Wavelet entropy and complexity analysis of cryptocurrencies dynamics

Victoria Vampa[1] and María T. Martín[2] and Lucila Calderón[1,2] and Aurelio F. Bariviera[3]

1 Universidad Nacional de La Plata, Facultad de Ingeniería, Argentina
2 Universidad Nacional de La Plata, Facultad de Ciencias Exactas, Argentina
3 Universitat Rovira i Virgili, Department of Business, Av. Universitat 1, 43204 Reus, Spain
aurelio.fernandez@urv.cat

Abstract. Cryptocurrencies emerged almost one decade ago, as an alternative peer-to- peer payment method and later as a synthetic financial asset. This paper examines the long memory properties in high frequency (5 minutes) time series of eight important cryptocurrencies. We perform a statistical analysis of two key financial characteristics of time series: return and volatility. We compute information theory quantifiers using a wavelet decomposition of the time series: wavelet entropy and wavelet statistical complexity of returns and volatility of each time series. We find two important features in the time series: (i) high frequency returns exhibit a trend toward a more efficient behavior, and (ii) high frequency volatility reflects a strong persistence in volatility. Both findings have important implications for portfolio managers, and investors in general.

Keywords: cryptocurrencies, wavelet entropy, statistical complexity, long memory.

Cite as:
Vampa, V., Martín, M. T., Calderón, L., & Bariviera, A. F. (2022). Wavelet Entropy and Complexity Analysis of Cryptocurrencies Dynamics. In M. del P. Rodríguez García, K. A. Cortez Alejandro, J. M. Merigó, A. Terceño-Gómez, M. T. Sorrosal Forradellas, & J. Kacprzyk (Eds.), Digital Era and Fuzzy Applications in Management and Economy (Vol. 384, pp. 25–35).
https://doi.org/10.1007/978-3-030-94485-8_2

Wavelet entropy and complexity analysis of cryptocurrencies dynamics

Abstract. Cryptocurrencies emerged almost one decade ago, as an alternative peer-to- peer payment method and later as a synthetic financial asset. This paper examines the long memory properties in high frequency (5 minutes) time series of eight important cryptocurrencies. We perform a statistical analysis of two key financial characteristics of time series: return and volatility. We compute information theory quantifiers using a wavelet decomposition of the time series: wavelet entropy and wavelet statistical complexity of returns and volatility of each time series. We find two important features in the time series: (i) high frequency returns exhibit a trend toward a more efficient behavior, and (ii) high frequency volatility reflects a strong persistence in volatility. Both findings have important implications for portfolio managers, and investors in general.

Keywords: cryptocurrencies, wavelet entropy, statistical complexity, long memory.

1 Introduction

Eleven years ago, a white paper (anonymously posted online) proposed a new paradigm for validating financial transactions [1]. This new financial clearing setting, called 'blockchain', replaces the central trusted authority by a community validation based on consensus. Such validation uses a dis- tributed ledger, which is replicated in servers within the network. The existence of several simultaneous encrypted copies of the transaction history produced robustness against data manipulation. Thus, it constitutes a safe alternative to validating transactions. The vehicle in such transactions is a new asset called cryptocurrency, due to its encrypted nature. Bitcoin price rocketed during 2017, reaching in December of that year almost $20,000 [2]. If we assume that an investor bought one bitcoin in June 2009 at around
$0.0001, he would have earned approximately 847% compounded annually. Such success encouraged other software developers to create new cryptocurrencies. Until 2016 Bitcoin was the unchallenged dominant player, with more than 90% of the market. As of July 2020, there are more than 5000 cryptocurrencies (coins and tokens), with a market capitalization of $272 billion and daily transactions of around $56 billion. Bitcoin represents 62% of the whole market [3].

This market is very dynamic with new coins (globally known as 'altcoins') entering the market every week, and some coins stopping trading because of lack of investors' attention.

Despite its name, cryptocurrencies are more an actively traded financial asset rather than a standard currency. Its daily use is really very limited. Coinmap [4] reports that there are only 22891 physical venues in the world accepting payments in cryptocurrencies, limiting their use as a medium of exchange. Volatility of crytocurrencies is ten times higher than any other financial assets such as fiat currencies, bonds, stocks or precious metals. Moreover, volatility exhibits long term memory [5]. Such situation challenges the use of cryptocurrencies use as store of value.

The aim of this paper is to revisit two important empirical features of cryptocurrencies time series: (i) the informational efficiency, and (ii) the persistence in volatility. The period under study is very relevant, because it covers the final period of the price bubble formed in the last quarter of 2017 and the aftermath of the bubble burst.

Our paper contributes to the existing literature in several ways. First, we study high frequency time series of the main cryptocurrencies, whereas most of the papers focused only on daily returns. Second, our technique (wavelet entropy) has been not used yet in this market. Third, we study not only high frequency returns, but also high frequency volatility in the period around the bubble burst of 2017-2018. This approach is relevant to provide an assessment of market risk.

The paper is organized as follows. Section 2 reviews the most relevant cryptocurrency literature. Section 3 presents the methodology. Section 4 describes data and discusses the main findings. Finally, Section 5 outlines the conclusion of our analysis.

2 Brief literature review

Scientific literature around cryptocurrencies covers different facets of the problem. From an economics point of view, an important discussion deals with the characterization of cryptocurrencies as "currencies". Yermack [6] argues that Bitcoin does not reach the currency status, as it performs poorly as a unit of account and as a store of value. Several papers [7,8] characterize bitcoin mainly as a speculative asset rather than a currency. it is widely recognized [9,10,11] that bitcoin (and generally most cryptocurrencies) con- form a unique and uncorrelated asset class, compared with traditional assets such as stocks, bonds, or commodities.

One important research line within financial economics is the study of the informational endowment of cryptocurrencies time series. The seminal paper by [12] defines informational efficiency as a situation when market prices fully reflect all available information. It is conventionally classified into three broad categories (weak, semi-strong, and strong), depending on the set of information considered as benchmark. This paper studies the weak form of informational efficiency, meaning that returns should follow a white noise.

Although the informational (in)efficiency of the cryptocurrency market is well-documented [13,14,15,5], most studies use daily data. Even though, there are papers that use intraday data [16,17,18,19], their focus is exclusively on bitcoin.

3 Wavelet entropy and wavelet complexity

Financial markets, and in our case cryptocurrencies' markets, register each transaction, including attributes such as time, price, and quantity of the operation. A close scrutiny on these data could be useful to extract information on the statistical characteristics of the price generating process in a given market.

Information theory-based quantifiers could be a suitable alternative to more traditional econometric methods in time series analysis. Entropy quantifiers have a long tradition in economics. Three papers dated in the 1960s were among the first to use information entropy to examine the predictability and temporal dependence in stock time series [20, 21, 22].

The celebrated Shannon entropy [23] constitutes a straightforward way to measure the degree of (dis)order in a system. Let $P = \{p_i \in \mathbb{R}; p_i \geq 0; i = 1, \dots, M\}$, be a discrete probability distribution, with $\sum_{i=1}^{M} p_i = 1$, then the Shannon entropy reads:

$$\boldsymbol{S}[P] = \sum_{i_1}^{M} p_i \, log \, p_i \tag{1}$$

This quantifier equals zero if the patterns are fully deterministic and reaches its maximum value for a uniform distribution.

Wavelet entropy is based on Shannon entropy and is computed by defining a probability distribution that arises from the multiresolution analysis on $L^2(\mathbb{R})$ [24]. Given a multiresolution analysis $\{V_j\}_{j \in Z}$, and being W_j the orthogonal complement of V_{j+1} in V_j with:

$$W_j \oplus V_j = V_{j+1} \tag{2}$$

it is obtained that:

$$L^2(\mathbb{R}) = \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \tag{3}$$

Then, the original signal $\mathrm{s} \in \mathrm{L}^2(\mathbb{R})$, could be written as:

$$\mathrm{s(t)} = \sum_{\mathrm{j} \in \mathbb{Z}} \sum_{\mathrm{k} \in \mathbb{Z}} \mathrm{d_j(k)} \psi_{\mathrm{j,k}}(\mathrm{t}) \tag{4}$$

where $\{\psi_{\mathrm{j,k}}\}$ for each given j, is a base on the space W_j.

The probability distribution is obtained from the concept of "wavelet energy":

$$\mathrm{E_j} \coloneqq \sum_{\mathrm{k} \in \mathbb{Z}} \left|\mathrm{d_j(k)}\right|^2 \tag{5}$$

Consequently, the total energy of a signal $s(t)$ can be obtained by adding up the energy of all levels j:

$$\mathrm{E_{total}} \coloneqq \sum_{\mathrm{j} \in \mathbb{Z}} \sum_{\mathrm{k} \in \mathbb{Z}} \left|\mathrm{d_j(k)}\right|^2 \tag{6}$$

Using equations (5) and (6), we define the energy probability distribution as:

$$\mathrm{p_j} \coloneqq \frac{\mathrm{E_j}}{\mathrm{E_{total}}} \tag{7}$$

Finally, the wavelet entropy $\mathrm{S_W[P]}$ is defined using these p_j in Equation (1).

Another information quantifier is the Statistical Complexity Measure (SCM), which is a global information quantifier [25], and will be denoted by $\mathrm{C[P]}$.

For a discrete probability distribution function $\mathrm{P} = \{\mathrm{p_i} \in \mathbb{R};\ \mathrm{p_i} \geq 0; \mathrm{i} = 1, \dots, \mathrm{M}\}$, associated with a time series, this functional $\mathbf{C}[P]$ is given by

$$C[P] = H[P].Q[P] \tag{8}$$

where $H[P]$ is the normalized Shannon entropy and Q[P] is the disequilibrium-distance.

$H[P]$ measures the amount of disorder and is defined as,

$$H[P] = \frac{S[P]}{S_{max}} \tag{9}$$

where S[P] is the Shannon entropy of Equation (1) and $S_{max} = S[P_e]$ corresponds to the uniform distribution $P_e = \{p_i = \frac{1}{M}; i = 1, \dots, M\}$, and so has the value $S_{max} = \log(M)$.

Q[P] is called the disequilibrium-distance. It is defined as:

$$Q[P] = Q_0.\ D[P, P_e] \tag{10}$$

where Q_0 is a normalization constant ($0 \leq Q \leq 1$) with its value equal to the inverse of the maximum possible value of the distance $D[P, P_e]$. This maximum distance is obtained when one of the components of P, say p_m is equal to one, while the remaining components are equal to zero. The disequilibrium-distance $\mathbf{Q}$ would be different from zero if there exist more likely states among the accessible ones.

Several distance-forms open several possibilities for the SCM [25]. In this work we consider the disequilibrium form for the complexity measure proposed by Lopez Ruiz, Mancini and Calbet (LMC-complexity measure) in which the Euclidean norm is considered:

$$D_E\left[P^{(1)}, P^{(2)}\right] = |P^{(1)} - P^{(2)}|_E^2 = \sum_{j=1}^{M}\left(p_j^{(1)} - p_j^{(2)}\right)^2 \tag{11}$$

with $P^{(l)} = \{p_i^{(l)} \in \mathbb{R}; p_i \geq 0; i = 1, \dots, M\}$, $l = 1{,}2$ are two discrete probability distributions.

Giving a measure of the complexity of a time series, $\mathbf{C}[P]$ quantifies the existence of correlational structures. In perfect order or total randomness of the signal, the value of $\mathbf{C}[P]$ is identically zero meaning the signal possesses no structure. A large range of possible stages may be realized between these two extreme instances and quantified by a nonzero $\mathbf{C}[P]$.

The statistical complexity does not only quantify randomness (or disorder) but also the degree of correlation between structures. It is a non-trivial function of entropy, and it is important to note that, for a given value of $H[P]$]$ there is a range of possible values for $\mathbf{C}[P]$ between a minimum and a maximum value [25].

Once these quantifiers $H[P]$ and $\mathbf{C}[P]$ are calculated, the results can be displayed in the $H\mathbf{C}$ plane. In this way, specific features associated with the dynamics of the series can be characterized.

In the present work, a new approach for the characterization of returns and volatility time series is presented. The information theory quantifiers $H_W[P]$ and $C_{LMC}[P]$ were used. The first one, $H_W[P]$, uses wavelet entropy $S_W[P]$ in Equation (9), while the complexity quantifier $C_{LMC}[P]$ is obtained from Equation (8) using the Euclidean distance D_E in Equation (11) to measure the disequilibrium-distance, Equation (10).

4 Data and results

The empirical analysis was performed using high frequency (5 minute) price data from eight of the main cryptocurrencies. In this way, signals $s(t)$ have returns and volatility values corresponding to a uniform time grid.
Data coverage spans from 03/12/2017 until 12/04/2018, comprising 29175 observations. Data were retrieved from Eikon Thomsom Reuters. The selection of the data span was justified because it included a period of boom and bust in the cryptocurrency market. During December 2017 there was an unprecedented increase in cryptocurrencies, followed by a crash during the first months of 2018. In this sense. Contrary to previous studies, we would like to focus our study during a time of market stress.
Following [5], we adopt a dynamic approach and compute the information quantifiers (wavelet entropy and LMC statistical complexity) using rolling windows. This procedure works as follows: we consider the first 128 observations (roughly half a day) and compute the quantifiers. Then we select the following 128 observations and compute their respective quantifiers. We repeat the algorithm until the end of the time series. We obtained 227 rolling windows.
Our study detects several interesting features in the cryptocurrency ecosystem, regarding both returns and volatility.

4.1 Cryptocurrencies' returns

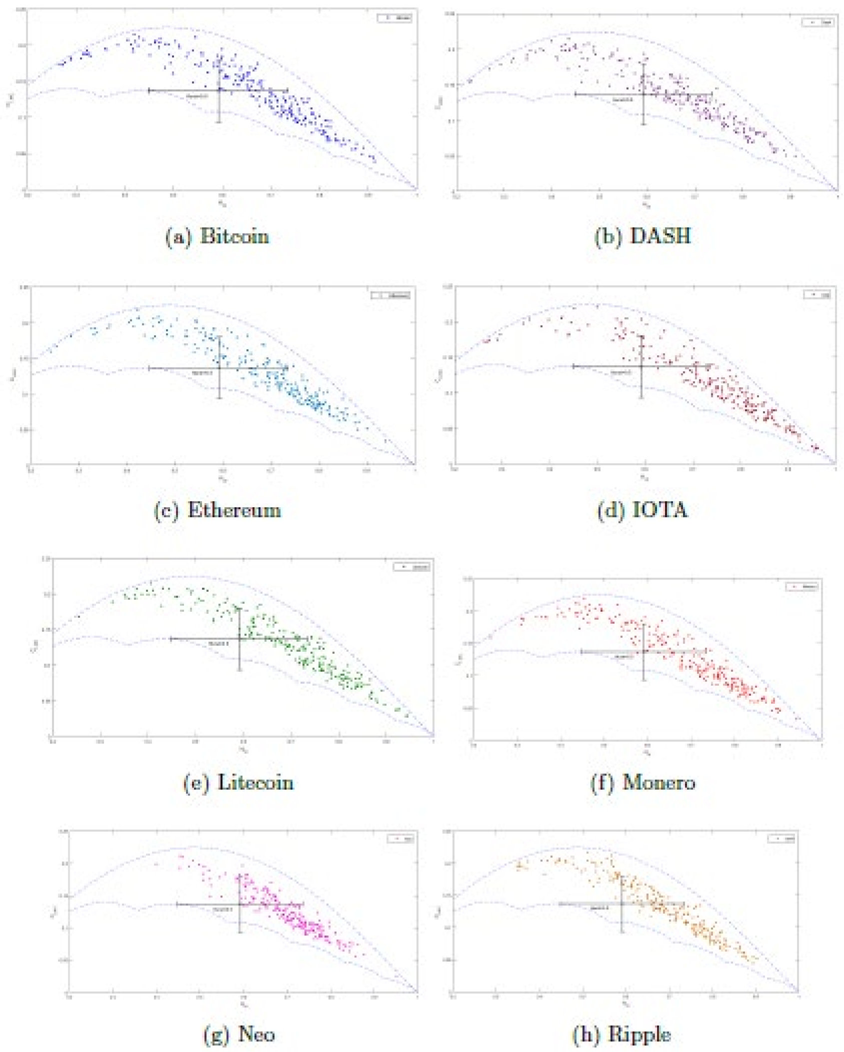
Our initial analysis explores the long memory of the time series, using the wavelet entropy-complexity plane. This representation allows to characterize different stochastic and chaotic dynamics. For the purpose of identification, we draw the confidence interval of the location of a Brownian motion with Hurst exponent equal to 0.5[1]
Figure 1 shows how the time series of all the cryptocurrencies occupy a broad extension of this plane. This situation means that the time series goes through different stages in stochastic dynamics. In fact, only a small proportion of the sliding windows lies in the area corresponding to Hurst=0.5. We can interpret that most of the time, the time series exhibit an informational inefficient behavior. One drawback of this representation is that it does not allow to identify how the position of the quantifiers evolve in time.
In order to overcome this situation, we present also Figure 2, which de- scribe the evolution of the wavelet entropy along the different sliding windows. This figure reveals that the informational efficiency, proxied by the wavelet entropy, varies significantly. There are periods when the time series exhibit a fairly random behavior, whereas in other periods the time series become informational inefficient. However, there is a clear trend towards a more efficient behavior.
Finally, this figure uncovers a remarkable variety in the cryptocurrency market. The movements in wavelet entropy are not completely synchronized for the different

---

[1] This confidence interval was obtained by simulating 1000 artificial time series using the function wfbm from Matlab and computing the average and standard deviation of wavelet entropy and statistical complexity

currencies under analysis. However, it emerges a trend to a closer comovement in this quantifier by the end of the examining period. One striking feature, that was not reported in previous literature, is that IOTA exhibits a dynamic that is apart from the other coins. Altogether these results have two important implications from a portfolio perspective. On the one hand, if cryptocurrencies display a convergence in the stochastic dynamics, means that the market is more closely integrated. On the other hand, the existence of some cryptocurrencies (in our study, IOTA), with mismatched dynamics could be used to exploit diversification benefits in portfolios.

Fig. 1. Wavelet entropy-complexity plane of cryptocurrencies' returns



(a) Bitcoin (b) DASH

(c) Ethereum (d) IOTA

(e) Litecoin (f) Monero

(g) Neo (h) Ripple

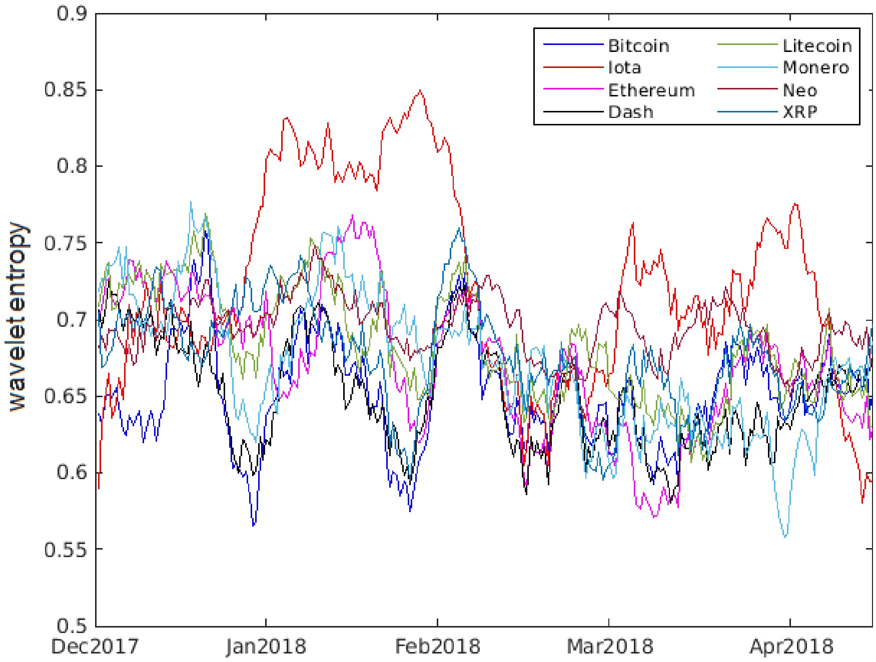


Fig. 2. Evolution of the wavelet entropy across time for the selected cryptocurrencies' returns

4.2 Cryptocurrencies' volatility

It has been previously reported by [11] that cryptocurrencies exhibit greater volatility than traditional assets such as stocks or bonds. In this section we attempt to measure the memory content in volatility time series. There several volatility measures, such as those proposed by Parkinson [28] and Garman and Klass [29]. A comparison of different alternative measured applied to stock markets can be found in [30]. In this paper, and following [5,31], volatility is defined as the logarithmic difference between the high and low values recorded within the 5 minutes interval of each observation. Considering that cryptocurrencies are, mainly, a speculative asset, it attracts short-term, who look to exploit price changes that occur minute to minute. Therefore, studying the volatility profile is crucial for developing a profitable trading strategy.

In this line, another important finding is that all the time series under analysis exhibit strong persistent volatility, reflecting Hurst exponents be- tween 0.8 and 0.9. Such situation means that periods of high (low) volatility are more likely to be followed by periods of high (low) volatility. Then, daily traders could formulate algorithmic trading strategies, in order to take advantage of uninformed traders.

From a theoretical point of view, this characteristic should be taken into account when modeling volatility. Such persistence favors the consideration of fractional integrated and hyperbolic GARCH models, such as those proposed by [26] or [27], to allow the presence of long-range dependence in volatility.

s.

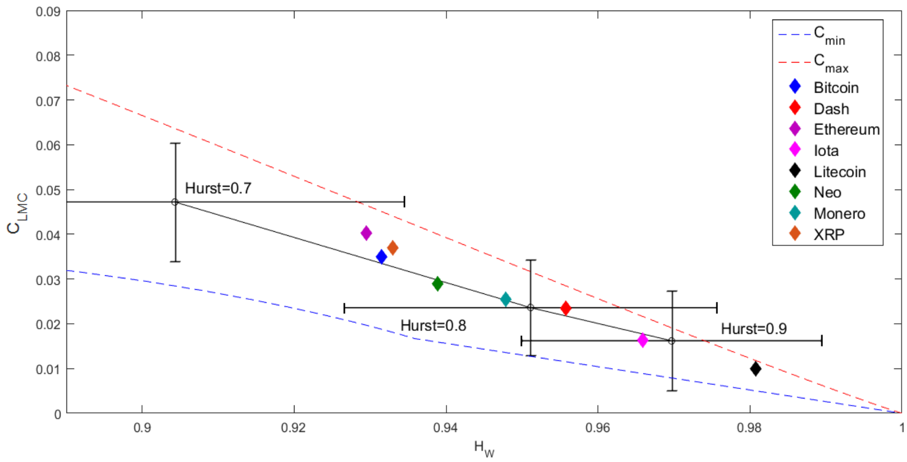


Fig. 3. Wavelet entropy-complexity Plane for cryptocurrencies' volatilities

5 Conclusions

This paper studies the persistence in the time series of eight cryptocurrencies return and volatility time series. Unlike previous literature, we study the stochastic behavior of high frequency time series. This analysis is justified in the existence of day traders, who by means of automatic intelligent algorithms, try to discover profitable trading patterns to be exploited in the very short run. Another difference with the previous literature is that we focus our analysis on a particular time frame from December 2017 until early April 2018. This period comprises a great boom and a subsequent bust in cryptocurrency prices. In this sense, the paper sheds light on the behavior of the long memory under stressing market conditions.

In this paper we compare the information endowment of the time series using a rolling window approach. We detect that, in the period under examination, the time series exhibit both efficient and inefficient behavior. However, there is a remarkable trend toward a more efficient behavior. We also argue that the market is heterogeneous in terms of stochastic dynamics, albeit it is moving towards a more synchronizing behavior.

A significant result is regarding volatility. We find that volatility is highly persistent for all the coins in our sample, and during the whole period. As a consequence, speculative traders could find this market particularly attractive, considering that they could test their trading algorithms seven days a week and at any time.

Finally, we detect that IOTA seems to have a distinct dynamic. These results could be important for portfolio managers, interested in forming port- folios with different coins, because they can control the overall risk by diversifying in different cryptocurrencies.

References

1. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system, https://bitcoin.org/bitcoin.pdf/, accessed: 2019-12-27 (2009).
2. Morris, D. Z.: Bitcoin hits a new record high, but stops short of $20,000, https://fortune.com/2017/12/17/ bitcoin-record-high-short-of-20000/, accessed: 2020-07-16 (2017).
3. Coinmarket: Crypto-Currency Market Capitalizations, https://coinmarketcap.com/currencies/, accessed: 2020-07-16 (2019).
4. Coinmap: All the cryptocurrency merchants and ATMs of the world in one map, https://coinmap.org/view/#/world/34.16181816/-13.18359375/2, accessed: 2020-07-18 (2020).
5. Bariviera, A. F.: The inefficiency of Bitcoin revisited: A dynamic approach, Economics Letters 161, 1-4 (2017).
6. Yermack, D.: Is bitcoin a real currency an economic appraisal, NBER Working Paper Series (2013). doi:10.3386.w19747, https://www.nber.orgpapers/w19747.pdf
7. Baur, D.G. , Hong, K., Lee, A.D.: Bitcoin medium of exchange or speculative assets, Journal of International Financial Markets, Institutions and Money 54, 177-189 (2018).
8. Corbet, S., Lucey, B., Peat, M., Vigne, S.: Bitcoin futures– what use are they, Economics Letters 172, 23-27 (2018).
9. 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?, Finance Research Letters 20, 192-198 (2017).
10. Corbet, S., Meegan, A., Larkin, C., Lucey, B., Yarovaya, L.: Exploring the dynamic relationships between cryptocurrencies and other financial assets, Economics Letters 165, 28-34 (2018).
11. Aslanidis, N., Bariviera, A.F., Martinez-Ibañez, O.: An analysis of cryptocurrencies conditional cross correlations, Finance Research Letters 31, 130-137 (2019).
12. Fama, E.F. :Efficient capital markets: A review of theory and empirical work, The Journal of Finance 25 (2), 383-417 (1970).
13. Urquhart, A.: The inefficiency of Bitcoin, Economics Letters 148, 80-82 (2016).
14. Nadarajah,S., Chu,J.: On the inefficiency of Bitcoin, Economics Letters 150, 6-9 (2017).
15. Bariviera, A.F., Basgall, M.J, Hasperué, W. Naiouf, M.: Some stylized facts of the Bitcoin market, Physica A 484, 82-90 (2017).
16. Takaishi, T., Adachi, T.: Taylor effect in bitcoin time series, Economics Letters 172, 5-7 (2018).
17. Corbet, S., Meegan, A., Larkin, C., Lucey, B., Yarovaya, L.: Exploring the dynamic relationships between cryptocurrencies and other financial assets, Economics Letters 165, 28-34 (2018).
18. Li, X., Li, S., Xu, C.: Price clustering in bitcoin market -- an extension, Finance Research Letters 32, 101072 (2020).
19. Sensoy, A.: The inefficiency of bitcoin revisited: A high-frequency analysis with alternative currencies, Finance Research Letters 28, 68-73 (2019).
20. Fama, E.F.: Tomorrow on the New York stock exchange, The Journal of Business 38 (3), 285-299 (1965).
21. Theil, H., Leenders, C.T.: Tomorrow on the Amsterdam stock exchange, The Journal of Business 38(3), 277-284 (1965).
22. Dryden, M.M.: Short-term forecasting of share prices: An Information Theory approach, Scottish Journal of Political Economy 15 (1), 227-249 (1968).

23. Shannon, C. E., Weaver, W.: The Mathematical Theory of Communication, University of Illinois Press, Champaign, IL (1949).
24. Chui, C.: An introduction on Wavelet Analysis, Academic Press, New York (1992).
25. Kowalski, A.M., Martín, M.T., Plastino, A., Rosso, O.A., Casas, M.: Distances in probability space and the statistical complexity setup, Entropy 13(6), 1055-1075 (2011).
26. Baillie, R., Bollerslev, T., Mikkelsen, H.: Fractionally integrated generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 74(1), 3-30 (1996).
27. Davidson, J.: Moment and memory properties of linear conditional heteroscedasticity models, and a new model, Journal of Business and Economic Statistics 22(1), 16-29 (2004).
28. Parkinson, M., The Extreme Value Method for Estimating the Variance of the Rate of Return, The Journal of Business, 53(1), 61–65 (1980).
29. Garman, M.B., Klass, M.J., On the estimation of security price volatilities from historical data. The Journal of Business, 53(1), 67–78 (1980).
30. Floros, C., Modelling volatility using high, low, open and closing prices: evidence from four S&P indices, International Research Journal of Finance and Economics, 28, 198-206 (2009).
31. Bariviera, A.F., Zunino, L., Rosso, O.A., An analysis of high-frequency cryptocurrencies prices dynamics using permutation-information-theory quantifiers, Chaos, 28(7), 075511 (2018).