

Article

Shannon Entropy: An Econophysical Approach to Cryptocurrency Portfolios

Noé Rodriguez-Rodriguez  and Octavio Miramontes * 

Departamento de Sistemas Complejos, Instituto de Física, Universidad Nacional Autónoma de México, Ciudad de México C.P. 04510, Mexico

* Correspondence: octavio@fisica.unam.mx

Abstract: Cryptocurrency markets have attracted many interest for global investors because of their novelty, wide on-line availability, increasing capitalization, and potential profits. In the econophysics tradition, we show that many of the most available cryptocurrencies have return statistics that do not follow Gaussian distributions, instead following heavy-tailed distributions. Entropy measures are applied, showing that portfolio diversification is a reasonable practice for decreasing return uncertainty.

Keywords: cryptocurrencies; econophysics; entropy; portfolio uncertainty; heavy-tailed distributions



Citation: Rodriguez-Rodriguez, N.; Miramontes, O. Shannon Entropy: An Econophysical Approach to Cryptocurrency Portfolios. *Entropy* **2022**, *24*, 1583. <https://doi.org/10.3390/e24111583>

Academic Editor: Stanisław Drożdż

Received: 28 September 2022

Accepted: 30 October 2022

Published: 1 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Financial mathematics and financial economics are two scientific disciplines that together are the backbone of modern financial theory. Both disciplines make extensive use of their own models and theories from mathematics and traditional economics, and are aimed mostly towards predictive analysis of markets. On the other hand, physics has historically had an important influence on economic theory; the formalism of thermal equilibrium was inspirational in the development of the theory of economic general equilibrium [1]. In recent years, there has been a renewed interest in the view that economic phenomena share many aspects of physical systems, and they are thus susceptible to being studied under the science of complex systems [2–4] and statistical physics [5], giving rise to novel research fields such as econophysics [6].

One of the first successful econophysical approaches was the discovery that stock market fluctuations are not Gaussian, and instead are heavy-tailed. Mandelbrot came to this conclusion when investigating cotton prices [7], followed later by Mantegna when characterizing the Milan stock market [8]. This discovery forced a serious rethinking of the view introduced by Louis Bachelier in the early 20th Century stating that price variations are random and statistically independent [9]; while historically important, this view not entirely appropriate. These prices are the outcome of the concurrent nonlinear action of many economic agents [10], and as such the fluctuations are usually correlated, meaning that the process is not a random Brownian walk at all. In this article, we explore fluctuations in the cryptocurrency market and show that all of the most common coins analyzed fail the Shapiro–Wilk normality test, and as such are best explained by heavy-tailed statistical models; for further reference, see [11,12].

On the other hand, entropy is a fundamental key quantity in physics in both thermodynamics and information theory. It is related to disorder and diversity, and has been previously used in finance, specifically in portfolio selection theory [13]. Acceptance of entropy in economic theory has historically seen a degree of reticence, for instance, that expressed by Paul Samuelson, an influential figure in 20th Century economics [14]. However, has gained wide acceptance lately [15,16], specially with the development of non-equilibrium thermodynamics and complex systems theory [17–20]. It is well known that portfolio diversification is a good strategy to minimize specific risks, and as such we explore

the use of entropy in the cryptocurrency market as a measure of return uncertainty and risk minimization [13,19,21]. It is worth mentioning that recent studies elsewhere have been interested in exploring the use of entropy in the cryptocurrency markets, where heavy-tailed distributions and multi-scaling phenomena have been identified as well [11,12].

When speaking of information theory, its most powerful tools are considered to be information measures (MI); although the most famous MI is Shannon's entropy, Fisher's information is worth mentioning. Shannon's entropy measures missing information based on the Probability Distribution Function (PDF). The counterpart of the Shannon entropy is the Fisher information, which increases as the Shannon entropy decreases. Furthermore, the Fisher information reflects the amount of positive information from the observer, and depends on the PDF as its first derivatives. For this reason, the Fisher entropy is a local measure, while the Shannon entropy is a global measure. There have been studies elsewhere [22] showing that from the market point of view, minimizing risk is equivalent to minimizing Fisher's information; the present study provides empirical results that contribute to these perspectives.

2. Methods, Data and Analysis

In this article, we use the historical daily prices of 18 cryptocurrencies as presented in CryptoDataDownload [23] spanning at least 3 years. These are: Basic Attention Token (BAT), Bitcoin Cash (BCH), Bitcoin (BTC), Dai (DAI), Eidoo (EDO), Eos (EOS), Ethereum Classic (ETC), Ethereum (ETH), Metaverse ETP (ETP), Litecoin (LTC), Neo (NEO), OMG Network (OMG), Tron (TRX), Stellar (XLM), Monero (XMR), Verge (XVG), Ripple (XRP), and Zcash (ZEC) [24,25].

We used the opening and closing prices of the cryptocurrencies, quoted in US Dollars from 16 October 2018 to 31 December 2021, for a total of 1172 observations per cryptocurrency, to calculate daily returns. First, the distributions of the daily returns of each cryptocurrency were statistically characterized through normality tests and parametric adjustments of heavy-tailed distributions. Subsequently, we performed an analysis in which entropy functions were derived, similarly as in Dionisio et al. (2006), [13], Ormos and Zibriezky (2014) [19], and Mahmoud and Naoui (2017) [26]. From the set of 18 cryptocurrencies, the assets were randomly selected to compose investment portfolios, where the only premise used was that the proportion invested in each asset was $\frac{1}{N}$, with N being the number of assets in the portfolio. In order to consistently compare entropy to standard deviation, normal entropy was used, as it is a function of variance. It is important to mention that the time period chosen in this study was arbitrary; while it includes fluctuations due to the COVID-19 pandemic (see [11,27,28]), it certainly does not include recent events that have influenced the behavior of the cryptocurrency markets, such as the war in Eastern Europe and the increased economic rivalry between the USA and China.

In the following subsections, we review several useful concepts related to entropy functions: the discrete entropy function, the continuous entropy function, entropy as a measure of uncertainty, comparison between entropy and variance, and investment portfolios.

The appendices contain the methods used for the analysis of this article. Because the Shannon entropy is only a function of the PDF for the continuous case, Appendix A.1 shows how the problem should be treated when the PDF is not known and only finite experimental data are available. Appendix A.1.1 provides a complete solution for entropy estimation by means of the equidistant histogram method. Appendix A.1.2 shows how the PDF is found via the Kernel method. Appendix A.1.3 explains how the PDF is found via the parametric method. Furthermore, Appendix A.1.4 explains the max log-likelihood method that is mentioned throughout the article. Finally, Appendix A.2 shows the functional form of the distributions used for the parametric fits.

2.1. Discrete Entropy Function

Let X be a discrete random variable, $\{A_1, A_2, A_3, \dots, A_n\}$ be a set of possible events, and the corresponding probabilities $p_X(x_i) = \Pr(X = A_i)$, with $p_X(x_i) \geq 0$ and $\sum_{i=1}^n p_X(x_i) = 1$.

The generalized discrete entropy function, or Rényi entropy [19,29], for the variable X is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^n p_X(x_i)^\alpha \right) \quad (1)$$

where α is the order of the entropy and $\alpha \geq 0$. This order can be considered as a bias parameter, where $\alpha \leq 1$ favors rare events and $\alpha \geq 1$ favors common events [30]. The base of the logarithm is 2.

When $\alpha = 1$, this indicates a special case of generalized entropy that assumes ergodicity and independence, which the generalized case does not. However, substituting into (1) results in division by zero. By means of L'Hôpital's rule, it can be shown that when α tends to 1, we have the Shannon entropy:

$$H_1(X) = - \sum_{i=1}^n p_X(x_i) \log(p_X(x_i)) \quad (2)$$

The Shannon entropy produces exponential equilibrium distributions, while generalized entropy produces power law distributions.

2.2. Continuous Entropy Function

Let X be a continuous random variable that takes values of $\mathbb{R}$ and let $p_X(x)$ be the density function of the random variable. Continuous entropy is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \ln \int p_X(x)^\alpha \quad (3)$$

Note that the logarithm bases of (1) and (3) are different. Although the entropy depends on the base, it can be shown that the value of the entropy changes only by a constant coefficient for different bases.

When $\alpha = 1$, we have the Shannon entropy for the continuous case

$$H_1(X) = - \int p_X(x) \ln(p_X(x)) dx \quad (4)$$

The properties of discrete and differential entropy are similar. The differences are that the discrete entropy is invariant under variable changes and the continuous entropy is not necessarily; furthermore, the continuous entropy can take negative values.

2.3. Entropy as a Measure of Uncertainty

According to Shannon (1948) [31], an uncertainty measure $H(p_X(x_1), p_X(x_2), \dots, p_X(x_n))$ must satisfy the following:

1. H must be continuous on $p_X(x_i)$, with $i = 1, \dots, n$.
2. If $p_X(x_i) = \frac{1}{n}$, H must be monotone increasing as a function of n .
3. If an option is split into two successive options, the original H must be the weighted sum of the individual values of H .

Shannon showed that a measure that satisfies all these properties is (2) multiplied by any positive constant (the constant sets the unit of measure). Among the properties that make it a good uncertainty choice are:

1. $H(X) = 0$ if and only if all but one of $p_X(x_i)$ are zero.
2. When $p_X(x_i) = \frac{1}{n}$, i.e., when the discrete probability distribution is constant, $H(X)$ is maximum and equal to $\log(n)$.
3. $H(X, Y) \leq H(X) + H(Y)$, where the equality holds if and only if X and Y are statistically independent, i.e., $p(x_i, y_j) = p(x_i)p(y_j)$.
4. Any change towards the equalization of the probabilities $p_X(x_i)$ increases H .

5. $H(X, Y) = H(X) + H(Y|X)$; thus, the uncertainty of the joint event $(Y|X)$ is the uncertainty of X plus the uncertainty of Y when X is known.
6. $H(Y) \geq H(Y|X)$, which implies that the uncertainty of Y is never increased by knowledge of X , and decreases unless X and Y are independent, in which case it does not change.

2.4. Comparison between Entropy and Variance

Ebrahimi et al. (1999) [32] showed that entropy can be related to higher order moments of a distribution, and thus can offer a better characterization of $p_X(x)$, as it uses more information about the probability distribution than the variance (which is only related to the second moment of a probability distribution).

The entropy measures the disparity of the density $p_X(x)$ of the uniform distribution. That is, it measures uncertainty in the sense of using $p_X(x)$ instead of the uniform distribution [26], while the variance measures an average of distances from the mean of the probability distribution. According to Ebrahimi et al. [32], both measures reflect concentration, although they use different metrics. The variance measures the concentration around the mean and the entropy measures the density diffusion regardless of the location of the concentration. Statistically speaking, entropy is not a mean-centered measure; it takes into account the entire empirical distribution without concentrating on a specific moment. In this way, the entire distribution of returns can be considered without focusing on one in particular [33]. The discrete entropy is positive and invariant under transformations, while the variance is not. In the continuous case, neither the entropy nor the variance are invariant under one-to-one transformations [13,26]. According to Pele et al. (2017) [34], entropy is strongly related to the tails of the distribution; this feature is important for distributions with heavy-tails or with an infinite second-order moment, where the variance is obsolete. Furthermore, the entropy can be estimated for any distribution without prior knowledge of its functional form. The authors mentioned above found that heavy-tailed distributions generate low entropy levels, while light-tailed distributions generate high entropy values.

2.5. Investment Portfolios

A portfolio or investment portfolio is simply a collection of assets. They are characterized by the value invested in each asset. Let w_i be the fraction invested in asset i with $i = 1, 2, \dots, n$; the required constraint is that

$$\sum_{i=1}^n w_i = 1 \quad (5)$$

We define the return R_i of a common share i during a certain period as

$$R_i = \frac{P_{1,i} - P_{0,i}}{P_{0,i}} \quad (6)$$

where $P_{0,i}$ is the price of stock i at the beginning of the period and $P_{1,i}$ is the price of i at the end of that period. This return is the historical return or ex post return, whereas the total portfolio return is simply the weighted average of the expected returns of the individual securities in the portfolio:

$$R_T = \sum_{i=1}^n w_i R_i \quad (7)$$

The risk of investment portfolios can be divided into specific and systematic risk. Systematic risk is inherent to market uncertainty, and is not diversifiable; usually, the price of an asset is affected by factors such as inflation, economic growth or recession, and fluctuations in the world financial markets. The specific or unsystematic risk corresponds to the risk of an asset or a small group of assets due to its specific characteristics, and is diversifiable. Forming portfolios can reduce specific or unsystematic risk.

Entropy can provide similar information if we define the average mutual information between two random variables X and Y , where one is the independent variable and the other is the dependent variable, as follows:

$$I(X, Y) = \sum_{i=1}^n \sum_{j=1}^m p_{X,Y}(x_i, y_j) \log \left(\frac{P_{X,Y}(x_i|y_j)}{P_X(x_i)P_Y(y_j)} \right) \quad (8)$$

for the discrete case and

$$I(X, Y) = \int_X \int_Y p_{X,Y}(x, y) \log \left(\frac{P_{X,Y}(x|y)}{P_X(x)P_Y(y)} \right) \quad (9)$$

for the continuous case. This measure reflects the correlation between X and Y . Thus $H(X) = I(X, Y) + H(X|Y)$, where $I(X, Y)$ can be compared to systematic risk, while $H(X|Y)$ can be compared to the specific risk.

In practice, standard risk measures such as CAPM beta or standard deviation are calculated on daily or monthly returns. We follow this methodology here as well. Because the returns of the values can take values of a continuous co-domain, we focus mainly on the differential entropy.

The assumption that data and residuals follow a normal distribution is common in both portfolio management and risk analysis. Therefore, in this context, in order to parametrically estimate the entropy of a normal distribution (because information theory measures cannot be directly compared to variance in metric terms), we use

$$NH(X) = \int p_X(x) \log(\sqrt{2\pi}\sigma) dx + \int p_X(x) \frac{(x - \bar{x})^2}{2\sigma^2} dx = \log(\sqrt{2\pi e}\sigma) \quad (10)$$

To calculate the empirical differential entropy, it is only necessary to estimate the Probability Distribution Function (PDF), for which there are essentially three methods: histograms, kernels, and parametrically (see Appendix A for details).

3. Results

We present our findings in three parts. First, it is shown that the daily returns of each cryptocurrency do not follow a normal distribution. Second, it is indicated that the daily returns have heavy tails. Finally, the behavior of the empirical and normal entropy is shown as a function of the number of assets in the portfolio. As far as we know, this is the first time that such a large number of cryptocurrencies has been analyzed together; however, see [11,12].

3.1. Normal Distribution and Q-Q Plots

Figure 1, made according to Equations (6) and (A5), shows as an example the histogram of the probability distribution of the daily returns of Bitcoin considering the daily opening and closing price. The blue curve shows the normal distribution with the variance and mean calculated from the empirical data. The red line shows the Kernel density estimator; the cross-validation method was used to determine the bandwidth and the Kernel that maximize the total Log-Likelihood of the data in X [35]. Although there are alternative principles, such as MaxEnt or FIM-minimization, for estimating the best PDF, log-likelihood was used for simplicity. Among the kernels used were Gaussian, Epanechnikov, Tophat, exponential, linear, and cosine. Bandwidth was optimized for values between 0 and 1. For all cryptocurrencies, a significant deviation of the distribution calculated with Kernels was found with respect to the normal distribution.

Quantiles are values that divide a probability distribution into equal intervals, with each interval having the same fraction of the total population. Q-Q plots are commonly used to visualize data and comparatively find the type of probability distribution to which a random variable can belong, for example, whether they are Gaussian, uniform, Paretian, exponential distributions, etc. To build a Q-Q plot, the quantiles of the base distribution, or

the “theoretical quantiles”, are plotted on the X axis, while the sample quantiles are plotted on the Y axis. If both sets of quantiles come from the same distribution, then the points form a more or less straight line.

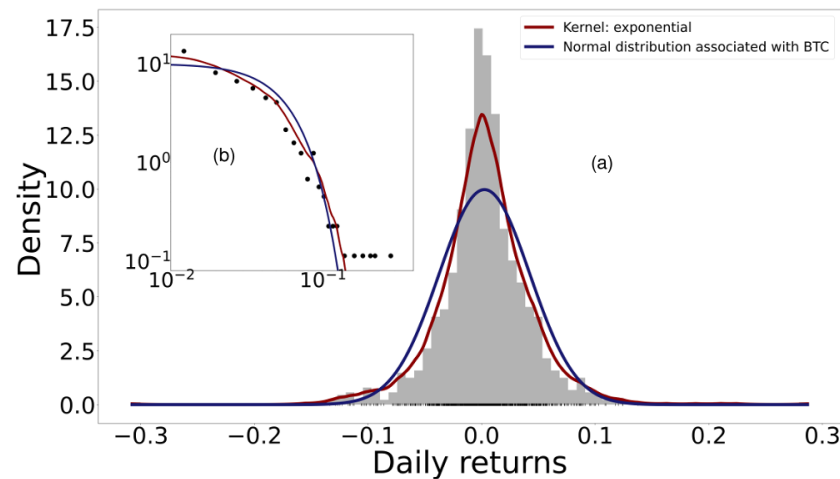


Figure 1. (a) Histogram of the probability distribution of the return of Bitcoin. The normal curve is shown in blue, associated with the dataset, with the curve estimated using the Kernel method in red. (b) The log–log plot of the same data is depicted in the inset. Notice that the normal distribution tail decays faster in comparison with the kernel, as expected.

Figure 2 shows the normal Q-Q plot of the empirical daily returns of Bitcoin. Across all cryptocurrencies, the normal distribution was found to capture the middles of the data well, but not the tails. In fact, the characteristic behavior of heavy-tailed distributions can be clearly seen by the “S” shape on the graph. This means that compared to a normal distribution there are more data located in the tail of the heavy-tailed distribution.

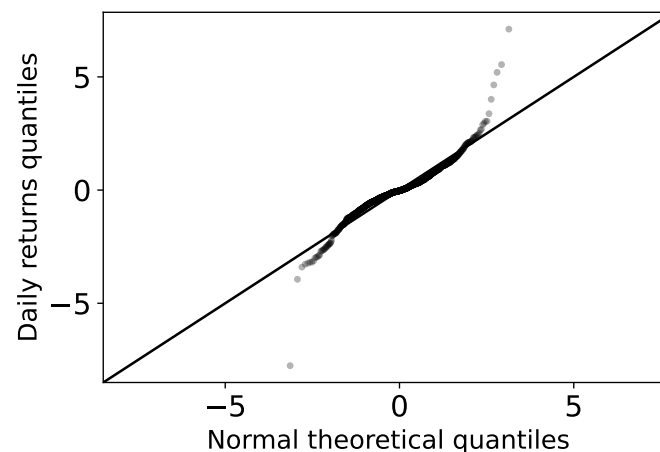


Figure 2. Normal Q-Q chart of daily Bitcoin performance. The theoretical normal quantiles are plotted against the observed quantiles.

The Shapiro–Wilks test posits the null hypothesis that a sample comes from a normal distribution. When applying this statistical test to the daily returns of all the cryptocurrencies studied, it was found that the p -value is at most on the order of 10^{-18} . It is established as common practice that if the p value is less than 0.05, the hypothesis that the sample is normal should be rejected. In our case, the value of p is very small (indeed, almost negligible) for all cryptocurrencies, which strongly indicates that the null hypothesis that daily returns follow a normal distribution [36] should be rejected. Therefore, we accept the alternative hypothesis that the returns do not present normality.

3.2. Heavy-Tailed Distributions

Nadarajah et al. (2015) [37] made adjustments to the exchange rate of various fiat currencies (money that is not backed by securities or physical assets such as gold, and is instead backed by the government that issues it) using flexible distributions. They used distributions such as Student's *t*, skewed Student's *t*, hyperbolic, generalized hyperbolic, generalized lambda, Skew-T, and Gaussian inverse normal to analyse the data. According to the previous section, we must limit ourselves to the study of distributions with long tails. Osterrieder (2016) [24] performed heavy-tailed fits for different cryptocurrencies using Student's *t*-test, Generalized Student's *t*, Hyperbolic, Generalized Hyperbolic, Gaussian Inverse Normal, and Asymmetric Gamma variance distribution. In both studies, they found that all the heavy-tailed distributions provided statistically similar results, although the best fit was found with the generalized hyperbolic distribution. In addition, they concluded that Student's *t*-test is a good option considering its simplicity. Similarly, Briere et al. (2017) [38] reached similar results when studying parametric adjustments to the returns of the seven most traded cryptocurrencies of 2015. Therefore, we make adjustments here to the daily returns with the Student's *t*-distribution and the normal distribution for comparison, although it is light-tailed.

Figure 3 shows the histogram of the probability distribution of Bitcoin's daily performance along with the adjustments of each distribution, meaning that the log-likelihood is maximized. The log-likelihood value for each fit is shown in Table 1. Note that the log-likelihood is always smaller for the normal distribution, while the heavy-tailed Student's *t*-distribution generates good fits.

Table 1. Log-likelihood of the fitted distributions.

Cryptocurrency	Normal	t-Student
BAT	1531	1667
BCH	1555	1815
BTC	2111	2250
DAI	4395	4871
EDO	1239	1518
EOS	1583	1794
ETC	1608	1861
ETH	1831	1915
ETP	1393	1594
LTC	2111	2250
NEO	1631	1740
OMG	1401	1570
TRX	1694	1829
XLM	1590	1808
XMR	1814	1947
XRP	1595	1884
XVG	1305	1396
ZEC	1637	1722

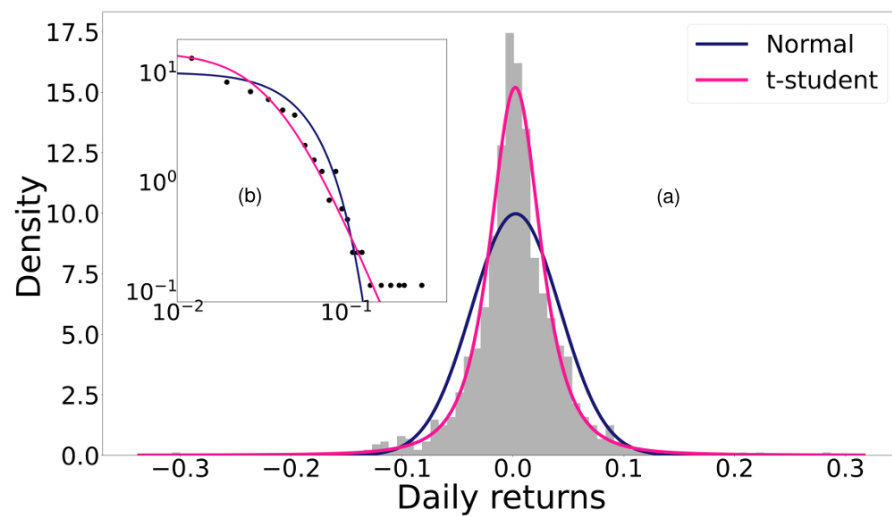


Figure 3. (a) Histogram of the probability distribution of the return of Bitcoin with the adjustments of a normal distribution and Student's *t*. (b) The log–log plot of the same data is depicted in the inset. Notice that the normal distribution tail decays faster in comparison with the Student's *t*-distribution, as expected.

3.3. Portfolio Entropy

A portfolio was generated by adding the 18 cryptocurrencies randomly. We began by adding an asset to the portfolio, then its normal entropy and empirical entropy were calculated using the histogram, Kernel, and parametric methods. Assets continued to be added randomly until the 18 cryptocurrencies studied had all been added. Every time a new asset was added to the portfolio, it was necessary to calculate the normal entropy and the empirical entropy of the total return; i.e., every time an asset was added, the distribution function of the total return was calculated again by means of the three methods presented. Figure 4 shows the adjustments when performing the previous procedure, while Figure 5 shows the normal and empirical entropy against the number of assets in the portfolio.

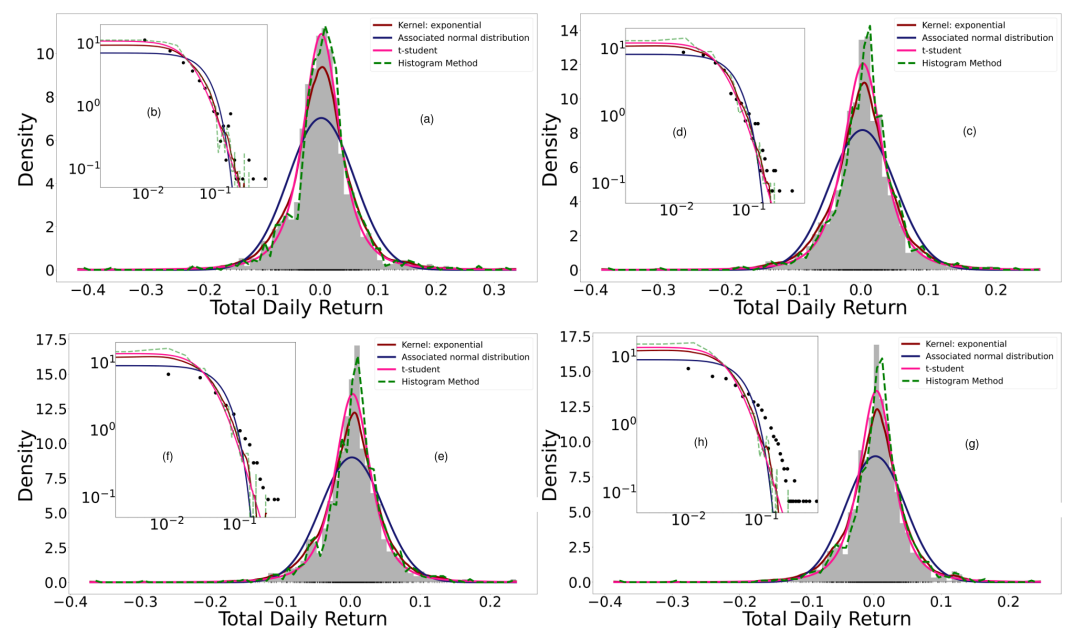


Figure 4. (a,c,e,g) correspond to the adjustments of the probability distribution of the total daily returns with $n \in \{3, 8, 13, 18\}$ assets in the portfolio, respectively, with the kernel, parametric, and histogram methods and the normal curve associated with the data. In (b,d,f,h), the same fits are shown in log–log scale.

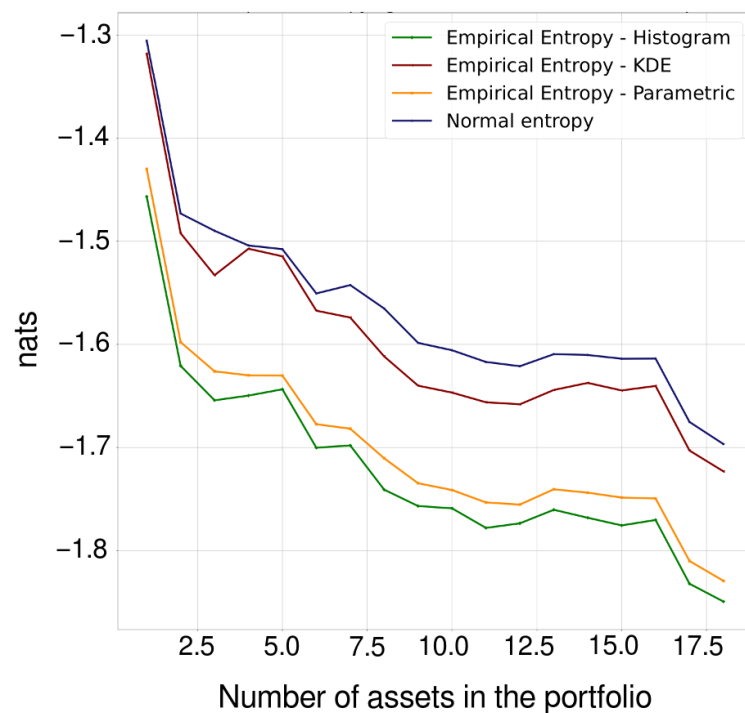


Figure 5. Normal and empirical entropy against the number of assets in the portfolio. The decrease in entropy as the number of assets in the portfolio increases verifies the diversification effect.

4. Discussion

The yields of the 18 cryptocurrencies were studied, with several of their statistical properties shown in detail. For this, the daily returns were adjusted to different heavy-tailed distributions such as the normal and Student's *t*-distributions. Our results show that the distributions of daily yields exhibit statistically significant leptokurtic return distributions with heavy tails, as expected. In particular, Student's *t*-distribution adequately describes the data and is a good option considering its simplicity; it can be useful for financial risk management, where it is necessary to calculate the value at risk (VaR) and the expected deficit (ES). The aforementioned results are useful for investment purposes as well. As far as we know, this is one of the few works to investigate the statistical properties of many cryptocurrencies together, beyond just Bitcoin or Ethereum, although see [11,12] as well.

Entropy is used here as a measure of uncertainty in portfolio management. Figure 5 shows that the Shannon entropy behaves similarly, although not exactly the same, as the variance, serving as a measure of risk and to verify the effect of diversification, both of which tend to decrease when including assets in the portfolio. As the number of assets in the portfolio increases, the possible number of states of the system, in this case the portfolio, progressively increases as well, and the uncertainty about this portfolio decreases according to the property $H(X, Y) \leq H(X) + H(Y)$. We should note that the normal entropy always takes values greater than the empirical entropy in all cases, which implies that the uncertainty is less than that which would be observed if the returns were normally distributed.

In the study by Ormos and Zibriezy (2014) [19], it was found that the histogram method for calculating the Shannon entropy was more efficient in terms of explanatory and predictive power and exhibited simplicity compared to the Kernel method. We observe that the fits of Figure 4 to the daily total return of the generated portfolios show that the histogram method presents overfitting compared to the empirical data. The small "jumps" observed in the tails with this method are evident in the log–log scaled plot. On the other hand, the best fit is parametric, followed by KDE. In these last two cases, the log-likelihood of the parametric method with Student's *t*-distribution is always greater than that of the Kernel Density Estimation (KDE) method, although the log-likelihood of both is always

greater than that of the normal distribution. Regardless, diversification is verified with any method. This corroborates and contributes to the empirical findings made by Dionisio et al. (2006) [13], Ormos and Zibriezky (2014) [19], and Mahmoud and Naoui (2017) [26], all of whom observed the diversification of the Shannon entropy, with the difference that these authors verified it in classical assets and did not characterize their distributions. Our results can be useful to investors, traders, and portfolio managers.

5. Conclusions

Based in our findings, we conclude that entropy observes the effect of diversification and is a more general measure of uncertainty than the standard deviation. This is the case for the following reasons. (i) It uses more information from the probability distribution, as it uses moments of higher order while the deviation standard only uses the second moment. (ii) It does not depend on any particular distribution, unlike the standard deviation, which eliminates the error introduced by fitting a normal distribution to returns. This becomes evident in non-symmetric distributions with additional non-normal moments. In this way, entropy can capture the complexity of systems without the need for hypotheses that could bias the results. (iii) Entropy is independent of the mean considering any distribution, meaning that it satisfies the first-order conditions. (iv) Entropy meets the requirements that good uncertainty measures must satisfy. (v) Entropy can be used for both metric and non-metric data (quantitative and qualitative data, respectively). For this reason, it can be used as a measure to complement traditional models, thanks to a mean and variance that are more restrictive, when dealing with assumptions that tend not to be verified empirically.

The drawbacks of the potential use of entropy as a measure of uncertainty in portfolios should be mentioned as well, one of them being that it is more complex compared to the more common standard deviation. On the other hand, entropy does not take into account the real values of the variables, meaning that care must be taken when using it in risk analysis and portfolio selection. In addition, there is always statistical bias in data measurements due to the degrees of freedom allowed in an experiment.

Although the present study does not address punctual prediction of the cryptocurrency market, it is illustrative to mention the relative success in this area of the use of nonlinear physics methods to make substantive predictions, among others by the “Prediction Company” founded by two pioneers of chaos theory in 1991 and associated with the Santa Fe Institute. This is the case of Doyné Farmer and Norm Packard [39]. In one of his best-known contributions [40], Farmer mentions the possibility of using “adaptive dynamics” to make predictions in natural time series, which follows the principles of neural networks, more currently known as known as “machine learning”. His study basically consists of the use of systems that learn to classify patterns (output) from a time series (input). The output would ideally be an investment strategy that would aim to maximize profits and the input would be the immediate, ideally real-time, time series of market behavior. Of course, there are recent applications of machine learning and other methods to the prediction of the temporal behavior of cryptocurrencies [41,42].

For the temporal prediction of financial markets, today there are more than 5000 algorithms created by various programmers who publish them in public software libraries [43]. Investors typically use several of these algorithms in combination, although the most popular are usually based on Average True Value (ATR), Relative Strength Index (RSI), Moving Average Convergence/Divergence (MACD), and Exponential Moving Average (EMA) [43]. As can be seen, the use of moving averages is common, as they essentially help to reduce noise or fluctuations in time intervals. Several authors [44,45] have coupled them with regression analysis to form ARMA (autoregressive moving average) models, which in turn are used together with skewed (non-normal) probability distribution functions because, as we have already confirmed here, the PDFs of the financial markets and cryptocurrencies are not normal. It is interesting to note that there are “forecasts” online that suggest the right times, in real time, to sell or buy cryptocurrencies, and that these are mostly based on moving average models [46].

Finally, it is illustrative to mention that cryptocurrency markets have qualitative and quantitative properties similar to traditional financial markets [44,45], which allows us to speak of generic properties between the two, that is, aspects of universality typical of complex systems. Studying the statistical behavior of cryptocurrency markets, a phenomenon we have witnessed from its birthday while having at hand the whole arsenal of contemporary mathematical tools, is a fantastic opportunity to understand the real-world dynamics of a complex economical system, the interactions of the agents involved, and the emergent collective properties that may influence our society. For example, in future cryptocurrency markets, financial markets, e-commerce markets, and energy markets [28,47–50]). We hope our contribution can be useful in understanding the statistical properties of contemporary market and portfolio modeling.

Author Contributions: N.R.-R. wrote the code for analyzing the data and prepared the tables and figures. Both authors designed the research and contributed equally to the writing of the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a grant (PRIDIF-2021) from the Physics Institute, National Autonomous University of Mexico (UNAM).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data analyzed in this article are freely available from the following source: <https://www.cryptodatadownload.com>, accessed on 1 January 2022.

Acknowledgments: N.R.-R. would like to thank the Physics Institute at UNAM in Mexico City for providing financial support.

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

Appendix A

Appendix A.1. Probability Density Function Estimation (PDF)

Let $x_1, x_2, \dots, x_n$ be the observations of the continuous random variable X and let $H_{\alpha,n}(X)$ be the sample estimate of $H_{\alpha}(X)$. The evaluation of the entropy is obtained by estimating the density function. The probability density function $p_X(x)$ is approximated by $p_{X,n}(x)$. Thus, the integral estimate of the entropy is

$$H_{\alpha,n}(X) = \frac{1}{1-\alpha} \ln \int_{A_n} p_{X,n}(x)^{\alpha} dx \quad (A1)$$

where $A_n = (\min(x), \max(x))$. For this reason, one of the difficulties in determining continuous entropy is that the underlying probability density function (or PDF) is unknown. There are three methods to overcome this problem: histogram-based, kernel-based, and parametric estimators [13,19].

Appendix A.1.1. Histogram

Let $b_n = (\max(x), \min(x))$ be the range of the sample values; we divide the range into k bins of equal width where the cuts are made at the points t_j . The width of each bin is $h = \frac{b_n}{k} = t_{j+1} - t_j$. The density function is estimated by

$$p_{X,n}(x) = \frac{v_j}{nh} \quad (A2)$$

if $x \in (t_j, t_{j+1})$, with v_j being the number of points that land in the j th bin. From (4) and (A1),

$$H_{1,n}(X) = \frac{1}{n} \sum_{j=1}^k v_j \ln \left(\frac{v_j}{nh} \right) \quad (A3)$$

The hyperparameter of this method is the number of bins, k . There are several methods to choose this parameter; for example Scott's (1979) [51] normal reference rule, the Freedman–Diaconis (1981) [52] rule, the square root rule, etc. In particular, the Freedman–Diaconis rule is quite robust and usually provides good results in practice. The method consists of minimizing the integral of the squared difference between the histogram and the density of the theoretical distribution. The bandwidth in this method is provided by

$$h = 2 \cdot IQR(x) \cdot n^{-1/3} \quad (A4)$$

where n is the number of observations and $IQR(x)$ is the interquartile range of the dataset. In this way, the optimal number of containers is chosen as

$$k = \frac{b_n \cdot n^{1/3}}{2 \cdot IQR(x)} \quad (A5)$$

Appendix A.1.2. Kernel Density Estimation

Let $x_1, \dots, x_n$ be a sample of independent and identically distributed random variables; then, an estimator of the distribution function is

$$p_{X_n}(x) = \frac{1}{nh} \sum_{i=1}^n \int_{-\infty}^x K\left(\frac{u - x_i}{h}\right) du \quad (A6)$$

where K is the kernel function and h is the bandwidth parameter. The function K must be a real function such that $K(x) \geq 0$, $K(x) = K(-x)$ for all x in $\mathbb{R}$, $\int_{-\infty}^{\infty} K(x) dx = 1$, and $\int_{-\infty}^{\infty} xK(x) dx = 0$. The uniform convergence to the theoretical distribution has been studied regardless of the form of the kernel used [34]. The hyperparameters are the Kernel and the bandwidth; the choice of Kernel determines how the influence of each observation is distributed.

Appendix A.1.3. Parametric Density Estimation

The set of probability distributions in a sample space indexed by a set Θ (parameter space) is denoted as $\mathcal{P}$. For each $\theta \in \Theta$, there is a P_θ member of $\mathcal{P}$ which is a distribution. Statistical models can be written as

$$\mathcal{P} = \{P_\theta | \theta \in \Theta\} \quad (A7)$$

These models are parametric if $\Theta \subseteq \mathbb{R}^k$ with $k \geq \in \mathbb{Z}^+$, otherwise we would have a non-parametric model like the estimators of kernels, in which case, the model consists of continuous distributions. Parametric methods assume that the particular shape of the probability distribution function is known and only its parameters need to be estimated; in the case of a normal distribution, only its mean and variance need to be estimated.

Appendix A.1.4. Maximum Likelihood Estimate

Consider a set $(x_1, x_2, \dots, x_n)$ of i.i.d. Because they are identically distributed, they must have the same probability function $f(X|\theta)$. The likelihood of the data having the parameter θ is

$$\mathcal{L}(\theta) = \prod_{i=1}^n f(x_i|\theta) \quad (A8)$$

For different θ , the probability of the data having those parameters is different. In maximum likelihood estimation (MLE), the goal is to find the parameters θ that maximize the likelihood function $\mathcal{L}$. The MLE assumes that (the arguments of the maximum, or *argmax*, are those points in the domain of a function at which it is maximized):

$$\hat{\theta} = \arg \max_{\theta} (\mathcal{L}(\theta)) \quad (A9)$$

The log-likelihood function is usually used:

$$\mathcal{LL}(\theta) = \log(\mathcal{L}(\theta)) = \log \prod_{i=1}^n f(x_i|\theta) = \sum_{i=1}^n \log f(x_i|\theta) \quad (\text{A10})$$

Because the logarithm is monotonically increasing, in practice the MLE works with

$$\hat{\theta} = \arg \max_{\theta} (\mathcal{LL}(\theta)) \quad (\text{A11})$$

For this work, the histogram, Kernel, and parametric methods were used to determine the PDFs while always using the Freedman–Diaconis rule. In the last two cases, the probability density functions were calculated by maximum likelihood estimation (MLE).

Appendix A.2. Parametric Distributions Used

Normal distribution

$$p_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (\text{A12})$$

for $-\infty \leq x \leq \infty$, $-\infty \leq \mu \leq \infty$ and $\sigma \geq 0$.

Non-standardized Student's t-distribution

$$p_X(x) = \frac{K(\nu)}{\sigma} \left[1 + \frac{(x-\mu)^2}{\sigma^2\nu} \right]^{-(1+\nu)/2} \quad (\text{A13})$$

for $-\infty \leq x \leq \infty$, $-\infty \leq \mu \leq \infty$ and $\sigma \geq 0$ u $\nu \geq 0$, where $K(\nu) = \sqrt{\nu}B(\nu/2, 1/2)$ and B is the beta function, defined by

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt \quad (\text{A14})$$

References

- Smith, E.; Foley, D.K. Classical thermodynamics and economic general equilibrium theory. *J. Econ. Dyn. Control* **2008**, *32*, 7–65. [\[CrossRef\]](#)
- Anderson, P.W.; Arrow, K.J.; Pines, D. *The Economy as An Evolving Complex System*; CRC Press: Boca Raton, FL, USA, 1988.
- Arthur, W.B. *Complexity and the Economy*; Oxford University Press: New York, NY, USA, 2014.
- Rosser, J.B. *Foundations and Applications of Complexity Economics*; Springer: Berlin/Heidelberg, Germany, 2021.
- Stanley, H.E.; Afanasyev, V.; Amaral, L.A.N.; Buldyrev, S.V.; Goldberger, A.L.; Havlin, S.; Leschhorn, H.; Maass, P.; Mantegna, R.N.; Peng, C.K.; et al. Anomalous fluctuations in the dynamics of complex systems: From DNA and physiology to econophysics. *Phys. A Stat. Mech. Appl.* **1996**, *224*, 302–321. [\[CrossRef\]](#)
- Mantegna, R.N.; Stanley, H.E. *Introduction to Econophysics: Correlations and Complexity in Finance*; Cambridge University Press: Cambridge, UK, 1999.
- Mandelbrot, B. The variation of some other speculative prices. *J. Bus.* **1967**, *40*, 393–413. [\[CrossRef\]](#)
- Mantegna, R.N. Lévy walks and enhanced diffusion in Milan stock exchange. *Phys. A Stat. Mech. Appl.* **1991**, *179*, 232–242. [\[CrossRef\]](#)
- Bachelier, L. Louis Bachelier's Theory of Speculation. In *Louis Bachelier's Theory of Speculation*; Princeton University Press: Princeton, NJ, USA, 2011.
- Bonanno, G.; Lillo, F.; Mantegna, R.N. Levels of complexity in financial markets. *Phys. A Stat. Mech. Appl.* **2001**, *299*, 16–27. [\[CrossRef\]](#)
- Wątopek, M.; Kwapien, J.; Drożdż, S. Financial return distributions: Past, present, and COVID-19. *Entropy* **2021**, *23*, 884. [\[CrossRef\]](#)
- Wątopek, M.; Drożdż, S.; Kwapien, J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* **2021**, *901*, 1–82. [\[CrossRef\]](#)
- Dionisio, A.; Menezes, R.; Mendes, D.A. An econophysics approach to analyse uncertainty in financial markets: an application to the Portuguese stock market. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **2006**, *50*, 161–164. [\[CrossRef\]](#)
- Samuelson, P. Gibbs in Economics. In *Proceedings of the Gibbs Symposium*, Yale University, New Haven, CT, USA, 15–17 May 1989; Caldi, D.G., Mostow, G.D., Eds.; American Mathematical Society: New Haven, CT, USA, 1990.
- Lauster, M.; Hoher, K.; Straub, D. A new approach to mathematical economics: On its structure as a homomorphism of Gibbs-Falkian thermodynamics. *J. Math. Anal. Appl.* **1995**, *193*, 772–794. [\[CrossRef\]](#)
- Jakimowicz, A. The role of entropy in the development of economics. *Entropy* **2020**, *22*, 452. [\[CrossRef\]](#)

17. Raine, A.; Foster, J.; Potts, J. The new entropy law and the economic process. *Ecol. Complex.* **2006**, *3*, 354–360. [\[CrossRef\]](#)
18. Prigogine, I. The rediscovery of value and the opening. In *The Evolutionary Foundations of Economics*; Dopfer, K., Ed.; Cambridge University Press: Cambridge, UK, 2005; p. 61.
19. Ormos, M.; Zibriczky, D. Entropy-based financial asset pricing. *PLoS ONE* **2014**, *9*, e115742. [\[CrossRef\]](#)
20. Almog, A.; Shmueli, E. Structural entropy: Monitoring correlation-based networks over time with application to financial markets. *Sci. Rep.* **2019**, *9*, 1–13. [\[CrossRef\]](#)
21. Goetzmann, W.N.; Brown, S.J.; Gruber, M.J.; Elton, E.J. *Modern Portfolio Theory and Investment Analysis*; John Wiley Sons: Hoboken, NJ, USA, 2014; Volume 237.
22. Makowski, M.; Piotrowski, E.W.; Frackiewicz, P.; Szopa, M. Transactional Interpretation for the Principle of Minimum Fisher Information. *Entropy* **2021**, *23*, 1464. [\[CrossRef\]](#)
23. CryptoDataDownload.com. Historical Data. Available online: <https://www.cryptodatadownload.com> (accessed on 1 January 2022).
24. Osterrieder, J. The statistics of bitcoin and cryptocurrencies. *SSRN Electron. J.* **2016**, 2872158, 1–12. [\[CrossRef\]](#)
25. Liang, J.; Li, L.; Zeng, D.; Zhao, Y. Correlation-based dynamics and systemic risk measures in the cryptocurrency market. In Proceedings of the 2018 IEEE International Conference on Intelligence and Security Informatics (ISI), Miami, FL, USA, 9–11 November 2018; pp. 43–48.
26. Mahmoud, I.; Naoui, K. Measuring systematic and specific risk: Approach mean-entropy. *Asian J. Empir. Res.* **2017**, *7*, 42–60. [\[CrossRef\]](#)
27. Neslihanoglu, S. Linearity extensions of the market model: A case of the top 10 cryptocurrency prices during the pre-COVID-19 and COVID-19 periods. *Financ. Innov.* **2021**, *7*, 1–27. [\[CrossRef\]](#)
28. Bazán-Palomino, W. Interdependence, contagion and speculative bubbles in cryptocurrency markets. *Financ. Res. Lett.* **2022**, *49*, 103132. [\[CrossRef\]](#)
29. Rényi, A. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*; University of California Press: Berkeley, CA, USA, 1961; Volume 4, pp. 547–562.
30. Tsallis, C.; Anteneodo, C.; Borland, L.; Osorio, R. Nonextensive statistical mechanics and economics. *Phys. A Stat. Mech. Appl.* **2003**, *324*, 89–100. [\[CrossRef\]](#)
31. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [\[CrossRef\]](#)
32. Ebrahimi, N.; Maasoumi, E.; Soofi, E.S. Ordering univariate distributions by entropy and variance. *J. Econom.* **1999**, *90*, 317–336. [\[CrossRef\]](#)
33. Deeva, G. Comparing entropy and beta as measures of risk in asset pricing. *Acta Univ. Agric. Silv. Mendel. Brunensis* **2017**, *65*, 1889–1894. [\[CrossRef\]](#)
34. Pele, D.T.; Lazar, E.; Dufour, A. Information Entropy and Measures of Market Risk. *Entropy* **2017**, *19*, 226. e19050226. [\[CrossRef\]](#)
35. Berrar, D. Cross-Validation. *Encycl. Bioinform. Comput. Biol.* **2019**, *1*, 542–554. [\[CrossRef\]](#)
36. Shaphiro, S.; Wilk, M. An analysis of variance test for normality. *Biometrika* **1965**, *52*, 591–611. [\[CrossRef\]](#)
37. Nadarajah, S.; Afuecheta, E.; Chan, S. A note on “Modelling exchange rate returns: Which flexible distribution to use?”. *Quant. Financ.* **2015**, *15*, 1777–1785. [\[CrossRef\]](#)
38. Briere, M.; Oosterlinck, K.; Szafarz, A. Virtual currency, tangible return: Portfolio diversification with Bitcoins. *J. Asset Manag.* **2015**, *16*, 365–373. [\[CrossRef\]](#)
39. Gwynne, P. Physicist who makes cash from chaos. *Phys. World* **2001**, *14*, 9. [\[CrossRef\]](#)
40. Farmer, J.D.; Sidorowich, J.J. Can New Approaches to Nonlinear Modeling Improve Economic Forecasts? In *The Economy as an Evolving Complex System*; Anderson, P.W., Arrow, K., Pines, D., Eds.; CRC Press: Boca Raton, FL, USA, 1988.
41. Cocco, L.; Tonelli, R.; Marchesi, M. Predictions of bitcoin prices through machine learning based frameworks. *PeerJ Comput. Sci.* **2021**, *7*, e413. [\[CrossRef\]](#)
42. Chen, W.; Xu, H.; Jia, L.; Gao, Y. Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. *Int. J. Forecast.* **2021**, *37*, 28–43. [\[CrossRef\]](#)
43. Tradingview.com. Indicators, Strategies and Libraries. Available online: <https://www.tradingview.com/scripts/indicator/> (accessed on 1 January 2022).
44. Lambert, P.; Laurent, S. Modelling skewness dynamics in series of financial data using skewed location-scale distributions. *Inst. De Stat. Louvain-La Discuss. Pap.* **2002**, 119.
45. Maleki, M.; Wraith, D.; Mahmoudi, M.R.; Contreras-Reyes, J.E. Asymmetric heavy-tailed vector auto-regressive processes with application to financial data. *J. Stat. Comput. Simul.* **2020**, *90*, 324–340. [\[CrossRef\]](#)
46. Tradingview.com. Indicators, Strategies and Libraries. Available online: <https://www.tradingview.com/symbols/BTCUSD/technical/> (accessed on 1 January 2022).
47. Fantazzini, D.; Calabrese, R. Crypto Exchanges and Credit Risk: Modeling and Forecasting the Probability of Closure. *J. Risk Financ. Manag.* **2021**, *14*, 516. [\[CrossRef\]](#)
48. Wu, J.; Zhou, Z.; Luo, Z. Research on Risk Prevention and Computer Modeling Optimization of Internet Financial Products. *J. Phys. Conf. Ser.* **2021**, 1744, 032108. [\[CrossRef\]](#)
49. Betz, F. In a Global Pandemic, Modeling the Central Bank Functions. *Mod. Econ.* **2021**, *12*, 623–658. [\[CrossRef\]](#)

-
50. Cebrián-Hernández, Á.; Jiménez-Rodríguez, E. Modeling of the bitcoin volatility through key financial environment variables: An application of conditional correlation MGARCH models. *Mathematics* **2021**, *9*, 267. [[CrossRef](#)]
 51. Scott, D.W. On optimal and data-based histograms. *Biometrika* **1979**, *66*, 605–610. [[CrossRef](#)]
 52. Freedman, D.; Diaconis, P. On the histogram as a density estimator: L 2 theory. *Probab. Theory Relat. Fields* **1981**, *57*, 453–476.