International Review of Economics and Finance 79 (2022) 387–407

Contents lists available at ScienceDirect

International Review of Economics and Finance

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

A cryptocurrency empirical study focused on evaluating their distribution functions

Carmen López-Martín [a,*], Raquel Arguedas-Sanz [a], Sonia Benito Muela [b]

[a] Department of Business and Accounting, Faculty of Economics and Business Administration, National Distance Education University (UNED), Senda Del Rey. 11, 28040, Madrid, Spain

[b] Department of Economic Analysis, Faculty of Economics and Business Administration, National Distance Education University (UNED), Senda Del Rey, 11, 28040, Madrid, Spain

ARTICLE INFO

JEL classification:
C46
C52
E17
G17

Keywords:
Cryptocurrencies
Distributions
Skewness
Fat tail
Risk management

ABSTRACT

This paper thoroughly examines the statistical properties of cryptocurrency returns, particularly focusing on studying which is the best statistical distribution for fitting this type of data. The preliminary statistical study reveals (i) high volatility, (ii) an inverse leverage effect, (iii) skewed distributions and (iv) high kurtosis. To capture the nonnormal characteristics observed in cryptocurrency data, we verified the goodness of fit of a large set of distributions, both symmetric and skewed distributions such as skewed Student-*t*, skewed generalized *t*, skewed generalized error and the inverse hyperbolic sign distributions. The results show that the skewed distributions outperform normal and Student-*t* distributions in fitting cryptocurrency data, although there is no one skewed distribution that systematically better fits the data. In addition, we compare these distributions in terms of their ability to forecast the market risk of cryptocurrencies. In line with the results obtained in the statistical analysis, we find that the skewed distributions provide better risk estimates than the normal and Student-*t* distributions, both in short and long positions, with SGED being the distribution that provides better results.

1. Introduction

Cryptocurrencies are one of the most disruptive financial innovations of the last decade (Feng et al., 2018). Katsiampa et al. (2018) define cryptocurrencies as “a digital asset designed to work as a medium of exchange using cryptography to secure the transactions without being subject to any government intervention”. The first and largest cryptocurrency in the world by market capitalization, Bitcoin, is considered the first large-scale implementation of blockchain technology. The Bitcoin blockchain was designed to allow fast and secure transactions and, at the same time, maintain the anonymity of users using a public record book that authenticates transactions between economic agents without the need for a central entity that proves the movement of funds. The original intention of blockchain development was not to create a new currency (Procházka, 2018) but to establish the principles of a functional decentralized cash payment system such as a peer-to-peer network for file sharing (Rosic, 2017).

Bitcoin has led the cryptocurrency industry since its inception in 2008 by a programmer (or group of programmers) under the pseudonym Satoshi Nakamoto (Nakamoto, 2008). Since then, many cryptocurrencies have been created, but it was not until Ethereum began in the summer of 2015 that a major step in the evolution of cryptocurrency took place with the integration of smart contracts

* Corresponding author.
E-mail addresses: carmen.lopez@cee.uned.es (C. López-Martín), rarguedas@cee.uned.es (R. Arguedas-Sanz), soniabm@cee.uned.es (S.B. Muela).

https://doi.org/10.1016/j.iref.2022.02.021
Received 20 June 2019; Received in revised form 30 April 2021; Accepted 12 February 2022
Available online 14 February 2022

onto the blockchain (Ray, 2019).

As of mid-November 2020, there were more than 7700 cryptocurrencies, with a total market capitalization of 538.765.720.190 US dollars (Bitcoin, Ethereum, Ripple and Litecoin represent an 80% market share), according to CoinMarketCap. This figure is similar to that reached at the end of 2017, a year in which the cryptocurrency market experienced greater growth. During some upper price peaks (for example, on January 4, 2018, when the market cap of the combined crypto markets was approximately $760 billion USD), the combined total market capitalization of cryptocurrencies was more than that of Google, which is the second largest company based on capitalization, and was comparable with the GDP of Switzerland, which is the 19th largest in the world according to the IMF (Haig, 2018).

The strong growth experienced by the price of cryptocurrencies since their creation has attracted the interest of many investors who demand these assets not so much for reasons of transaction but for reasons of investment. However, cryptocurrencies are primarily regarded as assets rather than currencies (Baek & Elbeck, 2015; Cheah & Fry, 2015; Dyhrberg, 2016).

The spectacular growth of cryptocurrencies' price since their introduction has attracted significant attention from the academic world. A large number of published papers focus on descriptive analysis of the Bitcoin network (see Zang & Lee, 2019 and Härdle et al., 2020, among others). Other studies analyse the efficiency of cryptocurrency markets, for instance, Lopez-Martin et al. (2021), Brauneis and Mestel (2018), Caporale et al. (2018), Tran and Leirvik (2019), Wei (2018) and Chaim and Laurini (2018, 2019). Although the results obtained in these papers are heterogeneous, they suggest that overall, the cryptocurrency markets are inefficient, although as in the case of Bitcoin, some authors argue that inefficiency tends to decrease (Bariviera et al., 2017; Köchling et al., 2019; Sensoy, 2019; Tran & Leirvik, 2019 and Vidal-Tomás et al., 2019).

Other papers study the determinants of the cryptocurrency price. On this topic, many studies conclude that cryptocurrencies' price depends neither on economic factors nor monetary factors but rather on speculative and supply and demand factors (see Bouoiyour & Selmi, 2015; Cheah & Fry, 2015; Ciaian et al., 2016; Eom et al., 2019). As stock markets are exposed to macroeconomic factors and government fiscal or monetary policies, the fact that cryptocurencies' prices may not depend on such factors opens the possibility for cryptocurrencies to be a source of diversification against the risk of stock markets. Along this line, some studies analyse the ability of cryptocurrencies *to diversify* assets and *hedge* against traditional asset risks (see Bouri, Jalkh, et al., 2017; Bouri, Molnár, et al., 2017; Brière et al., 2015; Dyhrberg, 2016; Eisl et al., 2015; Feng et al., 2018; García-Jorcano & Benito, 2020; Gkillas & Longin, 2019; Kang et al., 2019; Klein et al., 2018).[1] These papers describe a weak relationship between cryptocurrencies and stock markets, so they may act as a hedging asset against movements in the price of stocks.[2]

Other studies have focused on analysing the stylized facts of cryptocurrencies. For instance, Feng et al. (2018) point to the fact that this kind of data displays some characteristics of immature market assets, such as autocorrelated and nonstationary return series. Other papers indicate that cryptocurrency returns show (i) skewness and a degree of kurtosis (e.g., Zhang et al., 2018); (ii) high volatility (e.g., Chu et al., 2015; Chu et al., 2017; Katsiampa, 2017; Klein et al., 2018; Phillip et al., 2018); and (iii) a heavier tail than traditional currencies (Borri, 2019; Feng et al., 2018; Gkillas & Katsiampa, 2018; Osterrieder et al., 2017; Osterrieder & Lorenz, 2017; Phillip et al., 2018).

Our paper contributes to the existing literature by thoroughly examining the statistical properties of cryptocurrency returns, particularly which statistical distribution is the best for fitting this type of data. Regarding this last point, the literature is quite scarce. To the best of our knowledge, only the studies of Chan et al. (2017) and Osterrieder and Lorenz (2017) address this issue. Thus, in the present article, we try to fill this gap by checking the goodness of fit of a large set of distributions, both symmetric and skewed distributions. The distributions included in the comparison are the normal distribution, the Student-*t* distribution (symmetric), the skewed Student-*t* distribution (SSD) of Hansen (1994), the skewed generalized *t* distribution (SGT) of Theodossiou (1998), the skewed generalized error distribution (SGED) of Theodossiou (2001) and the inverse hyperbolic sign (IHS) of Johnson (1949).[3]

To study the dependence between two or more markets through copula analysis, it is necessary to previously model the marginal distribution of the analysed market returns. In this sense, the study carried out in this work is interesting not only for measuring market risk but also for studying the structural dependence between markets.

The results show that the skewed distributions outperformed the normal and Student-*t* distributions in fitting cryptocurrency data, although there is no skewed distribution that systematically better fits the data. Finally, we carry out an empirical application where we show that to estimate the market risk of cryptocurrencies, assuming asymmetric distributions improves the results with respect to the normal distribution and Student-*t* in both long and short positions, with the SGED providing better results.

The remainder of the paper is organized as follows: Section II describes the methodology. Section III presents the data and empirical results. In Section IV, we present an empirical application in portfolio management, and Section V discusses the conclusion.

2. Methodology

This paper evaluates the performance of several skewed and symmetric distributions in modelling the tail behaviour of daily cryptocurrency returns and in quantifying market risk. First, we fit several distributions, and then, we compare them in terms of their ability to measure market risk. In the next subsection, we show the distributions used for fitting cryptocurrency data and describe the measure used for quantifying market risk.

[1] Most of these applications have been done for Bitcoin.

[2] In this line, other authors study the dependence among several cryptocurrencies; see, for instance, Gkillas et al. (2018). In this study, the authors find a high dependence among cryptocurrencies, so they conclude that diversifying by investing in various cryptocurrencies may not be a good strategy.

[3] See Abad et al. (2014) for a detailed introduction on the skewed fat-tailed distributions.

2.1. Probability distribution

As previously mentioned, the empirical distribution of cryptocurrency returns has been documented to be asymmetric and exhibit a significant excess of kurtosis (fat tail and peakness). Therefore, assuming a normal distribution for risk management of these assets, particularly for estimating the Value at Risk (VaR), may not produce good results. Under this assumption, the size of the losses will be much higher than those predicted by a normal distribution.

The Student-*t* distribution can often account well for the excess kurtosis, but this distribution does not capture the skewness of the returns. Taking this fact into account, our interests lie in investigating the performance of the skewed distributions in fitting these data. In what follows, we describe the density functions of the skewed distributions used in this paper for fitting cryptocurrency returns.

2.1.1. Skewed student-t distribution (SSD) of Hansen

To model the skewness in the shape of the conditional return density, Hansen (1994) defined the SSD as:

$$f(z_t|v,\ \eta)=\begin{cases} bc\left[1+\dfrac{1}{\eta-2}\left(\dfrac{bz_t+\alpha}{1-\eta}\right)^2\right]^{-\frac{\eta+1}{2}} & if \quad z_t<-\left(\dfrac{a}{b}\right) \\ bc\left[1+\dfrac{1}{\eta-2}\left(\dfrac{bz_t+\alpha}{1+\eta}\right)^2\right]^{-\frac{\eta+1}{2}} & if \quad z_t\geq-\left(\dfrac{a}{b}\right) \end{cases} \tag{1}$$

The constants *a, b* and *c* are fixed as:

$$a=4\lambda c\left(\frac{\eta-2}{\eta-1}\right) \quad b^2=1+3\lambda^2-a^2 \quad c=\frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\sqrt{\pi\left(\eta-2\right)\Gamma\left(\frac{\eta}{2}\right)}}$$

where $z_t=\frac{r_t-\mu_t}{\sigma_t}$ are the standardized returns, $\Gamma(.)$ is the gamma function, λ is the skewness parameter, and η is a tail-thickness parameter. The parameters of the density function satisfy $|\lambda|<1$ and $\eta>2$.

2.1.2. Skewed generalized error distribution (SGED)

The SGED was proposed by Theodossiou (2001). The SGED probability density function is

$$f(z_t|v,\ \eta)f(z_t|\lambda,k)=\frac{C}{\sigma}\exp\left(-\frac{|z_t+\delta|^k}{(1+sign(z_t+\delta)\lambda)^k\theta^k}\right) \tag{2}$$

$$C=k\Big/\left(2\theta\Gamma\left(\frac{1}{k}\right)\right) \quad \delta=2\lambda AS(\lambda)^{-1} \quad \theta=\Gamma(1/k)^{0.5}\Gamma(3/k)^{-0.5}S(\lambda)^{-1} \quad \delta=2\lambda AS(\lambda)^{-1} \qquad S(\lambda)=\sqrt{1+3\lambda^2-4A^2\lambda^2}$$

$z_t=\frac{r_t-\mu_t}{\sigma_t}$ are the standardized returns, λ is a skewed parameter $|\lambda|<1$, k is the kurtosis parameter, and *sign* denotes the sign function.

2.1.3. Skewed generalized t distribution (SGT)

The SGT introduced by Theodossiou (1998) is a skewed extension of the generalized *t* distribution that was originally proposed by McDonald and Newey (1988). The SGT is a distribution that allows fitting for a very diverse level of skewness and kurtosis, and it has been used to model the unconditional distribution of daily returns for a variety of financial assets. Furthermore, SGT incorporates several well-known distributions, such as the generalized *t* distribution (McDonald & Newey, 1988), the SSD of Hansen (1994), the SGED of Theodossiou (2001), the normal distribution, the uniform distribution, the GED of Nelson (1991) and Student-*t* distribution. The SGT probability density function for the standardized residual is:

$$f(z_t|\lambda,\eta,k)=C\left(1+\frac{|z_t+\delta|^k}{\left(\frac{(\eta+1)}{k}\right)(1+sign(z_t+\delta)\lambda)^k\theta^k}\right)^{-\frac{\eta+1}{k}} \tag{3}$$

where

$$C=0,5k\left(\frac{\eta+1}{k}\right)^{-\frac{1}{k}}B\left(\frac{\eta}{k},\frac{1}{k}\right)^{-1}\theta^{-1}; \quad \theta=\frac{1}{\sqrt{g-\rho^2}};$$

$$\rho=2\lambda B\left(\frac{\eta}{k},\frac{1}{k}\right)^{-1}\left(\frac{\eta+1}{k}\right)^{\frac{1}{k}}B\left(\frac{\eta-1}{k},\frac{2}{k}\right)$$

$$g=\left(1+3\lambda^2\right)B\left(\frac{\eta}{k},\frac{1}{k}\right)^{-1}\left(\frac{\eta+1}{k}\right)^{\frac{1}{k}}B\left(\frac{\eta-1}{k},\frac{3}{k}\right);\quad \delta=\rho\theta$$

λ is the skewness parameter, $|\lambda|<1$; η is a tail-thickness parameter, $\eta>2$; k is a peakness parameter, $k>0$; *sign* is the sign function; B(.) is the beta function; δ is Pearson's skewness; and $z_t=\frac{r_t-\mu_t}{\sigma_t}$ is the standardized residual. The skewness parameter λ controls the rate of descent of the density around the mode of z_t. In the case of positive skewness ($\lambda>0$), the density function is skewed to the right. In contrast, the density function is skewed to the left with negative skewness ($\lambda<0$).

2.1.4. Inverse hyperbolic sign (IHS)

The IHS was proposed by Johnson (1949), and its density function is the following:

$$f(z_t|\lambda,k)=-\frac{k}{\sqrt{2\pi\left(\theta^2+(z_t+\delta)^2\right)}}\exp\left(-\frac{k^2}{2}\left(\ln\left((z_t+\delta)+\sqrt{\theta^2+(z_t+\delta)^2}\right)-(\lambda+\ln\theta)\right)^2\right) \tag{4}$$

where

$$\theta=\frac{1}{\sigma_w},\delta=\frac{\mu_w}{\sigma_w},\sigma_w=0.5\left(e^{2\lambda+k^{-2}}+e^{-2\lambda+k^{-2}}+2\right)^{0.5}\left(e^{k^{-2}}-1\right)$$

μ_w and σ_w are the mean and the standard deviation, respectively, of $w=\sinh(\lambda+x/k)$, sinh is the hyperbolic sine function, and x is a standard normal variable. Note that the negative value of k results in more leptokurtic distributions.

To analyse the goodness of fit of these distributions, we use several standard tests, such as the Kolmogorov-Smirnov test (KS) (Kolmogorov, 1933; Massey, 1951; Smirnov, 1939), and the Chi2 test of Pearson (1900). In addition, as the SGT distribution nets all the distributions considered in this paper (except IHS), we use the likelihood ratio test to evaluate which distribution provides the best fit to the data.

2.2. Measuring market risk

According to Jorion (2001), "Value at Risk (VaR) measure is defined as the worst expected loss over a given horizon under normal market conditions at a given level of confidence". Thus, VaR is a conditional quantile of the asset return loss distribution.

Let $X_1,X_2,\ldots,X_n$ be identically distributed independent random variables representing the financial returns. Using $F(x)$ to denote the cumulative distribution function, $F(x)=\Pr(X_t\leq x|\Omega_{t-1})$ conditioned on the information available at *t-1* (Ω_{t-1}). Assume that $\{X_t\}$ follows a stochastic process given by:

$$X_t=\mu_t+\sigma_t z_t \quad z_t\sim iid(0,1) \tag{5}$$

where μ_t is the conditional mean return; $\sigma_t^2=E\left(z_t^2|\Omega_{t-1}\right)$ and z_t has the conditional distribution function $G(z)$, $G(z)=P(z_t<z|\Omega_{t-1})$. The VaR with a given probability $\alpha\in(0,1)$, denoted by VaR(α), is defined as the α quantile of the probability distribution of financial returns:

$$F(VaR(\alpha))=\Pr(X_t\leq VaR(\alpha))=\alpha \tag{6}$$

To estimate this quantile, different methods have been developed: (i) nonparametric methods such as historical simulation methods; (ii) parametric methods and (iii) semiparametric methods such as filtered historical simulation (FHS), CaViar methods and methods based on extreme value theory (EVT). Among them, the parametric method is the most commonly used by financial institutions. In this method, the VaR of a portfolio is calculated as:

$$VaR_t(\alpha)=\mu_t+\sigma_t G^{-1}(\alpha) \tag{7}$$

where $G^{-1}(\alpha)$ is the percentile α of the distribution assumed for the innovations. In this study, we considered six types of distributions: (i) the normal distribution, (ii) the Student-*t* distribution, (ii) the skewed Student-*t* distribution (SSD) of Hansen, (iv) the skewed generalized error distribution (SGED) of Theodossiou (2001), (v) the skewed generalized *t* distribution (SGT) of Theodossiou (1998) and (vi) the inverse hyperbolic sine (HIS) of Johnson (1949).

To assess the performance in terms of VaR, we carry out a comparison using a two-stage selection approach. In the first stage, we use several accuracy tests, and in the second stage, and only for the remaining models, we calculate the loss function.

To check the accuracy of different VaR estimates, we use several standard tests, which are the most common procedures: the unconditional coverage (LR_{uc}) test (Kupiec, 1995), the conditional coverage (LR_{cc}) test, the independence test (LR_{ind}) of Christoffersen (1998), and the dynamic quantile (DQ) test of Engle and Manganelli (2004) (a detailed review of these tests can be found in Abad et al. (2014)).

To implement all these tests, the exception indicator (I_t) must be defined. If r_t represents the returns in time t and $VaR_t(\alpha)$ is the *VaR* obtained with a given probability $\alpha\in(0,1)$, we have an exception when $r_t<VaR_t(\alpha)$, and then, I_t is equal to one (zero otherwise).

The unconditional coverage test (LR_{uc}) has the null hypothesis $\hat{\alpha}=\alpha$, with a likelihood ratio statistic given by

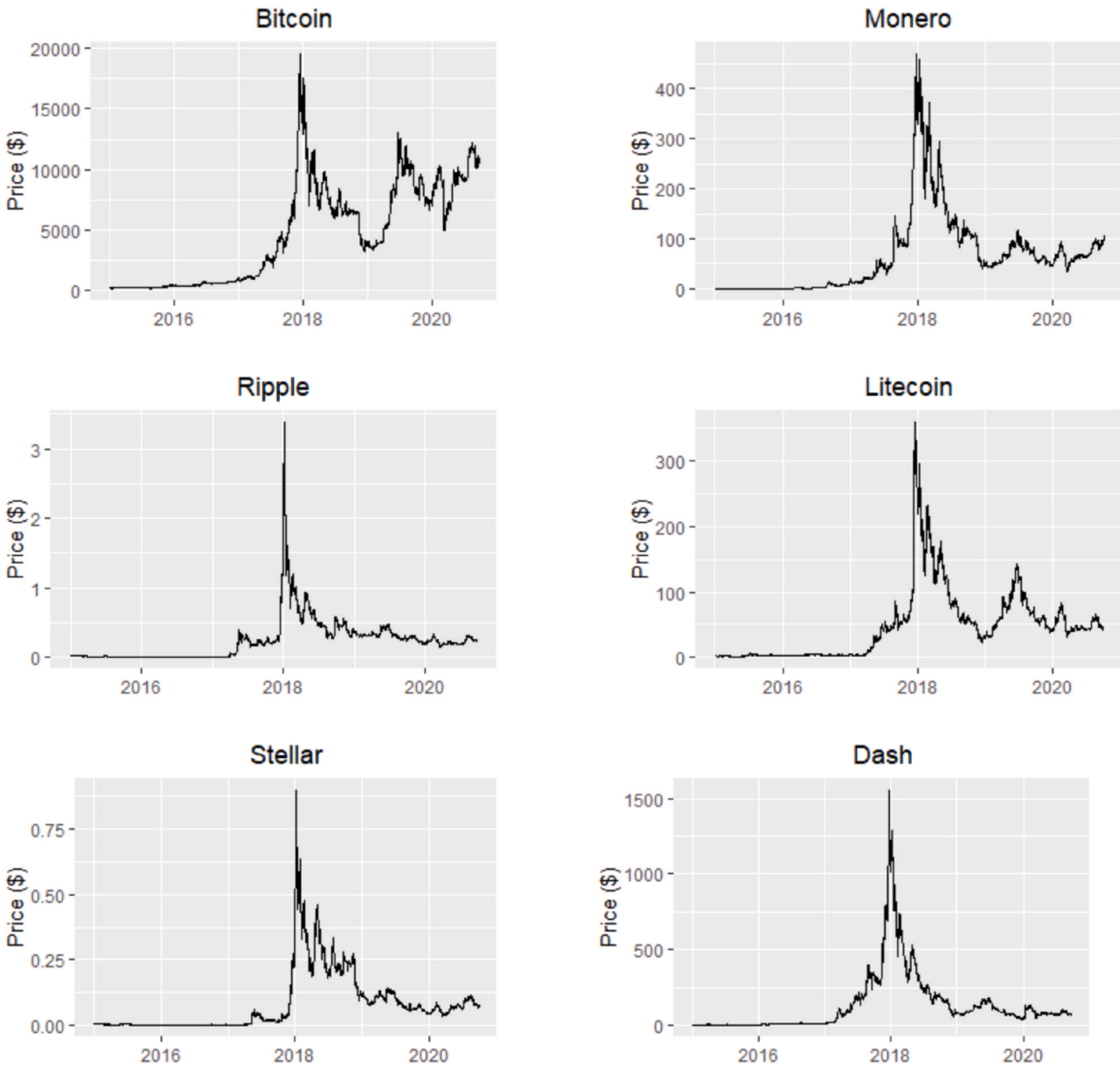


Fig. 1. Daily price of the cryptocurrencies.

$$LR_{uc} = 2log\frac{\widehat{\alpha}^x(1-\widehat{\alpha})^{N-x}}{\alpha^x(1-\alpha)^{N-x}} \sim \chi^2(1) \tag{8}$$

where $\widehat{\alpha}$ is the percentage of the exceptions; α is the expected percentage; and x is the number exception. The conditional coverage test (Christoffersen, 1998) jointly examines whether the percentage of exceptions is statistically equal to the expected percentage and the serial independence of I_t. The likelihood ratio statistic of the conditional coverage test is $LR_{cc} = LR_{uc} + LR_{ind}$, which is asymptotically $\chi^2(2)$-distributed and the LR_{ind} statistic is the likelihood ratio statistic for the hypothesis of serial independence against first-order Markov dependence.

Finally, the dynamic quantile test proposed by Engle and Manganelli (2004) examines whether the exception indicator is uncorrelated with any variable that belongs to the information set Ω_{t-1} available when the VaR is calculated. This is a Wald test of the hypothesis that all slopes are zero in a regression of the exception indicator variable on a constant, five lags and the VaR.

The backtesting procedures based on certain statistical tests present a disadvantage; they only show whether the VaR estimates are accurate, so this toolbox does not allow us to rank the models. Backtesting based on the loss function emphasizes the magnitude of the failure when an exception occurs. Lopez (1998, 1999), who is a pioneer in this area, proposed examining the distance between the observed returns and the forecasted VaR(α). This difference represents the loss that has not been covered. The loss function (LF)

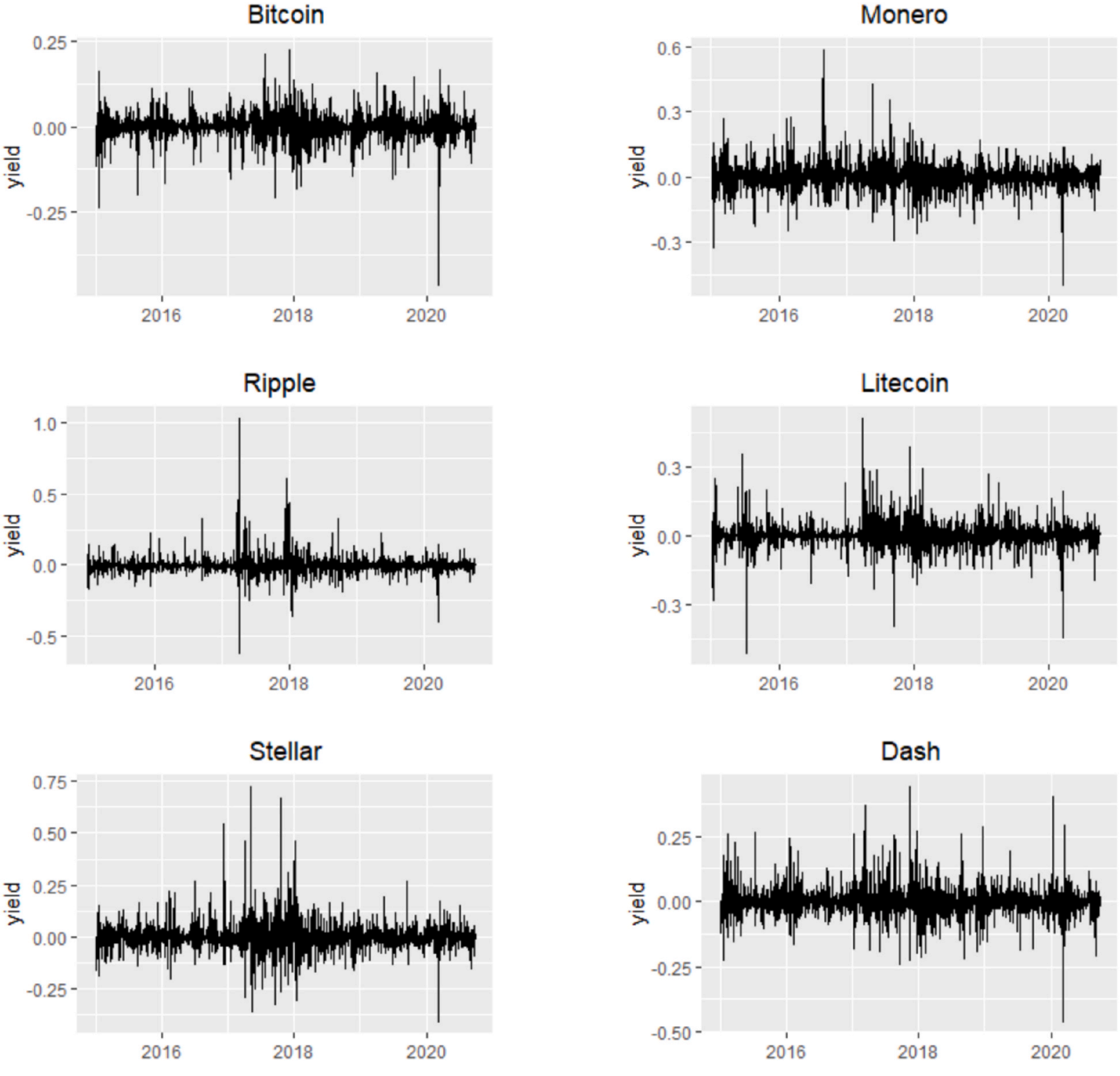


Fig. 2. Daily returns of the cryptocurrencies.

enables the financial manager to rank the models. The model that minimizes the total loss will be preferred to the other models. There are different formulations for loss functions (see Abad et al., 2014), and in this analysis, the linear loss function, which has the following specification, will be used:

$$\mathrm{LF} = \begin{cases} |VaR_t - r_t| & if \quad r_t < VaR_t \\ 0 & otherwise \end{cases} \tag{9}$$

3. Data and empirical results

3.1. Data

For our empirical analysis, we used data from six cryptocurrencies: Bitcoin Monero, Ripple, Litecoin, Dash and Stellar. These cryptocurrencies account for just over 67% of the market capitalization according to the data of CoinMarketCap.[4] All currencies are

[4] https://coinmarketcap.com/.

Table 1
Daily correlation, January 1, 2015 to September 30, 2020.

	Bitcoin	Monero	Ripple	Litecoin	Stellar	Dash
Bitcoin	1.00	0.57	0.39	0.64	0.42	0.54
Monero	0.57	1.00	0.35	0.47	0.41	0.50
Ripple	0.39	0.35	1.00	0.41	0.58	0.32
Litecoin	0.64	0.47	0.41	1.00	0.42	0.47
Stellar	0.42	0.41	0.58	0.42	1.00	0.35
Dash	0.54	0.50	0.32	0.47	0.35	1.00

Table 2
Daily log return statistics, January 1, 2015 to September 30, 2020.

	Mean	Median	Std. Dev.	Skewness	Kurtosis	Max.	Min.	Jarque-Bera
Bitcoin	0.002	0.002	0.039	−0.996	13.644	0.225	−0.465	16665*
Monero	0.003	0.000	0.064	0.589	8.979	0.585	−0.494	7190*
Ripple	0.001	−0.003	0.064	2.915	47.067	1.027	−0.616	197117*
Litecoin	0.001	0.000	0.057	0.400	14.203	0.510	−0.514	17739*
Stellar	0.001	−0.002	0.071	1.956	19.195	0.723	−0.410	33636*
Dash	0.002	−0.001	0.058	0.650	8.259	0.438	−0.459	6128*

The Jarque-Bera statistic is distributed as the Chi2 with two degrees of freedom. (*) denotes significance at the 1% level.

expressed in terms of the US dollar. The data consist of the closing daily price extracted from CoinMarketCap. These prices are transformed into returns by taking logarithmic differences. The complete sample nearly covers six years. The data period runs from January 01, 2015, to September 30, 2020.

Fig. 1 illustrates the price evolution of all currencies considered. In these plots, we can clearly observe the strong growth that the cryptocurrency price experienced during the first part of 2017 as well as the collapse of these currencies in the following months. Among the arguments that explain the fall of the market in 2018 is the lack of regulation and the frequent and allegations of fraud attacks on cryptocurrency platforms.[5] Fig. 1 also shows the effect of the health crisis caused by COVID-19 in March 2020. This crisis caused all cryptocurrencies to crash, producing losses of more than 25%. This fact can be observed clearly in Fig. 2, where we report the evolution returns.

Despite the health crisis and the lasting subsequent economic downturn, the value of some cryptocurrencies, such as Bitcoin, surprisingly recovered prior to March 2020. Finally, we observe that the range of fluctuation of Bitcoin is lower than that of the rest of the cryptocurrencies (see Fig. 2).

Table 1 provides the correlations of the daily returns for all currencies considered. Overall, the correlation detected between cryptocurrencies is very high, reducing the possibility of obtaining benefits derived from their diversification.

Further insights into the distributional properties of cryptocurrencies can be obtained by studying the basic descriptive statistics of the daily returns shown in Table 2. For all currencies considered, the mean return is positive, ranging from 0.1% to 0.3%. The standard deviation of the cryptocurrencies is also different, indicating that the cryptocurrencies do not show the same level of risk. Bitcoin, for example, is the safest currency, followed by Litecoin and Dash. Stellar has the largest dispersion of all cryptocurrencies.

The skewness statistic is positive for all currencies except for Bitcoin, which is negative. In all cases, the excess kurtosis statistic is very large, implying that the distributions of those returns have much thicker tails than the normal distribution. According to the Jarque-Bera test, we reject the hypothesis of normality in all cases.

The QQ plots displayed in Fig. 3 corroborate this result. According to these charts, we can say that the normal distribution does not fit well, neither in the centre nor in the tails of the probability distribution of the cryptocurrencies.

To study the autocorrelation of the daily returns, we use the Lung-Box test for different lags (see Table 3). This test is applied over the daily return and the squared return. Only in the case of Dash returns can the hypothesis of no autocorrelation not be rejected for any lags. This result is consistent with the efficient market hypothesis. For Bitcoin and Litecoin returns, the no autocorrelation hypothesis cannot be rejected for (1) and (5) lags, indicating some degree of efficiency. The rest of the cryptocurrencies, namely, Monero, Ripple and Stellar, seem to be negotiated in inefficient markets.

Table 3 also reports the Lung-Box test for the squared return. The hypothesis of no autocorrelation is rejected in all cases, a fact consistent with ARCH effects. This result justifies the appropriateness of using a GARCH framework to model conditional volatility.

This preliminary statistical study reveals the following facts about cryptocurrency returns: high volatility, skewed distributions, high kurtosis and positive autocorrelation for square returns. Concerning the autocorrelation of the returns, the results vary depending on the cryptocurrency analysed. Such stylized facts are similar to those obtained by other authors, such as Gangwal and Longin (2018) and Lopez-Martin et al. (2021).

[5] At the end of 2017 and the beginning of 2018, the European Union and US regulators warned of the risks involved in operating with these digital currencies as well as the fraudulent use that can be given to them, to launder money or finance terrorist activities, which resulted in an impact on the quote value of Bitcoin and other cryptocurrencies, although cryptocurrencies often operate outside the scope of national regulations (Auer & Claessens, 2018; Brühl, 2017).

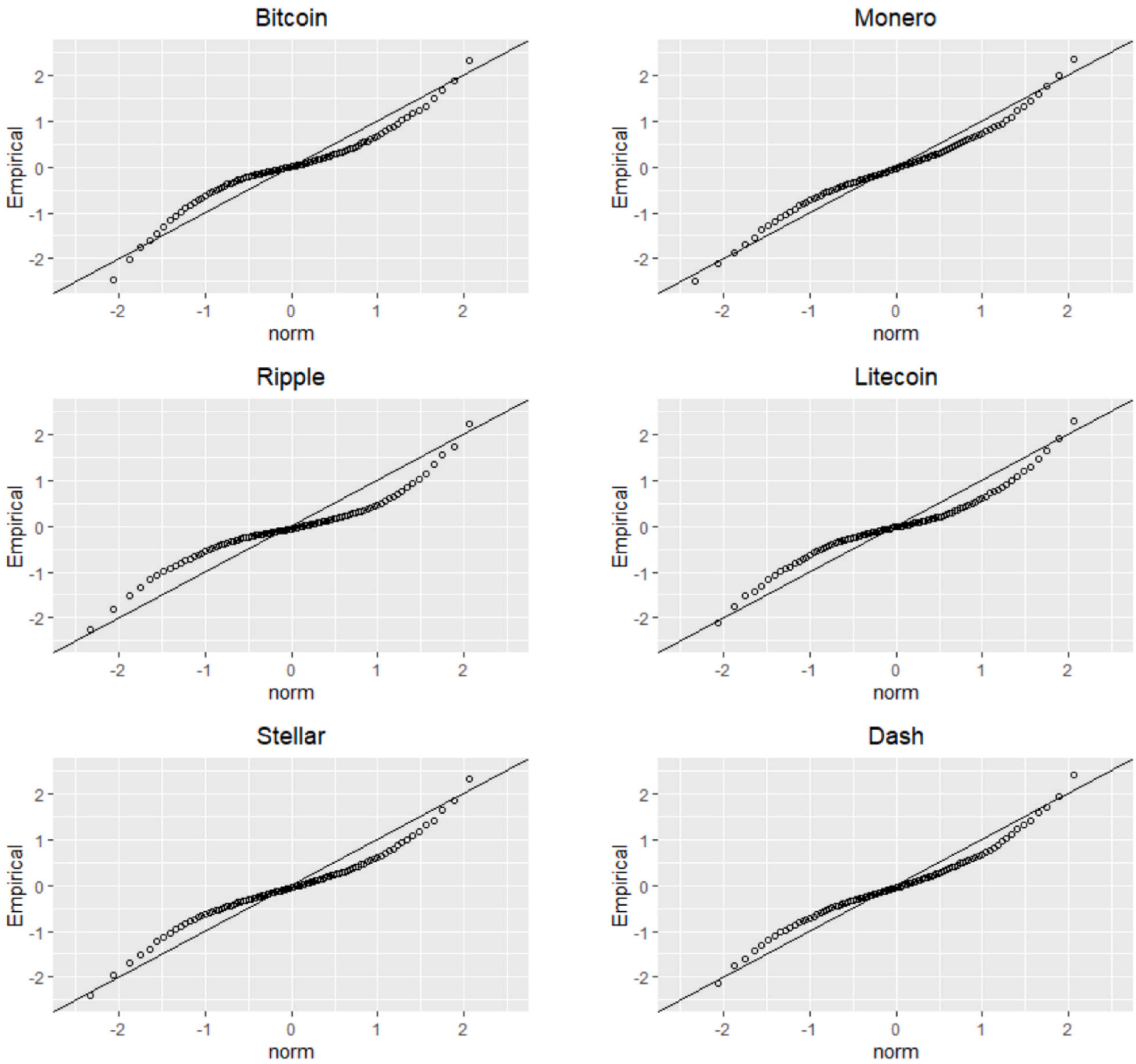


Fig. 3. Empirical percentile of the standardized returns against the percentile of the standard normal distribution.

3.2. Empirical results

3.2.1. Modelling conditional variance

In this section, we estimate the conditional standard deviation of the cryptocurrencies. To estimate the volatility of these currencies, we use two GARCH family models. The first model considered is GARCH (1,1), whose expression is as follows:

$$\sigma_t^2 \ = \ \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{13}$$

where $\omega,\ \alpha,\ \beta\ \geq 0$. The GARCH model captures some of the characteristics observed in financial returns such *volatility clustering* (Mandelbrot, 1963*)*. However, this model does not take the *leverage effect* (Black, 1976) into account. This effect is related to the possibility that volatility responds asymmetrically to surprises of different signs. To study this effect, we also estimate the EGARCH(1, 1) model, whose expression is as follows:

$$log\ \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \theta\left[\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| - E(|\varepsilon_t|)\right] + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \tag{14}$$

In this model, the parameter γ captures the leverage effect. If $\gamma > 0$, the positive surprises have more impact on volatility than

Table 3
Autocorrelation test for daily returns and squared returns.

	Returns			Squared Returns		
	Lung-Box (1)	Lung-Box (5)	Lung-Box (10)	Lung-Box (5)	Lung-Box (10)	Lung-Box (15)
Bitcoin	1.28 (0.25)	2.05 (0.84)	19.66 (0.03)	36.94 (0.00)	58.35 (0.00)	87.46 (0.00)
Monero	4.12 (0.04)	17.15 (0.00)	43.81 (0.00)	45.41 (0.00)	101.32 (0.00)	206.18 (0.00)
Ripple	0.34 (0.55)	34.46 (0.00)	55.84 (0.00)	175.53 (0.00)	269.01 (0.00)	307.13 (0.00)
Litecoin	0.18 (0.67)	6.92 (0.22)	28.79 (0.00)	36.67 (0.00)	79.78 (0.00)	109.08 (0.00)
Stellar	5.43 (0.02)	14.60 (0.01)	27.63 (0.00)	333.64 (0.00)	543.98 (0.00)	555.66 (0.00)
Dash	0.46 (0.49)	9.67 (0.08)	16.03 (0.10)	52.43 (0.00)	83.38 (0.00)	126.39 (0.00)

Note: In parentheses, we show the p-value of the Lung-Box test.

Table 4
Parameter estimations of the volatility models.

		ω	α	β	γ	v	Log-L	AIC	ARCH (5)	ARCH (7)	Sum Square Error
Bitcoin	**E**	−0.06	0.05*	0.98**	0.31**	2.94**	**4432**	**−4.215**	0.54	0.74	**0.091**
	G	0.00**	0.02*	0.86**		2.41**	4416	−4.203	0.99	9.99	0.113
Monero	**E**	−0.30**	0.06**	0.94**	0.31**	3.54**	**3171**	**−3.015**	0.25	0.37	**0.386**
	G	0.00**	0.01**	0.79**		3.52**	3166	−3.012	0.99	9.99	0.400
Ripple	E	−0.59**	−0.01	0.88**	0.75**	2.28**	3804	−3.620	0.95	0.98	9.306
	G	0.00	0.26*	0.67**		2.28**	**3805**	**−3.620**	0.99	9.99	**4.255**
Litecoin	**E**	−0.06**	0.02	0.98**	0.32**	2.34**	**3786**	**−3.601**	0.94	0.98	**0.555**
	G	0.00**	0.03*	0.87**		2.24**	3772	−3.590	0.99	1.00	0.766
Stellar	E	−0.39**	0.04	0.92**	0.39**	3.08**	3243	−3.085	0.92	0.95	1.182
	G	0.00**	0.02**	0.73**		3.04**	**3244**	**−3.086**	0.99	9.99	**1.074**
Dash	**E**	−0.35**	−0.01	0.93**	0.35**	3.19**	**3459**	**−3.283**	0.95	0.96	**0.240**
	G	0.00**	0.22**	0.76**		3.14**	3445	−3.278	0.99	9.99	0.254

Note: (E) EGARCH model; (G) GARCH model; γ parameter: captures the leverage effect; v: degrees of freedom; Log-L: log likelihood; AIC: Akaike information; Sum of squared errors: $\sum_{i=1}^{T}\left(\varepsilon_t^2 - \hat{\sigma}_t^2\right)^2$ where T is the size of the sample; ε_t are the innovations of the returns; and $\hat{\sigma}_t^2$ is the estimation of the conditional variance. ARCH(p) is the test for autoregressive conditional heteroskedasticity by Engle (1982) at the p lag. For each currency, we remark in bold the highest Log-L, the lowest AIC and the minimum sum of squared errors.

negative surprises; the reverse is true for a $\gamma < 0$. Overall, for stock returns, the γ parameter is negative, indicating that negative surprises have a higher impact on volatility than positive surprises. However, in the case of fixed income returns, the γ parameter is usually positive, as in this case, a negative surprise $(\varepsilon_{t-1} < 0)$ is good news for investors.

According to the evidence shown in the previous section, the volatility of the cryptocurrencies has been estimated assuming a Student-*t* distribution that has a heavier tail than the normal distribution.

Table 4 reports the estimation of parameters for two considered models of volatility and the test ARCH for autoregressive conditional heteroskedasticity at 5 and 7 lags. According to the ARCH test, both models adequately capture the volatility of the data. On the other hand, the parameters of the GARCH model are all positive and statistically significant. The persistence degree, given by $\alpha + \beta$, ranges depending on the cryptocurrency between 0.75 (Stellar) and 0.98 (Dash). For the EGARCH model, almost all parameters are statistically significant, including the γ parameter, which captures the leverage effect. Note also that the sign of this parameter is positive, indicating the existence of an *inverse leverage effect*. This means that in the case of cryptocurrencies, volatility tends to be higher after positive shocks.

The empirical literature dedicated to analysing the *leverage effect* in cryptocurrencies is very limited, and most papers are related to Bitcoin. Many controversies arise regarding this currency. Some authors find an *inverse leverage effect* (see Bouri, Azzi, & Dyhrberg, 2017; García-Jorcano & Benito, 2020; Klein et al., 2018), while other studies do not detect a significant *leverage effect* (Dyhrberg, 2016; Katsiampa, 2017; Takaishi, 2018; Tiwari et al. 2019).

In the case of Bitcoin, García-Jorcano and Benito (2020) argue that the *inverse leverage effect* could be explained by the fact that, in this period, Bitcoin behaved as a safe-haven asset, whose demand increases when uncertainty increases in traditional markets. If investors fleeing the risk of traditional markets seek Bitcoin as a safe-haven asset, then the price of Bitcoin will grow, thus transferring the volatility of the equity markets to the Bitcoin market. This argument was initially given by Baur (2012) to justify that gold volatility increases with

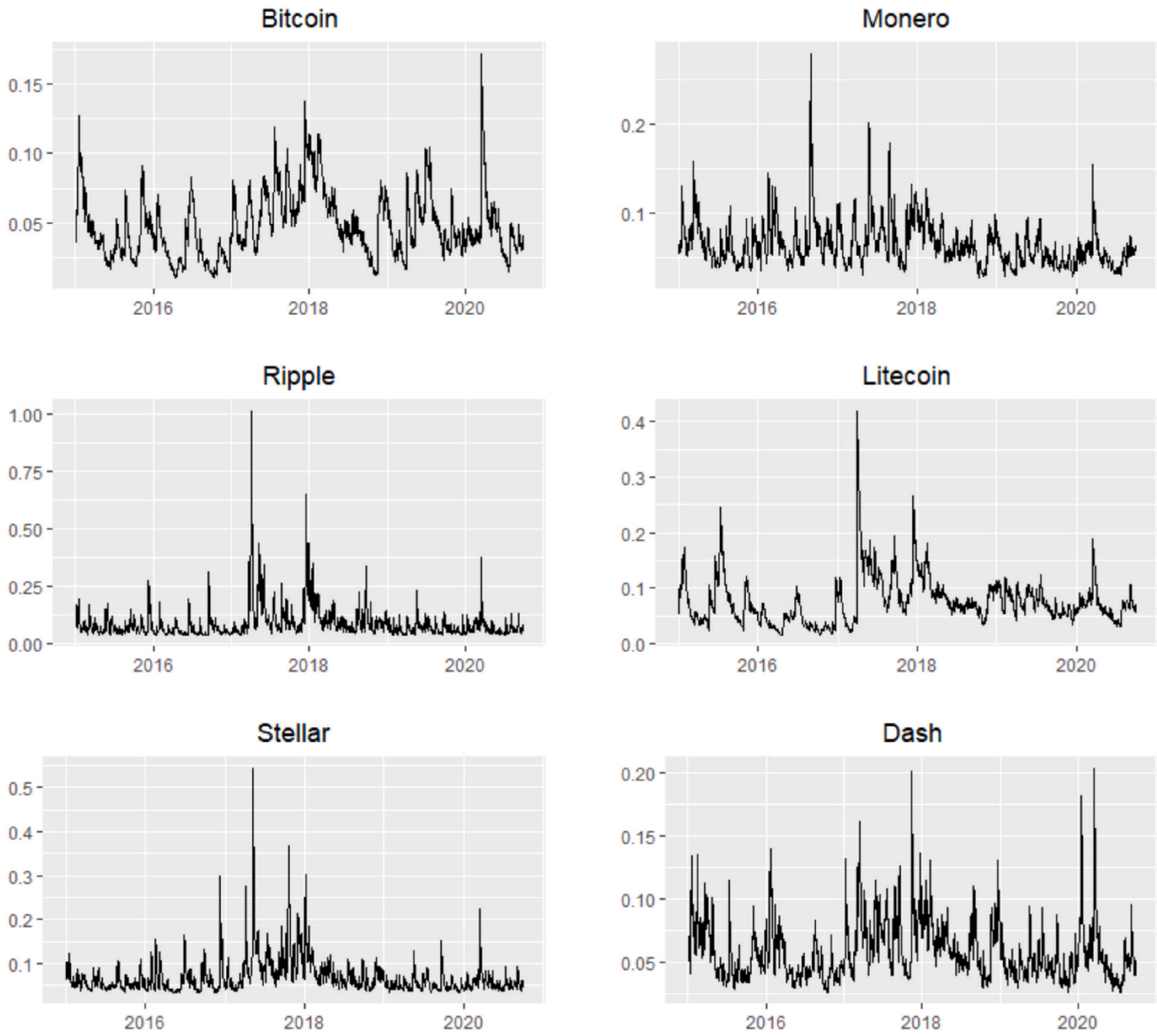


Fig. 4. Conditional standard deviation. For the estimation of the conditional standard deviation, we use an EGARCH model for Bitcoin, Monero, Litecoin and Dash and a GARCH model for Ripple and Stellar.

positive surprises. These authors use this same argument to justify what is observed in the Bitcoin market. The idea is not farfetched since in many articles, the role of Bitcoin as a risk-hedging asset has been proposed, playing a role similar to that played by gold.

Table 4 also reports the log likelihood (Log-L) and the Akaike information criterion (AIC) for both volatility models.[6] According to these criteria, the model that provides the best fit is the EGARCH model. Only in the case of Ripple and Stellar does the GARCH model appear to provide a better fitting.

Finally, to determine which model generates the best volatility forecasts, we compare ε_t^2 with the volatility obtained from both models $(\widehat{\sigma}_t^2)$. The sum of the squared errors is reported in the last column of Table 4. For all cryptocurrencies, except Ripple and Stellar, the EGARCH model provides the best forecasts.

According to the aforementioned results, we use an EGARCH(1,1) model to model the volatility of Bitcoin, Monero, Litecoin and Dash. For Ripple and Stellar, we use a GARCH(1,1) model.

Fig. 4 illustrates the conditional standard deviation of all currencies considered. The first that attracts our attention is the low volatility of Bitcoin compared to other currencies. Second, joint volatility peaks are also observed, which is consistent with the high correlation detected (Table 1). Finally, it should be noted that around March 2020, all markets exhibited a significant increase in volatility, coinciding with the global health crisis.

[6] Akaike (1974).

Table 5
Maximum likelihood estimates.

		μ	σ	λ	η	κ
Bitcoin	SGT	0.002 (0.001)*	0.039 (0.002)*	0.003 (0.011)	18.112 (8.763)*	0.774 (0.025)*
	SGED	0.002 (0.000)*	0.039 (0.001)*	0.000 (0.040)	2.006 (0.000)*	0.711 (0.022)*
	SSD	0.002 (0.001)*	0.353 (0.009)*	0.002 (0.014)	2.002 (0.000)*	0.827 (0.013)*
	IHS	0.002 (0.000)*	0.048 (0.001)*	−0.001 (0.021)		
	Student-*t*	0.002 (0.000)*	0.629 (0.016)*			
	Normal	0.002 (0.001)*	0.039 (0.001)*			
Monero	SGT	0.003 (0.001)*	0.065 (0.002)*	0.060 (0.021)*	4.287 (0.266)*	1.468 (0.057)*
	SGED	0.003 (0.000)*	0.062 (0.001)*	0.083 (0.002)*		0.946 (0.032)*
	SSD	0.003 (0.001)*	0.072 (0.002)*	0.051 (0.021)*	2.831 (0.065)*	
	IHS	0.003 (0.001)*	0.066 (0.001)*	0.078 (0.031)*		1.104 (0.030)*
	Student-*t*	0.001 (0.001) *	0.072 (0.002)*		2.835 (0.066)*	
	Normal	0.003 (0.001)*	0.064 (0.001)*			
Ripple	SGT	0.000 (0.001)	0.068 (0.002)*	0.065 (0.010) *	2.861 (0.067)*	1.134 (0.034)*
	SGED	0.000 (0.001)	0.056 (0.003)*	0.068 (0.014)*	2.015 (0.001)*	0.642 (0.023)*
	SSD	−0.001 (0.001)	0.284 (0.007)*	0.050 (0.015)*	2.001 (0.105)*	0.776 (0.010)*
	IHS	0.000 (0.001)	0.070 (0.002)*	0.093 (0.020)*		
	Student-*t*	−0.002 (0.001) *	1.393 (0.036)*			
	Normal	0.001 (0.001)	0.064 (0.001)*			
Litecoin	SGT	0.000 (0.001)	0.058 (0.002)*	0.003 (0.012)	6.222 (1.005)*	0.855 (0.027)*
	SGED	0.000 (0.007)	0.056 (0.006)*	0.003 (0.014)	2.041 (0.002)*	0.659 (0.023)*
	SSD	- 0.001 (0.001)	0.180 (0.005)*	−0.005 (0.014)	2.001 (0.001)*	0.801 (0.012)*
	IHS	0.001 (0.001)	0.068 (0.002)*	0.060 (0.020) *		
	Student-*t*	0.000 (0.001)	1.146 (0.029)*			
	Normal	0.001 (0.001)	0.057 (0.001)*			
Dash	SGT	0.002 (0.002)	0.059 (0.002)*	0.079 (0.018)*	4.046 (0.221)*	1.337 (0.049)*
	SGED	0.002 (0.002)	0.056 (0.002)*	0.092 (0.015)*	2.467 (0.030)*	0.863 (0.061)*
	SSD	0.002 (0.001)*	0.074 (0.002)*	0.075 (0.019)*	2.460 (0.030)*	1.009 (0.023)*
	IHS	0.002 (0.001)*	0.061 (0.001)*	0.121 (0.028)*		
	Student-*t*	- 0.001 (0.001)	0.074 (0.002)*			
	Normal	0.002 (0.001) *	0.058 (0.001)*			
Stellar	SGT	0.000 (0.001)	0.075 (0.002)*	0.070 (0.017)*	2.968 (0.080)*	1.489 (0.054)*
	SGED	0.000 (0.000)	0.065 (0.003)*	0.030 (0.007)*	2.234 (0.013)*	0.798 (0.034)*
	SSD	0.000 (0.001)	0.107 (0.003)*	0.069 (0.018)*	2.460 (0.030)*	0.946 (0.018)*
	IHS	0.001 (0.001)	0.072 (0.002)*	0.118 (0.027)*		
	Student-*t*	−0.001 (0.001)	0.074 (0.002)*			
	Normal	0.002 (0.001)*	0.058 (0.001)*			

Note: Parameter estimates of the SGT, SGED, SSD, HIS, Student-*t* and normal distributions. In brackets are standard errors. The dataset covers January 1, 2015 to September 30, 2020. μ, σ, λ and η are the estimated mean, standard deviation, skewness parameter, and tail-thickness parameter, respectively, and κ represents the peakness parameter. In the case of Student-*t*, η represents the degrees of freedom. An (*) denotes significance at the 5% level.

3.2.2. *Fitting distributions*

To capture the nonnormal characteristics observed in cryptocurrency data, we fit five distributions: Student-*t* distribution, which is symmetric, and four skewed distributions, namely, SGT, SGED, SSD and HIS. In addition, we include the normal distribution for comparative reasons.

The parameters of the distributions were estimated by the maximum likelihood method. Table 5 presents these estimates and the standard errors in brackets. As we expected, the unconditional mean is close to zero for all cryptocurrencies. Only for Bitcoin and Monero is the mean statistically significant. The standard deviation for all currencies is statistically significant. These estimates do not differ much from the data displayed in Table 2.

For all cryptocurrencies except Bitcoin and Litecoin, the skewness parameter is positive and statistically significant at the 5% level, which means that the distributions of these returns are skewed to the right. On the other hand, parameters κ and η, which control the peakness of the distribution around the mode and the tails of the distribution, respectively, are significant at the 5% level for all currencies considered. These results corroborate that the cryptocurrencies do not follow a normal distribution, display a higher degree of peakness and show heavier tails than the normal distribution. This result can be observed clearly in Fig. 5, where we display the histogram of the Bitcoin with the superimposed normal density function, the Student-*t* density function and the density functions of the SGT, SGED, SSD and IHS. We observe that the normal density function does not fit the histogram of Bitcoin well, neither in the centre nor in the tail. Only for very high quantiles does the fitting appear good. Unlike the normal distribution, the Student-*t* and skewed distributions appear to provide a good fit. The same is observed for the rest of the cryptocurrencies considered (see the Appendix).

Regarding Fig. 5 and the figures reported in the Appendix, we observe that the goodness of fit in the left tail provided by the skewed distribution is similar to that provided by Student-*t*, especially in high quantiles. However, in the right tail, we observe some differences. In this case, the density of the skewed distributions is above that of Student-*t*, although for very high quantiles, these functions

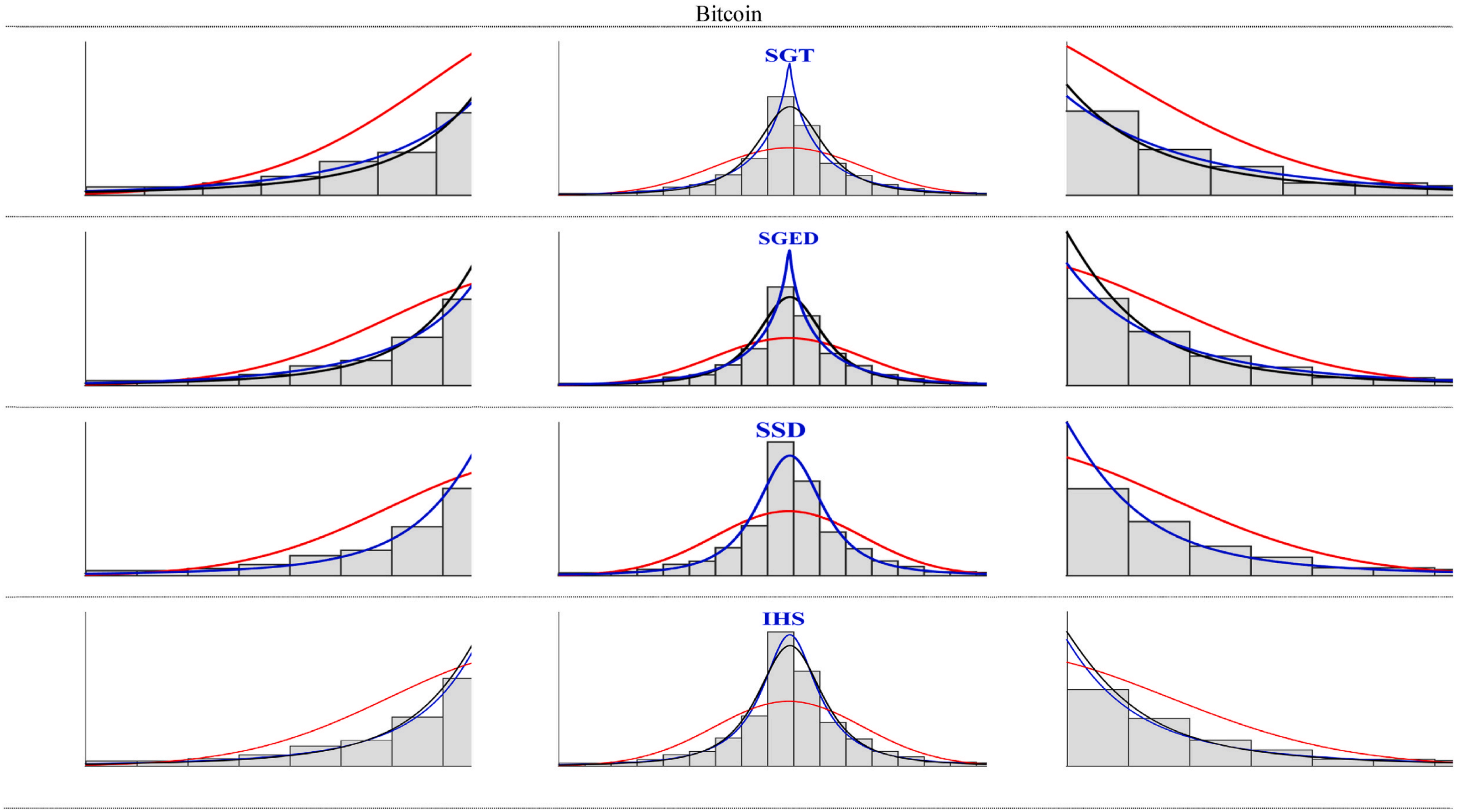


Fig. 5. Histogram of the Bitcoin with superimposed normal (red) and Student-*t* (black) distributions and skewed distributions (blue): SGT, SGED, SSD and IHS. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 6
Goodness of fit test.

	Log-L	LR_{NORMAL}	LR_{SGT}	AIC	Chi2	KS
Bitcoin						
SGT	4249.78	877.02**	–	−8489.6	0.03	0.19
SGED	4249.40	876.26**	0.76	**−8490.8**	0.05	0.19
SSD	4208.95	795.36**	81.66**	−8409.9	0.00	0.01
IHS	4227.87	–	–	−8447.7	0.00	0.07
Student	4208.98	–	81.60**	−8412.0	0.00	0.01
Normal	3811.27		–	−7618.5	0.00	0.00
Monero						
SGT	3047.31	524.20**	–	−6084.6	0.33	0.63
SGED	3035.92	501.42**	22.78**	−6063.8	0.19	0.07
SSD	3042.89	515.36**	8.84	−6077.8	0.01	0.55
IHS	3047.02	–	–	**−6086.0**	0.06	0.71
Student	3041.20	–	12.22**	−6076.4	0.02	0.43
Normal	2785.21		–	−5566.4	0.00	0.00
Ripple						
SGT	3634.5	1675.46**	–	**−7258.9**	0.10	0.57
SGED	3615.7	1637.96**	37.50**	−7223.4	0.00	0.13
SSD	3615.2	1636.88**	38.58**	−7222.4	0.01	0.02
IHS	3631.9	–	–	−7255.7	0.02	0.50
Student	3613.9	–	41.08**	−7221.9	0.00	0.01
Normal	2796.7		–	−5589.5	0.00	0.00
Litecoin						
SGT	3582.1	1108.20**	–	−7154.2	0.04	0.02
SGED	3581.2	1106.44**	1.76	**−7154.4**	0.02	0.01
SSD	3546.7	1037.44**	70.76**	−7085.4	0.01	0.01
IHS	3566.9	–	–	−7125.7	0.03	0.16
Student	3547.6	–	68.92**	−7089.2	0.03	0.02
Normal	3028.0		–	−6052.0	0.00	0.00
Dash						
SGT	3332.6	663.64**	–	**−6655.2**	0.26	0.98
SGED	3322.6	643.60**	20.04**	−6637.1	0.00	0.33
SSD	3325.7	649.94**	13.70**	−6643.5	0.09	0.42
IHS	3331.2	–	–	−6654.4	0.72	0.66
Student	3321.9	–	21.40**	−6637.8	0.05	0.37
Normal	3000.8		–	−5997.5	0.00	0.00
Stellar						
SGT	3074.7	996.38**	–	−6139.3	0.26	0.82
SGED	3049.3	945.57**	50.80**	−6090.5	0.00	0.01
SSD	3070.9	988.92**	7.46	−6133.9	0.09	0.74
IHS	3075.7	–	–	**−6143.5**	0.72	0.92
Student	3067.7	–	13.84**	−6129.5	1.00	0.60
Normal	2576.5		–	−5148.9	0.00	0.00

Note: Log-L is the maximum likelihood value. LR_{Normal} is the LR statistic from testing the null hypothesis that the daily returns are distributed as normal against SGT, SGED or SSD. LR_{SGT} is the LR statistic from testing the null hypothesis against the alternative distribution, the SGT. Columns 6 and 7 show the p-value of the Chi2 test (column 6) and p-value of the Kolmogorov-Smirnov test (column 7). AIC is the Akaike Information Criterion. An *(**) denotes significance at the 5% (1%) level.

tend to converge. This result may have implications in measuring market risk associated with long and short positions. It seems clear that in the case of short positions, the use of the skewed distribution is more appropriate.

To formally test the absence of normality, we used the Kolmogorov-Smirnov (KS) test (Kolmogorov, 1933; Massey, 1951; Smirnov, 1939) and Chi2 test. These tests analyse whether the asset return empirical distribution follows a theoretical distribution. For all considered currencies, the normality hypothesis is always rejected (see Table 6). In addition, as the normal distribution is nested within the SGT, SGED and SSD distributions, we use the log-likelihood ratio to test the null hypothesis of normality against that of SGT, SGED or SSD. Again, for all the cryptocurrencies considered, this statistic is quite large and statistically significant at the 1% level, providing evidence against the normality hypothesis (see Table 6).[7]

At this point, we are interested in determining the best distribution for fitting the currency return empirical distribution. As the SGT

[7] In this case, the null hypothesis stablishes that the returns follow a normal distribution. If this hypothesis is rejected, we find evidence in favour of skewed distributions such as the SGT, SGED and SSD.

Table 7
Accuracy tests for VaR estimations at 1% probability.

	Bitcoin	Monero	Ripple	Litecoin	Dash	Stellar
Long position						
SGT						
% exceptions	0.94	1.25	1.25	1.10	1.56	1.41
Number of tests that pass H_0	3	4	4	3	4	3
SGED						
% exceptions	0.31	1.41	1.25	0.00	1.56	1.25
Number of tests that pass H_0	4	4	4	0	4	3
SSD						
% exceptions	0.94	1.41	1.56	0.63	1.88	1.56
Number of tests that pass H_0	3	4	4	4	4	4
HIS						
% exceptions	0.47	0.78	1.25	0.47	1.41	1.41
Number of tests that pass H_0	4	3	4	4	4	3
Student-*t*						
% exceptions	1.10	0.78	1.25	0.94	1.41	1.41
Number of tests that pass H_0	3	3	4	3	4	3
Normal						
% exceptions	0.63	1.41	1.56	0.31	2.03	1.56
Number of tests that pass H_0	4	4	4	4	4	4
Short position						
SGT						
% exceptions	1.25	0.31	1.09	0.94	0.94	0.63
Number of tests that pass H_0	4	4	3	4	3	3
SGED						
% exceptions	0.63	1.09	1.09	0.31	0.78	0.63
Number of tests that pass H_0	4	4	3	4	3	3
SSD						
% exceptions	1.09	0.16	1.09	0.94	0.78	0.63
Number of tests that pass H_0	4	4	3	4	3	3
HIS						
% exceptions	0.63	0.16	1.09	0.31	0.78	0.63
Number of tests that pass H_0	4	4	3	4	3	3
Student-*t*						
% exceptions	1.25	0.78	1.71	0.94	1.25	0.78
Number of tests that pass H_0	4	4	3	4	4	3
Normal						
% exceptions	1.10	0.93	2.02	0.78	1.41	1.25
Number of tests that pass H_0	4	4	4	4	4	4

Note: The maximum losses at 1% probability in a long and short positions are given by the VaR(1%) and VaR(99%), respectively.

nets all the distributions considered in this paper (except IHS), we use the likelihood ratio test to evaluate which distribution is best in fitting the data. In the case of Bitcoin, the likelihood ratio test indicates rejection of the SSD, and Student-*t* in favour of the SGT.[8] However, in a comparison between SGED and SGT, the first distribution cannot be rejected. In addition, the Akaike information criterion (AIC) points to SGED as the best distribution in fitting Bitcoins returns. Note that this distribution is not rejected by the Chi2 and KS tests. Similar results are found for Litecoin, although in the case of this cryptocurrency, the SGED does not pass the Chi2 and KS tests.

For Monero and Stellar, the likelihood ratio test indicates rejection of the SGED, and Student-*t* in favour of the SGT. However, the SSD distribution cannot be rejected. For both cryptocurrencies, the Akaike information criterion (AIC) points to the IHS as the best distribution in fitting returns. Of these three distributions (SGT, SSD and IHS), only the IHS and SGT passed the KS and Chi2 tests. Therefore, we consider these two distributions to be the best in fitting Monero and Stellar returns.

For Ripple and Dash, the likelihood ratio test shows rejection of the SGED, SSD and Student-*t* in favour of the SGT. This result is

[8] In this case, the null hypothesis establishes that the returns follow an SSD distribution, and the alternative hypothesis is that the returns follow an SGT distribution.

Table 8
Accuracy tests for VaR estimations at 10% probability.

	Bitcoin	Monero	Ripple	Litecoin	Dash	Stellar
Long position						
SGT						
% exceptions	10.63	8.91	7.67	11.89	8.76	7.36
Number of tests that pass H_0	4	4	4	4	4	3
SGED						
% exceptions	6.72	7.03	6.57	5.48	7.82	6.10
Number of tests that pass H_0	4	4	3	3	4	2
SSD						
% exceptions	11.74	9.22	8.45	11.74	8.76	7.82
Number of tests that pass H_0	4	4	4	4	4	4
HIS						
% exceptions	9.86	8.44	7.67	8.92	8.29	7.04
Number of tests that pass H_0	4	4	4	4	4	4
Student-*t*						
% exceptions	11.09	8.44	7.67	11.58	8.76	7.36
Number of tests that pass H_0	4	4	4	4	4	4
Normal						
% exceptions	3.13	5.31	4.07	2.82	6.10	4.85
Number of tests that pass H_0	2	2	2	2	1	2
Short position						
SGT						
% exceptions	12.50	8.45	8.92	12.19	9.08	6.57
Number of tests that pass H_0	4	4	4	4	3	4
SGED						
% exceptions	7.03	6.73	7.51	5.94	7.82	5.16
Number of tests that pass H_0	4	4	4	2	4	2
SSD						
% exceptions	10.16	8.45	9.70	12.34	9.08	6.57
Number of tests that pass H_0	3	4	4	4	3	3
HIS						
% exceptions	10.78	7.98	9.55	10.00	9.08	6.42
Number of tests that pass H_0	4	4	4	4	4	4
Student-*t*						
% exceptions	12.50	9.23	10.17	12.50	9.23	7.36
Number of tests that pass H_0	4	4	4	4	4	4
Normal						
% exceptions	3.75	5.63	5.32	2.66	6.10	4.07
Number of tests that pass H_0	2	2	2	2	1	2

Note: The maximum losses at 10% probability in a long and short positions are given by the VaR(10%) and VaR(90%) respectively.

consistent with the Akaike information criterion, which in both cases points to the SGT as the best distribution in fitting these returns. It is worth noting that for these cryptocurrencies, the SGT distribution passes both the Chi2 and KS tests.

In summary, for all currencies considered, the normality hypothesis is always rejected; *second*, there is no evidence that the Student-*t* distribution fits the cryptocurrency data better than skewed distributions. In contrast, the likelihood ratio test indicates that the Student-*t* distribution is outperformed by the SGT distribution; and *third*, there is no one skewed distribution that fits the cryptocurrency data systematically better than the others. Thus, for Bitcoin and Litecoin, the best distribution is SGED, while for Ripple and Dash, the best distribution is SGT. For Monero and Stellar, the SGT and IHS distributions appear to be the best.

4. Implications for portfolio management

Considering the empirical findings discussed in Section 3, we would expect that in the case of cryptocurrencies, skewed distributions provide better market risk estimations than those provided by the normal and Student-*t* distributions.

To corroborate this fact, we compare these distributions in terms of their ability to forecast the market risk of the considered cryptocurrencies. For this purpose, we use the VaR measure, which is estimated using the parametric method described in Section 2.

Through the VaR measure, we quantified the losses associated with long and short positions for two confidence levels, 90% and

Table 9
Loss function.

	Bitcoin	Monero	Ripple	Litecoin	Dash	Stellar
1% probability						
Long position						
SGT	0.480	0.451	0.458	0.528	0.558	0.563
SGED	**0.275**	0.652	**0.436**	**–**	**0.492**	**0.504**
SSD	0.411	0.513	0.574	0.319	0.640	0.640
IHS	0.341	0.424	0.464	**0.135**	0.520	0.529
Student	0.488	**0.407**	0.470	0.474	0.494	0.508
Normal	0.324	0.520	0.561	0.047	0.653	0.653
Short position						
SGT	0.285	0.029	0.205	0.344	0.130	0.085
SGED	**0.199**	0.180	**0.186**	**0.133**	0.114	0.070
SSD	0.258	**0.020**	**0.186**	0.319	**0.069**	**0.028**
IHS	0.218	0.022	0.203	0.179	0.096	0.054
Student	0.329	0.053	0.305	0.408	0.208	0.152
Normal	0.302	0.129	0.403	0.315	0.365	0.291
10% probability						
Long position						
SGT	1.916	1.946	1.583	2.758	2.279	1.835
SGED	**1.311**	**1.699**	**1.466**	**1.617**	**2.020**	**–**
SSD	1.947	1.990	1.706	2.783	2.327	1.890
IHS	1.754	1.884	1.642	2.287	2.225	**1.791**
Student	1.960	1.900	1.652	2.716	2.260	1.825
Normal	–	–	–	–	–	–
Short position						
SGT	1.878	1.358	1.515	2.807	2.026	1.498
SGED	**1.078**	**1.038**	**1.348**	**1.491**	**1.703**	**–**
SSD	1.622	1.352	1.646	2.798	2.023	1.501
IHS	1.608	1.312	1.622	2.249	1.972	**1.459**
Student	1.932	1.430	1.713	2.849	2.103	1.571
Normal	–	–	–	–	–	–

Note: (–) denotes the case in which two or more accurate tests have been rejected.

99%. The sample, which runs from January 1, 2015 to September 30, 2020, is divided into a learning sample from January 1, 2015 to December 31, 2018 and a forecast sample from January 1, 2019 to the end of September 2020.

To evaluate the performance of the distributions, we use a two-stage procedure. First, we evaluate the accuracy of the VaR estimations for which we use several standard tests, and then, for distributions that pass the majority of the accuracy tests, we calculate Lopez's loss function. The best distribution is the one that provides the lowest losses.

Tables 7 and 8 show the percentage of exceptions and the number of accuracy tests that do not reject the null hypothesis, indicating that VaR estimates are accurate. For all currencies considered, all distributions, including the normal distribution, provide accurate VaR estimates in both long and short positions at the 99% confidence level (see Table 7). However, at a lower confidence level of 90%, a normal distribution fails to provide accurate VaR estimates. The rest of the distributions perform well in both long and short positions (Table 8).

Although all distributions provide accurate VaR estimates—except the normal distribution at a low confidence level—not all distributions offer the same results in terms of risk quantification. To quantify the magnitude of the losses in the cases in which the returns fall below VaR, we use Lopez's loss function (Eq. (9)). Table 9 shows these losses.

At the 90% confidence level, the SGED distribution obtains the best results by providing the lowest losses for all currencies in both long and short positions. For a 99% confidence level, in a long position, the SGED distribution provides the lowest losses in the majority of the cases, while in a short position, it is the SSD distribution in conjunction with SGED that provides the best results.

5. Conclusion

The paper contributes to the existing literature through a thorough examination of the statistical properties of cryptocurrency returns, in particular, by analysing which statistical distribution better fits this type of asset. Regarding this last issue, the literature is quite scarce. For this study, we used data from six cryptocurrencies: Bitcoin Monero, Ripple, Litecoin, Dash and Stellar. These cryptocurrencies account for just over 80% of the market capitalization according to the CoinMarketCap data.

The preliminary statistical study allows us to draw the following conclusions about cryptocurrency returns: high volatility, inverse leverage effect, skewed distributions, high kurtosis and positive autocorrelation in returns and squared returns.

To capture the nonnormal characteristics observed in cryptocurrency data, we tested the goodness of fit of a large set of distributions: symmetric and skewed distributions. The distributions included in the comparison are the normal distribution, the Student-*t* distribution (symmetric), the skewed Student-*t* distribution (SSD) of Hansen (1994), the skewed generalized t distribution (SGT) of Theodossiou (1998), the skewed generalized error distribution (SGED) of Theodossiou (2001) and the inverse hyperbolic sign (IHS) of Johnson (1949).

The results indicate that the skewed distributions outperform the normal and Student-*t* distributions in fitting cryptocurrency data, although there is no one skewed distribution that systematically better fits the data. Thus, for Bitcoin and Litecoin, the best distribution is SGED, while for Ripple and Dash, the best distribution is SGT. For Monero and Stellar, the SGT and IHS distributions appear to be the best.

Concerning the aforementioned findings, we would expect that the skewed distributions provide better market risk estimations than those provided by the normal and Student-*t* distributions. To corroborate this fact, we compare these distributions in terms of their ability to forecast the market risk of the cryptocurrencies. In line with the results obtained in the statistical analysis, we find that the skewed distributions provide better risk estimates than the normal and Student-*t* distributions, both in long and short positions, with the SGED distribution providing better results.

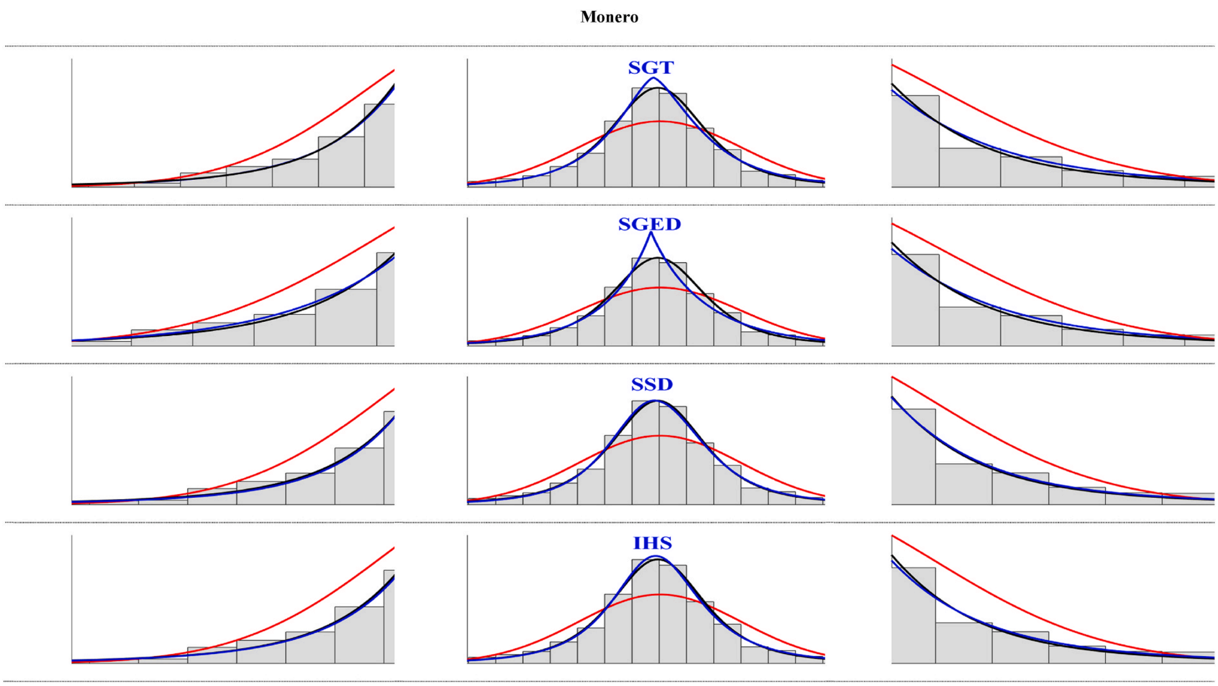
Author statement

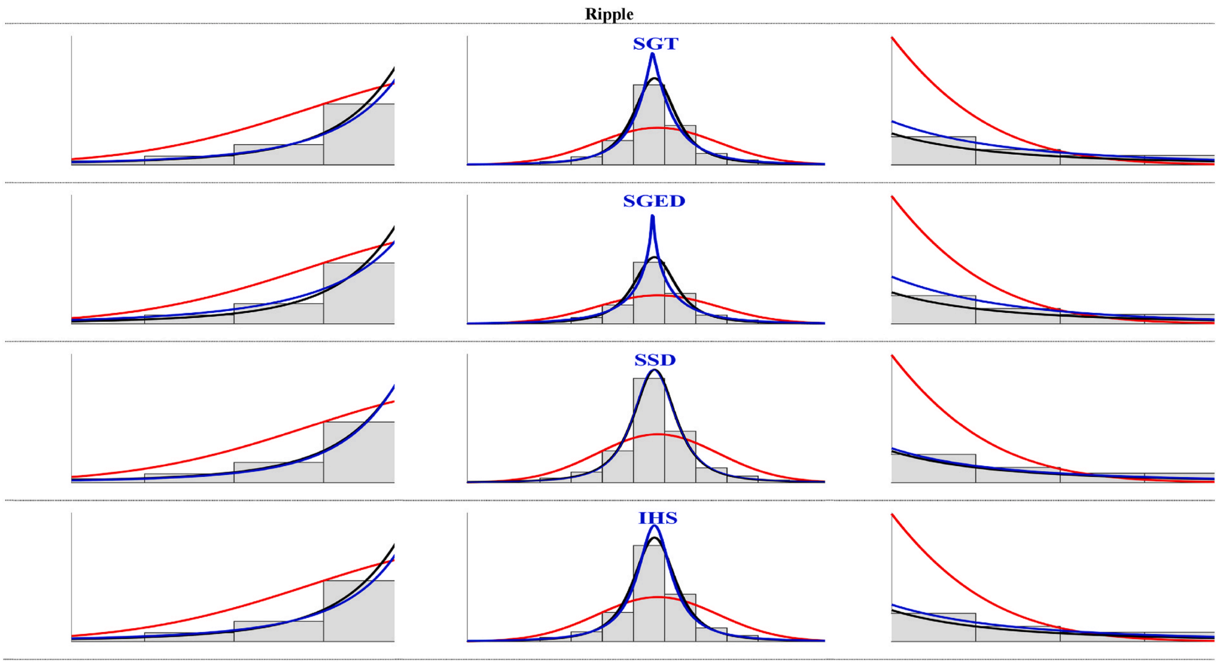
The authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

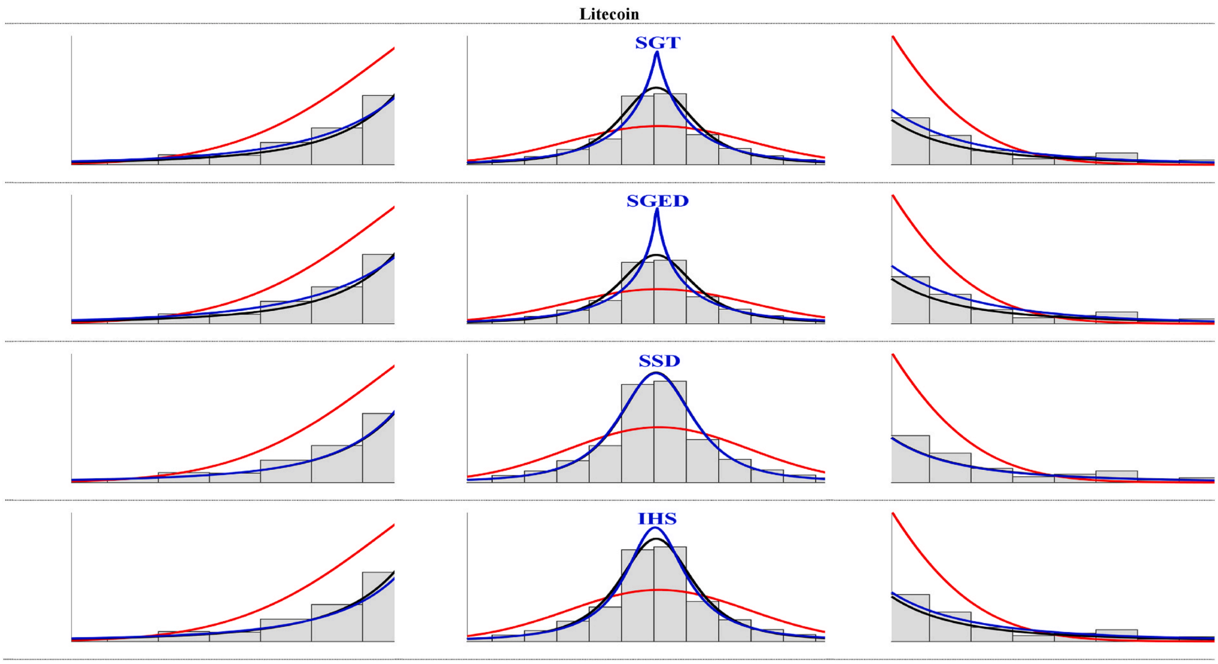
Acknowledgement

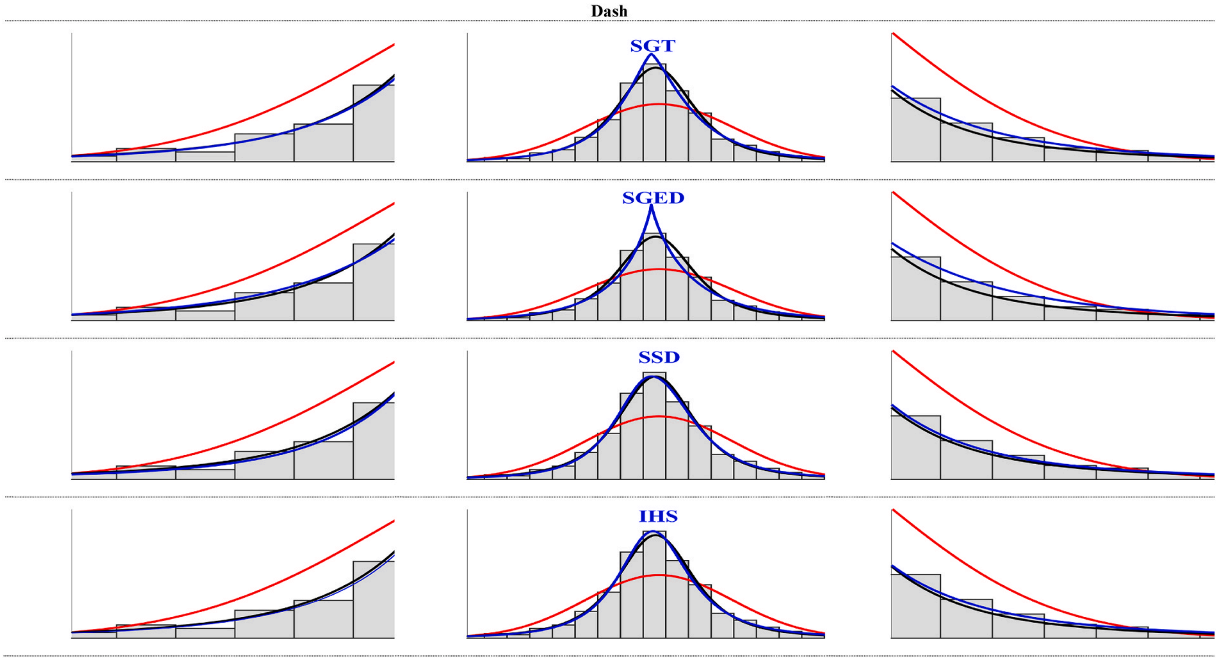
This work has been funded by the Spanish Ministerio de Ciencia e Innovación (PID2020-113367RB-I00).

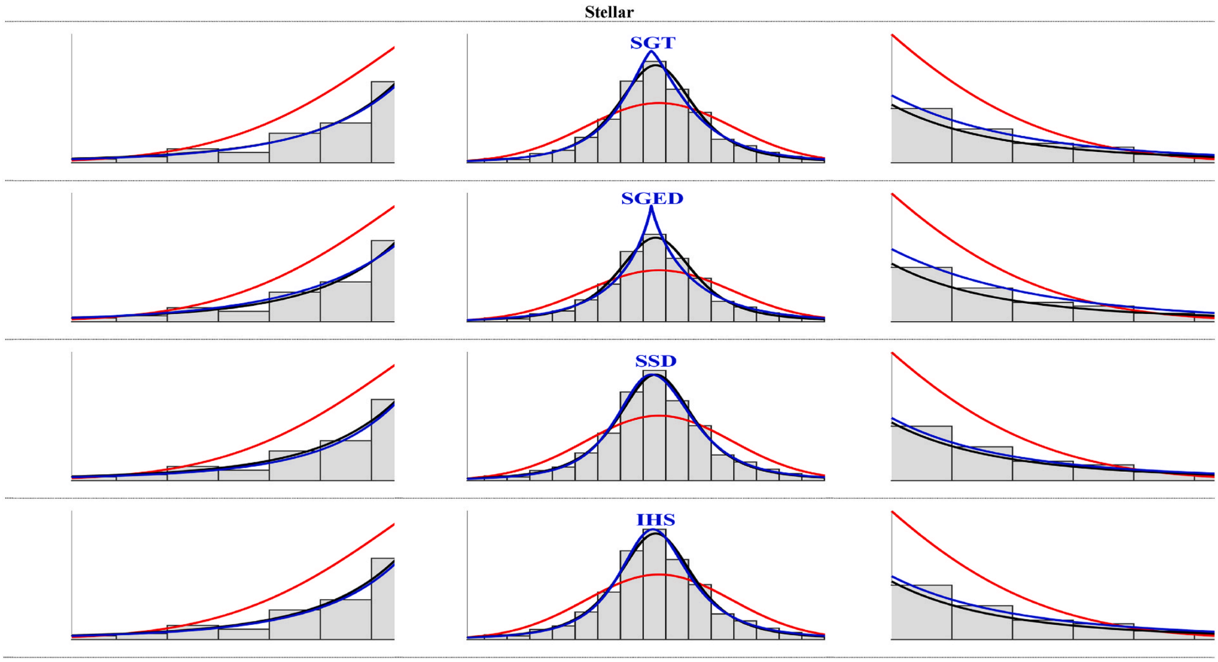
APPENDIX




Ripple
SGT
SGED
SSD
IHS


Litecoin
SGT
SGED
SSD
IHS





References

Abad, P., Benito, S., & López, C. (2014). A comprehensive review of Value at Risk methodologies. *The Spanish Review of Financial Economics, 12*(1), 15–32. https://doi.org/10.1016/j.srfe.2013.06.001

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.

Auer, R., & Claessens, S. (2018). Regulating cryptocurrencies: Assessing market reactions. *BIS quarterly review*, bank for international settlements, september. Available at: https://www.bis.org/publ/qtrpdf/r_qt1809f.htm.

Baek, C., & Elbeck, M. (2015). Bitcoins as an investment or speculative vehicle? A first look. *Applied Economics Letters, 22*(1), 30–34. https://doi.org/10.1080/13504851.2014.916379

Bariviera, A. F., Basgall, M. J., & Hasperué, W. (2017). Some stylized facts of the Bitcoin market. *Physica A, 484*, 82–90. https://doi.org/10.1016/j.physa.2017.04.159

Baur, D. (2012). Asymmetric volatility in the gold market. *Journal of Alternative Investments, 14*, 26–38. https://doi.org/10.3905/jai.2012.15.2.024

Black, F. (1976). Studies in stock price volatility changes. In *Proceedings of the 1976 business meeting of the business and economics statistics section* (pp. 177–181). American Association.

Borri, N. (2019). Conditional tail-risk in cryptocurrency markets. *Journal of Empirical Finance, 50*, 1–19. https://doi.org/10.1016/j.jempfin.2018.11.002

Bouoiyour, J., & Selmi, R. (2015). What does Bitcoin look like? *Annals of Economics and Finance, 16*(2), 449–492.

Bouri, E., Azzi, G., & Dyhrberg, A. H. (2017). On the return-volatility relationship in the Bitcoin market around the price crash of 2013. *Economics - ejournal, 11*, 1–17. https://doi.org/10.2139/ssrn.2869855

Bouri, E., Jalkh, N., Molnár, P., & Roubaud, D. (2017). Bitcoin for energy commodities before and after the december 2013 crash: Diversifier, hedge or safe haven? *Applied Economics, 49*(50), 5063–5073. https://doi.org/10.1080/00036846.2017.1299102

Bouri, E., Molnár, P., Azzi, G., Roubaud, D., & Hagfors, L. I. (2017). On the hedge and safe haven properties of bitcoin: Is it really more than a diversifier? *Finance Research Letters, 20*, 192–198. https://doi.org/10.1016/j.frl.2016.09.025

Brauneis, A., & Mestel, R. (2018). Price discovery of cryptocurrencies: Bitcoin and beyond. *Economics Letters, 165*, 58–61. https://doi.org/10.1016/j.econlet.2018.02.001

Brière, M., Oosterlinck, K., & Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with bitcoin. *Journal of Asset Management, 16*(6), 365–373.

Brühl, V. (2017). Virtual currencies, distributed ledgers and the future of financial services. *Intereconomics, 52*(6), 370–378. https://doi.org/10.1007/s10272-017-0706-3

Caporale, G. M., Gil-Alana, L., & Plastun, A. (2018). Persistence in the cryptocurrency market. *Research in International Business and Finance, 46*, 141–148. https://doi.org/10.1016/j.ribaf.2018.01.002

Chaim, P., & Laurini, M. P. (2018). Volatility and return jumps in bitcoin. *Economics Letters, 173*, 158–163. https://doi.org/10.1016/j.econlet.2018.10.011

Chaim, P., & Laurini, M. P. (2019). Nonlinear dependence in cryptocurrency markets. *The North American Journal of Economics and Finance, 48*, 32–47.

Chan, S., Chu, J., Nadarajah, S., & Osterrieder, J. (2017). A statistical analysis of cryptocurrencies. *Journal of Risk and Financial Management, 10*(2), 1–23. https://doi.org/10.3390/jrfm10020012

Cheah, E. T., & Fry, J. (2015). Speculative bubbles in bitcoin markets? An empirical investigation into the fundamental value of bitcoin. *Economics Letters, 130*, 32–36. https://doi.org/10.1016/j.econlet.2015.02.029

Christoffersen, P. (1998). Evaluating interval forecasting. *International Economic Review, 39*, 841–862. https://doi.org/10.2307/2527341

Chu, J., Chan, S., Nadarajah, S., & Osterrieder, J. (2017). GARCH modelling of cryptocurrencies. *Journal of Risk and Financial Management, 10*(4), 17. https://doi.org/10.3390/jrfm10040017

Chu, J., Nadarajah, S., & Chan, S. (2015). Statistical analysis of the exchange rate of bitcoin. *PLoS One, 10*(7), Article e0133678. https://doi.org/10.1371//journal.pone.0133678

Ciaian, P., Rajcaniova, M., & Kancs, d'A. (2016). The economics of Bitcoin price formation. *Applied Economics, 48*(19), 1799–1815. https://doi.org/10.1080/00036846.2015.1109038

Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar – a GARCH volatility analysis. *Finance Research Letters, 16*, 85–92. https://doi.org/10.1016/j.frl.2015.10.008

Eisl, A., Gasser, S., & Weinmayer, K. (2015). Caveat emptor: Does bitcoin improve portfolio diversification?. Available at https://doi.org/10.2139/ssrn.2408997.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of thevariance of UK inflation. *Econometrica, 50*, 987–1008.

Engle, R., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk byregression quantiles. *Journal of Business & Economic Statistics, 22*, 367–381.

Eom, C., Kaizoji, T., Kang, S. H., & Pichl, L. (2019). Bitcoin and investor sentiment: Statistical characteristics and predictability. *Physica A, 514*, 511–521. https://doi.org/10.1016/j.physa.2018.09.063

Feng, W., Wang, Y., & Zhang, Z. (2018). Can cryptocurrencies Be a safe haven: A tail risk perspective. *Analysis. Applied Economics, 50*(44), 4745–4762. https://doi.org/10.1080/00036846.2018.1466993

Gangwal, S., & Longin, F. (2018). Extreme movements in bitcoin prices: A study based on extreme value theory. ESSEC working paper https://doi.org/10.2139/ssrn.3245571.

García-Jorcano, L., & Benito, S. (2020). Studying the properties of the bitcoin as a diversifying and hedging asset through a copula analysis: Constant and time-varying. *Research in International Business and Finance, 54*, 1–24. https://doi.org/10.1016/j.ribaf.2020.101300

Gkillas, K., Bekiros, S., & Siriopoulos, C. (2018). Extreme correlation in cryptocurrency markets. Available at https://ssrn.com/abstract=3180934.

Gkillas, K., & Katsiampa, P. (2018). An application of extreme value theory to cryptocurrencies. *Economics Letters, 164*, 109–111. https://doi.org/10.1016/j.econlet.2018.01.020

Gkillas, K., & Longin, F. (2019). Is bitcoin the new digital gold? Evidence from extreme price movements in financial markets. *SSRN Electronic*. Available at https://doi.org/10.2139/ssrn.3245571.

Haig, S. (2018). The total market capitalization of all cryptocurrency markets is greater than that of Google. Available at: https://news.bitcoin.com/total-capitalization-of-the-crypto-markets-now-exceeds-750-billion/.

Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review, 35*, 705–730. https://doi.org/10.2307/2527081

Härdle, W. K., Campbell, R. H., & Raphael, C. G. (2020). Understanding cryptocurrencies. *Journal of Financial Econometrics, 18*(2), 181–208. https://doi.org/10.1093/jjfinec/nbz033

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translations. *Biometrika, 36*, 149–176. https://doi.org/10.1093/biomet/36.1-2.149

Jorion, P. (2001). *Value at risk: The new benchmark for managing financial risk*. Mc-Graw Hill.

Kang, S. H., Mclver, R. P., & Hernandez, J. A. (2019). Co-movements between bitcoin and gold: A wavelet coherence analysis. *Physica A, 536*, 1–9. https://doi.org/10.13140/RG.2.2.31672.78084

Katsiampa, P. (2017). Volatility estimation for bitcoin: A comparison of GARCH models. *Economics Letters, 158*, 3–6. https://doi.org/10.1016/j.econlet.2017.06.023

Katsiampa, P., Gkillas, K., & Longin, F. (2018). *Cryptocurrency market activity during extremely volatile periods*. https://doi.org/10.2139/ssrn.3220781

Klein, T., Hien, P. T., & Walther, T. (2018). Bitcoin is not the new gold: A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis, 59*, 105–116. https://doi.org/10.1016/j.irfa.2018.07.010

Köchling, G., Müller, J. P., & P, N. (2019). Does the introduction of futures improve the efficiency of Bitcoin? *Finance Research Letters, 30*, 367–370. https://doi.org/10.1016/j.frl.2018.11.006

Kolmogorov, A. (1933). Sulla Determinazione Empirica di una Leggi di Distribucione. *Giornale dell' Istituto Italiano degli Attuari, 4*, 1–11.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives, 2*, 73–84. https://doi.org/10.3905/jod.1995.407942

Lopez, J. A. (1998). Testing your risk tests. *Financial Survey, 18–20*.

Lopez, J. A. (1999). Methods for evaluating value-at-risk estimates. *Federal Reserve Bank of San Francisco Economic Review, 2*, 3–17.

Lopez-Martin, C., Benito Muela, S., & Arguedas, R. (2021). Efficiency in cryptocurrency markets: new evidence. *Eurasian Economic Review, 11*, 403–431. https://doi.org/10.1007/s40822-021-00182-5

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business, 36*(4), 394–419. http://www.jstor.org/stable/2350970.

Massey, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association, 46*, 68–78. https://doi.org/10.2307/2280095

McDonald, J., & Newey, W. (1988). Partially adaptive estimation of regression models via the generalized t distribution. *Econometric Theory, 4*(3), 428–457. Available at: https://search-proquest-com.ezproxy.uned.es/docview/36969879?accountid=14609.

Nakamoto, S. (2008). Bitcoin: A peer–to–peer electronic cash system. Available at: https://bitcoin.org/bitcoin.pdf.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*(2), 347–370. Available at: https://search-proquest-com.ezproxy.uned.es/docview/214871762?accountid=14609.

Osterrieder, J., & Lorenz, J. (2017). A statistical risk assessment of Bitcoin and its extreme tail behavior. *Annals of Financial Economics, 12*(1), 1750003. https://doi.org/10.1142/S2010495217500038

Osterrieder, J., Strika, M., & Lorenz, J. (2017). Bitcoin and cryptocurrencies—not for the faint-hearted. *International Finance and Banking, 4*(1), 56. https://doi.org/10.2139/ssrn.2867671

Pearson, K. (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. *Philosophy Magazine Series, 50*(5), 157–172.

Phillip, A., Chan, J., & Peiris, S. (2018). A new look at Cryptocurrencies. *Economics Letters, 163*, 6–9. https://doi.org/10.1016/j.econlet.2017.11.020

Procházka, D. (2018). Accounting for bitcoin and other cryptocurrencies under ifrs: A comparison and assessment of competing models. *International Journal of Digital Accounting Research, 18*, 161–188. https://doi.org/10.4192/1577-8517-v18_7

Ray, J. (2019). A next-generation smart contract and descentralized application plataform. White Paper. Available at: https://github.com/ethereum/wiki/wiki/White-Paper.

Rosic, A. (2017). What is cryptocurrency: Everything you need to know. https://blockgeeks.com/guides/what-is-cryptocurrency.

Sensoy, A. (2019). The inefficiency of bitcoin revisited: A high-frequency analysis with alternative currencies. *Finance Research Letters, 28*, 68–73. https://doi.org/10.1016/j.frl.2018.04.002

Smirnov, H. (1939). On the deviation of the empirical distribution function. *Recueil Mathématique (Matematiceskii Sbornik), N. S., 6*, 3–26.

Takaishi, T. (2018). Statistical properties and multifractality of Bitcoin. *Physica A: Statistical Mechanics and Its Applications, 506*, 507–519. https://doi.org/10.1016/j.physa.2018.04.046

Theodossiou, P. (1998). Financial data and skewed generalized t distribution. *Management Science, 44*, 1650–1661. https://doi.org/10.2139/ssrn.65037

Theodossiou, P. (2001). *Skewness and kurtosis in financial data and the pricing of options*. Rutgers University. Working Paper.

Tiwari, A. K., Kumar, S., & Pathak, R. (2019). Modelling the dynamics of bitcoin and Litecoin: GARCH versus stochastic volatility models. *Applied Economics, 51*(37), 4073–4082. https://doi.org/10.1080/00036846.2019.1588951

Tran, V. L., & Leirvik, T. (2019). Efficiency in the markets of crypto-currencies. *Finance Research Letters, 35*, 101382. https://doi.org/10.1016/j.frl.2019.101382

Vidal-Tomás, D., Ibáñez, A. M., & Farinós, J. (2019). Weak efficiency of the cryptocurrency market: A market portfolio approach. *Applied Economics Letters, 26*(19), 1627–1633. https://doi.org/10.1080/13504851.2019.1591583

Wei, W. C. (2018). Liquidity and market efficiency in cryptocurrencies. *Economics Letters, 168*, 21–24. https://doi.org/10.1016/j.econlet.2018.04.003

Zang, S., & Lee, J. H. (2019). Double-spending with a sybil attack in the bitcoin decentralized network. *IEEE Transactions on Industrial Informatics, 15*(10), 5715–5722.

Zhang, W., Wei, Z., Wang, P., Li, X., & Shen, D. (2018). Some stylized facts of the cryptocurrency market. *Applied Economics, 50*(55), 5950–5965. https://doi.org/10.1080/00036846.2018.1488076