protocols

Lending protocols (also called money markets, credit protocols or protocols for loanable funds) are a market that matches borrowers and lenders—users—who wish to gain interest on their savings, deposit their funds to the lending protocol and then it allows borrowers to lend available assets paying certain interest rate. Detailed explanation on how lending protocols work was provided in the paper [35]. Overall, lending protocols have attracted a lot of interest and become very popular among DeFi community—Ethereum-based lending protocols such as Aave [36], Compound [37], dYdX [38] and MakerDAO [39]. Credit protocols are one of the most popular use cases for AMM-based DEX data to be used as a price feed. Chainlink type of price databases are not always available for relatively new blockchains. In such cases, DEXs, acting as the only option for nascent chains, are used as substitutes for more robust oracle solutions. Considering the large TVL (total value locked in protocol) associated with the popularity of credit protocols and their growing functionality and complexity, it is vital to understand the safe settings of AMM pools that are used as a price information source.

In lending protocols, any user can anonymously borrow funds, but to be able to do so, they first need to provide some collateral asset. To ensure the safety and the solvency of protocol, the Loan-to-Value (LTV) parameter is used—this parameter shows how much a user can borrow relative to their collateral value (all loans in lending protocols are *overcollateralized*. For example, if user deposited 100 USD worth of collateral C and LTV parameter is 80%, then they can borrow up to 80% worth of the other asset B. More detailed explanations of lending protocols and their risk parameters can be found in [35–37,39].

In practice, lending protocols are the most frequent target for oracle manipulation attacks. An attacker tries to artificially increase their collateral value by compromising the oracle price information to be able to borrow more.

We assume a scenario where the attacker artificially increases the value of their collateral to be able to borrow more than their actual collateral value allows. In this case, attacker's profit can be formulated as follows:

$$Profit = (C \times LTV + C \times LTV \times \epsilon) - C \tag{16}$$

Here, ϵ is the target price manipulation fraction and C is the value of collateral—for convenience and normalization purposes, we consider it not as an absolute value but relative to the pool liquidity. This normalization allows us to generalize findings and give parameter recommendations for any pools regardless the size:

$$C = \frac{Collateral}{PoolLiquidity} \tag{17}$$

Because the attacker would need to give up their collateral in order to realize their profit from manipulation, we subtract the actual value of their collateral from the profit. From Equation (16), it is clear that the lower the LTV parameter, the more difficult it is to get the profit from an attack and the higher the ϵ should be. We can derive the value of the target manipulation price from the Equation (16)—we set the $Profit = 0$ and calculate the ϵ as:

$$\epsilon \geq \frac{1}{LTV} - 1 \tag{18}$$

Figure 1 shows the minimum manipulation target ϵ an attacker needs to achieve for the attack to be profitable given a certain LTV.

Next, after we know the minimum price target needed to make the attack profitable, we can calculate the total cost of an attack using the equations derived in Appendices A and B.

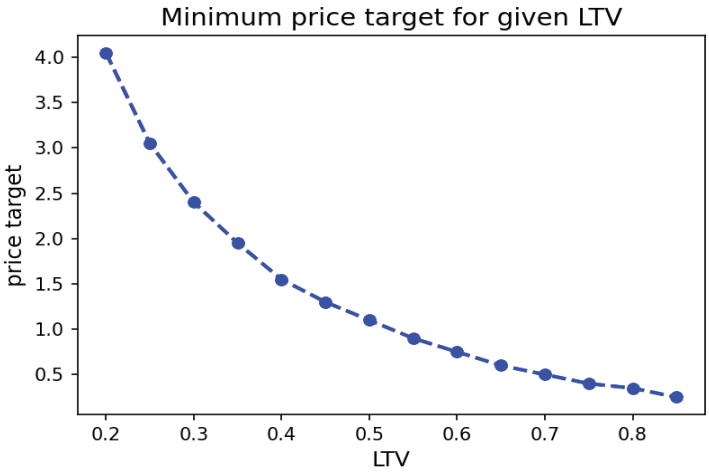


Figure 1. Minimum price target needed for given LTV to obtain the profit from the attack.

4.3. Attack Scenarios to Lending Protocol Using Constant Product AMM-Based TWAP Oracle

Knowing both the profit and attack cost equations, it is straightforward to simulate various attacks to lending protocols that are using any type of AMM as oracle. We looked at the two most popular types of AMM used in decentralized exchanges—constant product and stableswap—to obtain the full understanding about attacker's profits and manipulation capital needed.

In this section, we demonstrate how to calculate the cost of an attack under certain conditions using the terms and explanations shown above. Assumptions used in this example are as follows:

- Loan to Value of the target asset equals 40%.
- TWAP window equals 30 min.
- Time without arbitrage equals 1 min.

Note that we are not making any assumptions regarding pool liquidity given how Equation (16) was defined, which allows us to make calculations, irrespective of the pool liquidity.

Using the attack cost formula and simulating scenarios with varying attacker's collateral C, price target ϵ and pool's liquidity L, we arrive at the following profitability matrix shown in Figure 2, where the space in red indicates a loss (negative profit), the space in blue indicates a positive profit for the attacker and the white area indicates zero-profit scenarios.

From the results shown in Figure 2, we see that the attacker can theoretically reach a profit under almost every combination of events. Moreover, we see that an attack can be profitable by adjusting the requirements to the collateral and the manipulation target ϵ. The lower the ϵ, the higher the collateral the attacker needs to provide for the attack to become profitable and vice versa. These results make us question whether there is such a combination of ϵ and collateral that allows an attacker to obtain profit from an attack with minimum resources?

Figure 3 shows that, effectively, we can retrieve the attack's minimum cost through a specific combination of ϵ and collateral provided; let us call it an *optimal target*. For the example covered in this section, this point happens at collateral being around five times larger than the pool's liquidity and manipulation target ϵ being 4.7. Most importantly, the figure below shows that the total capital needed for a profitable attack is 9.3 times the liquidity in the pool used for the AMM.

The optimal target found above and amount of resources needed to reach that point can serve as a reference when deciding on the safety of an oracle.

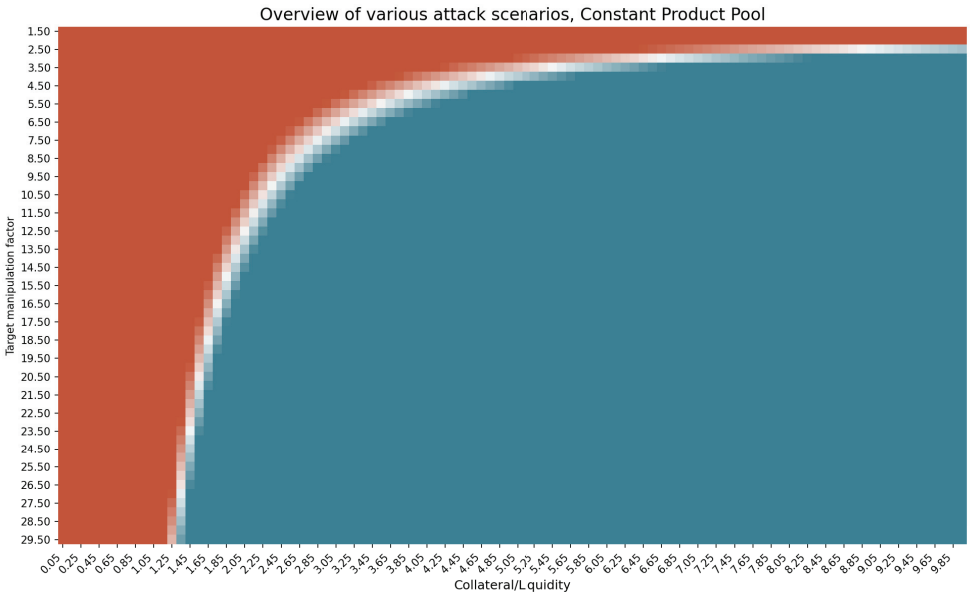


Figure 2. Overview of various attack scenarios in constant product market. The x-axis shows the amount of an attacker's collateral in terms of liquidity, and the y-axis is the target manipulation price ϵ. The space in red shows the non-profitable attack scenarios (when attack cost exceeds the profit). Blue areas show profitable attacks, while the space in white shows when $Profit - AttackCost$ is close to zero. LTV = 0.4 for all scenarios.

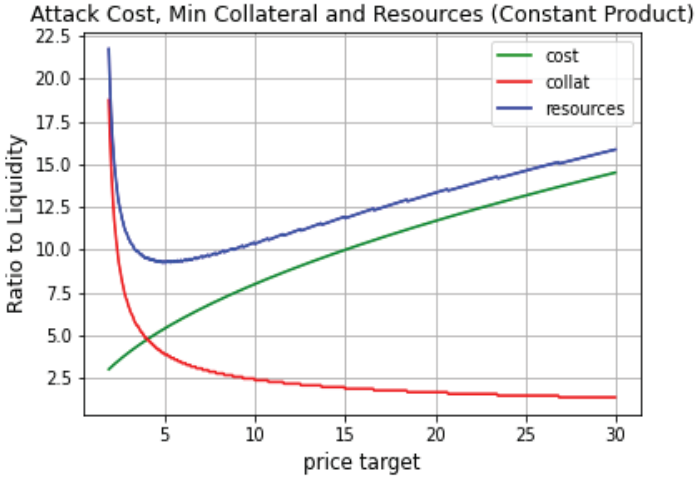


Figure 3. Attack cost, minimum collateral needed and the total resources needed for the profitable attack.

4.4. Attack Scenarios to Lending Protocol Using Stableswap AMM-Based TWAP Oracle

In the previous section, we showed how parameters need to be set for the constant product AMM oracle. In this section, we look at the attack cost and profit when using the stableswap pool as an oracle.

With the attack cost calculated for the stableswap in Appendices A and B, we can run simulations as in the previous section and produce the profitability matrix. Note that the

assumptions used within this section (except for the amplification factor, which is unique to the stableswap AMM) are the same as those used in the previous section. The following figure shows the profitability matrix using a stableswap with an amplification factor A of 30.

From Figure 4, it can be seen that the profitability space for an attack is larger in a stableswap AMM than a CPAMM. In other words, manipulating a stableswap-based TWAP is cheaper than a CPAMM-based TWAP oracle using the same assumptions.

Figure 5 shows the minimum cost of performing a profitable attack, indicating a significantly lower point of minimum cost for an attack in a stableswap AMM than in a constant product AMM. Moreover, another difference is the considerably slower growth rate of the attack cost as the manipulation target increases, which makes the total cost of the attack stagnate as the manipulation target increases. In contrast with constant product AMM, the total cost of an attack keeps increasing. For a stableswap AMM, this results in relatively cheap attack opportunities.

As a final note, a stableswap pool can be relatively stable (in terms of price) at a very unbalanced state (in terms of underlying reserves). At the extreme, we could have a situation where the pool is very close to the "knee" of the pricing curve (where the constant sum (linear) part of pricing curve meets the constant product part), and manipulation attacks become increasingly easier to perform given the aggressive nature of the stableswap curve. In other words, we cannot assume that the attack will take place from a 50:50 state or anything closer to that. The more unbalanced the pool at the start of the manipulation attack, the less resources needed to conduct the attack. Therefore, we do not recommend using stableswap pools as an oracle. Please also refer to Appendix B, where various attack scenarios in stableswap pool were shown under different LTV values—it is clear that this type of pool is much cheaper to manipulate comparing with the constant product market. Moreover, as we can see from Figure 5, once the *optimal price target* is reached, attacker does not need significantly more resources to manipulate price higher and to obtain even higher returns from the attack.

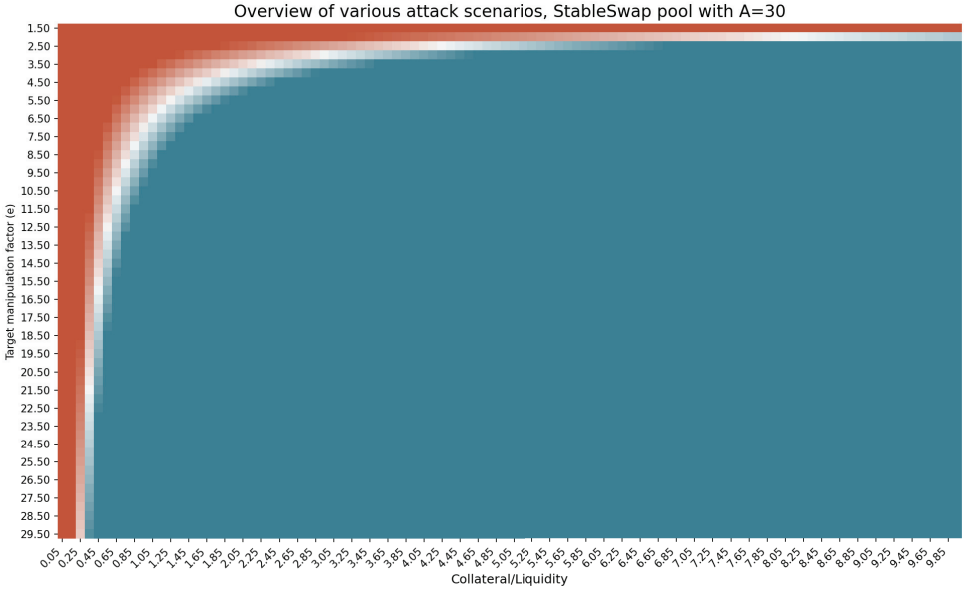


Figure 4. Overview of various attack scenarios in stableswap market. The x-axis shows the amount of attacker's collateral in terms of liquidity, and y-axis is the target manipulation price ϵ. The space in red shows the non-profitable attack scenarios (when attack cost exceeded the profit). Blue areas show profitable attacks, while the space in white shows when $Profit - AttackCost$ is close to zero. LTV = 0.4 for all scenarios.

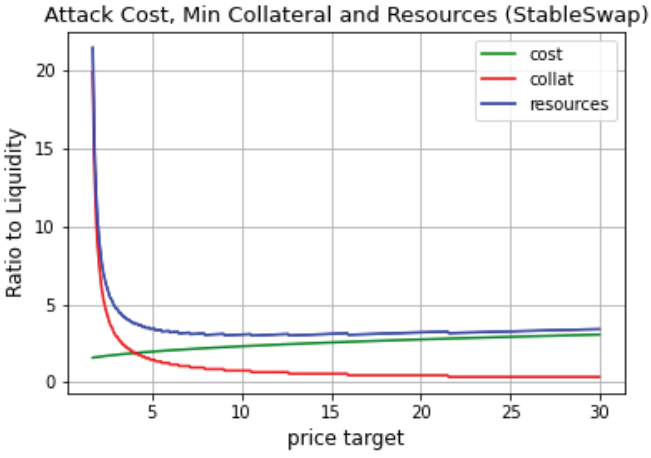


Figure 5. Attack cost, minimum collateral needed and the total resources needed for the profitable attack.

5. Conclusions and Discussion

DeFi protocols, as well as any other blockchain applications, function in a closed environment, and for their proper performance, a reliable data source (oracle) is needed. Currently, there are two different ways to fetch the data about assets' prices—either by using trust-based oracles (e.g., Chainlink) or by getting the prices directly from the decentralized exchange. In our research, we focused on understanding the mechanism of the latter option. Understanding the safety of a DEX-based oracle starts from the deep understanding on how DEXs work; nowadays, they function using the automated market-making (AMM) mechanisms and the asset's price discovery happens along the curve of the AMM cost function. We reviewed the most widely used AMM cost functions and derived the cost of an attack for them. The next step was to look at the various aggregation methods; because using the spot price directly from DEX can lead to cheap price manipulations, most of the DeFi applications aggregate historical spot prices over a certain window size to decrease the chance of an attack. Depending on the method implemented and the window size, the target manipulation price can be higher or lower. We have provided equations to estimate the target attack price based on the aggregation method. We then developed the algorithmic model to estimate the safety of a DEX-based oracle on the example of a lending protocol. A step-by-step algorithm considers protocol-specific, oracle-specific and DEX-specific parameters and provides the logic on how to proceed with deciding on the safety of an oracle. Although we used the lending protocol as an example of a DeFi application using a DEX-based oracle, the model we introduced can be easily generalized to other types of protocols by changing the protocol-specific parameter (LTV in our example).

Incidents that happen in the new field of decentralized finance often lead to the crisis of trust from users and have a large social impact on the entire industry. Despite the crucial role oracles play in decentralized finance, their underlying mechanics are still under-explored and poorly understood which resulted in several protocol exploits [13–15]. However, we see the growing interest from both academia and industry practitioners to improve the oracles' resistance to manipulation attacks—new AMM curves are being introduced [40–42], oracle research is growing and more protocols are aware of price manipulation attacks. There is still a lot of work that can be done to achieve the goal of a safe decentralized price oracle in every layer—protocols using oracles can improve their risk management strategies, the AMM cost function can contribute a lot to the safety of oracles, as we saw in Sections 4.3 and 4.4, where the different pricing curves result in the different costs of attack. Finding the optimal AMM cost function that would minimize the chances of manipulation

is not a trivial task, and new pricing curves are being proposed by academia [40,42] and implemented in practice [41]. We hope to see more work performed in this direction. More research can be conducted about information aggregation methods as well—for economic reasons, currently, protocols are using simple statistical methods such as the TWAP. Finding an optimal solution that is less sensitive to the outliers and at the same time has a high price precision and cheap gas cost is still an open question at the moment. Overall, oracles in decentralized finance remain one of the most important and under-researched topics in the field with a huge impact on the entire cryptocurrency system.

Author Contributions: Conceptualization, A.T.A. and M.A.B.; methodology, A.T.A. and M.A.B.; validation, A.T.A. and M.A.B.; formal analysis, A.T.A. and M.A.B.; investigation, A.T.A. and M.A.B.; resources, A.T.A. and M.A.B.; data curation, A.T.A.; writing—original draft preparation, A.T.A.; writing—review and editing, A.T.A. and M.A.B.; visualization, A.T.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to acknowledge Delphi Labs and Jonathan Erlich for their help and fruitful discussions.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LTV	Loan to Value
AMM	Automated Market Maker
TWAP	Time-Weighted Average Price
DEX	Decentralized Exchange

Appendix A. Attack Cost Calculation for Constant Product Market

In constant product AMMs, the liquidity is defined as follows:

$$x \times y = k \tag{A1}$$

If an attacker wants to move the spot price of an asset, they would need to swap against the pool, depending on whether they want to manipulate the price up or down. If we assume that they want to increase the price of y, they would need to sell some amount of x and receive some y tokens in return (adding x tokens from the pool, an attacker decreases its value). After the swap, the liquidity in the pool will be as follows:

$$(x + \Delta x)(y - \Delta y) = k = xy \tag{A2}$$

From here, we can express the Δy—how much of token y attacker would receive after making a swap:

$$\Delta y = y(\frac{\Delta x}{x + \Delta x}) \tag{A3}$$

Because the attacker wants to increase the price y by adding the Δx tokens and removing Δy tokens, we can express how big the change of manipulated price p_j would be:

$$p_j = \frac{x + \Delta x}{y - \Delta y} = \frac{x + \Delta x}{y - y(\frac{\Delta x}{x + \Delta x})} = \frac{(x + \Delta x)^2}{xy} = \frac{(x + \Delta x)^2}{k} \tag{A4}$$

From here, we can express the Δx—how many tokens x are needed to get the target p_j:

$$\Delta x = \sqrt{p_j \times y \times x} - x \tag{A5}$$

Now, when we know both how many tokens x will be needed to make a swap and how many tokens y we receive in exchange, we can easily calculate the total attack cost by subtracting $\Delta x - \Delta y$:

$$AC = (\sqrt{p_j \times y \times x} - x) - p_j \times y \frac{\sqrt{p_j \times y \times x} - x}{x + (\sqrt{p_j \times y \times x} - x)} \tag{A6}$$

Appendix B. Attack Cost Calculation for Stableswap Market

The stableswap formula first introduced by Curve protocol is as follows:

$$2^2 A(x+y) + D = 2^2 AD + \frac{D^3}{2^2 xy} \tag{A7}$$

To be able to calculate the cost of an attack, we first need to express the y—how it is valued in terms of token x. For this, we rearrange Equation (8) in the form of quadratic equation which allows us to easily obtain the y formula:

$$y = \frac{(1 - \frac{1}{A}) \times (\frac{D}{4} - x) + \sqrt{[(1 - \frac{1}{A}) \times \frac{D}{4} - x]^2 + \frac{4D^3}{16Ax}}}{2} \tag{A8}$$

The derivative of the function in Equation (A8) would stand for the price:

$$y' = 0.5 \frac{[-1 + \frac{1}{2}(-\frac{D^3}{Ax^2} - 2(1 - \frac{1}{A})D - x)]}{\sqrt{\frac{D^3}{Ax} + ((1 - \frac{1}{A})D - x)^2}} \tag{A9}$$

To calculate the number of tokens Δx an attacker needs to swap to move the price to a specific target, we numerically solve for x in the price formula obtained in the previous step.

After we find how many tokens Δx are needed to have the price y' (or p_j to be consistent with Appendix A), we use that number to figure out how many tokens Δy would be in the pool after the swap by using Equation (A8) with the known x.

Finally, we can calculate the attack cost by subtracting the difference in token numbers before and after the swap: $AC = \Delta x - \Delta y$.

References

1. Caldarelli, G. Who Is Contributing to Academic Research on Blockchain Oracles? A Bibliometric Analysis. *Preprints* **2021**. Available online: https://www.preprints.org/manuscript/202109.0135/v1 (accessed on 5 November 2022).
2. Zhou, L.; Xiong, X.; Ernstberger, J.; Chaliasos, S.; Wang, Z.; Wang, Y.; Qin, K.; Wattenhofer, R.; Song, D.X.; Gervais, A. SoK: Decentralized Finance (DeFi) Incidents. *arXiv* **2022**, arXiv:abs/2208.13035.
3. Wu, S.; Wang, D.; He, J.; Zhou, Y.; Wu, L.; Yuan, X.; He, Q.; Ren, K. DeFiRanger: Detecting Price Manipulation Attacks on DeFi Applications. *arXiv* **2021**, arXiv:abs/2104.15068.
4. Zhou, L.; Qin, K.; Cully, A.; Livshits, B.; Gervais, A. On the Just-In-Time Discovery of Profit-Generating Transactions in DeFi Protocols. In Proceedings of the 2021 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 24–27 May 2021; pp. 919–936.
5. Werner, S.M.; Perez, D.; Gudgeon, L.; Klages-Mundt, A.; Harz, D.; Knottenbelt, W.J. SoK: Decentralized Finance (DeFi). *arXiv* **2021**, arXiv:abs/2101.08778.
6. Eskandari, S.; Salehi, M.; Gu, W.C.; Clark, J. SoK: oracles from the ground truth to market manipulation. In Proceedings of the 3rd ACM Conference on Advances in Financial Technologies, Arlington Virginia, VA, USA, 26–28 September 2021.
7. Al-Breiki, H.; ur Rehman, M.H.; Salah, K.; Svetinovic, D. Trustworthy Blockchain Oracles: Review, Comparison, and Open Research Challenges. *IEEE Access* **2020**, *8*, 85675–85685. [CrossRef]
8. Lo, S.K.; Xu, X.; Staples, M.; Yao, L. Reliability analysis for blockchain oracles. *Comput. Electr. Eng.* **2020**, *83*, 106582. [CrossRef]

9. Liu, B.; Szalachowski, P. A First Look into DeFi Oracles. In Proceedings of the 2021 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS), Online, 23–26 August 2021, pp. 39–48.
10. Steve Ellis, A.J.; Nazarov, S. Chainlink: A Decentralized Oracle Network. Available online: https://research.chain.link/whitepaper-v1.pdf (accessed on 5 November 2022).
11. Chainlink 2.0: Next Steps in the Evolution of Decentralized Oracle Networks. Available online: https://chain.link/whitepaper (accessed on 5 November 2022).
12. Oraclize. A Scalable Architecture for On-Demand, Untrusted Delivery of Entropy. Available online: https://provable.xyz/papers/random_datasource-rev1.pdf (accessed on 5 November 2022).
13. bZx Hack II Full Disclosure (With Detailed Profit Analysis). Available online: https://peckshield.medium.com/bzx-hack-ii-full-disclosure-with-detailed-profit-analysis-8126eecc1360 (accessed on 5 November 2022).
14. Understanding the Cream Finance Hack. Available online: https://medium.com/@AndyPavia/swissblock-post-mortem-cream-finance-hack-7c1caff4335c (accessed on 5 November 2022).
15. Harvest Finance. Available online: https://rekt.news/harvest-finance-rekt/ (accessed on 5 November 2022).
16. Mackinga, T.; Nadahalli, T.; Wattenhofer, R. TWAP Oracle Attacks: Easier Done than Said? *Cryptology ePrint Archive*, Paper 2022/445, 2022. Available online: https://eprint.iacr.org/2022/445 (accessed on 5 November 2022).
17. Othman, A.; Sandholm, T. Automated market-making in the large: the gates hillman prediction market. In Proceedings of the 11th ACM conference on Electronic commerce, Cambridge, MA, USA, 7–11 June 2010.
18. Goel, S.; Pennock, D.M.; Reeves, D.M.; Yu, C. Yoopick: A Combinatorial Sports Prediction Market. In Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, Chicago, IL, USA, 13–17 July 2008; pp. 1880–1881.
19. Othman, A. Automated Market Making: Theory and Practice. Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 2012.
20. Artzner, P.; Delbaen, F.; Eber, J.M.; Heath, D. Coherent Measures of Risk. *Math. Financ.* **1999**, *9* 203–228. [CrossRef]
21. Hanson, R.D. Combinatorial Information Market Design. *Inf. Syst. Front.* **2003**, *5*, 107–119. [CrossRef]
22. Hanson, R.D. Logarithmic Markets Scoring Rules For Modular Combinatorial Information Aggregation. *J. Predict. Mark.* **2012**, *1*, 3–15. [CrossRef]
23. Othman, A.; Pennock, D.M.; Reeves, D.M.; Sandholm, T. A Practical Liquidity-Sensitive Automated Market Maker. *ACM Trans. Econ. Comput.* **2013**, *1*, 14:1–14:25. [CrossRef]
24. VC Carvalho, A.; Silveira, D.; Ely, R.A.; Cajueiro, D.O. *A Logarithmic Market Scoring Rule Agent-Based Model to Evaluate Prediction Markets*; SSRN: Amsterdam, The Netherlands, 2022.
25. Peterson, J.; Krug, J.; Zoltu, M.; Williams, A.K.; Alexander, S. Augur: A Decentralized Oracle and Prediction Market Platform. *arXiv* **2022**, arXiv:1501.01042.
26. Gnosis Whitepaper. Available online: https://github.com/gnosis/research/blob/master/gnosis-whitepaper.pdf (accessed on 5 November 2022).
27. Berg, J.A.; Fritsch, R.; Heimbach, L.; Wattenhofer, R. An Empirical Study of Market Inefficiencies in Uniswap and SushiSwap. *arXiv* **2022**, arXiv:2203.07774.
28. Hayden Adams, N.Z.; Robinson, D. Uniswap v2 Core. Available online: https://uniswap.org/whitepaper.pdf (accessed on 5 November 2022).
29. Adams, H.; Zinsmeister, N.; Salem, M.; Keefer, R.; Robinson, D. Uniswap v3 Core. Available online: https://uniswap.org/whitepaper-v3.pdf (accessed on 5 November 2022).
30. Angeris, G.; Chitra, T. Improved Price Oracles: Constant Function Market Makers. In Proceedings of the 2nd ACM Conference on Advances in Financial Technologies, New York, NY, USA, 21–23 October 2020.
31. Angeris, G.; Kao, H.T.; Chiang, R.; Noyes, C.R.; Chitra, T. An Analysis of Uniswap Markets. *Cryptoecon. Syst.* **2021**, *1*. [CrossRef]
32. Egorov, M. StableSwap-Efficient Mechanism for Stablecoin liquidity. Available online: https://wikibitimg.fx994.com/attach/2020/10/189869321/WBE189869321_21425.pdf (accessed on 5 November 2022).
33. Xu, J.; Paruch, K.; Cousaert, S.; Feng, Y. SoK: Decentralized Exchanges (DEX) with Automated Market Maker (AMM) protocols. *arXiv* **2021**, arXiv:abs/2103.12732.
34. Hoyte, D. Median Oracle. Available online: https://github.com/euler-xyz/median-oracle (accessed on 5 November 2022).
35. Gudgeon, L.; Werner, S.M.; Perez, D.; Knottenbelt, W.J. DeFi Protocols for Loanable Funds: Interest Rates, Liquidity and Market Efficiency. In Proceedings of the 2nd ACM Conference on Advances in Financial Technologies, New York, NY, USA, 21–23 October 2020.
36. Aave Protocol Whitepaper. Available online: https://github.com/aave/aave-protocol/blob/master/docs/Aave_Protocol_Whitepaper_v1_0.pdf (accessed on 5 November 2022).
37. R. Leshner, G.H. Compound: The Money Market Protocol. Available online: https://compound.finance/documents/Compound.Whitepaper.pdf (accessed on 5 November 2022).
38. Juliano, A. dYdX: A Standard for Decentralized Margin Trading and Derivatives. Available online: https://whitepaper.dydx.exchange (accessed on 5 November 2022).
39. The Maker Protocol: MakerDAO's Multi-Collateral Dai (MCD) System. Available online: https://makerdao.com/en/whitepaper (accessed on 5 November 2022).

40. Krishnamachari, B.; Feng, Q.; Grippo, E. Dynamic Curves for Decentralized Autonomous Cryptocurrency Exchanges. *arXiv* **2021**. [CrossRef]
41. Egorov, M. Automatic Market-Making with Dynamic Peg. 2021. Available online: https://classic.curve.fi/files/crypto-pools-paper.pdf (accessed on 5 November 2022).
42. Port, A.; Tiruviluamala, N. Mixing Constant Sum and Constant Product Market Makers. *arXiv* **2022**. [CrossRef]

entropy

MDPI

Article

An Edge-Supported Blockchain-Based Secure Authentication Method and a Cryptocurrency-Based Billing System for P2P Charging of Electric Vehicles

A. F. M. Suaib Akhter [1], Tawsif Zaman Arnob [2], Ekra Binta Noor [2], Selman Hizal [1] and Al-Sakib Khan Pathan [3,*]

[1] Department of Computer Engineering, Sakarya University of Applied Sciences, Serdivan 54050, Sakarya, Turkey
[2] Department of Mechanical and Production Engineering, Islamic University of Technology, Gazipur 1704, Bangladesh
[3] Department of Computer Science and Engineering, United International University (UIU), Dhaka 1212, Bangladesh
* Correspondence: sakib.pathan@gmail.com

Abstract: The popularity of electric vehicles (EVs) is constantly increasing, as they use relatively greener, sustainable energy. However, it is a fact that the charging stations for EVs are yet to meet the demand. It could be a great solution if a peer-to-peer (P2P) charging system could be initiated by anyone who wants to make their garage's charge points publicly available for commercial purposes, named a home charging station (HCS). In this work, our idea is to bring interested charging stations under a network of nodes and a blockchain-based management system, where the blockchain is responsible for ensuring the authenticity of both the charging stations and charge receiver. A cryptocurrency-based payment system has also been proposed to ensure transactions' security, integrity, transparency, and immutability. A reputation management system is applied to maintain the quality of service. Miners with high processing power are used to alleviate lagging during block creation, supported by edge servers. The proposed system has been implemented by using virtual machines. A theoretical analysis is presented to assess the compatibility and possible cost requirements to implement the system in a real-world scenario.

Keywords: blockchain; cryptocurrency; edge computing; electric vehicles; Ethereum; P2P charging

Citation: Akhter, A.F.M.S.; Arnob, T.Z.; Noor, E.B.; Hizal, S.; Pathan, A.-S.K. An Edge-Supported Blockchain-Based Secure Authentication Method and a Cryptocurrency-Based Billing System for P2P Charging of Electric Vehicles. *Entropy* **2022**, *24*, 1644. https://doi.org/10.3390/e24111644

Academic Editors: Stanisław Drożdż, Jarosław Kwapień and Marcin Wątorek

Received: 28 September 2022
Accepted: 10 November 2022
Published: 12 November 2022

1. Introduction

Vehicles with the potential of using renewable energy sources, such as electric vehicles (EVs), have caught worldwide attention in recent times [1–3]. These vehicles do not depend on fossil fuels but use other renewable energy sources to reduce gas emissions. As stated in [4,5], by the year 2040, it is projected that renewable energy is to come to equivalence with coal and natural gas-based electricity generation. Additionally, the EV stock is expected to reach at least 140 million by 2030. However, it will take a long period of time to integrate them efficiently within the infrastructure. At present, EV users are hesitant to use their vehicles for long drives, and potential customers go through the dilemma of choosing an EV over a traditional vehicle, as there are so few charging stations. Level 2 charging equipment can provide a vehicle with 10 to 20 miles of range for every hour of charging. With the necessary set-up, anyone can make their garage available to charge an EV for long-distance traveling [6]. This could provide an abundance of charging stations across the country within the existing infrastructure. Our idea in this setting is that the EVs and the home charging stations would be two different party nodes under a secure and reliable system with availability, security, preservation of privacy, and payment facilities.

Malicious operators can seriously threaten EVs' security and privacy through various malicious exploitations [7–9], e.g., privacy leakage, falsification, node impersonation, or

advertising fraudulent charging services. To provide secure charging services for EVs, many innovative mechanisms have been proposed so far, and some are even implemented to some extent [10,11]—e.g., trust mechanism and monetary approaches. However, the trust mechanism is not sustainable and susceptible to Sybil attacks and whitewashing attacks, and the monetary approach relies on trusted centers. Trusted centers may not only leak users' private information for profit, but also may be vulnerable to attacks. In this context, blockchain [12] offers a unique platform for secure energy transactions within a distributed network without trusted agents through the use of an immutable ledger, cryptocurrency, and the execution of smart contracts.

As we know from various recent works and interest shown by a wide range of researchers, blockchain technology can come in handy for various management systems. A blockchain is a decentralized, distributed, open ledger, and each node in the network has a copy of the ledger. It was developed as a peer-to-peer network without third-party intervention [13]. The blockchain's integrity is based on strong cryptography and hash functions that provide validation and chain blocks together on transactions, making it nearly impossible to tamper with a block or any individual transaction without being detected [14].

As the number of online systems has increased, we have witnessed that the threats of various types of cyber attacks have also increased significantly. Among the various types of attacks, unauthorized entities or malware-based attacks can cause fatal damage to the system [15]. Thus, a deficiency in the proper authentication process can make a P2P system vulnerable to various types of attacks. In our proposed system, a blockchain-based authentication system is used, so that before making an agreement, the entities (EVs and the HCS) can check the authenticity of each other. Again, while getting services from the HCS, a proper charging measurement system is essential to calculate the amount of the electricity that has been exchanged from the HCS to the EV. Moreover, it is also required to determine the number of bills to be paid. In our system, a smart meter is used to calculate the amount of charging, and the HCS would share that by using the blockchain.

Due to the popularity of cryptocurrency in the financial sector, researchers have also started utilizing it in various fields. In fact, the transparency, trustworthiness, worldwide availability, convenient exchange facilities, ease of access, minimum transaction cost, etc., encourage the business world to utilize cryptocurrency [16,17]. Hence, in this study, a cryptocurrency-based payment system has been used for the system: after calculating the amount (to be paid), the smart meter requests a transaction in the blockchain. The system will automatically deduct the amount from the EV, which will be credited to the HCS's account. After each transaction, the service receiver, i.e., the EV owner, can provide feedback about the service received. It will help the server suggest that HCS out of those nearby, as an HCS with a higher rating will come before one with a lower rating. As generating blocks for a blockchain requires high computational power, edge computing services have been used to perform the complex mathematical calculations required for mining.

In this work, we used a combination of multiple protocols, and the main contributions of this can be summarized as follows:

- A blockchain-based electric vehicle charging management system is proposed where an EV can receive charging (or recharge) when necessary from anywhere in the world. HCSs from anywhere in the world can join the network and can earn money by providing charging services to EVs. Management, searching, etc.—related services—would be provided by the blockchain.
- To avoid unauthorized access, malware, DDoS (distributed denial-of-service), or any other security attacks, a blockchain-supported authentication system is employed. Additionally, the system also preserves the privacy of the members.
- To remove confusion, miscalculations, etc., a smart billing system is proposed in this paper where two different agents are responsible for measuring the amount of charge exchanged between EVs and HCSs. Moreover, to avoid the hassle of payment, the

smart billing system (SBS) automatically calculates the amount to be paid and creates a transaction in the blockchain after the charging process is finished.
- To ensure safe, secure, transparent, and authentic payment, a cryptocurrency-based payment system is proposed, which was developed by using the Ethereum blockchain. The payment system is handled by the blockchain, and currency transfers will happen automatically.
- To ensure the transparency and quality of service (QoS), a reputation management system is also proposed, where EVs have the opportunity to express the grade of the service received.
- The system was developed using the Ethereum blockchain, with which the authentication, billing, payment, and reputation management system were simulated.

To explain the full system, the paper is organized as follows: Section 2 presents the motivations for the proposed system, together with notable research. Section 3 gives some preliminary knowledge on the issues and definitions that could help the general readers get useful information, and it establishes the importance of this work. Section 4 describes the system's architecture with its components and transactions. The implementation details are provided in Section 5. Then, Section 6 contains the performance analysis, and in Section 7, challenges and limitations are discussed. The paper concludes with Section 8 with future research directions.

2. Related Work and Motivation

The energy-sharing method is not new, yet the popularity of using EVs has unlocked a vast area of research. In this section, we will present some previously published energy-sharing-related methods.

In [18], Zhang et al. depicted a typical incentive-based approach in the smart grid environment and explored vehicle-to-vehicle (V2V) scenarios. This is a cloud-based energy trading process with a contract theory approach. Tushar et al. introduced an incentive game-based mechanism for distributed renewable energy management in a smart community [19] to improve the operator's profit and minimize total energy trading cost. Bera et al. [20] introduced a novel cooperative energy consumption system within communities in the smart grid to mitigate energy consumption costs for users and reduce the peak-to-average ratio. In [21], a global control scheme is proposed for electric energy micro-storage systems in smart communities to improve the local power quality of demanded and current power consumption globally.

Sharing energy between two peers requires ensuring the security services, including authenticity, privacy, integrity, attack prevention capability, etc. To provide these, in [22], a token-based decentralized energy trading system was shown that enables peers to perform transactions anonymously and securely. The system was developed by using multi-signature and anonymous encryption methods. Li, Z. et al. [23] provided a secure distributed energy trading market and designed a novel energy blockchain system in the industrial Internet of things (IIoT) environment. They implemented the system by using a consortium blockchain. Li, L. et al. [24] presented a novel announcement network named credit-coin that uses blockchain technology to protect vehicles' privacy and motivate users to broadcast traffic information. In [25], a pragmatic blockchain utilization case is introduced for machine-to-machine (M2M) transactions of energy within the housing society environment. Based on the lightning network and smart contract in the energy blockchain ecosystem, Huang et al. presented a decentralized security model for the enhancement of the security of trading between EVs and charging piles in the peer-to-peer (P2P) network [26].

A secure way to pay and proper pricing need to be ensured for a P2P EV charging system. To do that, Zou et al. designed a progressive second-price auction game mechanism for resolving large-scale EV charging cooperation problems. They have ensured incentive compatibility over a finite horizon in their work [27]. Mohammadi et al. depicts a distributed cooperative charging scheme for plug-in electric vehicles (PEVs) to minimize

the charging cost for PEV fleets with the integration of a receding horizon method [28]. Liu et al. [29] proposed a novel renewable energy pricing scheme for smart communities to reduce the total electricity bill of the residential users utilizing an advanced cross-entropy optimization method in smart home energy scheduling. A contract game-based direct-energy trading system is proposed by Zhang et al. [30] for modeling the decision-making process of electricity operators and consumers in vehicular edge computing networks. Yang et al. [31] presented a coordinate EV charging mechanism in a microgrid-powered setting via wind-powered generators through a Markov decision process (MDP) approach. Utilizing stochastic dynamic programming methods, Wu et al. [32] proposed smart-home energy management integrated with PEVs (plug-in electric vehicles) to address the problem of intermittent renewable energy supplies to minimize the electricity cost.

With the popularity of blockchains, some other projects have been found where a blockchain is primarily utilized for EV charging. For example, in [33], a blockchain was proposed to ensure a secure and trusted electricity trading solution. The authors of [34] utilized blockchain to create a trusted distributed environment for charge sharing. In [26], a blockchain was used for charging management, and in [7], it was used to store the trading records between EVs and charging stations.

A two-stage autonomous EV charging coordination method implemented on blockchain was shown by Ping et al. [35] to enable dependable EV charging coordination in the absence of a third-party coordinator. This mechanism also preserves the privacy of the users. Wang et al. developed an optimization model on a blockchain framework to manage the operation of crowdsourced energy systems (CESs) with peer-to-peer (P2P) energy trading transactions (ETTs) [36]. One of the new paradigms created by the decarbonization, decentralization, and digitalization of the energy supply chain (that enables direct exchange between energy users and producers) is depicted in [37]. Chen et al. proposed an energy trading framework that marries a blockchain and distributed optimization; the blockchain enables checks and balances among the participants and disables dishonesty [38].

Consider the above-mentioned papers. Our take is that they provide partial solutions in terms of P2P charge-sharing systems. A model is required where there will be secure communication and management, easy and time-saving payment facilities, and reliable quality of service. Hence, a complete solution is proposed where all the required features are available for the users. Additionally, in almost all of these previous work, there was no implementation to show the compatibility and validity of their approaches. Hence, there is indeed a gap in the existing literature. We brought forth a real-world implementation by using Ethereum blockchain to show our system's compatibility, to understand the behavior of the components and responses, and to collect important data from the system. The motivations for this work are presented in the next section.

Motivation

Electric vehicles are popular nowadays, and many are embracing the idea of an electricity-run car with the utmost interest. However, in reality, EV owners are still not confident enough or are often hesitant to go for long road trips. This is mainly due to the relatively small number of available charging stations compared to conventional fuel stations. If the car has a dual mode that runs both on electricity and traditional fuel, the problem is making a decision of how much fuel is to be always carried (how frequently the fuel tank needs to be filled up) and how frequently to charge the battery of the EV. This adds an extra layer of decision making for EV owners. In the usual case, an owner would like to run the vehicle on electricity, as it is designed to run in that way, which makes it often significantly costlier than many other regular cars. Thus, the issue we have here is that we need some kind of efficient charging mechanism or model to support long journeys by EVs. Some previous works tried to solve this issue in some ways, but as mentioned in the previous section, almost all of them fall short of the required efficiency, and some have not even given enough thought to this issue; i.e., models are available, but they are not

comprehensive or advanced enough. We found a gap in the existing literature on this issue, and that basically motivated us to devise our mechanism.

P2P systems can be used for this issue of charging, but they would require trust and security, as complete strangers can come to a charging station. In this case, a good way to think about getting a solution is that anyone who is registered under the blockchain would be allowed to get the service. The blockchain makes the system secure and trustworthy, preventing malicious attacks while recording each transaction. This would create an increase in the number of charging stations throughout a country, creating a source of income for charge-point providers without the intervention of a third party.

Furthermore, it can reduce the pressure on expanding the infrastructure for setting up new charging stations to meet the rising demands of EV charging. This P2P charging environment will enable an opportunity to see an increase in the frequency of EVs in sectors such as delivery and medical (ambulance), where EVs are still considered inefficient. As we aim to move on to a greener future, we need to exploit our infrastructure the most we can without consuming more areas to set up charging stations.

To minimize the hassles such as platform dependency of the payment system, uncontrolled pricing, and delays because of payment confirmation, in the proposed method, all those issues are solved by using blockchain-based cryptocurrency. A charging management system is responsible for measuring the transferred energy, and it calculates the price according to that. Then, the system can directly make a transaction in the blockchain, and the amount will be automatically deducted from the service receiver's account. In this way, the payment will be much easier, easy to handle, and time-saving for both parties.

A reputation management system is added to the proposed method where EVs can provide a review about the service received, more specifically, about the HCS. Upon receiving a review, the server will calculate the mean of overall ratings for each HCS, and while suggesting nearby HCSs to the EVs, HCSs with relatively higher ratings will get priority on the list. Moreover, as the HCSs have the authority to fix their own pricing, they can balance the ratings and pricing.

3. Prior Knowledge

3.1. Blockchain

It can be strongly said that blockchain is the future of secure currency management. Security, integrity, worldwide availability, preservation of privacy, immutability, transparency, etc., are the basic services provided by blockchains [39]. Distributed storage and decentralized storing system are considered as additional advantages. To understand the popularity, the amounts of money invested (over the years) by different industries for blockchain are presented in Figure 1 (adopted from [40]).

Protection from several types of attacks is another special feature of blockchains. Due to its decentralized and distributed storage technique, many typical attacks, including Sybil attacks, unknown source attacks, man in the middle (MITM) attacks, unauthorized entry, and DDoS (distributed denial-of-service), are not possible to perform on a blockchain. Additionally, consensus protocols put on another level of security on this, which can ensure the integrity and stability of the information.

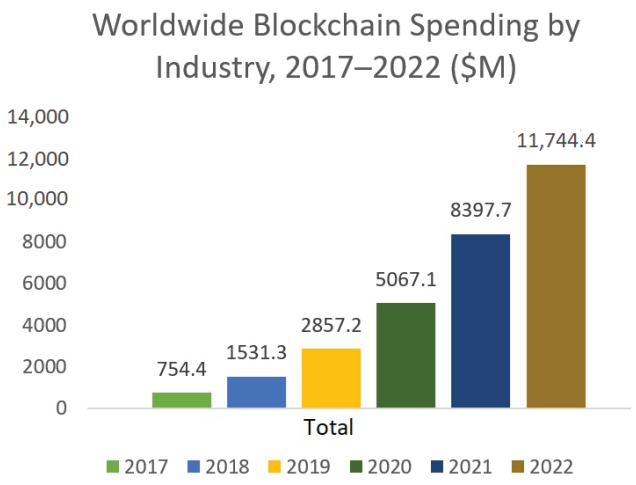


Figure 1. Worldwide blockchain spending by industry.

3.2. Ethereum

Ethereum was selected as the blockchain for the proposed method, which brings several advantages to the system. Although in terms of popularity, Ethereum loses first position to Bitcoin, it still has some very exclusive features that make it a popular choice for industries. The most relevant feature of Ethereum is the smart contract [41], which makes it a digital asset management system rather than just a money transfer system. Due to smart contracts, it is possible to manage the full system by using one platform, which is not possible for other types of blockchain, such as Bitcoin, Zcash, Dash, Peercoin, Ripple, Monero, and Multichain.

Ethereum supports more transactions per second than most of the other blockchains; again, Ethereum does not have any coin limit. On the other hand, while most of the blockchains support the latest scripting language by Bitcoin called Bitcoin Script [42], Ethereum supports multiple languages that are similar to the most popular languages. Examples include Solidity, which is similar to Java Script and C; Serpent, which is similar to Python; and LLL, which is similar to Lisp. Table 1 shows the advantages of using Ethereum over other blockchains.

Table 1. Comparison of the Ethereum blockchain with others.

Blockchain	Symbol	Scripting Language	Implementation Language	Average TPS [1]	SC [2] Support
Ethereum	ETH	Solidity, Serpent, LLL	Go-Ethereum, CPP-Ethereum, Py-Ethereum, EthereumJ, Parity	5.40	Yes
Bitcoin	BTC/XBT	Bitcoin Script	C++	3.50	No
Zcash	ZEC	Bitcoin Script	C++	0.06	No
Litecoin	LTC	Bitcoin Script	C++	0.35	No
Dash	DASH	Bitcoin Script	C++	0.07	No
Peercoin	PPC	Bitcoin Script	C++	0.01	No
Ripple	XRP	N/A	C++	10.75	No
Monero	XMR	N/A	C++	0.06	No
MultiChain	-	Bitcoin Script	C++	1000	No
Hyperledger	-	Go, Node.js, Java, C++, Python and more	Go, Python and More	Various	Yes

[1] TPS = transactions per second. [2] SC = smart contract.

3.3. Cryptocurrency

Cryptocurrency is a digital currency secured by cryptographic algorithms, and it provides high security, availability, transparency, etc. It can be said that cryptocurrency is the future of the economic world. In the proposed method, the payment system is managed by using *Ether*, which is a cryptocurrency supported by Ethereum blockchain. It is available all over the world, decentralized, cost-effective, time-saving transaction-wise facility, and self-governed; and its convenient money-exchange facilities make it one of the biggest currencies in the world. The use of *Ether* in the proposed method makes the payments automated, secured, easily accessible, time-saving, and hassle-free. In particcular, the EV does not have to wait after receiving charges, as the SBS automatically sends the billing information and makes the payment.

3.4. EDGE Computing

Due to consensus management, calculating complex cryptography and hash functions requires time to generate a block in the blockchain. On the other hand, as the proposed method would require mobile services, the computational workloads would be handed over to the edge computing servers. In fact, today's high-speed Internet, such as 5G and the upcoming 6G, will be highly efficient for accessing large amounts of data from edge servers. Thus, rather than arranging and spending a huge amount of money on a physical server setup, utilizing an edge server is proposed in this system.

4. System Structure

In this paper, a complete solution for an EV charging system is proposed. The system is comprised of three main protocols, which are: (1) The authentication protocol, (2) the smart billing system, and (3) the reputation management system. In this section, firstly, there are short descriptions of the components of the system, followed by detailed information about the communication, authentication, billing, and reputation management system.

4.1. Components

Four main components participate in this system. These are:

- Electric vehicles (EV).
- Home charging stations (HCS).
- Smart billing system (SBS).
- The blockchain and edge servers.

In Figure 2, visual representations of the components are given.

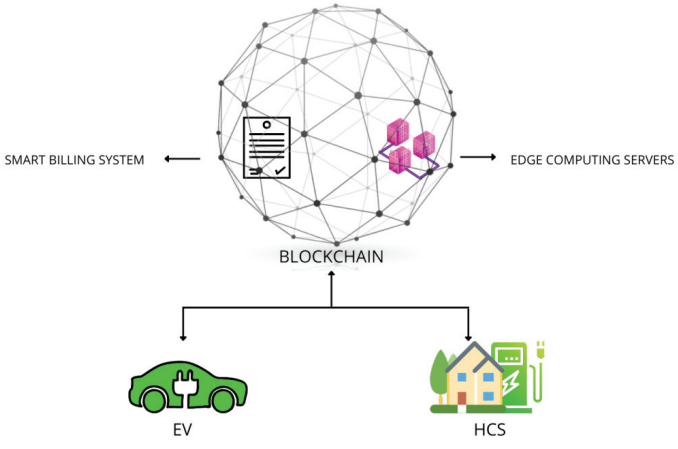


Figure 2. Components of the system.

4.1.1. Electric Vehicles (EV)

Vehicles that can utilize electricity (completely or partially) to store and later convert it to kinetic energy are known as EVs. In the proposed system, any EV can register with the system and get a charging facility from the registered HCS. To register with the system, a person with an EV charging facility has to provide their national identity information. The system will perform verification of the ID, phone number, and email address before accepting the person as a member. After receiving any service, EVs can send their feedback related to the service provider, and their ratings will be published publicly, which will help the EVs select the best service provider nearby.

4.1.2. Home Charging Stations (HCS)

A home charging station (HCS) is a station that is owned by an individual who may have an EV (or several), and that station is used to charge personal EVs or can be offered as a charging station (as a service) to other EVs commercially. HCS owners are also required to register with the system by providing detailed information: identity, charging equipment, facilities, capabilities, service time, location, etc. All the information is visible to the potential user before the decision to select the service. All the HCSs can choose their own pricing per kilowatt (KW), and the EV that selects a particular HCS will agree to that price. Every HCS will receive a rating point after providing any service, and the rating points will be available online. It will help the HCSs maintain the quality of service and the pricing level. Too-costly services may be given poor ratings by the service users (i.e., EVs' owners who have used the service).

4.1.3. Smart Billing System (SBS)

To calculate the amount of charge transferred from an HCS to an EV, a proper measurement system is used named SBS. SBS calculates the amount of charge in kilowatts (KW) and determines the amount to be paid according to the price asked by the HCS. In the proposed system, a smart meter is used to calculate the amount of charging, and the HCS will share it using the blockchain. As members of the blockchain, all the EVs and HCSs are connected to the blockchain by using a cryptocurrency. Once all calculations are done and fixed, the amount will be automatically deducted from the EV's account and then credited to the account of the HCS.

4.1.4. Blockchain

To join the proposed system, interested components (EVs, HCSs) are required to be registered by providing the necessary information and documentation. All the members will receive a pair of keys (public and private). The public key will be used as the member's identity, and all the communications will take place by using that. At the same time, the public key will hide the real identity of the member, and in this way, it can protect privacy as well. However, a typical blockchain has to go through a lot of complex calculations because of block generation and validation, which would require servers with high computational capabilities. Thus, an edge server is used to perform those calculations to minimize delays during transactions. With the authentication information, the blockchain is also responsible for storing all the outcomes of smart billing and reputation management system inside a transaction to ensure their security, integrity, availability, transparency, etc.

Any EV or HCS can register with the system by providing the required information and documents. After registration, it becomes a member of the blockchain and is able to perform transactions anytime over an Internet connection. An EV user that wants to charge its car can send a request for charging. Then, the system will suggest the nearby HCSs. Two factors will be applied while suggesting HCSs; one is the distance from the EV, and the other is the ratings of the HCSs. The EV can select the most suitable HCS among those on offer. When there is a mutual agreement between an EV and an HCS, the system will generate an ID for the transaction. In the future, information related to that charging will be identified uniquely by using that same ID. After receiving the service from an HCS, the

SBS will calculate the amount of energy (i.e., the amount of charge) transferred and send a transaction to the blockchain to store the information. At the same time, the system will deduct the payment from the EV and move it to the HCS's account. The flow of the system is illustrated in Figure 3.

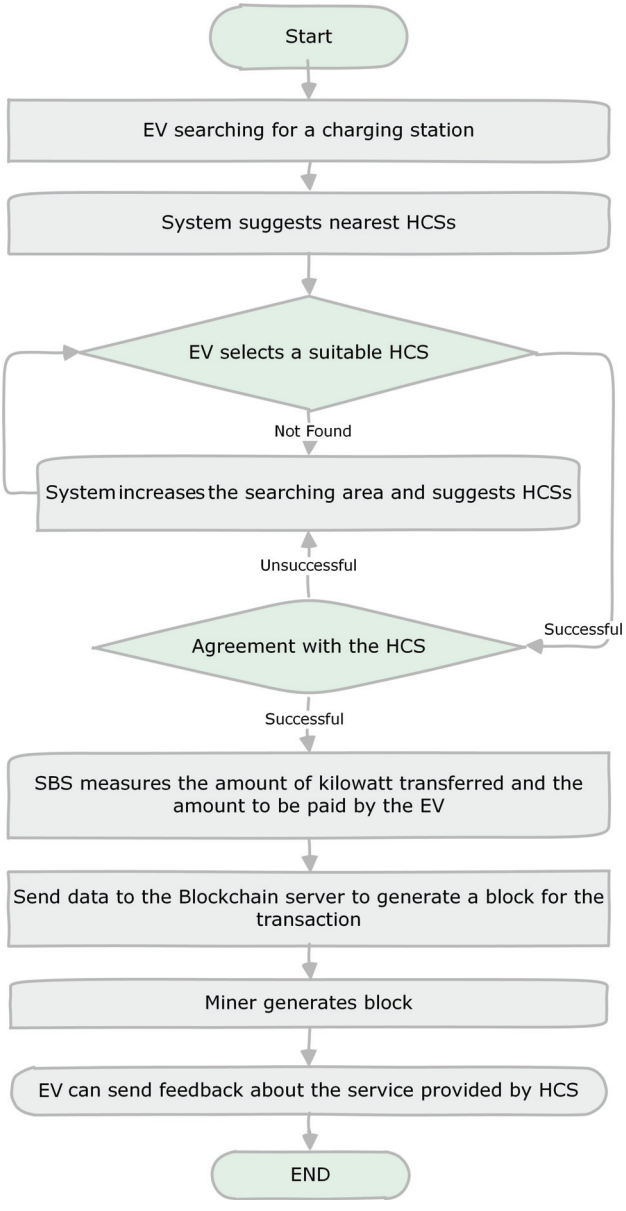


Figure 3. Workflow of the proposed system.

4.2. Communication System

To manage all the underlined protocols, several messages are passed among the blockchain's components and the servers. The message formats are presented in Figure 4. All the messages start with the *message type* field, which informs user about the kind of

information that resides inside the message and what possible actions are required to be taken.

Registration info. of the components. **Function name:** registration()

Message Type (4 bits)	Request for EV/HCS ? (2 bits)	Requester's Address (20 bytes)	Server Address (20 bytes)	Data (Legal documents, EV's License, driving license, etc.) (Flexible)

Authentication. **Function name:** reqAuthInfo() and verification()

Message Type (4 bits)	Message Type (Request / Response) (2 bits)	Response (Verified / Not Verified) (2 bits)	Sender's Address (20 bytes)	Receiver's Address (20 bytes)	Requested Address (20 bytes)	Optional Data

Searching for a nearby charge station. **Function name:** search()

Message Type (4 bits)	Requester's Address (20 bytes)	Server Address (20 bytes)	GPS information of the Requested Area (24 bytes)	Optional Data

Nearby HCSs (suggested). **Function name:** suggestion()

Message Type (4 bits)	Server Address (20 bytes)	Requester's / Receiver's Address (20 bytes)	List of nearby and available HCSs (Flexible)	Optional Data

Request to an HCS, Response to an EV. **Function name:** request() and response()

Message Type (4 bits)	Message Type (Request / Response) (2 bits)	Response Type (Request / Accept) (2 bits)	Sender's Address (20 bytes)	Receiver's Address (20 bytes)	Optional Data

Mutual agreement. **Function name:** mutual()

Message Type (4 bits)	Server Address (20 bytes)	EV's Address (20 bytes)	HCS's Address (20 bytes)	Transaction ID (32 bytes)	Optional Data

Completion/Termination of service. **Function name:** completion(), termination()

Message Type (4 bits)	EV's Address (20 bytes)	HCS's Address (20 bytes)	Transaction ID (32 bytes)	Status (Completion / Termination) (4 bits)	Optional Data

Amount of charge received/provided by the members. **Function name:** chargeReceived(), chargeProvided()

Message Type (4 bits)	Member Type (EV / HCS) (4 bits)	EV's Address (20 bytes)	HCS's Address (20 bytes)	Transaction ID (32 bytes)	Amount of Charge Transferred (20 bytes)	Optional Data

Amount of Ether to be paid. **Function name:** amount()

Message Type (4 bits)	Server Address (20 bytes)	Transaction ID (32 bytes)	Sender's Address (20 bytes)	Receiver's Address (20 bytes)	Amount (32 bytes)	Payment Status (4 bits)	Optional Data

Rating for a service. **Function name:** rating()

Message Type (4 bits)	Transaction ID (32 bytes)	Sender's Address (20 bytes)	Receiver's Address (20 bytes)	Rating (1 byte)	Comments (10 bytes)	Optional Data

Request for a block generation. **Function name**: transaction()

Message Type (4 bits)	Transaction ID (32 bytes)	Sender's Address (20 bytes)	Receiver's Address (20 bytes)	Amount (32 bytes)	Payment Status (4 bits)	Rating (1 byte)	Comments (10 bytes)	Optional Data

Figure 4. Message Formats.

According to the requirements, addresses, i.e., public keys, of EVs, HCSs, and servers, are added. However, to understand the format more clearly, sender/requester and receiver are shown on the figure. Completion and termination functions also do similar operations.

Some packet formats are applicable for multiple functions. For example, the same format is used for the *request()* and *response()* functions, where there is a field that informs the receiver about the type of message (request/response) and the status of the message. All the messages have a special field called optional data, which can be used for different purposes. Which function is used by which entity is presented in Table 2.

Table 2. Smart contract functions in the proposed system.

Responsibility	Function Name	Performed By		
		EVs	HCSs	Server
Registration info. of the components	registration()	✓	✓	-
Request for authentication info.	reqAuthInfo()	✓	✓	-
Verification of Authenticity	verification()	-	-	✓
Searching for a charge station nearby	search()	✓	-	-
Nearby HCSs (suggested)	suggestion()	-	-	✓
Request to an HCS	request()	✓	-	-
Response to an EV	response()	-	✓	-
Mutual agreement	mutual()	✓	✓	✓
Completion of service	completion()	✓	✓	-
Termination of a service	termination()	✓	✓	-
Amount of charge received by EV	chargeReceived()	✓	-	-
Amount of charge provided by HCS	chargeProvided()	-	✓	-
Amount of *Ether* to be paid	amount()	-	-	✓
Rating for a service	rating()	✓	-	-
Request for a block generation	transaction()	-	-	✓

4.3. Authentication Protocol

In the proposed system, blockchain is used to confirm the authenticity of the members, i.e., EVs and HCSs. All of them are required to be registered physically before getting services from the system. They receive a pair of keys after the registration, and later all the communications will take place with their public keys. Before generating any request of charging, the system checks the membership status of the EVs, and similarly, before suggesting nearby HCSs, the system checks the authenticity of the HCSs. Moreover, to ensure the authenticity of a particular EV or HCS, any of the members can send a request for authentication information of another component by sending a message to the server by using *reqAuthInfo()*. Then, the server will reply with the authenticity of the requested components. In this way, the authenticity of the components is ensured so that both parties can initiate a safe and secure connection. Moreover, using the public keys instead of real identities will protect their original identities and privacy.

4.4. Smart Billing System

The billing system for electric vehicles is a central hub that manages the exchange of electricity between the charging station and the EV. This proposed smart billing system (SBS) has two entities: a power management entity and a communication entity. The communication entity can also exchange information with other entities using wireless communication. The power management entities reside inside both EVs and HCS. This entity calculates and reports the amount of power to be charged from both sides. After the charging is finished, SBS verifies the amount of charge that has been transferred and sends it to the blockchain server as a transaction. The operational flow of charging and billing is illustrated in Figure 5.

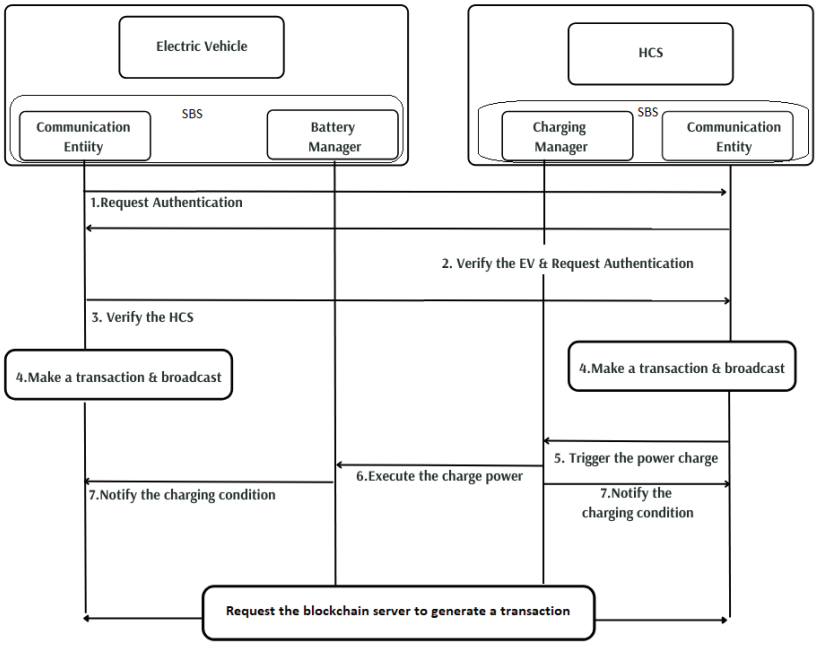


Figure 5. Operational flow of the billing system process.

4.5. Reputation Management System

A reputation management system was added to the system to ensure the quality of service. After receiving the charging service, the EV user can leave feedback about the received service. A rating from one to five can be provided, where five (5) indicates the best service and one (1) means the service is the worst possible. The server will calculate the mean of overall ratings for each HCS, and when suggesting nearby HCSs to the EVs, HCSs with relatively higher ratings will get priority on the list. Additionally, the EV users will be able to provide a comment describing the service received. After receiving the feedback from the EV, all the fields of the transaction are completed, and the server will generate the block from the transaction.

This will help HCSs to decide on their pricing, as they have the authority to fix their service rates; an HCS with a higher rating may ask for a higher price, and a newcomer may ask for a lower price to get good review scores. In this way, the reputation system will also be useful to create competition between the service providers, i.e., HCSs.

5. Implementation

To emulate the proposed blockchain-based P2P EV charging system, a virtual environment was created. Several virtual machines were prepared to represent EV, HCS, and blockchain servers. It was assumed that a specific amount of charge was transferred from an HCS to an EV, and the SBS requested a transaction in the blockchain. To simulate the blockchain, a blockchain testing platform called *Truffle* was used [43]. This platform provides a real blockchain with smart contract programming facilities. It provides *Ganache* [44], which simulates a real dummy of the Ethereum blockchain and additionally provides programming ability, customization, monitoring, debugging facilities, etc. The smart contract was written in Solidity programming language and deployed using Truffle. To develop the client side, a lightweight node server [45] was used with Node Packet Manager (NPM) [46].

The target of the implementation was to simulate the transactions, blockchain-based operations such as block generation, and cryptocurrency-based payment management in a real-world environment. Thus, the Ethereum blockchain was selected as the blockchain,

and *Ether* as the cryptocurrency. However, some of the transactions were not simulated to simplify the experimental analysis.

5.1. Experimental Setup

The following steps can explain the experimental setup:

- For the experiment, multiple virtual machines were used via VM VirtualBox 6.1. Four VMs were set up: two of them represented EVs (named EV_1 and EV_2), and the other two were HCSs (named HCS_1 and HCS_2).
- Another one was set up in the blockchain server named *BCS*. Truffle platform and the Ganache blockchain were set up in the *BCS*. Moreover, for web hosting and management, a lightweight node server [45] and Node Packet Manager (NPM) were used.
- All the EVs and HCSs were considered as full members of the blockchain, and in *Ganache*, they were registered. Before beginning, 100 were assigned virtual Ether, which is the currency used by the Ethereum blockchain.
- EVs and HCSs use Metamask [47] as an Ethereum wallet, by which they can connect to the blockchain. Simultaneously, EVs can pay and HCSs can receive money.
- The communication module of the SBS is prepared for the experiment, and to simplify the experiment, instead of the charge measurement system, the amount of charge transferred from HCS to EV was assumed.
- The *Truffle* framework supports multiple smart contract programming languages. In this experiment, *Solidity* programming language was used to manage communication, block generation, and so on. For each and every activity, a function is responsible. Details of the functions are shown in Table 2. The structure of the transferred messages is illustrated in Figure 4.
- EVs can provide feedback after getting services from HCSs, which would be helpful in maintaining the quality of services.

5.2. Deploying the Blockchain

To run the experimental setup, firstly, the *Ganache* blockchain was deployed in the *BCS* machine. By default, *Ganache* generates some public keys for users, and all the users receive 100 *Ether* transactions. Each member VM ($EV_{1,2}$, $HCS_{1,2}$) got a public key and used that as its public identity. Then, the members joined the blockchain by using the *Metamask* wallet.

During development, we kept the amount of charge open to receive manual entry so that it could receive user input rather than automatic calculation by the charging agent. After deploying the blockchain, we requested different amounts of charging in KWs manually. After receiving the entry from the members, the SBS module generated the amount to be paid and requested a blockchain transaction. Due to the simplification, the proposed system can generate the block almost instantly after the request and broadcast it to all the members. During the block generation process, the amount of cryptocurrency, i.e., *Ether*, is deducted from the EV's account and credited to the HCS's account. Additionally, the service receiver EV can provide a rating score (out of 5), and the server will calculate the mean of all the reputation scores received by each HCS. The rating score will be available publicly. By using the web interface, all the members can check the global (and also own) transaction histories, financial statements, and rating points (provided or received).

Multiple transactions were performed to analyze the system. After running the simulated system, data were prepared manually, and by using smart contracts, information was added to the *transaction()* function according to the message structure (presented in Figure 4). After inserting those data, a transaction was performed in the blockchain. Details of some of the performed transactions can be found in Figure 6.

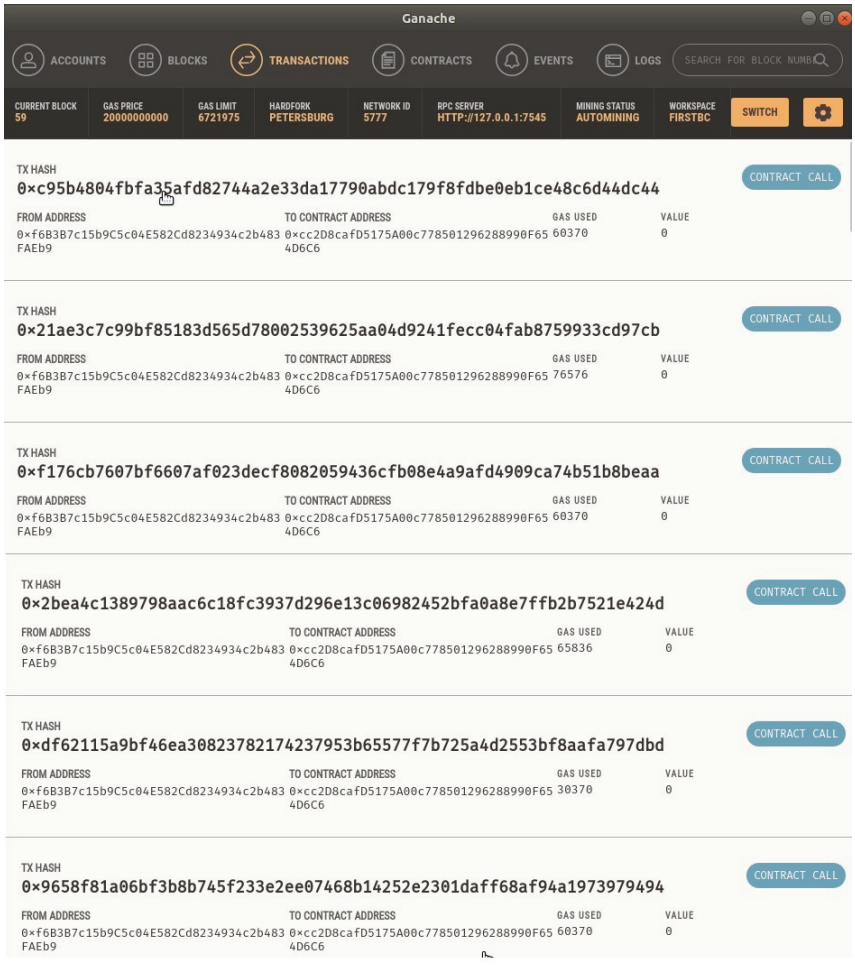


Figure 6. Transactions in the *Ganache* blockchain.

6. Performance Analysis

Our method aims to implement a blockchain-based P2P charging system where the payments will be exchanged using the modern money exchange solution called cryptocurrency. Due to that, the performance analysis section presents the feasibility analysis and the advantages that can be achieved from this proposed method.

6.1. Storage Overhead

In a typical Ethereum blockchain, near 2KB are required per transaction, and a block can accept 512 transactions per block [40]. The average block size for Ethereum is 83.557 KB [48]. Thus, each 512 charge exchange transaction will require almost 84KB of storage.

6.2. Computational Time

Elliptic Curve Cryptography (ECC) is the algorithm used in Ethereum, which is one of the strongest algorithms against cryptanalysis [49]. Another factor that consumes time is the consensus protocol. Ethereum generally uses the Proof-of-Work (PoW) method as consensus. If the consensus is PoW and the cryptography algorithm is ECC, it requires about 4 minutes to generate 40 blocks, and the difficulty is 32.49 KH [50]. The average block

generation times with different difficulties [50] are presented in Figure 7 (*n* is the difficulty of the consensus protocol).

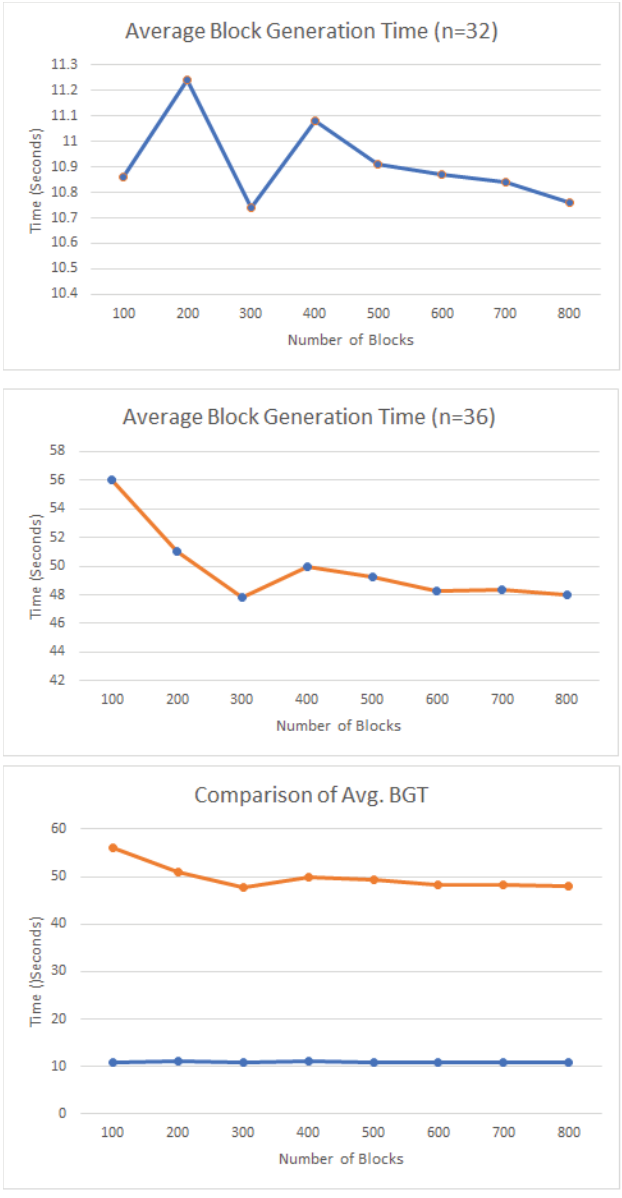


Figure 7. Average block generation time (BGT) when the difficulty of the consensus is 32 or 36, and a comparison.

6.3. Propagation Time

As the proposed method is designed to provide remote support, it requires propagation time to be minimized to maintain the efficiency of the system. However, today's high-speed Internet connections are sufficient to provide necessary services, i.e., reasonable propagation time. For example, by using a 5G internet connection, it is possible to transmit

50 Mbps to 1 Gbps, but the 6G connection will make it at least 100 times faster than that [51]. Thus, the main two components of the system, i.e., EVs and HCSs, are required to maintain a high-speed Internet connection to maintain high throughput.

7. Challenges and Limitations

Blockchain was first proposed in 2008, and it was utilized mainly to develop a currency exchange system for some time [13]. Later in 2014, Ethereum came up with the concept of smart contracts, which influenced researchers to utilize blockchains in different fields [52,53]. In fact, that worked like a catalyst for the creation of innovative applications and areas where they could be used. As a newly developed system, it still requires more experiments (and convincing proofs) to make it compatible with other systems.

The proposed method can also be considered as an effort to combine the P2P energy exchange with blockchain technology. Thus, while developing such a system, several challenges were faced. First of all, it is really difficult to develop a real-world system that can directly communicate with a blockchain. Thus, a simulation study was performed where VMs were considered EVs and HCSs. Secondly, there are very few way to learn and develop smart contracts. Thirdly, rather than popular languages, it supports newly developed languages such as Solidity, Serpent, and Yul, which makes the development phase more difficult. However, some simplified code was written using Solidity (which is an object-oriented, high-level language for implementing smart contracts) to simulate the proposed method. In spite of our efforts, it still requires improvement and optimization. Thirdly, to ensure the trustability of the blockchain, our mechanism uses a consensus protocol which could be highly time consuming (at times), and thus, it is considered a barrier while developing a system that requires high throughput.

While proposing our system, we have considered this above-mentioned issue about time consumption, and accordingly developed the system in such a way that the slowness of the block generation process would not harm the ongoing flow of the system. None of the components and none of the protocols have to wait for block generation, but rather, they can just perform the required actions. After all the communications, tasks, and transactions (starting from *searching()* and ending with *feedback()*) are completed, the server initiates the block generation process. Before the block generation, the charging process and money calculation are performed. Hence, EVs can leave the site just after receiving the charges. As the payment is done after the block generation, the HCSs are required to wait. However, as the system is secure using the blockchain, there is no confusion left regarding receiving the payment and getting extraordinary services, such as security, integrity, transparency, and availability; and HCSs can endure the delay and identify the service taker. We have mentioned the registration system under the blockchain to identify who would be allowed to get the service. While this gives the solution to this issue, a more efficient method can be searched for to enhance the performance of the blockchain. In the future, we plan to also find out more suitable ways to increase the throughput and more optimized outcomes from the currently designed system.

As blockchain-based systems have become popular recently, they were not developed to support all kinds of systems (just yet). Hence, there are several practical issues while adapting blockchain to new systems. Scalability is a critical limitation of the blockchains because of their decentralized and distributed nature. As all the members store the block information, the system requires a huge amount of storage space compared to a typical centralized system. Again, to ensure security, several cryptographic functions are required to be used for encryption, decryption, hashing, etc., which would require high computational power, and that would make the blockchain system expensive. Moreover, systems with low computational power require substantial computational time, which minimizes the throughput of the system. However, to mitigate these problems, several online solutions, such as edge/fog computing, would help the blockchain-based system, as they can virtually provide high computational and storage capacity and can perform complex cryptographic calculations in short periods of time.

8. Conclusions and Future Direction

With the gradual increase in environment-friendly electric vehicles, the availability of charging stations should be ensured. However, it is not an easy task anyway to make them available everywhere, especially in rural areas. Our proposed method largely solves this issue, as it uses an approach of enhanced peer-to-peer charging of EVs, which would increase the availability of charging stations without much change to the existing infrastructure. It will use the same areas but will employ a mechanism to make bonds between the EVs and charging stations anywhere. The rating system can be very useful for keeping the market prices in check and ensuring the quality of the service.

The transactions without a third party are made safe via integrated blockchain to secure the environment for the member nodes of the blockchain for trade. Moreover, a cryptocurrency-based payment system makes the system easy, automated, hassle-free, time-saving, durable, environment-friendly, immutable, and available worldwide. Furthermore, enhanced technological support structures, such as edge computing, and high-speed 5G/6G Internet, can be easily combined with the system to make it more efficient. In the future, we plan to integrate the proposed system with real EVs and HCSs and collect real-world data from those, and make necessary arrangements to enhance the quality of service.

Author Contributions: Conceptualisation, A.F.M.S.A., T.Z.A., E.B.N., S.H. and A.-S.K.P.; Methodology, T.Z.A. and E.B.N.; Software, A.F.M.S.A. and E.B.N.; Validation, A.F.M.S.A., S.H. and A.-S.K.P.; Formal Analysis, A.F.M.S.A., T.Z.A. and E.B.N.; Investigation, A.F.M.S.A. and S.H.; Resources, T.Z.A., E.B.N., S.H. and A.-S.K.P.; writing—original draft preparation, A.F.M.S.A., T.Z.A., E.B.N. and S.H.; writing—review and editing, A.F.M.S.A., T.Z.A. and A.-S.K.P. visualization, A.F.M.S.A., E.B.N. and A.-S.K.P.; supervision, S.H., A.-S.K.P. and A.F.M.S.A.; project administration, A.F.M.S.A., S.H. and A.-S.K.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liang, X.; Du, X.; Wang, G.; Han, Z. A deep reinforcement learning network for traffic light cycle control. *IEEE Trans. Veh. Technol.* **2019**, *68*, 1243–1253. [CrossRef]
2. Xu, Q.; Su, Z.; Zheng, Q.; Luo, M.; Dong, B. Secure content delivery with edge nodes to save caching resources for mobile users in green cities. *IEEE Trans. Ind. Inform.* **2017**, *14*, 2550–2559. [CrossRef]
3. Fu, T.; Liu, P.; Ding, Y.; Zhang, Y. Secure and efficient large content broadcasting in mobile social networks. *IEEE Access* **2018**, *6*, 42108–42118. [CrossRef]
4. Su, Z.; Hui, Y.; Xu, Q.; Yang, T.; Liu, J.; Jia, Y. An edge caching scheme to distribute content in vehicular networks. *IEEE Trans. Veh. Technol.* **2018**, *67*, 5346–5356. [CrossRef]
5. Global Status Report. Renewable Energy Policy Network. 2016. Available online: https://www.iea.org/reports/global-ev-outlook-2020 (accessed on 1 November 2022).
6. How to Choose a Home EV Charger. Available online: https://www.chargepoint.com/resources/how-choose-home-ev-charger/ (accessed on 24 October 2022).
7. Su, Z.; Xu, Q.; Luo, J.; Pu, H.; Peng, Y.; Lu, R. A secure content caching scheme for disaster backup in fog computing enabled mobile social networks. *IEEE Trans. Ind. Inform.* **2018**, *14*, 4579–4589. [CrossRef]
8. Saxena, N.; Grijalva, S.; Chukwuka, V.; Vasilakos, A.V. Network security and privacy challenges in smart vehicle-to-grid. *IEEE Wirel. Commun.* **2017**, *24*, 88–98. [CrossRef]
9. Paul, S.; Ni, Z. Vulnerability analysis for simultaneous attack in smart grid security. In Proceedings of the 2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 23–26 April 2017; pp. 1–5.
10. Wang, J.; Tang, J.; Yang, D.; Wang, E.; Xue, G. Quality-aware and fine-grained incentive mechanisms for mobile crowdsensing. In Proceedings of the 2016 IEEE 36th International Conference on Distributed Computing Systems (ICDCS), Nara, Japan, 27–30 June 2016; pp. 354–363.
11. Wu, Y.; Qian, L.P.; Mao, H.; Yang, X.; Zhou, H.; Shen, X. Optimal power allocation and scheduling for non-orthogonal multiple access relay-assisted networks. *IEEE Trans. Mob. Comput.* **2018**, *17*, 2591–2606. [CrossRef]

12. Novo, O. Blockchain meets IoT: An architecture for scalable access management in IoT. *IEEE Internet Things J.* **2018**, *5*, 1184–1195. [CrossRef]
13. Ahmed, M.; Zubair, S.; Akhter, A.S.; Ullah, A.S.B. Yet another investigation on blockchain in smart healthcare. *Int. J. Agil. Syst. Manag.* **2021**, *14*, 614–634. [CrossRef]
14. Esposito, C.; De Santis, A.; Tortora, G.; Chang, H.; Choo, K.K.R. Blockchain: A panacea for healthcare cloud-based data security and privacy? *IEEE Cloud Comput.* **2018**, *5*, 31–37. [CrossRef]
15. Akhter, A.; Ahmed, M.; Shah, A.; Anwar, A.; Kayes, A.; Zengin, A. A Blockchain-Based Authentication Protocol for Cooperative Vehicular Ad Hoc Network. *Sensors* **2021**, *21*, 1273. [CrossRef] [PubMed]
16. Shahbazi, Z.; Byun, Y.C. Knowledge Discovery on Cryptocurrency Exchange Rate Prediction Using Machine Learning Pipelines. *Sensors* **2022**, *22*, 1740. [CrossRef]
17. Akhter, A.F.M.S.; Ahmed, M.; Anwar, A.; Shah, A.F.M.S.; Pathan, A.S.K.; Zengin, A. Blockchain in vehicular ad hoc networks: Applications, challenges and solutions. *Int. J. Sens. Netw.* **2022**, *40*, 94–130. [CrossRef]
18. Zhang, K.; Mao, Y.; Leng, S.; Maharjan, S.; Zhang, Y.; Vinel, A.; Jonsson, M. Incentive-driven energy trading in the smart grid. *IEEE Access* **2016**, *4*, 1243–1257. [CrossRef]
19. Tushar, W.; Chai, B.; Yuen, C.; Smith, D.B.; Wood, K.L.; Yang, Z.; Poor, H.V. Three-party energy management with distributed energy resources in smart grid. *IEEE Trans. Ind. Electron.* **2014**, *62*, 2487–2498. [CrossRef]
20. Bera, S.; Misra, S.; Chatterjee, D. C2C: Community-based cooperative energy consumption in smart grid. *IEEE Trans. Smart Grid* **2017**, *9*, 4262–4269. [CrossRef]
21. Milanés-Montero, M.I.; Barrero-González, F.; Pando-Acedo, J.; González-Romera, E.; Romero-Cadaval, E.; Moreno-Munoz, A. Smart community electric energy micro-storage systems with active functions. *IEEE Trans. Ind. Appl.* **2018**, *54*, 1975–1982. [CrossRef]
22. Aitzhan, N.Z.; Svetinovic, D. Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Trans. Dependable Secur. Comput.* **2016**, *15*, 840–852. [CrossRef]
23. Li, Z.; Kang, J.; Yu, R.; Ye, D.; Deng, Q.; Zhang, Y. Consortium blockchain for secure energy trading in industrial internet of things. *IEEE Trans. Ind. Inform.* **2017**, *14*, 3690–3700. [CrossRef]
24. Li, L.; Liu, J.; Cheng, L.; Qiu, S.; Wang, W.; Zhang, X.; Zhang, Z. Creditcoin: A privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 2204–2220. [CrossRef]
25. Mattila, J.; Seppälä, T.; Naucler, C.; Stahl, R.; Tikkanen, M.; Bådenlid, A.; Seppälä, J. *Industrial Blockchain Platforms: An Exercise in Use Case Development in the Energy Industry*; Technical Report, ETLA Working Papers; The Research Institute of the Finnish Economy (ETLA): Helsinki, Finland, 2016.
26. Huang, X.; Xu, C.; Wang, P.; Liu, H. LNSC: A security model for electric vehicle and charging pile management based on blockchain ecosystem. *IEEE Access* **2018**, *6*, 13565–13574. [CrossRef]
27. Zou, S.; Ma, Z.; Liu, X.; Hiskens, I. An efficient game for coordinating electric vehicle charging. *IEEE Trans. Autom. Control* **2016**, *62*, 2374–2389. [CrossRef]
28. Mohammadi, J.; Hug, G.; Kar, S. A fully distributed cooperative charging approach for plug-in electric vehicles. *IEEE Trans. Smart Grid* **2016**, *9*, 3507–3518. [CrossRef]
29. Liu, Y.; Hu, S. Renewable energy pricing driven scheduling in distributed smart community systems. *IEEE Trans. Parallel Distrib. Syst.* **2016**, *28*, 1445–1456. [CrossRef]
30. Zhang, B.; Jiang, C.; Yu, J.L.; Han, Z. A contract game for direct energy trading in smart grid. *IEEE Trans. Smart Grid* **2016**, *9*, 2873–2884. [CrossRef]
31. Yang, Y.; Jia, Q.S.; Deconinck, G.; Guan, X.; Qiu, Z.; Hu, Z. Distributed coordination of EV charging with renewable energy in a microgrid of buildings. *IEEE Trans. Smart Grid* **2017**, *9*, 6253–6264. [CrossRef]
32. Wu, X.; Hu, X.; Yin, X.; Moura, S.J. Stochastic optimal energy management of smart home with PEV energy storage. *IEEE Trans. Smart Grid* **2016**, *9*, 2065–2075. [CrossRef]
33. Luo, Y.; Jin, H.; Li, P. A blockchain future for secure clinical data sharing: A position paper. In Proceedings of the ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization, Richardson, TX, USA, 27 March 2019; pp. 23–27.
34. Noor, S.; Yang, W.; Guo, M.; van Dam, K.H.; Wang, X. Energy demand side management within micro-grid networks enhanced by blockchain. *Appl. Energy* **2018**, *228*, 1385–1398. [CrossRef]
35. Ping, J.; Yan, Z.; Chen, S. A two-stage autonomous EV charging coordination method enabled by blockchain. *J. Mod. Power Syst. Clean Energy* **2020**, *9*, 104–113. [CrossRef]
36. Wang, S.; Taha, A.F.; Wang, J.; Kvaternik, K.; Hahn, A. Energy crowdsourcing and peer-to-peer energy trading in blockchain-enabled smart grids. *IEEE Trans. Syst. Man, Cybern. Syst.* **2019**, *49*, 1612–1623. [CrossRef]
37. Esmat, A.; de Vos, M.; Ghiassi-Farrokhfal, Y.; Palensky, P.; Epema, D. A novel decentralized platform for peer-to-peer energy trading market with blockchain technology. *Appl. Energy* **2021**, *282*, 116123. [CrossRef]
38. Chen, S.; Shen, Z.; Zhang, L.; Yan, Z.; Li, C.; Zhang, N.; Wu, J. A trusted energy trading framework by marrying blockchain and optimization. *Adv. Appl. Energy* **2021**, *2*, 100029. [CrossRef]
39. Akhter, A.; Ahmed, M.; Shah, A.; Anwar, A.; Zengin, A. A Secured Privacy-Preserving Multi-Level Blockchain Framework for Cluster Based VANET. *Sustainability* **2021**, *13*, 400. [CrossRef]

40. Storage Needs for Blockchain Technology- Point of View. 2018. Available online: https://www.ibm.com/downloads/cas/LA8XBQGR (accessed on 8 September 2022) .
41. Ahmed, M.; Pathan, A.S.K. Blockchain: Can It Be Trusted? *Computer* **2020**, *53*, 31–35. [CrossRef]
42. The Bitcoin Wiki. Bitcoin Script Examples. Available online: https://en.bitcoin.it/wiki/Script-Script_examples (accessed on 8 September 2022).
43. Truffle Suite. Available online: https://www.trufflesuite.com/ (accessed on 8 April 2020).
44. Ganache. Available online: https://www.trufflesuite.com/ganache (accessed on 8 April 2020).
45. GitHub Lightweight Node Server. Available online: https://github.com/johnpapa/lite-servers (accessed on 8 April 2020).
46. NPM (Software). Available online: https://en.wikipedia.org/wiki/Npm_(software) (accessed on 8 April 2020).
47. Metamask. Available online: https://metamask.io/ (accessed on 8 April 2020).
48. Ethereum Average Block Size. Available online: https://ycharts.com/indicators/ethereum_average_block_size (accessed on 22 September 2022).
49. Ethereum Glossary. Available online: https://ethereum.org/en/glossary/ (accessed on 8 December 2020).
50. Kim, H.; Jang, J.; Park, S.; Lee, H.N. Error-Correction Code Proof-of-Work on Ethereum. *IEEE Access* **2021**, *9*, 135942–135952. [CrossRef]
51. What Is the difference between 4G and 5G? Available online: https://justaskthales.com/en/difference-4g-5g/ (accessed on 8 January 2021).
52. Akhter, A.F.M.S.; Shah, A.F.M.S.; Ahmed, M.; Moustafa, N. A Secured Message Transmission Protocol for Vehicular Ad Hoc Networks. *Comput. Mater. Contin.* **2021**, *68*, 229–246. [CrossRef]
53. Ahmed, M.; Moustafa, N.; Akhter, A.S.; Razzak, I.; Surid, E.; Anwar, A.; Shah, A.S.; Zengin, A. A Blockchain-Based Emergency Message Transmission Protocol for Cooperative VANET. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 19624–19633. [CrossRef]

entropy

MDPI

Article

Observing Cryptocurrencies through Robust Anomaly Scores

Geumil Bae [1] and Jang Ho Kim [2,3,*]

[1] Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Korea
[2] Department of Industrial and Management Systems Engineering, Kyung Hee University, Yongin-si 17104, Korea
[3] Department of Big Data Analytics, Graduate School, Kyung Hee University, Yongin-si 17104, Korea
* Correspondence: janghokim@khu.ac.kr

Abstract: The cryptocurrency market is understood as being more volatile than traditional asset classes. Therefore, modeling the volatility of cryptocurrencies is important for making investment decisions. However, large swings in the market might be normal for cryptocurrencies due to their inherent volatility. Deviations, along with correlations of asset returns, must be considered for measuring the degree of market anomaly. This paper demonstrates the use of robust Mahalanobis distances based on shrinkage estimators and minimum covariance determinant for observing anomaly scores of cryptocurrencies. Our analysis shows that anomaly scores are a critical complement to volatility measures for understanding the cryptocurrency market. The use of anomaly scores is further demonstrated through portfolio optimization and scenario analysis.

Keywords: cryptocurrency; anomaly score; Mahalanobis distance; minimum covariance determinant; shrinkage estimators

Citation: Bae, G.; Kim, J.H. Observing Cryptocurrencies through Robust Anomaly Scores. *Entropy* **2022**, *24*, 1643. https://doi.org/10.3390/e24111643

Academic Editors: Stanisław Drożdż, Marcin Wątorek and Jarosław Kwapień

Received: 27 September 2022
Accepted: 9 November 2022
Published: 11 November 2022

1. Introduction

Even though investments in cryptocurrencies were initially viewed as risky bets, increased participation by individuals as well as institutions have been transforming those views, as the cryptocurrency market is now perceived as a new asset class for many investors. With these revolutions, there have been numerous studies with the overall objective of understanding the cryptocurrency market [1–3] or, more specifically, cryptocurrencies as investment assets. Much effort has been put into analyzing diversification effects and evaluating cryptocurrencies as an asset class [4–8]. Others have focused on diversification across cryptocurrencies [9] and cross-correlation among themselves [10,11]. Predicting price movement of cryptocurrencies using social media data [12], the economic and political uncertainty of the crypto market [13,14], and liquidity of cryptocurrencies [15,16] have also been studied. Due to large swings in cryptocurrencies, analyses also focus on the volatility of the market. Many studies have investigated risk factors of cryptocurrencies [17–20] and examined models for volatility forecasting [21,22], including forecasts of daily value-at-risk [23].

In this article, the risk and volatility of the cryptocurrency market are further examined but from a macro view of observing anomaly scores of market movements. The analysis is based on viewing cryptocurrencies through anomaly scores measured with Mahalanobis distance and its robust variations. The analysis has two significant contributions. First, while cryptocurrencies are generally understood as being more volatile compared to traditional assets, observing anomaly scores provides a standardized framework for identifying unlikely or outlier events, where anomaly calculations incorporate mean, variance, and correlations. Furthermore, anomaly scores are analyzed through cryptocurrency returns as well as risk factors, and robust formulations are proposed to handle extreme outliers in cryptocurrencies. Second, anomaly scores can enhance portfolio management and scenario

analysis of cryptocurrencies. Anomaly scores can act as an indicator of abnormal market conditions, and they can also portray a statistical picture of historical events that provide a medium for measuring historical likelihood as well as estimated likelihood of future scenarios. Performing scenario analyses using risk factors allows a more intuitive and rational interaction with the crypto market. Overall, the analysis provides a practical example of analyzing the crypto market from a more macro perspective that is a valuable complement to volatility analysis of cryptocurrencies.

2. Methodology

Anomaly scores of market movements are measured with Mahalanobis distance (MD), which is a multivariate extension of z-scores, and it is computed by standardizing the deviation from mean with the covariance matrix:

$$\mathrm{MD}(r) = \sqrt{(r-\mu)^T \Sigma^{-1} (r-\mu)}$$

where $\mu \in \mathbb{R}^n$ is the mean vector and $\Sigma \in \mathbb{R}^{n \times n}$ is the covariance matrix of a random vector $r \in \mathbb{R}^n$. While it assumes an elliptical distribution, MD has been shown to be effective in analyzing risks of financial markets [24]. Based on MD, the anomaly score is defined as:

$$\mathrm{A}(r) = \mathrm{MD}(r)/\sqrt{n}$$

to correct for an increase in MD caused by a larger number of variables n for measuring distance [25].

Due to frequent spikes in cryptocurrency movements, MD becomes sensitive to the choice of investment period; mean and covariance used in the calculation of MD are highly sensitive to outliers in price movements. Therefore, in this study, robust MDs were proposed for examining cryptocurrencies. The first proposed robust approach is taken from portfolio optimization where shrinkage estimators are used for computing MD. Mean vector was estimated with the Bayes–Stein estimator [26], and covariance matrix was shrunk using Ledoit and Wolf's [27] approach with a diagonal target. The second robust approach employed in this study is minimum covariance determinant for computing first and second moments without outliers in returns [28]. Even though cryptocurrency returns are not normally distributed [29], the robust MD methods provide a framework for comparing robust anomaly scores. In Section 4, the empirical results compare MD when mean and covariance are estimated from either the entire period (i.e., finding distance relative to the overall movement) or the most recent 104 weeks (i.e., finding distance relative to the market condition during the most recent two years, since March 2000).

More importantly, anomaly scores were initially measured with the price movements of the top cryptocurrencies, and the analysis was repeated with the risk factors of cryptocurrencies. Risk factors are especially important for managing investment portfolios because risk exposure of a portfolio can be effectively measured with underlying factors, whereas financial assets often display high cross-correlations [30,31]. Recently, Liu and Tsyvinski [20] performed a comprehensive empirical asset pricing analysis on cryptocurrencies and found that cryptocurrency returns are exposed to network factors such as number of transactions or number of wallet users, but not production factors such as electricity and computing costs. Thus, network factors were chosen in our analysis to measure anomaly scores of the cryptocurrency market. Principal component analysis is not included in our experiment because principal components that explain much of the variance in cryptocurrencies are highly correlated with the more volatile currencies since 2018, such as Dogecoin. In particular, we found that the top three principal components explained over 70% of variance, where the first principal component is highly correlated with the equally weighted return of the cryptocurrencies and the second principal component is highly correlated with Dogecoin.

3. Data

Two sets of data were used in our analysis: price data of cryptocurrencies and data capturing network effects in cryptocurrencies. Closing prices of cryptocurrencies were retrieved from CoinMarketCap (coinmarketcap.com) [32] and Coin Metrics (coinmetrics.io). Daily price data denominated in USD were collected from 1 January 2018 to 28 February 2022, and converted into weekly returns (weekly returns were used to mitigate any inconsistency in time for computing daily closing price). The analysis focused on the crypto market since 2018 because cryptocurrency funds reveal distinct characteristics in the post-ICO (initial coin offering) bubble period [33] and the market has generally become more mature following the ICO bubble [34]. Crypto markets starting from 2018 can be distinguished as the post-ICO bubble period [33], and ICOs, along with ICO-related events, such as regulatory bans, are observed to cause sensitivity in the market [35].

Empirical tests were performed with seven different sets of weekly returns that ended in different days of the week (i.e., the first set of weekly returns end every Monday, the second set of weekly returns end every Tuesday, and so on). We focused on the top 40 currencies with the largest market capitalization, which account for over 90% of market capitalization valued in USD of the top 500 cryptocurrencies (as of 27 February 2022). Filtering the cryptocurrencies with price data available since the beginning of 2018 results in 15 currencies: Bitcoin (BTC), Ethereum (ETH), Tether (USDT), BNB, XRP, Cardano (ADA), Dogecoin (DOGE), Litecoin (LTC), Chainlink (LINK), Tron (TRX), Bitcoin Cash (BCH), Decentraland (MANA), Stellar (XLM), Ethereum Classic (ETC), and Filecoin (FIL). Since the crypto market contains many more cryptocurrencies, we also used the CCi30 index in our analysis, which is an index of the top 30 cryptocurrencies (the CCi30 index has been used as a representative index of the crypto market for analyzing liquidity [36], herding behavior [37,38], and dynamics of cryptocurrencies [39]).

Based on the analysis in [20], four factors of network effect in cryptocurrencies were collected: number of active addresses (address), number of transactions (transaction), number of transfers (transfer), and number of unique wallets (wallet). The number of wallets was obtained from Blockchain.com for its users, and the other three factors were obtained from data for Bitcoin from Coin Metrics (coinmetrics.io). Weekly growth of these four metrics were computed to match the weekly return periods of cryptocurrencies (the descriptive statistics of the 15 cryptocurrencies and the four network factors are included in the Appendix A, and further details, such as the effect of days of the week, are presented in [20]).

4. Empirical Results on Anomaly Score

4.1. Volatility of Cryptocurrencies

In this section, the volatility of the cryptocurrency market is observed prior to computing anomaly scores. Figure 1 shows the annualized 30-day rolling standard deviation of the top 15 cryptocurrencies and the value-weighted CCi30 index. Figure 2 shows the historical values of the Crypto Volatility index (CVI), which begins in April of 2019, but the x-axis has been scaled to match other figures. Anomaly score outcomes in the following sections were compared with these volatility measures to distinguish the additional value provided by anomaly scores.

4.2. Anomaly Scores of Cryptocurrencies

The first set of anomaly scores were computed directly from the top cryptocurrencies without the use of factors. In order to resolve sensitivity in MD, shrinkage estimators of the mean vector [26] and covariance matrix [27] were used. Shrinkage estimators combine unbiased estimators, such as sample mean, with another component with more structure. A shrinkage estimator for the vector of expected return can be expressed as:

$$\hat{\mu}_{shrink} = (1 - \alpha)\hat{\mu} + \alpha\mu_0 1$$

where $\hat{\mu} \in \mathbb{R}^n$ is the sample mean, $\mu_0 \in \mathbb{R}$ is the shrinkage target, $1 \in \mathbb{R}^n$ is the vector of ones, and α is the shrinkage intensity. Similarly, a shrinkage estimator for the covariance matrix of returns can be written as:

$$\hat{\Sigma}_{shrink} = (1 - \alpha)S + \alpha\Sigma_0$$

where $S \in \mathbb{R}^{n \times n}$ is the sample covariance matrix and $\Sigma_0 \in \mathbb{R}^{n \times n}$ is the target. In our analysis, the shrinkage target μ_0 was set as the expected return of the portfolio with lowest risk (minimum variance portfolio) and Σ_0 was set as a scaled identity matrix. Even though the sample estimates can be sensitive to the estimation period, shrinking them toward a shrinkage target improves robustness [40].

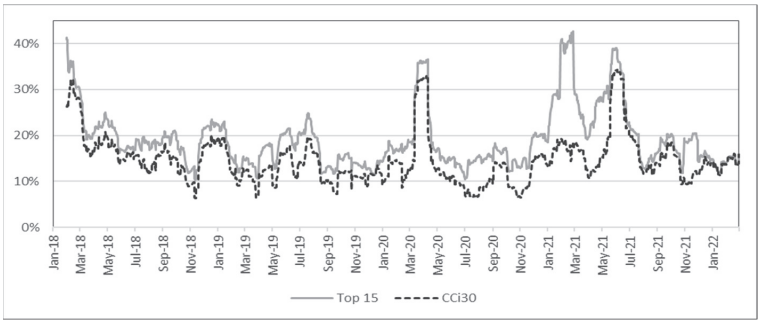


Figure 1. Historical annualized standard deviation (30-day rolling).

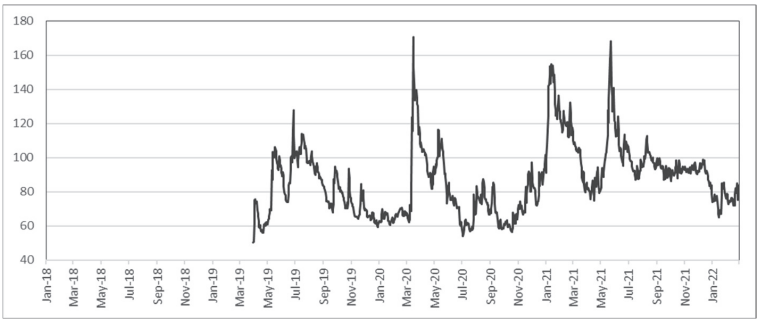


Figure 2. Historical values of Crypto Volatility index.

These shrinkage estimators are frequently applied in portfolio optimization to mitigate sensitivity in the performance of optimal allocations [41,42]. Figure 3 shows anomaly scores when mean and covariance are estimated from the entire period (from January 2018 to February 2022) or from only the last 104 weeks (from March 2000 to February 2020).

Several observations are noteworthy in Figure 3. For each of the seven figures, anomaly scores are not sensitive to the estimation period, and the results are very similar between the scores based on the market condition during January 2018 to February 2022 and the condition during March 2000 to February 2022. This clearly shows the strength of using shrinkage estimators (in contrast, Figure A2 demonstrates the high sensitivity of using non-robust MD for measuring anomaly scores). Moreover, a comparison of the seven graphs in Figure 3 shows that the overall trend and spikes in anomaly scores are fairly robust to the choice of weekly return calculations. For all seven graphs, high anomaly scores are cited between late 2020 and mid-2021, followed by a short spike from around October to November of 2021. Even though there are spikes between late 2018 and mid-2019, the overall anomaly scores are relatively low from the beginning of 2018 until late 2020.

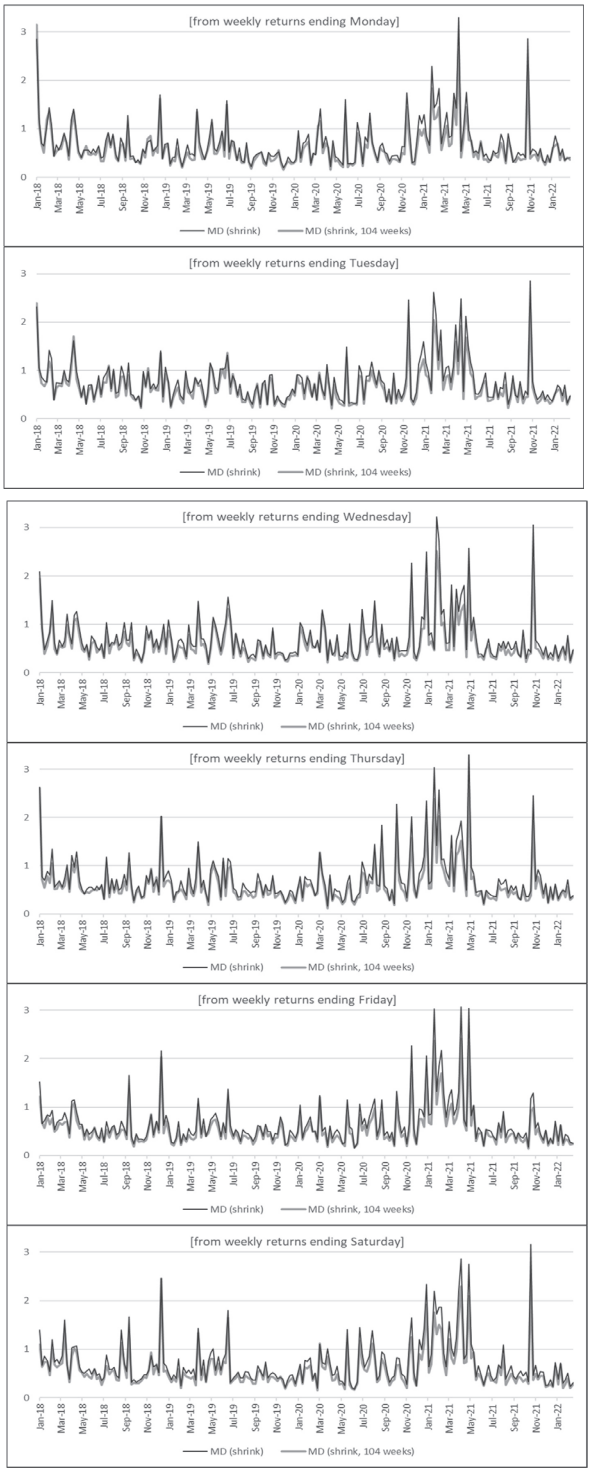


Figure 3. *Cont*.

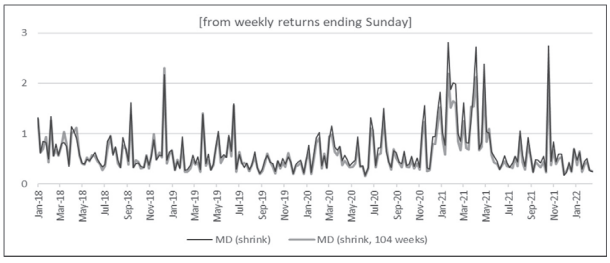


Figure 3. Anomaly scores from top cryptocurrencies (with shrinkage estimators).

When compared with the volatility measures from Section 4.1, anomaly scores show that the high volatility periods during early-to-mid 2021 are also reflected in the anomaly scores. However, more importantly, the market movement during March to May of 2020 was rather *normal*, whereas the condition from October to November of 2021 was *abnormal*. It must be clarified that a *normal* period based on anomaly scores does not necessarily reflect a less volatile period. Since anomaly scores show squared distances from the mean that are standardized by the covariance matrix, a cryptocurrency with high volatility on average will not necessarily have a large anomaly score simply because it deviates much from the mean. This is the key reason why anomaly scores are not a substitute for market volatility but an essential complement for analyzing market movements. For example, high anomaly scores from October to November 2021 were caused by a large spike in Decentraland (MANA), which increased more than five times in less than two months. Further analysis shows that the high anomaly was not only a result of large returns but also due to changes in cross-correlation that were captured by MD. In fact, this was a period when metaverse cryptocurrencies were soaring and anomaly scores were able to capture this new wave in the market.

4.3. Anomaly Scores from Risk Factors

Next, anomaly scores of the cryptocurrency market were further observed using the risk factors of the crypto market. Among several studies on cryptocurrency factors, Liu and Tsyvinski [20] performed comprehensive experiments to show the significance of network factors. While network factors do not provide a complete factor model for explaining the returns and risks of cryptocurrencies, it is worth analyzing with the factors that have been identified so far as being significant.

Weekly growths of four network factors (address, transaction, transfer, and wallet) were used for computing anomaly scores, and minimum covariance determinant (MCD) was chosen for robust MD calculations. The main idea of MCD is to find a sub-sample without outliers and the sub-sample is used for computing the sample mean and covariance [43]. Shrinkage estimators are often applied when the number of variables is large, so MCD was used in our experiment for estimating robust anomaly scores when there were only a few factors [44].

Here, returns were calculated for every week ending Sunday, following [20], and also because Figure 3 shows no substantial disparity among the seven graphs. In Figure 4, the anomaly scores either based on the entire period or only based on the last 104 weeks are almost identical; the robustness of MCD is also evident, similar to the robustness of shrinkage estimators in Figure 3. Additionally, the high volatility from March to May 2020 in Figures 1 and 2 is not noticeable in Figure 4, which matches the anomaly results in Figure 3. While there is a large spike in March 2019 in Figure 4, this is due to a sudden decrease in the numbers of transactions and transfers (see Figure A1). Even though these factors are not able to fully describe cryptocurrency returns or risks, the main purpose of the analysis using risk factors is to demonstrate its use in scenario analysis, as demonstrated in Section 5.2.

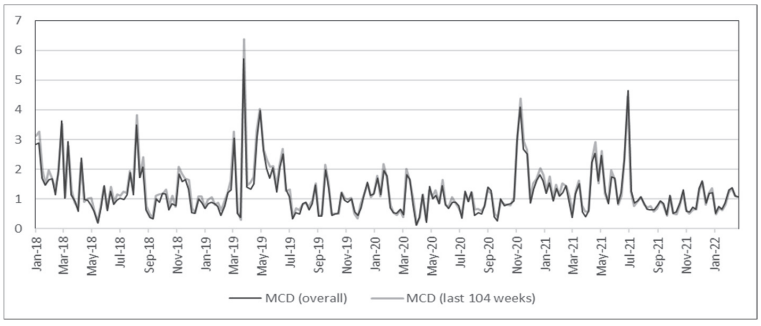
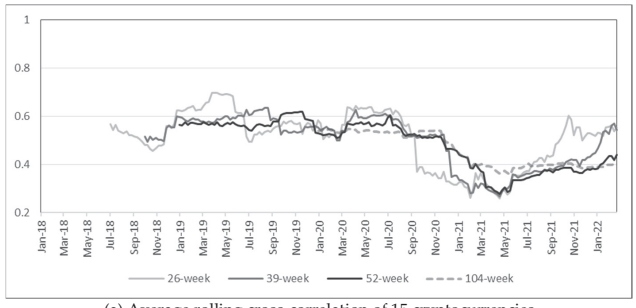


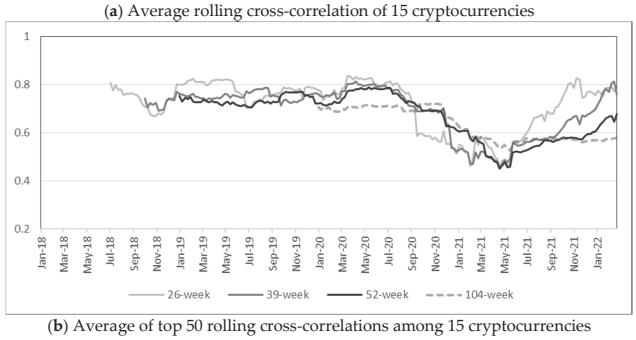
Figure 4. Anomaly scores from network factors (with MCD).

4.4. Further Discussion

One major distinction between measuring risk with volatility and anomaly score is that anomaly scores based on MD accounts for correlation among assets. Figure 5 plots cross-correlations among 15 cryptocurrencies for various rolling windows. The average cross-correlation is greater than 0.4 for most of the period in Panel (a), and a relatively high cross-correlation seems to be the norm due to inherent similarities among cryptocurrencies. In Panel (b), which plots the average among the top 50 cross-correlation values among 15 cryptocurrencies, the average cross-correlation is above 0.6 for most of the period. Nonetheless, there are noticeable drops in early 2021 for all the plots in Figure 5. In other words, cross-correlations among cryptocurrencies are relatively stable until late 2020 but inconsistency is observed in early 2021, which coincides with high anomaly scores. Even though average cross-correlations were more volatile when computed with daily returns as shown in Figure 6, lower cross-correlations in early 2021 are still observed, and it is especially evident from Panel (b) that the highest correlations show a significant drop in early 2021.

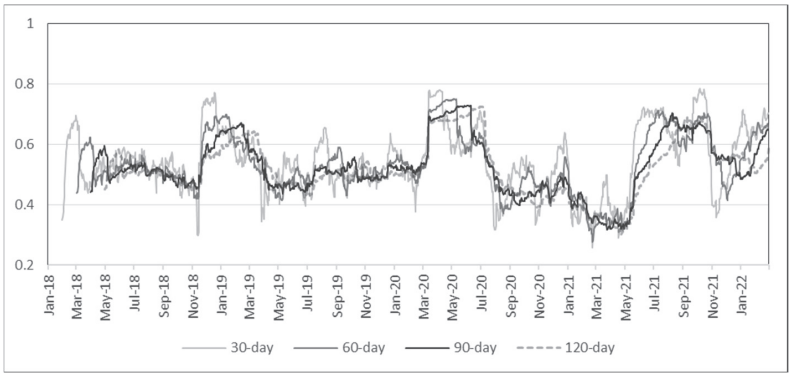


(**a**) Average rolling cross-correlation of 15 cryptocurrencies

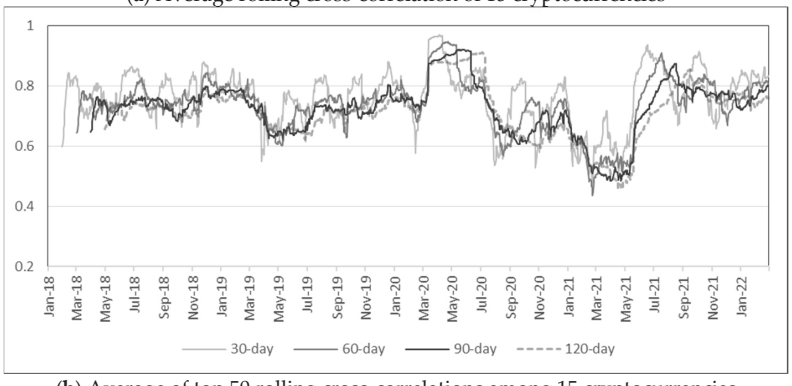


(**b**) Average of top 50 rolling cross-correlations among 15 cryptocurrencies

Figure 5. Rolling cross-correlations of weekly returns (rolling windows = 26, 39, 52, 104 weeks).



(**a**) Average rolling cross-correlation of 15 cryptocurrencies



(**b**) Average of top 50 rolling cross-correlations among 15 cryptocurrencies

Figure 6. Rolling cross-correlations of daily returns (rolling windows = 30, 60, 90, 120 days).

5. Portfolio Analysis Based on Anomaly Scores

As we have demonstrated so far in this study, anomaly scores provide another dimension for analyzing the risks of cryptocurrencies. Even though market anomaly provides valuable insights on its own, it can further enhance portfolio optimization and scenario analysis for investment in cryptocurrencies.

5.1. Incorporating Anomaly Scores into Portfolio Management

We first demonstrated how anomaly scores can be incorporated into portfolio optimization to form portfolios with lower volatility. Even when forming a diversified portfolio among cryptocurrencies, its volatility as measured by standard deviation is too high compared to traditional assets, because each cryptocurrency is volatile on its own and the correlation among cryptocurrencies are relatively high, as already discussed in Figures 5 and 6. However, anomaly scores can help reduce portfolio volatility. Anomaly scores reflect abnormal market movements, so avoiding these periods reduces portfolio volatility even when forming a portfolio that only invests in cryptocurrencies.

For this backtest, rolling optimization was performed with weekly re-optimization and a lookback period of either 52 or 104 weeks. In order to focus on portfolio models with low risk, global minimum-variance (GMV) and risk-parity (equal risk contribution) models were used for optimizing portfolio weights. These are two popular models for forming an investment portfolio based on investment risk rather than expected return.

The GMV portfolio model finds the optimal weights $\omega \in \mathbb{R}^n$ with the smallest risk in the mean-variance optimization framework [40,45] and is written as:

$$\min_{\omega} \frac{1}{2}\omega^T \Sigma\, \omega$$

where $\Sigma \in \mathbb{R}^{n \times n}$ is the covariance matrix of returns for n assets. The risk-parity formulation can be written as:

$$\min_{\omega} \sum_{i=1}^{n}\sum_{j=1}^{n}\left(RC(\omega_i) - RC(\omega_j)\right)^2 \quad \text{where } RC(\omega_i) = \omega_i \frac{\partial \sigma(\omega)}{\partial \omega_i}$$

that minimizes discrepancies among risk contributions (RC) of each asset, where RC is measured with respect to the standard deviation σ of a portfolio [46]. The feasible portfolios were restricted to non-negative weights that sum to one, which is the most basic setting in portfolio construction [47].

On each re-optimization date, the portfolio strategy decided not to invest in cryptocurrencies (i.e., sell all positions) if the anomaly score was above a certain pre-determined limit (e.g., 1 or 2), and ex ante anomaly scores with shrinkage were computed each time from either previous 52-week or 104-week returns. A 52-week lookback period results in portfolio performance from January 2019 to February 2022, and a 104-week lookback provides performance from January 2020 to February 2022. Portfolios were constructed with no-shorting constraints, and USDT was excluded in the backtest because it had negative expected returns during this period.

Table 1 presents weekly standard deviations, annualized standard deviations, and the number of weeks over limit for several anomaly limits. The third column shows results for an equally weighted portfolio of the top 14 cryptocurrencies. The annualized volatility was above 90% without incorporating anomaly scores, but decreased to below 50% with an anomaly limit of 0.5. GMV, and risk-parity portfolios had lower standard deviation compared to the equally weighted portfolio. In particular, GMV had the lowest risk and the annualized volatility was near 40% when an anomaly limit of 0.5 was imposed. Therefore, portfolios with annualized volatility above 80% are unreasonably risky for all rational investors, which is the case without any anomaly limit, but reducing volatility to 40% may provide a viable investment option for investors with minimal risk aversion.

Table 1. Risk performance of portfolios based on various anomaly limits.

Portfolio Model		**Equally Weighted**	**Global Minimum-Variance**		**Risk-Parity**	
Lookback period (weeks)		-	52	104	52	104
Without anomaly limit	std	0.128	0.112	0.114	0.125	0.137
	std (annual.) *	0.923	0.808	0.825	0.900	0.989
	number of weeks over limit	0	0	0	0	0
	total number of weeks	164	164	112	164	112
Anomaly limit = 2.0	std	0.123	0.110	0.099	0.120	0.121
	std (annual.) *	0.889	0.792	0.712	0.869	0.874
	number of weeks over limit	10	10	13	10	13
	total number of weeks	164	164	112	164	112
Anomaly limit = 1.0	std	0.104	0.089	0.088	0.101	0.107
	std (annual.) *	0.749	0.638	0.632	0.732	0.773
	number of weeks over limit	41	41	37	41	37
	total number of weeks	164	164	112	164	112
Anomaly limit = 0.5	std	0.068	0.056	0.064	0.067	0.071
	std (annual.) *	0.492	0.407	0.462	0.481	0.510
	number of weeks over limit	110	110	76	110	76
	total number of weeks	164	164	112	164	112

* Standard deviation is annualized by multiplying $\sqrt{52}$.

5.2. Scenario Analysis of Cryptocurrencies

A major significance of using factors for computing MD is its effectiveness in performing scenario analysis [48]. The two most important components of scenario analysis are the construction of meaningful scenarios and the probability of occurrence for the scenarios. Even though it is difficult to construct scenarios directly at the cryptocurrency level (e.g., it is challenging to form an outlook on short-term returns for a certain currency), it is more intuitive to form a logical outlook on risk factors such as the growth in total transactions or users. Furthermore, since the likelihood of a scenario is proportional to $e^{-MD/2}$, these values can be rescaled to sum to one when estimating the probability of several scenarios [48].

Here, an example is presented to demonstrate how scenarios can be formed with cryptocurrency factors when anomaly scores are computed with robust MD. Table 2 shows mean and standard deviation of weekly growth for the four factors, and the growth in weekly transactions appear to be near zero on average since the beginning of 2018. Suppose scenario analysis is performed based on the view that transactions are going to increase in the coming week; consider growth in transactions to be realized within the set {0.001%, 0.5%, 1.0%, 1.5%, . . . , 10.0%}. Thus, 11 scenarios are generated where transaction takes one of the 11 values, whereas the growth of the other three factors are assumed to stay unchanged (i.e., mean values from Table 2). The advantage of scenario generation from factors is clearly evident in this case. Expressing market outlook through growth in the number of transactions is intuitive even for an investor not familiar with the cryptocurrency market. More rational and detailed views can be expressed with factors.

Next, anomaly scores of these scenarios provide the likelihood (probability) of occurrence for each scenario, and Figure 7 plots the likelihood for the 11 scenarios in this example. The probability of growth in weekly transaction being at least 6% is less than 5%. Thus, even though scenarios are included for cases with large transaction growth, incorporating likelihood through anomaly scores controls the influence on future outcome that are considered outliers. Finally, based on the scenario analysis of traditional assets proposed by [48], the scenarios for the crypto market can be performed as summarized in Figure 8 by applying machine learning models to identify significant factors for efficiently forming rational outlook. These scenarios can be combined with anomaly scores for simulating portfolios invested in cryptocurrencies.

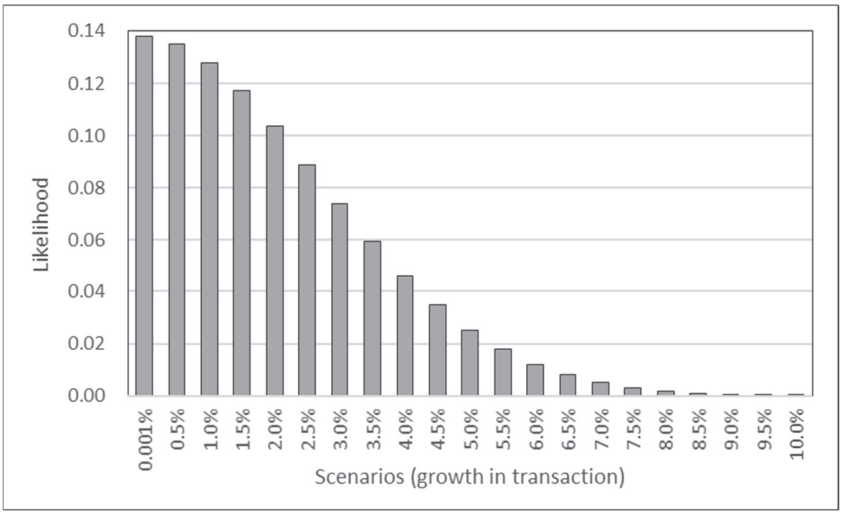


Figure 7. Likelihood of example scenarios on transaction.

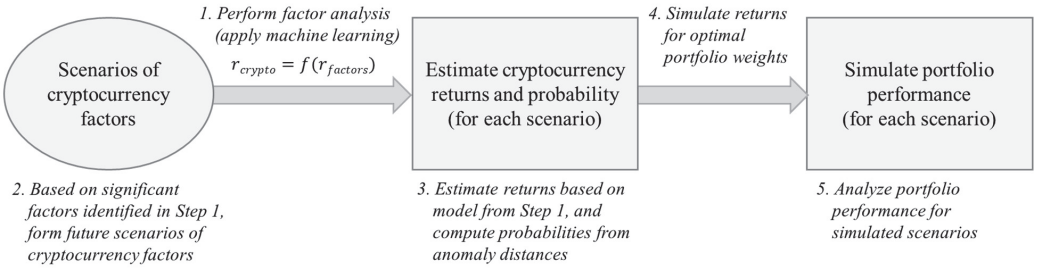


Figure 8. Framework for performing scenario analysis of the crypto market.

Table 2. Statistics of weekly growth (from January 2018 to February 2022).

	Mean	Std
Address	0.00764	0.1173
Transaction	0.00001	0.0896
Transfer	0.00356	0.0806
Wallet	0.00547	0.0040

6. Conclusions

In this article, the use of anomaly scores is illustrated for analyzing the cryptocurrency market. In addition to analyzing the volatility of the cryptocurrency market, anomaly scores of the market provide a complement to the analysis because anomalies are measured by deviation relative to variance and correlation. Specifically, robust Mahalanobis distance based on shrinkage estimators and minimum covariance determinant are shown to produce robust anomaly scores of cryptocurrencies that offer details of market anomalies that are not necessarily explained by standard volatility measures. With the use of anomaly scores as a complement to traditional volatility analyses, investment in cryptocurrencies can be further managed through a detailed understanding of normal or abnormal behavior of cryptocurrencies. Future research can be directed towards analyzing the underlying cause of the discrepancies between traditional volatility measures and the robust anomaly scores proposed in this study. One of the current shortcomings is the limited findings related to risk factors of cryptocurrencies and access to various data. Extended research into risk factors of cryptocurrencies will contribute to computing anomaly scores. Finally, further insight into abnormal behavior in cryptocurrencies will not only provide effective tools for managing investment in the crypto market but also become extremely valuable for investors expanding their assets with cryptocurrencies.

Author Contributions: Conceptualization, G.B. and J.H.K.; methodology, G.B. and J.H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education (MOE, Korea) and the National Research Foundation of Korea (NRF), and partly supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No.RS-2022-00155911, Artificial Intelligence Convergence Innovation Human Resources Development (Kyung Hee University)).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Descriptive statistics of the top 15 cryptocurrencies (weekly returns ending Sunday from 1 January 2018 to 28 February 2022).

	BTC	**ETH**	**USDT**	**BNB**	**XRP**	**ADA**	**DOGE**	**LTC**	**LINK**	**TRX**	**BCH**	**MANA**	**XLM**	**ETC**	**FIL**
mean	0.010	0.016	0.000	0.032	0.011	0.015	0.046	0.007	0.035	0.025	0.007	0.043	0.011	0.017	0.019
std	0.105	0.143	0.006	0.191	0.193	0.175	0.376	0.148	0.216	0.292	0.196	0.299	0.185	0.210	0.211
corr	BTC	ETH	USDT	BNB	XRP	ADA	DOGE	LTC	LINK	TRX	BCH	MANA	XLM	ETC	FIL
BTC	1	0.79	0.08	0.61	0.51	0.64	0.30	0.80	0.59	0.45	0.74	0.41	0.59	0.56	0.44
ETH		1	0.10	0.62	0.53	0.71	0.32	0.78	0.63	0.53	0.76	0.45	0.67	0.67	0.44
USDT			1	−0.02	0.01	0.04	0.01	0.09	0.04	−0.12	0.18	−0.01	−0.04	0.03	0.05
BNB				1	0.45	0.58	0.24	0.60	0.51	0.63	0.50	0.41	0.54	0.51	0.38
XRP					1	0.52	0.38	0.55	0.44	0.41	0.54	0.29	0.69	0.51	0.35
ADA						1	0.34	0.64	0.52	0.46	0.65	0.35	0.72	0.57	0.39
DOGE							1	0.33	0.24	0.28	0.34	0.22	0.29	0.42	0.08
LTC								1	0.59	0.45	0.79	0.38	0.56	0.74	0.42
LINK									1	0.56	0.51	0.39	0.56	0.48	0.33
TRX										1	0.41	0.40	0.58	0.44	0.32
BCH											1	0.35	0.59	0.73	0.40
MANA												1	0.39	0.32	0.24
XLM													1	0.56	0.38
ETC														1	0.33
FIL															1

Table A2. Descriptive statistics of the network factors (weekly growth ending Sunday from 1 January 2018 to 28 February 2022).

	Address	**Transaction**	**Transfer**	**Wallet**
mean	0.006	0.004	0.004	0.006
std	0.121	0.109	0.092	0.004

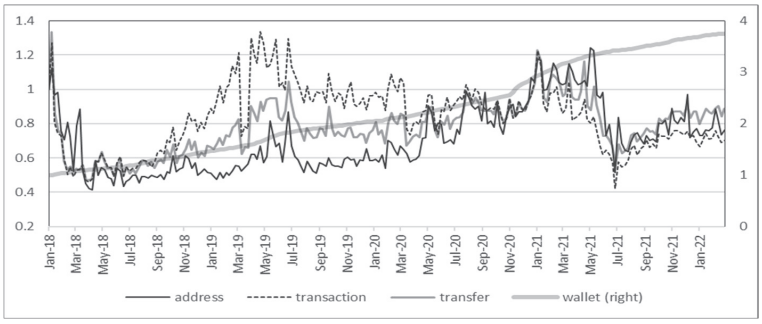


Figure A1. Weekly growth of the network factors (initial value: 1).

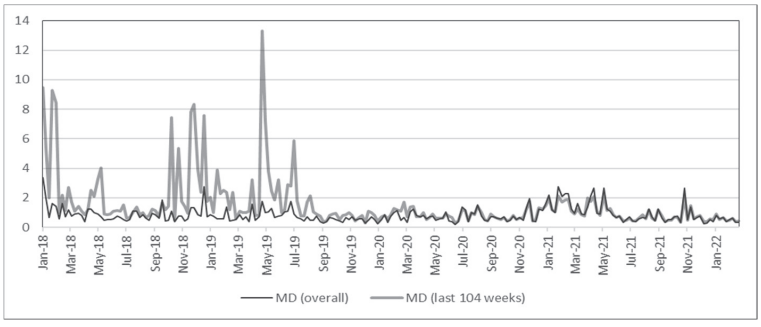


Figure A2. Anomaly scores from top cryptocurrencies (MD).

References

1. Fang, F.; Ventre, C.; Basios, M.; Kanthan, L.; Martinez-Rego, D.; Wu, F.; Li, L. Cryptocurrency trading: A comprehensive survey. *Financ. Innov.* **2022**, *8*, 1–59. [CrossRef]
2. Kim, K.; Lee, M. The impact of the COVID-19 pandemic on the unpredictable dynamics of the cryptocurrency market. *Entropy* **2021**, *23*, 1234. [PubMed]
3. Wątorek, M.; Drożdż, S.; Kwapień, J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* **2021**, *901*, 1–82. [CrossRef]
4. Chuen, D.L.K.; Guo, L.; Wang, Y. Cryptocurrency: A new investment opportunity? *J. Altern. Investig.* **2017**, *20*, 16–40. [CrossRef]
5. Bouri, E.; Molnár, P.; Azzi, G.; Roubaud, D.; Hagfors, L.I. On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Financ. Res. Lett.* **2017**, *20*, 192–198. [CrossRef]
6. Bianchi, D. Cryptocurrencies as an asset class? An empirical assessment. *J. Altern. Investig.* **2020**, *23*, 162–179. [CrossRef]
7. Sifat, I. On cryptocurrencies as an independent asset class: Long-horizon and COVID-19 pandemic era decoupling from global sentiments. *Financ. Res. Lett.* **2021**, *43*, 102013. [CrossRef]
8. Kim, J.H. Analyzing diversification benefits of cryptocurrencies through backfill simulation. *Financ. Res. Lett.* **2022**, *50*, 103238. [CrossRef]
9. Liu, W. Portfolio diversification across cryptocurrencies. *Financ. Res. Lett.* **2019**, *29*, 200–205. [CrossRef]
10. Kwapień, J.; Wątorek, M.; Drożdż, S. Cryptocurrency market consolidation in 2020–2021. *Entropy* **2021**, *23*, 1674. [CrossRef]
11. Nguyen, A.P.N.; Mai, T.T.; Bezbradica, M.; Crane, M. The cryptocurrency market in transition before and after COVID-19: An opportunity for investors? *Entropy* **2022**, *24*, 1317. [CrossRef]
12. Valencia, F.; Gómez-Espinosa, A.; Valdés-Aguirre, B. Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *Entropy* **2019**, *21*, 589. [CrossRef]
13. Demir, E.; Gozgor, G.; Lau, C.K.M.; Vigne, S.A. Does economic policy uncertainty predict the Bitcoin returns? An empirical investigation. *Financ. Res. Lett.* **2018**, *26*, 145–149. [CrossRef]
14. Colon, F.; Kim, C.; Kim, H.; Kim, W. The effect of political and economic uncertainty on the cryptocurrency market. *Financ. Res. Lett.* **2021**, *39*, 101621. [CrossRef]
15. Leirvik, T. Cryptocurrency returns and the volatility of liquidity. *Financ. Res. Lett.* **2021**, *44*, 102031. [CrossRef]
16. Brauneis, A.; Mestel, R.; Theissen, E. What drives the liquidity of cryptocurrencies? A long-term analysis. *Financ. Res. Lett.* **2021**, *39*, 101537. [CrossRef]
17. Gregoriou, A. Cryptocurrencies and asset pricing. *Appl. Econ. Lett.* **2019**, *26*, 995–998. [CrossRef]
18. Li, J.; Yi, G. Toward a factor structure in crypto asset returns. *J. Altern. Investig.* **2019**, *21*, 56–66. [CrossRef]
19. Zhang, W.; Li, Y.; Xiong, X.; Wang, P. Downside risk and the cross-section of cryptocurrency returns. *J. Bank. Financ.* **2021**, *133*, 106246. [CrossRef]
20. Liu, Y.; Tsyvinski, A. Risks and returns of cryptocurrency. *Rev. Financ. Stud.* **2021**, *34*, 2689–2727. [CrossRef]
21. Bouri, E.; Lau, C.K.M.; Lucey, B.; Roubaud, D. Trading volume and the predictability of return and volatility in the cryptocurrency market. *Financ. Res. Lett.* **2019**, *29*, 340–346. [CrossRef]
22. Ftiti, Z.; Louhichi, W.; Ben Ameur, H. Cryptocurrency volatility forecasting: What can we learn from the first wave of the COVID-19 outbreak? *Ann. Oper. Res.* **2021**, 1–26. [CrossRef] [PubMed]
23. Pele, D.T.; Mazurencu-Marinescu-Pele, M. Using high-frequency entropy to forecast bitcoin's daily value at risk. *Entropy* **2019**, *21*, 102. [CrossRef] [PubMed]
24. Kritzman, M.; Li, Y. Skulls, financial turbulence, and risk management. *Financ. Anal. J.* **2010**, *66*, 30–41. [CrossRef]
25. Golub, B.; Greenberg, D.; Ratcliffe, R. Market-driven scenarios: An approach for plausible scenario construction. *J. Portf. Manag.* **2018**, *44*, 6–20. [CrossRef]
26. Jorion, P. Bayes-Stein estimation for portfolio analysis. *J. Financ. Quant. Anal.* **1986**, *21*, 279–292. [CrossRef]
27. Ledoit, O.; Wolf, M. A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.* **2004**, *88*, 365–411. [CrossRef]
28. Leys, C.; Klein, O.; Dominicy, Y.; Ley, C. Detecting multivariate outliers: Use a robust variant of Mahalanobis distance. *J. Exp. Soc. Psychol.* **2018**, *74*, 150–156. [CrossRef]
29. Wątorek, M.; Kwapień, J.; Drożdż, S. Financial return distributions: Past, present, and COVID-19. *Entropy* **2021**, *23*, 884. [CrossRef]
30. Ang, A. *Asset Management: A Systematic Approach to Factor Investing*; Oxford University Press: Oxford, UK, 2014.
31. Madhavan, A.; Sobczyk, A.; Ang, A. What's in your benchmark? A factor analysis of major market indexes. *J. Portf. Manag.* **2018**, *44*, 46–59. [CrossRef]
32. Vidal-Tomás, D. Which cryptocurrency data sources should scholars use? *Int. Rev. Financ. Anal.* **2022**, *81*, 102061. [CrossRef]
33. Bianchi, D.; Babiak, M. On the performance of cryptocurrency funds. *J. Bank. Financ.* **2022**, *138*, 106467. [CrossRef]
34. Vidal-Tomás, D. The entry and exit dynamics of the cryptocurrency market. *Res. Int. Bus. Financ.* **2021**, *58*, 101504. [CrossRef]
35. Momtaz, P.P. Initial coin offerings. *PLoS ONE* **2020**, *15*, e0233018. [CrossRef]
36. Manahov, V. Cryptocurrency liquidity during extreme price movements: Is there a problem with virtual money? *Quant. Financ.* **2021**, *21*, 341–360. [CrossRef]
37. Ajaz, T.; Kumar, A.S. Herding in crypto-currency markets. *Ann. Financ. Econ.* **2018**, *13*, 1850006. [CrossRef]
38. Zhao, Y.; Liu, N.; Li, W. Industry herding in crypto assets. *Int. Rev. Financ. Anal.* **2022**, *84*, 102335. [CrossRef]

39. Vidal-Tomás, D. The new crypto niche: NFTs, play-to-earn, and metaverse tokens. *Financ. Res. Lett.* **2022**, *47*, 102742. [CrossRef]
40. Kim, W.C.; Kim, J.H.; Fabozzi, F.J. *Robust Equity Portfolio Management + Website: Formulations, Implementations, and Properties Using MATLAB*; Wiley: Hoboken, NJ, USA, 2016.
41. Disatnik, D.J.; Benninga, S. Shrinking the covariance matrix. *J. Portf. Manag.* **2007**, *33*, 55–63. [CrossRef]
42. Novais, R.G.; Wanke, P.; Antunes, J.; Tan, Y. Portfolio optimization with a mean-entropy-mutual information model. *Entropy* **2022**, *24*, 369. [CrossRef]
43. Rousseeuw, P.J. Least median of squares regression. *J. Am. Stat. Assoc.* **1984**, *79*, 871–880. [CrossRef]
44. Hubert, M.; Debruyne, M. Minimum covariance determinant. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 36–43. [CrossRef]
45. Markowitz, H.M. Portfolio selection. *J. Financ.* **1952**, *7*, 77–91.
46. Qian, E. Risk parity and diversification. *J. Investig.* **2011**, *20*, 119–127.
47. Kim, J.H.; Lee, Y.; Kim, W.C.; Fabozzi, F.J. Mean–Variance Optimization for Asset Allocation. *J. Portf. Manag.* **2021**, *47*, 24–40. [CrossRef]
48. Czasonis, M.; Kritzman, M.; Pamir, B.; Turkington, D. Enhanced scenario analysis. *J. Portf. Manag.* **2020**, *46*, 69–79. [CrossRef]

entropy

MDPI

Article

Dynamic Linkage between Bitcoin and Traditional Financial Assets: A Comparative Analysis of Different Time Frequencies

Panpan Wang [1,*], Xiaoxing Liu [1] and Sixu Wu [2]

[1] School of Economics and Management, Southeast University, Nanjing 211189, China
[2] School of Urban and Regional Science, East China Normal University, Shanghai 200241, China
* Correspondence: tjwangpanpan@163.com

Abstract: This study employs the ADCC-GARCH approach to investigate the dynamic correlation between bitcoin and 14 major financial assets in different time-frequency dimensions over the period 2013–2021, for which the risk diversification, hedging and safe-haven properties of bitcoin for those traditional assets are further examined. The results show that, first, bitcoin is positively linked to risk assets, including stock, bond and commodity, and negatively linked to the U.S. dollar, which is a safe-haven asset, so bitcoin is closer in nature to a risk asset than a safe-haven asset. Second, the high short-term volatility and speculative nature of the bitcoin market makes its long-term correlation with other assets stronger than the short-term. Third, the positive linkage between the prices of bitcoin and risk assets increases sharply under extreme shocks (e.g., the outbreak of COVID-19 in early 2020). Fourth, bitcoin can hedge against the U.S. dollar, and in the long term, bitcoin can hedge against the Chinese stock market and act as a safe haven for the U.S. stock market and crude oil. However, for most other traditional assets, bitcoin is only an effective diversifier.

Keywords: bitcoin; ADCC-GARCH; diversifier; hedge; safe haven

Citation: Wang, P.; Liu, X.; Wu, S. Dynamic Linkage between Bitcoin and Traditional Financial Assets: A Comparative Analysis of Different Time Frequencies. *Entropy* **2022**, *24*, 1565. https://doi.org/10.3390/e24111565

Academic Editors: Stanisław Drożdż, Jarosław Kwapień, Marcin Wątorek and Pierpaolo Vivo

Received: 14 September 2022
Accepted: 28 October 2022
Published: 30 October 2022

1. Introduction

Digital cryptocurrencies have rapidly entered the public view in recent years, and their market trading scale continues to expand. As a representative species in the cryptocurrency market, bitcoin has exhibited dramatic volatility since its inception, and its price fluctuations have long been a concern for both academia and practitioners. Due to the soaring price of bitcoin in recent years, more investors around the world are entering the bitcoin market, expecting to make profits while lacking a deep understanding of the price formation mechanism of bitcoin and its asset properties, thus, facing huge investment risks. To establish an analytical framework about the price formation mechanism of bitcoin, it is first necessary to define whether bitcoin is a risk or a safe-haven asset. Some argue that because bitcoin is completely decentralized and not controlled by a traditional central bank and because bitcoin supply is limited by its own protocol design to a fixed total of 21 million coins, bitcoin has a similar anti-inflation value to gold and is a safe-haven asset. However, there are also arguments that the bitcoin market is highly speculative and that there is a clear positive correlation between the prices of bitcoin and various risk assets, thus, making it more of a risk asset in nature.

Is bitcoin a risk asset or a safe-haven asset? What are the linkages between bitcoin and major global assets? What are the dynamics of these linkages over time? This paper aims to explore the asset properties of bitcoin from the perspective of its linkage with traditional financial assets. We use the asymmetric dynamic conditional correlation (ADCC)-GARCH approach to examine the dynamic correlations between bitcoin and various traditional assets in different time frequency dimensions and further explore bitcoin's diversification, hedging and safe-haven properties for each asset based on the dynamic correlations between bitcoin and these assets. Our analysis not only helps to further clarify the price

formation mechanism of bitcoin and its role in portfolio management and helps investors to reasonably hold digital cryptocurrencies for investment, but also helps policymakers improve the dynamic monitoring and risk management of the cryptocurrency market represented by bitcoin.

The remainder of the paper is structured as follows. Section 2 reviews the relevant literature. Section 3 uses the ADCC-GARCH approach to quantitatively measure the dynamic correlation between bitcoin and various traditional financial assets in different time-frequency dimensions. Section 4 further identifies bitcoin's risk diversification, hedging and safe-haven capabilities for each traditional asset. Section 5 concludes.

2. Literature Review

Bitcoin is a digital currency and payment system created by Satoshi Nakamoto [1]. As the first decentralized digital cryptocurrency, bitcoin's price and popularity have risen rapidly since its introduction in 2009. With the growing popularity of bitcoin worldwide, the economic and financial properties of bitcoin have begun to attract the attention of researchers [2–4]. The relevant literature focuses on whether bitcoin is a currency or an asset, what kind of asset bitcoin is, and what kind of risk–return properties bitcoin has.

The earlier literature focused on whether bitcoin was a currency or an asset. Undeniably, there are some commonalities between bitcoin and currency, but from the perspective of monetary function, bitcoin can only be used as a medium of exchange and not as a unit of account or a storage of value [5]; therefore, bitcoin does not have a complete form of currency. Glaser et al. [6] focused on the liquidity of bitcoin when it functions as a medium of exchange, and argued that the convertibility between bitcoin and traditional currencies gives bitcoin liquidity, but the limited supply of bitcoin limits its liquidity. Böhme et al. [7] argued that the liquidity of bitcoin can be significantly weakened due to the frequent delays in bitcoin transactions. However, because of the anonymity of user identities, the bitcoin protocol does not restrict international transfer operations to countries that are on watch lists or embargoed, which provides bitcoin with higher flexibility and liquidity than deposit currencies in supporting international transfers [7]. In terms of the attitudes of bitcoin holders, since most bitcoin users view bitcoin as an investment tool rather than a transactional payment tool [6], the market value of bitcoin is much higher than the size of the economic transactions it facilitates [5], making bitcoin more of a speculative investment tool than a currency. Luther and Salter [8] examined bitcoin's ability to replace traditional currencies based on the increase in bitcoin app downloads after the Cyprus bailout announcement and found that the rise in bitcoin app downloads was insignificant, suggesting that bitcoin is not replacing the currencies of those countries whose domestic banks are in trouble.

After determining that bitcoin is more of an asset, scholars began comparing bitcoin to traditional assets in an attempt to generalize which asset, or class of assets, bitcoin is more comparable. Bitcoin is often analogized to gold in the literature, and is even referred to as digital gold or the new gold [9]. The similarities between bitcoin and gold are that both have a much higher market value than their intrinsic value, and both derive their market value from scarcity of supply and high mining costs; both have no national attributes, and their supply is not controlled by the government; gold was used as a medium of exchange during the gold standard period but was abandoned later due to lack of liquidity, and bitcoin is likely to face similar problems in the future as the scale of bitcoin users expands. However, there are also differences between bitcoin and gold; for example, people use gold mainly because of its function as a store of value, while the instability of bitcoin prices makes it difficult for it to perform value storage effectively [10]. Klein et al. [11] compared the return volatility of bitcoin and gold and their respective correlations with other asset prices and found that while both prices respond asymmetrically to market shocks, their respective correlations with other asset prices differ significantly, especially during market downturns. Shahzad et al. [12] compared the safe-haven, hedging and diversification properties of bitcoin and gold for the G7 stock markets and found that gold outperforms bitcoin in terms of safe-haven and hedging effectiveness

and can provide higher conditional diversification benefits for stock investments than bitcoin, while Thampanya et al. [13] found that neither bitcoin nor gold is a good hedge for the Thailand stock market. Furthermore, Whelan [14] drew an analogy between bitcoin and the U.S. dollar, arguing that both are used as a medium of exchange, but the difference is that the dollar is backed by the government, whereas bitcoin is a private currency issued by the private sector, and, thus, the supply and governance mechanisms for both assets are different. Dyhrberg [10] examined whether bitcoin is more comparative to gold or the U.S. dollar, and found that the behavioral characteristics of bitcoin prices have both dollar- and gold-like components because bitcoin's property falls between a pure medium of exchange and a pure store of value. As such, bitcoin can be classified as an asset that is between the U.S. dollar and gold, and can be used as an effective tool for portfolio management.

More recent studies have begun to focus on the risk–return properties of bitcoin and its risk diversification, hedging and safe haven properties for traditional assets. Although bitcoin has a higher volatility than traditional assets [15], the inclusion of bitcoin in a portfolio may still improve the portfolio's risk–return performance [16–19]. Eisl et al. [20] used a CVaR approach to find that although the inclusion of bitcoin raises the conditional VaR of the portfolio, this additional risk is fully compensated by high returns, which ultimately improves the risk–return ratio. Dyhrberg [21] used daily frequency data to test the hedging effect of bitcoin on U.K. equities and the USD exchange rate and found that bitcoin can be used as a hedge for the FTSE index as well as the USD/EUR and USD/GBP exchange rates over the 2010–2015 period. Yang et al. [22] used a time-frequency domain approach to find that bitcoin has the capability to hedge against the currency market in the long term. Using a daily frequency sample over 2011–2015, Bouri et al. [23] found that bitcoin prices are negatively correlated with the Nikkei 225, MSCI Pacific and commodity indices and, therefore, have the ability to hedge against price fluctuations in these assets. Chan et al. [24] found that bitcoin can be used as an effective hedge for the U.S., European, Canadian, Japanese and Chinese stock markets in the monthly frequency dimension over the period 2010–2017. Wang et al. [25] found that cryptocurrencies are a safe haven rather than a hedge for most international stock indices, and the safe-haven properties are more pronounced in developed markets or markets with larger market capitalizations and higher liquidity. Shahzad et al. [26] showed that the safe-haven role of bitcoin for Chinese, U.S. and other developed and developing stock markets is time-varying and varies across different stock markets. Urquhart and Zhang [27] used high-frequency data to examine bitcoin's hedging and safe-haven capabilities for foreign exchange and found that bitcoin can be a hedge for the CHF, EUR and GBP, a diversifier for the AUD, CAD and JPY, and a safe haven for the CAD, CHF and GBP in times of extreme market turmoil. Wang et al. [28] examined the mean and volatility spillovers between bitcoin and six major financial assets in China and found that bitcoin can hedge China's stock, bond, and monetary markets and can serve as a safe haven for China's monetary market. Smales [29] argued that bitcoin's high volatility, low liquidity and high transaction costs (in terms of time and fees) compared to other assets preclude it from being considered a safe haven until its market matures. Kwapień et al. [3] analyzed detrended cross-correlations between cryptocurrency markets and some traditional markets (including stock, commodity and forex markets) and found that the levels of cross-correlations become higher in turbulent periods.

After the outbreak of the COVID-19 pandemic, the price dynamics and portfolio performance of cryptocurrencies during the pandemic attracted widespread attention [30]. Wątorek et al. [4] pointed out that the COVID-19 pandemic has had a significant impact on the cryptocurrency market, transforming cryptocurrencies from a hedging instrument for investors fleeing traditional markets into a part of the global market which is closely linked to traditional financial instruments including currency, stock and commodity. Using a network connectedness model, Balcilar et al. [31] found increasing risk spillovers between cryptocurrencies and global emerging stock markets following the COVID-19 pandemic outbreak, and that cryptocurrencies cannot serve as a diversifier for emerging stock markets in both the short and long term. Caferra and Vidal-Tomas [32] used the wavelet coherence

approach and Markov switching autoregressive model to find that cryptocurrencies have some hedging properties against stock markets in response to shocks from the COVID-19 pandemic. Using the COVID-19 outbreak as a quasi-experiment, Grobys [33] used a difference-in-differences approach to test bitcoin's performance in hedging U.S. stock market risk and found that bitcoin performs poorly in hedging the tail risk of the U.S. market. Conlon and McGee [2] evaluated the safe-haven capability of bitcoin against traditional assets during the sharp decline in global financial markets following the outbreak of COVID-19, and found that bitcoin is not a safe haven for the S&P 500 and that including bitcoin in an equity portfolio at this time would substantially increase the portfolio's downward risk exposure. Conlon et al. [34] further found that bitcoin is not a safe haven for most international stock markets except for China's CSI 300 index. Wen et al. [35] used time-varying parameter vector autoregression to find that bitcoin is not a safe haven for crude oil and stocks during the COVID-19 pandemic. Dutta et al. [36] also found that bitcoin is only a diversifier rather than a safe haven for crude oil during the COVID-19 pandemic.

Although there has been a rich literature on the economic and financial properties of bitcoin, several gaps still remain. First, the sample of relevant studies does not cover a wide enough range of asset classes, and, therefore, they fail to provide a comprehensive dissection of the linkage between bitcoin and various major global assets and the diversification, hedging and safe-haven properties of bitcoin for these assets. Most of the existing studies focus on analyzing the linkage between bitcoin and a specific class of financial assets and bitcoin's diversification, hedging and safe-haven properties for that class of assets, including stocks [24–26,31–34], commodities [36], currencies [22,27], etc. Although the sample selected by Wang et al. [28] covers multiple asset classes, these assets are all Chinese assets and are not globally representative. Second, most of the existing studies have examined the risk–return characteristics of bitcoin and its correlation with other assets in a frequency-specific sample, while few have conducted comparative analyses at different time frequencies. The high short-term volatility and speculative nature of the bitcoin market is likely to impair the short-term correlation between bitcoin and other assets, resulting in the linkage between bitcoin and other assets exhibiting very different characteristics in different time frequency dimensions, and may even lead to changes in bitcoin's risk diversification, hedging and safe-haven properties in different time frequency dimensions.

We selected the prices of bitcoin and 14 major financial assets covering stock, bond, commodity and currency over the period 2013–2021 as a sample to test the dynamic linkage of bitcoin with each asset and the linkage's variation in different time frequency dimensions, and to further identify the risk diversification, hedging and safe-haven properties of bitcoin for various assets. We first used the ADCC-GARCH approach to quantitatively measure the dynamic correlation between bitcoin and other traditional assets. Based on the estimated dynamic conditional correlation (DCC) series, we further adopted Ratner and Chiu's [37] approach to identify the risk diversification, hedging and safe-haven properties of bitcoin for various types of assets to assess the extent to which bitcoin can be used as a diversifier, hedge or safe haven for those assets. As a result, we show that: (i) bitcoin is positively linked to risk assets, including stocks, bonds and commodities, and negatively linked to the U.S. dollar, which is a typical safe-haven asset, so bitcoin is closer in nature to a risk asset than a safe-haven asset; (ii) the high short-term volatility and speculative nature of the bitcoin market makes its long-term correlation with other asset prices stronger than the short-term correlation; (iii) the positive linkage between the prices of bitcoin and risk assets increases sharply under extreme shocks (e.g., the outbreak of COVID-19 in early 2020); (iv) bitcoin can hedge against the U.S. dollar, and in the long term, bitcoin can hedge against the Chinese stock market and act as a safe haven for the U.S. stock market and crude oil. However, for most other traditional assets, bitcoin is only an effective diversifier.

The contribution of this study is twofold. First, our selected sample covers four major asset classes—stock, bond, commodity and currency—which specifically include 14 major representative global financial assets. Based on this sample, we provide a comprehensive analysis of the linkage between bitcoin and major global assets as well as bitcoin's diver-

sification, hedging and safe haven properties for these assets. Second, we also perform a comparative analysis in different time-frequency dimensions, comparing the dynamic correlation between bitcoin and each asset in daily, weekly and semi-monthly frequency dimensions as well as the differences in bitcoin's risk diversification, hedging and safe-haven capabilities for other assets in different time-frequency dimensions.

3. The Dynamic Correlation between Bitcoin and Traditional Assets

3.1. ADCC-GARCH Model

The aim of this study is to examine the dynamic linkages between bitcoin and various traditional financial assets at different frequencies and to explore the risk diversification, hedging and safe-haven properties of bitcoin for each asset based on the dynamic linkages between bitcoin and them. To capture the time-varying correlation between bitcoin and other traditional financial assets, we employed the DCC-GARCH approach. The DCC-GARCH method can estimate the time-varying conditional correlation coefficient and has the advantage of portraying the dynamic relationship between variables. The DCC model was first introduced by Engle [38] to allow for time-varying correlation between variables. Cappiello et al. [39] further introduced an asymmetric version of the DCC-GARCH (i.e., ADCC-GARCH) to address the effect of asymmetric information on time-varying correlations. In this study, the ADCC model of Cappiello et al. [39] was used to model the volatility dynamics and conditional correlation between bitcoin and other assets.

Let r_t be a $n \times 1$ vector of asset returns. The AR(1) process for r_t conditioned on the information set I_{t-1} can be written as follows:

$$r_t = \mu + \varphi r_{t-1} + \varepsilon_t \tag{1}$$

The residuals are modeled as:

$$\varepsilon_t = H_t^{1/2} z_t \tag{2}$$

H_t is the conditional covariance matrix of r_t, and z_t is a $n \times 1$ i.i.d. random vector of errors. Engle's [38] DCC model is estimated in two steps, with the GARCH parameters estimated in the first step and the conditional correlation in the second step, where:

$$H_t = D_t R_t D_t \tag{3}$$

where H_t is a $n \times n$ conditional covariance matrix, R_t is the conditional correlation matrix, and D_t is the diagonal matrix with time-varying standard deviations on the diagonal.

$$D_t = diag\left(h_{1,t}^{1/2}, \ldots, h_{n,t}^{1/2}\right) \tag{4}$$

$$R_t = diag\left(q_{1,t}^{-1/2}, \ldots, q_{n,t}^{-1/2}\right) Q_t diag\left(q_{1,t}^{-1/2}, \ldots, q_{n,t}^{-1/2}\right) \tag{5}$$

The expression for h is a univariate GARCH. For the GARCH(1,1) model, the elements of H_t can be written as follows:

$$h_{i,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1} \tag{6}$$

Q_t is a symmetric positive definite matrix that can be written in the following form:

$$Q_t = (1 - \theta_1 - \theta_2)\overline{Q} + \theta_1 z_t z_{t-1}' + \theta_2 Q_{t-1} \tag{7}$$

where $\overline{Q}$ is the $n \times n$ unconditional correlation matrix of the standardized residuals $z_{i,t}$ ($z_{i,t} = \varepsilon_{i,t} / \sqrt{h_{i,t}}$). The parameters θ_1 and θ_2 are non-negative and are related to the

exponential smoothing process used to construct the dynamic conditional correlations. The DCC model is mean-reverting as long as $\theta_1 + \theta_2 < 1$. The correlation is estimated as:

$$\rho_{i,j,t} = \frac{q_{i,j,t}}{\sqrt{q_{i,i,t} q_{j,j,t}}} \tag{8}$$

Since the above DCC model does not allow for asymmetries and asset-specific news impact parameter, Cappiello et al. [39] developed the ADCC model to incorporate asymmetric effects and asset-specific news impact. For the ADDC model, the dynamics of Q is of the following form:

$$Q_t = \left(\overline{Q} - A'\overline{Q}A - B'\overline{Q}B - G'\overline{Q}^{-}G\right) + A'z_{t-1}z'_{t-1}A + B'Q_{t-1}B + G'z_t^{-}z_t'^{-}G \tag{9}$$

where A, B and G are $n \times n$ parameter matrices and z_t^- is the zero-threshold standardized error, which is equal to z_t when less than zero, and zero otherwise. $\overline{Q}$ and $\overline{Q}^-$ are the unconditional matrices of z_t and z_t^-, respectively.

3.2. Data and Descriptive Statistics

We selected daily data of bitcoin and 14 asset prices from 6 June 2013, to 2 August 2021 as the sample, with the number of observations being 1703. Data were obtained from the Wind and Yahoo Finance databases. These assets cover four categories, namely, stock, bond, commodity and currency. The stock sample included the MSCI world index, S&P 500, FTSE 100, DAX 30, Nikkei 225 and SSEC. The bond sample includes the US bond index, non-U.S. bond index and emerging markets bond index, which are measured by the prices of ETFs that track the three indices. The commodity sample includes the S&P GSCI, CRB commodity index, Brent oil and gold. The currency sample is the U.S. dollar index.

Figure 1 plots the time series of the prices of bitcoin versus each asset. The correlation between bitcoin and most asset prices is not stable, with some periods moving in the same direction and some moving in the opposite direction. For example, bitcoin was positively correlated with the S&P 500 during 2017–2018 and then became negatively correlated in 2019. This suggests that the linkages between bitcoin and these traditional asset prices are time-varying, thus, necessitating the use of an (asymmetric) DCC model to more accurately capture the dynamic correlation between bitcoin and each asset. The daily percentage returns of all asset prices are calculated based on the following equation:

$$r_t = 100 * (\ln P_t - \ln P_{t-1}) \tag{10}$$

where r_t represents the return of each asset and P_t represents the original asset price. Table 1 presents descriptive statistics for the return of each asset. Bitcoin had a much higher mean return than other assets, which reflected the long-term upward trend in bitcoin prices over the sample period. Bitcoin also had a much higher standard deviation than other assets, indicating that bitcoin prices were extremely volatile. The level of skewness in all asset returns was not high, except for the emerging markets bond index, which exhibited a clear left skew. Most asset returns had high kurtosis. The Jarque–Bera (J-B) test significantly rejected the assumption of normality of the distribution for all series.

GARCH modeling requires data stationarity to ensure the validity of the estimation, and having ARCH effects is also a prerequisite for GARCH modeling. We used the augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests to perform unit root tests on the return series (i.e., the first-order difference of asset price) and the Portmanteau test to perform an ARCH effect test on the return series, with the results reported in Tables 2 and 3, respectively. All return series were stationary at the 1% significance level and had significant ARCH effects at lags of order 5, 10 and 15, thus, satisfying the GARCH modeling conditions.

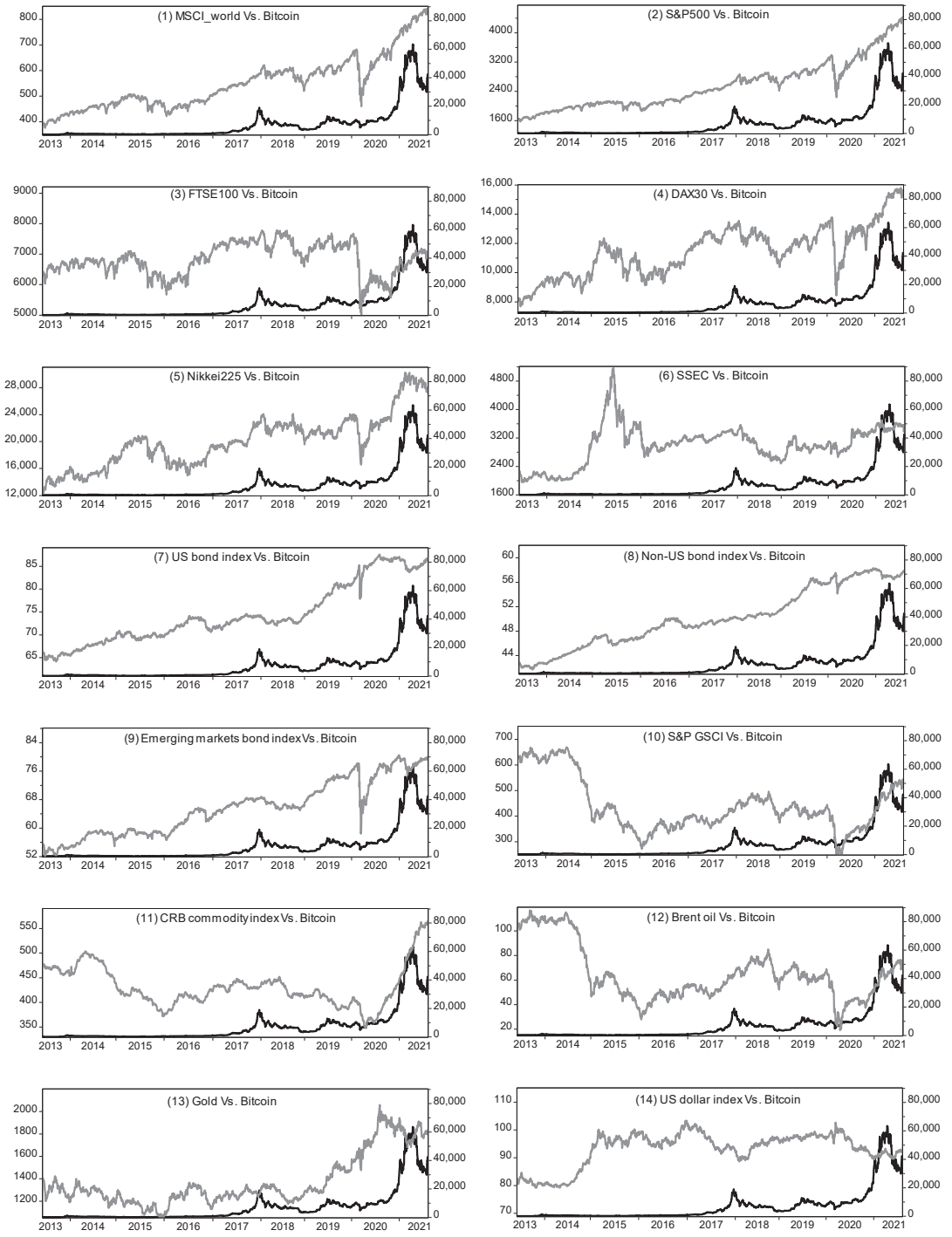


Figure 1. Time series of the prices of bitcoin versus various traditional financial assets. In each graph, the bitcoin price series is marked with a black line, corresponding to the right axis; the traditional asset price series is marked with a gray line, corresponding to the left axis.

Table 1. Descriptive statistics.

Variable	Mean	Median	Max.	Min.	Std. Dev.	Skewness	Kurtosis	J-B	*p* Value
Bitcoin	0.339	0.224	48.478	−49.397	5.685	−0.20	12.78	6795	0.0000
MSCI_world	0.045	0.083	7.793	−9.642	0.955	−1.44	23.52	30,451	0.0000
S&P500	0.059	0.084	8.968	−12.765	1.159	−1.23	24.35	32,771	0.0000
FTSE100	0.006	0.050	8.667	−11.512	1.114	−0.78	15.36	11,007	0.0000
DAX30	0.038	0.091	10.414	−13.055	1.350	−0.70	13.52	7989	0.0000
Nikkei225	0.045	0.060	7.731	−8.253	1.388	−0.12	6.75	1004	0.0000
SSEC	0.025	0.070	7.548	−8.873	1.499	−0.95	9.49	3249	0.0000
U.S. bond index	0.016	0.023	4.133	−5.592	0.316	−1.51	94.07	589,195	0.0000
Non-U.S. bond index	0.018	0.019	0.897	−2.320	0.213	−1.77	21.48	25,118	0.0000
Emerging markets bond index	0.021	0.049	4.345	−8.381	0.578	−3.93	63.28	262,235	0.0000
S&P GSCI	−0.009	0.061	7.115	−12.523	1.461	−0.93	11.82	5772	0.0000
CRB commodity index	0.010	0.000	3.726	−3.192	0.425	0.17	11.60	5255	0.0000
Brent oil	−0.020	0.083	21.115	−27.976	2.732	−0.85	21.61	24,769	0.0000
Gold	0.015	0.016	6.287	−6.120	1.044	−0.09	6.79	1020	0.0000
U.S. dollar index	0.006	0.003	2.495	−2.142	0.449	0.13	5.37	403	0.0000

Note: The J-B statistic is used to test the normality of the distribution of variables, and its null hypothesis is that the variable follows a normal distribution. The *p* value corresponds to the J-B test.

Table 2. Unit root test.

Variables	ADF			PP		
	Statistic	*p* Value	Result	Statistic	*p* Value	Result
Bitcoin	−21.28	0.0000	I(0)	−44.20	0.0001	I(0)
MSCI_world	−23.58	0.0000	I(0)	−41.42	0.0000	I(0)
S&P500	−27.57	0.0000	I(0)	−48.47	0.0001	I(0)
FTSE100	−41.24	0.0000	I(0)	−41.26	0.0000	I(0)
DAX30	−40.08	0.0000	I(0)	−40.07	0.0000	I(0)
Nikkei225	−42.12	0.0000	I(0)	−42.22	0.0000	I(0)
SSEC	−39.97	0.0000	I(0)	−39.98	0.0000	I(0)
U.S. bond index	−18.10	0.0000	I(0)	−42.74	0.0000	I(0)
Non-U.S. bond index	−41.59	0.0000	I(0)	−41.59	0.0000	I(0)
Emerging markets bond index	−14.82	0.0000	I(0)	−37.19	0.0000	I(0)
S&P GSCI	−41.67	0.0000	I(0)	−41.83	0.0000	I(0)
CRB commodity index	−20.19	0.0000	I(0)	−41.53	0.0000	I(0)
Brent oil	−40.07	0.0000	I(0)	−40.12	0.0000	I(0)
Gold	−40.85	0.0000	I(0)	−40.86	0.0000	I(0)
U.S. dollar index	−41.13	0.0000	I(0)	−41.13	0.0000	I(0)

Table 3. ARCH effect test.

Variables	ARCH(5)		ARCH(10)		ARCH(15)	
	Statistic	*p* Value	Statistic	*p* Value	Statistic	*p* Value
Bitcoin	89.13	0.0000	98.218	0.0000	105.45	0.0000
MSCI_world	580.03	0.0000	592.67	0.0000	628.76	0.0000
S&P500	714.77	0.0000	735.71	0.0000	807.19	0.0000
FTSE100	280.37	0.0000	409.82	0.0000	450.28	0.0000
DAX30	174.22	0.0000	298.34	0.0000	331.67	0.0000
Nikkei225	154.97	0.0000	213.87	0.0000	222.62	0.0000
SSEC	183.46	0.0000	193.38	0.0000	217.98	0.0000
U.S. bond index	377.26	0.0000	505.76	0.0000	536.61	0.0000
Non-U.S. bond index	445.02	0.0000	520.17	0.0000	532.95	0.0000
Emerging markets bond index	539.35	0.0000	705.25	0.0000	756.29	0.0000
S&P GSCI	140.46	0.0000	226.2	0.0000	241.21	0.0000
CRB commodity index	132.10	0.0000	134.20	0.0000	137.72	0.0000
Brent oil	117.79	0.0000	265.92	0.0000	311.46	0.0000
Gold	57.90	0.0000	98.66	0.0000	101.70	0.0000
U.S. dollar index	48.41	0.0000	60.45	0.0000	73.482	0.0000

Note: This table reports the results of the ARCH effect test at lags of order 5, 10 and 15. The null hypothesis for this test is "no ARCH effect".

Given the extremely high short-term volatility of bitcoin prices, bitcoin's short-term correlation with other assets is likely to be disturbed by its sharp short-term volatility. Intuitively, the long-term correlation between bitcoin and other asset prices is likely to be more

stable than the short-term correlation. To this end, we not only performed ADCC-GARCH analysis on daily frequency data of bitcoin and other asset prices, but also further performed ADCC-GARCH analysis on weekly and semi-monthly frequency data and then compared the results at different time frequencies to examine the differences in the linkage between bitcoin and various assets at different time frequencies. The weekly and semi-monthly frequency samples were obtained by taking the weekly and semi-monthly end-of-period values of the daily frequency samples, respectively, and the number of observations for both samples was 422 and 196, respectively. In addition to the weekly and semi-monthly frequencies, we also established an ADCC-GARCH model for the monthly frequency sample. However, since the number of observations for the monthly sample was only 97, the algorithm could not converge when performing ADCC-GARCH estimation. Therefore, it was abandoned.

3.3. ADCC-GARCH Estimation Results

The mean equation of the GARCH model was set to be AR(1) with the intercept term included, and the parameters of the intercept term and AR(1) were denoted by μ and φ, respectively. The variance equation was set to be GARCH(1,1), and the parameters of its intercept term, ARCH term and GARCH term were denoted by ω, α and β, respectively. The dynamic correlation parameters of the ADCC model were denoted by a, b and g, and υ denoted the joint distribution parameter of the model. Since all return series were not normally distributed, the multivariate joint t-distribution was selected for the distribution function.

Table 4 reports the ADCC-GARCH estimation results for the daily frequency sample of bitcoin and each asset price. In the variance equation, the coefficients of the ARCH and GARCH terms for all assets were significantly positive at least at the 5% level, indicating that the GARCH(1,1) setting was plausible. The coefficients of the GARCH term for all assets were much larger than the coefficients of the ARCH term, indicating that the conditional variance was more influenced by its prior period value and less sensitive to the previous period's return volatility, which showed that the price movements of these assets exhibited volatility clustering. From the ADCC estimation results, a was not negative, indicating that the standardized residuals with one lag had a positive effect on the dynamic correlation coefficient; b was close to 1, indicating that the dynamic correlation between bitcoin and other assets had strong persistence; and the sum of a and b was less than 1, ensuring that the conditional covariance matrix was positive definite and mean-reverting. ARCH tests were further performed on the residual terms of the ADCC-GARCH estimation results, and the results showed no significant ARCH effect. The above results indicated that the ADCC-GARCH estimation results were reliable. In addition to the daily frequency sample, we also performed ADCC-GARCH modeling for the weekly and semi-monthly frequency samples.

Table 4. Results of ADCC-GARCH estimation.

	Mean Equation		**Variance Equation**			**ADCC Parameters**			
	μ	φ	ω	α	β	a	b	g	υ
MSCI_world	0.0686 ***	0.1441 ***	0.0374 ***	0.1544 ***	0.7928 ***	0.0063	0.9849 ***	0.0000	4.0000 ***
	(0.0006)	(0.0000)	(0.0063)	(0.0001)	(0.0000)	(0.1662)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0078)	(0.9388)	(0.0157)	(0.0012)	(0.0000)				
S&P500	0.0996 ***	−0.0500 *	0.0733 ***	0.2741 ***	0.6756 ***	0.0066	0.9838 ***	0.0000	4.0000 ***
	(0.0000)	(0.0698)	(0.0004)	(0.0007)	(0.0000)	(0.8233)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0072)	(0.9385)	(0.0154)	(0.0011)	(0.0000)				
FTSE100	0.0212	0.0458	0.0507 **	0.1210 ***	0.8338 ***	0.0084	0.9552	0.0000	4.0000 ***
	(0.3310)	(0.1072)	(0.0197)	(0.0009)	(0.0000)	(0.8788)	(0.4951)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0081)	(0.9390)	(0.0162)	(0.0011)	(0.0000)				

Table 4. *Cont.*

	Mean Equation		Variance Equation			ADCC Parameters			
	μ	φ	ω	α	β	a	b	g	v
DAX30	0.0616 **	0.0208	0.0375 **	0.0783 ***	0.9000 ***	0.0078	0.9637 ***	0.0000	4.0000 ***
	(0.0243)	(0.4427)	(0.0164)	(0.0001)	(0.0000)	(0.2637)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0079)	(0.9388)	(0.0158)	(0.0012)	(0.0000)				
Nikkei225	0.0760 ***	−0.0367	0.0882 **	0.1146 ***	0.8396 ***	0.0030	0.9878 ***	0.0000	4.0000 ***
	(0.0069)	(0.1783)	(0.0318)	(0.0001)	(0.0000)	(0.3440)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0079)	(0.9388)	(0.0158)	(0.0012)	(0.0000)				
SSEC	0.0314	0.0266	0.0188	0.0778 ***	0.9177 ***	0.0069	0.9780 ***	0.0031	4.0000 ***
	(0.2363)	(0.3553)	(0.1164)	(0.0039)	(0.0000)	(0.1485)	(0.0000)	(0.5676)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 ***	0.1775 ***	0.7911 ***				
	(0.0081)	(0.9387)	(0.0158)	(0.0012)	(0.0000)				
U.S. bond index	0.0193 ***	−0.0025	0.0101 **	0.1479 **	0.6820 ***	0.0041	0.9383 ***	0.0000	4.1864 ***
	(0.0002)	(0.9266)	(0.0210)	(0.0117)	(0.0000)	(0.4438)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0080)	(0.9388)	(0.0158)	(0.0012)	(0.0000)				
Non-U.S. bond index	0.0238 ***	0.0388	0.0045 **	0.1619 **	0.7337 ***	0.0085	0.9324 ***	0.0000	4.0111 ***
	(0.0000)	(0.2261)	(0.0274)	(0.0102)	(0.0000)	(0.1982)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0080)	(0.9389)	(0.0158)	(0.0012)	(0.0000)				
Emerging markets bond index	0.0305 ***	0.0491	0.0103 ***	0.2980 ***	0.7010 ***	0.0058	0.9792 ***	0.0000	4.0000 ***
	(0.0003)	(0.1751)	(0.0025)	(0.0003)	(0.0000)	(0.3347)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0079)	(0.9387)	(0.0158)	(0.0012)	(0.0000)				
S&P GSCI	0.0174	0.0060	0.0333 **	0.0688 ***	0.9165 ***	0.0000	0.9927 ***	0.0009	4.0000 ***
	(0.5484)	(0.8154)	(0.0386)	(0.0000)	(0.0000)	(1.0000)	(0.0000)	(0.5566)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0079)	(0.9388)	(0.0158)	(0.0012)	(0.0000)				
CRB commodity index	0.0054	0.0847 ***	0.0103 *	0.0410 **	0.9015 ***	0.0024	0.9825 ***	0.0000	4.0000 ***
	(0.6288)	(0.0028)	(0.0999)	(0.0262)	(0.0000)	(0.5244)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0085)	(0.9387)	(0.0158)	(0.0012)	(0.0000)				
Brent oil	0.0478	−0.0190	0.0952 **	0.1205 ***	0.8761 ***	0.0000	0.9927	0.0010	4.0000 ***
	(0.3078)	(0.4767)	(0.0247)	(0.0000)	(0.0000)	(1.0000)	(0.5850)	(0.9954)	(0.0000)
Bitcoin	0.2708 **	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0104)	(0.9376)	(0.0141)	(0.0013)	(0.0000)				
Gold	0.0087	0.0016	0.0076 **	0.0271 ***	0.9657 ***	0.0066	0.9780 ***	0.0000	4.0000 ***
	(0.7094)	(0.9521)	(0.0265)	(0.0000)	(0.0000)	(0.1151)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0080)	(0.9389)	(0.0157)	(0.0012)	(0.0000)				
U.S. dollar index	0.0047	−0.0060	0.0012 *	0.0355 ***	0.9589 ***	0.0076	0.9736 ***	0.0000	4.3210 ***
	(0.6250)	(0.8119)	(0.0778)	(0.0000)	(0.0000)	(0.1743)	(0.0000)	(1.0000)	(0.0000)
Bitcoin	0.2708 ***	−0.0025	1.7513 **	0.1775 ***	0.7911 ***				
	(0.0080)	(0.9389)	(0.0158)	(0.0012)	(0.0000)				

Note: This table reports the results of ADCC-GARCH estimation for the daily frequency sample. The p values are in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Figure 2 shows the trend of dynamic correlation coefficients between bitcoin and various assets, including daily, weekly and semi-monthly frequencies. The correlation coefficients between bitcoin and each asset all exhibited significant time variability, suggesting that using an ADCC-GARCH approach was necessary to capture the dynamic correlation between bitcoin and each asset. We classified these assets into stock, bond, commodity and currency to further analyze the dynamic correlation between bitcoin and different classes of assets.

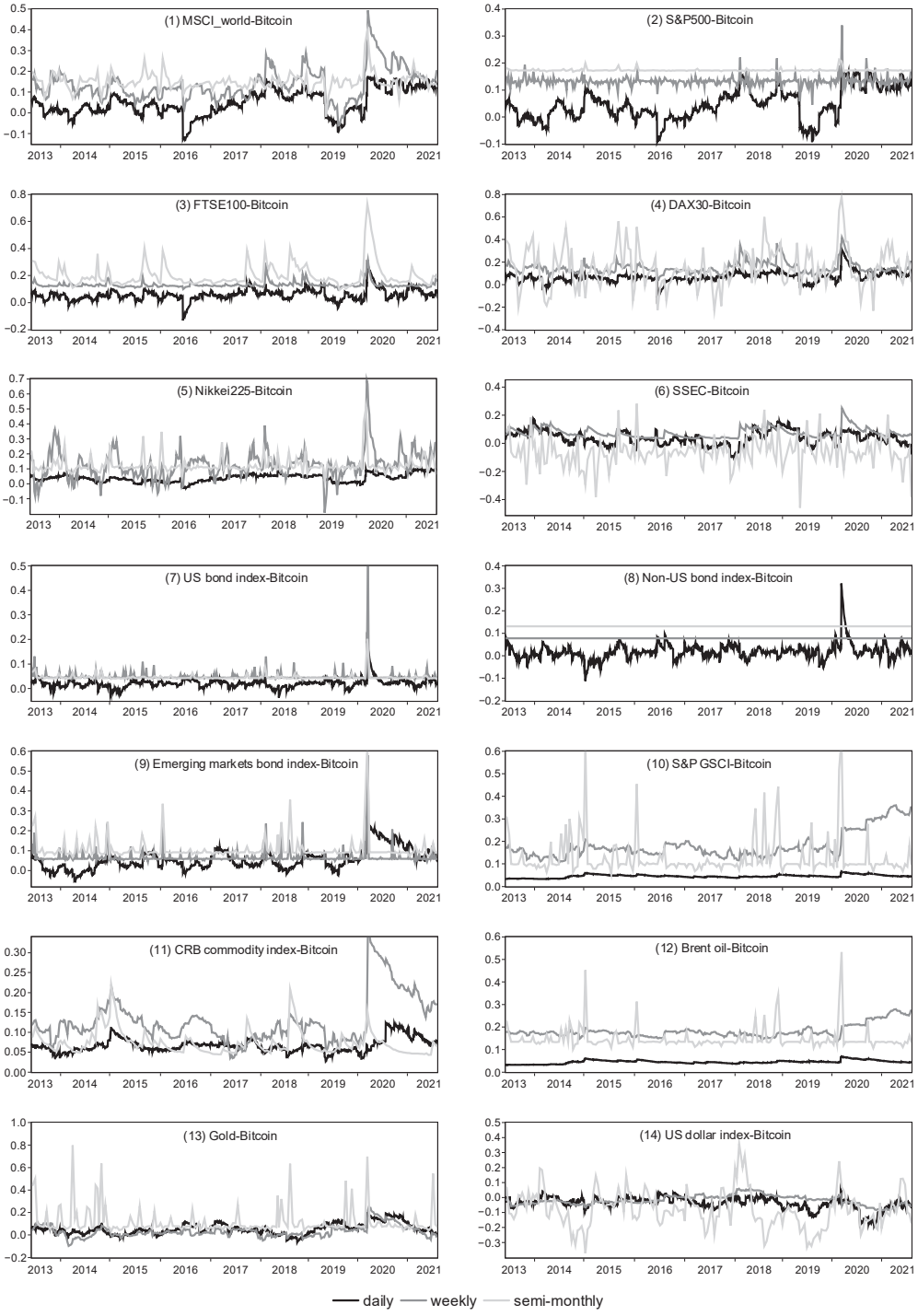


Figure 2. Daily, weekly and semi-monthly dynamic correlation coefficients between bitcoin and other assets.

Figure 2, panels (1)–(6), show the trend of dynamic correlation coefficients between bitcoin and representative stock indices. Firstly, in the daily frequency dimension, the correlation coefficients between bitcoin and all stock indices were low and largely fluctuated around 0, showing no sustained positive or negative correlation. The reason for this may be that bitcoin's high short-term volatility undermines its short-term correlation with other assets. Secondly, as the frequency changed from high to low, bitcoin began to show a significant positive correlation with most stock indices. In both the weekly and semi-monthly frequency dimensions, the correlation coefficients between bitcoin and global, U.S., U.K., German and Japanese stock indices were consistently positive in most periods, and the magnitude of the coefficients was also significantly higher. In particular, the dynamic correlation coefficients of bitcoin with global, U.S. and U.K. stock indices showed a clear "semimonthly frequency > weekly frequency > daily frequency". This showed that bitcoin had a weak correlation with major stock prices in the short term, but a more stable positive correlation in the long term. Thirdly, bitcoin's linkage with SSEC had different characteristics from its linkage with stock indices of developed countries. While the correlation between bitcoin and SSEC in the daily frequency dimension fluctuated around 0 over the full sample interval, the daily frequency correlation between the two was negative for most periods before 2017, which was consistent with the findings of Wang et al. [28]. Furthermore, the dynamic correlation coefficient between bitcoin and SSEC in the semi-monthly frequency dimension was negative in most periods, indicating that bitcoin and SSEC are negatively correlated in the long term. Unlike developed countries, China's financial markets have long been subject to capital controls, resulting in relatively limited channels for investors to invest abroad. In this context, when there is a long-term downward trend in the Chinese stock market, investors tend to enter the cryptocurrency (e.g., bitcoin) market to hedge their domestic stock investment losses, which is a possible reason for the negative correlation between bitcoin and the SSEC index in the long run. Fourth, the bitcoin-stock linkage increased sharply when subjected to exogenous extreme shocks. This observation was consistent with Kwapień et al. [3], who also found that the level of correlation between the cryptocurrency market and the stock market becomes higher during turbulent periods. Following the outbreak of COVID-19 in early 2020, the dynamic correlation coefficients between bitcoin and all stock indices rose rapidly, with the weekly frequency correlation coefficients of bitcoin with global, U.S., Japanese and Chinese stock indices rising sharply to approximately 0.5, 0.35, 0.7 and 0.25, respectively, and the semi-monthly frequency correlation coefficients with the U.K. and German stock indices both rising sharply to levels close to 0.8. The epidemic shock has led to a rapid rise in uncertainty and a sudden drop in investor risk appetite, causing investors to be less willing to hold not only traditional risk assets such as stocks, but also bitcoin, which has led to a sharp decline in both bitcoin and stock markets and a sharp increase in the positive linkage between bitcoin and the underlying stock indices.

Figure 2, panels (7)–(9), show the trend of dynamic correlation coefficients between bitcoin and representative bond indices. Firstly, similar to the dynamic correlation coefficient between bitcoin and stock prices, the dynamic correlation coefficient between bitcoin and bond prices exhibited a gradual increase from the short term (high frequency) to the long term (low frequency). In the daily frequency dimension, the dynamic correlation coefficients between bitcoin and the U.S. bond index, the non-U.S. bond index and the emerging markets bond index all fluctuated in a small range around 0; however, in the weekly and semi-monthly frequency dimensions, the coefficients were not only consistently positive but also significantly higher in magnitude. In particular, the dynamic correlation coefficients between bitcoin and both the non-U.S. bond index and emerging markets bond index showed a clear "semimonthly frequency > weekly frequency > daily frequency". Secondly, in the comparable time-frequency dimension, the correlation between bitcoin and bond prices was lower than its correlation with stock prices, indicating that bitcoin is more closely linked to the stock market than its linkage to the bond market. Thirdly, the linkage between bitcoin and bond prices also exhibited a sharp enhancement in response to extreme

shocks. The outbreak of COVID-19 in early 2020 caused both bitcoin and bond prices to fall significantly, resulting in the weekly frequency correlation coefficients of bitcoin with the U.S. bond index, the daily frequency correlation coefficients with the non-U.S. bond index and the semi-monthly frequency correlation coefficients with the emerging markets bond index rising sharply to over 0.5, 0.3 and 0.6, respectively.

Figure 2, panels (10)–(13), show the trend of dynamic correlation coefficients between bitcoin and representative commodity prices. Firstly, with the exception of gold, the dynamic correlation coefficients between bitcoin prices and the S&P GSCI, CRB commodity index and oil prices were positive for all periods and all time-frequency dimensions, indicating a persistent positive linkage between bitcoin and major commodity prices. Secondly, similar to the dynamic correlation coefficient between bitcoin and stocks/bonds, the dynamic correlation coefficients between bitcoin and the three commodities other than gold showed a gradual increase from the short term (high frequency) to the long term (low frequency). In the daily frequency dimension, the dynamic correlation coefficients between bitcoin and the three commodities, although consistently positive, were at a low level of below 0.1 for most periods, while in the weekly and semi-monthly frequency dimensions, the positive correlation coefficients were significantly higher. The correlation between bitcoin and gold on both daily and weekly frequencies fluctuated basically in a small range around 0. However, the correlation between the two on semi-monthly frequencies, increased. Thirdly, the bitcoin–commodity market linkage also exhibited a sharp increase in response to extreme shocks, which was consistent with Kwapień et al. [3], who showed that the level of correlation between the cryptocurrency market and the commodity market becomes higher during turbulent periods. Following the outbreak of COVID-19 in early 2020, both bitcoin and commodity prices fell rapidly, with bitcoin's semi-monthly frequency correlation coefficients with the S&P GSCI, oil prices and gold prices rising sharply to over 0.6, 0.5 and 0.6, respectively, and its weekly frequency correlation coefficient with the CRB commodity index rising rapidly to approximately 0.35.

Figure 2, panel (14), shows the trend of the dynamic correlation coefficient between bitcoin and the U.S. dollar index. At all frequencies, the dynamic correlation coefficient between bitcoin and the U.S. dollar index was negative in most periods, indicating that bitcoin price has an inverse linkage to the U.S. dollar index and that bitcoin can be used as an effective hedge against dollar depreciation. Bitcoin had a weak negative correlation with the U.S. dollar index on the daily and weekly frequencies, but showed a strong negative correlation for most periods on the semi-monthly frequency, peaking at nearly -0.4. The reason for the negative correlation between bitcoin and the dollar index may be interpreted in two ways. First, since bitcoin is denominated in USD, a dollar depreciation will cause bitcoin to become cheaper, thus, increasing demand for bitcoin and driving its price upward. Second, as deduced from the previous results regarding the predominantly positive correlation between bitcoin and stock/bond/commodity prices, bitcoin is closer in nature to a risk asset, while the U.S. dollar is typically a safe-haven asset, so the two prices naturally exhibit a negative correlation.

In summary, the linkage between bitcoin and various assets varies by asset class and time frequency, and can undergo significant structural changes in response to exogenous shocks in international financial markets. In terms of asset classes, bitcoin was positively correlated with risk assets including stock, commodity and bond, and bitcoin's positive correlation with stock or commodity was stronger than its positive correlation with bond; bitcoin had a significant negative correlation with the U.S. dollar, a typical safe-haven asset. As such, bitcoin is closer in nature to a risk asset than a safe-haven asset. In terms of time frequency, the long-term correlation between bitcoin and various asset prices was significantly stronger than the short-term correlation, mainly because the short-term high volatility and speculative nature of the bitcoin market undermine its short-term correlation with other assets. Finally, the linkage between bitcoin and the risk assets can increase sharply in response to exogenous extreme shocks. For example, after the outbreak of COVID-19 in early 2020, the plunge in investor risk appetite led to a sharp decline in the

prices of bitcoin and risk assets (including stocks, bonds and commodities), at which point the positive linkage between bitcoin and those risk assets rose rapidly. This also suggests that the outbreak of the COVID-19 pandemic has accelerated the integration of bitcoin with traditional financial markets, transforming it into part of a global market that is increasingly correlated with traditional assets [4].

4. Bitcoin's Risk Diversification, Hedging and Safe-Haven Properties for Other Assets

The DCC coefficients estimated in the previous section can characterize the dynamic linkage between bitcoin and other assets and provide a basis for further identifying the risk diversification, hedging and safe-haven properties of bitcoin for those assets.

4.1. Criteria for Distinguishing an Asset's Risk Diversification, Hedging and Safe-Haven Properties

We first established the criteria by which bitcoin is considered a diversifier, hedge or a safe haven. We followed Baur and Lucey [40] and Ratner and Chiu [37] to define a diversifier, hedge and safe haven, as these have become standard in the literature. Following Baur and Lucey [40] and Ratner and Chiu [37], we then distinguished bitcoin's risk diversification, hedging and safe-haven properties. The risk diversification property means that bitcoin is positively (but not perfectly) correlated with the returns of other assets. The hedging property includes a weak hedge, which means that bitcoin is uncorrelated with the returns of other assets, and a strong hedge, which means that bitcoin is negatively correlated with the returns of other assets. The safe-haven property includes a weak safe haven, which means that bitcoin is uncorrelated with the returns of other assets in times of market turmoil, and a strong safe haven, which means that bitcoin is negatively correlated with the returns of other assets in times of market turmoil. It is worth noting that our definition of a safe haven is ex post facto; that is, an asset (such as bitcoin) is considered a safe haven if investors show a preference for it during a market stress, which is of course only a necessary but not sufficient condition for an asset to be a safe-haven asset.

4.2. Identification Approach

The DCC coefficient between bitcoin and each asset is regressed on the lower quantile dummy variables of the corresponding asset returns to specifically identify the diversification, hedging and safe-haven properties of bitcoin for that asset. The regression equation is as follows:

$$DCC_t = m_0 + m_1 D(r_{other\ asset} q_{10}) + m_2 D(r_{other\ asset} q_5) + m_3 D(r_{other\ asset} q_1) + \varepsilon_t \quad (11)$$

where $r_{other\ asset}$ is the return on the corresponding asset, and $D(r_{other\ asset} q_{10})$, $D(r_{other\ asset} q_5)$ and $D(r_{other\ asset} q_1)$ denote the 10%, 5%, and 1% lower quartile dummy variables for the return on that asset, respectively. Figure 3 shows the 10%, 5% and 1% lower quartile values of daily returns for 14 traditional financial assets. The observations below the lower quartile values comprise the analysis period for our test of bitcoin's safe-haven properties. For observations below the 10% lower quantile value, we set $D(r_{other\ asset} q_{10}) = 1$; for observations below the 5% lower quantile value, we set $D(r_{other\ asset} q_5) = 1$; and for observations below the 1% lower quantile value, we set $D(r_{other\ asset} q_1) = 1$.

If m_0 is significantly positive, then bitcoin has risk diversification capability for that asset; if m_0 is equal to 0, then bitcoin has weak hedging capability for that asset; if m_0 is negative, then bitcoin has strong hedging capability for that asset. If m_1, m_2 and m_3 are not significantly nonzero, then bitcoin has a weak safe-haven capability for that asset; if m_1, m_2 and m_3 are significantly negative, then bitcoin has a strong safe-haven capability for that asset.

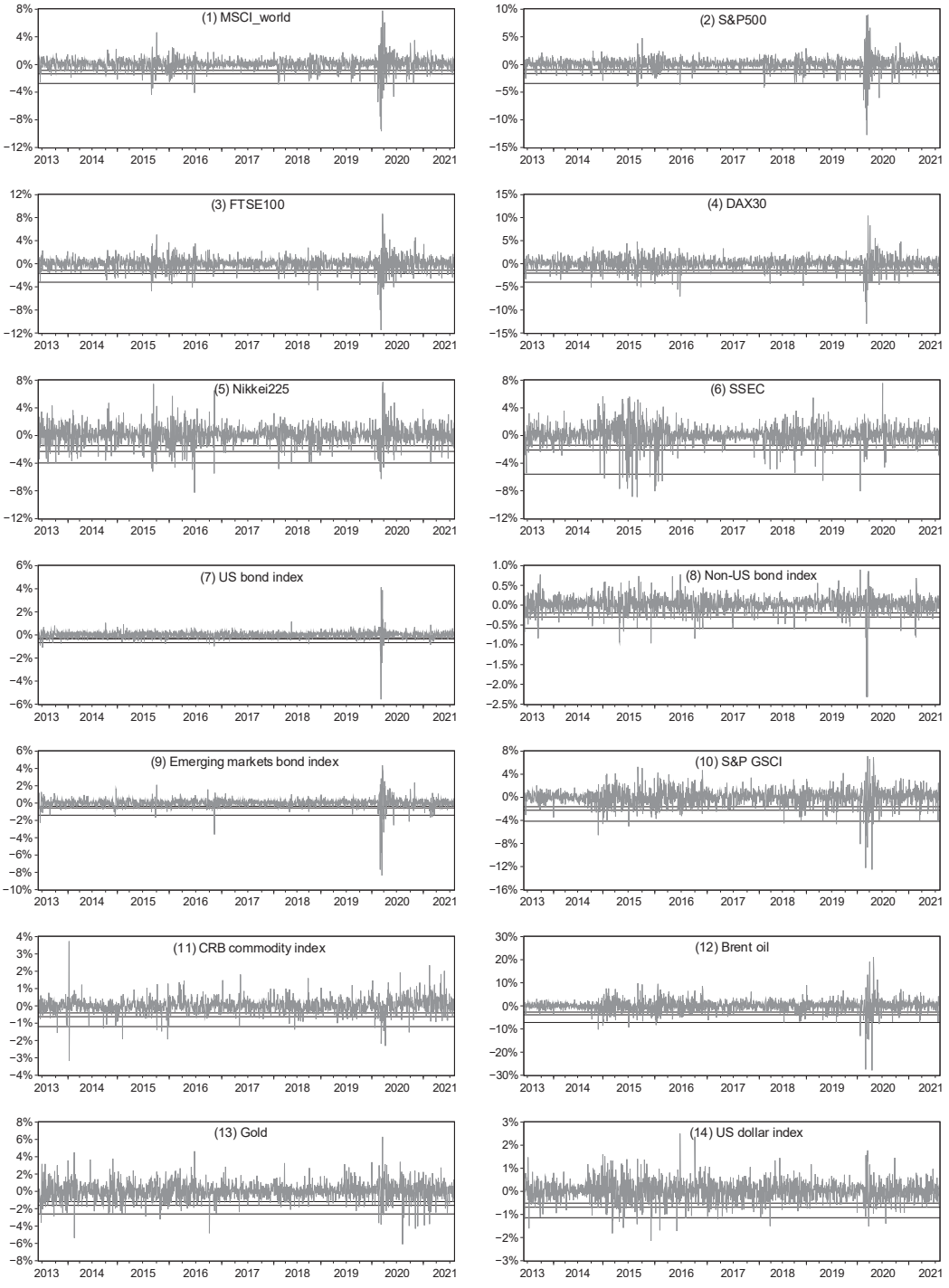


Figure 3. The 10%, 5% and 1% lower quartile values of 14 traditional asset returns (daily frequency). The three horizontal lines from top to bottom in each graph indicate the 10%, 5% and 1% lower quartile values, respectively.

4.3. Regression Results

We extracted the pairwise DCC series between bitcoin and other assets from the ADCC-GARCH estimation results and then regressed the DCC series based on Equation (11) to assess the diversification, hedging and safe-haven properties of bitcoin for each asset. For example, by regressing the DCC series between bitcoin and the MSCI world index on a constant (m_0) and three dummy variables (m_1, m_2 and m_3), we can assess the diversification or hedging capability of bitcoin for the global stock market based on the value and significance level of m_0 and the safe-haven capability of bitcoin for the global stock market based on the values and significance levels of m_1, m_2 and m_3. Tables 5–7 report the regression results for the daily, weekly and semi-monthly DCC series, respectively.

Table 5. Regression results for the risk diversification, hedging and safe-haven properties of bitcoin for other assets (daily frequency).

	10% **Quantile** (m_1)	5% **Quantile** (m_2)	1% **Quantile** (m_3)	m_0
MSCI_world	0.0049	0.0052	0.0141	0.0481 ***
	(0.4753)	(0.6080)	(0.3900)	(0.0000)
S&P500	0.0146 **	−0.0031	0.0247	0.0455 ***
	(0.0229)	(0.7416)	(0.1046)	(0.0000)
FTSE100	−0.0039	0.0067	0.0586 ***	0.0572 ***
	(0.4001)	(0.3302)	(0.0000)	(0.0000)
DAX30	0.0030	−0.0031	0.0406 ***	0.0679 ***
	(0.5659)	(0.6855)	(0.0010)	(0.0000)
Nikkei225	0.0007	−0.0025	0.0046	0.0390 ***
	(0.7960)	(0.5561)	(0.5064)	(0.0000)
SSEC	0.0076	−0.0056	−0.0176	0.0391 ***
	(0.1571)	(0.4772)	(0.1721)	(0.0000)
U.S. bond index	−0.0009	−0.0023	0.0240 ***	0.0212 ***
	(0.6568)	(0.4624)	(0.0000)	(0.0000)
Non-U.S. bond index	−0.0051	0.0031	0.0198 **	0.0184 ***
	(0.1968)	(0.5963)	(0.0363)	(0.0000)
Emerging markets bond index	0.0105 *	−0.0070	0.0916 ***	0.0532 ***
	(0.0632)	(0.3981)	(0.0000)	(0.0000)
S&P GSCI	0.0023 ***	0.0008	0.0043 **	0.0451 ***
	(0.0017)	(0.4658)	(0.0159)	(0.0000)
CRB commodity index	0.0003	−0.0027	−0.0025	0.0671 ***
	(0.8803)	(0.3298)	(0.5879)	(0.0000)
Brent oil	0.0022 ***	0.0032 ***	0.0039 **	0.0445 ***
	(0.0086)	(0.0073)	(0.0461)	(0.0000)
Gold	−0.0031	0.0091	0.0431 ***	0.0543 ***
	(0.5665)	(0.2536)	(0.0009)	(0.0000)
U.S. dollar index	−0.0016	0.0060	0.0072	−0.0373 ***
	(0.7287)	(0.3643)	(0.5032)	(0.0000)

Note: p values are in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table 6. Regression results for the risk diversification, hedging and safe-haven properties of bitcoin for other assets (weekly frequency).

	10% **Quantile** (m_1)	5% **Quantile** (m_2)	1% **Quantile** (m_3)	m_0
MSCI_world	0.0263	−0.0138	0.0125	0.1357 ***
	(0.1283)	(0.5845)	(0.7494)	(0.0000)
S&P500	0.0009	0.0174 ***	−0.0206 **	0.1318 ***
	(0.8267)	(0.0037)	(0.0268)	(0.0000)
FTSE100	0.0025	−0.0021	0.0501 ***	0.1314 ***
	(0.6155)	(0.7775)	(0.0000)	(0.0000)
DAX30	−0.0005	0.0193	0.0519 *	0.1489 ***
	(0.9693)	(0.2701)	(0.0580)	(0.0000)
Nikkei225	−0.0215	−0.0080	0.0473	0.1290 ***
	(0.2985)	(0.7904)	(0.3139)	(0.0000)
SSEC	0.0070	−0.0112	−0.0050	0.0743 ***
	(0.3758)	(0.3320)	(0.7807)	(0.0000)
U.S. bond index	0.0010	−0.0031	0.0071	0.0491 ***
	(0.8678)	(0.7188)	(0.6011)	(0.0000)

Table 6. *Cont.*

	10% Quantile (m_1)	5% Quantile (m_2)	1% Quantile (m_3)	m_0
Non-U.S. bond index	3.53×10^{-9}	-9.88×10^{-8}	4.40×10^{-7} ***	0.0772 ***
	(0.9610)	(0.3474)	(0.0075)	(0.0000)
Emerging markets bond index	0.0032	−0.0026	0.0755 ***	0.0665 ***
	(0.6917)	(0.8243)	(0.0001)	(0.0000)
S&P GSCI	−0.0109	0.0276	−0.0164	0.1805 ***
	(0.4083)	(0.1503)	(0.5828)	(0.0000)
CRB commodity index	0.0019	0.0024	0.0691 **	0.1274 ***
	(0.8851)	(0.9008)	(0.0195)	(0.0000)
Brent oil	−0.0124 *	0.0141	−0.0066	0.1828 ***
	(0.0571)	(0.1367)	(0.6563)	(0.0000)
Gold	0.0034	0.0162	0.0543 **	0.0303 ***
	(0.7694)	(0.3356)	(0.0395)	(0.0000)
U.S. dollar index	0.0058	−0.0041	−0.0084	−0.0182 ***
	(0.4148)	(0.6935)	(0.6027)	(0.0000)

Note: *p* values are in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table 7. Regression results for the risk diversification, hedging and safe-haven properties of bitcoin for other assets (semi-monthly frequency).

	10% Quantile (m_1)	5% Quantile (m_2)	1% Quantile (m_3)	m_0
MSCI_world	0.0025	−0.0006	0.0942 ***	0.1505 ***
	(0.8494)	(0.9734)	(0.0032)	(0.0000)
S&P500	−0.0002	0.0100 ***	−0.0107 ***	0.1730 ***
	(0.8938)	(0.0000)	(0.0052)	(0.0000)
FTSE100	0.0215	−0.0177	0.1584 **	0.1990 ***
	(0.4422)	(0.6645)	(0.0210)	(0.0000)
DAX30	0.0922 *	−0.1048	0.3512 ***	0.1428 ***
	(0.0881)	(0.1836)	(0.0079)	(0.0000)
Nikkei225	0.0126	0.0064	0.1188 ***	0.1203 ***
	(0.4645)	(0.7992)	(0.0050)	(0.0000)
SSEC	0.0726 **	−0.0299	−0.0981	−0.0671 ***
	(0.0271)	(0.5299)	(0.2172)	(0.0000)
U.S. bond index	−0.0026	0.0098	−0.0110	0.0453 ***
	(0.5480)	(0.1217)	(0.2998)	(0.0000)
Non-U.S. bond index	-2.44×10^{-9}	-5.92×10^{-8}	-3.97×10^{-8}	0.1307 ***
	(0.9245)	(0.1166)	(0.5259)	(0.0000)
Emerging markets bond index	−0.0095	0.0956 ***	0.0199	0.1025 ***
	(0.5822)	(0.0002)	(0.6362)	(0.0000)
S&P GSCI	0.0166	−0.0015	0.4531 ***	0.1210 ***
	(0.5002)	(0.9665)	(0.0000)	(0.0000)
CRB commodity index	0.0028	−0.0067	0.0659 ***	0.0736 ***
	(0.7768)	(0.6464)	(0.0074)	(0.0000)
Brent oil	0.0180	−0.0072	0.2724 ***	0.1429 ***
	(0.1671)	(0.7030)	(0.0000)	(0.0000)
Gold	−0.0393	−0.0011	0.1826 *	0.1183 ***
	(0.3131)	(0.9847)	(0.0551)	(0.0000)
U.S. dollar index	0.0506	−0.0428	0.0929	−0.0825 ***
	(0.2098)	(0.4673)	(0.3438)	(0.0000)

Note: *p* values are in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

The regression results for the daily frequency sample (Table 5) show that bitcoin cannot be considered a strong safe haven for all financial assets, implying that investors holding bitcoin cannot protect against extreme volatility in the prices of those assets. For the S&P 500, FTSE 100, DAX30, U.S. bond index, non-U.S. bond index, emerging markets bond index, S&P GSCI, oil and gold, the coefficients of the dummy variables characterizing the lower quartile values of their returns (m_1, m_2 and m_3) are all significantly positive in at least one case, which can only mean that bitcoin is an effective risk diversifier at the corresponding lower quartile levels of these asset returns. In terms of hedging properties, only the constant term (m_0) of the U.S. dollar index is significantly negative across all assets, implying that bitcoin is an effective tool to hedge against the movement of the USD exchange rate and that USD investors in the FX market tend to use bitcoin to hedge their currency portfolios. This is consistent with the findings of Dyhrberg [21], who also shows that bitcoin can be

used as a hedge for the USD/EUR and USD/GBP exchange rates. All assets other than the U.S. dollar have significantly positive m_0, so for these assets, bitcoin is only an effective risk diversifier. The main reason why bitcoin has risk diversification capabilities for most assets is that the mechanism of bitcoin price formation is more influenced by the bitcoin market's own supply and demand and investors' preferences for cryptocurrencies and less correlated with global macroeconomic and financial developments.

The regression results for the weekly frequency sample (Table 6) showed that in terms of hedging properties, only the U.S. dollar index had a significantly negative m_0, while all other assets had a significantly positive m_0. This suggests that in the weekly frequency dimension, bitcoin is still only a strong hedge against the U.S. dollar exchange rate and only an effective risk diversifier for other assets. In terms of safe-haven properties, m_3 was significantly negative for the S&P 500, and m_1 was significantly negative for crude oil, suggesting that bitcoin can be viewed to some extent as a safe haven against extreme volatility in the U.S. stock market (at the lower 1% quartile) and crude oil market (at the lower 10% quartile). This also means that investors have a tendency to put their money into the bitcoin market when there is a crisis in the U.S. stock market and the crude oil market. A possible explanation for why investors choose bitcoin as a safe haven in some cases, is that bitcoin operates as a decentralized cryptocurrency that is completely independent of any central institution. When traditional financial markets are under downward pressure, investors choose to seek shelter in the bitcoin market, which is independent of the traditional financial system and its underlying technical architecture.

The regression results for the semi-monthly frequency sample (Table 7) were similar to those for the weekly frequency samples. However, there were two differences. First, in terms of hedging properties, on the semi-monthly frequency, bitcoin not only had the ability to hedge the U.S. dollar index but also had a significant hedging effect on the SSEC. This suggests that bitcoin has the ability to hedge against declines in China's stock market in the long term. Moreover, the observation that bitcoin can hedge the Chinese stock market in the long term is also mentioned in the literature; for example, Chan et al. [24] found that bitcoin could serve as an effective hedge for the Chinese stock market in the monthly frequency dimension over the period 2010–2017. The likely reason is, that due to the existence of cross-border capital flow controls in China, when its domestic stock market is under long-term downward pressure, investors may seek to enter the bitcoin market to hedge against declines in the domestic stock market, as they have limited access to foreign investment. Second, in terms of safe-haven properties, on the semi-monthly frequency, bitcoin no longer has a strong hedging effect on the crude oil market but can still be considered to some extent as a safe haven against the risk of extreme volatility in the U.S. stock market at the low 1% quartile. Combining the weekly and semi-monthly frequency regression results, it was found that among the global stock markets, bitcoin had a significant safe-haven effect only for the U.S. stock market, which echoed the results found by Wang et al. [25] that the safe-haven properties of cryptocurrencies are more pronounced in developed markets or markets with larger market capitalization and higher liquidity.

In summary, bitcoin's risk diversification, hedging and safe-haven properties vary across time frequency dimensions, so distinguishing the holding period matters to bitcoin holders. For example, bitcoin does not have hedging ability against the Chinese stock market on the daily and weekly frequencies, but starts to produce a significant hedging effect on the semi-monthly frequency. In addition, bitcoin has no safe-haven properties for all assets on the daily frequency, but is starting to exhibit safe-haven properties for some assets (i.e., S&P 500 and crude oil) on the weekly and semi-monthly frequencies. One explanation for the differences in bitcoin's hedging and safe-haven properties at different time frequencies is that bitcoin's high short-term volatility and its strong speculative properties undermine its hedging and safe-haven properties in the short term and may even compromise its hedging and safe-haven properties in the long term. Another explanation is that bitcoin's hedging and safe-haven properties at different frequencies may be driven

by different factors, as price formation of bitcoin may be influenced by different factors in the long and short term.

5. Conclusions

Using bitcoin and 14 global financial asset price data covering stock, bond, commercial and currency for the period 2013–2021, this study applied the ADCC-GARCH approach to test the dynamic correlation between bitcoin and each asset at different time frequencies, and further identified the risk diversification, hedging and safe-haven properties of bitcoin for those traditional assets. The main findings are as follows:

(i) Bitcoin is positively linked to risk assets, including stocks, bonds and commodities, and negatively linked to the U.S. dollar, which is a typical safe-haven asset. Therefore, bitcoin is closer in nature to a risk asset than a safe-haven asset;
(ii) The high short-term volatility and speculative nature of the bitcoin market makes its long-term correlation with other assets stronger than the short-term correlation;
(iii) The positive linkage between bitcoin and risk assets increases sharply under extreme shocks (e.g., the outbreak of COVID-19 in early 2020);
(iv) Bitcoin can hedge against the U.S. dollar, and in the long term, bitcoin can hedge against the Chinese stock market and act as a safe haven for the U.S. stock market and crude oil. However, for most other traditional assets, bitcoin is only an effective diversifier.

Our conclusions provide useful insights for market participants and policymakers. First, because bitcoin is closer in nature to a risk asset, investors should allocate to bitcoin as a risk diversifier for traditional risk assets such as stock, bond and commodity, rather than as a hedge, especially in times of extreme exogenous shocks. Second, the short-term high volatility and speculative nature of the bitcoin market leads to great uncertainty in the short-term price of bitcoin, while also undermining its short-term correlation with major financial assets, making bitcoin's diversification, hedging and safe-haven properties vary across different time-frequency dimensions. This reminds bitcoin holders that it is important to distinguish between bitcoin holding periods. Investors who enter the bitcoin market should opt for long-term holdings as much as possible. Short-term speculation could expose them to significant investment risk and would likely result in large capital losses. Third, as uncertainties in global financial markets further increase in the post-epidemic era, policymakers and investors should keep paying attention to potential structural changes in the linkage between bitcoin and major asset prices under exogenous extreme shocks and the financial risks they may trigger. Finally, as the market for bitcoin trading is immature and the price is extremely unstable, individual investors should be discouraged from entering the cryptocurrency market represented by bitcoin, to protect the safety of their property.

Author Contributions: Conceptualization, P.W.; methodology, P.W., X.L. and S.W.; software, P.W., X.L. and S.W.; validation, P.W., X.L. and S.W.; formal analysis, P.W., X.L. and S.W.; investigation, P.W.; resources, P.W., X.L. and S.W.; data curation, P.W., X.L. and S.W.; writing—original draft preparation, P.W.; writing—review and editing, P.W., X.L. and S.W.; visualization, P.W., X.L. and S.W.; supervision, P.W., X.L. and S.W.; project administration, P.W. and X.L.; funding acquisition, P.W. and X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Zhejiang Provincial Philosophy and Social Science Planning Project of China (grant number 23NDJC023Z); the National Key Research and Development Program of China (grant number 2021QY2100); the National Natural Science Foundation of China (grant number 72173018); and the Zhejiang Provincial Natural Science Foundation of China (grant number LQ21G030005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available on request.

Acknowledgments: We would like to thank two anonymous referees and the subject editor of this journal for helpful comments and suggestions on earlier versions of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. *Decentralized Bus. Rev.* **2008**, 21260. Available online: https://www.ussc.gov/sites/default/files/pdf/training/annual-national-training-seminar/2018/Emerging_Tech_Bitcoin_Crypto.pdf (accessed on 7 September 2022).
2. Conlon, T.; McGee, R. Safe haven or risky hazard? Bitcoin during the COVID-19 bear market. *Financ. Res. Lett.* **2020**, *35*, 101607. [CrossRef] [PubMed]
3. Kwapień, J.; Wątorek, M.; Drożdż, S. Cryptocurrency market consolidation in 2020–2021. *Entropy* **2021**, *23*, 1674. [CrossRef] [PubMed]
4. Wątorek, M.; Drożdż, S.; Kwapień, J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* **2021**, *901*, 1–82. [CrossRef]
5. Yermack, D. Is Bitcoin a real currency? An economic appraisal. In *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data*; Chuen, D.L.K., Ed.; Academic Press: Cambridge, MA, USA, 2015; pp. 31–43.
6. Glaser, F.; Zimmermann, K.; Haferkorn, M.; Weber, M.C.; Siering, M. Bitcoin—Asset or Currency? *Revealing Users' Hidden Intentions.* 2014. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2425247 (accessed on 7 September 2022).
7. Böhme, R.; Christin, N.; Edelman, B.; Moore, T. Bitcoin: Economics, technology, and governance. *J. Econ. Perspect.* **2015**, *29*, 213–238. [CrossRef]
8. Luther, W.J.; Salter, A.W. Bitcoin and the bailout. *Q. Rev. Econ. Financ.* **2017**, *66*, 50–56. [CrossRef]
9. Popper, N. *Digital Gold: The Untold Story of Bitcoin*; Penguin: London, UK, 2015.
10. Dyhrberg, A.H. Bitcoin, gold and the dollar—A GARCH volatility analysis. *Financ. Res. Lett.* **2016**, *16*, 85–92. [CrossRef]
11. Klein, T.; Thu, H.P.; Walther, T. Bitcoin is not the New Gold–A comparison of volatility, correlation, and portfolio performance. *Int. Rev. Financ. Anal.* **2018**, *59*, 105–116. [CrossRef]
12. Shahzad, S.J.H.; Bouri, E.; Roubaud, D.; Kristoufek, L. Safe haven, hedge and diversification for G7 stock markets: Gold versus bitcoin. *Econ. Model.* **2020**, *87*, 212–224. [CrossRef]
13. Thampanya, N.; Nasir, M.A.; Huynh, T.L.D. Asymmetric correlation and hedging effectiveness of gold & cryptocurrencies: From pre-industrial to the 4th industrial revolution. *Technol. Forecast. Soc. Chang.* **2020**, *159*, 120195. [CrossRef]
14. Whelan, K. How is Bitcoin Different from the Dollar? *2013.* Available online: https://www.forbes.com/sites/karlwhelan/2013/11/19/how-is-bitcoin-different-from-the-dollar (accessed on 7 September 2022).
15. Valstad, O.C.A.; Vagstad, K. A Bit Risky? A Comparison between Bitcoin and Other Assets Using an Intraday Value at Risk Approach. Master's Thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2014.
16. Briere, M.; Oosterlinck, K.; Szafarz, A. Virtual currency, tangible return: Portfolio diversification with bitcoin. *J. Asset Manag.* **2015**, *16*, 365–373. [CrossRef]
17. Corbet, S.; Meegan, A.; Larkin, C.; Lucey, B.; Yarovaya, L. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Econ. Lett.* **2018**, *165*, 28–34. [CrossRef]
18. Guesmi, K.; Saadi, S.; Abid, I.; Ftiti, Z. Portfolio diversification with virtual currency: Evidence from bitcoin. *Int. Rev. Financ. Anal.* **2019**, *63*, 431–437. [CrossRef]
19. Gil-Alana, L.A.; Abakah, E.J.A.; Rojo, M.F.R. Cryptocurrencies and stock market indices. Are they related? *Res. Int. Bus. Financ.* **2020**, *51*, 101063. [CrossRef]
20. Eisl, A.; Gasser, S.M.; Weinmayer, K. Caveat Emptor: Does Bitcoin Improve Portfolio Diversification? *2015.* Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2408997 (accessed on 7 September 2022).
21. Dyhrberg, A.H. Hedging capabilities of bitcoin. Is it the virtual gold? *Financ. Res. Lett.* **2016**, *16*, 139–144. [CrossRef]
22. Yang, C.; Wang, X.; Gao, W. Is Bitcoin a better hedging and safe-haven investment than traditional assets against currencies? Evidence from the time-frequency domain approach. *N. Am. Econ. Financ.* **2022**, *62*, 101747. [CrossRef]
23. Bouri, E.; Molnár, P.; Azzi, G.; Roubaud, D.; Hagfors, L.I. On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Financ. Res. Lett.* **2017**, *20*, 192–198. [CrossRef]
24. Chan, W.H.; Le, M.; Wu, Y.W. Holding Bitcoin longer: The dynamic hedging abilities of Bitcoin. *Q. Rev. Econ. Financ.* **2019**, *71*, 107–113. [CrossRef]
25. Wang, P.; Zhang, W.; Li, X.; Shen, D. Is cryptocurrency a hedge or a safe haven for international indices? A comprehensive and dynamic perspective. *Financ. Res. Lett.* **2019**, *31*, 1–18. [CrossRef]
26. Shahzad, S.J.H.; Bouri, E.; Roubaud, D.; Kristoufek, L.; Lucey, B. Is Bitcoin a better safe-haven investment than gold and commodities? *Int. Rev. Financ. Anal.* **2019**, *63*, 322–330. [CrossRef]
27. Urquhart, A.; Zhang, H. Is Bitcoin a hedge or safe haven for currencies? An intraday analysis. *Int. Rev. Financ. Anal.* **2019**, *63*, 49–57. [CrossRef]
28. Wang, G.; Tang, Y.; Xie, C.; Chen, S. Is bitcoin a safe haven or a hedging asset? Evidence from China. *J. Manag. Sci. Eng.* **2019**, *4*, 173–188. [CrossRef]
29. Smales, L.A. Bitcoin as a safe haven: Is it even worth considering? *Financ. Res. Lett.* **2019**, *30*, 385–393. [CrossRef]
30. James, N. Dynamics, behaviours, and anomaly persistence in cryptocurrencies and equities surrounding COVID-19. *Physica A* **2021**, *570*, 125831. [CrossRef]

31. Balcilar, M.; Ozdemir, H.; Agan, B. Effects of COVID-19 on cryptocurrency and emerging market connectedness: Empirical evidence from quantile, frequency, and lasso networks. *Physica A* **2022**, *604*, 127885. [CrossRef]
32. Caferra, R.; Vidal-Tomás, D. Who raised from the abyss? A comparison between cryptocurrency and stock market dynamics during the COVID-19 pandemic. *Financ. Res. Lett.* **2021**, *43*, 101954. [CrossRef]
33. Grobys, K. When Bitcoin has the flu: On Bitcoin's performance to hedge equity risk in the early wake of the COVID-19 outbreak. *Appl. Econ. Lett.* **2021**, *28*, 860–865. [CrossRef]
34. 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. Financ.* **2020**, *54*, 101248. [CrossRef]
35. Wen, F.; Tong, X.; Ren, X. Gold or Bitcoin, which is the safe haven during the COVID-19 pandemic? *Int. Rev. Financ. Anal.* **2022**, *81*, 102121. [CrossRef]
36. Dutta, A.; Das, D.; Jana, R.K.; Vo, X.V. COVID-19 and oil market crash: Revisiting the safe haven property of gold and Bitcoin. *Resour. Policy* **2020**, *69*, 101816. [CrossRef]
37. Ratner, M.; Chiu, C.C.J. Hedging stock sector risk with credit default swaps. *Int. Rev. Financ. Anal.* **2013**, *30*, 18–25. [CrossRef]
38. Engle, R. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econ. Stat.* **2002**, *20*, 339–350. [CrossRef]
39. Cappiello, L.; Engle, R.F.; Sheppard, K. Asymmetric dynamics in the correlations of global equity and bond returns. *J. Financ. Econom.* **2006**, *4*, 537–572. [CrossRef]
40. Baur, D.G.; Lucey, B.M. Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financ. Rev.* **2010**, *45*, 217–229. [CrossRef]

entropy

MDPI

Article

Public-Key Cryptosystems and Bounded Distance Decoding of Linear Codes

Selda Çalkavur

Math Department, Faculty of Arts and Science, Kocaeli University, Kocaeli 41380, Turkey; selda.calkavur@kocaeli.edu.tr

Abstract: Error-correcting codes form an important topic in information theory. They are used to correct errors that occur during transmission on a noisy channel. An important method for correcting errors is bounded distance decoding. The public-key cryptosystem is a cryptographic protocol that has two different keys. One of them is a public-key that can be known by everyone, and the other is the private-key only known to the user of the system. The data encrypted with the public-key of a given user can only be decrypted by this user with his or her private-key. In this paper, we propose a public-key cryptosystem based on the error-correcting codes. The decryption is performed by using the bounded distance decoding of the code. For a given code length, dimension, and error-correcting capacity, the new system allows dealing with larger plaintext than other code based public-key cryptosystems.

Keywords: public-key cryptosystem; error correcting code; bounded distance decoding

Citation: Çalkavur, S. Public-Key Cryptosystems and Bounded Distance Decoding of Linear Codes. *Entropy* **2022**, *24*, 498. https://doi.org/10.3390/e24040498

Academic Editors: Stanisław Drożdż, Jarosław Kwapień and Marcin Wątorek

Received: 9 March 2022
Accepted: 31 March 2022
Published: 1 April 2022

1. Introduction

Public-key cryptosystems or asymmetric cryptosystems have been a subject of study since 1976. These systems consider two different keys, which are called public-key and private-key. These keys are not completely independent of each other. There must be a mathematical relationship as factoring, discrete logarithm, etc. [1,2]. The public-key cryptosystem was first introduced in 1976 by Diffie and Hellman [3]. Rivest, Shamir and Adleman's paper, known as the RSA cryptosystem [4], also present a public-key cryptosystem. The RSA cryptosystem was based on the factorization integers [5]. Merkle and Hellman [6] suggested a cryptosystem based on the difficulty of the integer packing "knapsack" problem.

The first public-key cryptosystem based on the error-correcting codes was presented by R. J. McEliece in 1978 [7]. He has employed error correcting codes, in particular binary Goppa codes, with a known decoding algorithm to construct the system. The generator matrix G plays an important role. The most important property of McEliece's cryptosystem is its large key size. Niederreiter suggested another code-based public-key cryptosystem that is based on the syndrome decoding of linear codes [8]. This system is used for the parity-check matrix H of a linear code. Thus, it is also the dual version of McEliece's cryptosystem. If it is used with exactly the same parameters [9], McEliece's cryptosystem and Niederreiter's cryptosystem offer an equivalent security. Li et al. [10] proposed new classes of trapdoor functions to solve the bounded distance decoding problem in lattices. Moreover, a lot of cryptosystems have been presented by using linear codes after McEliece's and Niederreiter's schemes. The use of subcodes of generalized Reed–Solomon codes was introduced by Berger and Loidreau [11]. Berlekamp et al. [12] studied the complexity of the decoding of arbitrary linear codes. Krouk [13] proposed a different class of public-key cryptosystems. Sidelnikov [14] introduced the use of Reed–Muller codes for cryptosystems. Berger et al. [15] and Misoczki-Barreto [16] proposed using quasi-cyclic and quasi-dyadic codes to shorten the McEliece key. The original parameters of the McEliece cryptosystem have been broken [17], but the general system is still considered safe.

In this study, we propose a public-key cryptosystem based on the error-correcting codes using a known bounded distance decoding method. We present the encryption and decryption algorithms by inspiring both McEliece's and Niederreiter's cryptosystems.

- McEliece's system has been constructed based on linear codes over $\mathbb{F}_2$.
- Both Niederreiter's and our system have been constructed based on linear codes over $\mathbb{F}_q$.
- However, in our cryptosystem, since it is easier to generate the pieces of keys, the encryption, decryption, and key generation are more effective than Niederreiter's cryptosystem.
- Another difference of our system from Niederreiter's is the use of the bounded distance decoding method, which corrects errors and guarantees unique decoding.
- It is impossible to find the private-key with public-key by an attacker in our public-key cryptosystem.
- Similarly, even if an enemy knows the public-key and ciphertext, he/she cannot calculate the plaintext.

These conditions ensure the new system is safe. Moreover, we consider some possible attacks in this paper. So, we analyze its security and performance, and we calculate some important parameters for our cryptosystem. When we compared it with McEliece's and Niederreiter's cryptosystems, we can say that our system performs better as regards encryption speed.

The rest of the paper is organized as follows. The next section gives the necessary background on coding theory and cryptography. Section 3 introduces the new public-key cryptosystem. Section 4 analyzes its security and examines some possible attacks. Section 5 compares it to the other code-based public-key cryptosystems. Section 6 concludes the paper.

2. Preliminaries

In this section, we remind of some important topics [18,19] that are necessary for the paper.

2.1. Linear Codes

Definition 1 (Linear Code). *A linear code C of length n and dimension k is a subspace of $(\mathbb{F}_q)^n$, where $\mathbb{F}_q$ is the finite field with q elements, q is a prime power, and k and n are positive integers such that $k \leq n$. It is denoted by an $[n,k]$-code. The error-correcting capacity of C is the maximum number t of errors that C can skillfully decode. All vectors of $(\mathbb{F}_q)^n$ that are orthogonal to every codeword of C consist of the dual code $C^{\perp}$ which is an $[n, n-k]$-code.*

Definition 2 (Hamming Weight). *The Hamming weight $w(x)$ of a vector x in $(\mathbb{F}_q)^n$ is the number of non-zero entries of x.*

Definition 3 (Generator Matrix). *A generator matrix G of C is the rows that are a basis of C. G is also a $k \times n$ matrix.*

Definition 4 (Parity-Check Matrix). *A parity-check matrix H for a linear code C is an $(n-k) \times n$ matrix which is a generator matrix for its dual code $C^{\perp}$.*

2.2. Coset Decoding

Definition 5. *Let C be an $[n,k]$-code over $\mathbb{F}_q$ and u be any vector in $(\mathbb{F}_q)^n$. The coset of C is defined as follows.*

$$u + C = \{u + x | x \in C\}. \tag{1}$$

Theorem 1 (Lagrange). *Suppose C is an $[n,k]$-code over $\mathbb{F}_q$. Then,*

(i) Every vector of $(\mathbb{F}_q)^n$ is in some coset of C;

(ii) Every coset contains exactly q^k vectors;

(iii) *Two cosets either are disjointed or coincided;*
(iv) *C contains exactly* q^{n-k} *cosets.*

Definition 6 (Coset Leader)**.** *The coset leader is the vector having a minimum weight in a coset. If a coset contains more than one vector which has the minimum weight, then it is chosen at random as the coset leader.*

Definition 7 (Syndrome Decoding)**.** *Consider H is a parity-check matrix of an* $[n, k]$*-code C. In this case,*

$$S(y) = yH^T \tag{2}$$

is called the syndrome of y, where y is any vector of $(\mathbb{F}_q)^n$*, the* $1 \times (n-k)$ *row vector. Moreover,*

$$S(y) = 0 \Longrightarrow y \in C. \tag{3}$$

Lemma 1. *Two vectors u and v are in the same coset of C if and only if they have the same syndrome.*

Corollary 1. *There is a one-to-one correspondence between cosets and syndromes.*

2.3. Public-Key Cryptosystems

A cryptosystem is an application of cryptographic methods and ensures the information security services. The cryptosystems can be examined under two titles as the public-key and private-key. Each person has a pair of keys; one is the public-key, and the other is the private-key. The public-key is accessible to the other users; however, the private-key should be stored so that only the owner can access it. Any person can send an encrypted message using the public-key, but only the private-key, which is a pair of public-keys, can decrypt the encrypted message. There is always the mathematical relationship between the public-key and private-key in the public-key cryptosystems. The hardness of two mathematical problems, as integer factoring and discrete logarithm, are used to generate these keys. So, it is impossible to obtain the private-key using the public-key.

The Diffie–Hellman cryptosystem [3] and RSA cryptosystem [4] are pioneers of public-key cryptosystems. However, McEliece [7] and Niederreiter [8] are the first founders of the code-based public-key cryptosystems.

2.4. McEliece's Public-Key Cryptosystem

McEliece's public-key cryptosystem is the first system based on the algebraic block codes; it was presented in 1978 [7]. In order to construct his cryptosystem, it used a binary $(n, k, 2t+1)$ Goppa code C. It is clear that n is the code length, k is the code dimension, and t is the error-correcting capacity of C. The encryption and decryption algorithms are as follows.

Private-key: G, S, P; where G is a $k \times n$ generator matrix, S is any $k \times k$ non-singular matrix, and P is any $n \times n$ permutation matrix.
Public-key: $G' = SGP$ and t.
Plaintexts: k bit vectors m over $\mathbb{F}_2$.
Encryption:

$$c = mG' + e, \tag{4}$$

where e is an n-bit error vector with Hamming weight t. So, c is the n-bit ciphertext.
Decryption:

$$cP^{-1} = (mS)G + eP^{-1} \tag{5}$$

since

$$c = mSGP. \tag{6}$$

It is used as the fast decoding algorithm for C to correct the error eP^{-1}; then, it is found mS and therefore m.

2.5. Niederreiter's Public-Key Cryptosystem

Niederreiter [8] proposed a knapsack-type public-key cryptosystem which is based on $(n, k, 2t+1)$ linear code C over $\mathbb{F}_q$.

Private-key: H, M, and P, where H is an $(n-k) \times n$ parity-check matrix of C, M is any $(n-k) \times (n-k)$ non-singular matrix, and P is any $n \times n$ permutation matrix, all over $\mathbb{F}_q$.
Public-key: $H' = MHP$ and t.
Plaintexts: n-dimensional vectors m over $\mathbb{F}_q$ with weight t.
Encryption: $c = mH'^T$, c is the ciphertext of dimension $n-k$.
Decryption:

$$c(M^T)^{-1} = (mP^T)H^T \tag{7}$$

since

$$c = m(MHP)^T. \tag{8}$$

It is used as the fast decoding algorithm for C to obtain mP^T and m.

3. The System

The construction of our public-key cryptosystem is based on $[n, k, 2t+1]$-code over $\mathbb{F}_q$. The syndrome-decoding procedure is used for decryption. The public-key and private-key are constructed by each user as follows.

(1) Select a generator $k \times n$ matrix G of a linear $[n, k, 2t+1]$-code C over $\mathbb{F}_q$, where t is the error-correcting capability.
(2) Construct a parity-check $(n-k) \times n$ matrix H from G for the code C.
(3) Select any non-zero syndrom vector h which has weight t and dimension $(n-k)$.
(4) Select a random non-singular $(n-k) \times (n-k)$ matrix M over $\mathbb{F}_q$.
(5) Calculate $n \times (n-k)$ matrix $H' = H^T \cdot M$, where H^T is denoted by the transpose of H.
(6) The public-key is (H', h).
(7) The private-key is (G, H, M).

Encryption:
Message: n dimension vector m over $\mathbb{F}_q$ with weight t.
Cryptogram: $c = mH' + h$
Decryption:

(1) Calculate $c' = cM^{-1}$;
(2) Obtain m by syndrome decoding c' in the code C.

Decryption is correct, since

$$w(hM^{-1}) = w(h), \tag{9}$$

it can be computed

$$c' = cM^{-1} = (mH' + h)M^{-1} = mH'M^{-1} + hM^{-1} \tag{10}$$

$$cM^{-1} = mH^T MM^{-1} + hM^{-1} \tag{11}$$

$$cM^{-1} = mH^T + hM^{-1} \tag{12}$$

$$cM^{-1} - hM^{-1} = mH^T \tag{13}$$

and the procedure of syndrome decoding may be effectively used.

Example 1. *Consider an* $[4, 2, 3]$*-code* C *over* $\mathbb{F}_3$*. The generator matrix* G *and parity-check matrix* H *are*

$$G = \begin{pmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 2 & 1 \end{pmatrix}, \tag{14}$$

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}. \tag{15}$$

$C = \{0000, 0121, 0212, 1022, 1110, 1201, 2011, 2102, 2220\}$. Select any non-singular matrix $M = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$. The syndromes and coset leaders of C are as follows.

Syndromes	Coset Leaders
(00)	(0000)
(11)	(1000)
(12)	(0100)
(10)	(0010)
(01)	(0001)
(22)	(2000)
(21)	(0200)
(20)	(0020)
(02)	(0002)

The size of different cosets of C is

$$3^{4-2} = 3^2 = 9.$$

So, there are also nine syndrome vectors, which are $\{00, 11, 12, 10, 01, 22, 21, 20, 02\}$. Calculate the matrix

$$H' = H^T \cdot M = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 2 & 2 \\ 1 & 2 \\ 2 & 0 \end{pmatrix} \tag{16}$$

and

$$M^{-1} = \begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix}. \tag{17}$$

Let h be the syndrome vector (20). Since $d = 3$, C is the corrected $t = 1$ error. So, the public-key is

$$H' = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, h = (20) \tag{18}$$

and the private-key is

$$(G, H, M) = (\begin{pmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 2 & 1 \end{pmatrix}), \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}). \tag{19}$$

Encryption: Let the message vector be $m = (1000)$ and $h = (20)$. The cryptogram is

$$c = mH' + h = (1000) \cdot \begin{pmatrix} 0 & 2 \\ 2 & 2 \\ 1 & 2 \\ 2 & 0 \end{pmatrix} + (20) = (02) + (20) = (22). \tag{20}$$

Decryption: Calculate

$$c' = cM^{-1} = (22) \cdot \begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix} = (12). \tag{21}$$

Since

$$c = mH' + h \tag{22}$$

and

$$H' = H^T \cdot M, \tag{23}$$

c' is also equal to

$$c' = (mH' + h)M^{-1} = mH'M^{-1} + hM^{-1} = mH^T MM^{-1} + hM^{-1} \tag{24}$$

$$c' = mH^T + hM^{-1}. \tag{25}$$

So,

$$(12) = (m_1 m_2 m_3 m_4) \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} + (20) \cdot \begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix}. \tag{26}$$

$$(12) = (m_1 + m_2 + m_3, m_1 + 2m_2 + m_4) + (01) \tag{27}$$

$$(12) - (01) = (m_1 + m_2 + m_3, m_1 + 2m_2 + m_4) \tag{28}$$

$$(11) = (m_1 + m_2 + m_3, m_1 + 2m_2 + m_4). \tag{29}$$

We get the message $m = (1000)$ by solving the linear system.

Proposition 1. *The size of the plaintext is* $\log_q \binom{n}{t}$.

Proof. The plaintext is an $n - q$ tuple word of weight t. These are the integers between 1 and $\binom{n}{t}$ to the set of words of weight t and length n. Therefore, the size of the plaintext is $\log_q \binom{n}{t}$. □

Proposition 2. *The size of the ciphertext is* $(n - k)$.

Proof. Since the ciphertext is a $(n - k) - q$ tuple word, the proof is clear. □

Corollary 2. *The transmission rate of the new system is*

$$\frac{\log_q \binom{n}{t}}{(n-k)}.$$

Proof. The proportion of the number of information symbols to the number of transmitted symbols gives the transmission rate. So, it is

$$\frac{\log_q \binom{n}{t}}{(n-k)}.$$

□

Proposition 3. *Given a syndrome vector* y *of weight* w*, the number of eligible* h*'s is* $\binom{w}{t}(q-1)^t$.

Proof. It is known that the weight of h is t, and h is non-zero. Thus, the number of non-zero vectors of weight t among the vectors of w is $\binom{w}{t}(q-1)^t$. □

Example 2. *Let C be the extended binary Hamming code of parameters* $[8, 4, 4]$. *Its packing radius is 1. We examine some properties of the public-key cryptosystem based on C. The size of the plaintext is*

$$\log_2 \binom{8}{1} = \log_2 8 = 3.$$

The size of the ciphertext is

$$8 - 4 = 4.$$

The transmission rate is

$$\frac{\log_2 \binom{8}{1}}{(8-4)} = 0,75.$$

4. Security Comments

In this section, we examine the security of the new system. We recommend using a linear $[n, k, 2t + 1]$-code over $\mathbb{F}_q$. The decryption method is based on the bounded distance decoding task. In order to be a secure public-key cryptosystem, the following conditions should be implemented.

- The size of the public-key should be fairly small. In our cryptosystem, this size is $(n - k)$, which is reasonably small.
- The encryption, decryption, and key generation should be effective. It is computationally simple to create the public-key and private-key. Thus, the encryption and decryption algorithms are too efficient.
- It should be impossible to reach the plaintext by an attacker.
- The system should be resistant to all possible attacks. Now, we discuss these attacks for the new system.

4.1. Algebraic Attack

The security of a public-key cryptosystem depends on the security of the private-key. So, the first attack will be factorization H' to find the private-keys G, H, and M. If the code parameters n, k, d are large enough, this attack is impracticable, because it is difficult to recover the factors of H'. This means the security is ensured with the private-key. The security of the new system is also based on decoding in the code H', while H' is not only non-equivalent to the code H in the cryptosystem, but after multiplying H^T by M from the right, the error-correction capability of public-key H' is unknown. Furthermore, the vector h is secret. Thus, the best attack may not carry out the complete decoding.

4.2. Generic Attack

The second attack is to reach m from c without using the private-key. The plaintext is an n-q tuple word of weight t. We require an useful algorithm that matchs the integers between 1 and $\binom{n}{t}$ to the set of words of weight t and length n and vice versa, since the plaintext is a n-q tuple word of weight t. In this case, the attacker will try to repeatedly select n bits at random from an $(n - k)$-bit ciphertext vector and guess m based on the n selected bits, which is impossible. So, our cryptosystem is strong to all possible attacks. At the same time, the described system presents a general access, which is not for the specific cryptosystem.

Moreover, the probability of no error in the constructing of this system is

$$(1 - \frac{t}{n-k})^n.$$

Consider the Goppa code, which has the parameters

$$n = 1024, k = 524, t = 50.$$

In the public-key cryptosystem constructing based on this code, the probability of no error is

$$(1-\frac{50}{500})^{1024}=(0,9)^{1024}.$$

It is a very small number.

5. Comparison with the Other Public-Key Cryptosystems

In this section, we compare our system with the other code-based cryptosystem for an $[n,k,d]$-code C over $\mathbb{F}_q$, where $d \geq 2t+1$. We denote by S, R, T, and K, respectively, the size of plaintext, ciphertext, the transmission rate, and the dimension of the public-key.

The new system is a further development of the McEliece and Niederreiter cryptosystems. McEliece's system is constructed based on binary linear codes, but both Niederreiter's and our new system are constructed based on linear codes over $\mathbb{F}_q$. Especially, we use the bounded distance decoding to construct our system. In the new system, as the public-key is smaller than McEliece's cryptosystem, it is more useful in industry. Moreover, as it is seen in Table 1, the plaintext is a word of small weight, which is one of the coset leaders, and the number of operations involved during the encryption is less than McEliece's cryptosystem. Furthermore, it is seen that the public-keys in our system and Niederreiter's system are equivalent. However, our system is more effective than Niederreiter's cryptosystem, since it is easier to generate the pieces of keys. This condition increases the security. It is impossible to reach the private-key with public-key by an attacker in the new system. In addition, the plaintext cannot be calculated even if the public-key and ciphertext are known by an enemy cryptanalyst. When the transmission rates of systems are compared, it is noticed that the proposed system has the bigger magnitude. That is, the encryption is faster than the others. So, it is more reliable by means of security.

Table 1. Comparison with other schemes.

System	[7]	[8]	This Paper
S	k	n	n
R	k	$n-k$	$n-k$
T	$\frac{k}{n}$	$\frac{\log_2 \binom{n}{t}}{(n-k)}$	$\frac{\log_q \binom{n}{t}}{(n-k)}$
K	k	$n-k$	$n-k$

6. Conclusions

We presented a new public-key cryptosystem based on error-correcting codes in this study. This system refers to the class of cryptosystems based on the bounded distance decoding task. The sizes of the plaintext and ciphertext of the system are calculated. Therefore, the transmission rate is given. The possible attacks are considered. It is determined that the new system stands well when compared with known systems.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: I would like to thank to Patrick Solé for his valuable comments, which helped to improve the content of the paper.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Krasnobayev, V.A.; Yanko, A.S.; Koshman, S.A. A Method for arithmetic comparison of data represented in a residue number system. *Cybern. Syst. Anal.* **2016**, *52*, 145–150. [CrossRef]
2. Menezes, A.J.; van Oorschot, P.C.; Vanstone, S.A. *Handbook of Applied Cryptography*; CRC Press: Boca Raton, FL, USA, 1997.
3. Diffie, W.; Hellman, M.E. New Directions in Cryptography. *IEEE Trans. Inf. Theory* **1976**, *IT-22*, 644–654. [CrossRef]
4. Rivest, R.L.; Shamir, A.; Adleman, L. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM* **1978**, *26*, 96–99. [CrossRef]
5. Markku-Juhani, O. Saarinen, Linearization attacks against syndrome based hashes. In Proceedings of the 8th International Conference on Cryptology in India, Chennai, India, 9–13 December 2007.
6. Merkle, R.C.; Hellman, M.E. Hiding information and signatures in trapdoor kanpsacks. *IEEE Trans. Inform. Theory* **1978**, *24*, 525–530. [CrossRef]
7. McEliece, R.J. *A Public-Key Cryptosystem Based on Algebraic Coding Theory*; DSN Progress Report; Jet Propulsion Labaratory: Pasadena, CA, USA, 1978; pp. 42–44.
8. Niederreiter, H. Knapsack-type cryptosystems and algebraic coding theory. *Probl. Control. Inf. Theory* **1986**, *15*, 19–34.
9. Li, Y.X.; Deng, R.H.; Wang, X.M. On the equivalence of mceeliece's and niederreiter's public-key cryptosystems. *IEEE Trans. Inf. Theory* **1994**, *40*, 271.
10. Li, Z.; Ling, S.; Xing, C.; Yeo, S.L. On the Bounded Distance Decoding Problem for Lattices Constructed and Their Cryptographic Applications. *IEEE Trans. Inf. Theory* **2020**, *66*, 2588–2598. [CrossRef]
11. Berger, T.P.; Loidreau, P. How to mask the structure of codes for a cryptographic use. *Des. Codes Cryptogr.* **2005**, *35*, 63–79. [CrossRef]
12. Berlekamp, E.; McEliece, R.; van Tilborg, H. On the inherent intractability of certain coding problems (corresp.). *IEEE Trans. Inf. Theory* **1978**, *24*, 384–386. [CrossRef]
13. Krouk, E. A New Public-Key Cryptosystem. In Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques, Konstanz, Germany, 11–15 May 1997; pp. 285–286.
14. Sidelnikov, V.M. A public-key cryptosystem based on binary reed-muller codes. *Discret. Math. Appl.* **1994**, *4*, 191–208. [CrossRef]
15. Berger, T.P.; Cayrel, P.-L.; Gaborit, P.; Otmani, A. Reducing key length of the mceliece cryptosystem. In Proceedings of the Second International Conference on Cryptology in Africa, Gammarth, Tunisia, 21–25 June 2009; pp. 77–97.
16. Misoczki, R.; Barreto, P. Compact mceliece keys from goppa codes. In *Selected Areas in Cryptography*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 376–392.
17. Canteaut, A.; Chabaud, F. A new algorithm for finding minimum-weight words in a linear code: Application to McEliece cyptosystem and to narrow-sense BCH codes of length 511. *IEEE Trans. Inf. Theory* **1998**, *44*, 367–378. [CrossRef]
18. Hill, R. *A First Course in Coding Theory*; Oxford University: Oxford, UK, 1986.
19. Véron, P. Code based cryptography and steganography. In Proceedings of the 5th International Conference, CAI 2013, Porquerolles, France, 3–6 September 2013.

entropy

MDPI

Article

Is Bitcoin's Carbon Footprint Persistent? Multifractal Evidence and Policy Implications

Bikramaditya Ghosh [1,*] and Elie Bouri [2,*]

[1] Symbiosis Institute of Business Management (SIBM), Symbiosis International (Deemed University) (SIU), Bengaluru 560017, India
[2] School of Business, Lebanese American University, Beirut P.O. Box 13-5053, Lebanon
* Correspondence: b.ghosh@sibm.edu.in (B.G.); elie.elbouri@lau.edu.lb (E.B.); Tel.: +91-95350-15777 (B.G.)

Abstract: The Bitcoin mining process is energy intensive, which can hamper the much-desired ecological balance. Given that the persistence of high levels of energy consumption of Bitcoin could have permanent policy implications, we examine the presence of long memory in the daily data of the Bitcoin Energy Consumption Index (BECI) (BECI upper bound, BECI lower bound, and BECI average) covering the period 25 February 2017 to 25 January 2022. Employing fractionally integrated GARCH (FIGARCH) and multifractal detrended fluctuation analysis (MFDFA) models to estimate the order of fractional integrating parameter and compute the Hurst exponent, which measures long memory, this study shows that distant series observations are strongly autocorrelated and long memory exists in most cases, although mean-reversion is observed at the first difference of the data series. Such evidence for the profound presence of long memory suggests the suitability of applying permanent policies regarding the use of alternate energy for mining; otherwise, transitory policy would quickly become obsolete. We also suggest the replacement of 'proof-of-work' with 'proof-of-space' or 'proof-of-stake', although with a trade-off (possible security breach) to reduce the carbon footprint, the implementation of direct tax on mining volume, or the mandatory use of carbon credits to restrict the environmental damage.

Keywords: Bitcoin carbon footprint; Bitcoin mining; energy consumption; FIGARCH; MFDFA; long memory; Hurst exponent; permanent policy

Citation: Ghosh, B.; Bouri, E. Is Bitcoin's Carbon Footprint Persistent? Multifractal Evidence and Policy Implications. *Entropy* **2022**, *24*, 647. https://doi.org/10.3390/e24050647

Academic Editors: Stanisław Drożdż, Jarosław Kwapień and Marcin Wątorek

Received: 18 March 2022
Accepted: 1 May 2022
Published: 5 May 2022

1. Introduction

Bitcoin is a celebrated yet controversial digital currency that continues to attract much attention from users, investors, and regulators across the globe. It is a completely decentralised digital currency without a regulator, with transactions recorded in a publicly distributed ledger called a blockchain [1,2]. New transactions are bucketed into 'blocks' and written onto the end of a 'chain' of pre-existing blocks representing old transactions, hence the name 'blockchain'. Despite wide price fluctuations and periods of booms and busts, Bitcoin holds a major volume in the cryptocurrency domain. Notably, new Bitcoin is introduced into circulation via a process called 'mining', through which transactions are validated for a blockchain. Successful miners are rewarded newly minted Bitcoin for synchronising Bitcoin transactions after solving a complex hashing puzzle. In the process of such proof-of-work mining, new Bitcoins are issued at intervals of almost 10 min, and finding a single block of Bitcoin involves approximately 10 hash computations. While less energy-intensive mechanisms of mining, such as proof-of-space or proof-of-stake, have recently emerged to secure transactions on blockchain by enabling computer networks to collaborate, their application cannot guarantee security and raises significant technological issues. This is why proof-of-work remains the most popular mechanism of mining.

Bitcoin miners operate specialised mining devices with increasingly advanced hardware, such as application-specific integrated circuits (ASICs). They generally use multiple

machines to synchronise Bitcoin transactions and optimise their odds of getting the mining reward, which requires repeatedly running and cooling multiple mining machines. Notably, enormous energy resources are wasted in the Bitcoin mining process. In fact, the mining process consumes huge amounts of electricity [3], and the resulting electricity consumption has been measured at 110.53 TWh per year, exceeding the energy consumption of the Netherlands, inducing a carbon footprint of 36.95 megatons of CO_2 per year, comparable to that of New Zealand. Rising Bitcoin prices make mining very lucrative and attractive, which leads to more electricity consumption [1,4] and greater carbon footprints.

Previous studies show that electricity consumption has a direct and positive relationship with CO_2 emissions [5], and a study of Chinese and Russian electricity markets finds that Bitcoin price volatility is positively correlated with the utility market pricing volatility [6]. The public transaction record (blockchain) is also very energy intensive. However, proof-of-work consumes significantly more energy than proof-of-stake. The number of miners may decline over time or move to more energy-efficient machines [7].

In this paper, we contribute to the above debate on the carbon footprint of Bitcoin mining by examining the long memory process in Bitcoin electricity consumption that reflects the CO_2 emissions of the mining process. Specifically, we use daily data on the Bitcoin Energy Consumption Index (BECI) over the period 25 February 2017 to 25 January 2022 and apply fractionally integrated GARCH (FIGARCH) models and multifractal detrended fluctuation analysis (MFDFA).

Accordingly, we extend the above literature, which remains silent on whether Bitcoin's carbon footprint, measured by energy consumption, exhibits a long memory process. Interestingly, the true long memory process has many facets, which make its application to Bitcoin electricity consumption very informative. Its implications matter to the choice of the most suitable policies that should be applied to address the carbon footprint of Bitcoin mining. For linking stationary long memory and the types of policy (transitory versus permanent), Belbute & Pereira [6,8] argue that if emissions are stationary, then transitory policies (i.e., promotion of energy efficiency, switching from fossil fuel to green energy etc.) will have only transitory effects and fade away in the long-term. Conversely, if emissions are non-stationary, then transitory policies will have a lasting permanent effect [6,8]. Concerns over the energy consumption of Bitcoin mining are indicated by McCook [9], who includes mining-rig procurement and cooling calculations, and argues that Bitcoin is less harmful to the environment than gold mining. Bitcoin's carbon footprint is comparable to that of Ireland [10], and Mora et al. [11] confirm that the estimated CO_2 emissions from Bitcoin could make the globe warmer by 2 °C. Howson [12] expresses concern about the carbon footprint of Bitcoin, while Krause and Tolaymat [10] show that the mining of 4 cryptocurrencies (Bitcoin, Ethereum, Litecoin, and Monero) generated 3–15 million tonnes of CO_2 emissions over the period 1 January 2016 to 30 June 2018. Sedlmeir [13] points to the huge energy consumption of blockchains, especially on the basis of the number of transactions they operate.

In fact, a strong statistical dependence of a mean-reverting time series indicates long memory, long-range dependence or simply persistence [14–16]. Generally, the dependence becomes weaker with time but not in the presence of long memory. Non-stationary time series also show evidence of persistence, sometimes even more strongly than stationary series. Thus, mean reversion holds the key to true long memory. Past and present values are connected by a fractionally integrating parameter, d, which must be empirically calibrated. A partially long memory exists when $d \neq 0$, since a significant mean reversion happens at first difference. For such cases, permanent policy changes are recommended to address the carbon footprint of Bitcoin mining. However, for pure long memory ($d = 0$), the effects generated by transitory policy shocks persist for a long time, and thereby the type of long memory indicates the preferred type of policy that should be adopted by regulators and policymakers to address the carbon footprint of Bitcoin mining.

Following this introduction, Section 2 describes the data and methodology. Section 3 presents and discusses the results. Section 4 concludes and offers policy implications.

2. Data and Methodology

2.1. Data

We used daily data from the Bitcoin Energy Consumption Index (BECI) over the period 25 February 2017 to 25 January 2022, according to the data availability from Digiconomist. BECI data have been recently used in academia [12,17]. They consist of daily data covering three series, BECI upper bound (BECI UB), BECI lower bound (BECI LB), and BECI average. Accordingly, in this work, 1825 daily observations were used for each of the three series. The plots of the three series are shown in Figure 1, in which an increase in the three series is observed from around the second quarter of 2021, which coincides with the spike in the price of Bitcoin.

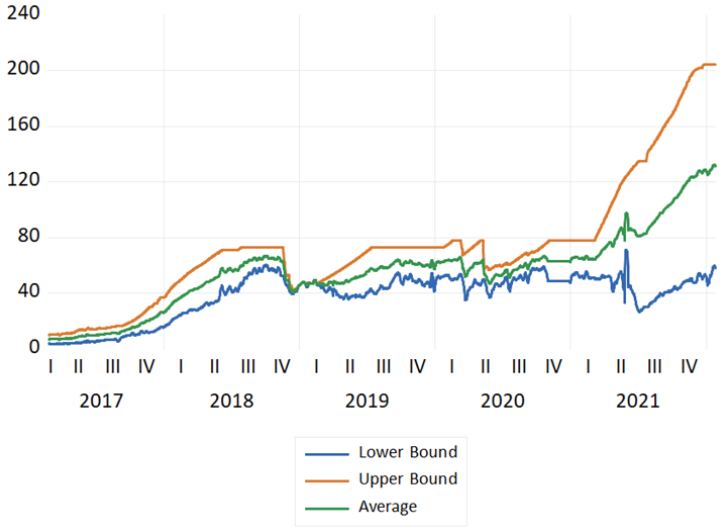


Figure 1. Plots of BECI UB, LB, and average during the study period (25 February 2017 to 25 January 2022).

Considering both BECI upper bound (BECI UB) and BECI lower bound (BECI LB) for electricity consumptions, we argue the following. The BECI UB is defined as the break-even point of mining revenues and electricity costs; therefore, it is more sensitive to the economic parameters. In contrast, the BECI LB is a state where all miners use the most efficient hardware, which makes it more stable and reliable for our current study [13,14] examining the long memory traits of the energy consumption series. Therefore, our analysis emphasises the lower bound results.

Our analysis involved the application of 17 windows to the three series, with the length of each sliding window being 200 days. Given that each series consists of 3400 (17*200) daily observations, a total of 51(17*3) windows (with 10,200 total observations, 3400*3) were considered in our calculation for the sample period 25 February 2017 to 25 January 2022. The choice of sliding window-based estimation procedure is backed by the academic literature [18], which points to the suitability of the application of an increasing window size in a dynamic model for estimating long memory. A sliding window approach to modelling is suitable and an increasing estimation window leads to an increase in the estimation accuracy when calibrating the long memory [19].

Each series observation in the BECI indices is expressed in terawatt-hours (TWh), the standard unit of electricity consumption. The three BECI series are significantly mean-reverting at first difference. Therefore, all the calculations and analyses were conducted at first difference (Δ). Long-range dependence or long memory was characterised by a slow, power law decay of the autocorrelation function (ACF).

2.2. Methodology

Methodologically, we used a pluralistic approach combining two different methods to calibrate 'd' and 'H'. The FIGARCH is an extension of the famed GARCH family as described by Baillie (1996), which is consistent with EMH [20]. The MFDFA is an extension of detrended fluctuation analysis (DFA) as described by Kantelhardt (2008), which is consistent with FMH [21,22].

Firstly, we employed a FIGARCH model to uncover evidence of long memory in BECI indices. FIGARCH considers conditional heteroscedasticity and is comparable to ARCH, but allows for long memory in the conditional variance. It is preferred over autoregressive fractionally integrated moving average (ARFIMA) models because it can detect mean-reverting long memory. Usually, financial time series have d = 1 (fractional integrating parameter), which is consistent for log closing prices of various tradeable securities. Furthermore, it is perfectly in harmony with the efficient market hypothesis (EMH), which concludes that closing levels are martingales, and log returns are martingale differences (usually first difference). Martingales are sequences of random variables with the future expectation equaling the present value. Squared returns typically carry a fractional value of *d*.

Consider a time series, such as the first level difference of each of the three BECI indices:

$$\Delta \text{BECI}_t = \mu + \varepsilon_t \quad \text{with } \varepsilon_t = \nu_t \sigma_t^2 \tag{1}$$

where ν_t is a serially uncorrelated process with zero-mean and unit variance; σ_t is a time-varying measurable function with respect to the information set available at time t−1 (ψ− (t−1)); and σ_t^2 is the time dependent conditional variance of ΔBECI_t. The FIGARCH model of Baillie and his co-researchers [20] is given by:

$$(1 - \beta_1 L)\sigma_t^2 = \omega_0 + [1 - \beta_1 L - \alpha_1 L(1 - L)^d]\varepsilon_t^2 \tag{2}$$

where, $0 \leq d \leq 1$ is the fractional differential (long memory) parameter; L is the lag operator; β(L) is a finite order lag polynomial with the roots assumed to be situated outside the unit circle; and $\propto_k$ represents the autoregressive coefficient of an ARFIMA (1, *d*, 0) model. Unlike ordinary ARCH and GARCH, the FIGARCH model does not reach a constant level quickly. It is reduced to a standard GARCH when d = 0 and to an integrated GARCH (IGARCH) when *d* = 1.

Secondly, we used multifractal detrended fluctuation analysis (MF-DFA) to find h(q) value, where 'h' is the Hurst Exponent and 'q' is the order [23,24]. To this end, we relied on Espen Ihlen's algorithm in MATLAB 13 [25]. It involves a five-step process, as follows:

i. Determining the profile

$$Y(i) = \sum_{k=1}^{i} [x_k - \langle x \rangle] \tag{3}$$

where, x_k is the series, and mean subtraction occurs. Further, $i = 1, 2, \ldots, N$

ii. Dividing the profile: To divide the profile $Y(i)$ N numbers of non-overlapping series of the same length 's'. Since N may not be a multiple of the time scale 's', $2N_s$ was considered.

iii. Calculation of the local trend: Local trend finding for each $2N_s$ segments are carried out by a least-square fit procedure & finding the variance in this process.

$$F^2(S, v) \equiv 1/\text{s} \sum_{i=1}^{s} \{Y[(v-1)s + i] - y_v(i)\}^2 \tag{4}$$

where $y_v(i)$ is the curve fitting polynomial is segment v.

iv. Averaging across all segments to find qth order fluctuation function:

$$F_q(S) \equiv \{\frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(s, v)]^{q/2}\}^{1/q} \tag{5}$$

where q can be any real number, but not zero. It is interesting to note that $q = 2$ coincides with the standard DFA process. Research suggests that extremely large q values (-10 or $+10$) increase the error in the multifractal spectrum tails [26]; therefore, $q = 5$ was used to calibrate such series, which is recommended by another research work [27].

v. Determination of the scaling property of the fluctuation function:

$$F_q\ (\mathrm{s}) \sim s^{H(q)} \tag{6}$$

where $H(q)$ represents the generalised Hurst exponent of the underlying series.

To understand this process intuitively, we referred to Mandelbrot's research, according to which, scaling exponents are unique in nature and depend upon time. Hence, monofractal is not a full proof. It depicts an incorrect narrative. Stochastic time series, such as Bitcoin energy consumption (BECI), have multiple dimensions that add further complexity. For this reason, multifractal is preferred over monofractal. Asset returns tend to deviate from the normal distribution. Moreover, they tend to obey Lévy stable condition. In other words, α ranges from 0 to 2, where $\alpha = 2$ satisfies the condition for Gaussian distribution. Thus, [28] reformulate the Rescaled Range Analysis (R/S) approach proposed by Hurst in 1951. The Hurst exponent expresses $H = 1/\alpha$; when $\alpha = 2$, it becomes stochastic; i.e., it follows a Brownian motion. The legacy of fractals was investigated by a group of researchers [29] who constructed the mathematical formulae to measure the impact of multifractality in a noisy time series. Time series with consistent noises can be transformed into 'random walk' series by subtracting the mean value [25]. Ihlen (2012) integrates it further. According to his process, calculation for the root mean square variation (RMS) is crucial. RMS values are calculated for the localised areas with clear patterns or trends. Finally, all these RMS values are summarised. These RMS samples usually exhibit 'power law' characteristics. In technical terms, this process is known as detrended fluctuation analysis (DFA). The exponent for this relation is the Hurst exponent [29]. Kantelhardt (2002) formulated MFDFA formally for calibration. Ihlen (2012) extended this calculation to the qth-order, suggesting the multifractal detrended fluctuation analysis (MFDFA). Multifractal power law has more than one exponent. Further, Power law relationship and persistent pattern in most cases are two important facets of time series [30,31], such as BECI. The Hurst exponent and fractal dimensions do change from monofractal to multifractal, with the latter being more reliable [32].

Since our data points were 200 for each sliding panel, we altered the segments and scale in the MATLAB code proposed by Ihlen (2012). We took segment = 200 and scale = 4. In the first loop, samples 0–200 (Window 1) were taken. In the second loop, samples 100–300 (Window 2) were considered; the third loop considered 200–400 (Window 3), etc. A polynomial trend fit in each loop was conducted. Quadratic and cubic polynomials were used in this code. We obtained values for the fifth order ($q = 5$) Hurst exponent, and considered it for interpretations as suggested by Kantelhardt. It has recently been proved that a window size of 288, with four sub-windows having 72 observations each, works well through MFDFA [33]; therefore, our window size qualified for a robust calibration.

3. Results

Looking at Table 1, the first differences in the three BECI indices are non-normally distributed, as evidenced by the Jarque–Bera statistics, with low kurtosis points to low volatility and high persistence, which can be confirmed by both the fractional integrating parameter and the Hurst exponent (H). The results of the augmented Dickey–Fuller (ADF) test show that a significant level of stationarity is achieved at first difference.

Table 2 gives the ranges of d and H and their interpretation, indicating the difference between intermediate memory tending towards short and long memory.

Table 1. Summary statistics of the first difference of BECI indices.

	Mean	Max.	Min.	Std. Dev.	Kurtosis	Jarque-Bera	ADF Test
BECI-LB	0.0261	0.818	−0.801	0.200	6.871	124.51	−16.451 *
BECI-UB	0.0349	0.5128	−0.587	0.139	5.892	72.61	−13.017 *
BECI Average	0.0305	0.5293	−0.463	0.131	6.098	86.18	−12.071 *

Notes: The sample period is 25th February 2017 to 25th January 2022. BECI upper bound (BECI UB), BECI lower bound (BECI LB). * Indicates statistical significance at the 1% level.

Table 2. Ranges of d and H and their interpretations.

Ranges of 'd'	Ranges of 'H'	Interpretation
−0.5 < d < 0	0 < H < 0.5	Intermediate memory tending towards short memory
0 < d < 0.5	0.5 < H < 1	Long memory, autoregression decays

Notes: d is the fractional differential (long memory) parameter. H stands for Hurst exponent, which measures the extent of long memory in time series.

Table 3 shows that all the estimated d parameters are below the 0.5 level. This shows that there are no cases of non-stationarity ($d = 1$). At the same time, there is no purely stationary case ($d = 0$). All cases are mean-reverting ($d < 1$), but they have different degrees of decaying autocorrelations.

Table 3. d and H values for BECI UB, LB and Average- FIGARCH method.

Window Number	Sliding Observations	BECI UB d	BECI UB H	BECI LB d	BECI LB H	BECI Average d	BECI Average H
1	0–200	0.36	0.86	0.50	1.00	0.43	0.93
2	100–300	0.45	0.95	0.41	0.91	0.43	0.93
3	200–400	0.35	0.85	0.30	0.80	0.33	0.83
4	300–500	0.48	0.98	0.41	0.91	0.45	0.95
5	400–600	0.44	0.94	0.41	0.91	0.42	0.92
6	500–700	0.44	0.94	0.31	0.81	0.37	0.87
7	600–800	0.42	0.92	0.29	0.79	0.36	0.86
8	700–900	0.45	0.95	−0.05	0.45	0.20	0.70
9	800–1000	0.41	0.91	0.13	0.63	0.27	0.77
10	900–1100	0.46	0.96	0.43	0.93	0.44	0.94
11	1000–1200	0.29	0.79	0.48	0.98	0.39	0.89
12	1100–1300	0.50	1.00	0.38	0.88	0.44	0.94
13	1200–1400	0.33	0.83	0.47	0.97	0.40	0.90
14	1300–1500	0.43	0.93	0.40	0.90	0.42	0.92
15	1400–1600	0.41	0.91	0.48	0.98	0.45	0.95
16	1500–1700	0.43	0.93	0.43	0.93	0.43	0.93
17	1600–1800	0.43	0.93	0.37	0.87	0.40	0.90

Note: This table shows evidence of more observations having long memory using FIGARCH, but of various degrees. d is the fractional differential (long memory) parameter. H stands for Hurst exponent, which measures the extent of long memory in time series. BECI (Bitcoin Energy Consumption Index), BECI UB (BECI upper bound), BECI LB (BECI lower bound), and BECI Average. The sample period is 25 February 2017 to 25 January 2022.

3.1. Results from the MFDFA

The results from the MFDFA investigation revealed evidence of mean-reverting long memory process across BECI UB, LB, and average (see Table 4). Further, no pure short memory was detected ($d = 0$). The effect of long memory was found to be consistent across most sliding windows, barring one (window 10). The persistent pattern was observed for 17/17 (100%) cases in BECI UB and 16/17 (94%) cases in BECI LB. Most importantly, all these cases recorded a Hurst exponent significantly larger than 0.5, indicating a consistently higher intensity of long memory. Extremely strong long memory (H > 0.85) was found in more than 90% cases through this method. Multifractal or singularity spectrum was used to describe the fractal dimension (d) having the same HÖlder exponent. Here, all of the empirical values of 'd' ranged between (−0.07 to +0.50). Both Figures 2 and 3 (chosen randomly out of 51 Windows in consideration), exhibit similar multifractal spectrums, peak around Hq = 1.1 to 1.3. This is more consistent with FMH rather than EMH, which would suggest a peak around Hq = 0.5. The spectrum in each case ranges from 0.6 to 0.8 on the lower side and from 1.6 to 1.7 on the higher side. Small fluctuations are clearly persistent from these two figures. The lower limit of the spectrums in most cases (barring one) ended around Hq > 0.5, indicating the presence of long memory throughout. Another interesting observation is related to the width of the multifractal spectrum. In all 51 windows the width (multifractal) was around 1, exhibiting clear asymmetry (more for BECI LB). This indicates that the degree of complexity in BECI is quite high. This finding remains consistent with existing research [34]. However, future research may further investigate the differences between the left side (small fluctuations from individual cascades reaching white noise faster) and the right side (contracts far sooner to monofractal) [35].

Table 4. d and H values for BECI UB, LB and Average- MFDFA method.

Window Number	Sliding Observations	BECI UB d	BECI UB H	BECI LB d	BECI LB H	BECI Average d	BECI Average H
1	0–200	0.40	0.90	0.27	0.77	0.34	0.84
2	100–300	0.40	0.90	0.35	0.85	0.38	0.88
3	200–400	0.48	0.98	0.45	0.95	0.47	0.97
4	300–500	0.45	0.95	0.34	0.84	0.40	0.90
5	400–600	0.35	0.85	0.39	0.89	0.37	0.87
6	500–700	0.34	0.84	0.44	0.94	0.39	0.89
7	600–800	0.32	0.82	0.09	0.59	0.21	0.71
8	700–900	0.49	0.99	0.45	0.95	0.47	0.97
9	800–1000	0.48	0.98	0.41	0.91	0.45	0.95
10	900–1100	0.35	0.85	−0.07	0.43	0.14	0.64
11	1000–1200	0.40	0.90	0.12	0.62	0.26	0.76
12	1100–1300	0.21	0.71	0.09	0.59	0.15	0.65
13	1200–1400	0.47	0.97	0.27	0.77	0.37	0.87
14	1300–1500	0.32	0.82	0.2	0.7	0.26	0.76
15	1400–1600	0.50	1.00	0.02	0.52	0.26	0.76
16	1500–1700	0.39	0.89	0.26	0.76	0.33	0.83
17	1600–1800	0.47	0.97	0.27	0.77	0.37	0.87

Note: This table shows evidence of more observations having long memory using MFDFA (q = 5th order), but of various degrees. d is the fractional differential (long memory) parameter. H stands for Hurst exponent, which measures the extent of long memory in time series. BECI (Bitcoin Energy Consumption Index), BECI UB (BECI upper bound), BECI LB (BECI lower bound), and BECI Average. The sample period is 25 February 2017 to 25 January 2022.

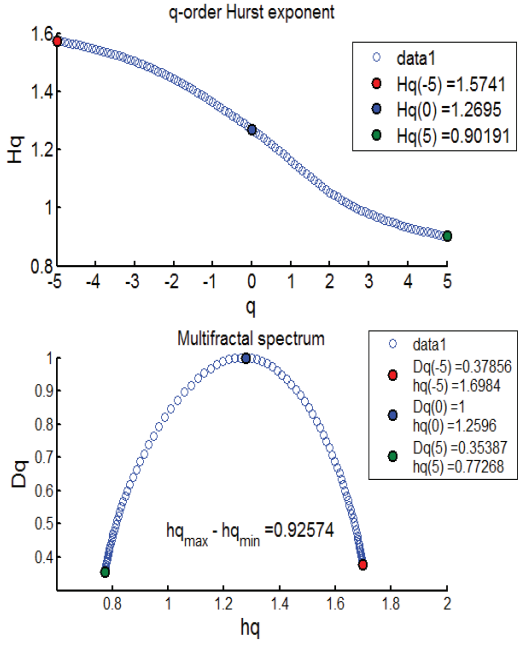


Figure 2. BECI UB window 1 exhibiting the Multifractal Spectrum.

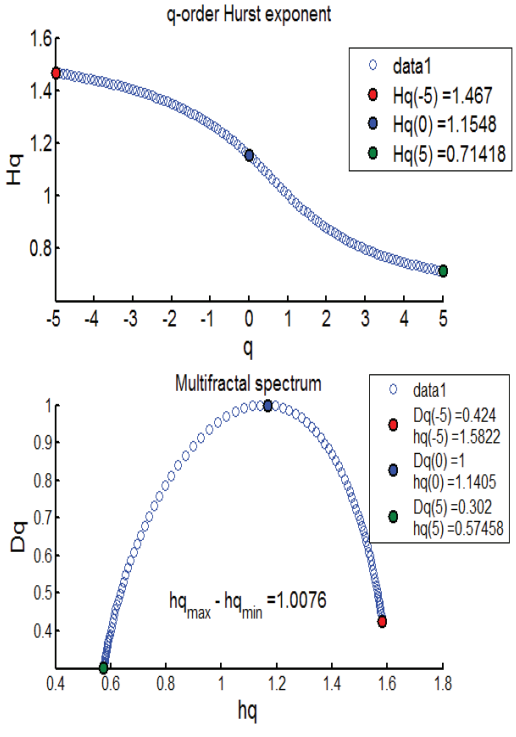


Figure 3. BECI LB window 14 exhibiting the Multifractal Spectrum.

3.2. Results from the FIGARCH

Both BECI UB d and BECI LB d indicated a completely stationary process with long memory (see Table 3). The BECI average likewise exhibits a long memory process. However, no cases of short memory ($d = 0$) were detected. Logically, the effect of long memory changes with the sliding window. The results are consistent in terms of the Hurst exponent (except for one instance, where BECI LB showed anti-persistence). BECI UB displayed a persistent pattern in 17/17 cases, whereas BECI LB showed persistent values in 16/17 cases. The BECI average showed a clear persistent pattern in 17/17 cases with the H value being higher than 0.5. A notable 88% of the BECI UB indicated extremely long memory ($H > 0.85$), whereas BECI average stood second, with 82%, followed by BECI LB, with 76%. Given that hardware costs are substantial and electricity prices are not constant (globally), upper-bound-based calibration is sensitive to various cost matrices [13]. Accordingly, the BECI UB cannot be taken too seriously when making policy decisions. The BECI LB can be considered a better proxy of electricity consumption, and thus suitable for measuring the carbon footprint of the Bitcoin market.

3.3. Overall Results Analysis

It can be concluded that Bitcoin energy consumption (represented by the BECI LB) exhibits strong traces of long memory, with 90% of the feature (i.e., Hurst exponent) scoring over 0.85 with the same argument. Both the fractional integration parameter, d, and the measurement of long memory, H, provide enough evidence of true long memory with mean-reverting traits in the Bitcoin carbon footprint (as represented by BECI LB). Hence, Bitcoin's electricity consumption-led carbon footprint has an overall persistent pattern, with varying degrees across different window sizes. These findings complement the existing literature [36–39], and provide evidence for the suitability of applying permanent policy implications to address the carbon footprint of Bitcoin mining. The MFDFA findings were consistent with FMH and were found to be more realistic; they uncover a higher degree of long memory over FIGARCH (the traditional model in accordance with EMH).

4. Conclusions and Policy Implications

In this paper, we examine the long memory process of the three series of the Bitcoin Energy Consumption Index (BECI) from a pluralistic viewpoint (using FIGARCH and MFDFA) to make inferences regarding the carbon footprint of Bitcoin mining and the possible long-lasting impact of various energy efficiency programs.

Using daily data from 25 February 2017 to 25 January 2022 and FIGARCH model, the results show evidence for the presence of long memory in most series, although with varying degrees. Since the three BECI indices are stationary at first difference ($d < 1$ across all windows), transitory policy changes to reduce the carbon footprint cannot be sustained in the long-term, given that intuitively random policy shocks do not sustain for a longer period in any mean-reverting time series (stationary). Conversely, a permanent policy implementation would have a much longer-lasting effect. Accordingly, permanent policy implementation becomes an evident course of action. Furthermore, finding alternative energy sources, or applying carbon-footprint reduction policies to Bitcoin mining, becomes imperative.

Some policy alternatives emerge in the light of these findings. Specific types of low-energy-consuming hardware for mining Bitcoin using proof-of-stake should replace the energy-consuming proof-of-work scheme by way of a declared policy. Accordingly, selective miners with proof of low energy usage could be allowed to conduct the mining. Proof-of-space, requiring a defined amount of memory, can be put to use by policymakers, which might be far less energy consuming than even proof-of-stake. However, security issues remain for both proof-of-space and proof-of-stake, indicating a clear trade-off. Although some researchers argue that private or consortium blockchains could be used instead of a public blockchain, it is worth noting that if the major participants reach such an agreement, private or consortium blockchain transactions would become perfectly

editable (if they use proof-of-stake or proof-of-space). This presents a major drawback from a security perspective (especially for financial and other sensitive cases) but would make the blockchain system more energy efficient. Permanent environmental policies are crucial in this regard since they would be more effective due to the persistent pattern of the carbon footprint. Accordingly, green energy could be made mandatory for Bitcoin mining. Alternatively, direct tax could be levied on mining volume to further restrict or contain it.

As Bitcoin is contributing to the production of an environmental crisis, future research could consider whether the entire cryptocurrency universe, including major cryptocurrencies other than Bitcoin, is also responsible for high energy consumption and a large carbon footprint. Carbon credits could be made mandatory (they are voluntary as of now) for cryptocurrency miners, forcing responsible mining. Intuitively, efficient and relatively mature markets have seasoned underlying energy efficiency, which propels them to stay afloat in the long term. Currently, most cryptocurrencies suffer from excess energy consumption. Potentially, an alternate technology other than blockchain may be needed to take Bitcoin to a truly mature stage, given that blockchain seems to suffer from extreme consumption of energy, inability to correct its protocol midway (thus abandoning the entire chain), and relatively low performance (5 transactions/second compared to 1700 transactions/second by Visa) [40]. Moreover, the movement of Bitcoin too is not consistent with the EMH [41], making it even more difficult to predict. Hopefully, the tight competition for Bitcoin mining rewards will spur a technological evolution and thereby a genuine solution to the energy-intensive mining process of Bitcoin.

5. Limitations & Future Scope of Study

We found some limitations and future scope of study during our investigation. First, there is a need to modify Ihlen (2012) code, in order for Fq to remain well within the segment. Second, conclusive evidence behind the selection of qth order (2 or 5) must be sought. Third, the asymmetry coefficient could be calibrated separately for both sides, as they have different meanings.

Author Contributions: Conceptualisation, B.G. and E.B.; methodology, B.G.; software, B.G.; validation, B.G. and E.B.; formal analysis, B.G.; investigation, B.G.; data curation, B.G.; writing—original draft preparation, B.G. and E.B.; writing—review and editing, E.B.; visualisation, E.B.; supervision, E.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available on request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bouri, E.; Jalkh, N.; Molnár, P.; Roubaud, D. Bitcoin for energy commodities before and after the December 2013 crash: Diversifier, hedge or safe heaven? *Appl. Econ.* **2017**, *49*, 5063–5073. [CrossRef]
2. O'Dwyer, K.J.; Malone, D. Bitcoin mining and its energy footprint. *IET Conf. Publ.* **2014**, *2014*, 280–285.
3. Corbet, S.; Lucey, B.M.; Yarovaya, L. The Financial Market Effects of Cryptocurrency Energy Usage. 2019. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3412194 (accessed on 17 March 2022).
4. Howson, P. Tackling climate change with blockchain. *Nat. Clim. Chang.* **2019**, *9*, 644–645. [CrossRef]
5. Karathanasopoulos, A.; Dunis, C.; Khalil, S. Modelling, forecasting and trading with a new sliding window approach: The crack spread example. *Quant. Financ.* **2016**, *16*, 1875–1886. [CrossRef]
6. Belbute, J.M.; Pereira, A.M. Do Global CO2 Emissions from Fuel Consumption Exhibit Long Memory? A Fractional Integration Analysis (Issue 165). 2015. Available online: https://economics.wm.edu/wp/cwm_wp165.pdf (accessed on 11 January 2022).
7. Vranken, H. Sustainability of bitcoin and blockchains. *Curr. Opin. Environ. Sustain.* **2017**, *28*, 1–9. [CrossRef]
8. Belbute, J.M.; Pereira, A.M. Do global CO_2 emissions from fossil-fuel consumption exhibit long memory? A fractional-integration analysis. *Appl. Econ.* **2017**, *49*, 4055–4070. [CrossRef]

9. McCook, H. The Cost & Sustainability of Bitcoin. 2018. Available online: https://cryptowords.github.io/the-cost-and-stability-of-bitcoin (accessed on 10 January 2022).
10. Krause, M.J.; Tolaymat, T. Quantification of energy and carbon costs for mining cryptocurrencies. *Nat. Sustain.* **2018**, *1*, 711–718. [CrossRef]
11. Mora, C.; Rollins, R.L.; Taladay, K.; Kantar, M.B.; Chock, M.K.; Shimada, M.; Franklin, E.C. Bitcoin emissions alone could push global warming above 2 °C. *Nat. Clim. Chang.* **2018**, *8*, 931–933. [CrossRef]
12. Howson, P.; de Vries, A. Preying on the poor? Opportunities and challenges for tackling the social and environmental threats of cryptocurrencies for vulnerable and low-income communities. *Energy Res. Soc. Sci.* **2022**, *84*, 102394. [CrossRef]
13. Sedlmeir, J.; Buhl, H.U.; Fridgen, G.; Keller, R. The Energy Consumption of Blockchain Technology: Beyond Myth. *Bus. Inf. Syst. Eng.* **2020**, *62*, 599–608. [CrossRef]
14. Stoll, C.; Klaaßen, L.; Gallersdörfer, U. The Carbon Footprint of Bitcoin. *Joule* **2019**, *3*, 1647–1661. [CrossRef]
15. GHG PROTOCOL. GHG Protocol Scope 2 Guidance: An amendment to the GHG Protocol Corporate Standard. In GHG Protocol Scope 2 Guidance. 2015. Available online: https://ghgprotocol.org/sites/default/files/standards/Scope%202%20Guidance_Final_Sept26.pdf (accessed on 11 January 2022).
16. WBCSD; WRI. A Corporate Accounting and Reporting Standard. In *Greenhouse Gas Protocol*; World Resources Institute: Washington, DC, USA, 2012.
17. De Vries, A. Bitcoin's energy consumption is underestimated: A market dynamics approach. *Energy Res. Soc. Sci.* **2020**, *70*, 101721. [CrossRef]
18. Hanapi, A.L.M.; Othman, M.; Sokkalingam, R.; Sakidin, H. Developed A Hybrid Sliding Window and GARCH Model for Forecasting of Crude Palm Oil Prices in Malaysia. *J. Phys. Conf. Ser.* **2018**, *1123*, 1–8.
19. Vera-Valdés, J.E. On long memory origins and forecast horizons. *J. Forecast.* **2020**, *39*, 811–826. [CrossRef]
20. Baillie, R.; Bollerslev, T.; Mikkelsen, H.O. Fractionally integrated generalized autoregressive conditional heteroscedasticity. *J. Econom.* **1996**, *74*, 3–30. [CrossRef]
21. Kantelhardt, J.W. Fractal and Multifractal Time Series, 1–59. 2008. Available online: http://arxiv.org/abs/0804.0747 (accessed on 10 January 2022).
22. Kantelhardt, J.W.; Zschiegner, S.A.; Koschielny-Bunde, E.S.; Havlin, A.; Bunde, H.E.S. Multi-fractal detrended fluctuation analysis of nonstationary time series. *Physica A* **2002**, *316*, 87–114. [CrossRef]
23. Ramos-Requena, J.P.; Trinidad-Segovia, J.E.; Sánchez-Granero, M.A. Introducing Hurst exponent in pair trading. *Phys. A Stat. Mech. Its Appl.* **2017**, *488*, 39–45. [CrossRef]
24. Morales, R.; Di Matteo, T.; Gramatica, R.; Aste, T. Dynamical generalized Hurst exponent as a tool to monitor unstable periods in financial time series. *Phys. A Stat. Mech. Its Appl.* **2012**, *391*, 3180–3189. [CrossRef]
25. Ihlen, E.A.F. Introduction to multifractal detrended fluctuation analysis in matlab. *Front. Physiol.* **2012**, *3*, 141. [CrossRef]
26. Kaulakys, B.; Alaburda, M. Modeling the inverse cubic distributions by nonlinear stochastic differential equations. In Proceedings of the 21st International Conference on Noise and Fluctuations, Toronto, ON, Canada, 12–16 June 2011; pp. 499–502. [CrossRef]
27. Thompson, J.R.; Wilson, J.R. Multifractal detrended fluctuation analysis: Practical applications to financial time series. *Math. Comput. Simul.* **2016**, *126*, 63–88. [CrossRef]
28. Mandelbrot, B.B.; Fisher, A.; Calvet, L. *A Multifractal Model of Asset Returns*; Working Papers—Yale School of Management's Economics Research Network; Yale University: New Haven, CT, USA, 1997; Volume 1. Available online: https://users.math.yale.edu/~{}bbm3/web_pdfs/Cowles1164.pdf (accessed on 11 January 2022).
29. Watkins, N.W.; Franzke, C. A brief history of long memory: Hurst, Mandelbrot and the road to Road to ARFIMA, 1951–1980. *Entropy* **2017**, *19*, 437. [CrossRef]
30. Bella, G.; Massidda, C.; Mattana, P. The relationship among CO2 emissions, electricity power consumption and GDP in OECD countries. *J. Policy Modeling* **2014**, *36*, 970–985. [CrossRef]
31. Bouri, E.; Gil-Alana, L.A.; Gupta, R.; Roubaud, D. Modelling Long Memory Volatility in the Bitcoin Market: Evidence of Persistence and Structural Breaks. *Int. J. Finance Econ.* **2019**, *24*, 412–426. [CrossRef]
32. Peters, E.E. *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*; John Wiley & Sons: Hoboken, NJ, USA, 1994; Volume 24.
33. Rydin Gorjão, L.; Hassan, G.; Kurths, J.; Witthaut, D. MFDFA: Efficient multifractal detrended fluctuation analysis in python. *Comput. Phys. Commun.* **2022**, *273*, 108254. [CrossRef]
34. Drozdz, S.; Kowalski, R.; Oświęcimka, P.; Rak, R.; Gębarowski, R. Dynamical variety of shapes in financial multifractality. *Complexity* **2018**, *2018*, 7015721. [CrossRef]
35. Drożdż, S.; Oświęcimka, P. Detecting and interpreting distortions in hierarchical organization of complex time series. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **2015**, *91*, 030902(R). [CrossRef]
36. Hayes, A.S. Bitcoin price and its marginal cost of production: Support for a fundamental value. *Appl. Econ. Lett.* **2019**, *26*, 554–560. [CrossRef]
37. Lo, Y.C.; Medda, F. Bitcoin mining: Converting computing power into cash flow. *Appl. Econ. Lett.* **2019**, *26*, 1171–1176. [CrossRef]
38. Wang, L.; Sarker, P.K.; Bouri, E. Short- and Long-Term Interactions between Bitcoin and Economic Variables: Evidence from the US. *Comput Econ* **2022**. [CrossRef]
39. Ghosh, B.; Bouri, E. Long memory and fractality in the universe of volatility indices. *Complexity* **2022**, *22*, 6728432. [CrossRef]

40. Drozdz, S.; Kwapień, J.; Oświecimka, P.; Stanisz, T.; Watorek, M. Complexity in economic and social systems: Cryptocurrency market at around COVID-19. *Entropy* **2020**, *22*, 1043. [CrossRef] [PubMed]
41. Fama, E.F. Efficient capital markets: A review of theory and empirical work. *J. Financ.* **1970**, *25*, 383–417. [CrossRef]

entropy

MDPI

Article

Predicting Bitcoin (BTC) Price in the Context of Economic Theories: A Machine Learning Approach

Sahar Erfanian [1], Yewang Zhou [1,*], Amar Razzaq [1,*], Azhar Abbas [2,*], Asif Ali Safeer [1] and Teng Li [3,4]

[1] Business School, Huanggang Normal University, No. 146 Xinggang 2nd Road, City Development Zone, Huanggang 438000, China
[2] Institute of Agricultural and Resource Economics, University of Agriculture, Faisalabad 38000, Pakistan
[3] Faculty of Behavioral and Social Sciences, University of Groningen, 9712 TG Groningen, The Netherlands
[4] College of Economics and Management, Huazhong Agricultural University, No. 1 Shizishan Street, Hongshan District, Wuhan 430070, China
* Correspondence: zhouyewang@hgnu.edu.cn (Y.Z.); amar.razzaq@hotmail.com (A.R.); azhar.abbas@uaf.edu.pk (A.A.)

Abstract: Bitcoin (BTC)—the first cryptocurrency—is a decentralized network used to make private, anonymous, peer-to-peer transactions worldwide, yet there are numerous issues in its pricing due to its arbitrary nature, thus limiting its use due to skepticism among businesses and households. However, there is a vast scope of machine learning approaches to predict future prices precisely. One of the major problems with previous research on BTC price predictions is that they are primarily empirical research lacking sufficient analytical support to back up the claims. Therefore, this study aims to solve the BTC price prediction problem in the context of both macroeconomic and microeconomic theories by applying new machine learning methods. Previous work, however, shows mixed evidence of the superiority of machine learning over statistical analysis and vice versa, so more research is needed. This paper applies comparative approaches, including ordinary least squares (OLS), Ensemble learning, support vector regression (SVR), and multilayer perceptron (MLP), to investigate whether the macroeconomic, microeconomic, technical, and blockchain indicators based on economic theories predict the BTC price or not. The findings point out that some technical indicators are significant short-run BTC price predictors, thus confirming the validity of technical analysis. Moreover, macroeconomic and blockchain indicators are found to be significant long-term predictors, implying that supply, demand, and cost-based pricing theories are the underlying theories of BTC price prediction. Likewise, SVR is found to be superior to other machine learning and traditional models. This research's innovation is looking at BTC price prediction through theoretical aspects. The overall findings show that SVR is superior to other machine learning models and traditional models. This paper has several contributions. It can contribute to international finance to be used as a reference for setting asset pricing and improved investment decision-making. It also contributes to the economics of BTC price prediction by introducing its theoretical background. Moreover, as the authors still doubt whether machine learning can beat the traditional methods in BTC price prediction, this research contributes to machine learning configuration and helping developers use it as a benchmark.

Keywords: AI; business development; information processing; volatility; precision; financial development

Citation: Erfanian, S.; Zhou, Y.; Razzaq, A.; Abbas, A.; Safeer, A.A.; Li, T. Predicting Bitcoin (BTC) Price in the Context of Economic Theories: A Machine Learning Approach. *Entropy* **2022**, *24*, 1487. https://doi.org/10.3390/e24101487

Academic Editors: Stanisław Drożdż, Marcin Wątorek and Jarosław Kwapień

Received: 24 August 2022
Accepted: 13 October 2022
Published: 18 October 2022

1. Introduction

Cryptocurrency is a private system that enables trades between individuals without a central and intermediate agency. In early 2009, Bitcoin (BTC) was valued for the first time at US$0.08. The currency fluctuated for more than four years until the price touched $1110 in 2013. Due to high volatility and massive fluctuations in prices in cryptocurrencies, accurate price predictions are a complex and challenging task. That is mainly because the costs of

cryptocurrency move unpredictably and chaotically. Machine learning techniques may help bring in some methodology that will lead to better solutions to the problem. In the last several years, there has been an increasing interest in using machine learning techniques in different areas of science [1,2], particularly cryptocurrency price forecasting [3]. For instance, Dutta et al. [4] used macroeconomic indicators, including interest rates, S&P500 market returns, US bond yields, and gold price level as predictive variables for daily BTC prices. The results show that macroeconomic indicators have short-term predictability power. Wang and Vergne [5] investigated macroeconomic indicators, namely supply growth, defined as BTCs in circulation, to see their effect on BTC return. They found that an increase in supply is positively associated with weekly returns. Conrad et al. [6] found that S&P500 volatility has a significantly positive effect on long-term BTC volatility.

Jang and Lee [7] investigated the effect of blockchain information, including average block size, miner revenue, mining difficulty, and hash rate, on BTC prices. Their results proved that the recent volatility in BTC prices stems from the blockchain information indicators. Wang and Vergne [5] investigated blockchain information indicators, including several unique collaborators contributing code to the project, the number of proposals merged in the core codebase, the number of issues raised by the community about the code, and fixed the developer's number of forks on BTC returns. They found a positive and significant relationship between blockchain information variables and weekly returns. Therefore, the first research question arises: (1) What are the significant variables as short-term or long-term BTC price predictors? In addition, much previous research on BTC price predictions with machine learning is conducted either using machine learning techniques or conventional statistical analysis without enough theoretical and analytical support. This study investigates whether the macroeconomic, microeconomic, and blockchain information indicators based on economic theories predict the BTC price. According to these considerations, the second research question is: (2) What are the underlying economic theories of BTC price predictors?

There is not enough available literature on BTC price prediction on Google Scholar compared to stocks: around 400 papers about BTC price prediction problems with machine learning algorithms. There are almost 5500 papers about stock price prediction with machine learning algorithms. Also, according to the existing literature, some research on the BTC price prediction problem shows that machine learning outperforms conventional statistical analysis. At the same time, some still believe that traditional models can predict the BTC price better. For instance, Chen et al. [8] applied machine learning techniques models, including random forest, XGBoost, quadratic discriminant analysis, SVM, and LSTM, and statistical methods, including logistic regression and linear discriminant analysis, to predict high-frequency BTC price. They found that Statistical methods achieve an accuracy of 66%, outperforming more complicated machine learning algorithms for daily BTC price prediction. However, machine learning for BTC's 5-min interval price prediction is superior to statistical methods, with accuracy reaching 67.2%. Pang et al. [9] compared neural network models, sentiment data models, and conventional technical indicators and decision trees to predict BTC prices. The analysis found that the robust neural network models offer better accuracy in predicting BTC prices. Therefore, more research should show whether machine learning algorithms are superior to statistical analysis. Hence, the third research questions are: (3) Are machine learning algorithms superior to traditional methods for BTC price prediction? What machine learning model performs better? What are the best feature selection techniques?

The research innovation herein is looking at BTC price prediction through theoretical aspects. The overall findings show that SVR is superior to other machine learning models and traditional models. This paper has several contributions. It can contribute to international finance to be used as a reference for setting asset pricing and improved investment decision-making. It will be helpful for central bankers, traders, investors, and portfolio managers. Also, it contributes to the economics of BTC price prediction by introducing its theoretical background. Moreover, as the authors still doubt whether machine learning can

beat the traditional methods in BTC price prediction, this research contributes to machine learning configuration and helping developers to use it as a benchmark. The rest of the paper is as follows. In the literature review section, there is an overview of existing work and differences from the current work. After that, the methodologies used in this research are briefly explained. Subsequently, the results and discussion are presented. In the end, the paper is concluded.

2. Literature Review

The interaction between demand and supply, which determines the price, is critical in economic theory. The theory contrasts the supply side, i.e., the number of coins available in the market, with the demand side, i.e., investors willing to buy. It is the investors or the consumers who are considered the key player. It is assumed that trading in BTC is a reseller market. Reselling to generate profit is the most important in the market. The investors who buy the asset, keep it for a while, and then sell it at a later date are the ones who represent the demand side of this market. BTCs are known for their decentralization as the nodes in the markets are anonymous. Miners are rewarded with BTCs instead of their service for making available the computing power. The miners manage the supply side of BTC, and hence they can be terms as the suppliers as per the whitepaper and the blueprint for BTC, the total supply of BTC will be restricted to 21 million. It is ensured that the mining is gradual and not with large influxes.

In addition, Antoniou et al. [10] describe technical analysis as "part of how traders learn about fundamentals." The technical analysis predicts future market behavior by studying past market data, such as volume and price. It is based on the premise that historical data can assist in giving future directions. Similarly, Wang and Vergne [5] found a positive correlation between the volume of BTC trading and returns generated. The stated study results concur, proving that technical analysis affects BTC prices.

2.1. Underlying Theory of the Macroeconomic Indicators: Demand and Supply Theory

The quantity theory of money is a concept in monetary economics that holds that money's supply and demand determine the price level. Using this paradigm, Buchholz et al. [11] highlighted how the forces of supply and demand are the main factors influencing the price of Bitcoin. Additionally, utilizing the Keynesian theory of speculative demand for money framework, NaiFovino, et al. [12] and Ciaian et al. [13] highlighted the association between macrofinancial indicators and Bitcoin prices. According to the hypothesis, people who trade in currencies do so to avoid suffering a capital loss on their investments in bonds and other financial assets. A rise in interest rates lowers the value of economic assets, resulting in a loss on the investment of financial assets [14].

Kristoufek [15] extended the research to study the impact of some macroeconomic indicators on the BTC price prediction. He found that Bitcoin appreciates in the long run if it is used more for trade, i.e., non-exchange transactions.

2.2. Underlying Theory of the Microeconomic Indicators: Microstructure Theory

The theoretical frameworks of the microstructure approach developed by Lyons [16] imply that the market information structure is asymmetric, which means not all market participants know about the market information. Some agents have their private information, not necessarily about fundamentals. Lyons found that order flow is the most critical determinant of exchange rate determination in the short run. According to Lyons [16], order flow can be measured as the number of buyer-initiated orders minus the number of seller-initiated orders in the market. In microeconomics, supply and demand is an economic market price determination model [17,18]. Theory and empirical evidence suggest that, for an asset with a given cash flow, the higher its market liquidity, the lower its expected return (e.g., [19,20]). Market liquidity affects asset prices and expected returns. In the Bitcoin market, the bid–ask spread factor as a proxy for market liquidity. As more and more buy and sell orders are placed, overall supply and demand become more and more

apparent. Some empirical studies also showed the short-term predictability of the Bitcoin microstructure. For example, Dyhrberg et al. [21] investigated the liquidity and transaction costs of Bitcoin markets as a microstructure analysis of Bitcoin. Scaillet et al. [22] showed the bid–ask spread has significant predicting power over jumps in Bitcoin price. In another study, Guo et al. [23] made a short-term prediction of BTC price fluctuations (measured with volatility) using buy and sell orders.

The private information in the BTC market is different from the stock market. In stock market trading, private information is referred to an improved understanding of a firm or company's prospects and provides a better evaluation of a potential cash flow. When a particular group of traders is made accessible to private information, it helps to create a clear-cut distinction between a BTC market and a stock market. However, it is essential to note that, like the stock market, BTC entertains an uninformed group of traders who enter the market for liquidity only. The questions here follow: What if there remains no future cash flow available for discounting or there remains no asset for valuation? In such a scenario, what exactly would private information provide?

It is indicated that the valuation of BTC is strongly dependent on the level of confidence of its traders. Hence, private information announces great estimation and prediction of the value that a BTC can potentially gain. These types of evaluations are dependent on the consumption of BTC and their usages. Private information like this adds to the prices of BTC and stimulates its demand. Since BTC has a fixed supply, private information helps increase the demand, increasing the prices in the global market. Data provided by the order book covers all the causes of demand and supply conditions of an asset in the form of bids and asks, which are implemented as trades ultimately. The data here provide an insight into the market's microstructure and an internal overview, which might not be easy to comprehend otherwise. Bid and ask price are two essential components of private information. The bid price refers to the highest price that a potential buyer of BTC is willing to pay. It is also referred to as the buying price for the exchange. When demand for BTC is high, the bid price increases, which means trading volume affects the bid price.

Ask price is the lowest price a seller wants to accept BTC. If the demand falls, there is a fall in the asking price as well. Ask prices are generally higher in comparison to bid prices. Therefore, the difference between these two prices, called the spread, is precisely the profit extracted in these exchanges. BTC prices are highly volatile, which causes extreme fluctuations along with the spreads, which is why sellers enter this market after a great deal of negotiation with the investors and traders to initiate a bidding war. Once that happens, this buying pressure will force an increase price.

2.3. Underlying Theory of Blockchain Information Indicators: Cost-Based Pricing Theory

According to Noble and Gruca [24], the cost price of any service or product can be computed based on a predefined profit margin percent calculated over the total cost. The primary focus of the cost-based pricing theory focuses on the variable cost and fixed cost components classified as part of the internal cost. This pricing theory is crucial to BTCs miners as it helps them compute from which cost price is the mining activity more profitable. Blockchain information is one of the critical considerations of BTC's cost price, as per Wang and Vergne [5]. The mining hardware efficiency can be improved significantly using the right technology resulting in a reduced cost of mining the BTC and a lower price. The lower cost and lower price will lead to increased demand, resulting in ultimately improved return on the overall investment in BTC. Extra hashing power can be achieved for the global mining network on blockchain information which contradicts the lower cost of mining as the difficulty level increases leading to higher mining costs and higher prices for BTCs, resulting in reduced demand and lesser returns.

By developing a cost-of-production model for valuing Bitcoin, Hayes [25] showed that the three factors of computational power, rate of coin production, and mining difficulty used might account for more than 84% of relative value formation. Increasing the difficulty will result in fewer units produced for a given amount of hash power, increasing the relative

cost of production. Similar to this, reducing the block reward will result in fewer units. The marginal cost of production is reduced with improved mining hardware energy efficiency, drop-in electricity charges globally, or reduced mining difficulty. With improvement in technical processes, the efficiency of the mining process also improves, which leads to a reduction in the cost of production, which in turn puts downward pressure on prices. In another study, Hayes [26] back-tests a marginal cost of production model applied to value Bitcoin. The author applied vector autoregression (VAR) and traditional regression models on the historical data from 29 June 2013, through 27 April 2018, when the mining difficulty changes, i.e., every two weeks. Results demonstrate that the marginal cost of production is important in explaining Bitcoin pricing in the long run (considering every two weeks a long run prediction).

The block size limits the number of transactions verified with each block, resulting in more computation power for verifying larger blocks. This increased need for more computational power will increase the cryptocurrency price in line with what has been discussed. By definition, hash rate means the quantum of processing and computing power that the mining process contributes to the network. The value of hash rate is referred to provide the value of the network power. Thus computed, this value is used to correct the mining difficulty, i.e., to increase or decrease it and thereby correspondingly increase or decrease the BTC price.

The average block time of the network is evaluated after n number of blocks, and if it is higher than the expected block time, then the level of difficulty of the proof of work algorithm is declined. On the contrary, if the average block time is less than expected, the difficulty level will increase, which is in line with the concept of economics called the law of diminishing marginal utility. The speed with which the things are made available, then the value decreases over time. In terms of BTC terminology, the faster the rate of unit formation, the lower the price of the coin goes.

Difficulty is changed based on the time it took to discover 2016 previous blocks. If a block is found every 10 min (finding 2016 blocks will take exactly 2 weeks). The more (or less) time was spent on finding the previous 2016 blocks the more will difficulty be lowered (raised). Because mining is still lucrative despite the difficulties adjusting higher and the margins becoming somewhat slimmer, more miners are encouraged to join. more miners joining the effort means that the network is growing, which is good for Bitcoin's price in the long run. This cycle keeps going until a sizable part of the miners can no longer keep up. Some are compelled to sell a growing proportion of the newly created Bitcoins, which finally depletes their treasuries. This causes an increased supply of Bitcoins for sale on the market. They eventually give up and cease mining. The difficulty is then adjusted downward when the hash rate declines.

2.4. Application of Machine Learning in Real-World Problem Solving

Artificial intelligence (AI) is a relatively new trend in science that wants to bring about fundamental changes in people's lives. AI is a little challenging to define, but it can be said that it combines different sciences to make machines more intelligent. One of the most popular subfields of artificial intelligence is machine learning, which is hotly debated. Everyone feels the impact of the learning machine every day in daily life. Simply machine Learning is a science that teaches machines how to learn new things from themselves. Machine learning is one of the modern human inventions that has contributed to the progress of various industries and businesses and has also been very influential in the individual lives of human beings [27]. Machine learning is a subset of artificial intelligence that focuses on learning from the database to build intelligent computer systems. At present, machine learning has been used in various fields and industries. For example, machine learning has been used to diagnose and treat diseases [28], image processing [29], classification [30], and more. Support vector regression can be used in many areas, such as dynamic response prediction of magnetorheological elastomer base isolator [31], thermal spring back of hot press forming [32], text classification [33], etc.

2.5. Related Work and Research Gap

Thus far, empirical studies do not demonstrate a clear advantage for the emerging techniques of using machine learning algorithms to predict the BTC price. Research in this area is insufficient [34,35]. Therefore, this study will help to show the significance of machine learning methods in BTC price prediction problems. Also, some research shows machine learning outperforms statistical analysis, and some still believe in the superiority of conventional statistical analysis. Table 1 presents some related work on the BTC price prediction problem. The current research differs from previous studies in terms of completeness and comprehensiveness, and the comparative analysis in the current study has not been conducted before. In addition, a variety of indicators, including macroeconomic indicators, microstructure indicators, blockchain information, and technical indicators, have been used to analyze the significant variables as BTC price predictors.

Table 1. Overview of research published on BTC price prediction.

Reference	Year	Methodologies	Data Categorization	Findings
Chen et al. [8]	2020	Logistic Regression and Linear Discriminant Analysis, Random Forest, XGBoost, Quadratic Discriminant Analysis, SVM, and Long Short-term Memory	Blockchain Information, Macroeconomic Indicators	Statistical methods outperform machine learning for BTC daily price prediction, while, Machine learning for BTC's 5-min interval price prediction is superior to statistical methods,
Aggarwal et al. [36]	2020	SVM and decomposition (CEEMD)	technical indicators	The proposed method for short-term, midterm, and long term-prediction has a predictability power
Dutta et al. [4]	2020	Gated Recurring Unit, simple neural network (NN), LSTM	Blockchain Information, Macroeconomic Indicators, Technical Indicators	GRU outperforms the NN and LSTM for daily price prediction
Jiang, X. [37]	2019	MLP, LSTM, Gated Recurrent Network	Technical Indicators	
Munim et al. [38]	2019	ARIMA and neural network autoregression (NNAR)	Technical Indicators	ARIMA outperforms NNAR in daily price prediction
Huang et al. [39]	2019	A tree-based predictive mod buy and-hold strategy	Technical Indicators,	A tree-based predictive model for daily return outperform a buy and-hold strategy
Shen et al. [40]	2019	GARCH model, SMA, RNN	Technical Indicators	RNN method outperforms the GARCH model and SMA model for daily return prediction
Mangla et al. [41]	2019	Logistic regression, SVM, RNN, and ARIMA	Technical Indicators	ARIMA is better for next-day prediction, RNN better for weekly
Siami-Namini and Namin [42]	2018	ARIMA, long short-term memory (LSTM)	Technical Indicators	LSTM is superior to ARIMA for daily prediction
Jang and Lee [7]	2017	Bayesian neural networks (BNNs), SVR, and linear regression	Blockchain Information and macroeconomic indicators	BNN outperforms SVR and linear regression

Table 1. *Cont.*

Reference	Year	Methodologies	Data Categorization	Findings
Pichl and Kaizoji [43]	2017	Multilayer Perceptron	Technical Indicators	HARRVJ neural network captures well the dynamics of daily Realized Volatility as aggregated on the 5-min grid.
Indera et al. [44]	2017	MLP-based NARX	Technical Indicators	NARX has predictive power for daily price
Current Work	2022	OLS, MLP, ENSEMBLE, and SVR	Technical indicators, macroeconomic indicators, microstructure indicators, and blockchain information indicators	SVR beats the other models Macroeconomics and blockchain information have long term predictivity power There is no feature selection to improve the model

In the existing literature, there is no comprehensive work in which almost all categories of indicators are investigated. Most of the works regarding BTC price prediction are empirical analyses. However, the current study first looks at the BTC price prediction problem from the perspective of economic theories, including demand and supply theory, microstructure theory, and Cost-based pricing theory. It then identifies the associated variables affecting the BTC price. After that, we empirically prove the predictability power of the attributes through emerging machine learning models and traditional methods.

3. Materials and Methods

This research applies a traditional OLS method [45] and some machine learning methods for the BTC price prediction problem, including Ensemble learning, SVR, and MLP multilayer perceptron, which are briefly explained.

3.1. Multilayer Perceptron

Rosenblatt [46] introduced a multilayer perceptron (MLP) concept with a single perceptron in 1958, consisting of the input layer, middle layers, and output layer. The input layer is a connection between outer space with the network. The middle layers are called hidden layers. Because there is no connection with the outside world, its values are not observed in the training set. The number of neurons in the input layer corresponds to the number of input parameters. Neurons in the hidden layer can be determined by the "trial and error" method. The output layer includes neurons according to our desired output, e.g., the forecasted value in the forecasting problems. A set of weights connects the neurons (see Figure 1).

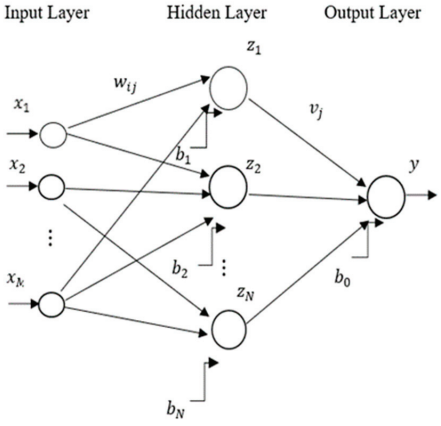


Figure 1. The structure of the three-layer perceptron.

The output value y of a three-layer perceptron can be formulated as:

$$y = \varphi_2(\sum_{j=1}^{N} v_j z_j + b_0) \tag{1}$$

where N is the number of neurons in the hidden layer, v_j is the weight of the second layer, z_j is the output of neuron j, b_0 is the bias of the output neuron and φ_2 is the activation function of the output neuron. Several activation functions have been used in MLP models, such as scaled conjugate gradient (SCG), Levenberg–Marquardt (LM), gradient descent with adaptive learning rate (GDA), gradient descent with momentum (GDM), and others. The output value of neuron j in the hidden layer is given by:

$$z_j = \varphi_1\left(\sum_{i=1}^{M} w_{ij} x_i + b_j\right) j = 1, \ldots, N \tag{2}$$

where M is the number of inputs, w_{ij} are the weights of the first layer, x_i are inputs and b_j is the bias of neuron j, and φ_1 is the activation function of hidden layers. The reason behind choosing MLP is that they are fast to train and can afford hidden layer size 256 instead of 32–64. Also, colossal variance gives a strong ensemble with a single model type.

3.2. Support Vector Regression

Support vector regression (SVR) is an emerging nonlinear regression method based on statistical learning theory having a more stable solution than traditional neural network models. Adopting the structural risk minimization principle in SVM reduces overfitting and local minima issues. In SVR, the nonlinear regression problem is transformed into a linear regression problem by mapping the input data into a high dimensional feature space by applying kernel functions [47]. Consider a set of data $(x_i, y_i)_{i=1}^{m} \subset \mathbb{R}^m \times \mathbb{R}$ where x_i is a vector of inputs, y_i represents the scalar output. In the nonlinear regression case, the linear estimation function can be formulated as $f(x) = \langle w, \phi(x) \rangle + b$ where, $w \in \mathbb{R}^m$ is weight vector, $\phi(x)$ is the mapping function, $\langle \cdot, \cdot \rangle$ denotes the dot product in the feature space, and b is a constant. Several cost functions can be used in SVR, including Huber's Gaussian, ε-insensitive, and Laplacian. The robust ε-insensitive loss function introduced by Vapnik [48] is the most frequently used function, which can be formulated as follows:

$$L_\varepsilon(f(x) - y) = \begin{cases} |f(x) - y| \; if \; |f(x) - y| \; \geq \varepsilon \\ 0 \; otherwise \end{cases} \tag{3}$$

where ε is the tube radius around the regression function $f(x)$, affecting the number of support vectors used to construct the regression function. The cost of errors on the points inside the tube is zero. Figure 2 shows a schematic diagram of the nonlinear regression by SVR.

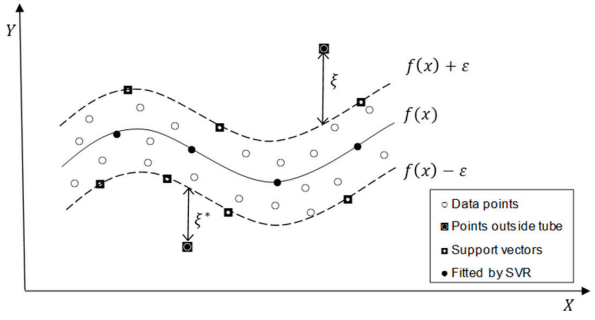


Figure 2. A schematic diagram of the nonlinear regression by SVR based on the ε-insensitive loss function in the feature space.

The SVR performs linear regression in the feature space using the ε-insensitive loss function by minimizing the empirical risk $R_{emp} = \frac{1}{n}\sum_{i=1}^{n} L_\varepsilon(f(x)-y)$ as well as minimizing the regularization term, $\|w\|^2$ to reduce the model complexity (flatness). The slack variables ξ_i and ξ_i^* represents the deviation of training samples out of the ε-insensitive zone. The optimal regression function can be obtained [47]:

$$min\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{k}(\xi_i + \xi_i^*) \tag{4}$$

$$s.t. y_i - \langle w, \phi(x_i)\rangle - b \le \varepsilon + \xi_i \tag{5}$$

$$\langle w, \phi(x_i)\rangle + b - y_i \le \varepsilon + \xi_i^* \tag{6}$$

$$\xi_i, \xi_i^* \ge 0 \tag{7}$$

where C is the regularization constant determining the trade-off between the empirical risk and the regularization term. The above optimization problem can be solved by using Lagrangian multipliers α_i^* and α_i and Karush–Kuhn–Tucker conditions as the following form:

$$max -\varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) + \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)y_i - \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)\, K\langle x_i, x_j\rangle \tag{8}$$

$$s.t. \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) = 0 \tag{9}$$

$$0 \le \alpha_i \le C,\ i = 1, \ldots, n \tag{10}$$

$$0 \le \alpha_i^* \le C,\ i = 1, \ldots, n \tag{11}$$

where $K\langle x_i, x_j\rangle$ is the kernel function which is defined as the inner product of $\phi(x_i)$ and $\phi(x_j)$ in the feature space. After solving the optimization problem, the optimal form of the regression function can be obtained as [47]:

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K\,\langle x, x_i\rangle + b \tag{12}$$

By setting the parameters C and ε and the kernel parameters, the estimation accuracy can be obtained. The reason for choosing SVR is that it is robust to outliers. The decision model can be easily updated. It has excellent generalization capacity with high prediction accuracy, and its implementation is straightforward.

3.3. Ensemble Method

Various experiences show no specific training algorithm in machine learning methods that can be the best and most accurate for all applications. Each algorithm is a particular model based on certain assumptions. Sometimes these assumptions are met, and sometimes they are violated. Therefore, no algorithm alone can succeed in all situations. Ensemble methods have been introduced to solve this problem. The primary motivation for developing the Ensemble method is to reduce the error rate. Forecasting error using the Ensemble approach, a group of techniques is much lower than using a single model. When combining independent and different classifiers, the likelihood of making the right decision is strengthened since each of these classifiers will perform better than a random guess.

Hansen and Salamon [49] presented deploying multiple models on regression. They proved that someone could show that the overall error E decreases uniformly concerning N with the N independent classifier with a probability of error $e < 0.5$. Also, the overall performance is significantly reduced if someone uses dependent categories. The methodology consists of two consecutive steps: The training and testing phases. As shown in Figure 3, several predictive models are produced using training samples in the training phase. Predictive models would combine to predict the next step or the testing phase.

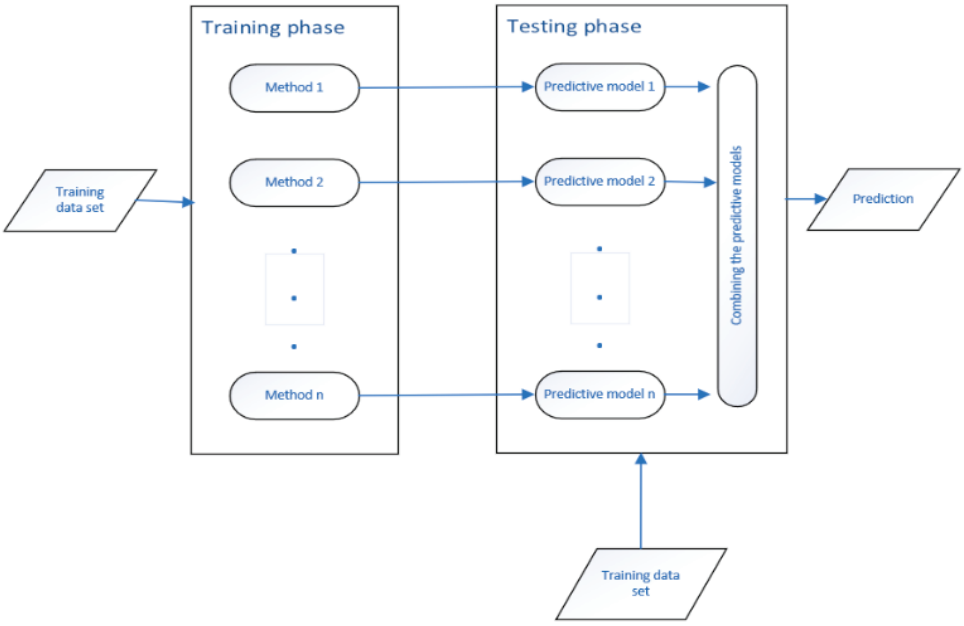


Figure 3. Ensemble method flowchart.

Some popular ensemble methods are Boosting, Bagging, and Blending, of which the Bagging approach will be used in this research. There are two main reasons to choose an Ensemble model: performance and robustness. The Ensemble model can make better forecasts and do better than any single model. An Ensemble model reduces the spread or distribution of the estimates and model accuracy.

3.4. Feature Selection Methods

Feature selection, variable selection, or attribute selection plays an essential role in classification problems. It reduces the number of attributes by excluding the irrelevant and redundant ones to achieve the lower complexity model (see Figure 4). The more uncomplicated and faster models with fewer variables are desirable in machine learning models. Feature selection is an essential part of the machine learning process, leading to overfitting. Overfitting happens when the model learns details and noises made by too many variables, and then the model will not generalize well when presented with new data.

Figure 4. Feature selection in one glance (each color is representing one feature).

In this research, some feature selections, such as principal component analysis (PCA), particle swarm optimization (PSO), evolutionary search, genetic search, best-first search, and variance inflation factor (VIF), are used.

3.5. Model Evaluation

A model evaluation metric quantifies a predictive model's performance, typically involving training a model on a dataset, using the model to make predictions on a "test dataset" not used during training, then comparing the predictions to the expected values in the test dataset. Different authors use different metrics to compare their models. Table 2 shows the evaluation metrics used in this study. In all formulas, y_t $\hat{y}_t$ T is the target value, output value, and the size of a test dataset in out-of-sample or out-of-fold prediction.

Table 2. Common types of evaluation metrics.

Accuracy Metrics	Formula
R^2 [50]	$R^2 = 1 - \sum_{t=1}^{T} (\hat{y}_t - y_t)^2 / \sum_{t=1}^{T} \hat{y}_t{}^2$ T is the size of a test dataset in out of sample prediction
Pearson's r	$r = \sum_{t=1}^{T} \hat{y}_t y_t / \sqrt{\sum_{t=1}^{T} \hat{y}_t{}^2 \sqrt{y_t}}$
Root Mean Square Error (RMSE) [51]	$\mathrm{RMSE} = \sqrt{\sum_{t=1}^{T} (\hat{y}_t - y_t)^2 / T}$

3.6. Model Validation

One of the more used statistical analyses, cross-validation, helps assess and validate the machine learning model's performance. The key intention behind evaluating the model is to see whether or not one can check if the trained model is generalizable. As part of the K-fold cross-validation process, the entire data set is first split into several folds. After that, the model is trained on all folds but one and the test model on the remaining fold. The test is reiterated multiple times until the model tests all the folds. Finally, the average scores obtained in every fold are taken as the final metrics. Predictions are made on the test sets that were not used to train the model during the process. These predictions are called 'out of fold predictions,' a type of 'out of the sample' forecast. In contrast to the simple

train-test, the method discussed prevents overfitting and helps in a more robust model evaluation form.

Cross-validation on a rolling basis is a method that is used for cross-validating the time series models. According to Kuhn and Johnson [52], the value of $k = 10$ is expected. The repeated K-fold cross-validation method replicates the entire process multiple times. For instance, if ten-fold cross-validation were repeated five times, it would result in 50 times out-of-fold predictions, estimating the model's efficacy. The ten times K-fold cross-validation is a prevalent method to Kuhn and Johnson [52]. As depicted in Figure 5, the process starts with a small subset of data for training. Subsequently, the forecast for the later data point finally, the data point is for checking the accuracy. The same forecasted data point is included in the following training data set basis on which the next data points are predicted.

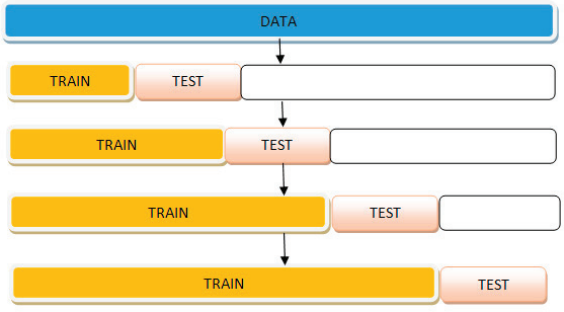


Figure 5. Cross-validation on a rolling basis.

4. Results and Discussion

This section consists of three parts. In the first part, a multilinear regression model is built for the BTC price prediction problem on monthly BTC prices from 18 August 2010 to 17 September 2018. Data includes macroeconomic and blockchain information indicators. The second part presents two comparative approaches: feature-based and category-based comparative analysis consisting of OLS, Ensemble methods, SVR, and MLP for the BTC price prediction problem on a daily data set from 11 October 2016 to 12 June 2017. Data is composed of macroeconomic, microeconomic, and technical indicators. All predictions in this part are out-of-fold predictions.

During the k-fold cross-validation process, predictions are made on test sets comprised of data not used to train the model. These predictions are called out-of-fold predictions, a type of out-of-sample predictions. Another analysis similar to the second part is described in the third part on different BTC datasets, including macroeconomic, microeconomic, blockchain information, and technical indicators from 1 January 2018 to 5 June 2018. For validation of results in this research, three metrics, namely RMSE, R^2, and Pearson r, have been used to compare the out-of-sample and out-of-fold predictive models under the T-test at the significance level of 0.05. The k-fold cross-validation with $k = 10$ (so-called cross-validation on a rolling basis) is used to construct a high-performance model and have robust results. Results are averaged on 100 prediction trials.

4.1. The BTC Price Prediction Problem Using OLS

According to the theoretical analysis regarding demand and supply theory, macroeconomic indicators have long-term predictability power on BTC prices. For the empirical analysis, a multilinear regression model is built for the BTC price prediction problem (model 1 in the Appendix A) on monthly BTC prices from 18 August 2010 to 17 September 2018, including macroeconomic and blockchain information indicators.

4.1.1. Data Description

Monthly BTCUSD transactions occurring on the significant BTC exchanges, available at blockchain.com from 18 August 2010 and ending on 17 September 2018, including 24 variables, have been examined. Dependent variables can be categorized into Macroeconomic indicators and Blockchain information indicators obtained via provided API at blockchain.com (see Table 3). Some descriptive statistics, including minimum, maximum, mean, and standard deviation, have been calculated and shown in Table A1 in the Appendix A.1 to describe or summarize the data.

Table 3. Data categorization.

Indicator Category	Indicator Name
Macroeconomic indicators	Market capitalization, BTCs in circulation, US federal funds rate, S&P500 stock market index, Nasdaq composite, DJIA stock market index, WTI, gold-fixing price, breakeven inflation rate,
Blockchain information indicators	Hash rate, mining difficulty, number of transactions per block, block size, average block size, median confirmation time, orphan blocks, cost per transaction, transaction fees, estimated transaction value (BTC), estimated transaction value (USD), total output value

4.1.2. Feature Selection

First, data cleaning, including estimating outliers (extreme values) and missing values, has been applied to the raw data to build a better data set. After that, VIF is applied to the data set to deal with multicollinearity. Table A2 in the Appendix A.1 shows variables, namely, market capitalization, transactions per block, Hash Rate, mining difficulty, cost per transaction, total transactions per day, Nasdaq Composite, Dow Jones Industrial Average, and S&P 500, which have a VIF greater than 10. Instead of dropping variables, the entire sample period has been tested in nine models with different combinations of variables.

4.1.3. OLS Regression for BTC Price Prediction

Table A3 in the Appendix A.1 shows the results of nine regression models built to avoid multicollinearity. The variables in quotes are the variables with a high correlation. They are added to the rest of the variables to build a new regression model. The response variable in each model is the BTC price. The value in parentheses represents the results of the *t*-test for the null hypothesis-rejecting variables, based on a *p*-value of 0.05. The R^2 from regression models is relatively high, suggesting that, for example, approximately 73% of the variation in BTC prices in model "9" is determined by the variables in the model. Due to the t-statistics and p-value, all models are statistically significant. By looking at the coefficients, which are not tiny, it is evident that all variables are economically significant for the models.

The regression analysis showed that the significant macroeconomic indicators in all models for monthly BTC price are market capitalization, Nasdaq Composite, Dow Jones Industrial Average, and S&P500. Therefore, macroeconomic indicators have long-term predictive power on BTC prices as expected a priori and the t-statistic indicates the significance of the results. Also, blockchain information indicators, including the block size, cost per transaction, mining difficulty, hash rate, transaction fees, and estimated transaction value, verify that the supply and demand theory is the underlying theory of predictors. Therefore, blockchain information indicators have a long-term predictive power on BTC prices as expected a priori. The t-statistic indicates that it is highly statistically significant that blockchain information indicators influence the price confirming that the cost-based pricing theory is underlying the predictors. Empirical results answer the first and second research questions. (1) What are the significant variables as short-term or long-term BTC price predictors? (2) What are the underlying economic theories of BTC price predictors?

4.2. Proposed Comparative Analysis for Dataset 1

According to the theoretical analysis regarding demand and supply theory, macroeconomic indicators do not have short-term predictability power on BTC prices. For the empirical analysis, a comparative machine learning model, including OLS, Ensemble methods, SVR, and MLP for the BTC price prediction problem on data sets from 11 October 2016, to 12 June 2017, including macroeconomic, microeconomic, and technical indicators. Feature selections, namely Best First Search, PSO Search, and Evolutionary Search, are applied to the data. The price prediction model is described in the Appendix A (model 2).

4.2.1. Data Description

Daily BTC/USD transactions occurring on the Bitfinex exchange, obtained via provided API at bitfinex.com (accessed on 2 October 2019) from 11 October 2016, to 12 June 2017, including 22 independent variables, have been examined. Dependent variables can be categorized into three groups; Macroeconomic indicators, obtained at fred.stlouisfed.org, and microeconomic and technical indicators extracted from bitfinex.com. Table 4 shows the specification for each group. Some descriptive statistics, including minimum, maximum, mean, and standard deviation, have been calculated and shown in Table A4 in the Appendix A.2 to describe or summarize the data.

Table 4. Data categorization.

Indicator Category	Indicator Name
Macro-Economic Indicators	Trade-weighted US Dollar Index, gold-fixing price, DJIA Index, Brent Crude oil price, WTI
Microeconomic Indicators	Trades per minute, bid/ask sum, bid–ask spread, buy/sell signals,
Technical Indicators	volume, MTM, CCI, SMA

4.2.2. Feature-Based Comparative Analysis

This section applies the comparative analysis to different datasets containing the indicators chosen by different feature selection techniques, including VIF, genetic search, evolutionary search, and best-first search. Table A5 in the Appendix A.2 shows the different features chosen by various methods. The comparison is conducted under the *T*-test at the significance level of 0.05 by WEKA software (version 3.9.4, developed at the University of Waikato, New Zealand). To evaluate the predictive machine learning models' performance and have robust results, the 10-fold cross-validation on a rolling basis evaluation technique is used, and each model is repeated ten times. Therefore, the average results of 100 prediction trials, including the forecasting ability of models, namely RMSE and Pearson's *r*, are shown in Tables 5 and 6. The standard deviation is shown in parenthesis.

Table 5. RMSE of different models on different data sets.

Model Indicators	OLS	Ensemble Methods (Bagging)	SVR	MLP
All indicators	8.86 (2.36)	9.04 (1.97)	8.68 (2.48)	9.30 (2.20)
PCA Reduction	8.79 (1.98)	11.45 (2.48)	8.59 (2.09)	11.67 (2.31)
VIF	15.97 (3.03)	13.92 (3.00)	16.01 (3.18)	15.28 (4.57)
Genetic Search	8.77 (2.23)	9.45 (2.05)	8.67 (2.27)	10.11 (2.39)
Evolutionary Search	8.72 (1.98)	9.00 (2.06)	8.68 (2.13)	9.56 (2.39)
Best First	8.80 (2.23)	9.40 (2.07)	8.68 (2.26)	10.08 (2.49)

Table 6. Pearson's *r* of different models on different indicators.

Model Indicators	OLS	Ensemble Methods (Bagging)	SVR	MLP
All Indicators	0.88 (0.08)	0.88 (0.07)	0.89 (0.08)	0.89(0.07)
PCA	0.88 (0.06)	0.88 (0.07)	0.89 (0.07)	0.80(0.09)
VIF	0.56 (0.15)	0.68 (0.17)	0.55 (0.15)	0.72(0.15)
Genetic Search	0.88 (0.07)	0.87 (0.07)	0.88 (0.07)	0.87(0.07)
Evolutionary Search	0.88 (0.07)	0.88 (0.07)	0.89 (0.07)	0.88(0.06)
Best First Search	0.88 (0.07)	0.87 (0.07)	0.88 (0.07)	0.87(0.07)

According to Tables 5 and 6, the SVR performs better on the attributes made by PCA. Thus, one can use a combination of SVR and PCA to boost the model. No feature selection can improve the models. The VIF method is the worst feature selection method among the mentioned feature selection methods due to the poor prediction results. Different models are compared to identify the best model for each data set, except for VIF data (due to the not promising forecasting results). Table 7 summarizes the model's comparisons, showing that the SVR model has the best accuracy and the MLP has the worst accuracy.

Table 7. Order of the models in terms of the accuracy.

Indicators	Models
All Indicators	SVR, OLS, Ensemble methods, and MLP
PCA	SVR, OLS, Ensemble methods, and MLP
Genetic Search	SVR, OLS, Ensemble methods, and MLP
Evolutionary Search	SVR, OLS, Ensemble methods, and MLP
Best First Search	SVR, OLS, Ensemble methods, and MLP

4.2.3. Category-Based Comparative Analysis

This section applies the comparative analysis to different datasets containing different categories such as macroeconomic, microeconomic, and technical indicators. Comparison is conducted under the *T*-test at the significance level of 0.05 by WEKA software. To evaluate the predictive machine learning models' performance and have robust results, the 10-fold cross-validation on a rolling basis evaluation technique is used, and each model is repeated ten times. Therefore, the average results of 100 prediction trials, including the forecasting ability of models, namely RMSE and Pearson's *r*, are shown in Tables 8 and 9. The standard deviation is represented in parenthesis.

Table 8. RMSE of different models on different indicators.

Model Indicators	OLS	Ensemble Methods (Bagging)	SVR	MLP
All indicators	8.86 (2.36)	9.04 (1.97)	8.68 (2.48)	9.30 (2.20)
Macroeconomic indicators	19.27 (3.55)	18.54 (3.97)	19.25 (3.79)	20.74 (4.42)
Microeconomic indicators	18.42 (3.76)	16.04 (2.83)	18.76 (3.99)	17.35 (4.02)
Technical indicators	8.72 (2.10)	9.05 (2.14)	8.68 (2.17)	9.61 (2.39)

Table 9. Pearson's *r* of different models on different indicators.

Model Indicators	OLS	Ensemble Methods (Bagging)	SVR	MLP
All Indicators	0.88 (0.08)	0.88 (0.07)	0.89 (0.08)	0.89 (0.07)
Macroeconomic Indicators	0.06 (0.19)	0.25 (0.29)	0.09 (0.27)	0.25 (0.22)
Microeconomic Indicators	0.33 (0.19)	0.53 (0.23)	0.27 (0.21)	0.61 (0.20)
Technical Indicators	0.88 (0.07)	0.88 (0.07)	0.88 (0.07)	0.88 (0.07)

According to Tables 8 and 9, technical indicators impact prediction results in OLS and SVR models. The Ensemble methods and MLP models have the best accuracy on the data, including all variables. Prediction using technical indicators or using all indicators has nearly the same accuracy. In addition, all models applied on the macroeconomic and microeconomic indicators have bad accuracy with a very low Pearson's *r* and high RMSE. Therefore, it is not recommended to be used. The order of indicators according to their impact on prediction is shown in Table 10. Models applied to all attributes, and technical indicators are compared in Table 11. In both cases, the SVR model outperforms other models. Also, MLP is considered the worst model.

Table 10. The order of indicators according to their impact on prediction.

Models	The Order of Indicators according to Their Impact on Prediction
OLS	Technical indicators, all indicators, microeconomic indicators, macroeconomic indicators
Ensemble methods	All indicators, technical indicators, microeconomic indicators, macroeconomic indicators
SVR	Technical indicators, all indicators, microeconomic indicators, macroeconomic indicators
MLP	All indicators, technical indicators, microeconomic indicators, macroeconomic indicators

Table 11. The order of the models in terms of accuracy.

Indicators	Models
All Indicators	SVR, OLS, Ensemble methods, and MLP
Technical Indicators	SVR, OLS, Ensemble methods, and MLP

The category-based comparative analysis showed that macroeconomic indicators (trade-weighted US dollar index, gold fixing price, DJIA index, Brent crude oil price, and WTI) are not significant predictors for short-term BTC price. Microeconomic indicators are also not significant except for the MLP model. In addition, technical indicators, namely volume, MTM, CCI, and SMA, predict the price with nearly the same accuracy as the prediction model using all indicators. Therefore, the recommendation is to use technical analysis to predict the short-term BTC price. These empirical results answer the first and second research questions. (1) What are the significant variables as short-term or long-term BTC price predictors? (2) What are the underlying economic theories of BTC price predictors? To answer the third research question (What machine learning model performs better? What are the best feature selection techniques?), empirical results showed that the SVR model in feature-based and category-based comparative analyses outperform other models. Also, in terms of data preparation, no feature selection improved the model, and VIF turned out to be the worst feature selection.

4.3. Proposed Comparative Analysis for Dataset 2

According to the theoretical analysis regarding demand and supply theory and cost-based pricing theory, macroeconomic and blockchain information indicators do not have short-term predictability power on BTC prices. For the empirical analysis, a comparative machine learning model, including OLS, Ensemble methods, SVR, and MLP for the BTC price prediction problem on datasets from 1 January 2018 to 5 June 2018, including macroeconomic, microeconomic, technical indicators, and blockchain information indicators. Feature selections, namely best first search, PSO search, and evolutionary search, are applied to the data. The price prediction model is described in the Appendix A (model 3).

4.3.1. Data Description

Daily BTCUSD transactions occurring on the Bitfinex exchange obtained via provided API at bitfinex.com from 1 January 2018, to 5 June 2018, including 17 independent variables, have been examined. Dependent variables can be categorized into macroeconomic variables, extracted from macrotrends.net (accessed on 2 October 2019), microeconomic, technical indicators, and Blockchain information indicators obtained from data.BTCity.org. Table 12 shows the specification for each group. Some descriptive statistics, including minimum, maximum, mean, and standard deviation, have been calculated and shown in Table A6 in the Appendix A.3 to describe or summarize the data.

Table 12. Data categorization.

Indicator Category	Indicator Name
Macroeconomic indicators	S&P500 index, Nasdaq Composite, DJIA index, CAC 40 Index, WTI, gold fixing price
Microeconomic indicators	Bid–ask spread (10BTC), ask sum (10%), bid sum (10%), trades per minute
Technical indicators	Volatility, volume, SMA
Blockchain information indicators	Hash rate, mining difficulty, number of transactions per block, block time

4.3.2. Feature-Based Comparative Analysis

This section applies the comparative analysis to different datasets containing the indicators chosen by different feature selection techniques, including best-first search, evolutionary search, PSO search, and PCA dimension reduction methods. Table A7 in the Appendix A.3 presents the different features chosen by other methods. For the analysis, machine learning models, including OLS, Ensemble methods (bagging), SVR (with a polynomial kernel), and MLP (with one hidden layer and nine neurons), have been applied to different datasets, which include the indicators selected by other feature selections. The aim is to specify the best feature selection method and determine the best machine learning method. To evaluate the predictive machine learning models' performance and have robust results, the 10-fold cross-validation on a rolling basis evaluation technique is used, and each model is repeated ten times. Therefore, the average results of 100 prediction trials, including the forecasting ability of models, namely RMSE and Pearson's *r*, are shown in Tables 13 and 14.

Table 13. RMSE of different models on different datasets.

Model Indicators	OLS	Ensemble Methods (Bagging)	SVR	MLP
All Indicators	157.36 (30.24)	160.06 (36.52)	154.49 (31.53)	163.37 (44.62)
Best First	161.36 (34.57)	162.85 (38.69)	158.87 (36.20)	164.16 (40.37)
PCA Reduction	160.48 (34.38)	178.77 (40.04)	160.26 (33.52)	179.77 (45.12)
PSO Search	160.70 (29.31)	162.90 (37.43)	158.06 (34.26)	175.40 (43.50)
Evolutionary Search	161.03 (31.97)	162.43 (34.99)	160.00 (38.76)	169.70 (49.65)

Table 14. Pearson's *r* of different models on different data sets.

Model Indicators	OLS	Ensemble Methods (Bagging)	SVR	MLP
All Indicators	0.77 (0.13)	0.74 (0.14)	0.76 (0.13)	0.77 (0.12)
Best First Search	0.74 (0.16)	0.72 (0.16)	0.74 (0.14)	0.77 (0.16)
PCA Reduction	0.76 (0.13)	0.65 (0.17)	0.74 (0.13)	0.76 (0.12)
PSO Search	0.75 (0.14)	0.72 (0.16)	0.75 (0.13)	0.73 (0.17)
Evolutionary Search	0.74 (0.15)	0.74 (0.13)	0.74 (0.14)	0.77 (0.16)

According to Tables 13 and 14, all models applied to all indicators have the best accuracy than those applied to the other datasets. Therefore, it can be concluded that no feature selection improves the model's accuracy. Compared to those applied to the different datasets, all models applied to data reduced by PCA have the lowest accuracy. Therefore, it can be concluded that the PCA reduction method is not a promising feature selection method for this research data. Different models are compared together for each data set to identify the best model. Table 15 summarizes the model's comparisons, showing that the SVR model has the best accuracy for all datasets, and the MLP has the least accuracy.

Table 15. Order of the models in terms of accuracy.

Datasets	Models
All Indicators	SVR, OLS, Ensemble methods, and MLP
Best First Search	SVR, OLS, Ensemble methods, and MLP
PCA Reduction	SVR, OLS, Ensemble methods, and MLP
PSO Search	SVR, OLS, Ensemble methods, and MLP
Evolutionary Search	SVR, OLS, Ensemble methods, and MLP

4.3.3. Category-Based Comparative Analysis

OLS, Ensemble methods, SVR, and MLP are applied to economic and technical indicators. The aim is to see which indicators can be selected as better predictive indicators. Also, different models are compared on the same data to find a more accurate model. To evaluate the predictive machine learning models' performance and have robust results, the 10-fold cross-validation on a rolling basis evaluation technique is used, and each model is repeated ten times. Therefore, the average results of 100 prediction trials, including the forecasting ability of models, namely RMSE and Pearson's *r*, are shown in Tables 16 and 17.

Table 16. RMSE of different models on different indicators.

Model Indicators	OLS	Ensemble Learning	SVR	MLP
All indicators	157.36 (30.24)	160.06 (36.52)	154.49 (31.53)	174.37 (44.62)
Blockchain information indicators	242.29 (46.77)	243.07 (48.78)	248.09 (51.72)	281.13 (60.84)
Macroeconomic indicators	251.56 (46.90)	230.01 (43.84)	249.30 (47.05)	262.18 (59.76)
Microeconomic indicators	198.61 (36.65)	193.00 (36.62)	197.99 (37.74)	205.60 (48.95)
Technical indicators	173.07 (41.11)	161.97 (38.69)	172.72 (40.78)	191.32 (52.98)

Table 17. Pearson's *r* of different models on different indicators.

Models Indicators	OLS	Ensemble Learning	SVR	MLP
All indicators	0.75 (0.13)	0.74 (0.14)	0.76 (0.13)	0.77 (0.12)
Blockchain information indicators	0.11 (0.27)	0.10 (0.24)	−0.01 (0.25)	−0.04 (0.26)
Macroeconomic indicators	−0.00 (0.25)	0.23 (0.34)	0.07 (0.32)	0.21 (0.31)
Microeconomic indicators	0.57 (0.23)	0.58 (0.21)	0.57 (0.22)	0.60 (0.23)
Technical indicators	0.68 (0.16)	0.73 (0.14)	0.69 (0.16)	0.69 (0.16)

According to Tables 16 and 17, all models applied to all indicators have the best accuracy. Therefore, it is recommended that the combination of technical, microeconomics, macroeconomic, and Blockchain information indicators work better for price prediction than each indicator category alone. Moreover, technical indicators are also considered good predictors. However, prediction slightly improves by combining with other variables.

Blockchain information and macroeconomic indicators are considered bad predictive indicators due to the very low Pearson's *r* and high RMSE. The order of indicators according to their impact on prediction is shown in Table 18. Models applied on all indicators and technical indicators are compared in Table 19. In both cases, the SVR model outperforms other models. Also, MLP is considered the worst model.

Table 18. The order of indicators according to their impact on prediction.

Models	Order of Indicators according to Their Impact on Prediction
OLS	All indicators, technical indicators, microeconomic indicators, blockchain information indicators, macroeconomic indicators
Ensemble methods	All indicators, technical indicators, microeconomic indicators, macroeconomic indicators, Blockchain information indicators
SVR	All indicators, technical indicators, microeconomic indicators, Blockchain information indicators, macroeconomic indicators
MLP	All indicators, technical indicators, microeconomic indicators, macroeconomic indicators, Blockchain information indicator

Table 19. The order of the models in terms of accuracy.

Indicators	Models
All indicators	SVR, OLS, Ensemble methods, and MLP
Technical indicators	SVR, OLS, Ensemble methods, and MLP

The results of the category-based comparative analysis showed that macroeconomic indicators (trade-weighted US dollar index, gold-fixing price, DJIA index, Brent crude oil price, and WTI) are not significant predictors. Also, the Blockchain information indicators, including hash rate, mining difficulty, number of transactions per block, and block time, are not significant predictors for short-term BTC price. Also, microeconomic indicators, including trades per minute, bid/ask sum, bid–ask spread, and buy/sell signals, are not significant for the BTC price prediction except for the MLP model. Since the technical indicators have nearly the same results as all indicators, the recommendation is to use the technical analysis to predict the short-term BTC price. These empirical results answer the first and second research questions. (1) what are the significant variables as short-term or long-term BTC price predictors? (2) What are the underlying economic theories of BTC price predictors? To answer the third research questions (What machine learning model performs better? What are the best feature selection techniques?), empirical results showed that the SVR model in feature-based and category-based comparative analyses outperform the other models. In terms of data preparation, no feature selection improved the model, and PCA dimension reduction turned out to be the worst feature selection.

5. Conclusions

Today, international finance is a multi-trillion-dollar sector that needs a secure and stable mechanism that cryptocurrencies are currently inching. Cryptocurrencies were developed under Blockchain technology. In contrast with the traditional central authority systems wherein the sole control lies under one organization, Blockchain technology has a diversified approach. This paper applied several machine learning models to the BTC price prediction model on different data sets to verify the theoretical analysis and answer the research questions. A multilinear regression model to monthly BTC prices showed that macroeconomic and Blockchain information indicators are significant long-term predictors. That verifies that supply and demand and cost-based pricing theory are underlying BTC price predictors. These empirical results answer the first and second research questions. (1) What are the significant variables as short-term or long-term BTC price predictors? (2) What are the underlying economic theories of BTC price predictors? In addition, the empirical results showed that SVR is the best machine learning model, and

no feature selection technique is proven to be the best, which answers the third research questions (Are machine learning algorithms superior to traditional methods for BTC price prediction? What machine learning model performs better? What are the best feature selection techniques?).

The conclusions are relevant to central bankers, investors, asset managers, etc., who are generally interested in information about which indicators provide reliable, accurate forecasts of BTC price. The study can be used to set asset pricing and improve investment decision-making. Therefore, it provides a significant opportunity to contribute to international finance since the results have significant implications for the future decisions of asset managers. In time series prediction, the correlation between independent variables and dependent variables differs from time to time. Consequently, reestimating prediction models is not unlikely. This study has used many data categories composing macroeconomic, microstructure, Blockchain information, and technical indicators to make a wide-ranging work.

In this study, attributes are selected based on economic theories. Macroeconomic indicators are chosen based on the supply and demand theory. Microstructure theory is the underlying theory of microeconomic indicators. Also, Blockchain information indicators are selected according to the cost-based pricing theory. Previous studies are mostly empirical research in which the focus is on the prediction methods. After describing the price movement from the perspective of economic theories, the empirical results confirmed the theoretical analysis. This study compared methodologies to predict short-term and long-term BTC prices. The conclusion is also helpful for machine learning developers to understand the configuration of machine learning prediction models and use it as benchmarks. According to the literature review, the authors still doubt whether machine learning can beat the traditional methods for BTC price prediction. Therefore, this study is evidence of the superiority of machine learning.

This research has some suggestions for future work, which are as follows. In this research, only a few critical feature selection methods have been applied to data sets. Many other attribute selection techniques, including ranker search, Tabu search, and many more, can be examined to improve the model. Other research can compare trending models, such as recurrent neural networks (RNN) to SVR. According to this research, a correct prediction of BTC prices can be profitable; therefore, it can diversify a portfolio. Further studies can be conducted to examine the portfolio return by adding BTC to a portfolio to determine the right amount of BTC to keep. Future research can predict other cryptocurrencies, including Ethereum and Ripple. In addition, some other indicators, such as "news," can be investigated in other studies.

Author Contributions: Conceptualization, S.E. and Y.Z.; methodology, S.E., A.R., A.A. and T.L. software, Y.Z. and A.A.S.; validation, A.R., A.A. and Y.Z.; formal analysis, S.E. and A.A.S.; investigation, A.R., T.L. and A.A.; resources, Y.Z. and T.L.; data curation, A.R. and A.A.S.; writing—original draft preparation, S.E., Y.Z. and T.L.; writing—review and editing, A.A., A.R. and T.L.; visualization, Y.Z., A.R. and A.A.; supervision, Y.Z. and T.L.; project administration, S.E. and A.A.S.; funding acquisition, Y.Z., A.R. and T.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The datasets used and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1 Model 1. OLS Model Description

The purpose is to find a model that can approximate a target function, which can be written as:

$$r_{i+1} = \alpha + \beta_1 X_{i,1} + \beta_2 X_{i,2} + . + \beta_N X_{i,N} + \varepsilon_i \quad \text{(A1)}$$

where r_i is the BTC price at month $i + 1$. $X_{i,1}$ *to* $X_{i,N}$ are attributes at day/month i which is described as follows:

$X_{i,1}$: BTCs in Circulation in month i is the total number of mined BTC currently circulating on the network.

$X_{i,2}$: Market capitalization in the month i is calculated by multiplying the total number of BTCs in circulation by the BTC price.

$X_{i,3}$: Block size in month i imposes a limit on the number of transactions that can be verified on each block. As a result of such a mechanism, larger blocks require more processing power and longer extraction time.

$X_{i,4}$: Average block size in month i.

$X_{i,5}$: Orphaned blocks in the month i are blocks that are not accepted into the blockchain network, which is created due to the delay in receiving a block, at which point another miner responds to the same block. Orphan blocks are valid, but do not register any transaction and have been rejected by the chain.

$X_{i,6}$: Number of transactions per block in month i are the transactions that happen in a block, and as soon as a block is solved, it is not possible to extend the block by adding more transactions.

$X_{i,7}$: Median confirmation time in month i is the median time for dealing with miners' fees enclosed in a mined block and superimposed to the public ledger.

$X_{i,8}$: Hash rate in day/month i is the speed at which computational operations are completed to mine a BTC block.

$X_{i,9}$: Mining difficulty in month i is a measure of how difficult it is to mine a BTC block, or in more technical terms, to find a hash below a given target.

$X_{i,10}$: Transaction fees in month i are paid when cryptocurrencies are transferred to another wallet. Processing transactions on the blockchain takes effort, and these fees are used to compensate the miners and validators who help keep things running smoothly.

$X_{i,11}$: Cost per transaction in month i is calculated as miners' revenue divided by the number of transactions.

$X_{i,12}$: Unique addresses in month i are installment addresses that have a non-zero adjust. This metric is one way of understanding day-by-day utilization of the BTC arrangement.

$X_{i,13}$: Total BTC transactions in month i.

$X_{i,14}$: Transaction volume excluding popular addresses in month i is the total number of transactions excluding those involving the network's 100 most popular addresses.

$X_{i,15}$: Total output value in month i is the total value of all transaction outputs, including coins, returned to the sender as change.

$X_{i,16}$: Estimated transactions value in month i is the total estimated value in BTC transactions on the blockchain, which does not include coins returned as change.

$X_{i,17}$: Nasdaq Composite is a stock market index of the common stocks and similar securities listed on the Nasdaq stock market.

$X_{i,18}$: Dow Jones Industrial Average (DJIA) index in month i is a stock market index that measures the stock performance of 30 large companies listed on stock exchanges in the United States.

$X_{i,19}$: S&P500 stock market index in month i is a stock market index that measures the stock performance of 500 large companies listed on stock exchanges in the United States.

$X_{i,20}$: Gold-fixing price in month i is the setting of the gold price that takes place via a dedicated conference line. The price continues to be set twice daily at 10:30 and 15:00 London GMT in US dollars.

$X_{i,21}$: West Texas Intermediate crude oil (WTI) prices or Texas light sweet in month i is a benchmark in oil pricing, refined mainly in the Midwest and Gulf Coast regions in the United States.

$X_{i,22}$: US federal funds rate in month i is the interest rate at which depositors trade federal funds with each other at night. When a depository institution has a surplus in its reserve accounts, it can lend to other banks that need those funds. In other words, a bank with extra cash can lend it to another bank with a liquidity problem, and thus the cash balance of a bank with a problem with cash increases rapidly.

$X_{i,23}$: Breakeven inflation rate in month i is a measure of expected inflation, the difference between a nominal bond's yield and an inflation-linked bond with the same maturity.

Table A1. Descriptive statistics.

Indicators	Min	Max	Mean	Std. Dev
BTCs in Circulation	4,002,626.667	17,213,768.33	12,440,527.97	3,742,297.323
Market Capitalization	280,390.572	2.55×10^{11}	23,117,501,754	49,243,267,354
Block Size	1	179,101.0913	47,145.66415	54,242.32369
Average Block Size	0.01	1.054375	0.409168031	0.354286668
Orphaned Block	0	2.071428571	0.361061508	0.556278003
Transactions Per Block	1.625	2208.7575	760.5240094	666.8726595
Median Confirmation Time	6.201875	16.96133333	9.397754898	2.300005847
Hash Rate	0.01	49,050,545.4	3,657,003.427	9,072,465.431
Mining Difficulty	797.7186667	6.32×10^{12}	4.79166×10^{11}	1.18908×10^{12}
Transaction Fee	0.056875	591.31625	62.77254092	105.0417027
Cost Per Transaction	1.242	117.1433333	20.31455332	25.94717866
Unique Addresses	513.6666667	825,390.9375	224,455.5763	207,905.1755
Total Transactions Per Day	464.0666667	358,831.0625	11,5921.8508	100,973.4369
Transaction Volume Excluding Popular Addresses	464.0666667	341,004.75	107,356.0276	101,413.1972
Total Output Value	63,281.56267	11,338,010.91	1,650,410.836	1,539,818.478
Estimated Transaction Value	27,539.66667	997,305.9375	209,259.0939	130,674.3663
Nasdaq Composite	2286.248	7882.400667	4435.937597	1480.271196
Dow Jones Industrial Average	10,576.508	25,807.52933	16,784.42003	3980.41153
S&P 500	1119.546667	2855.994	1880.952363	472.923476
Gold Price Index	1072.293333	1773.213333	1361.324703	184.7161172
Crude Oil WTI	30.485	110.3573333	74.41217956	23.3697773
US Federal Funds Rate	0.067142857	1.915333333	0.392972284	0.491999589
Breakeven Inflation Rate	1.302857143	2.586666667	2.011131634	0.298349738

Table A2. VIF for choosing attributes.

Variables	VIF
BTCs in Circulation	7.98
Market Capitalization	27.44 *
Block Size	7.68
Average Block Size	5.07
Orphaned Block	1.5
Transactions Per Block	39.11 *
Median Confirmation Time	1.73
Hash Rate	52.36 *
Mining Difficulty	51.45 *
Transaction Fee	3.53
Cost Per Transaction	33.88 *
Unique Addresses	9.75
Total Transactions Per Day	48.67 *

Table A2. *Cont.*

Variables	VIF
Transaction Volume Excluding Popular Addresses	8.44
Total Output Value	2.4
Estimated Transaction Value	2.49
Nasdaq Composite	11.86 *
Dow Jones Industrial Average	23.71 *
S&P 500	44.93 *
Gold Price Index	2.32
Crude Oil WTI	2.16
US Federal Funds Rate	1.99
Breakeven Inflation Rate	2.97

'*' VIF greater than 10.

Table A3. OLS regression results.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
BTCs in Circulation									
Block Size	0.268 ** (0.039)				0.521 * (0.164)	0.418 * (0.196)	0.408 ** (0.038)	0.443 ** (0.021)	0.436 ** (0.024)
Transaction Fees	0.131 *** (0.001)	0.167 ** (0.05)		0.158 ** (0.049)			0.166 *** (0.002)		0.155 ** (0.003)
Unique Addresses					1.021 *** (0.184)				
Total Number of Transactions			−0.023 . (0.012)						
Estimated Transaction Value	−0.096 ** (0.03)	−0.192 ** (0.071)		−0.179 * (0.070)	−0.149 * (0.062)	−0.242 ** (0.071)			−0.213 *** (0.004)
Cost Per Transaction	0.781 ** (0.07)								
Mining Difficulty		0.327 * (0.124)							
Market Capitalization			1.00 *** (0.002)						
Hash Rate				0.397 ** (0.126)					
Nasdaq Composite							0.809 . (0.1)		
Dow Jones Industrial Average								0.277 ** (0.024)	
S&P500									0.081 ** (0.038)
Adjusted ^{R2}	0.91	0.67	0.81	0.73	0.68	0. 67	0.73	0.78	0.73
Residual Standard Error	0.044	0.1	0.001	0.023	0.087	0.10	0.099	0.089	0.089
p-value	$<2.2 \times 10^{-16}$	5.56×10^{-11}	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$	1.073×10^{-14}	7.28×10^{-10}	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-15}$

'***' Significant at the 0.001 level, '**' Significant at the 0.01 level, '*' Significant at the 0.05 level, '.' Significant at the 0.1 level.

Appendix A.2 Model 2. Model Description

The purpose is to find a model that can approximate a target function by navigating the space of possible hypotheses (e.g., for ANN models, the space of hypotheses includes

the network topology and hyperparameters) to predict the price changes for one day ahead. The target function can be written as:

$$\hat{\Delta p}_{i+1} = f(\Delta X_{i1}, \Delta X_{i2}, \Delta X_{i3}, \ldots, \Delta X_{in}) \tag{A2}$$

where $\hat{\Delta p}_{i+1}$ are the BTC price changes at day $i+1$. ΔX_{i1} *to* ΔX_{in} are attributes at day i that might affect the price changes, which are described as follows:

X_{i1} : Trade-weighted US dollar index or broad index (TWEXB) on day i is a measure of the value of the United States dollar relative to other world currencies.

X_{i2} : Gold-fixing price on day i is the setting of the price of gold that takes place via a dedicated conference line. The price continues to be set twice daily at 10:30 and 15:00 London GMT in US dollars.

X_{i3} : Dow Jones Industrial Average (DJIA) index on the day i is a stock market index that measures the stock performance of 30 large companies listed on stock exchanges in the United States.

X_{i4} : Brent Crude oil price on day i is a primary trading classification of sweet light crude oil from the North Sea that is an important benchmark that defines the prices for oil worldwide.

X_{i5} : West Texas Intermediate crude oil (WTI) prices or Texas light sweet, on day i is a benchmark in oil pricing, refined mainly in the Midwest and Gulf Coast regions in the United States.

X_{i6} : Trades per minute on the day i is the number of BTCs traded in a minute.

X_{i7} : Ask sum (5%) on day i, calculated as the amount of BTC on the order books waiting to be sold within a 5% range from the BTC price.

X_{i8} : Bid sum (5%) on day i, calculated as the amount of USD on the order books waiting to buy BTC within a 5% range from the BTC price.

X_{i9} : Bid–ask spread (10BTC) on day i is spread with 10 BTC slippage, i.e., with 10 BTC worth of orders removed from bids and from asks, which is calculated as $\frac{ask_{min} - bid_{max}}{ask_{min}} \times 100$.

X_{i10} : Bid–ask spread (100BTC) on day i, i.e., with 10 BTC worth of orders removed from bids and from asks, which is calculated as

$$\frac{ask_{min} - bid_{max}}{ask_{min}} \times 100.$$

X_{i11} : Buy0BTC on day i, defined as buy orders with an amount of less than 1 BTC.

X_{i12} : Sell0BTC on day i, defined as sell orders with an amount of less than 1 BTC.

X_{i13} : Buy1BTC on day i, defined as buy orders with an amount of 1 BTC.

X_{i14} : Sell1BTC on day i, defined as sell orders with an amount of 1 BTC.

X_{i15} : Buy5BTC on day i, defined as buy orders with an amount of 5 BTC.

X_{i16} : Sell5BTC on day i, defined as sell orders with an amount of 5 BTC.

X_{i17} : Buy10BTC on day i, defined as buy orders with an amount of 10 BTC.

X_{i18} : Sell10BTC on day i, defined as sell orders with an amount of 10 BTC.

X_{i19} : Momentum (MTM) (10 days) on day i is the difference between the price of BTC on day i and the BTC price on $i - N$th day which is $N = 10$ in this model.

X_{i20} : Commodity Channel Index (CCI), on day i, compares the price of BTC against its simple moving average and mean deviation of the price.

X_{i21} : Volume on day i is the number of BTCs traded during a given period, which is one day in our model.

X_{i22} : Simple moving average (SMA) on day i, calculated by adding recent prices and then dividing that by the number of periods, is five days for this model.

Table A4. Descriptive statistics.

Indicators	Min	Max	Mean	Std. Dev
TWEXB	91.45	96.87	93.90	1.25
Gold Fixing Price	1130.55	1304.55	1228.23	42.67
DJIA	17,888.28	21,271.97	20,055.42	951.69
Brent Crude Oil Price	41.61	56.34	51.27	3.57
WTI	43.29	54.48	50.13	2.84
Trades Per Minute	0.92	63.17	11.71	10.43
Ask Sum (5%)	750.97	6067.64	2737.32	1065.50
Bid Sum (5%)	567.64	5667.86	2378.06	989.71
Bid–Ask Spread (10BTC)	0.04	0.66	0.17	0.11
Bid–Ask Spread (100BTC)	0.30	2.90	0.78	0.45
Buy0BTC	767.00	41,552.00	8257.62	7121.62
Sell0BTC	559.00	49,411.00	8630.88	8091.82
Buy1BTC	160.00	7583.00	1820.29	1436.78
Sell1BTC	179.00	9272.00	1781.89	1600.48
Buy5BTC	35.00	2055.00	332.16	301.08
Sell5BTC	25.00	2553.00	354.77	365.67
Buy10BTC	1.00	685.00	94.53	86.32
Sell10BTC	2.00	838.00	93.58	109.68
Momentum	85.96	120.56	97.87	5.60
CCI	−351.04	524.63	87.53	111.17
Volume	1,538,729.58	134,500,681.52	22,138,504.67	22,622,391.54
SMA	631.17	2867.59	1150.82	508.79

Table A5. Chosen attributes by different feature selection techniques.

Attributes	VIF	Genetic Search	Evolutionary Search	Best First Search
TWEXB	✓			✓
Gold-Fixing Price	✓			
DJIA	✓			
Brent Crude Oil Price	✓	✓		✓
WTI				
Volume				
Trades Per Minute	✓			
Ask sum (5BTC)	✓			
Bid Sum (5BTC)				
Bid–Ask Spread (10BTC)	✓	✓	✓	✓
Bid–Ask Spread (100BTC)	✓	✓		✓
Buy0BTC				
Sell0BTC	✓			
Buy1BTC				
Sell1BTC	✓			
Buy5BTC				
Sell5BTC				
Buy10BTC	✓		✓	
Sum5BTCPrice		✓	✓	✓
Sell10BTC	✓		✓	
MTM	✓			
CCI	✓	✓	✓	✓

Appendix A.3 Model 3. Model Description

The purpose is to find a model that can approximate a target function, which can be written as:

$$\hat{\Delta p}_{i+1} = f(\Delta X_{i1}, \Delta X_{i2}, \Delta X_{i3}, \ldots, \Delta X_{in}) \quad \text{(A3)}$$

where $\hat{\Delta p}_{i+1}$ are the BTC price changes at day $i+1$. ΔX_{i1} *to* ΔX_{in} are attributes at day i that might affect the price changes, which are described as follows.

X_{i1} : S&P500 stock market index on day i is a stock market index that measures the stock performance of 500 large companies listed on stock exchanges in the United States.

X_{i2} : Dow Jones Industrial Average (DJIA) index on day i is a stock market index that measures the stock performance of 30 large companies listed on stock exchanges in the United States

X_{i3} : CAC 40 stock market index on day i is a stock market index representing a capitalization-weighted measure of the 40 most significant stocks among the 100 most oversized market caps on the Euronext Paris.

X_{i4} : West Texas Intermediate crude oil (WTI) prices or Texas light sweet on day i is a benchmark in oil pricing, refined mainly in the Midwest and Gulf Coast regions in the United States.

X_{i5} : Nasdaq Composite on day i is a stock market index of the common stocks and similar securities listed on the Nasdaq stock market.

X_{i6} : Gold-fixing price on day i is the setting of the gold price that takes place via a dedicated conference line. The price continues to be set twice daily at 10:30 and 15:00 London GMT in US dollars.

X_{i7} : Bid–ask spread (10BTC), on day i is spread with 10 BTC slippage, that is with 10 BTC worth of orders removed from bids and from asks, which is calculated as $\frac{ask_{min} - bid_{max}}{ask_{min}} \times 100$.

X_{i8} : Ask sum (10%) on day i, calculated as the amount of BTC on the order books waiting to be sold within a 10% range from the BTC price.

X_{i9} : Bid sum (10%) on day i, calculated as the amount of USD on the order books waiting to buy BTC within a 10% range from the BTC price.

X_{i10} : Trades per minute on day i are the number of BTCs traded in a minute.

X_{i11} : Volatility on day i is the changes in market prices over a specified period. The faster prices change, the higher the volatility. It can be measured and calculated based on historical prices and can be used for trend identification.

X_{i12} : Volume on day i is the number of BTCs traded during a given period, which is one day in our model.

X_{i13} : Simple moving average (SMA) on day i, calculated by adding recent prices and then dividing that by the number of periods, which is five days for this model.

X_{i14} : Hash rate on day i is the speed at which computational operations are completed to mine a BTC block.

X_{i15} : Mining difficulty on day i is a measure of how difficult it is to mine a BTC block, or in more technical terms, to find a hash below a given target.

X_{i16} : Number of transactions per block on day i are the transactions that happen in a block, and as soon as a block is solved, it is not possible to extend the block by adding in more transactions.

X_{i17} : Block time on day i is an average time to mine a block in minutes.

Table A6. Descriptive statistics.

Data	Min	Max	Mean	Std. Dev
S&P500 Index	2581.00	2872.87	2711.50	63.73
Nasdaq Composite	6777.16	7637.86	7246.83	186.26
DJIA Index	23,533.20	26,616.71	24,842.04	676.33
CAC 40 Index	1425.12	5640.10	5297.19	527.43
WTI	59.19	72.24	65.03	3.34
Gold Fixing Price	1285.85	1360.25	1324.94	16.97
Bid/Ask Spread (10BTC)	0.21	0.68	0.36	0.12
Ask Sum (10%)	5.62×10^6	2.32×10^7	1.21×10^7	3.62×10^6
Bid Sum (10%)	9.71×10^6	2.65×10^7	1.49×10^7	3.34×10^6
Trades Per Minute	10.50	94.21	31.59	14.22
Volatility	7.80	154.64	39.35	25.15
Volume	3144.45	70961.37	3144.45	8830.17
SMA	1498.466429	14,907.4622	9030.497881	2036.55495
Hash Rate	1.63×10^{18}	9.42×10^{18}	3.89×10^{18}	1.12×10^{18}
Mining Difficulty	1.93×10^{12}	4.94×10^{12}	3.30×10^{12}	7.27×10^{11}
Number of Transactions Per Block	1.35×10^5	4.25×10^5	2.10×10^5	5.16×10^4
Block Time	7.48	12.22	9.34	0.86

Table A7. Attributes selected by different feature selection methods.

Attributes	Best First Search	PSO Search	Evolutionary Search
S&P500 Index			✓
Nasdaq Composite			
DJIA Index		✓	
CAC 40 Index	✓	✓	✓
WTI	✓		✓
Gold Fixing Price	✓		✓
Bid–Ask Spread (10BTC)	✓	✓	
Ask Sum within (10BTC)			✓
Bid Sum within (10BTC)			
Trades Per Minute	✓	✓	✓
Volatility	✓	✓	✓
Volume			
SMA	✓	✓	✓
Hash rate			
Mining Difficulty			
Number of Transactions			
Block Time	✓	✓	✓

References

1. Rosenberg, J.M.; Krist, C. Combining machine learning and qualitative methods to elaborate students' ideas about the generality of their model-based explanations. *J. Sci. Educ. Technol.* **2021**, *30*, 255–267. [CrossRef]
2. Bertolini, R.; Finch, S.J.; Nehm, R.H. Testing the impact of novel assessment sources and machine learning methods on predictive outcome modeling in undergraduate biology. *J. Sci. Educ. Technol.* **2021**, *30*, 193–209. [CrossRef]
3. Ashayer, A. *Modeling and Prediction of Cryptocurrency Prices Using Machine Learning Techniques*; East Carolina University Greenville: Greenville, NC, USA, 2019.
4. Dutta, A.; Kumar, S.; Basu, M. A gated recurrent unit approach to bitcoin price prediction. *J. Risk Financ. Manag.* **2020**, *13*, 23. [CrossRef]
5. Wang, S.; Vergne, J.-P. Buzz factor or innovation potential: What explains cryptocurrencies' returns? *PLoS ONE* **2017**, *12*, e0169556.
6. Conrad, C.; Custovic, A.; Ghysels, E. Long-and short-term cryptocurrency volatility components: A GARCH-MIDAS analysis. *J. Risk Financ. Manag.* **2018**, *11*, 23. [CrossRef]

7. Jang, H.; Lee, J. An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. *Ieee Access* **2017**, *6*, 5427–5437. [CrossRef]
8. Chen, Z.; Li, C.; Sun, W. Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *J. Comput. Appl. Math.* **2020**, *365*, 112395. [CrossRef]
9. Pang, Y.; Sundararaj, G.; Ren, J. Cryptocurrency price prediction using time series and social sentiment data. In Proceedings of the 6th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies, Auckland, New Zealand, 2–5 December 2019; pp. 35–41.
10. Antoniou, A.; Ergul, N.; Holmes, P.; Priestley, R. Technical analysis, trading volume and market efficiency: Evidence from an emerging market. *Appl. Financ. Econ.* **1997**, *7*, 361–365. [CrossRef]
11. Buchholz, M.; Delaney, J.; Warren, J.; Parker, J. Bits and bets, information, price volatility, and demand for Bitcoin. *Economics* **2012**, *312*, 2–48.
12. Nai Fovino, I.; Steri, G.; Fontana, A.; Ciaian, P.; Kancs, D.; Nordvik, J. *On Virtual and Crypto Currencies: A General Overview, from the Technological Aspects to the Economic Implications*; JRC Technical Report JRC9997: Ispra, Italy, 2015.
13. Ciaian, P.; Rajcaniova, M.; Kancs, D.A. Virtual relationships: Short- and long-run evidence from BitCoin and altcoin markets. *J. Int. Financ. Mark. Inst. Money* **2018**, *52*, 173–195. [CrossRef]
14. Keynes, J.M. *The General Theory of Employment, Interest, and Money*; Harcourt, Brace & World, Inc.: New York, NY, USA, 1936.
15. Kristoufek, L. What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PLoS ONE* **2015**, *10*, e0123923. [CrossRef]
16. Lyons, R.K. *The Microstructure Approach to Exchange Rates*; MIT Press: Cambridge, MA, USA, 2001; Volume 333.
17. Marshall, A. *Principles of Economics*; Palgrave Macmillan: London, UK, 2013.
18. Zhang, Y.-C. Toward a theory of marginally efficient markets. *Phys. A Stat. Mech. Its Appl.* **1999**, *269*, 30–44. [CrossRef]
19. Amihud, Y.; Mendelson, H. Asset pricing and the bid-ask spread. *J. Financ. Econ.* **1986**, *17*, 223–249. [CrossRef]
20. Reinganum, M.R. Market microstructure and asset pricing: An empirical investigation of NYSE and NASDAQ securities. *J. Financ. Econ.* **1990**, *28*, 127–147. [CrossRef]
21. Dyhrberg, A.H.; Foley, S.; Svec, J. How investible is Bitcoin? Analyzing the liquidity and transaction costs of Bitcoin markets. *Econ. Lett.* **2018**, *171*, 140–143. [CrossRef]
22. Scaillet, O.; Treccani, A.; Trevisan, C. High-frequency jump analysis of the bitcoin market. *J. Financ. Econom.* **2020**, *18*, 209–232.
23. Guo, T.; Bifet, A.; Antulov-Fantulin, N. Bitcoin volatility forecasting with a glimpse into buy and sell orders. In Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM), Singapore, 17–20 November 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 989–994.
24. Noble, P.M.; Gruca, T.S. Industrial pricing: Theory and managerial practice. *Mark. Sci.* **1999**, *18*, 435–454. [CrossRef]
25. Hayes, A.S. Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telemat. Inform.* **2017**, *34*, 1308–1321. [CrossRef]
26. Hayes, A.S. Bitcoin price and its marginal cost of production: Support for a fundamental value. *Appl. Econ. Lett.* **2019**, *26*, 554–560. [CrossRef]
27. Yu, Y.; Rashidi, M.; Samali, B.; Yousefi, A.M.; Wang, W. Multi-image-feature-based hierarchical concrete crack identification framework using optimized SVM multi-classifiers and D–S fusion algorithm for bridge structures. *Remote Sens.* **2021**, *13*, 240. [CrossRef]
28. Mohan, S.; Thirumalai, C.; Srivastava, G. Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access* **2019**, *7*, 81542–81554. [CrossRef]
29. Li, C.; He, Y.; Xiao, D.; Luo, Z.; Fan, J.; Liu, P.X. A novel hybrid approach of ABC with SCA for the parameter optimization of SVR in blind image quality assessment. *Neural Comput. Appl.* **2022**, *34*, 4165–4191. [CrossRef]
30. Ebrahimpour, Z.; Wan, W.; Khoojine, A.S.; Hou, L. Twin hyper-ellipsoidal support vector machine for binary classification. *IEEE Access* **2020**, *8*, 87341–87353. [CrossRef]
31. Yu, Y.; Li, Y.; Li, J.; Gu, X. Self-adaptive step fruit fly algorithm optimized support vector regression model for dynamic response prediction of magnetorheological elastomer base isolator. *Neurocomputing* **2016**, *211*, 41–52. [CrossRef]
32. Chun, M.W.; Huat, N.C.; Pauline, O. Application of machine learning algorithm to prediction of thermal spring back of hot press forming. *Res. Prog. Mech. Manuf. Eng.* **2022**, *3*, 875–883.
33. Shafiabady, N.; Lee, L.H.; Rajkumar, R.; Kallimani, V.; Akram, N.A.; Isa, D. Using unsupervised clustering approach to train the Support Vector Machine for text classification. *Neurocomputing* **2016**, *211*, 4–10. [CrossRef]
34. Erfanian, S.; Ziaullah, M.; Tahir, M.A.; Ma, D. How does justice matter in developing supply chain trust and improving information sharing-an empirical study in Pakistan. *Int. J. Manuf. Technol. Manag.* **2021**, *35*, 354–368. [CrossRef]
35. Razzaq, A.; Tang, Y.; Qing, P. Towards Sustainable Diets: Understanding the Cognitive Mechanism of Consumer Acceptance of Biofortified Foods and the Role of Nutrition Information. *Int. J. Envion. Res. Pub. Health* **2021**, *18*, 1175. [CrossRef]
36. Aggarwal, D.; Chandrasekaran, S.; Annamalai, B. A complete empirical ensemble mode decomposition and support vector machine-based approach to predict Bitcoin prices. *J. Behav. Exp. Financ.* **2020**, *27*, 100335. [CrossRef]
37. Jiang, X. Bitcoin price prediction based on deep learning methods. *J. Math. Financ.* **2019**, *10*, 132–139. [CrossRef]
38. Munim, Z.H.; Shakil, M.H.; Alon, I. Next-day bitcoin price forecast. *J. Risk Financ. Manag.* **2019**, *12*, 103. [CrossRef]

39. Huang, J.-Z.; Huang, W.; Ni, J. Predicting bitcoin returns using high-dimensional technical indicators. *J. Financ. Data Sci.* **2019**, *5*, 140–155. [CrossRef]
40. Shen, Z.; Wan, Q.; Leatham, D.J. Bitcoin Return Volatility Forecasting: A Comparative Study of GARCH Model and Machine Learning Model. *J. Risk Financ. Manag.* **2019**, *14*, 337. [CrossRef]
41. Mangla, N.; Bhat, A.; Avabratha, G.; Bhat, N. Bitcoin price prediction using machine learning. *Int. J.l Inf. Comput. Sci.* **2019**, *6*, 318–320.
42. Siami-Namini, S.; Namin, A.S. Forecasting economics and financial time series: ARIMA vs. LSTM. *arXiv* **2018**, arXiv:1803.06386.
43. Pichl, L.; Kaizoji, T. Volatility analysis of bitcoin. *Quant. Financ. Econ.* **2017**, *1*, 474–485. [CrossRef]
44. Indera, N.; Yassin, I.; Zabidi, A.; Rizman, Z. Non-linear autoregressive with exogeneous input (NARX) Bitcoin price prediction model using PSO-optimized parameters and moving average technical indicators. *J. Fundam. Appl. Sci.* **2017**, *9*, 791–808. [CrossRef]
45. Razzaq, A.; Liu, H.; Zhou, Y.; Xiao, M.; Qing, P. The Competitiveness, Bargaining Power, and Contract Choice in Agricultural Water Markets in Pakistan: Implications for Price Discrimination and Environmental Sustainability. *Front. Environ. Sci.* **2022**, *10*, 670. [CrossRef]
46. Rosenblatt, F. *Two Theorems of Statistical Separability in the Perceptron*; United States Department of Commerce: Washington, DC, USA, 1958.
47. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]
48. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1999.
49. Hansen, L.K.; Salamon, P. Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 993–1001. [CrossRef]
50. Wright, S. Correlation and Causation. *J. Agric. Resour.* **1921**, *20*, 557–585.
51. Hyndman, R.J.; Koehler, A.B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [CrossRef]
52. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 26.

entropy

MDPI

Review

Blockchain Technology, Cryptocurrency: Entropy-Based Perspective

Feng Liu [1,2,*], Hao-Yang Fan [3] and Jia-Yin Qi [2,*]

[1] School of Computer Science and Technology, East China Normal University, Shanghai 200062, China
[2] Institute of Artificial Intelligence and Change Management, Shanghai University of International Business and Economics, Shanghai 201620, China
[3] School of Statistics and Information, Shanghai University of International Business and Economics, Shanghai 201620, China; 21349074@suibe.edu.cn
* Correspondence: 52205901024@stu.ecnu.edu.cn or lsttoy@163.com (F.L.); ai@suibe.edu.cn or qijiayin@139.com (J.-Y.Q.)

Abstract: The large-scale application of blockchain technology is an expected to be an inevitable trend. This study revolves around published papers and articles related to blockchain technology, relevance analysis and sorting through the retrieved documents with six core layers of blockchain: Application Layer, Contract Layer, Actuator Layer, Consensus Layer, Network Layer and Data Layer. Based on the analysis results, this study found that China's research is more towards the preference and application of landing and industry and smart cities with blockchain as the underlying technology. International research is more focused on the research of finance as the underlying technology of blockchain and tries to combine crypto assets with real industries, such as crypted assets and payment systems for traditional industries. This paper studies the impact of monetary entropy on cryptocurrencies in smart cities and uses the monetary entropy formula to measure the crypto-economic entropy. We use Kolmogorov entropy to describe the degree of chaos in the cryptocurrency market in a smart city. The study illustrates the current status of blockchain technology and applications from the perspective of cryptocurrency in a smart city. We find that smart cities and cryptocurrencies have a mutually reinforcing effect.

Keywords: blockchain technology; cryptocurrency; Kolmogorov entropy; DAO; metaverse

Citation: Liu, F.; Fan, H.-Y.; Qi, J.-Y. Blockchain Technology, Cryptocurrency: Entropy-Based Perspective. *Entropy* **2022**, *24*, 557. https://doi.org/10.3390/e24040557

Academic Editors: Stanisław Drożdż, Jarosław Kwapień and Marcin Wątorek

Received: 21 March 2022
Accepted: 13 April 2022
Published: 15 April 2022

1. Introduction

Since Nakamoto published 'Bitcoin: A peer-to-peer electronic cash system' in 2008, the development of Bitcoin has been up and down, but its underlying blockchain technology has received more and more attention in recent years. In 2019, China decided to take blockchain as an important breakthrough for independent innovation of core technologies, and to accelerate the development of blockchain technology and industrial innovation in smart cities with blockchain technology. The main application scenarios could be found in areas like people's livelihood services, urban governance, industrial economy, and ecological livability. The most eye-catching international use of blockchain technology is the rapid development of cryptocurrencies. This study will analyze the development direction of cryptocurrency from the perspective of world political economy, Bitcoin futures ETF, NFT, DAO, Web3.0. At the same time, along with the rising inflation in the United States and the impact of the COVID-19, the American public's attention to encrypted assets has increased and global encrypted assets have reached the level of trillions of dollars. The international application of blockchain is more focused on the financial field and focused on the virtualization industry. This study uses two scenarios of smart city and cryptocurrency to correspond to China's and other countries' concerns, respectively, to describe the development of blockchain technology in China and other countries, and they can get useful inspiration by comparing different development paths to promote the

development of industry and blockchain technology. This paper uses entropy to analyze the performance of cryptocurrency. The cryptocurrency market and smart city represent a chaotic state, and we introduce Kolmogorov entropy to measure the degree of chaos, which is the direction of our future research. The study is a qualitative study, and we will do more quantitative studies in the future to measure specific monetary entropy, crypto-economic entropy, urban entropy, Kolmogorov entropy etc.

The rest of this article is: Section 2 Entropy in cryptocurrency markets; Section 3 Study of blockchain technology; Section 4 Solutions about blockchain technology; Section 5 Cryptocurrency and entropy; Section 6 Discussion; Finally, Section 7 Concludes this work.

2. Entropy in Cryptocurrency Markets

Thermodynamics-related theories in physics have driven the development of economics in the past 100 years, especially when the theory of 'entropy' in thermodynamics emerged, which greatly promoted the development of modern economic theories. For example, entropic economics, complex economics, and quantum economics, etc. This shows that 'entropy' has bridged the gap between economics and physics and has had a dramatic impact on mainstream economic theory. The concept of 'entropy' originated in the 19th century, and first indicated that part of the energy of a steam engine could not be transformed into useful work due to friction and other reasons, and 'entropy' measured the missing energy in this part. The first mathematical definition of 'entropy' is shown below [1].

$$C = \frac{1}{T}q \tag{1}$$

$$\Delta c = \left(\frac{1}{T_2} - \frac{1}{T_1}\right)q \tag{2}$$

where Δc represents the change in entropy and q is the heat transferred from an object with temperature T_1 to other object with temperature T_2. A slightly different definition of entropy, being a measure of the molecular disorder of the system, was formulated by Boltzmann. It has the following form:

$$C = K\ \ln(m) \tag{3}$$

where K is the Boltzmann constant, while m is the number of microscopic states.

'Entropy' in thermodynamics can be used in the social sciences as a general measure of the disorder in a system. For example, the concept of 'corporate entropy' is used in management and organizational sciences, and should be understood as a loss of productive energy. The entropy in an organization is always growing, just like the thermodynamic entropy in the universe. As the concept of 'entropy' continues to evolve, the various definitions of entropy can be made more specific and applied to specific financial scenarios. John Bryant in his book provides a careful mathematical description of 'entropy' in the economy with the following expression shown in, it has the following form:

$$c = \ln\left(\frac{v}{L}\right) \tag{4}$$

where v represents the volume of economic activity and L represents the constrained level of that activity. The change in the entropy of the economy per unit of time can be expressed in a more precise formula.

$$dC = (w - wn + 1)\frac{dv}{v} = \left(1 + \frac{1-n}{r}\right)\frac{dv}{v} \tag{5}$$

where dv/v represents the growth rate of volume flow, w is the lifetime factor, n represents the elasticity index, and r is the natural rate of return. The factor $(w - wn + 1)$ is called the marginal entropy index, and the integration of Equation (5) yields the entropy generation per unit time using the following mathematical form:

$$c = (w - wn + 1)\ln(v) + c_0 = \left(1 + \frac{1-n}{r}\right)\ln(v) + c_0 \tag{6}$$

Equation (6) can be used to describe the monetary entropy, in which case the rate of return approximates the long-run average or natural level of the velocity of money circulation.

Kolmogorov entropy is an important quantity to characterize chaotic systems. In different types of dynamical systems, the value of K is different, and in systems with chaotic motion, the value of K is greater than zero. the larger the value of K, then the greater the rate of loss of information. the formula for Kolmogorov entropy is shown below [2].

$$k = -\lim_{\tau \to 0} \lim_{\varepsilon \to 0} \lim_{d \to \infty} \frac{1}{d\tau} \sum_{i_1, \ldots, i_d} p(i_1, \ldots, i_d) \ln p(i_1, \ldots, i_d) \quad (7)$$

where $p(i_1, \ldots, i_d)$ is the joint probability and ε and d are fixed values. Equation (7) can portray the degree of chaos of cryptocurrency and smart cities. Samet Gunay andand Kerem [3] showed that cryptocurrency markets are not random but chaotic. In the present, this means that the short-term prediction of the cryptocurrency market may be achievable, but it is completely unpredictable in the long-term prediction.

R. Fistola and R. A. La Rocca [4] studied the measurement of urban entropy, which is a complex system, and divided the urban system into five subsystems, each of which contains several influences that are used to measure the entropy of the city. In addition, it is beneficial to keep the entropy of a city within a reasonable range, but when the entropy is too low or too high, this can lead to a 'fragile' city or reduce the ability to sustain development. Dehouche [5] uses an approximate entropy approach to verify the reasons for the exponential and persistent fluctuations of the bitcoin price, using data such as the daily price of bitcoin, the price of gold, and the SandP 500 index, and calculating their standard deviations. Pele, DT and Marinescu-Pele [6] used the entropy of bitcoin's daily returns to predict the daily value-at-risk of bitcoin and demonstrated that entropy outperforms the classical GARCH model, and the following conclusions are drawn: There is a strong positive correlation between the daily log price of bitcoin and the intra-day return entropy, indicating that entropy has predictive power for bitcoin price. Grilli and Domenico [7] introduced the concept of Boltzmann entropy into cryptographic digital currencies and used Boltzmann entropy to predict the price change trend of cryptographic digital currencies.

The entropy of the cryptocurrency market and the traditional currency market will be affected by various factors, as shown in Table 1 such as inflation rate, fiscal deficit level, interest rate volatility, and high cost of currency management in the traditional currency market. For the cryptocurrency market, the basic blockchain peer-to-peer and decentralized technology of cryptocurrency saves a lot of operation and management costs, meanwhile, along with the emergence of NFT, Dao, Web 3.0, metaverse in the cryptocurrency market, they are constantly optimizing the ecological environment of the crypto market.

Table 1. The factors of influencing the entropy values.

	Traditional Monetary Market	**Cryptocurrency Market**
The factors of influencing the entropy values	Inflation rate	NFT
	Fiscal deficit level	DAO
	Interest rate volatility	Web 3.0
	The cost of currency management	Metaverse

In economic reform, 'entropy' can be used as a tool for future monetary reform, in this paper mentioned cryptocurrency and smart city, in which includes a large number of 'entropy', such as: 'monetary entropy ', 'education entropy', 'transportation entropy', 'ecological entropy', etc., when cryptocurrencies appear in the monetary market as well as transforming the urban governance model to 'smart city' mode is transformed, it is in reducing the degree of chaos within the original system, reducing the lost part of the system operation, improving the operation efficiency, and then introducing 'entropy'

into the currency market and city construction for the global crypto market and city governance model.

In this paper, we introduce the concept of Kolmogorov entropy to smart cities and cryptocurrency, and use Kolmogorov entropy to measure the degree of disorder in the monetary market of smart cities as a way to speculate whether the smart cities are developing stably.

3. Study of Blockchain Technology

This study on blockchain technology focuses on the six layers of the blockchain, which are: Application Layer, Contract Layer, the Actuator Layer, Consensus Layer, Network Layer and Data Layer. Detailed technical research of studies is shown in Table 2.

Table 2. Different science and technology concerns for each layer at China and outside China.

Blockchain Architecture		Research in China	Research Outside China
Application layer	Programmable currency	Tokenized Open Finance [8]	Investment [9]
		Trading account traceability [10]	Payment [11,12]
			Anti-money Laundering [13]
			Identity [14]
	Programmable finance	Trading System [15]	Business Economy [16]
		E-commerce platform [17–19]	Finance [20,21]
		Sharing Economy [22]	
	Programmable society	Supply Chain [23]	Social Governance [24,25]
		Social Governance [26–34]	Identity management [35]
		Education [36–40]	Business Process [41]
		Taxation [42–45] Medical [46] Rights Protection [47] Intellectual Property [48,49]	
Contract layer	Smart Algorithm	State Machine Network [50] Domain State Machine Vector clock state machine	Business process modelling notation [51] Reparo Protocal [52]
	Smart Contract	Double auction pricing contract	E-Bidding [53]
	Smart Scripts	Smart scripts incorporating machine learning [54]	Picture Hash [55]
Incentive layer	Game model	Reputation model based on multiple games	Reputation incentive model [56]
	Incentives	Stochastic equity proof mechanism [57]	Authoritative participation and internal drive incentive mechanism [58]
Consensus layer	PoS DPoS	Multi-group proxy consensus mechanismMG-DPoS	Improving consensus in proof of stake protocols [59]
	PoW	estPoW	Proof-of-Trust (PoT) [60]
		Zero-determinant strategy based on game theory [61]	Randomly Elected Blockchain (REBC) [62]
	PBFT	Dynamic consensus mechanism	Open Business Environment-PBFT [63]

Table 2. *Cont.*

Blockchain Architecture		Research in China	Research Outside China
Network layer	P2P Network	Network layer node configuration trading subsystem [64]	IoT trust management [65]
	Spread mechanism	Blockchain network protocol based on IPv6 [66]	Dual-Channel Parallel Broadcast model (DCPB) [67]
		Trusted service quality model [68]	
	Authentication mechanism	SDN control layer security mechanism construction method [69]	Network is under the influence of such attacks [70]
Data layer	Data Block Asymmetric encryption	BigchainDB [71]	Health data based on an extension of permissioned blockchain [72]
		Block chain data safe transmission method based on SCTP protocol [73]	Graphic data encryption solution based on private blockchain
	Chain Structure	Two-layer blockchain user trust negotiation model [74]	Complex networks modelling framework [75]
	Timestamp Hash Merkle Tree	Attack detection model based on Merkle hash tree structure [76]	Data Structure of Streaming Authentication Based on Merkle Tree [77]

For the Application Level, the China's studies focus on social industries such as e-commerce, education, taxation, medical care, intellectual property and social governance. This paper [78] believes that the introduction of blockchain, edge computing and other technologies under the distributed architecture computing network can provide digital power for smart earth applications. International studies focus more on the economic and financial fields such as investment, payment, identification, anti-money laundering, business process re-engineering, finance, etc. For the Contract Layer, the Chinese studies focus on tying up the industrial chain through the contract state machine, to implement its designed smart contract into the industry. This paper [79] studies a new type of decentralized threshold signature protocol. By combining distributed the key generation protocol and BLS signature, a set of threshold signature protocol with fixed signature length that can be participated by multiple parties is designed. The international studies focus more on the blockchain smart contract itself, through the design of the contract layer protocol to serve the blockchain network. For the Actuator Layer, both Chinese and international studies focus on the design of reputation-based game model. In terms of preferences for technical scenarios' landing. This paper [80] proposes a two-party elliptic curve digital signature algorithm suitable for blockchain. Through the mathematical logic of the given signature algorithm and its security model, it is integrated into the blockchain for evaluation, which proves the feasibility of the scheme. Chinese studies focus on the integration of actuation and data sharing scenarios, while international studies focus on the blockchain network improvements. For the Consensus Layer, both Chinese and international studies prefer the improved consensus algorithm for existing PoW, (D) PoS and PBFT. Few new consensus algorithms have been proposed. For the Network Layer, China preferences is for efficiency improvements through external software and services, while the international preference is more focused on technical improvements to the blockchain network layer. For the Data Layer, Chinese preference is to apply data in the security field, while the international focus is more on research and improvements of the blockchain data layer, such as data structure, data expansion and chain structure improvement. For example, in the case of ensuring data security and credibility with the help of the blockchain double-chain structure, the endogeneity is improved [81]. The problem of low efficiency of data interaction between middle and platform, etc.

To sum up, based on the six levels of analysis mentioned above, we can clearly see the similarities and differences between Chinese and foreign research directions: At the Application Level, China's attention is paid more to the application of actual scenarios, while international attention is paid to the application in the financial field; at the Contract Layer, China's focus on the application of smart contracts in the industrial chain, while the international research on the contract itself has been strengthened; in Actuator Layer, both Chinese and international studies focus on the design of game model; in Consensus Layer, both are researching How to improve the consensus mechanism; in Network Layer, China attaches great importance to external software and services, while the international community pays more attention to improving the blockchain network itself; In Data Layer, domestic data is used more in the field of security, while the international community pays more attention to the research of the data layer itself.

4. Solutions about Blockchain Technology

Based on the research differentiation in China and internationally, we can condense the following table conclusions on the six-layer structure of the blockchain, as shown in Table 3.

Table 3. Differences Between China and International Research using blockchain technology.

Blockchain Technology	Differences between Chinese and International Research	
	In China	**Outside China**
Application Layer	Focused on the real economy	Focused on the virtual economy
Contract Layer	Linked to the industrial chain	Blockchain contract protocol improvements
Actuator Layer	Reputation-based game model design	
Consensus Layer	Improved consensus algorithm based on existing PoW, PoS, DPoS, PBFT	
Network Layer	Improvements through external software and services	Improvements of the network layer itself
Data Layer	Application to Security	Improvements of the data layer itself

By analyzing the domestic and international research of blockchain, we can make relevant references. Considering that Chinese blockchain research is more linked to the industrial chain and rooted in the real economy applications. In terms of the core algorithms, we can learn more from international research of improvements of the contract, consensus and network layers. We can also take full use of China's achievements in the security and privacy design. Thus, by combining the advantages of both to achieve and implement the whole.

5. Cryptocurrency and Entropy

Cryptocurrency is one of the most important scenarios for the application of blockchain in the smart city, and with the development of DEFI and cryptocurrency, which the relationship will become closer and closer, and cryptocurrency may play an important role in the international financial market of the smart city. As shown in Figure 1, which analyses the factors that affect the development of cryptocurrencies.

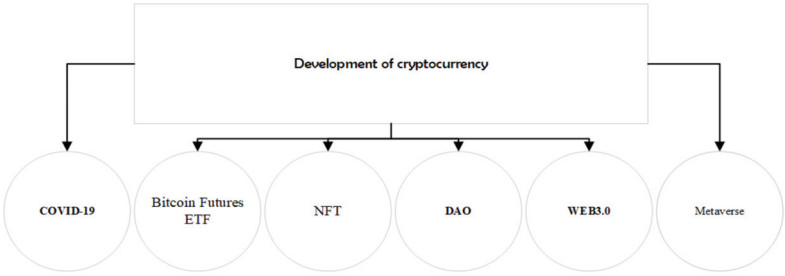


Figure 1. Development of cryptocurrency.

5.1. The Global Epidemic and Inflation Will Accelerate People's Interest in Encrypted Assets

The COVID-19 epidemic in 2020 and the recession that came with it hit individuals, small businesses and governments in a short period of time, but the consequences of this are also complex, providing opportunities for innovation on the one hand, causing social unrest and economic hardship on the other hand. In terms of monetary policy, the Federal Reserve is moving constantly and its balance sheet is expanding at a very fast pace. The Fed's fiscal deficit reached $3.1 trillion in 2020 and is expected to reach $3 trillion in 2021, creating insecure expectations for the population. The rapid changes in society as a whole during this period forced investors to react accordingly and new investment trends are developing. Based on this, our survey shows a significant increase in interest in safe-haven assets compared to the results of 2019, and among the safe-haven assets are cryptocurrencies such as Bitcoin, a phenomenon that is within the forecast range. During the COVID-19 epidemic in 2020, which set back the entire traditional financial market. Cryptocurrencies played a role as a hedge against risk, with Bitcoin being ideal for risk averse individuals in the face of downward pressure in financial markets [82]. Further studies found that bitcoin in both domestic and cross-border entities served to achieve diversification benefits and risk mitigation, acting as a 'safe haven' [83–85]. The cryptocurrency market shows higher levels of cross-correlations with the others during the COVID-19 periods, in which it is strongly cross-correlated itself [86]. From these results, it can be concluded that the COVID-19 promoted bitcoin investments, 63% of the bitcoin investors were influenced by the COVID-19 epidemic in the past year, and it has boosted the price of bitcoin [87].

Based on the analysis of the socio-political and economic status quo, the traditional financial market is vulnerable to severe shocks, and the market is unstable. The Fed accelerates money printing and inflation, and the issue of asset preservation has received more and more attention. Cryptocurrency has quickly attracted attention due to its excellent properties such as security, limitedness, and easy liquidity. Once investors make profits in the encrypted market, they will attract more investment. This is based on the external unstable financial environment make cryptocurrency investment market boom.

5.2. Bitcoin Futures ETF

Several Bitcoin ETF applications have been rejected since 2013, citing market manipulation, fraud and failure to protect consumers. On 19 October 2021 at 9:30 am EST, the SEC approved the issuance of a bitcoin futures ETF, which will probably lower the investment threshold for bitcoin after the adoption of the bitcoin futures ETF. With the caveat that the purchase of a bitcoin futures ETF is an indirect investment in bitcoin, not a direct investment. The purchase here is simply a futures contract, which is less effective at tracking the price of bitcoin. For better tracking of the price of bitcoin, a spot ETF would be a better option, but is not available at this time as the liquidity of bitcoin may not be sufficient for the liquidity required by institutions. Bitcoin futures ETFs are regulated to track the spot bitcoin price using CME regulated futures contracts, which can be regulated by traditional exchanges for added security. However, there are two disadvantages to bitcoin futures ETFs: one is that futures have a premium and discount problem that if the bitcoin price

changes by 1% in an hour, the bitcoin futures ETF may drop by 2%, which is a significant deviation, meaning that futures ETFs do not track the bitcoin price well enough; the other is that investment costs are high, as futures contracts have an expiry date, and near the expiry date. This is the cost of shifting the futures ETFs, which can eat up some of the profits, so bitcoin futures ETFs are not suitable for long-term holdings, and can be used for short-term investments. In reality, Bitcoin also requires the retention of an associated Bitcoin wallet as well as registration with a cryptocurrency exchange, which remains unknown to those unfamiliar with the space and requires a degree of self-education. With a Bitcoin futures ETF, investors do not need to worry about private keys, storage or security. They own shares in the ETF, just like their shares, and have access to the cryptocurrency market without having to buy and hold cryptocurrency. Bitcoin futures ETFs are managed by companies that buy and hold actual bitcoins, with prices linked to the bitcoins held in the fund. The company lists the ETF on a traditional stock exchange and we, as investors, can trade the ETF just like any other stock.

In conclusion, for investors who are reluctant to invest in bitcoin directly due to price, security and regulation. buying a bitcoin futures ETF is an opportunity to invest in cryptocurrencies, which adds to the focus on cryptocurrencies in the traditional financial markets, which is why I believe that bitcoin futures ETFs are an important option for some investors in the future.

5.3. NFT

The NFT market has been very popular in the past two or three years. The concept of NFT originated from Ethereum. NFT is a form of cryptocurrency [88]. NFT can also be regarded as the ownership certification for identifying virtual assets. They can be largely subdivided into several small parts. NFT is a subdivision that cannot be carried out. Such tokens can be bound with virtual or digital assets to form their unique identifiers. After forming NFTs, they can be freely traded in the market, which greatly promotes the development of virtual currencies and stimulates the prosperity of the DAPP market. By using NFT on smart contracts, creators can easily prove ownership including 'picture', 'video', and 'artwork', so that intellectual property rights can be properly protected. NFT no matter the transaction many times, its original creators have drawn royalties, which has also spurred the development of digital artworks in disguise. The real-time data as of the time when the author wrote the paper shows that the current total transaction value of NFT are close to 16.4 billion US dollars as shown in Figure 2 (Data from: nonfungible.com (accessed on 30 December 2021), and its growing market and large returns have attracted people's attention.

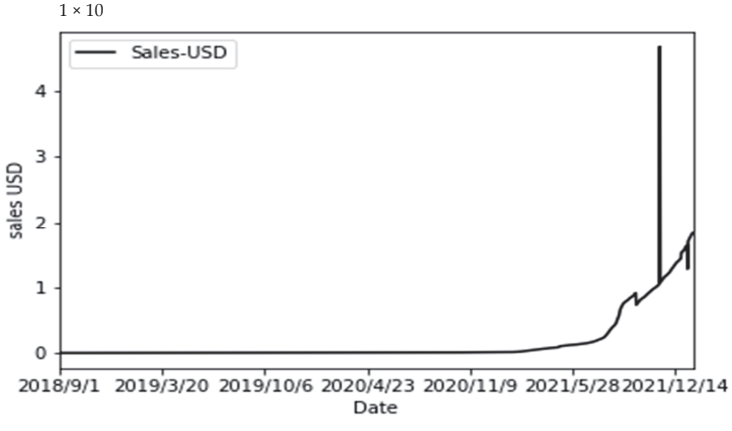


Figure 2. Total NFT transaction volume.

In the past 24 h, the transaction volume of NFT was US$49,858,590.46 [89], and the transaction volume of cryptocurrency was US$102,099,049,616 [90]. The transaction volume of NFT accounted for about 0.05% of the transaction volume of cryptocurrency. It can be seen that NFT is still developing at the initial stage of 2020, but there is a significant increase compared to the same period in 2020, and its growth can also be seen from other indicators. In April 2020, the number of initial transactions was 25,729, the number of secondary transactions was 8589, and the current number of initial transactions is 14,852,246 times, and the number of secondary transactions was 13,359,460 times. In addition, the active level of investment can also be seen from the number of active wallets. As shown in Figure 3 (Data from: nonfungible.com (30 December 2021), the number of wallets investing in NFT-related products in the past three months has increased from about 97,000 to about 200,000 now, which also reflects given the latest investment trends in cryptocurrencies, NFT-related products will continue to attract a large influx of cryptocurrencies.

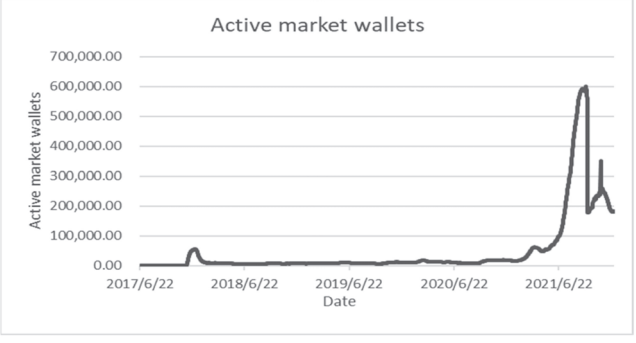


Figure 3. The number of active wallets related to investing in NFTs.

In addition, the concept of the 'metaverse' rises in 2021, and 2021 is also known as the first year of the 'metaverse'. With the renaming of Facebook and the listing of the Roblox game company, the concept of the Metaverse instantly aroused thousands of waves, leading to the influx of other Internet giants, which also led to the accelerated development of the industry and the entry of capital, and it would quickly integrate human resources. This further stimulated the market to have a high valuation for the Metaverse, and finally formed a market outlet industry. This way, investment and speculation in the market will come one after another. Now the total value of the Metaverse-related industries has reached about 714 million USD. In the future, running the metaverse will need to include a large number of NFTs, cryptocurrencies and NFT-related products will be used in the future virtual world. It can be used, leased and bought and sold in the metaverse, which requires a large amount of cryptocurrencies to promote the market's investment interest in cryptocurrencies.

In a word, NFT will enhance the development of the entire encrypted market [91], such as: existing encrypted games; one can continue to buy and sell NFT products in the game; protect digital collectibles, protect digital assets and intellectual property rights; finally, prosper the entire metaverse. From the current point of view, NFT-related products are similar to the process of re-enclosure. Redistributing resources and wealth will attract a large number of people to participate, which also requires the use of cryptocurrencies, and it will be one of the investment directions of cryptocurrencies in the future.

5.4. DAO

DAO is a decentralized autonomous organization. The decision of the organization is made by the group. The organization will also issue its own tokens (but not necessarily every organization will issue it), some of tokens held or those participating in the DAO. The project, which determines the size of the voting power, DAO can be a company

where everyone is the boss, both consumers and owners, which actually shortens the distance between the market and the organization, and through DAO efficient voting, decision-making. It can shorten the mechanism of decision-making and execution, thereby improving the efficiency of the organization. The business model of DAO can be very promising in the future. NFT is used to illustrate that the current NFT market has a phenomenon of oversupply and lack of development potential. If DAO appears at this time, its development can be promoted. For example, Party-DAO, which focuses on investing in NFT projects, its goal is to gather the power of retail investors to obtain funds through crowdfunding to buy NFTs. After the purchase is successful, the internal tokens will be distributed proportionally to record the NFT purchase share. If you want to sell NFTs, you need to initiate a vote within the organization. If the support exceeds 50%, a public auction will be held, and the proceeds will be divided equally according to the proportion. This not only promotes the trading of NFTs, but also promotes more cash flow into the cryptocurrency market, so DAO is a "good medicine" to improve the entire decentralized finance.

As shown in Table 4 DAO companies are divided into eight categories according to their functions and business. DAO-operating systems: The main task is to help users create DAO; Investments DAOs: To raise funds to invest in projects recognized by DAO internal members, if there is profit, it will be divided proportionally, otherwise the risk will be shared; Grants DAO: If you want to improve the internal DAO, you need to initiate a proposal as a member. If the proposal is passed, you can get a bonus to complete your own ideas; Collector DAOs: It belongs to the collection DAO, that is, only invests in NFT projects, and the profits are divided proportionally, otherwise you need to share the risk; protocol DAOs: Mainly do cryptocurrency lending business on the blockchain; Service DAOs: Provide services to DAO, such as fundraising to buy NFT, or analyze data; Social DAOs: Mainly do online discussions and interactions; Media DAOs: Blockchain news.

Table 4. DAO Landscape around world.

DAO Landscape	Examples
DAOoperating systems	DAOStack, colony, Orca etc.
Investments DAOs	theLao, BitDao, Metacartel etc.
Grants DAO	Uniswap Grants, Aave Grants etc.
Collector DAOs	PleasrDao, herstoryDAO etc.
protocolDAOs	Curve, AAVE, Sushi etc.
SerivceDAOs	partyDAO, metaverseDAO etc.
SocialDAOs	FWB, seedclub etc.
Media DAOs	GCR etc.

If you want to join any of the above organizations, you need to hold some of the cryptocurrencies specified by the organization or participate in the project together. These are two main ways to join the organization. No matter which method is used, it is inseparable from cryptocurrencies. It is also an important choice in investing in cryptocurrencies.

5.5. Web 3.0

Web3.0 is a newly proposed technical means; it is based on the blockchain technology to develop the next generation of a brand-new and efficient network world. In the era of web 1.0, that is, the early Internet, you could only browse information related to web pages; In the era of web 2.0, mobile phone terminals appeared, allowing users to interact with the platform a lot. Absolute management and supervision rights; In the future web3.0 era, there is no central agency for review, smart contracts are cornerstone of web3.0 operation, users have absolute control over the privacy of data, and users can combine their own actual situation. Data are sold for profit. Web3.0 is not a direct invest cryptocurrency; it is more like participating in the maintenance of the network together. As a part of it, the cryptocurrency reward obtained by contributing one's own strength, which will

mobilize the power of global netizens to jointly the future network world created. Although cryptocurrency investment is not directly reflected here, as long as you participate in it, you can get the corresponding governance token rewards, which is another way to invest.

For example: 'brave' browser, which now has 42 million users worldwide, is based on the concept of web3.0, and its focus on privacy protection, which will not be interfered by any advertisements. One can also choose to watch advertisements, you will be rewarded with bat tokens. This can form a win-win situation for users, advertisers and the platform.

The emergence of Web3.0 will be of significant epoch-making, it will change the operation mode of the entire encrypted economy and the development and profit model of the entire Internet company. Based on blockchain technology in the era of web3.0, every user's attention has been mobilized. This way of operation will also be more private, secure, and efficient.

5.6. Metaverse

In the science fiction "*Snow Crash*" published in the 1990s, a virtual world was constructed, which is a concept of "metaverse" in today's view. In 2021, the concept of "metaverse" will rise. Metaverse' first year. There are many definitions of the metaverse now, it is necessary to build a virtual world with a strong sense of experience. The future Metaverse will not disappear with the collapse of an Internet company, and the economic system in the metaverse will not disappear. To be interconnected with the economic system of the real world, more and more internet giants are optimistic about the development potential of the Metaverse. With the renaming of Facebook and the listing of the Roblox game company, the concept of the Metaverse instantly aroused thousands of waves. Leading to other Internet Giants are also pouring to speed up the layout of the Metaverse, which will lead to the accelerated development of the industry and the entry of capital, and it will further accelerate the integration of manpower and resources, which will further enhance the market's value expectations for the Metaverse, and then form a market outlet industry. In this way, investment and speculation in the market will continue to flow, and the total value of the industry related to the Metaverse has now reached about 714 million US dollars. In the future, running the metaverse will probably need to include a large number of NFTs, cryptocurrencies and NFT-related products will be used in the future virtual world. It can be used, leased and bought and sold in the metaverse, which requires the use of a large number of cryptocurrencies to promote the market's investment interest in cryptocurrencies.

Since many companies do not currently have data related to each other, the entire data will be integrated in the future metaverse. This step may become a revolution in the "interface" between the real world and the digital world. Since the current metaverse is still in the early stage of development, there will be huge changes in the future. Metaverse-related projects are also being explored in the industry. At present, NVIDIA has created the NVIDIA Omniverse platform, which can use digital twin technology to build a virtual factory in a virtual world, and real products can be tested. Its data can be Synchronized with the real factory, which can liberate productivity, reduce costs and increase efficiency.

The metaverse must bring about drastic changes, which require continuous innovation of a large number of technical means. At the same time, virtual currency is booming, which also promotes the development of the virtual world. However, since the construction of the metaverse and the entire virtual world is still in its infancy, it should be subject to the supervision of relevant international organizations. It is possible to explore the use of legal digital currency as the transaction currency of the metaverse.

Based on the monetary entropy formula, this study gives the concise measurement Equation (8). The six influencing factors, such as inflation, Bitcoin futures ETF, NFT, DAO, Web 3.0 and Metaverse, are measured separately in Equation (8). Ce_i represents the entropy value of the 6 influencing factors, w is the lifetime coefficient, N represents the elastic index, V stands for the volume of economic activity, C_o is a constant.

$$Ce_i = (W_i - W_i N_i) \ln(V_i) + C_o \tag{8}$$

Equation (9), Ce represents the measurement of the overall development of crypto-economy in smart cities under the combined effect of six factors such as NFT, DAO, WEB3.0, etc. After we get the crypto-economy entropy, we can infer the financial market operation in the smart city.

$$Ce = C(Ce_1 \dots Ce_n) \tag{9}$$

After analyzing the changes in the world's political economy, NFT, DAO, WEB3.0, bitcoin futures ETF, Metaverse, it can be seen the entire traditional financial market is changing, online world is also undergoing restructuring. At the same time, with the development of communication technology, it will reshape a new economy patterns and social formations. Blockchain technology enables everyone to have the opportunity to integrate into economic and social development. However, because of it is still in the early stage of development and the development rules are not perfect, it is necessary to improve supervision capabilities to meet the challenges of the ever-changing encrypted economy in the future of smart city.

6. Discussion

In this paper, we introduce a smart city as a scenario and use Kolmogorov entropy to calculate the degree of chaos of cryptocurrency and smart city. Smart cities applying cryptocurrency can introduce monetary entropy, which may be one of the future research directions. In addition, smart cities can also use entropy to measure the development of the whole urban system. The following are the main development directions of smart cities.

The construction of smart cities is the development direction of major cities in China in these years, along with the rapid development of information and communication technology (ICT) such as artificial intelligence, 5G, blockchain and Internet of Things, which make the technical means of building smart cities more and more perfect, but at the same time, cities generate a large amount of data and huge data transmission tasks, which leads to the problem of information security, and blockchain technology can rebuild trust in society and solve the sharing problem. However, the application of blockchain technology in smart cities is a difficult system project, requiring the construction of a secure and credible smart city data infrastructure, a city-level multi-level blockchain public service platform, and a breakthrough with a focus on government service innovation to promote the gradual implementation of blockchain applications. In the subsequent development of blockchain technology applications in smart cities, importance should be attached to the construction of blockchain underlying architecture and infrastructure, putting the building of blockchain basic service platforms in the first place, deep integration with technologies such as big data, artificial intelligence and the Internet of Things, and further improvement of regulatory and standard systems. What we want to solve the openness, exchange, integration, sharing and security of data resources that are isolated and segregated because of category, industry, sector and geography. Based on the analysis of blockchain technology in Part II III, as shown in Figure 4, we can divide the development of smart cities into four directions—people's livelihood services, urban governance, industrial economy and ecological livability.

6.1. People's Livelihood Services

The main application scenarios of blockchain technology in the field of livelihood services include: smart healthcare and smart education, as shown in Figure 5.

In recent years there have been numerous problems in the medical field and the masses have not been able to get better solutions to their medical problems. The use of blockchain technology can-to a certain extent, improve the current system construction in the medical field-enhance its efficiency and further promote the application of Internet+ medicine. As an important industry application scene, the medical field has seen a year-on-year increase in the proportion of major enterprises, government departments and investment institutions in China, and outside China making strategic investments in its layout, constantly accelerating the application of blockchain technology in the medical

industry. Smart healthcare which includes electronic health cases, information sharing, anti-counterfeiting of drugs, and digital currency payments. We use technology to promote the development of medical informatization and ensure the storage and sharing of medical data, which includes all aspects of hospital and patient information. For example, policy data storage and sharing, medical and health records storage on the chain, etc.

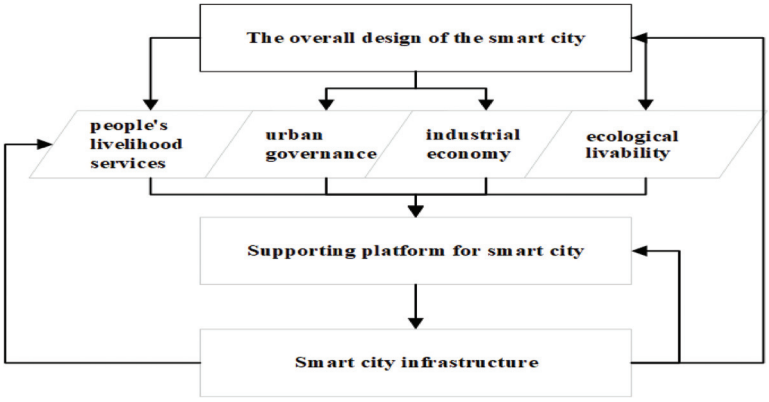


Figure 4. Smart city construction based on blockchain technology.

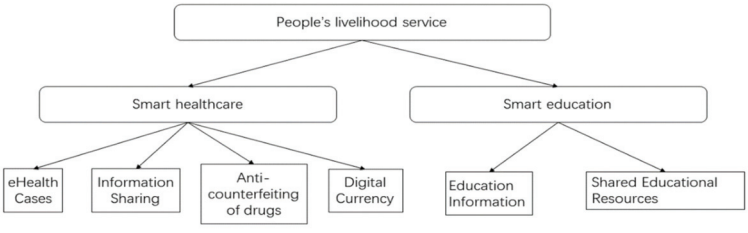


Figure 5. People's livelihood service structure diagram.

Education has come a long way in the last 20 years, but there is still a long way to go before education is fully modernized. Technologies like blockchain can accelerate the process of modernizing education. Distributed ledgers, artificial intelligence, and electronic devices are slowly becoming the future direction of choice for educational tools. When blockchain is used in education, blockchain technology can enhance the transparency of education, such as submitting assignments and checking grades and learning progress, and it can improve the motivation of students to learn, and scholarships can be awarded using cryptocurrency. The main focus of smart education is to preserve information data of teachers, students and educational institutions, and to share resources, and to build efficient online learning communities through smart contracts, so that a series of tasks such as uploading, authenticating, flowing and sharing educational resources can be executed automatically, and reducing the cost of sharing resources, improving the efficiency of resource sharing and monitoring the community ecosystem in real time. The new ecosystem of 'blockchain + education' is formed based on the characteristics of blockchain technology such as efficiency and transparency, and it helps to innovate the education industry [92], as shown in Figure 6.

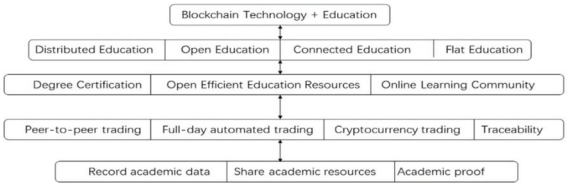


Figure 6. A new ecology of education based on blockchain technology.

Blockchain technology can be used in education as a record for storing distributed learning; it can provide a trusted certificate system for online education; it can be used as a smart contract to complete educational contracts and depository; blockchain technology can be used as a copyright tool to mark academic achievements; and it can be a decentralized global knowledge base, and a knowledge currency [93]. In terms of practical applications, the MT Media Lab (MIT Media Lab) has released 'Blockcerts', a blockchain certificate project, an open standard for digital academic certificates based on the Bitcoin blockchain. 'Blockcerts' provides a decentralized authentication system. Because it relies on the most secure Bitcoin blockchain, its credentials are tamper-proof and verifiable. In addition, 'Blockcerts' can be used to issue any type of credentials, including professional certificates, transcripts, credits, or degrees [94]; the Holberton School, a software education institution in San Francisco, uses blockchain technology to record academic credentials for its schools and will start to share information on academic credentials on the blockchain starting in 2017. In addition to providing evidence of academic achievement, blocks can be used as a basis for measuring an individual's intellectual wealth [95]. By analogy with Bitcoin, a block that records one's academic achievements can also be used as a "knowledge currency". In other words, the concept of "Learning is Earning" is used to promote education [96], and the above knowledge currency will also become a token (cryptocurrency) within the DAO organization, where students can earn cryptocurrency in the form of questions, answers, and posts, which will also motivate students to learn. This also proves the argument that 'learning is earning'.

6.2. Urban Governance

The key application scenarios of blockchain technology in the field of urban governance include: smart government, smart transportation, as shown in Figure 7.

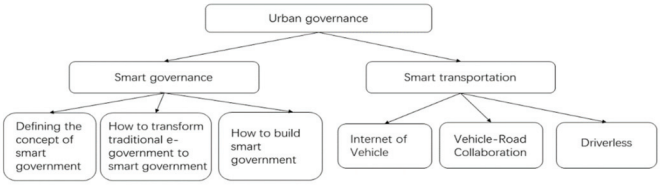


Figure 7. Urban governance structure diagram.

So-called smart government is an e-government based on blockchain, cloud computing and other technologies. In these years, the development of e-government has encountered various problems that still need to be solved, such as low efficiency, data cannot be shared and so on. The first is to define the concept of e-government, which is understood by academics as the use of advanced information technology for collaborative governance of society and the provision of new and efficient services to the public to meet the changing social needs [97,98]. The second, how to develop from traditional government to e-government, which research shows can be developed in steps development and finally reach the level of e-government [99]. The third, how to build smart government specifically, some scholars have proposed a new government service system using GIS and cloud computing technology, which can visualize the operational information of the city [100].

This concept has been expanded, and some scholars have combined the concept of borderless with e-government to form, whose goal is to efficiently meet the needs of the public, means to integrate the process of government affairs across borders, and to collaborate organizations and information across borders weakening the organization of government functions. The goal is to efficiently meet the needs of the public by integrating government affairs processes across boundaries, collaborating organizations and information across boundaries, and weakening the organizational boundaries of government functions so that the entire government functions are truly integrated under the perspective of public services [101]. Borderless wisdom government is supported by blockchain technology, and the information of each department is stored in a distributed manner, and the mutual trust problem is solved between the nodes using consensus mechanisms, which finally forms a borderless government service system and improves the security and efficiency of government services.

Smart transportation, as one of the cores of smart cities, involves many integrated technologies such as Internet of Things, cloud computing, and big data, which enable the coordinated operation of people, vehicles, and roads [102]. With the characteristics of blockchain technology such as polycentricity, security and trustworthiness, and smart contracts, it can realize the construction of a more efficient transportation network, and solve the government, enterprise data sharing and Intelligent management of infrastructure and other issues. On the basis of ensuring open and transparent data circulation, data security is ensured to improve the efficiency of intelligent transportation operation. However, policy failure and control hijacking caused by information security have become major hidden dangers for the promotion and application of new technologies. The application of blockchain in Telematics technology can achieve more secure, reliable data storage and authentication through data encryption and consensus mechanisms. Providing sustainable information services and effective manner, ensuring data security to safeguard telematics technology security. At the same time, through the establishment of alliances and contracts, the effective and seamless integration of information collected by smart transportation field terminals, intelligent vehicle information, manual control commands, and road infrastructure information can be realized. Thus, effectively solving traffic congestion, parking difficulties, and other traffic hotspot problems [103].

Take electric car charging piles as an example; at present, all countries in the world advocate reducing carbon emissions and using more clean energy, and the demand for electric cars is gradually increasing, but the number of charging piles is limited and the configuration is unbalanced. How to let users choose the right one for charging from the limited number of charging piles with uneven space distribution is an urgent problem to be solved. Users can log into the APP to check the distribution of available charging piles and then make a choice according to their own wishes. The whole transaction process is completed by a central processing entity in the background. Such an approach does not consider the variability of individual travelers' needs. Therefore, by introducing the concept of blockchain and adopting a decentralized smart charging contract, we can effectively help users select the most convenient parking/charging location and choose the service completely independently. The specific architecture includes four layers: user layer, in-vehicle information interaction layer, smart contract layer, and target layer [104]. The figure below shows a new management model for new transportation, where multiple types of blockchains are managed collaboratively, as shown in Figure 8. For example, German energy giant Innogy and IoT platform company Slock.it have partnered to launch a blockchain-based peer-to-peer charging project for electric vehicles. Instead of signing any power supply contract with the power company, users can simply install the Share and Charge APP on their smartphones and complete user verification to charge at Innogy's charging posts across Europe, with tariffs automatically determined by a backend program in real-time based on the prevailing and local grid load. Thanks to blockchain technology, the entire charging and tariff optimization process is fully traceable and searchable, thus significantly reducing trust costs. When charging is needed, an available charging station

nearby is found from the app and payment is made according to the price in the smart contract. However, this type of charging is not yet popular, even in Germany, where Ethereum wallets are only an option for some people, but it will still accelerate the connection between cryptocurrencies and real life

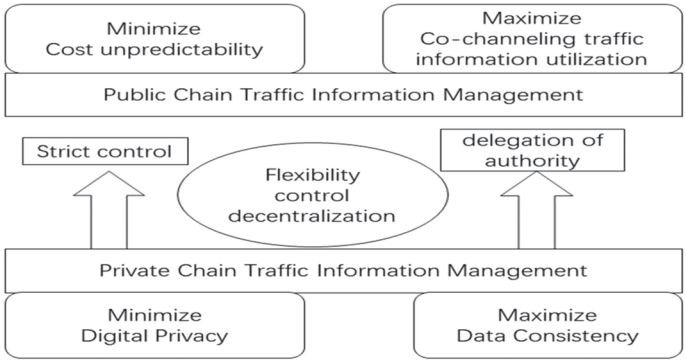


Figure 8. Blockchain collaborative traffic information management framework.

6.3. Industrial Economy

The key application scenarios of blockchain technology in the industrial economy include: intelligent Internet of things, intelligent industry, as shown in Figure 9.

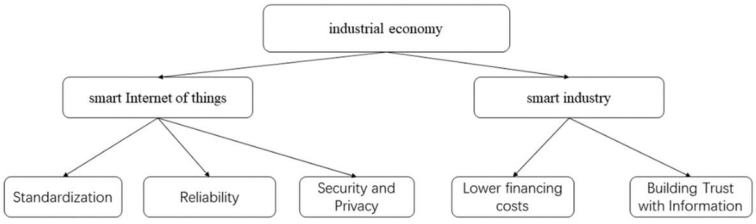


Figure 9. Industrial economy structure diagram.

As an emerging technology, IoT has gradually penetrated into every aspect of life. The security of blockchain technology makes IoT + blockchain be more and more attention. Using blockchain technology can solve the original problems of IoT such as no standardization, information security. Internet of things is an Internet-based and can make all the ordinary physical objects connected to the information carrier. After the development of wired and wireless networks, especially to the 5G era, it can realize the network interconnection of people and people, people and things, things and things. With the development of artificial intelligence, big data, cloud computing and other technologies to promote the development of the Internet of things, in the future, the Internet of things still need to pay attention to the standardization issue. About the Internet of things, there is no standardized construction policy, there is no unified communication protocol, communication interface, etc., there is the same protocol to make the device interconnection, and each company wants to implement their own communication protocol, in order to be the formation of industry development barriers, but now is still in the early stage of industry development. Additionally, the reliability issue. As the current IoT architecture is to aggregate all data into a central control system, all the data have the risk of being controlled and modified at will, and in the process of data transmission may go through multiple links, the authenticity and integrity of the data cannot be guaranteed, and then, security and privacy issues. In the field of IoT, the centralized service architecture stores and forwards all monitoring data and signals through a central server, and a large amount of user data information is stored

in the central server. Although IoT operators keep emphasizing that they can effectively protect user data and privacy, security breaches and privacy leaks still occur, making a large number of users unable to trust that their privacy is secure [105].

During the long-term development and evolution of IoT, the following 9 industry pain points have been encountered: device security, personal privacy, architectural rigidity, communication compatibility and multi-subject collaboration 5 major pain points. The improvements of blockchain technology for IoT are: cost reduction, privacy protection, identity identification, traceability, and cross-subject collaboration. The current development status is: leading IoT companies introduce blockchain technology in large numbers, for example, IBM launched blockchain services for Bluemix cloud platform back in 2016, Amazon chose to start cooperation with DCG, a digital currency company; from the perspective of traditional power companies, they mainly invest in different pilot projects by cooperating with startups, setting up subsidiaries, or even buying startups to create distributed energy systems and peer-to-peer energy trading platforms. Including Sweden's state-run power company VattenFall (Waterfall Power), which invested in a startup (PowerPeers) in Amsterdam, the Netherlands, and to build an energy-sharing platform that allows consumers to freely choose their power channels, and Germany's Rheinland (RWE), which partnered with startup Slock.it to launch 's BlockCharge EV charging project, these platforms have terminal payment systems that support the use of cryptocurrencies, which is also driving the development of cryptocurrencies in practical applications [106]. The main application scenarios: sensor data deposition and traceability; new sharing economy; energy trading; charging cars and charging piles, communication and intelligence for drones. In the following, the new 'sharing' economy is used as an example, as shown in Figure 10. The whole blockchain network is built based on the blockchain. Based on the smart contract system, the asset owner sets the rent, deposit and related rules to complete the binding of various 'locks' with the asset, and the end of user pays the corresponding rent and deposit to the asset owner through the APP to obtain the key and then obtain the right to use the asset. At the end of the use, return the item and get back the deposit, the payment system here will try to use cryptocurrency.

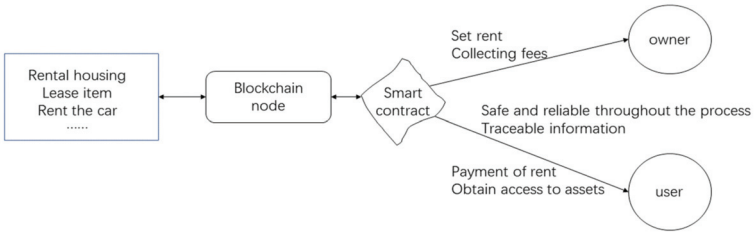


Figure 10. New sharing economy structure diagram.

Smart industry is mainly involved in the supply chain finance with the back ground of blockchain, using blockchain as a way of information transfer and data taking, which can effectively reduce the cost of trust and loan taking. reducing the cost of enterprise financing, focusing on solve the problems of financing difficulties and high capital costs of upstream and downstream SMEs [107]. The project supply chain information platform is optimized by combining blockchain technology in engineering projects, and integrated with project management function information integration and project information collection system to build an engineering project information integration management platform based on smart industry [108]. They are also involved in industrial equipment identity management, equipment access control, equipment registration management, and equipment operation status supervision, etc. The execution of blockchain smart contracts is used to obtain and verify equipment identity, and blockchain technology can guarantee the security of relevant data and ensure that the enterprises of the industrial chain can access credible and consistent equipment operation data. The distributed storage, tamper-evident and

encryption algorithms of blockchain technology can be used to realize data exchange among various entities of the industrial Internet, and the production and manufacturing data of each enterprise can be stored on the blockchain, while realizing the sharing of data of other enterprises.

6.4. Ecological Livability

The key application scenarios of blockchain technology in the field of ecological livability include: smart energy, smart new retail, as shown in Figure 11.

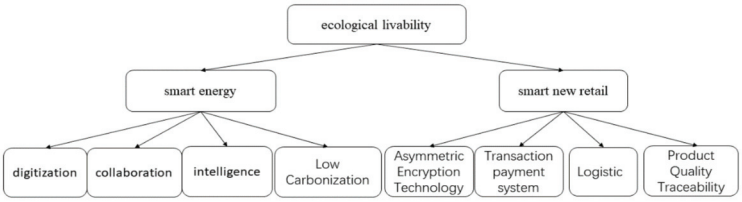


Figure 11. Ecological livability structure diagram.

With the development of industrialization in China, energy, as the support of national life, has received more and more attention in terms of security, efficiency and trust in its development and transaction process. Combining blockchain technology with energy can optimize the traditional energy transaction mode, improve transaction efficiency, and promote the healthy development of energy [109]. Some scholars have elaborated the current situation and challenges faced by blockchain technology in the application of integrated energy system, analyzed the key issues to be solved in the construction of energy blockchain system, and made an outlook on the future development of integrated energy based on blockchain [110], and some scholars have also conducted research on the optimization of energy costs [111,112]. However, there are still many challenges when applying blockchain technology. Firstly, the energy interaction information of the energy system is dynamically changing, and the data throughput is significantly more than the blockchain application scenario of transaction settlement, which causes difficulties for the efficient operation of the system, and even the communication delay and information blockage. Secondly, the consensus mechanism is wasteful of resources for the high energy demand, and it is needed to reduce the energy consumption in the future. The data of the energy production chain is generated and stored locally by each energy vendor, and the data of each energy vendor is disconnected and unconnected. It can be combined with IoT devices to realize safe storage and sharing of production data by applying blockchain technology based on data collection, improving monitoring accuracy, mining data value, and creating an information basis for government supervision.

There are already a large number of use cases in real-world applications, for ex ample, the US energy company LO3 Energy partnered with Siemens Digital Grid in April 2016, and the Bitcoin development company Consensus Systems to create the Brooklyn Microgrid-an interactive grid platform based on the blockchain system the project is the world's first energy marketplace based on blockchain technology. This microgrid project enables peer-to-peer electricity trading for residents between communities, allowing users to access data related to electricity generation and consumption in real time through smart meters and buy or sell electricity and energy to others through the blockchain; Power Ledger was founded in Perth, Australia, by Ledger Assets, an Australian blockchain software company. Power Ledger uses blockchain-based software to build a P2P system for trading surplus solar power to the grid. Unlike the PoW (proof of work) mechanism used by Bitcoin, Power Ledger uses a POS (proof of stake) mechanism, and the blockchain was developed by Ledger Assets and is called Ecochain [113]. The above cases are representative cases, where the Power Ledger platform has modular and scalable features, mainly through three aspects of sustainable energy aspects of business, and any individual module can be extended

on demand, namely energy trading and traceability, VVP model, and environmental commodity trading [114], as shown in Figure 12.

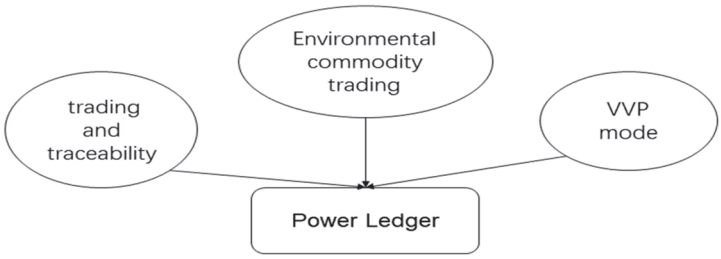


Figure 12. Power ledger module.

The energy trading and traceability product is called xGrid. xGrid allows trading renewable energy across the grid or behind the meter, improving the relationship between customers, retailers, and the distribution network to make it more efficient. Specifically, xGrid supports customers in selling energy from solar panels to other energy consumers connected to the same grid. From generation to consumption, it manages the settlement of energy transactions between the two parties, allowing dynamic price discovery at intervals as short as five minutes. Users can customize their profiles to sell power in their preferred way, and the VPP model software will detect when wholesale market prices are peaking. Managing stored power, helping customers maximize their return on investment while supporting the grid with clean solar energy. The environmental commodity trading product, called TraceX, is a digital marketplace for trading and settling environmental commodities, such as Renewable Energy Certificates (RECs) and Time-based Environmental Certificates (T-EACs).

With the increasing maturity of blockchain technology, this has brought new opportunities for the development of various fields in China. The main problems faced in China's new retail supply chain are also found, namely, transaction payment security, logistics information and commodity traceability. It was shown that the asymmetric encryption technology in blockchain technology achieves the upgrading and optimization of the retail supply chain, while blockchain technology makes it possible to realize the decentralization of transaction payments, thus improving the effectiveness of logistics information and the integrity of commodity traceability [115]. Using the distributed bookkeeping and non-tampering characteristics of blockchain, the supply chain data of commodities are stored in a distributed manner, access rights and encryption verification are set, the data exchange, and sharing in the supply chain process, and the business operation process simplified to realize the overall efficiency improvement and optimization of the supply chain. Using blockchain technology to build a digital supply chain, especially the supply chain management in cross-border trade, the distributed bookkeeping, and tamper-evident characteristics of blockchain are used to distribute the supply chain data of commodities for storage, authorized access and encryption verification, simplifying the data exchange and sharing in the supply chain process and the business operation process. Realizing the overall efficiency improvement and optimization of the supply chain.

The above four aspects do not cover all of the smart city, but can give a direction to the blockchain-based urban governance. It can be seen that when the industry is combined with blockchain it can bring security and efficiency improvement, moreover, it can reduce the cost of governance.

7. Conclusions

This study analyses the development direction of a smart city based on blockchain technology and cryptocurrency and uses a literature review to compare the six areas of blockchain in China and other countries. We included the hot spots and research preferences

of the application layer, contract layer, consensus layer, incentive layer, network layer and data layer of the blockchain, and found that China's research is more preference application of landing and industry, while international research is more preferred to the research of the underlying technology of finance and blockchain. Based on the analysis of cryptocurrencies and smart city, cryptocurrencies will get more market support in the future, and currently 'metaverse', 'DAO', 'NFT', 'Web 3.0' are developing rapidly, cryptocurrencies may be combined with smart cities to reshape the whole financial industry and network world. This paper introduces the concept of Kolmogorov entropy to cryptocurrencies and smart cities to measure the level of disorder within the system. This study is a qualitative study, and we will do more quantitative studies in the future to measure specific monetary entropy and Kolmogorov entropy, crypto-economic entropy, urban entropy, etc. The study may help scholars interested in blockchain to learn the basic knowledge of blockchain, and at the same time helps to understand the global hotspots of entropy and the research direction of entropy.

Author Contributions: Conceptualization, F.L. and J.-Y.Q.; methodology, F.L.; investigation, F.L. and H.-Y.F.; resources, F.L.; data curation, F.L.; supervision, F.L. and J.-Y.Q.; project administration, J.-Y.Q.; funding acquisition, J.-Y.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by 2019, Digital Transformation in China and Germany: Strategies, Structures and Solutions for Ageing Societies, GZ 1570.

Institutional Review Board Statement: No applicable.

Informed Consent Statement: No applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Thanks to the blockchain technology team of the Cross-Innovation Laboratory of East China Normal University for their fully support. At the same time, We also thank the reviewers for their valuable time.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Jakimowicz, A. The Role of Entropy in the Development of Economics. *Entropy* **2020**, *22*, 452. [CrossRef] [PubMed]
2. Grassberger, P.; Procaccia, I. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. A* **1983**, *28*, 2591–2593. [CrossRef]
3. Gunay, S.; Kaşkaloğlu, K. Seeking a Chaotic Order in the Cryptocurrency Market. *Math. Comput. Appl.* **2019**, *24*, 36. [CrossRef]
4. Fistola, R.; La Rocca, R.A. The Sustainable City and the Smart City: Measuring urban entropy first. *WIT Trans. Ecol. Environ.* **2014**, *191*, 537–548.
5. Dehouche, N. Revisiting the Volatility of Bitcoin with Approximate Entropy. *Cogent Econ. Financ.* **2021**, *10*, 2013588. [CrossRef]
6. Pele, D.T.; Mazurencu, M. Using High-Frequency Entropy to Forecast Bitcoin's Daily Value at Risk. *Entropy* **2019**, *21*, 102. [CrossRef]
7. Santoro, D.; Grilli, L. A Statistical Ensemble Based Approach for Entropy in Cryptocurrencies Markets. In *13th Chaotic Modeling and Simulation International Conference*; International Society for the Advancement of Science and Technology: San Diego, CA, USA, 2020.
8. Wu, T. Advantages, constraints and promotion strategies of open finance based on blockchain. *Econ. Asp.* **2020**, *2*, 91–98.
9. Zhu, L.; Gao, F.; Shen, M.; Li, Y.; Zheng, B.; Mao, H.; Wu, Z. Summary of research on blockchain privacy protection% Survey on Privacy Preserving Techniques for Blockchain Technology. *Comput. Res. Dev.* **2017**, *54*, 2170–2186.
10. Wan, G.; Sun, T. Research on the Ideal, Reality and Regulatory Countermeasures of "Blockchain + Securities". *Shanghai Financ.* **2017**, *57*, 58–64.
11. Pu, D.; Fan, C.; Liang, H. Construction and application of e-commerce platform system based on blockchain perspective. *China Circ. Econ.* **2018**, *32*, 44–51.
12. Zhou, M.; Qin, W. Reconstruction of e-commerce credit system in the era of blockchain 3.0. *Learn. Pract.* **2019**, *11*, 47–56.
13. Li, D. Research on the Coupling Development of Digital Economy, Cross-border E-commerce and Digital Trade—Also on the Application of Blockchain Technology in the Three. *Theor. Discuss* **2020**, *1*, 115–121.
14. Yan, Z. A new sharing economy model based on blockchain technology. *Soc. Sci. Res.* **2020**, *1*, 94–101.
15. Zhang, X. Optimization of supply chain management model based on blockchain. *China Circ. Econ.* **2018**, *32*, 44–52.

16. Zhang, X.; Zhang, H.; Guo, X.; Wen, Z. Research and analysis of electronic voting and election system based on blockchain. *Electron. Technol. Appl.* **2017**, *43*, 132–135.
17. Lu, M. Research on the Reconstruction of Blockchain Technology and Social Credit System. *Lanzhou Acad. J.* **2020**, *3*, 80–88.
18. Shi, D. On the value and risk of blockchain technology for digital copyright governance. *Sci. Technol.* **2019**, *6*, 111–120.
19. Yan, Z.; Liu, L.; Li, Q. Cooperative Evolution of Parallel Society under Blockchain System. *China Sci. Technol. Forum* **2018**, *6*, 50–58.
20. Liu, Y.; Zhang, Y.; Wu, Y.; Zheng, C. Blockchain technology and document archive management: Two-way thinking of technology and management. *Arch. Newsl.* **2020**, *1*, 4–12.
21. Lin, M.; Zhang, Z. "Blockchain + Production" Promotes Enterprise Green Production—A New Thinking on the Hand of Government. *Econ. Trends* **2019**, *5*, 42–56.
22. Zhao, J.; Meng, T. Technology Empowerment: How Blockchain Reshapes Governance Structure and Model. *Contemp. World Soc.* **2019**, *3*, 187–194.
23. Gao, Q.; Zhang, J. Blockchain and Transformation of Global Economic Governance—Based on the Perspective of Con-structing a Global Justice Economic Order. *Academia* **2019**, *9*, 21–36.
24. Yang, H. The new globalization process after the volatility and the new issues of global governance—Transnational coupling of blockchain, human mobility and authority. *Acad. Circ.* **2018**, *239*, 150–160.
25. Jin, Y. Demand analysis and technical framework of blockchain + education. *China Audio-Vis. Educ.* **2021**, *2017-9*, 62–68.
26. Xu, J.; Wang, J. Trust Building of Academic Publishing Based on Blockchain Technology. *Publ. Sci.* **2017**, *25*, 19–24.
27. Xu, T. Application and Challenge of Blockchain Technology in Education and Teaching. *Mod. Educ. Technol.* **2017**, *1*, 110–116.
28. Li, Z.; Qiu, T.; Li, K. Research on Credit Banking System Based on Blockchain Technology. *Mod. Educ. Technol.* **2019**, *7*, 120–126.
29. Chen, X. Blockchain technology and its application in audit experiment teaching. *China Audio-Vis. Educ.* **2019**, *6*, 131–132.
30. Yang, Y.; Du, J.; Luo, X. Analysis of the impact of blockchain technology on both parties of tax collection. *Tax Res.* **2019**, *409*, 116–120.
31. Ren, C. Research on tax collection and management model based on blockchain technology. *Tax Res.* **2018**, *406*, 92–99.
32. Li, W.; Liu, H.; Deng, X. Application of Blockchain Technology to Promote Credit Management of Tax Payment in my country. *Tax Res.* **2018**, *407*, 80–84.
33. Tang, X.; Zhou, H. Construction of tax governance framework based on blockchain technology. *Tax Res.* **2018**, *11*, 98–104.
34. Dong, D.; Wang, X. Research on electronic medical record sharing based on blockchain. *Comput. Technol. Dev.* **2019**, *29*, 121–125.
35. Zhu, J. Blockchain finance consumer rights protection: The path of experimental regulation. *Nanjing Soc. Sci.* **2018**, *374*, 106–111.
36. Ding, X. Thoughts on the integration of blockchain into copyright insurance. *Nanjing Soc. Sci.* **2018**, *9*, 120–125.
37. Zhang, L.; Zhou, Y.; Xue, L. Intellectual Property Management and Policy Research of Strategic Emerging Industries Based on Blockchain Technology. *China Sci. Technol. Forum* **2018**, *12*, 120–126.
38. Huang, B.; Deng, X.; Yu, Z.; Wang, C.; Guo, D.; Yang, Z. A Data Synchronization Method Based on the Operation of Block-Chain State Unit Network. China Patent CN108667928A, 16 October 2018.
39. Zhang, Y.; Wang, Y.; Yang, Z.; Yang, R. Blockchain smart contract script design based on machine learning. *Inf. Eng.* **2019**, *5*, 4–10.
40. Hao, S.; Xu, W.; Tang, Z. Research on scientific data sharing blockchain model and implementation mechanism. *Inf. Theory Pract.* **2018**, *41*, 57–62.
41. Tang, C.; Yang, Z.; Zheng, Z.; Cheng, Z.Y. Analysis and Optimization of Game Dilemma in PoW Consensus Algorithm. *J. Autom.* **2017**, *43*, 1520–1531.
42. Zhu, X. Blockchain-Based Transaction Buffer/Acceleration Method and Blockchain Transaction Processing System. China patent CN108805702A, 13 November 2018.
43. Lin, J. Promote the development of blockchain technology through the integration of IPv6 technology. *Comput. Prod. Circ.* **2019**, *6*, 81.
44. Wang, L.; Zhao, X. Research on service composition strategy in cloud computing environment based on blockchain mechanism. *J. Comput. Appl.* **2019**, *36*, 81–86.
45. Weng, J.; Weng, J.; Liu, J.; Wei, K.; Luo, W. Blockchain-Based Software-Defined Network Control Layer Security Mechanism Construction Method. China Patent CN107222478A, 29 September 2017.
46. Jiao, T.; Nie, T.; Shen, D.; Li, X. Blockchain database: A queryable and tamper-proof database. *J. Softw.* **2019**, *30*, 2671–2685.
47. Deng, E. A Method for Secure Transmission of Blockchain Data Based on SCTP Protocol. China Patent CN107104977A, 28 August 2017.
48. Yang, M.; Zhang, S.; Zhang, H.; Liu, N.; Gan, B. User trust negotiation model based on two-layer blockchain in heterogeneous alliance system. *J. Appl. Sci.* **2019**, *37*, 244–252.
49. Wang, M.; Duan, M. Merkle Hash Tree Structure Based Blockchain Second Preimage Attack. *Inf. Netw. Secur.* **2018**, *1*, 38–44.
50. Andriole, S.J. Blockchain, Cryptocurrency, and Cybersecurity. *IT Prof.* **2020**, *22*, 13–16. [CrossRef]
51. Rezaeibagha, F.; Mu, Y. Efficient Micropayment of Cryptocurrency from Blockchains. *Comput. J.* **2019**, *62*, 507–517. [CrossRef]
52. Zhong, L.; Wu, Q.; Xie, J.; Guan, Z.; Qin, B. A secure large-scale instant payment system based on blockchain. *Comput. Secur.* **2019**, *84*, 349–364. [CrossRef]
53. Yin, H.H.S.; Langenheldt, K.; Harlev, M.; Mukkamala, R.R.; Vatrapu, R. Regulating Cryptocurrencies: A Supervised Machine Learning Approach to De-Anonymizing the Bitcoin Blockchain. *J. Manag. Inf. Syst.* **2019**, *36*, 37–73. [CrossRef]

54. Juhasz, P.L.; Steger, J.; Kondor, D.; Vattay, G. A Bayesian approach to identify Bitcoin users. *PLoS ONE* **2018**, *13*, e0207000. [CrossRef]
55. Ante, L. A place next to Satoshi: Foundations of blockchain and cryptocurrency research in business and economics. *Scien-tometrics* **2020**, *124*, 1305–1333. [CrossRef]
56. Eyal, I. Blockchain Technology: Transforming Libertarian Cryptocurrency Dreams to Finance and Banking Realities. *Computer* **2017**, *50*, 38–49. [CrossRef]
57. Hegadekatti, K.; Yatish, S.G. The Programmable Economy: Envisaging an Entire Planned Economic System as a Single Computer through Blockchain Networks. *MPRA Pap.* **2017**, *58*, 32. [CrossRef]
58. Aste, T.; Tasca, P.; Matteo, T.D. Blockchain Technologies: The Foreseeable Impact on Society and Industry. *Computer* **2017**, *50*, 18–28. [CrossRef]
59. Campbell-Verduyn, M. Introduction to special section on blockchains and financial globalization. *Glob. Netw.* **2019**, *19*, 283–290. [CrossRef]
60. Abe, R.; Watanabe, H.; Ohashi, S.; Fujimura, S.; Nakadaira, A. Storage Protocol for Securing Blockchain Transparency. In Proceedings of the 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo, Japan, 23–27 July 2018.
61. Wang, S.; Ouyang, L.; Yuan, Y.; Ni, X.; Han, X.; Wang, F.Y. Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *49*, 2266–2277. [CrossRef]
62. Thyagarajan, S.A.K.; Bhat, A.; Magri, B.; Tschudi, D.; Kate, A. Reparo: Publicly Verifiable Layer to Repair Blockchains. *arXiv* **2020**, arXiv:2001.00486.
63. Manimaran, P.; Dhanalakshmi, R. Blockchain-Based Smart Contract for E-Bidding System. In Proceedings of the International Conference on Intelligent Communication & Computational Techniques, Jaipur, India, 28–29 September 2019.
64. Lamberti, F.; Gatteschi, V.; Demartini, C.; Pelissier, M.; Gomez, A.; Santamaria, V. Blockchains Can Work for Car Insurance: Using Smart Contracts and Sensors to Provide On-Demand Coverage. *IEEE Consum. Electron. Mag.* **2018**, *7*, 72–81. [CrossRef]
65. Wang, E.K.; Liang, Z.; Chen, C.M.; Kumari, S.; Khan, M.K. PoRX: A reputation incentive scheme for blockchain consensus of IIoT. *Future Gener. Comput. Syst.* **2020**, *102*, 140–151. [CrossRef]
66. Okada, H.; Yamasaki, S.; Bracamonte, V. Proposed classification of blockchains based on authority and incentive dimensions. In Proceedings of the International Conference on Advanced Communication Technology, Pyeongchang, Korea, 19–22 February 2017.
67. Wu, W.; Gao, Z. An Improved Blockchain Consensus Mechanism Based on Open Business Environment. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *428*, 012043. [CrossRef]
68. Zou, J.; Ye, B.; Qu, L.; Wang, Y.; Orgun, M.A.; Li, L. A Proof-of-Trust Consensus Protocol for Enhancing Accountability in Crowdsourcing Services. *IEEE Trans. Serv. Comput.* **2018**, *12*, 429–445. [CrossRef]
69. Feng, L.; Hui, Z.; Tsai, W.T.; Sun, S. System architecture for high-performance permissioned blockchains. *Front. Comput. Sci.* **2019**, *13*, 1151–1165. [CrossRef]
70. Leonardos, S.; Reijsbergen, D.; Piliouras, G. Weighted Voting on the Blockchain: Improving Consensus in Proof of Stake Protocols. In Proceedings of the IEEE International Conference on Blockchain and Cryptocurrency (ICBC 2019), Seoul, Korea, 15–17 May 2019.
71. Yang, D.; Jeon, S.; Doh, I.; Chae, K. Randomly Elected Blockchain System based on Grouping Verifiers for Efficiency and Security. In Proceedings of the 2020 22nd International Conference on Advanced Communication Technology (ICACT), Pyeongchang, Korea, 16–19 February 2020.
72. Chicarino, V.; Albuquerque, C.; Jesus, E.; Rocha, A. On the detection of selfish mining and stalker attacks in blockchain networks. *Ann. Telecommun.* **2020**, *75*, 143–152. [CrossRef]
73. Frahat, R.T.; Monowar, M.M.; Buhari, S.M. Secure and Scalable Trust Management Model for IoT P2P Network. In Proceedings of the 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 1–3 May 2019.
74. Ito, K.; Tago, K.; Jin, Q. i-Blockchain: A Blockchain-Empowered Individual-Centric Framework for Privacy-Preserved Use of Personal Health Data. In Proceedings of the 2018 9th International Conference on Information Technology in Medicine and Education (ITME), Hangzhou, China, 19–21 October 2018.
75. Ferretti, S.; D'Angelo, G. On the Ethereum Blockchain Structure: A Complex Networks Theory Perspective. *Concurr. Comput. Pract. Exp.* **2020**, *32*, e5493. [CrossRef]
76. Liu, F.; Yang, C.; Yu, X.; Qi, J. Spectrograph Convolutional Neural Network for Decentralized Double Differential Privacy. *Inf. Network Secur.* **2022**, *22*, 152–159.
77. Liu, F.; Zhang, J.; Zhou, J.; Li, M.; Kong, D.; Yang, J.; Qi, J.; Zhou, A. Blockchain Cross-Chain Asset Interaction Protocol Based on Improved Hash Time Lock. *Comput. Sci.* **2022**, *49*, 336–344.
78. Xu, J.; Wei, L.; Zhang, Y.; Wang, A.; Zhou, F.; Gao, C.Z. Dynamic Fully Homomorphic encryption-based Merkle Tree for lightweight streaming authenticated data structures. *J. Netw. Comput. Appl.* **2018**, *107*, 113–124. [CrossRef]
79. Waqas, P.; Byun, Y. A Blockchain-Based Secure Image Encryption Scheme for the Industrial Internet of Things. *Entropy* **2020**, *22*, 175.

80. Lv, T.; Liu, F. Research on computing power network under the background of digital economy. *J. Beijing Jiaotong Univ. (Soc. Sci. Ed.)* **2021**, *20*, 11–18.
81. Liu, F.; Wang, Y.; Yang, J.; Zhou, A.; Qi, J. A blockchain-based high-threshold signature protocol integrating DKG and BLS. *Comput. Sci.* **2021**, *48*, 46–53.
82. Liu, F.; Yang, J.; Qi, J. Blockchain Two-Party Elliptic Curve Based on Hash Proof System. *Inf. Netw. Secur.* **2021**, *1*, 89–96.
83. Liu, F.; Yang, J.; Li, Z.; Qi, J. Research on an Endogenous Data Security Interaction Protocol for Dual-Central-Two-chain Architecture. *J. East China Norm. Univ. (Nat. Sci. Ed.)* **2020**, *5*, 44–55.
84. Dyhrberg, A.H. Bitcoin, gold and the dollar—A GARCH volatility analysis. *Financ. Res. Lett.* **2016**, *16*, 85–92. [CrossRef]
85. Huang, Y.; Duan, K.; Mishra, T. Is Bitcoin really more than a diversifier? A pre- and post-COVID-19 analysis. *Financ. Res. Lett.* **2021**, *43*, 102016. [CrossRef]
86. Mariana, C.D.; Ekaputra, I.A.; Husodo, Z.A. Are Bitcoin and Ethereum safe-havens for stocks during the COVID-19 pandemic? *Financ. Res. Lett.* **2021**, *38*, 101798. [CrossRef]
87. Disli, M.; Nagayev, R.; Salim, K.; Rizkiah, S.K.; Aysan, A.F. In search of safe haven assets during COVID-19 pandemic: An empirical analysis of different investor types. *Res. Int. Bus. Financ.* **2021**, *58*, 101461. [CrossRef]
88. Kwapień, J.; Wątorek, M.; Drożdż, S. Cryptocurrency Market Consolidation in 2020–2021. *Entropy* **2021**, *23*, 1674. [CrossRef]
89. "The Transaction Volume of NFT" [EB/OL]. Available online: https://nonfungible.com/market/history (accessed on 30 December 2021).
90. Wang, Q.; Li, R.; Wang, Q.; Chen, S. Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges. *arXiv* **2021**, arXiv:2105.07447. [CrossRef]
91. Mnif, E.; Jarboui, A.; Mouakhar, K. How the cryptocurrency market has performed during COVID 19? A multifractal analysis. *Finance Research Letters* **2020**, *36*, 101647. [CrossRef]
92. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. *Decentralized Bus. Rev.* **2008**, 21260.
93. Li, Q.; Zhang, X. Blockchain:Promoting openness and credibility of education with technology. *J. Distance Educ.* **2017**, *35*, 36–44.
94. Schmidt, P. 'Blockcerts': Open Standard for Blockchain Certificates Created by MIT[EB/OL]. Available online: https/www.8btc.com/article/107456 (accessed on 30 December 2021).
95. Sharples, M.; Domingue, J. The Blockchain and Kudos: A Distributed System for Educational Record, Reputation and Reward. In *European Conference on Technology Enhanced Learning*; Springer: Cham, Switzerland, 2016; pp. 490–496.
96. Wang, A. Blockchain Technology and Its Applications [EB/OL]. Available online: https://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1020&context=asars (accessed on 30 December 2021).
97. Fei, J.; Jia, H. The path selection of public service platform provided by government APP from the perspective of smart government. *E-Government* **2015**, *9*, 31–37.
98. Yu, S.; Yang, D.; Wang, J.; Zhang, Y.; Wang, J. Big data-based smart government portal: From concept to practice. *E-Government* **2013**, *5*, 64–74.
99. Luo, X.; Yu, B.; Yao, M. Analysis of e-government development stages from the perspective of information chain. *Libr. Stud.* **2014**, *6*, 35–40.
100. Lv, Z.H.; Li, X.M.; Wang, W.X.; Zhang, B.; Hu, J.; Feng, S. Government affairs service platform for smart city. *Future Gener. Comput. Syst. Int. J. ESci.* **2018**, *81*, 443–451. [CrossRef]
101. Hu, M.; Ma, J. Research on the mechanism of promoting borderless smart government in the perspective of information collaboration. *Intell. Data Work.* **2019**, *40*, 44–51.
102. Yuan, Y.K.; Zhang, Y.; Wei, T. Review of key technologies and applications of intelligent transportation. *Electron. Technol. Appl.* **2015**, *41*, 9–12.
103. Zhang, H.-L. Application scenarios and challenges of blockchain technology in the transportation field. *China Transp. Informatiz.* **2020**, *12*, 96–98.
104. Wang, M.; Gong, Z. 'Converging Applications of Blockchain Technology in the Field of Intelligent Transportation' [EB/OL]. Available online: https://www.sohu.com/a/237977066_649849 (accessed on 30 December 2021).
105. Feng, Y.; Wang, F. Blockchain-based application of Internet of things technology. *Internet Things Technol.* **2021**, *11*, 120–122.
106. 'Blockchain in the Internet of Things' [EB/OL]. Available online: https://blog.csdn.net//bigtree_3721/article/details/79517416 (accessed on 30 December 2021).
107. Na, L. New generation of information technology to enable intelligent development of coal industry. *Coal Eng.* **2020**, *52*, 193–196.
108. Yang, D.Q.; Yue, A.; Yang, R. Research on integrated management of engineering project information under smart construction-application based on blockchain technology. *Constr. Econ.* **2019**, *40*, 80–85.
109. Yu, S.B.; Zheng, D.D. Application and prospect of blockchain technology in the field of energy and power. *Huadian Technol.* **2020**, *42*, 17–23.
110. Zhang, Y.; Wang, L.; Wu, J.; Yuan, R.; Li, M. Blockchain and integrated energy systems: Applications and perspectives. *China Sci. Found.* **2020**, *34*, 31–37.
111. Huang, X.; Zhang, Y.; Li, D.; Han, L. An optimal scheduling algorithm for hybrid EV charging scenario using consortium blockchains. *Future Gener. Comput. Syst.* **2019**, *91*, 555–562. [CrossRef]
112. Andoni, M.; Robu, V.; Flynn, D.; Abram, S.; Geach, D.; Jenkins, D.; McCallum, P.; Peacoack, A. Blockchain technology in the energy sector:A systematic review of challenges and opportunities. *Renew. Sustain. Energy Rev.* **2019**, *100*, 143–174. [CrossRef]

113. 'Trends and Cutting-Edge Applications of Energy Blockchain, IoT Technologies in Smart Grid' [EB/OL]. Available online: https://www.cnblogs.com/newstart/p/10594471.html (accessed on 30 December 2021).
114. 'Why Does Power Ledger Favor Solana?' [EB/OL]. Available online: https://weibo.com//ttarticle/p/show?id=2309404668673431306451&sudaref=www.baidu.com (accessed on 30 December 2021).
115. Wang, Q.; Chen, Z.; Wang, P. Application of blockchain technology in new retail supply chain. *Bus. Econ. Res.* **2020**, *14*, 34–36.

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
www.mdpi.com

Entropy Editorial Office
E-mail: entropy@mdpi.com
www.mdpi.com/journal/entropy

MDPI

Academic Open
Access Publishing

www.mdpi.com

ISBN 978-3-0365-8575-8