Research in International Business and Finance 66 (2023) 102021

Contents lists available at ScienceDirect

Research in International Business and Finance

journal homepage: www.elsevier.com/locate/ribaf

# A Fractal and Comparative View of the Memory of Bitcoin and S&P 500 Returns☆

Klaus Grobys [a,b]

[a] Innovation and Entrepreneurship (InnoLab), University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland
[b] Finance Research Group, School of Accounting and Finance, University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland

ARTICLE INFO

*JEL Classification:*
C22
G12
G13
G14
O10

*Keywords:*
Bitcoin
Fractals
Fractality
Hurst exponent
Memory
S&P 500
Statistical self-affine
Pareto distributions
Power laws
Second moment
Variance

ABSTRACT

The majority of previous studies used autocorrelation-based methodologies to explore the dependency structure for Bitcoin, but this paper follows Benoit Mandelbrot in taking a fractal point of view. This perspective showed that Bitcoin and S&P 500 returns exhibit fractal-like behavior. Additional evidence suggested that the infinite variance hypothesis cannot be rejected for either asset supporting Mandelbrot's (1963) early study on cotton price changes. This result held across non-overlapping subsamples. Following Mandelbrot (2008), Hurst exponents were estimated using rescaled/range analysis. The key findings are that (a) Bitcoin returns exhibit a higher level of persistence than S&P 500 returns across various subsamples, (b) the level of persistence in Bitcoin returns did not change over time, (c) the S&P 500 moved from efficiency in the first subsample to inefficiency in the ex-post June 17, 2018, period, (d) even if it was assumed that the variance of S&P 500 returns was finite, the kurtosis remained statistically undefined. The study concluded that the correlation-based methods used to explore the S&P 500 universe result in misleading answers.

## 1. Introduction

Using distributed ledgers and blockchain technology, cryptocurrencies are human-engineered systems of digital currencies secured by cryptography. Bitcoin, the first traded cryptocurrency—and still the most dominant cryptocurrency (in terms of its market capitalization corresponding to more than USD 500 billion) has been the subject of a considerable body of research.[1] An early and often-cited study by Urquhart (2016) investigated the market efficiency of Bitcoin. Using a sample from August 1, 2010 to July 31, 2016, the study found that Bitcoin was inefficient over the full sample. However, a sample split test showed that Bitcoin was efficient in the latter period (e.g., August 1, 2013, to July 31, 2016). Therefore, Urquhart (2016) concluded that Bitcoin was an inefficient market but may be in the process of moving towards an efficient market. Unsurprisingly, the market (in)efficiency of Bitcoin has sparked an intense

☆ The author is thankful for having received valuable comments from the participants of the 2022 Cryptocurrency Research Conference. The author is especially thankful to Larisa Yarovaya for insightful comments. Finally, the author is grateful for having received useful comments from two anonymous reviewers.

*E-mail address:* klaus.grobys@uwasa.fi.

[1] The market capitalization of Bitcoin is about USD 500 billion as of May 25, 2023; see https://coinmarketcap.com/.

https://doi.org/10.1016/j.ribaf.2023.102021
Received 20 December 2022; Received in revised form 6 June 2023; Accepted 9 June 2023
Available online 15 June 2023

debate in the academic literature.

One one hand, a strand of literature argues that the Bitcoin market is efficient. For instance, noting that powering to an odd integer does not lead to any loss of information, Nadarajah and Chu (2017) used the odd integer power of Bitcoin returns to explore the efficiency of the Bitcoin market. Their results showed that the transformed Bitcoin returns exhibited efficiency. Using detrended fluctuation analysis (DFA) in a sliding windows approach, Bariviera (2017) argued that Bitcoin returns were a white noise process in the ex-post 2014 period and thus advocated market efficiency. Using permutation entropy, Sensoy (2019) compared the time-varying weak-form efficiency of Bitcoin prices in terms of the U.S. dollar (BTCUSD) versus the euro (BTCEUR) at a high-frequency and found that the BTCUSD and BTCEUR markets have become more informationally efficient at the intraday level since the beginning of 2016.[2] Furthermore, Sensoy (2019) found that pricing efficiency depended on the frequency; that is, the higher the frequency, the lower the pricing efficiency.

On the other hand, another stream of literature has emerged that either postulates that the Bitcoin market is inefficient or finds mixed evidence. For instance, employing the efficiency index of Kristoufek and Vosvrda (2013), a study by Kristoufek (2018) found strong evidence that the Bitcoin market remained mostly inefficient between the 2010 and 2017 sample periods. Similarly, Cheah et al. (2018) and Bouri at al. (2017) argued that the Bitcoin market was moderate to highly inefficient due to its price persistence. Zargar and Kumar (2019) examined the informational inefficiency of Bitcoin using data at different frequencies. Their findings indicated that daily Bitcoin returns appeared to follow a memory-less stochastic process, indicating efficiency, whereas Bitcoin returns exhibited informational inefficiency at higher frequencies. Employing several long-range dependence estimators, Tiwari et al. (2018) revisited the informational efficiency of Bitcoin and found that the Bitcoin market was efficient, with exceptions during the periods of April to August 2013 and August to November 2016. Evaluating the adaptive market hypothesis for the Bitcoin market, Khuntia and Pattanayak (2018) argued that the conclusion that Bitcoin price movements were either efficient or inefficient, according to the EMH framework, was not practically true due to the existence of behavioral bias and the creation of events that could change efficiency.

Overall, there is still no consensus regarding whether the Bitcoin market exhibits market efficiency. However, clarifying this issue is important for at least the following reasons. First, the market capitalization of Bitcoin has increased by a considerable margin, and in this regard, Fry and Cheah (2016, p. 350) pointed out that "from an economic perspective the sums of money involved are substantial." Second, market efficiency in human-engineered digital systems−such as cryptocurrencies−would be in stark contrast to the overwhelming amount of literature documenting the inefficiency of traditional asset markets. If virtual currencies were more efficient than traditional asset markets, policy-makers would have incentives to adopt this new technology. Third, PricewaterhouseCoopers recently highlighted that the total assets under management (AuM) of crypto hedge funds globally increased from about USD 1 billion in 2018 to more than USD 2 billion in 2019, suggesting an increased interest in using cryptocurrencies as investment tools.[3] As hedge funds rely on exploiting market inefficiencies, it is important to assess whether cryptocurrency market inefficiencies exist.

In view of this largely inconsistent literature, the purpose of this study is to comparatively investigate whether Bitcoin returns and S&P 500 returns exhibit long-term memory. The presence of persistence in prices would indicate market inefficiency. This study uses the S&P 500 as the archetype for a traditional, important and long-standing asset market, as opposed to Bitcoin–a human-engineered, digital, and decentralized system based on blockchain technology. As many methodologies have been used in the literature to explore market efficiency, contradictory results could be a statistical artefact of inaccurate methodologies. An important question that arises is: Which methods are correct? Thus, we begin the analysis by exploring whether Bitcoin returns or S&P 500 returns exhibit fractal-like behavior. Employing the most recent 10-year sample of weekly data from July 28, 2013, until May 14, 2023, we cluster Bitcoin and S&P 500 returns into two independent sample: an earlier subsample (e.g., July 28, 2013, until June 17, 2018) and a later subsample (e.g., June 24, 2018 until May 14, 2023). For each subsample and asset market, we test whether the asset markets exhibit Paretian tails. This is a very important issue, because if Paretian tails do not exhibit a finite variance (variance of variance), autocorrelation-based methods deliver wrong (misleading) results.

Doing so means employing power law functions and maximum likelihood estimation (MLE) to estimate the corresponding power law exponent for each asset market and subsample. Then, we use Clauset et al.'s (2009) proposed goodness-of-fit test to test the plausibility of the power law null hypothesis. To account for unknown dependency structures in the data, we employ blocks bootstraps in line with Grobys and Junttila (2021) and estimate the bootstrapped standard deviations used for the hypothesis tests. To explore (a) whether Bitcoin and S&P 500 returns exhibit long-term dependencies and (b) whether potential dependency structures are subject to change, we employ rescaled range (R/S) analysis, which was derived and detailed by Mandelbrot (1963, 1969, 1971, 1972) and Mandelbrot and Wallis (1969).

This study makes several important contributions. First, this paper contributes to the wide strand of literature exploring the market efficiency of the Bitcoin market, which has yielded inconclusive results. As stated earlier, an often-cited study is that by Urquhart (2016), analyzed the market efficiency of Bitcoin over a 6 year period. Using autocorrelation-based methods, such as the Ljung–Box test, Urquhart found that Bitcoin exhibited market efficiency in the later subsample.[4] However, using different methodologies, Cheah et al. (2018) and Bouri et al. (2017) concluded that the Bitcoin market is moderate to highly inefficient. As the conclusions drawn in the literature appear to be dependent on the chosen methodologies, this study extends this strand of literature by, first, identifying which methodologies are most accurate for statistical inferences given the market characteristics of Bitcoin. Second, in doing so this study considers Bitcoin in a comparative manner with the S&P 500, which is considered a highly efficient market index. Third,the study

[2] Moreover, the BTCUSD market appears to be slightly more efficient than the BTCEUR market during the sample period.

[3] See https://www.pwc.com/gx/en/financial-services/pdf/pwc-elwood-annual-crypto-hedge-fund-report-may-2020.pdf.

[4] Evidence for Bitcoin's market efficiency has been confirmed in studies of Nadarajah and Chu (2017), Bariviera (2017), and Sensoy (2019).

explores whether potential dependencies were subject to change over time.Fourth, it accounts for a considerably longer time span than previous studies.

Next, the present research adds to the recent stream of finance literature by using power laws to model financial returns and volatility. In an often-cited literature review, Lux and Alfarano (2016) summarized the most influential studies on power laws in financial economics. Although in the seminal study on cotton price changes, Mandelbrot (1963) concluded that the variance in cotton price changes is infinite, Lux and Alfarano (2016) highlighted that subsequent studies (conducted in the wake of Mandelbrot (1963) highly influential paper) raised doubts about the validity of the infinite variance hypothesis by questioning the stability-under-aggregation property of these estimates. Lux and Alfarano (2016) documented that the pertinent literature gradually converged to the insight of an exponent close to $\alpha \approx 3$ suggesting, in turn, that the theoretical variance is defined. To the best of our knowledge, this study is the first to test the economically important infinite variance hypothesis for Bitcoin. However, this study also takes the stance of the vast majority of studies by assuming that variance exists for the returns of both asset markets, the S&P 500 and Bitcoin. Noting Taleb's (2020) argument that operating with variances is allowed only if kurtosis exists, this study is the first to examine whether the fourth moments for S&P 500 returns and Bitcoin returns exist. This is an important issue to clarify because even if we assumed that the variance exists, we would not observe its true value in finite samples, provided kurtosis is undefined (Taleb, 2020).[5]

Finally, the present research extends the literature by investigating long-term dependency in asset prices. Based on the seminal work of Mandelbrot (1963, 1969, 1971, 1972) and Mandelbrot and Wallis (1969), the scaled range (e.g., R/S) statistic is a metric for measuring the dependency structure for data that is non-Gaussian distributed. For instance, in traditional financial research contexts, the Hurst exponent derived from Mandelbrot's R/S analysis has been applied to explore dependency structures in foreign exchange rates (e.g., Raimundo and Okamoto, 2018; Da Silva, Matsushita, Gleria and Figueiredo, 2007; Muniandy et al., 2001), and stocks (e.g., Grobys et al., 2023; Eom et al., 2008; Lahmiri, 2016; Rejichi and Aloui, 2012). The present research adds to this literature by re-examining the market efficiency of the S&P 500 in the most recent decade. This is an important issue because recent economic events, such as the Covid-19 outbreak and the ongoing Russian-Ukrainian conflict had severe economic consequences. Can the S&P 500 maintain its efficiency despite the central bank's extreme economic policy actions?[6] Moreover, investigating the persistence of cryptocurrencies has been the subject of recent studies by Bariviera (2017) and Wu and Chen (2020). The present study adds to that literature by (a) exploring the persistence of Bitcoin returns in a comparative manner with S&P 500 returns, which is often considered a nearly efficient market, and by (b) using two-sample $z$-tests applied to nonoverlapping subsamples to explore whether the level of persistence has changed over time. To the best of our knowledge, this is the first study to employ $z$-tests to nonoverlapping subsamples to investigate whether persistence in Bitcoin returns was subject to change in the most recent decade.

Using a 10-year period of weekly data for S&P 500 returns and Bitcoin returns, the overall sample was divided into two nonoverlapping subsamples of equal length. The results indicate that we cannot reject the power law null hypothesis for Bitcoin returns or S&P 500 returns for the whole sample or for most subsamples. Although Wu and Chen (2020) found that the tail index for Bitcoin returns ranges from 1.69-3.11 for positive Bitcoin returns and 1.32–2.96 for negative Bitcoin returns, which could, according to the authors, suggest the presence of infinite autocovariances, we argue that the infinite variance hypothesis cannot be rejected for any asset market returns. This result holds across all subsamples. Additional hypothesis tests show that even if we assume that the variance for S&P 500 returns is finite, the theoretical fourth moment of S&P 500 returns is not defined. This issue is very serious, and Fama (1963) pointed out that.

> "… the infinite variance assumption of the stable Paretian model has extreme implications. From a purely statistical standpoint, if the population variance of the distribution of first differences is infinite, the sample variance is probably a meaningless measure of dispersion. Moreover, if the variance is infinite, other statistical tools (e.g., least-squares regression) which are based on the assumption of finite variance will, at best, be considerably weakened and may in fact give very misleading answers." (Fama, 1963, p. 421)

Therefore, the results of this study cast doubts on the validity of the reported research results of an enormous stream of finance literature. The replication failure of asset pricing studies, as documented by Hou et al. (2020), could be a manifestation of drawing conclusions derived from correlation-based methodologies using data that exhibit undefined kurtoses.

Finally, in contrast to previous studies that showed either that Bitcoin is efficient in the later subsample or moves toward efficiency

[5] Other relevant studies using power laws for modeling the dynamics of financial returns is Warusawitharana (2019) who analyzed whether models with time-varying volatility could help explain the observed tail distribution of stock returns. Moreover, Gabaix et al. (2003, 2006) provided a sound theoretical framework for the power law property of stock returns. Another recent stream of literature used power laws to model the dynamics of the second moment. For instance, Grobys et al. (2021) investigated the volatility processes of stablecoins and their potential stochastic interdependencies with Bitcoin volatility, whereas Grobys (2021) tested the power law null hypothesis for the realized variances of five different asset markets. Notably, that study found that the variance of variance did not exist for any of those asset markets, rendering reported results derived from $t$-statistics using standard econometric methodologies sample-specific. Moreover, his findings indicate that the realized variance of the S&P 500 is more prone to extreme events than Bitcoin's realized variance.

[6] In the lecture "The Economic Machine," Ray Dalio detailed the mechanism of how stock prices could be inflated due to the central bank's monetary policies (see https://www.youtube.com/watch?v=PHe0bXAIuk0). Specifically, he argued that the central bank is forced to print money when it reaches a regime of zero interest rates. However, the central bank uses the printed money to buy financial assets and government bonds, which drives up asset prices. Inflated assets priced due to excessive money printing could also have an impact on the efficiency of market indices such as the S&P 500.

(Urquhart, 2016; Nadarajah and Chu, 2017; Bariviera, 2017; Sensoy, 2019; Wu and Chen, 2020), the results of this study do not support this claim. Specifically, a commonality that we identify is that R/S analysis and DFA suggest that the market for Bitcoin is more inefficient than the S&P 500, which is manifested in Hurst exponents exhibiting a considerably higher economic magnitude for Bitcoin returns regardless of the subsample. Moreover, the results derived from the R/S analysis show that the level of persistence in Bitcoin returns has not changed over time; that is, Bitcoin in the earlier sample is as inefficient as in the most recent subsample. Unexpectedly, the S&P 500 returns in the most recent subsample show a statistically significantly higher level of persistence than in the earlier subsample, suggesting that the efficiency of the S&P 500 has decreased in the last decade.

This study is organized as follows. The next two sections describe the data and methodologies. The fourth section reports the results. The fifth section provides a discussion of the findings and the last section concludes.

## 2. Data

Weekly data on Bitcoin and the S&P 500 covering the period July 28, 2013, until May 14, 2023, were downloaded from investing.com. The whole data sample consists of 512 weekly observations. Using weekly data to investigate cryptocurrencies is in line with previous studies that pointed out that monthly data would not provide enough observations (e.g., Grobys and Juttila, 2021; Shen et al., 2020; Platanakis and Urquhart, 2020). The overall sample was split into two subsamples of equal length. The first subsample is from July 28, 2013, until June 17, 2018, whereas the second subsample is from June 24, 2018, until May 14, 2023. For both data we compute the absolute returns. The descriptive statistics for return data and absolute return data are reported for the whole sample and the two subsamples in Tables A.1 to A.3 in the appendix. Unsurprisingly, in Tables A.1 –A.3 we see that the sample mean returns and standard deviations are considerably larger for Bitcoin returns than for S&P 500 returns, regardless of which sample is considered. In addition, the economic magnitudes for the maxima and minima are larger for the Bitcoin returns for all samples. The kurtosis values for all samples exceed 3, which indicates that the distributions governing the return-generating data exhibit fatter tails than implied by the normal distribution.

## 3. Statistical analysis

### 3.1. Do Bitcoin and S&P 500 Returns Exhibit Paretian Tails? Evidence from Power Law Models

As the high kurtosis values of Bitcoin and S&P 500 returns possibly indicate the presence of fat tails, we explore whether the return-generating processes exhibit fractal behavior. This is an important issue because fractality implies that standard statistical tools, such as autocorrelation-based methods, cannot be used, as these methodologies are designed for Gaussian or thin-tailed data-generating processes. Working with empirical data, fractality typically manifests in the presence of Paretian tails. Thus, to comparatively explore whether Bitcoin or S&P 500 returns exhibit fractal-like behavior, we employ the absolute amount of Bitcoin and S&P 500 returns, denoted here as, $|ret_{BTC}|$ and $|ret_{S\&P500}|$, respectively. Then, we estimate the following power-law models for the whole sample and each subsample separately:

$$p(x) = Cx^{-\alpha} \tag{1}$$

where $C = (\alpha - 1)x_{MIN}^{\alpha-1}$ with $\alpha \in \{\mathbb{R}_+|\alpha > 1\}$, and $x \in \{|ret_{BTC}|, |ret_{S\&P500}|\}$; that is, $x$ is either the absolute amount of Bitcoin returns or the absolute amount of the S&P 500 returns, provided $x \in \{\mathbb{R}_+|x_{MIN} \leq x < \infty\}$, and $x_{MIN}$ is the corresponding minimum value governed by the corresponding power law process, whereas $\alpha$ is the magnitude of the specific tail exponent.[7] It is important to note that Taleb (2020, p. 34) emphasized that the tail exponent $\alpha$ of a power law function captures, via extrapolation, the low-probability deviation not seen in the data, which, however, plays a disproportionately large share in determining the mean. Next, it can be shown that the conditional expectation, $E[X|X \geq x_{MIN}]$, is then given by

$$E[X|X \geq x_{MIN}] = \int_{x_{MIN}}^{\infty} xp(x)dx = \frac{(\alpha - 1)}{(\alpha - 2)}x_{MIN}, \tag{2}$$

whereas the conditional second moment $E[X^2X \geq x_{MIN}]$, or the conditional variance, is defined as follows:

$$E[X^2|X \geq x_{MIN}] = \int_{x_{MIN}}^{\infty} x^2p(x)dx = \frac{(\alpha - 1)}{(\alpha - 3)}x_{MIN}^2 \tag{3}$$

Higher conditional moments of order $k$ are analogously defined as follows:

$$E[X^k|X \geq x_{MIN}] = \frac{(\alpha - 1)}{(\alpha - 1 - k)}x_{MIN}^k \tag{4}$$

[7] This study follows the notation in Clauset et al. (2009).

From Eqs. (3) and (4), we see that the conditional mean exists only for $\alpha > 2$, whereas the conditional variance exists only for $\alpha > 3$. Following White et al. (2008) and Clauset et al. (2009), who concluded that MLE performs best for estimating power law exponents, tail exponents are estimated for each data (sub)sample as follows:

$$\widehat{\alpha} = 1 + n\left(\sum_{i=1}^{N} \ln\left(\frac{x_i}{x_{MIN}}\right)\right)^{-1}, \tag{5}$$

where $\widehat{\alpha}$ denotes the MLE estimator, $n$ is the number of observations exceeding $x_{MIN}$, and the other notation is as before. Clauset et al. (2009) showed that $\widehat{\alpha}$ is normally distributed with an estimated standard error $\widehat{\sigma} = (\widehat{\alpha} - 1)/n^{1/2}$.

### 3.2. Is the Power Law Null Hypothesis Reasonable? Evidence from Goodness-of-Fit Tests

As noted in Clauset et al. (2009), an essential issue is how to determine the corresponding values for $\widehat{\alpha}$ and $\widehat{x}_{MIN}$ to accurately estimate the probability density functions. Note from Eq. (5) that the MLE estimator depends on the chosen $x_{MIN}$, and as a result, there are possibly many MLE estimators from which to choose. Clauset et al. (2009) showed that it is common practice to employ the $\widehat{\alpha}/x_{MIN}$-plot and to choose the value for $x_{MIN}$, beyond which $\widehat{\alpha}$ is stable. Unfortunately, this approach appears to be subjective and can be sensitive to noise or fluctuations in the tail of the distribution. Therefore, Clauset et al. (2009) proposed a goodness-of-fit (GoF) test based on the optimized Kolmogorov–Smirnov (KS) distance $D$ between the power-law model and the empirical data. The implementation of this GoF test is detailed in Clauset et al., (2009, p. 675−678) and can be summarized as follows. First, $D$ measures the maximum distance between the cumulative density functions (CDFs) of the data and the fitted model and is given by the following equation:

$$D = MAX_{x \geq x_{MIN}} |S(x) - P(x)|, \tag{6}$$

where $S(x)$ is the CDF of the data for the observation with a value of at least $x_{MIN}$, and $P(x)$ is the CDF for the power law model that best fits the data in the region $x \geq x_{MIN}$. The estimate $\widehat{x}_{MIN}$ is then the value of $x_{MIN}$ that minimizes $D$. Using the parameter vector $(\widehat{\alpha}, \widehat{x}_{MIN})$, which is optimal with respect to $D$, Clauset et al.'s (2009) GoF test generates a $p$-value that quantifies the plausibility of the power-law null hypothesis. Specifically, their proposed test compares the estimated $D$ with distance measurements for comparable synthetic data sets drawn from the hypothesized model using $(\widehat{\alpha}, \widehat{x}_{MIN})$, and the $p$-value is defined as the fraction of the synthetic distances that are larger than the empirical distance. For a given significance level of 5%, the power law null hypothesis is not rejected, as the difference between the empirical data and the model can be attributed to statistical fluctuations alone.

### 3.3. Employing Blocks Bootstrap for Estimating Standard Deviations for Two-Sample Tests

Although Clauset et al. (2009) showed that $\widehat{\alpha}$ is normally distributed with an estimated standard error $\widehat{\sigma} = (\widehat{\alpha} - 1)/n^{1/2}$, it is important to note that standard error $\widehat{\sigma}$ is derived from the assumption of independent distributed data. As volatility clustering−which is a stylized fact for financial assets−suggests some form of dependency, the uncertainty of $\widehat{\alpha}$ is underestimated. To address this issue, we implement a block bootstrap procedure, as proposed by Grobys and Junttila (2021). Denoting the selected block length as $m$, we implement a block bootstrap procedure such that $E[m] = T^{1/2}$. Then, from a given data vector $\boldsymbol{x}$, blocks $m_j$ are randomly drawn, which are distributed as a geometric distribution $m_j \sim GEO(p)$ with $E[m_j] = \frac{(1-p)}{p}$. For instance, for the overall data sample, we employ $E[m_j] = 23$ implying $p = 0.0417$, whereas $E[m_j] = 16$ implying $p = 0.0588$ for subsample data block bootstraps. Using this procedure, the blocks drawn from a given data vector $\boldsymbol{x}$ vary in length. Randomly drawn blocks $m_j$ from data vector $\boldsymbol{x}$ are stacked in vector $\boldsymbol{x}^b$ as follows:

$$\boldsymbol{x}_i^b = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ \vdots \end{bmatrix}$$

The procedure is stopped when the length of the artificial vector $\boldsymbol{x}^b$ exhibits a length exceeding $T$. Observations exceeding $T$ are cut off; that is, every artificial data vector $\boldsymbol{x}^b$ has the same length as the original data vector $\boldsymbol{x}$. Using this block bootstrap procedure, for each data vector, $B = 1,000$ artificial data vectors are constructed,

$$\begin{bmatrix} \boldsymbol{x}^1 & \boldsymbol{x}^2 & \cdots & \boldsymbol{x}^B \end{bmatrix},$$

and point estimates for α are obtained for each bootstrapped data vector $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^B$ using Clauset et al.'s (2009) approach such that,

$$\begin{bmatrix} \widehat{a}^1 & \widehat{a}^2 & \cdots & \widehat{a}^B \end{bmatrix}.$$

Finally, the corresponding bootstrapped standard error $\widehat{\sigma}_{BOOT}$ is computed for the bootstrapped $\widehat{a}^1, \widehat{a}^2, \ldots, \widehat{a}^B$ for each given data vector. According to Grobys and Junttila (2021), this approach is robust to unknown dependency structures in the data, which are commonly observed for financial assets.

Next, the question arises as to whether S&P 500 and Bitcoin returns exhibit distinct tail risks. Thus, using standard deviations derived from blocks bootstraps, we test the following null hypothesis:

$$H_0 : (\alpha_{S\&P500} - \alpha_{BTC}) = 0 \quad \text{vs.} \quad H_1 : (\alpha_{S\&P500} - \alpha_{BTC}) \neq 0,$$

where $\alpha_{S\&P500}$ ($\alpha_{BTC}$) are the tail exponents for S&P 500 (Bitcoin) returns. We use a two-sided test and a significance level of 5%. Rejection of the null hypothesis would suggest that the data-generating processes of Bitcoin returns and S&P 500 returns are exposed to distinct tail risks. The corresponding two-sample $z$-test statistic is defined as follows:

$$z = \frac{(\alpha_{S\&P500} - \alpha_{BTC})}{\sqrt{\frac{\sigma^2_{\alpha_{S\&P500}}}{n_{S\&P500}} + \frac{\sigma^2_{\alpha_{BTC}}}{n_{BTC}}}},$$

where $\sigma^2_{\alpha_{S\&P500}}$ ($\sigma^2_{\alpha_{BTC}}$) denotes the variance of the tail exponent of the S&P 500 (Bitcoin), estimated via block bootstrap, as described in the previous section, $n_{S\&P500}$ ($n_{BTC}$) denotes the number of observations in relative terms, and the other notation is as previously. Note that, because we have the same number of sample observations, the $z$-statistic simplifies to

$$z = \frac{(\alpha_{S\&P500} - \alpha_{BTC})}{\sqrt{\left(\sigma^2_{\alpha_{S\&P500}} + \sigma^2_{\alpha_{BTC}}\right)/2}}$$

Next, we investigate whether the tail risks for each asset are subject to change over time. To explore this issue, we test the following null hypothesis for each asset:

$$H_0 : (\alpha_{1st} - \alpha_{2nd}) = 0 \quad \text{vs.} \quad H_1 : (\alpha_{1st} - \alpha_{2nd}) \neq 0,$$

where $\alpha_{1st}$ ($\alpha_{2nd}$) denotes the tail exponent for the first (second) subsample. We use a two-sided test and a significance level of 5%. Rejection of the null hypothesis would suggest that the tail risk for a given asset has changed over time. The corresponding two-sample $z$-test statistic is defined as follows:

$$z = \frac{(\alpha_{1st} - \alpha_{2nd})}{\sqrt{\left(\sigma^2_{\alpha_{1st}} + \sigma^2_{\alpha_{2nd}}\right)/2}},$$

where $\sigma^2_{\alpha_{1st}}$ ($\sigma^2_{\alpha_{2nd}}$) denotes the variance for the first (second) subsample derived from block bootstraps for a given asset, and all the other notation is as previously.

### *3.4. Are the Theoretical Variances Defined? Testing the Infinite Variance Hypothesis*

Next, we test the infinite variance hypothesis. This is an important issue because Fama (1963) pointed out that

> "… the infinite variance assumption of the stable Paretian model has extreme implications. From a purely statistical standpoint, if the population variance of the distribution of first differences is infinite, the sample variance is probably a meaningless measure of dispersion. Moreover, if the variance is infinite, other statistical tools (e.g., least-squares regression) which are based on the assumption of finite variance will, at best, be considerably weakened and may in fact give very misleading answers." (Fama, 1963, p. 421)

Note that Mandelbrot (1963) was the first to show that cotton price changes do not exhibit a theoretically defined variance. However, Lux and Alfarano (2016) argued that other studies raised doubts about the validity of the infinite variance hypothesis by questioning the stability-under-aggregation property of these estimates. The authors argued that the pertinent literature gradually converged to the insight of an exponent significantly larger than 2 and mostly close to $\approx 3$ suggesting, in turn, that the theoretical variance is defined. To shed new light on this issue, we examine whether the infinite variance hypothesis holds for S&P 500 and Bitcoin returns; that is, we test the following hypothesis for all (sub)samples:

$$H_0 : (\alpha - 3) \leq 0 \quad \text{vs.} \quad H_1 : (\alpha - 3) > 0,$$

where $\alpha$ is the power law exponent for a given asset based on a given (sub)sample. In doing so, we assess the statistical significance via standard $t$-tests and $p$-values derived from block bootstraps, as detailed in Section 3.3. We use a one-sided test and a significance level of 5% (with a corresponding critical value of 1.6450). Rejection of the null hypothesis would confirm previous studies in which the infinite variance hypothesis was rejected (Lux and Alfarano, 2016).

### *3.5. Analyzing the Long-Run Memories of Bitcoin and the S&P 500: Evidence from Hurst Exponents and Subsample Tests*

To explore whether Bitcoin returns exhibit long-term memory or whether some potential dependency structure was subject to change, we employ the R/S analysis to derive the Hurst exponent. Mandelbrot (2008, p.298) pointed out that "[o]ne of its principal virtues is that, in contrast to many common statistical tests, it makes no assumption about how the original data are organized–a critical point when studying something like stock prices for which evidence abounds that the conventional assumptions are flatly wrong."

The R/S statistic was derived and detailed in Mandelbrot (1963, 1969, 1971, 1972) and Mandelbrot and Wallis (1969) and can be, according to Mandelbrot (2008, p.298–299), summarized as

$$R/S_k = \frac{MAX_{1\le k\le n}\sum_{j=1}^{k}(r_j - \bar{r}_n) - MIN_{1\le k\le n}\sum_{j=1}^{k}(r_j - \bar{r}_n)}{\left[\frac{1}{n}\sum_j (r_j - \bar{r}_n)^2\right]^{1/2}}, \quad (7)$$

where $\bar{r}_n$ is the average return over $n$ days. Then, for each shorter time period, the difference between the return $r_j$ over that period and the average return, $\bar{r}_n$, is calculated, and a running total of all differences as the time periods increase up to period $k$, where $k \in \{4, 8, 16, 32, 64, 128, 256\}$. Then, the maximum and the minimum of all the differences are computed. Subtracting the minimum from the maximum gives us an estimate of the range from the peak to the trough in the accumulated deviations, which is the numerator in Eq. (7). The denominator is the conventional measure of the standard deviation in the data series.[8]

If the data were independent, the ratio between the numerator and the denominator should be, according to Mandelbrot, expected to be 1:2, corresponding to a Hurst exponent of $H = 0.50$. Moreover, $H > 0.50$ implies long-term dependence, a long-memory of the stochastic process; that is, the data are persistent. Whereas $H < 0.50$ implies anti-persistence, which is characterized by the tendency to keep back on themselves. According to the theory, $(R/S)_k \approx Ck^H$ is distributed as a power law. The value of the Hurst exponent $H$ can be calculated using the following log-log regression model:

$$\ln(R/S)_k = \ln(C) + H\ln(k) + u, \quad (8)$$

where $u$ is assumed to be distributed as $u \sim iid(0, \sigma_u)$. Next, the question arises as to whether the long-term memory of S&P 500 and Bitcoin returns exhibit distincts features across the subsamples. Thus, using standard deviations and point estimates derived from log-log regressions, we test the following null hypothesis for a given subsample:

$$H_0 : (H_{S\&P500} - H_{BTC}) = 0 \quad \text{vs.} \quad H_1 : (H_{S\&P500} - H_{BTC}) \neq 0,$$

where $H_{S\&P500}$ ($H_{BTC}$) denotes the Hurst exponent for S&P 500 (Bitcoin) returns for a given subsample. We use a two-sided test and a significance level of 5%. Rejection of the null hypothesis would suggest that the data-generating processes of Bitcoin and S&P 500 returns exhibit distinct memory features. As the subsamples have equal lengths, the corresponding two-sample $z$-test statistic is defined as follows:

$$z = \frac{(H_{S\&P500} - H_{BTC})}{\sqrt{(\sigma^2_{H_{S\&P500}} + \sigma^2_{H_{BTC}})/2}},$$

where $\sigma^2_{H_{S\&P500}}$ ($\sigma^2_{H_{BTC}}$) is the variance of the Hurst exponent of S&P 500 returns (Bitcoin returns) for a given subsample. Next, we investigate whether the memory process for each asset was subject to change over time. To explore this issue, we test the following null hypothesis for each asset:

$$H_0 : (H_{1st} - H_{2nd}) = 0 \quad \text{vs.} \quad H_1 : (H_{1st} - H_{2nd}) \neq 0,$$

where $H_{1st}$ ($H_{2nd}$) denotes the Hurst exponent for the first (second) subsample of a given asset. We use a two-sided test and a significance level of 5%. Rejection of the null hypothesis would suggest that the memory process of a given asset has changed over time. The corresponding two-sample $z$-test statistic is defined as follows:

$$z = \frac{(H_{1st} - H_{2nd})}{\sqrt{\sigma^2_{H_{1st}} + \sigma^2_{H_{2nd}})/2}},$$

where $\sigma^2_{H_{1st}}$ ($\sigma^2_{H_{2nd}}$) denotes the variance of the Hurst exponent of the first (second) subsample for a given asset.

[8] Note that this approach is in line with Grobys et al. (2023) who explores the dependency structure of S&P 500 survivor stocks. To analyze whether the long-term memory of survivor firms had changed across time, the authors split the overall sample into two independent non-overlapping samples of equal length. It is important to stress that non-overlapping sample are necessary to implement two-sample z-statistics used in the current research.

### 3.6. Robustness Checks

#### 3.6.1. One-Sigma Test to Differentiate Between Thin- and Fat-Tailed Distributions

Following recent studies (e.g., Grobys 2021, 2023, Grobys et al., 2023), weemploy the GoF test proposed by Clauset et al. (2009) to test the plausibility of the power law null hypothesis. Note that this GoF test employs the power law model as the null model. The reader might wonder whether the power law hypothesis holds if the power law model serves as the alternative model. Taleb (2020) argued that a common feature of fat-tailed distributions is that more observations are within one standard deviation than predicted by the normal distribution. In fact, for a normal distribution, such as $T \longrightarrow \infty$, we would expect a fraction of 0.6827 of the observations to be within one standard deviation from the mean. In this regard, Taleb (2020) highlighted:

> "The probability of an event staying within one standard deviation of the mean is 68%. As the tails fatten, to mimic what happens in financial markets for example, the probability of an event staying within one standard deviation of the mean rises to between 75% and 95%. So note that as we fatten the tails we get higher peaks, smaller shoulders, and a higher incidence of a very large deviation." (Taleb, 2020, p. 23).

Inspired by Taleb's (2020) notion of one-sigma events, we propose what we term the *one-sigma test*, which is constructed using the normal distribution as the null model. Specifically, let us denote $y_{1t}$ as drawing $t$ from some normal distribution. According theory, $(y_{1t} - \mu_{y_1})/\sigma_{y_1} \sim N(0,1)$ and defining $y_{2t} = (y_{1t} - \mu_{y_1})/\sigma_{y_1}$ and $p(y_{2t})$ as the corresponding probability function, it is well-known that $\int_{-1}^{1} p(y_{2t}) = 0.6827$ as $T \longrightarrow \infty$. However, in finite samples, $\mu_{y_1}$ and $\sigma_{y_1}$ are estimated from a given data vector, and $\int_{-1}^{1} p(y_{2t})$ will exhibit variation, which depends on the sample size $T$. Let us denote the reference statistic for evaluating the fraction of observations that are within one standard deviation from the mean of some normal distribution as $\lambda_0 = \int_{-1}^{1} p(y_{2t})$. Then, we employ a simulation experiment in which we simulate 100,000 time series vectors, and each time series vector has the dimension $500 \times 1$ consisting of random drawings $t$ from the standard normal.[9] For each simulated vector, we calculate $\lambda_0$. The simulated distribution of $\lambda_0$ allows us to determine the corresponding 95% probability interval. If, for a given data sample, the estimated $\lambda$ is outside the 95% probability interval for $\lambda_0$, we would reject the normal distribution as the underlying data-generating process. We implement the one-sigma test for theS&P 500 and Bitcoin returns. Therefore, we test the following hypothesis pair:

$H_0 : \lambda \sim$thin-tailed distribution vs. $H_1 : \lambda \sim$fat-tailed distribution.

We use a significance level of 5%, and the critical values for $\lambda_0$ are derived via a simulation, as described previously.

#### 3.6.2. Detrended Fluctuation Analysis to Reexamine the Long-Term Memory of S&P 500 and Bitcoin Returns

In this study, we follow Mandelbrot (2008), who advocated using R/S analysis to estimate the Hurst exponent. However, Bassingthwaighte and Raymond (1994) found that R/S analysis tends to give biased estimates of the Hurst exponent, which are too low for $H > 0.72$ and too high for $H < 0.72$.[10] A more recent study by Bryce and Sprague (2012) pointed out that although DFA, as proposed first by Peng et al. (1994), has become a popular tool for estimating the Hurst exponent, it introduces (a) uncontrolled bias, (b) is computationally more expensive than the R/S statistic, and (c) cannot provide generic or useful protection against nonstationaries. Therefore, the literature is inconclusive regarding which estimation procedure should be preferred.

Nevertheless, to check the robustness of our findings, we re-examine the memory features using DFA which has become a popular tool to for estimating Hurst exponents. DFA can be summarized as follows: First, a given data series $x_t$ is converted to the mean-centered cumulative sum:

$$\widetilde{x}_t = \sum_{t=1}^{T} x_t$$

Analogous to R/S analysis, different time scales $k$ are defined; that is, $k \in \{4, 8, 16, 32, 64, 128, 256\}$. Depending on the defined time scale, the data are split into epochs, and for each epoch $s$, a time series regression is used to detrend the data. For instance, employing $k = 256$ means that the weekly data for $\widetilde{x}_t$ are split into two nonoverlapping epochs. For each epoch $s$, the following regression is employed:

$$\widetilde{x}_t = \gamma_0 + \gamma_1 t + e_t,$$

where $t = 1, \ldots, 128$ for the first epoch and $t = 129, \ldots, 256$ for the second epoch. Then, for each epoch $s$, the root mean squared error (RMSE) is computed as

$$RMSE_S = \sqrt{\frac{1}{T_s} \sum_{t}^{T_s} \widehat{e}_t},$$

where $T_s = 256$. Finally, the estimates for $RMSE_S$ are averaged for each time scale $k$, giving us.

[9] Note that we employ $\approx 500$ observations in the overall sample.

[10] Notethat the estimates for $H$ in the present research do not reach 0.72, and therefore, the results based on R/S analysis can be considered reliable.

$\overline{RMSE}_k$. According to theory, the following relation holds:

$$\overline{RMSE}_k = ck^H$$

The Hurst exponent is then estimated by computing a linear fit between log-scales and log-$\overline{RMSE}_k$.

Analogous to R/S analysis, if the data were independent, the ratio between the numerator and the denominator should be, according to the theory, 1:2, corresponding to a Hurst exponent of $H = 0.50$. Moreover, $H > 0.50$ implies long-term dependence, that is, a long memory of the stochastic process in which the data are persistent; however, $H < 0.50$ implies antipersistence, which is characterized by the tendency to keep back on themselves.

### *3.6.3. Is the Theoretical Kurtosis Defined?*

In Section 3.4, we tested whether theoretical variances are defined; in this section, we address the question of whether the theoretical kurtoses are defined. This is an important issue because Lux and Alfarano (2016) argued that the pertinent literature gradually converges to the insight that the theoretical variances of most financial asset classes are well-defined, whereas the returns on speculative investments such as venture capital and R&D investments are perhaps subject to infinite theoretical variances. However, Taleb (2020) argued:

> "In 2009 I took 55 years of data and looked at how much of the kurtosis (a function of the fourth moment) came from the largest observation [...] For a Gaussian the maximum contribution over the same time span should be around.008 ± .0028. For the S&P 500 it was about 80%. This tells us that we don't know anything about the kurtosis of these securities. Its sample error is huge; or it may not exist so the measurement is heavily sample dependent. If we don't know anything about the fourth moment, we know nothing about the stability of the second moment. It means we are not in a class of distribution that allows us to work with the variance, even if it exists. Science is hard; quantitative finance is hard too." (Taleb, 2020, p. 50)

Therefore, it is of fundamental importance to examine whether the fourth moments of S&P 500 and Bitcoin returns exist. From Eq. (4), we observe that the theoretical fourth moment is defined for $\alpha > 5$. Thus, we test the following hypthesis for S&P 500 and Bitcoin returns and for all (sub)samples:

$$H_0 : (5 - \alpha) \leq 0 \quad \text{vs.} \quad H_1 : (5 - \alpha)\rangle 0$$

In doing so, we assess the statistical significance via standard *t*-tests and *p*-values derived from blocks bootstraps, as detailed in Section 3.3. Rejection of the null hypothesis would confirm Taleb's (2020) argument in the sense that kurtosis is undefined. In

**Table 1**
Estimated power lawexponents for S&P 500 returns and Bitcoin returns.

| | $\hat{\alpha}$ | $\hat{\sigma}_{BOOT}$ | $\hat{x}_{\min}$ | $N$ (%) | *p*-value (GoF) |
|---|---|---|---|---|---|
| Panel A: Estimates for the power law exponents for the whole sample | | | | | |
| S&P 500 | 3.4784 | 0.8148 | 2.6000 | 96 (18.75%) | 0.4020 |
| BTC | 3.8181 | 1.0766 | 21.6300 | 44 (8.59%) | 0.5240 |
| Panel B: Estimates for the power law exponents for the first sample | | | | | |
| | $\hat{\alpha}$ | $\hat{\sigma}_{BOOT}$ | $\hat{x}_{\min}$ | $N$ (%) | *p*-value (GoF) |
| S&P 500 | 4.2048 | 1.6178 | 3.0300 | 21 (8.20%) | 0.5860 |
| BTC | 3.2422 | 0.8635 | 9.8800 | 88 (34.38%) | 0.0220 |
| Panel C: Estimates for the power law exponents for the second sample | | | | | |
| | $\hat{\alpha}$ | $\hat{\sigma}_{BOOT}$ | $\hat{x}_{\min}$ | $N$ (%) | *p*-value (GoF) |
| S&P 500 | 3.1294 | 0.8024 | 2.4100 | 73 (28.52%) | 0.1460 |
| BTC | 4.0921 | 1.8760 | 20.6200 | 17 (6.64%) | 0.4490 |

We employ the absolute amount of Bitcoin and S&P 500 returns denoted here as, $|ret_{BTC}|$ and $|ret_{S\&P500}|$ and estimate the following power law models forthe whole sample and subsamples:
$p(x) = Cx^{-\alpha}$,
where $C = (\alpha - 1)x_{MIN}^{\alpha-1}$ with $\alpha \in \{\mathbb{R}_+ | \alpha > 1\}$, $x \in \{|ret_{BTC}|, |ret_{S\&P500}|\}$, that is, $x$ is either the absolute amount of Bitcoin returns or the absolute amount of S&P 500 returns, provided $x \in \{\mathbb{R}_+ | x_{MIN} \leq x < \infty\}$, $x_{MIN}$ is the corresponding minimum value governed by the power law process, and $\alpha$ is the magnitude of the specific tail exponent. The tail exponents are estimated for each (sub)sample as:
$\hat{\alpha} = 1 + n\left(\sum_{i=1}^{N} \ln\left(\frac{x_i}{x_{MIN}}\right)\right)^{-1}$,
where $\hat{\alpha}$ denotes the MLE estimator, $n$ is the number of observations exceeding $x_{MIN}$. To test the plausibility of the power law model, we employ Clauset et al.'s (2009) GoF test, as detailed in Section 3.2. The estimated standard deviations are derived from block bootstraps, as detailed in Section 3.3. The period for the whole sample is from July 28, 2013, to May 14, 2023 corresponding to 512 weekly observations. The first subsample is from July 28, 2013, to June 17, 2018, whereas the second subsample is from June 24, 2018, to May 14, 2023. Both subsamples have equal length and comprise 256 weekly observations.

evaluating statistical significances, we use a one-sided test and a standard significance level of 5% corresponding to a critical value of 1.6450.

## 4. Results

### 4.1. Main Results

Using Clauset et al.'s (2009) approach to estimate the power law exponents in association with block bootstraps, we see from Table 1 that the whole sample (e.g., July 28, 2013, to May 14, 2023) estimates for α are $\widehat{\alpha} = 3.48$ for $|ret_{S\&P500}|$ and $\widehat{\alpha} = 3.81$ for $|ret_{BTC}|$, respectively. Whereas 18.75% of $|ret_{S\&P500}|$ are governed by a power law process; the corresponding figure is 8.59% for $|ret_{BTC}|$. This result indicates that about twice as much observations of $|ret_{S\&P500}|$ are governed by a power law as of $|ret_{BTC}|$. Clauset et al.'s (2009) GoF tests do not reject the power law models, which implies that both asset return processes exhibit significant fractal-like behavior. The evidence is qualitatively the same for the second subsample (e.g., June 24, 2018, to May 14, 2023). However, for the first subsample (e.g., July 28, 2013 to June 17, 2018), the estimates for α are $\widehat{\alpha} = 4.2048$ for $|ret_{S\&P500}|$ and $\widehat{\alpha} = 3.2422$ for $|ret_{BTC}|$, respectively. Whereas Clauset et al.'s (2009) GoF test does not reject the power law model for $|ret_{S\&P500}|$ for the second subsample, it is rejected for Bitcoin. This finding is surprising because it indicates that Bitcoin does not exhibit power law behavior despite generating extreme events manifested in kurtosis values exceeding 3 by a substantial margin (see Table A.2).

Next, for the hypothesis testing, we first estimate the standard deviations for the power law exponents for both data vectors (e.g., $|ret_{S\&P500}|$ and $|ret_{BTC}|$) via block bootstrap. The distributional properties of the block bootstrapped power law exponents are reported in Table A.4, whereas the corresponding blocks bootstrapped standard deviations are reported in Table 1 denoted as $\widehat{\sigma}_{BOOT}$. Using block bootstrapped estimates for tail indices and standard deviations, we test the following null hypothesis:

$H_0 : (\alpha_{S\&P500} - \alpha_{BTC}) = 0 \quad \text{vs.} \quad H_1 : (\alpha_{S\&P500} - \alpha_{BTC}) \neq 0.$

The results are reported in Table 2. Strikingly, in Table 2, we observe that using two-sample z-tests and a significance level of 5%, the estimated power law exponents for $|ret_{S\&P500}|$ and $|ret_{BTC}|$ are not statistically different, regardless of which (sub)sample we consider because the corresponding z-test statistics are in absolute terms clearly below the critical value of |1.96|. This result indicates that the tail risks, measured in terms of the economic magnitude of power law exponents, are not statistically different from each other for $|ret_{S\&P500}|$ and $|ret_{BTC}|$ which is a novel finding.

We continue to explore whether the tail risks measured by the economic magnitude of the power law exponents were subject to change over time for a given asset. To do so, we test the following hypothesis for both assets:

$$H_0 : (\alpha_{1st} - \alpha_{2nd}) = 0 \quad \text{vs.} \quad H_1 : (\alpha_{1st} - \alpha_{2nd}) \neq 0.$$

**Table 2**
Two-sample z-test for power-law exponents.

| | S&P 500 | BTC | $(\alpha_{S\&P500} - \alpha_{BTC})$ (z-statistic) |
|---|---|---|---|
| 1. Sample | 4.2048 | 3.2422 | 0.9626 |
| (Std.Dev) | (1.6178) | (0.8635) | (0.7423) |
| 2. Sample | 3.1294 | 4.0921 | -0.9627 |
| (Std.Dev) | (0.8024) | (1.8760) | (-0.6673) |
| $(\alpha_{1st} - \alpha_{2nd})$ | 1.0754 | -0.8499 | |
| (z-statistic) | (0.8422) | (-0.5820) | |

Using standard deviations derived from block bootstraps, as detailed in Section 3.3, we test the following null hypothesis:
$H_0 : (\alpha_{S\&P500} - \alpha_{BTC}) = 0$ vs. $H_1 : (\alpha_{S\&P500} - \alpha_{BTC}) \neq 0$,
where $\widehat{\alpha}_{S\&P500}$ ($\widehat{\alpha}_{BTC}$) denotes the estimate for $\alpha$ for S&P 500 (Bitcoin) returns. Rejection of the null hypothesis would suggest that the data-generating processes of Bitcoin returns and S&P 500 returns are exposed to different tail risks. The corresponding two-sample z-test statistic is defined as:

$$z = \frac{(\alpha_{S\&P500} - \alpha_{BTC})}{\sqrt{\sigma^2_{\alpha_{S\&P500}} + \sigma^2_{\alpha_{BTC}})/2}},$$

where $\sigma^2_{\alpha_{S\&P500}}$ ($\sigma^2_{\alpha_{BTC}}$) denotes the variance of the tail exponent of the S&P 500 (Bitcoin), estimated via block bootstrap, as described in Section 3.3. To explore whether power law exponents have changed over time, we test the following null hypothesis for each asset:
$H_0 : (\alpha_{1st} - \alpha_{2nd}) = 0$ vs. $H_1 : (\alpha_{1st} - \alpha_{2nd}) \neq 0$,
where $\alpha_{1st}$ ($\alpha_{2nd}$) denotes the tail exponent for the first (second) subsample. Rejection of the null hypothesis would suggest that the tail risk for the respective asset has changed over time. The corresponding two-sample z-test statistic is defined as:

$$z = \frac{(\alpha_{1st} - \alpha_{2nd})}{\sqrt{\sigma^2_{\alpha_{1st}} + \sigma^2_{\alpha_{2nd}})/2}},$$

where $\sigma^2_{\alpha_{1st}}$ ($\sigma^2_{\alpha_{2nd}}$) denotes the variance for the first (second) subsample derived from block bootstraps for a given asset, and all the other notation is as previously. The period for the whole sample is from July 28, 2013, to May 14, 2023 corresponding to 512 weekly observations. The first subsample is from July 28, 2013, to June 17, 2018, whereas the second subsample is from June 24, 2018, to May 14, 2023. Both subsamples have equal length and comprise 256 weekly observations.

Again, using two-sample z-tests and a significance level of 5%, we observe in Table 2 that the estimated power law exponents for $|ret_{S\&P500}|$ and $|ret_{BTC}|$ are not statistically different across the samples because the corresponding z-test statistics are in absolute terms clearly below the critical value of $|1.96|$. This result indicates that the tail risks, measured in terms of the economic magnitude of power law exponents, are stable across the subsamples.

Next, we examine the economically important infinite variance hypothesis. To do so, we test the following hypothesis for both assets and all (sub)samples:

$$H_0 : (\alpha - 3) \leq 0 \qquad \text{vs.} \qquad H_1 : (\alpha - 3)\rangle 0$$

The results are reported in Table 3. Interestingly, we observe from Panel A of Table 3 that the estimated test statistics are clearly below the critical value of 1.645, indicating that using a 5% significance level, we cannot reject the hypothesis that the theoretical variances for $|ret_{S\&P500}|$ and $|ret_{BTC}|$ do not exist. This result holds for all (sub)samples and both assets. The p-values derived from block bootstraps, reported in Panel B of Table 3, strongly confirm this evidence because all p-values are well above 0.05, regardless of which sample is considered.

Because we cannot reject the infinite variance hypothesis for $|ret_{S\&P500}|$ and $|ret_{BTC}|$, we recall that Fama (1963, p. 421) commented that correlation-based methodologies "which are based on the assumption of finite variance will, at best, be considerably weakened and may in fact give very misleading answers." This is a serious issue for financial research. Next, to examine the dependency structure of S&P 500 and Bitcoin returns, we follow Mandelbrot (2008) and use R/S analysis to estimate the Hurst exponents via log-log regressions. Employing point estimates for Hurst exponents and standard deviation via log-log regressions for both assets and both subsamples, we test the following null hypothesis for both subsamples:

$$H_0 : (H_{S\&P500} - H_{BTC}) = 0 \qquad \text{vs.} \qquad H_1 : (H_{S\&P500} - H_{BTC}) \neq 0$$

The results are reported in Table 4. Using two-sample z-tests and a significance level of 5%, we observe from Table 4 that the null hypothesis is rejected for both subsamples because the z-test statistics clearly exceed in absolute terms the critical value of $|1.96|$. This result implies that the memory of Bitcoin returns exhibits a higher level of persistence than the memory of S&P 500 returns. Moreover, the higher level of persistence does not vanish over time, because this result holds across both subsamples.

We then investigate whether the memory process for each asset was subject to change over time. To explore this issue, we test the following null hypothesis for each asset:

$$H_0 : (H_{1st} - H_{2nd}) = 0 \qquad \text{vs.} \qquad H_1 : (H_{1st} - H_{2nd}) \neq 0$$

The results are reported in Table 4. Using two-sample z-tests and a significance level of 5%, we observe from Table 4 that we can reject the null hypothesis only for the Hurst exponents of S&P 500 returns because the z-test statistic for testing the Hurst exponents of Bitcoin returns does not exceed in absolute terms the critical value of $|1.96|$. Note also from Table 4 that the estimated Hurst exponent for S&P 500 returns is 0.5424 in the first subsample with an estimated standard deviation of 0.0302. This implies that the Hurst exponent for S&P 500 returns was not statistically different from 0.50 in the first subsample (e.g., t-statistic 1.4040). A statistically significant increase in the economic magnitude of the Hurst exponent—as we move from the first to the second subsample—implies that S&P 500 returns exhibit a higher level of persistence in the second subsample. This evidence suggests that the S&P 500 moved from market efficiency to market inefficiency, which is a surprising finding.

### *4.2. Are the Results Robust?*

Inspired by Taleb's (2020) notion concerning one-sigma events, we proposed a one-sigma test that is constructed using the normal distribution as the null model. Simulating 100,000 time series vectors with a dimension of 500× 1 and consisting of random drawings from the standard normal, we computed the distribution of $\lambda_0$, where $\lambda_0 = \int_{-1}^{1} p(y_{2t})$, with $y_{2t} = (y_{1t} - \mu_{y_1})/\sigma_{y_1}$ and $p(y_{2t})$ is the corresponding probability function and $y_{2t} \sim N(\mu_{y_1}, \sigma_{y_1})$. The distributional properties of the one-sigma statistic $\lambda_0$ for the normal distribution are reported in Table A.5. Using a one-sided test and a significance level of 5%, we observe from Table A.5 that we would reject the normal distribution if the fraction of observations that is within one standard deviation from the mean exceeds the critical value of 0.7160. Recall that the exact value, as $T \to \infty$, should be $\lambda_0 = 0.6827$. Testing the hypothesis:

$H_0 : \lambda \sim$thin-tailed distribution vs. $H_1 : \lambda \sim$fat-tailed distribution,

we find that $\hat{\lambda} = 0.8004$ for S&P 500 returns and $\hat{\lambda} = 0.8240$ for Bitcoin returns. Specifically, for S&P 500 we observe that 402 out of 512 observations are within one standard deviation corresponding to 80.04%, whereas for Bitcoin we observe that 412 out of 512 observations are within one standard deviation corresponding to 82.40%. Thus, both asset markets exhibit statistically significant fat tails which supports the main analysis and our argument to use (a) power laws to model the dynamics of the return distributions and to use (b) R/S analysis to model the dependency structures of the returns.

Another way to implement the one-sigma test is to use a version of Pearson's chi-square test, defined as,

$$\lambda = \frac{(\#_{x \leq 1\sigma} - 0.6827T)^2}{0.6827T} + \frac{(\#_{x > 1\sigma} - (T - 0.6827T))^2}{(T - 0.6827T)},$$

where $_{x \leq 1\sigma}$ denotes the number of observations some distribution that are within one standard-deviation from the mean, $_{x > 1\sigma}$ denotes the remaining observations, 0.6827 is the expected frequency under the assumption of normality, and $T$ is the sample size. In an early study, Pearson (1900) showed that this type of test statistic is distributed as $\chi^2(1)$ as $T \to \infty$. As $_{x \leq 1\sigma}$ ($_{x > 1\sigma}$) corresponds to 402 (110)

**Table 3**
Testing the infinite variance hypothesis.

| Panel A. Testing the infinite variance hypothesis using standard *t*-tests | | | |
|---|---|---|---|
| | Whole sample | First subsample | Second subsample |
| S&P 500 | 0.5871 | 0.7447 | 1.4075 |
| BTC | 0.7599 | 1.4386 | 0.5821 |
| **Panel B. Testing the infinite variance hypothesis using bootstrapped *p*-values** | | | |
| | Whole sample | First subsample | Second subsample |
| S&P 500 | 0.1840 | 0.3710 | 0.5050 |
| BTC | 0.2070 | 0.4930 | 0.3860 |

We test for all (sub)samples the infinite variance hypothesis using the following hypothesis pair:
$H_0 : (\alpha - 3) \leq 0$ vs. $H_1 : (\alpha - 3) > 0$.
The statistical significance is assessed via standard *t*-tests and *p*-values derived from block bootstraps, as detailed in Section 3.3. In evaluating statistical significances, we use a one-sided test and a standard significance level of 5% corresponding to a critical value of 1.645. Panel A reports the *t*-test statistics, whereas Panel B reports the *p*-values derived from block bootstraps. The period for the whole sample is from July 28, 2013, to May 14, 2023 corresponding to 512 weekly observations. The first subsample is from July 28, 2013, to June 17, 2018, whereas the second subsample is from June 24, 2018, to May 14, 2023. Both subsamples have equal length and comprise 256 weekly observations.

**Table 4**
Testing the memory processes using rescaled/range analysis.

| | S&P 500 | BTC | $(H_{S\&P500} - H_{BTC})$ (*z*-statistic) |
|---|---|---|---|
| 1. Sample | 0.5424 | 0.6903 | -0.1479*** |
| (Std.Dev) | (0.0302) | (0.0221) | (-5.5892) |
| 2. Sample | 0.6219 | 0.6641 | -0.0422** |
| (Std.Dev) | (0.0175) | (0.0175) | (2.4137) |
| $(H_{1st} - H_{2nd})$ | -0.0795*** | 0.0262 | |
| (*z*-statistic) | (-3.2211) | (1.3124) | |

We use R/S analysis to derive Hurst exponents for S&P 500 and Bitcoin returns for both subsamples separately. According to Mandelbrot (2008, p.298-299), the R/S statistic is summarized as:

$$R/S_k = MAX_{1 \leq k \leq n} \frac{\sum_{j=1}^{k} (r_j - \bar{r}_n) - MIN_{1 \leq k \leq n} \sum_{j=1}^{k} (r_j - \bar{r}_n)}{\left[\frac{1}{n} \sum_j (r_j - \bar{r}_n)^2\right]^{1/2}},$$

where $\bar{r}_n$ is the average return over $n$ days. Then, for each shorter time period, the difference between the return $r_j$ over that period and the average return, $\bar{r}_n$ is calculated, and a running total of all the differences as the time-periods lengthen up to a period $k$, where $k \in \{4, 8, 16, 32, 64, 128, 256\}$. Then the maximum and the minimum of all those differences is computed. Subtracting the minimum from the maximum gives us an estimate of the range from peak to trough in the accumulated deviations which is the numerator in the equation above, whereas the denominator is the conventional measure of the standard deviation in the data series. The estimated value of the Hurst exponent $H$ can be retrieved via the following log-log regression model:
$\ln(R/S)_k = \ln(C) + H \ln(k) + u$,
where $u$ is assumed to be distributed as $u \sim iid(0, \sigma_u)$. The period for the whole sample is from July 28, 2013, to May 14, 2023 corresponding to 512 weekly observations. The first subsample is from July 28, 2013, to June 17, 2018, whereas the second subsample is from June 24, 2018, to May 14, 2023. Both subsamples have equal length and comprise 256 weekly observations.

observations for the S&P 500, we estimate $\hat{\lambda} = 24.8112$. As $24.8112 >> 3.8410 = \chi^2_{0.95}(1)$, the null hypothesis of normality is clearly rejected on a 5% level. The corresponding test statistic for Bitcoin is $\hat{\lambda} = 35.1723$, and thus, the statistical conclusion is the same as for the S&P 500. Overall, the conclusions remain unchanged.

Next, we reexamine the memory features using DFA instead of R/S analysis, which has become a popular tool for estimating Hurst exponents. Using the DFA framework, we employ log-log regressions for both assets and both subsamples and, again, test the following null hypothesis for both subsamples:

$$H_0 : (H_{S\&P500} - H_{BTC}) = 0 \quad \text{vs.} \quad H_1 : (H_{S\&P500} - H_{BTC}) \neq 0$$

The results are reported in Table A.6. Using two-sample *z*-tests and a significance level of 5%, we observe from Table 6 that the null hypothesis is rejected for both subsamples because the *z*-test statistics clearly exceeded in absolute terms the critical value of $|1.96|$. This result implies that the memory of Bitcoin returns exhibits a higher level of persistence than the memory of S&P 500 returns. This result strongly confirmed our previous findings derived from R/S analysis. Moreover, we test the hypothesis pair,

$H_0 : (H_{1st} - H_{2nd}) = 0$ vs. $H_1 : (H_{1st} - H_{2nd}) \neq 0$.

The results shown in Table 6 indicate that DFA and R/S analysis are qualitatively the same for the memory features of S&P 500 returns, in the sense that the level of persistence of S&P 500 returns increased statistically significantly over time. Unlike the results derived from R/S analysis, however, the DFA results suggest that Bitcoin returns exhibited a lower level of persistence in the second subsample. Overall, the commonality we find is that DFA and R/S analysis suggest that the market for Bitcoin is more inefficient than the S&P 500, which is manifested in Hurst exponents exhibiting a considerably higher economic magnitude for Bitcoin returns, regardless of the subsample.

Finally, we test whether kurtosis is defined for $|ret_{S\&P500}|$ and $|ret_{BTC}|$ as well as for all (sub)samples . Specifically, we test the following hypothesis:

$H_0 : (5 - \alpha) \leq 0$ vs. $H_1 : (5 - \alpha)\rangle 0$

The results are reported in Table A.7. Interestingly, considering $|ret_{S\&P500}|$, we observe from Panel A of Table A.7 that the estimated test statistics are clearly above the critical value of 1.645 for the whole sample and the second subsample. The estimated *p*-values reported in Panel B of Table A.7 strongly confirm this result, as they are below $p = 0.05$. This finding suggests that even if we assume that population variance exists, we are not allowed to operate with correlation-based methods because, in finite samples, we cannot expect to observe the true value of the variance (Taleb, 2020).

However, considering $|ret_{BTC}|$, the results reported in panels A and B of Table A.7 suggest that the fourth moment of $|ret_{BTC}|$ exists because the null hypothesis cannot be rejected for the whole sample and the second subsample. Recall that using Clauset et al.`s (2009) GoF test, the power law hypothesis is rejected for the first subsample; thus, rejecting the null hypothesis cannot be interpreted in a reasonable manner. This finding is interesting because it suggests that if we assumed that the population variance exists, as indicated by a point estimate satisfying $\hat{\alpha} > 3$, we would be allowed to operate with correlation-based methods for evaluating Bitcoin returns, as we would expect to observe the true value of the variance in large samples.

In contrast, regarding $|ret_{S\&P500}|$, Taleb (2020) posited that scholars cannot argue with "Gaussian behavior" if kurtosis is infinite, even when lower moments exist and even for a power law exponent corresponding to $\alpha \approx 3$. Because the central limit operates very slowly, $n$ of the order of $10^6$ is required to become acceptable. As pointed out by Taleb (2020), the problem is that we do not have very many data in the history of financial markets; therefore, we are not in a research environment that allows us to use Gaussian methodologies.

## 5. Discussion

### 5.1. Limitations

One may wonder whether Bitcoin as a single asset is comparable with the S&P 500, which is a stock index consisting of 500 large cap companies. In this regard, it is important to note that the cryptocurrency market is highly concentrated and highly correlated.[11] Specifically, when this research was conducted, more 20,000 cryptocurrencies were traded at more than 600 exchanges. Bitcoin, as a single asset corresponding to only $\approx 0.0005\%$ of the total number of cryptocurrencies, comprises $\approx 50\%$ of the overall market capitalization. For the traditional Pareto 80/20 distribution, 1% of the (cumulative) total of largest observations comprises $\approx 50\%$ of the cumulative total. This means that the market capitalization in the cryptocurrency market is so concentrated that the theoretical mean does not exist.[12] This means that the cryptocurrency market exhibits an extremely high level of concentration. This study adds to the recent literature exploring the S&P 500 and Bitcoin. These studies include, among others, Conlon and McGee (2020), who investigated the safe haven properties of Bitcoin during the Covid-19 bear market. The authors found that Bitcoin is not a safe haven because the price for Bitcoin decreased in price in lockstep with the S&P 500 during the bear market. Moreover, Pal and Mitra (2019) explored the possibility of hedging Bitcoin prices with the S&P 500. The authors' findings suggest that Bitcoin can be hedged with the S&P 500. Another recent study by Klein et al. (2018) compared Bitcoin with traditional assets, including the S&P 500, and focused on volatility, correlation, and portfolio diversification. The authors' findings indicated that Bitcoin returns have an asymmetric response to market shocks and Bitcoin and do not resemble any other conventional asset from an econometric perspective. Although these studies drew conclusions from standard econometric models, the present research took a fractal view in comparing the memory and potential changes in the memory of Bitcoin and the S&P 500. We argue that in the presence of fat tails, dependency structures should be investigated via R/S analysis as it does not require any distributional assumptions. However, future research should compare this issue with a crypto-index instead of Bitcoin.[13]

[11] For instance, the study of Borri (2019) documents that cryptocurrencies are highly correlated assets.

[12] Since the power-law exponent of the traditional Pareto 80/20 is $\alpha \approx 2.0$, it follows that for the market capitalization in the cryptocurrency market $\alpha << 2.0$. From Eq. (2) we see that the theoretical mean is undefined for a distribution exhibiting such an extraordinary high level of concentration.

[13] Note that today, Ethereum is the second largest cryptocurrency in terms of market capitalization. That is, a cryptocurrency market index consisting solely of Bitcoin and Ethereum would correspond to 70% of the overall cryptocurrency market capitalization. As Ethereum did not exist in 2013, which is the starting point for the present research sample, such an index could not be computed. Furthermore, by the end of 2013, about 50 different cryptocurrencies were traded, whereas the market dominance of Bitcoin exceeded 99% according to tradingview.com, which implies that a value-weighted basket of cryptocurrencies would be virtually solely determined by Bitcoin price changes.

Next, our findings indicate that the fourth moment of $|ret_{BTC}|$ exists. However, this finding could be subject to sample specificity. Note that Taleb (2010) stated:

> "I have learned a few tricks from experience: whichever exponent I try to measure will likely be overestimated (recall that a higher exponent implies a smaller role for large deviations) – what you see is likely to be less Black Swannish than what you don't see. I call this the masquerade problem. Let's say I generate a process that has an exponent of 1.7. You do not see what is inside the engine, only the data coming out. If I ask you what the exponent is, odds are that you will compute something like 2.4. You would do so even if you had a million data points. The reason is that it takes a long time for some fractal processes to reveal their properties, and you underestimate the severity of the shock." (Taleb, 2010, p. 266).

Specifically, our finding that the fourth moment of $|ret_{BTC}|$ is defined could be an artefact of parameter *under*estimation. This means that although we employed block bootstraps, it is possible that we underestimated the power law exponent for $|ret_{BTC}|$ simply because we lacked data. Future research is needed to elaborate on this issue.

Finally, fractality also implies that power law behavior does not change over different time scales, as pointed out by Mandelbrot (2008). Research is still needed to explore the economic magnitudes of the power law exponents for $|ret_{BTC}|$ or $|ret_{S\&P500}|$ on various time scales. However, this is beyond the scope of this paper and is therefore left for future research.

### 5.2. Implications

An early and often-cited study by Urquhart (2016) argued that Bitcoin is in the process of moving toward an efficient market. Although the view that Bitcoin exhibits market efficiency is supported by numerous studies (e.g., Nadarajah and Chu, 2017; Bariviera, 2017; Sensoy, 2019), we did not find such evidence. The results we derived from the R/S analysis show that Bitcoin returns exhibit a level of persistence that is not in line with an efficient market. The R/S analysis showed, moreover, that the level of persistence did not chang over time. Therefore, the results of this study are in line with the literature arguing that the Bitcoin market is inefficient (e.g., Cheah et. al, 2018; Bouri et al., 2017; Zargar and Kumar, 2019; Tiwari et al., 2018; Khuntia and Pattanayak, 2018). What are the practical implications of return persistence? As observed by Mandelbrot (2008), Edgar E. Peters, chief investment officer of PanAgora Asset Management, found that high-tech stocks had high dependence manifested in higher $H$ values than other stocks, which makes them a better bet for investors due to more readily perceived price trends. Similarly, this study showed that Bitcoin returns exhibit high dependence, manifested in high $H$ values across subsamples. The continued price persistence of Bitcoin could be one reason Bitcoin attracts speculators. Future research is encouraged to explore this issue in more detail and for other cryptocurrencies.

What are the implications of power law behavior from a risk management perspective? In a recent study, Grobys (2023) derived the concept of cofractality, which measures the codependence between the parts of the distributions governed by a power law process. As correlation is not defined for processes for which the variance is not defined, cofractality might be used to diversify portfolios using S&P 500 or Bitcoin returns. As this topic is beyond the scope of this study, future research is needed to explore this issue in more detail.

Finally, contrary to other studies that raised doubts about the validity of the infinite variance hypothesis, as shown by Lux and Alfarano (2016), we confirmed Mandelbrot's (1963) early study on cotton price changes by finding that we cannot reject the infinite variance hypothesis for $|ret_{S\&P500}|$ or $|ret_{BTC}|$. This result held for thewhole sample and the subsamples. An important implication, therefore, is that conclusions derived from standard statistical models based on the assumption of finite variances should be treated with caution as they potentially result in misleading answers (Fama, 1963). Due to the high level of replication failures in finance studies (Hou, Xue, and Zhang, 2020), this is certainly an important issue that future research needs to take into account.[14]

## 6. Conclusion

Although the majority of previous studies relied on autocorrelation-based metrics to explore dependency structures for Bitcoin, this study took a fractal view on this issue. We compared the long-term memory of Bitcoin and the S&P 500, which was used in the present research as an archetype example of an efficient asset market. In doing so, we followed Mandelbrot by assuming that the asset markets are subject to fractality. Evidence of fractality renders statistical inference based on traditional methodologies derived from Gaussian models invalid, especially if the economic magnitude of the power law exponent is low.

In fact, testing whether Bitcoin returns or S&P returns are governed by power laws provided strong evidence for both markets exhibit strong power law behavior. Testing the infinite variance hypothesis (Mandelbrot, 1963), the evidence suggested that theoretical variances for $|ret_{S\&P500}|$ or $|ret_{BTC}|$ are undefined. As a result, the reported research results in the previous literature derived from correlation-based methodologies were inevitably sample-specific. As pointed out, this is a serious issue, ad noted by Fama (1963):

> "… the infinite variance assumption of the stable Paretian model has extreme implications. From a purely statistical standpoint, if the population variance of the distribution of first differences is infinite, the sample variance is probably a meaningless measure of dispersion. Moreover, if the variance is infinite, other statistical tools (e.g., least-squares regression) which are based on the assumption of finite variance will, at best, be considerably weakened and may in fact give very misleading answers." (Fama, 1963, p. 421)

[14] Note that a recent study of Grobys (2021) documents that replication failure of finance studies could be a manifestation of using correlation-based methods in research environments where the variance is sample-specific.

Due to the evidence of fractality, we made use of the R/S statistic because Mandelbrot (2008) pointed out that one of the R/S statistic's principal virtues is that, in contrast to many common statistical tests, it does not require any assumption on how the original data are organized, which is a critical point when examining financial assets.

Hurst exponents estimated via applications of Mandelbrot (2008) R/S analysis showed that (a) Bitcoin returns exhibit a significantly higher level of persistence than S&P 500 returns and that (b) the level of persistence, measured in terms of the economic magnitude of the estimated Hurst exponents, has not changed over time. Although the former result was strongly supported by applying DFA, DFA showed that the level of persistence in Bitcoin returns decreased slightly over time. However, the results derived from both statistical methodologies (e.g., R/S analysis and DFA) suggest that S&P 500 returns exhibited a statistically significantly higher level of persistence in the later subsample than in the earlier subsample. This finding is counterintuitive and interesting because it suggests that the efficiency of the S&P 500 has decreased in the last decade. Future research is encouraged to elaborate more on this issue.

## Author statement

I conducted the research on my own.

## Data Availability

Data will be made available on request.

## Appendix

**Table A1**
Descriptive statistics for return data covering the sample period July 28, 2013, until May 14, 2023.

| | S&P 500 | BTC | \|S&P500\| | \|*BTC*\| |
|---|---|---|---|---|
| Mean | 0.2046 | 1.9119 | 1.6121 | 8.5741 |
| Median | 0.3400 | 0.8000 | 1.2100 | 5.5150 |
| Maximum | 12.1000 | 118.0800 | 14.9800 | 118.0800 |
| Minimum | -14.9800 | -42.8000 | 0.0000 | 0.0000 |
| Std. Dev. | 2.3097 | 13.3068 | 1.6651 | 10.3477 |
| Skewness | -0.6524 | 1.9594 | 3.0019 | 4.2132 |
| Kurtosis | 9.9094 | 18.1401 | 17.8494 | 34.4536 |
| Jarque-Bera | 1054.7770 | 5217.7060 | 5473.0730 | 22,620.5000 |
| Probability | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Observations | 512 | 512 | 512 | 512 |

**Table A2**
Descriptive statistics for return data covering the sample period July 28, 2013, until June 17, 2018.

| | S&P 500 | BTC | \|S&P500\| | \|*BTC*\| |
|---|---|---|---|---|
| Mean | 0.2040 | 2.7384 | 1.2139 | 9.9499 |
| Median | 0.2700 | 1.3700 | 0.9150 | 6.0900 |
| Maximum | 4.3000 | 118.0800 | 5.9600 | 118.0800 |
| Minimum | -5.9600 | -42.8000 | 0.0000 | 0.0000 |
| Std. Dev. | 1.6322 | 15.9312 | 1.1074 | 12.7257 |
| Skewness | -0.7404 | 2.2383 | 1.6152 | 4.0444 |
| Kurtosis | 4.9469 | 16.4727 | 6.2488 | 27.9549 |
| Jarque-Bera | 63.8156 | 2149.8990 | 223.8920 | 7340.5300 |
| Probability | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Observations | 256 | 256 | 256 | 256 |

Weekly data on Bitcoin and the S&P 500 covering the period July 28, 2013, until May 14, 2023 were downloaded from investing.com. The whole data sample consists of 512 weekly observations. The overall sample is split into two subsamples of equal length: The first subsample is from July 28, 2013, until June 17, 2018, whereas the second subsample is from June 24, 2018, until May 14, 2023. For both data series we compute the absolute returns. The descriptive statistics for return data and absolute return data are reported for the first subsample in this table.

**Table A3**
Descriptive statistics for return data covering the sample June 24, 2018, until May 14, 2023.

| | S&P 500 | BTC | \|S&P500\| | \|*BTC*\| |
|---|---|---|---|---|
| Mean | 0.2052 | 1.0855 | 2.0103 | 7.1983 |
| Median | 0.5100 | 0.7000 | 1.5150 | 5.2950 |
| Maximum | 12.1000 | 31.5000 | 14.9800 | 41.6900 |

(*continued on next page*)

**Table A3** (*continued*)

| | S&P 500 | BTC | \|S&P500\| | \|BTC\| |
|---|---|---|---|---|
| Minimum | -14.9800 | -41.6900 | 0.0200 | 0.0400 |
| Std. Dev. | 2.8331 | 9.9830 | 2.0028 | 6.9874 |
| Skewness | -0.5679 | -0.1733 | 2.7723 | 1.6448 |
| Kurtosis | 8.2449 | 4.9211 | 14.2322 | 6.1678 |
| Jarque-Bera | 307.1904 | 40.6475 | 1673.6570 | 222.4619 |
| Probability | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Observations | 256 | 256 | 256 | 256 |

Weekly data on Bitcoin and the S&P 500 covering the period July 28, 2013, until May 14, 2023 were downloaded from investing.com. The whole data sample consists of 512 weekly observations. The overall sample is split into two subsamples of equal length: The first subsample is from July 28, 2013, until June 17, 2018, whereas the second subsample is from June 24, 2018, until May 14, 2023. For both data we compute the absolute returns. The descriptive statistics for return data and absolute return data are reported for the second subsample in this table.

**Table A4**
Distributional properties of blocks bootstrapped power-law exponents.

| | S&P 500 whole sample | S&P 500 first sub-sample | S&P 500 s subsample | BTC whole sample | BTC first sub-sample | BTC second subsample |
|---|---|---|---|---|---|---|
| Min | 2.3722 | 2.3177 | 2.1954 | 2.1344 | 1.8380 | 2.0224 |
| < 2.50% | 2.5247 | 2.5328 | 2.4217 | 2.5708 | 2.1787 | 2.2039 |
| < 5.00% | 2.6322 | 2.6132 | 2.4786 | 2.6360 | 2.2626 | 2.2529 |
| Median | 3.3807 | 3.2694 | 2.9938 | 3.4539 | 3.0212 | 3.3047 |
| > 5% | 4.4137 | 7.1538 | 3.9772 | 6.0317 | 4.7097 | 7.6547 |
| > 97.50% | 4.7534 | 7.7971 | 4.6103 | 6.6318 | 5.7383 | 8.2599 |
| Kurtosis | 75.8183 | -0.2686 | 46.3695 | 2.2960 | 2.7957 | 0.4052 |
| Skewness | 6.6249 | 0.8680 | 5.7267 | 1.4313 | 1.4757 | 1.0354 |

Weekly data on Bitcoin and the S&P 500 covering the period July 28, 2013, until May 14, 2023 were downloaded from investing.com. The whole data sample consists of 512 weekly observations. The overall sample is split into two subsamples of equal length: The first subsample is from July 28, 2013, until June 17, 2018, whereas the second subsample is from June 24, 2018, until May 14, 2023. For each data series (e.g., whole same and subsamples), we implement a block bootstrap procedure, as proposed by Grobys and Junttila (2021). Denoting the selected block length as $m$, we implement a blocks bootstrap procedure such that $E[m] = T^{1/2}$. Then from each data vector $\boldsymbol{x}$, we randomly draw blocks $m_j$ which are distributed as a geometric distribution $m_j \sim GEO(p)$ with $E[m_j] = \frac{(1-p)}{p}$. For instance, for the overall sample, we employ $E[m_j] = 23$, implying $p = 0.0417$, whereas $E[m_j] = 16$ implying $p = 0.0588$ for subsample blocks bootstraps. Using this procedure, the blocks drawn from data vectors $\boldsymbol{x}$ vary in lengths. Randomly drawn blocks $m_j$ from data vector $\boldsymbol{x}$ are stacked in vector $\boldsymbol{x}^B$ as:

$$\boldsymbol{x}_i^b = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ \vdots \end{bmatrix}$$

The procedure is stopped when the length of the artificial vector $\boldsymbol{x}^b$ exhibits a length exceeding $T$. Observations exceeding $T$ are cut off; that is, every artificial data vector $\boldsymbol{x}^b$ has the same length as the original data vector $\boldsymbol{x}$. Using this block bootstrap procedure, for each data vector, $B = 1,000$ artificial data vectors are constructed,

$$[\boldsymbol{x}^1 \quad \boldsymbol{x}^2 \quad \dots \quad \boldsymbol{x}^B],$$

and point estimates for α are obtained for each bootstrapped data vector $\boldsymbol{x}^1, \boldsymbol{x}^2, \dots, \quad \boldsymbol{x}^B$ using Clauset et al.'s (2009) approach so that,

$$[\widehat{\alpha}^1 \quad \widehat{\alpha}^2 \quad \dots \quad \widehat{\alpha}^B]$$

Finally, the corresponding bootstrapped standard error $\widehat{\sigma}_{BOOT}$ is computed for the bootstrapped $\widehat{\alpha}^1, \widehat{\alpha}^2, \dots, \quad \widehat{\alpha}^B$ for each given data vector. This table reports the descriptive statistics of block bootstrapped estimated power law exponents for each data sample.

**Table A5**
Distributional properties of the one-sigma statistic for the normal distribution.

| | |
|---|---|
| Min | 0.5980 |
| < 1% | 0.6340 |
| < 2.5% | 0.6420 |
| < 5.0% | 0.6480 |
| Median | 0.6820 |
| > 95% | 0.7160 |
| > 97.5% | 0.7240 |
| > 99% | 0.7300 |
| Max | 0.7760 |
| Mean | 0.6827 |
| Std.Dev | 0.0208 |

(*continued on next page*)

**Table A5** (*continued*)

| | |
|---|---|
| Kurtosis | -0.0047 |
| Skewness | -0.0314 |

Denoting $y_{2t}$ as a drawing from some normal distribution and defining $y_{2t} = (y_{1t} - \bar{y})/\sigma$ where $p(y_{1t})$ is the corresponding probability function and $y_{1t} \sim N(\mu_{y_1}, \sigma_{y_1})$, it follows that $y_{2t} \sim N(0,1)$. Then, $\int_{-1}^{1} p(y_{2t}) = 0.6827$as $T \longrightarrow \infty$. In finite samples $\bar{y}$ and $\sigma$ are estimated from a given data vector and the fraction $\int_{-1}^{1} p(y_{2t})$ will exhibit variation depending on the sample size $T$. Denoting the reference statistic for evaluating the fraction of observations that are within one standard- deviation of some distribution as $\lambda_0 = \int_{-1}^{1} p(y_{2t})$, we employ a simulation experiment where we simulate 100,000 time series vectors are drawn from a standard normal distribution, whereas each time series vector has the dimension of 500×1. For each simulated vector, we calculate $\lambda_0$ and determine the corresponding 95% probability interval. This table reports the distributional properties of the one-sigma statistic for the normal distribution.

**Table A6**
Testing the memory processes using detrended fluctuation analysis.

| | S&P 500 | BTC | $(H_{S\&P500} - H_{BTC})$($z$-statistic) |
|---|---|---|---|
| 1. Sample | 0.4160 | 0.7074 | -0.2914*** |
| (Std.Dev) | (0.0304) | (0.0212) | (-11.1192) |
| 2. Sample | 0.5150 | 0.6062 | -0.0912*** |
| (Std.Dev) | (0.0216) | (0.0327) | (-3.2911) |
| $(H_{1st} - H_{2nd})$ | -0.0990*** | 0.1012*** | |
| ($z$-statistic) | (-3.7543) | (3.6724) | |

We use DFA to derive the Hurst exponents. First, data series $x_t$ is converted to the mean-centered cumulative sum:
$\widetilde{x}_t = \sum_{t=1}^{T} x_t$.
Analogous to R/S analysis, different time scales $k$ are defined; that is, $k \in \{4, 8, 16, 32, 64, 128, 256\}$. Depending on the defined time scale, data is split into epochs and for each epoch $s$, a time series regression is used to detrend the data. For instance, employing $k = 256$ means that weekly data for $\widetilde{x}_t$ is split into two nonoverlapping epochs. For each epoch $s$, the following regression is employed:
$\widetilde{x}_t = \gamma_0 + \gamma_1 t + e_t$,
where $t = 1, \ldots, 128$ for the first epoch and $t = 129, \ldots, 256$ for the second epoch. Then, for each respective epoch $s$, the root mean squared error (RMSE) is computed as:
$RMSE_S = \sqrt{\frac{1}{T_s} \sum_t^{T_s} \widehat{e}_t}$,
where $T_s = 256$. Finally, the estimates for $RMSE_S$ are averaged for each time scale $k$, giving us
$\overline{RMSE}_k$. According to theory, the following relation holds:
$\overline{RMSE}_k = ck^H$.
The Hurst exponent is then estimated by computing a linear fit between log-scales and log-$\overline{RMSE}_k$. The period for the whole sample is from July 28, 2013, to May 14, 2023 corresponding to 512 weekly observations. The first subsample is from July 28, 2013, to June 17, 2018, whereas the second subsample is from June 24, 2018, to May 14, 2023. Both subsamples have equal length and comprise 256 weekly observations.

**Table A7**
Testing the finite kurtosis hypothesis.

| Panel A. Testing the finite-kurtosis-hypothesis using standard $t$-tests |
|---|

(*continued on next page*)

**Table A7** (*continued*)

| | Whole sample | First subsample | Second subsample |
|---|---|---|---|
| S&P 500 | 1.8675 * * | 0.4915 | 2.3313 * * |
| BTC | 1.0978 | 2.0357 * * | 0.4840 |
| Panel B. Testing the finite-kurtosis-hypothesis using bootstrapped *p*-values | | | |
| | Whole sample | First subsample | Second subsample |
| S&P 500 | 0.0190 | 0.3360 | 0.0170 |
| BTC | 0.1300 | 0.0410 | 0.3220 |

We test for all (sub)samples whether the kurtosis exists using the following hypothesis pair:
$H_0 : (5 - \alpha) \leq 0$ vs. $H_1 : (5 - \alpha) > 0$.
The statistical significance is assessed via standard *t*-tests and *p*-values derived from block bootstraps, as detailed in Section 3.3. In evaluating statistical significances, we use a one-sided test and a standard significance level of 5% with corresponding critical value of 1.6450. Panel A reports the *t*-test statistics, whereas Panel B reports the *p*-values derived from block bootstraps. The period for the whole sample is from July 28, 2013, to May 14, 2023 corresponding to 512 weekly observations. The first subsample is from July 28, 2013, to June 17, 2018, whereas the second subsample is from June 24, 2018, to May 14, 2023. Both subsamples have equal length and comprise 256 weekly observations.

## References


Bariviera, A.F., 2017. The inefficiency of Bitcoin revisited: a dynamic approach. Econ. Lett. 161, 1–4.
Bassingthwaighte, J.B., Raymond, G.M., 1994. Evaluating rescaled range analysis for time series. Ann. Biomed. Eng. 22, 432–444.
Borri, N., 2019. Conditional tail-risk in cryptocurrency markets. J. Empir. Financ. 50, 1–19.
Bouri E., Azzi G., Dyhrberg A.H. , 2017. On the return-volatility relationship in the Bitcoin market around the price crash of 2013, Economics Discussion Papers, No 2016–41, Kiel Institute for the World Economy.
Bryce, R.M., Sprague, K.B., 2012. Revisiting detrended fluctuation analysis. Sci. Rep. 2, 315.
Cheah, E.-T., Mishra, T., Parhi, M., Zhuang, Zhang, 2018. Long Memory Interdependency and Inefficiency in Bitcoin markets. Econ. Lett. 167, 18–25.
Clauset, A., Shalizi, C.R., Newman, M.E.J., 2009. Power law distributions in empirical data. SIAM Rev. 51, 661–703.
Conlon, T., McGee, R., 2020. Safe haven or risky hazard? Bitcoin during the COVID-19 bear market. Financ. Res. Lett. 35, 101607.
Da Silva, S., Matsushita, R., Gleria, I., Figueiredo, A., 2007. Hurst exponents, power laws, and efficiency in the Brazilian foreign exchange market. Econ. Bull. 7, 1–11.
Eom, C., Choi, S., Oh, G., Jung, W.-S., 2008. Hurst exponent and prediction based on weak-form efficient market hypothesis of stock markets. Phys. A: Stat. Mech. its Appl. 387, 4630–4636.
Fama, E.F., 1963. Mandelbrot and the stable Paretian hypothesis. J. Bus. 36, 420–429.
Fry, J., Cheah, E.T., 2016. Negative bubbles and shocks in Cryptocurrency markets. Int. Rev. Financ. Anal. 47, 343–352.
Gabaix, X., Parameswaran, G., Vasiliki, P., Eugene, S.H., 2003. A theory of power-law distributions in financial market fluctuations. Nature 423, 267–270.
Gabaix, X., Parameswaran, G., Vasiliki, P., Eugene, S.H., 2006. Institutional investors and stock market volatility. Q. J. Econ. 121, 461–504.
Grobys, K., 2021. What do we know about the second moment of financial markets? Int. Rev. Financ. Anal. 78, 101891.
Grobys, K., 2023. Correlation versus co-fractality: evidence from foreign-exchange-rate variances. Int. Rev. Financ. Anal. 86, 102531.
Grobys, K., Junttila, J.-P., 2021. Speculation and lottery-like demand in cryptocurrency markets. J. Int. Financ. Mark., Inst. Money 71, 101289.
Grobys, K., Kolari, J.W., Junttila, J.-P., Sapkota, N., 2021. On the stability of stablecoins. J. Empir. Financ. 64, 207–223.
Grobys, K., Han, Y., Kolari, J.W., 2023. Better not forget: on the memory of S&P 500 survivor stock companies. J. Risk Financ. Manag. 16 (2), 126.
Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. Rev. Financ. Stud. 33 (5), 2019–2133.
Khuntia, S., Pattanayak, J.K., 2018. Adaptive market hypothesis and evolving predictability of bitcoin. Econ. Lett. 167, 26–28.
Klein, T., Thu, H.P., Walther, T., 2018. Bitcoin is not the new gold–a comparison of volatility, correlation, and portfolio performance. Int. Rev. Financ. Anal. 59, 105–116.
Kristoufek, L., 2018. On Bitcoin markets (in)efficiency and its evolution. Phys. A: Stat. Mech. its Appl. 503, 257–262.
Lahmiri, S., 2016. Clustering of Casablanca stock market based on hurst exponent estimates. Phys. A: Stat. Mech. its Appl. 456, 310–318.
Lux, T., Alfarano, S., 2016. Financial power laws: empirical evidence, models, and mechanisms. Chaos, Solitons Fractals 88, 3–18.
Mandelbrot, B.B., 1963. The variation of certain speculative prices. J. Bus. 36, 394–419.
Mandelbrot, B.B., 1969. Long-run linearity, locally Gaussian process, H-Spectra and infinite variances. Int. Econ. Rev. 10, 82–111.
Mandelbrot, B.B., 1971. When can price be arbitraged efficiently? a limit to the validity of the random walk and martingale models. Rev. Econ. Stat. 53, 225–236.
Mandelbrot, B.B., 1972. Statistical methodology for nonperiodic cycles from covariance to R/S analysis. Ann. Econ. Soc. Meas. 1, 259–290.
Mandelbrot, B.B. , 2008. The (Mis)Behavior of Markets. A Fractal View of Risk, Ruin and Reward.Profile Books, London, UK.
Mandelbrot, B.B., Wallis, J.R., 1969. Robustness of the rescaled range R/S in the measurement of noncyclic long run Statistical dependence. Water Resour. Res. 5, 967–988.
Muniandy, S.V., Lim, S.C., Murugan, R., 2001. Inhomogeneous scaling behaviors in Malaysian foreign currency exchange rates. Phys. A: Stat. Mech. its Appl. 301, 407–428.
Nadarajah, S., Chu, J., 2017. On the inefficiency of Bitcoin. Econ. Lett. 150, 6–9.
Pal, D., Mitra, S.K., 2019. Hedging bitcoin with other financial assets. Financ. Res. Lett. 30, 30–36.
Pearson, K.F.R.S., 1900. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Lond., Edinb., Dublin Philos. Mag. J. Sci. 50 (302), 157–175.
Peng, C.K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., Goldberger, A.L., 1994. Mosaic organization of DNA nucleotides. *Phys. Rev.* E 49 (2), 1685–1689.
Platanakis, E., Urquhart, A., 2020. Should investors include Bitcoin in their portfolios? a portfolio theory approach. Br. Account. Rev. 52 (4), 100837.
Raimundo, M.S., Okamoto Jr., J., 2018. Application of hurst exponent (H) and the R/S analysis in the classification of FOREX securities. Int. J. Model. Optim. 8, 116–124.
Rejichi, I.Z., Aloui, C., 2012. Hurst exponent behavior and assessment of the MENA stock markets efficiency. Res. Int. Bus. Financ. 26, 353–370.
Sensoy, A., 2019. The inefficiency of Bitcoin revisited: a high-frequency analysis with alternative currencies. Financ. Res. Lett. 28, 68–73.
Shen, D., Urquhart, A., Wang, P., 2020. A three-factor pricing model for cryptocurrencies. Financ. Res. Lett. 34, 101248.

Taleb, N.N., 2010. The Black Swan. Random House,, New York, NY.
Taleb, N.N., 2020. Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistemology, and Applications. Papers and Commentary. Academic Press,, STEM.
Tiwari, A.K., Jana, R.K., Das, D., Roubaud, D., 2018. Informational efficiency of Bitcoin—an extension. Econ. Lett. 163, 106–109.
Urquhart, A., 2016. The inefficiency of Bitcoin. Econ. Lett. 148, 80–82.
Warusawitharana, M., 2019. Time-varying volatility and the power law distribution of stock returns. J. Empir. Financ. 49, 123–141.
White, E., Enquist, B., Green, J.L., 2008. On estimating the exponent of power law frequency distributions. Ecology 89, 905–912.
Wu, L., Chen, S., 2020. Long memory and efficiency of Bitcoin under heavy tails. Appl. Econ. 52, 5298–5309.
Zargar, F.N., Kumar, D., 2019. Informational inefficiency of Bitcoin: A study based on high-frequency data. Res. Int. Bus. Financ. 47, 344–353.