



Semi-nonparametric risk assessment with cryptocurrencies

Inés Jiménez^a, Andrés Mora-Valencia^b, Javier Perote^{a,*}

^a Department of Economics and Economic History and IME, Faculty of Economics and Business, University of Salamanca, Campus Miguel de Unamuno (Edif. F.E.S.), 37007, Salamanca, Spain

^b School of Management, Universidad de los Andes, Calle 21 No. 1-20, Bogotá, 111711, Colombia

ARTICLE INFO

JEL classification:

C01
C12
C13
C14
C16
C53

Keywords:

Gram Charlier series
Value-at-Risk
Expected shortfall
Median shortfall
Backtesting
Cryptocurrencies

ABSTRACT

This paper establishes a brand-new perspective of analyzing the risk of crypto assets through a semi-nonparametric approach, discussing its theoretical advantages and testing its performance compared to parametric approaches and in terms of backtesting techniques and different risk measures: Value-at-Risk, Expected Shortfall and Median Shortfall. Our comprehensive analysis for six cryptocurrencies shows that flexible semi-nonparametric approaches outperform risk measures of most crypto assets (particularly Bitcoin) and tend to provide the most conservative risk assessment. Furthermore, we propose the Median Shortfall as a robust-to-outliers and reliable risk measure for cryptocurrencies and discuss on the choice of the appropriate probability levels according to the assumed distribution. The evidence supports that Median Shortfall at 98.31 % and 98.51 % confidence levels as accurate alternatives to Value-at-Risk at 99 % and Expected Shortfall at 97.5 %.

1. Introduction

Cryptocurrencies have plunged from attracting attention and being known by only a few, to being a global economic landmark studied by regulators, academics, policymakers, governments and institutions. Just 12 short years since their inception, cryptos are an extraordinary fast-growing industry and a threat to the stability of the full monetary and financial system (Dyhrberg, 2016b).

It was in 2008 when the first paper about Bitcoin (BTC), the first fully developed virtual currency, appeared (Nakamoto, 2008). On September 27th, 2020, there were more than eight thousand cryptocurrencies, over three hundred billion dollars on the market (www.coinmarketcap.com). Despite the number of cryptocurrencies that exist, merely a few are relevant today. At that time, Bitcoin represents 59.14 % out of the total market followed by Ethereum (ETH) with 11.67 %. Just two cryptocurrencies give a total number of more than 70 % equivalent to more than two hundred and thirty billion dollars of market capitalization.

Cryptocurrencies own specific characteristics which make them different from fiat currencies. They have a decentralized structure where no regulatory or financial institutions are required. The algorithms transact around the blockchain concept and attest about the transactions made itself, replacing financial intermediaries and central banks on monetary issues. Blockchain technology is based on cryptographic tests whose transactions are irreversible, avoiding errors in the blockchain or misusing money from any users. This system acts such as peer-to-peer concept, so in this way, the confirmation of all the financial transactions is made in a decentralized way. Thus, every node interacts within a network, which searches by majority agreement the addition of a new block to the blockchain.

* Corresponding author.

E-mail addresses: inesjimenez@usal.es (I. Jiménez), a.mora262@uniandes.edu.co (A. Mora-Valencia), perote@usal.es (J. Perote).

This structure provides performance and own capabilities unknown in comparison to the traditional economy. Moreover, they provide global access to certain resources and financial products, being a way to assure equal opportunities to the users regardless of nationality or heritage.

Until recently, the term cryptocurrency was exclusively used, but today it would be more correct to include the term 'cryptoasset' commonly-accepted when refer to cryptocurrencies, tokens, and any other asset from that kind (Rauchs et al., 2019). In this paper, we will continue to talk about them as cryptocurrencies although we consider them not an alternative transaction system, but a speculative investment (Glaser et al., 2014; Yermack, 2015). In fact, many papers have studied the short- and long-term determinants for Bitcoin demand e.g. (Dwyer, 2015; De la Horra et al., 2019) and numerous authors have explained the relationship with other financial assets (Dyhrberg, 2016b; Corbet et al., 2018) among others. However, perhaps the major area of concern to policymakers, economists, practitioners, and researchers is riskiness, volatility or setting proper forecasting patterns. The latter has great practical value since cryptocurrencies exhibit much higher volatility than any other asset because they are extremely dependent on news and speculations as in mainstream finances, apart from the lack of maturity they have since they are in their early life-cycle stage (Baek and Elbeck, 2015). For that reason, this paper is focused on this area.

Volatility analysis has especially based on Bitcoin and has been analyzed by different authors (Baek and Elbeck, 2015; Balcilar et al., 2017; Bouri et al., 2017; Blau, 2018; Ardía et al., 2019; Borri, 2019). Some of them, are focused on studying volatility properties with different GARCH-types processes such as EGARCH (Dyhrberg, 2016a), TGARCH (Bouoiyour and Selmi, 2016), CGARCH (Katsiampa, 2017), FIGARCH (Lahmiri et al., 2018), and measuring relative performance of different GARCH types (Chu et al., 2017).

Although there is no consensus on which the best process is, because explaining the extreme volatility of Bitcoin and other cryptos involve not only conditional heteroskedastic processes but also the modelling of the full shape of the density of the underlying stochastic process. This density should account for skewness, highly kurtosis, and extreme jumps at the far end of the tails due to presence of extreme values. Along the line of Cerqueti et al. (2020), where the Skewed non-Gaussian GARCH models are analyzed for cryptocurrencies, we propose the use of the semi-nonparametric (SNP) approach, which is based on the asymptotic properties of Gram-Charlier (GC) series for approximating frequency functions featuring all these stylized facts with a flexible parametric structure. Although such series expansions have a long history in mathematics and statistics – see Hald (2000) for a revision on the origins – their applications to economics and finance did not appear until the last decades e.g. (Sargan, 1976; Corrado and Su, 1996; Mauleón and Perote, 2000), introduced this approach to econometrics, asset pricing and risk assessment respectively.

Recently, this SNP approach has been proved to be useful for representing asset returns with high volatility and extreme values, e.g. commodity ETFs (Del Brio et al., 2020) or Bitcoin (Jiménez et al., 2020). However, to the best of our knowledge, this approach has never been applied to other cryptocurrencies or altcoins, which seem to be even more volatile. Furthermore, in the presence of high instability, the estimation of GC densities might produce unreliable results if estimation algorithms fail to converge to global optima. To solve this problem, positive transformations, e.g. (León et al., 2009; Níguez et al., 2012) are worth implementing. This paper covers these two gaps in the literature on cryptocurrencies risk assessment by providing a comprehensive analysis on the performance of GC densities with two versions: GC (Type A) and Positive GC for a wide variety of cryptocurrencies, compared to other parametric alternatives (Gaussian, Student's-t, and Skewed-t). Model performance is tested for different risk measures and tests including Model Confidence Sets (MCS) procedures (Hansen et al., 2011), implemented for the first time in this framework.

The Value-at-Risk (VaR) is the most widely used measure to assess financial risk because its simplicity in both calculation and interpretation – it is just a quantile of the return/loss distribution. In spite of this, the VaR has been criticized for not capturing extreme events properly and not being a coherent risk measure (i.e., does not always meet subadditivity property). These issues are particularly important for cryptocurrencies since they are extremely volatile, and they can be used to portfolio diversification.

In contrast to VaR, the expected shortfall (ES) considers the tail risk beyond the quantile and it is defined as the expected loss given that the VaR bound has been exceeded. From the year 2016, this measure has been proposed by the Basel Committee of Banking Supervision (BCBS) shifting 99 %-VaR to 97.5 %-ES. Although ES is coherent to portfolio risk diversification (i.e., the risk of the portfolio does not exceed the sum of individual risks), it does not fulfil elicibility property thus hindering the ES backtesting procedures. In this paper, we perform a simple implicit approach (Emmer et al., 2015; Kratz et al., 2018) that has been recently applied for Bitcoin (Acereda et al., 2019; Jiménez et al., 2020) and other assets (Díaz et al., 2019), but, as far as we know, never before for other cryptocurrencies. The implementation of these quite novel techniques allows the testing of model performance in terms of ES but based on multinomial VaR for different confidence levels.

Recent literature has focused on risk assessment with VaR and ES for a different assets, including cryptocurrencies (Gkillas and Katsiampa, 2018; Stavroyiannis, 2018; Trucíos, 2019; Liu et al., 2020), some of them using backtesting techniques (Acerbi and Szekely, 2014; Fischer et al., 2018). Though both VaR and ES are the best-known risk measures and the most commonly used, however, they have some caveats. Whereas VaR does not have the ability to capture "tail risk", ES is very sensitive to extreme values. For this reason, median shortfall (MS) has been proposed as a new alternative (So and Wong, 2012; Barnard et al., 2018; Fischer et al., 2018). Although Jiménez et al. (2020) introduce this measure in the analysis of Bitcoin returns, in this paper we propose its use to alternative cryptocurrencies, which seems to be even more volatile. Like ES, MS considers the tail risk beyond VaR but with the robustness property of the median statistic. This is why this measure is appropriate for extremely heavy-tailed data as infinite mean models, where ES may not be calculated in certain distribution models (Nešlehová et al., 2006).

Although all cryptocurrencies exhibit common general features, each of them has particularities that make it different from the others. Therefore, we conjecture that the flexible SNP approach becomes appropriate to tackle all of them, but all their special features at the same time. In this paper we consider probably the six most relevant cryptocurrencies: Bitcoin (BTC), Litecoin (LTC), Ripple (XRP), Monero (XMR), Stellar (XLM), and Ethereum (ETH), which in turn can be divided into two groups: Bitcoin and altcoins, that refers to the alternative cryptocurrencies different from Bitcoin, as suggested by different authors (Cagli, 2019; Nguyen et al., 2019).

Furthermore, the model performance, according to the results of our study, is highly dependent on the risk measure employed and the criteria established by regulators and risk managers.

In summary, our main contributions are: (i) the risk analysis of a wide group of crypto assets in terms of backtesting techniques and for a variety of risk measures and densities; (ii) the introduction of the SNP approach as a general methodology for cryptocurrencies and considering not only alternative expansion lengths but also positive transformations, as well as the GJR-GARCH for conditional variance; (iii) the discussion on the MS appropriateness to tackle the sensitivity of risk measures at the very far end of the tail of the cryptocurrency distribution; and (iv) the comprehensive analysis for risk assessment for different criteria and methodologies, including multinomial tests, relative pairwise model performance and model ranking in terms of MCS. Our analyses show that the SNP seems to provide conservative but accurate risk measures and MS with a proper probability level (98.31 % and 98.51 %) is a robust-to-outliers alternative to 99 %-VaR and 97.5 %-ES in the presence of extreme volatility.

The remainder of the paper is structured as follows: Section 2 reviews the SNP methodology for modelling the full shape of risky assets and outlines the model and the methodology for the risk assessment of cryptocurrencies on the basis of backtesting techniques and different risk measures. Section 3 presents the results of the SNP approach compared to other parametric alternatives and different risk measures. Section 4 introduces a brief discussion on the impact of the choice of the risk measurement model on regulatory capital, and Section 5 concludes.

2. Model and methodology

2.1. SNP approach

2.1.1. Gram Charlier (Type A)

The use of the SNP approach for modelling non-Gaussian stochastic processes is sustained by the property of GC (Type A) series as the least squares approximation of an arbitrary integrable function $f(x)$ in terms of an infinite series of orthogonal polynomials $\{H_r(x)\}_{r=0}^{\infty}$. These polynomials, where derived by (Chebyshev, 1855; Hermite, 1859) as the series satisfying

$$f(x) \approx \sum_{r=0}^{\infty} a_r H_r(x) \quad (1)$$

where

$$a_r = \frac{\int_{-\infty}^{\infty} f(x) H_r(x) \omega(x) dx}{\int_{-\infty}^{\infty} H_r(x)^2 \omega(x) dx} \quad (2)$$

and

$$\int_{-\infty}^{\infty} H_r(x) H_s(x) \omega(x) dx = 0 \text{ for } r \neq s \quad (3)$$

This seminal result is the basis of different convergent series, e.g., Fourier, Legendre, Maclaurin, or Laguerre for different choices of the weighting function $\omega(x)$. Particularly, if the Gaussian probability function (pdf) is chosen, i.e.

$$\omega(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) \quad (4)$$

then, the so-called Chebyshev-Hermite (CH) polynomials are obtained, such that

$$\frac{d^r \omega(x)}{dx^r} = (-1)^r H_r(x) \omega(x) \quad (5)$$

A closed formula for the CH polynomial of order r is

$$H_r(x) = r! \sum_{s=0}^{\lfloor r/2 \rfloor} \frac{(-1)^s x^{r-2s}}{s! (r-2s)! 2^s} \quad (6)$$

Besides the orthogonality property in Eq. (3) CH polynomials satisfy that

$$\int_{-\infty}^{\infty} H_r(x)^2 \omega(x) dx = r! \quad (7)$$

and thus, the GC series of Type A is traditionally expressed as

$$f(x) = \sum_{r=0}^{\infty} c_r H_r(x) \omega(x) \quad (8)$$

where

$$c_r = \frac{1}{r!} \int_{-\infty}^{\infty} f(x) H_r(x) dx \quad (9)$$

This series allows to approximate empirical frequency functions to any degree of accuracy by considering a sufficiently large expansion, as well as to measure departures from the Gaussian distribution, which is frequently found in real series, particularly those related to economic and financial data. However, the truncation of the series in a term n requires to constrain the allowable range for $\{c_r\}_{r=1}^n$ such that $f(x)$ has a positive domain (Jondeau and Rockinger, 2001; León et al., 2009; Níguez and Perote, 2012). Although for most series implementing constraints or positivity restrictions is not necessary, provided that appropriate maximum likelihood (ML) estimation algorithms and checks for avoiding local optima are implemented. On the other hand, the condition $c_0 = 1$ ensures that the GC expansion integrates to one and, by convention, it is also assumed that $H_0(x) = 1$.

Although under such conditions the (truncated) GC expansion in Eq. (10) characterizes a stochastic process in terms of its pdf,

$$g(x) = \sum_{r=0}^n c_r H_r(x) \omega(x) \quad (10)$$

it is also convenient to formulate other alternative characterizations such as its cumulative distribution function (cdf) – see e.g. Trespalacios et al. (2020)

$$G(\phi) = Prob[x \leq \phi] = \int_{-\infty}^{\phi} \omega(x) dx - \omega(\phi) \sum_{r=1}^n c_r H_{r-1}(\phi) \quad (11)$$

or the moment generating function (mgf) – (Níguez et al., 2012),

$$M(t) = \int_{-\infty}^{\infty} e^{tx} g(x) dx = e^{\frac{t^2}{2}} \sum_{r=0}^n c_r t^r \quad (12)$$

The latter function allows to directly compute the distribution moments, although they can also be obtained as

$$E \left[x^r \right] = \sum_{s=0}^r D_s c_s \frac{r!}{2^{\frac{r-s}{2}} \left(\frac{r-s}{2} \right)!} \quad (13)$$

where D_s is a dummy scoring one if $r - s$ is even and zero otherwise. This implies that moments are finite for all order r and the moment of order r even (odd) depends linearly on all c_s with s even (odd) and such that $s \leq r$, i.e. if $c_s = 0 \forall s$ odd then the distribution is symmetric. That relationship among its moments and parameters is a major advantage of this expansion, which makes straightforward interpretation; Thus, c_1 and c_2 correspond to the mean and variance and c_3 and c_4 capture skewness and excess kurtosis, respectively. The higher-order even parameters have an impact on the higher-order moments and thus to capture extreme values and jumps at the distribution tails the inclusion of at least c_6 and c_8 is suitable – see Mauleón and Perote (2000). On the other hand, in order to avoid overparameterization and keep parsimony the only odd parameter that is usually considered is c_3 (also note that higher-order odd parameters are usually insignificant and may yield model identification problems).

The model estimation follows two steps –see Del Brio et al. (2011):

Step 1: AR(1)-GJR-GARCH(1,1) is estimated under Gaussian distribution – i.e. Quasi ML (QML) estimation.

Step 2: Standardized residuals obtained in the previous step are used to estimate SNP parameters.

2.1.2. Positive Gram-Charlier

For the purpose of controlling for possible estimation problems due to extreme values and following (León et al., 2009; Níguez and Perote, 2012), a Positive Gram-Charlier version has been also considered. This approach, built on the same philosophy of the Gallant and Nychka (1987) transformation, yields to the Positive GC pdf in Eq. (14):

$$g^*(x) = \frac{1}{w(n)} \left(1 + \sum_{r=1}^n c_r^2 H_r(x)^2 \right) \omega(x) \quad (14)$$

where $\omega(x)$ denotes the standard Gaussian pdf in Eq. (4), c_s^2 parameters and CH polynomials are squared, and w is the constant that ensures that the density integrates up to one :

$$w(n) = 1 + \sum_{r=1}^n c_r^2 r! \quad (15)$$

It is clear that Eqs. (14) and (15) render a well-defined symmetric pdf in the full parametric space and in virtue of the Eqs. (3) and (7) it may be also characterized by the cdf in Eq. (16) – see Níguez and Perote (2012) for more details:

$$G^*(\phi) = Prob[x \leq \phi] = \int_{-\infty}^{\phi} \omega(x) dx - \frac{\omega(\phi)}{w(n)} \sum_{s=1}^n \sum_{i=0}^{s-1} c_s^2 \frac{s!}{(s-i)!} H_{s-1}(\phi) H_{s-i-1}(\phi) \quad (16)$$

Although this type of distribution assures positiveness in the whole domain, it is at the expense of losing the linear relationship between non-central moments and parameters and being symmetric distribution. The latter feature may be solved by considering the positive SNP family in León et al. (2009) but also by taking into account an asymmetric model for the conditional variance as the

GJR-GARCH. In the empirical model described in Section 2.2 we opt by this latter solution.

As in the previous case, Gram Charlier Type A, it is estimated in two steps with ease.

2.2. The model

Let us consider that losses (L_t) follow an AR(1)-GJR-GARCH(1,1) model for conditional mean (μ_t) and variance (σ_t^2), and with an innovation term distributed as *iid* with (standardized) pdf H :

$$L_t = \mu_t + \varepsilon_t \quad (17)$$

$$\varepsilon_t = \sigma_t z_t \quad (18)$$

$$z_t \sim H(0, 1), \quad (19)$$

$$\mu_t = \varphi_0 + \varphi_1 \mu_{t-1} \quad (20)$$

$$\sigma_t^2 = \varpi + \alpha \varepsilon_{t-1}^2 + I_{t-1} \gamma \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (21)$$

where $|\varphi_1| < 1$, $\varpi \geq 0$, $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \gamma \geq 0$ and $I_{t-1} = \begin{cases} 0 & \text{if } \varepsilon_{t-1} > 0 \\ 1 & \text{if } \varepsilon_{t-1} < 0 \end{cases}$

GJR-GARCH model (Glosten et al., 1993) is an extension of the GARCH model ($\gamma = 0$) which considers asymmetry in the conditional variance where the negative stocks have a stronger impact than positive ones.

In this research, we have considered different (standardized) density specifications for H : Gaussian, Student's t, Skewed-t, and two alternatives for SNP: Gram-Charlier type A (SNP, hereafter) with c_3 and c_4 parameters (i.e., just incorporating skewness and kurtosis parameters as in most applications of this approach in finance literature) and SNP with c_3 , c_4 , c_6 and c_8 parameters, in order to compare the impact of enlarging the expansion length, especially to analyze the far end – i.e. expansions are truncated in $n = 4$ and $n = 8$. Note that the inclusion of high-order parameters seems recommendable given the high proportion of extreme values exhibited by cryptocurrencies and for the sake of model identification a single odd parameter is required. Furthermore, Positive Gram Charlier (henceforth refereed as PGC) with c_2 , c_4 , c_6 is also included in the backtesting procedures in order to make a comparison between SNP alternatives. In this case, since the distribution is symmetric, the inclusion of c_3 is worthless.

As the model in Eqs. (17)–(21) represents a location-scale model, we assume $c_1 = 0$ and $c_2 = 0$ since mean and variance are explicitly modelled as denoted by Eqs. (20) and (21). Regarding the rest of the parameters of the SNP models, as previously mentioned, c_3 and c_4 capture skewness (since $c_r = 0$ for all r odd and $r \neq 3$) and excess kurtosis, and c_6 and c_8 incorporate the jumps in the probabilistic mass at the far end of the tails due to the presence of extreme values. Therefore, the SNP densities are built on the following polynomials:

$$H_0(z_t) = 1, H_3(z_t) = z_t^3 - 3z_t, H_4(z_t) = z_t^4 - 6z_t^2 + 3, H_6(z_t) = z_t^6 - 15z_t^4 + 45z_t^2 - 15 \text{ and } H_8(z_t) = z_t^8 - 28z_t^6 + 210z_t^4 - 420z_t^2 + 105. \quad (22)$$

In spite of Gaussian, SNP and PGC densities described in Eqs. (4), (10) and (14) and (15), respectively, we also consider the Skewed-t (Fernández and Steel, 1998) and its symmetric Student's t pdfs in Eqs. (23) and (24) as natural competitors to account for excess kurtosis and skewness of cryptocurrencies.

$$st(z_t) = \begin{cases} -\frac{2}{\xi + \frac{1}{\xi}} t(\xi z_t) & \text{for } z_t < 0 \\ \frac{2}{\xi + \frac{1}{\xi}} t(\xi z_t) & \text{for } z_t \geq 0 \end{cases}, \quad (23)$$

where

$$t(z_t) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi(v-2)}\Gamma\left(\frac{v}{2}\right)} \left(1 + \frac{z_t^2}{v-2}\right)^{-\frac{v+1}{2}} \quad (24)$$

v is the number of degrees of freedom (df) or shape parameter, Γ is the gamma function and ξ is the skew parameter. However, we conjecture that the extreme volatility and abundant extreme values exhibited for these assets require a more flexible approach as that of the SNP density. Another sound alternative would be the Generalized Error Distribution (GED), recently applied by Cerqueti et al (2020) in this framework, although we argue that the GC is even more flexible and it has the asymptotic property of approximating the true distribution. Even more, we argue that the advantage of using an appropriate density specification is not only on the estimation of the mean-variance model – since Engle and González-Rivera (1991) showed that QML estimation is consistent even if observations were non-Gaussian – but for providing accurate risk measures.

2.3. Risk measures

2.3.1. Value at Risk

The VaR represents the maximum loss for a given probability level α in a holding period. In this research, we are focusing on the right tail since we considered losses instead of returns and α corresponds to a high level of probability close to 100 %.

$$VaR_\alpha = \inf\{x | G_L(x) \geq \alpha\} \quad (25)$$

where $G_L(x)$ is the cumulative distribution function of losses, L_t and thus the VaR is the quantile of the loss distribution for the α probability level. If $\phi_{t+1}(\alpha)$ is the estimated α -quantile of the standardized distribution H in Eq. (19), then the VaR computed in t for $t+1$ at probability α is:

$$VaR_{t+1,\alpha} = \mu_{t+1} + \sigma_{t+1}\phi_{t+1}(\alpha) \quad (26)$$

where μ_{t+1} and σ_{t+1} are the one-step ahead forecasted conditional mean and standard deviation (volatility), respectively, obtained from the AR(1)-GJR-GARCH(1,1) in Eqs. (20) and (21). Therefore, the VaR for the SNP/PGC distribution can be obtained from its cdf in Eqs. (11)/(16).

2.3.2. Expected shortfall

ES, also known as Conditional VaR (CVaR) or Tail VaR (TVaR), is the expected loss conditioned on the fact that VaR is exceeded and thus provides more information of the tail shape beyond VaR. Therefore, ES for variable z_t with pdf denoted by $g(z_t)$, can be defined by,

$$ES_\alpha = [z_t | z_t > \phi_t(\alpha)] = \int_{\phi_t(\alpha)}^{\infty} z_t g(z_t) dz_t \quad (27)$$

and for the SNP density in Eq. (10) becomes

$$ES_\alpha = \frac{\omega(\phi_t(\alpha))}{1-\alpha} \left[1 + \sum_{r=3}^n c_r [H_r(\phi_t(\alpha)) + rH_{r-2}(\phi_t(\alpha))] \right] \quad (28)$$

– see Del Brio et al. (2020) for further details. Furthermore, the ES can be computed in terms of the continuously aggregated quantiles beyond the VaR and thus for the time horizon $t+1$ may be expressed as

$$ES_{t+1,\alpha} = \mu_{t+1} + \sigma_{t+1} \left(\frac{1}{1-\alpha} \int_a^1 \phi_u du \right) \quad (29)$$

where ϕ_u is the quantile function of variable z_{t+1} . Consequently, Emmer et al. (2015) propose approximating ES by averaging N quantiles corresponding to confidence levels $\alpha_j = \alpha + \frac{j-1}{N}(1-\alpha)$, $\forall j = 1, \dots, N$, $N \in \mathbb{N}$. Particularly, in our application, we employ an approximation of Eq. (28) for $N = 8$, since at least a minimum value of $N \geq 4$ has been argued to provide reliable results (Kratz et al., 2018). However, eight levels of VaR are preferable for obtaining more accurate results in the tests.

$$ES_\alpha \approx \frac{1}{N} \sum_{j=1}^N \phi_{\alpha + \frac{j-1}{N}(1-\alpha)} \quad (30)$$

Although it is not an exact method, it is an accurate and easy to compute approximation, e.g., for $\alpha = 0.975$ it consists of computing independently eight levels ($N = 8$) of VaR: 97.5 %-VaR, 97.8125 %-VaR, 98.125 %-VaR, 98.4375 %-VaR, 98.75 %-VaR, 98.0625 %-VaR, 99.375 %-VaR and 99.6875 %-VaR. These levels are averaged for obtaining the 97.5 %-ES approximation.

2.3.3. Median shortfall

MS is the median loss of the tail risk defined by a given VaR, thus it provides a measure of information beyond VaR that is not sensitive to outliers. Like VaR, it is a quantile-based measure and it may be straightforwardly computed, as suggested by Kou and Peng (2016, 2014),

$$MS_\alpha = VaR_{1+\frac{\alpha}{2}} \quad (31)$$

According to the equation above the 99 %-MS and 97.5 %-MS is estimated as 99.5 %-VaR and 98.75 %-VaR.

2.4. Backtesting procedures

Backtesting techniques are procedures used to examine forecasting model performance in terms of different score variables (Campbell, 2007). We implement these procedures for risk management purposes and thus we compare the VaR, ES, and MS performance to predict ex-post real losses of cryptocurrencies.

2.4.1. VaR backtesting

Backtesting procedures split the total sample T in a rolling estimation window of size W_E and a testing window of size $W_T = T - W_E$

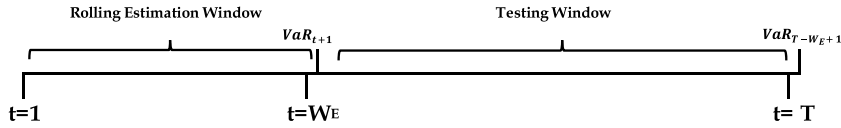


Fig. 1. Backtesting procedure.

(see Fig. 1). The one-day-ahead VaR is forecasted along the testing window by rolling the estimation window (adding a new observation but keeping W_E unchanged) and the predicted VaR obtained in every estimation window is compared to the actual losses for the same period t (Danielsson, 2011)

It is said that a violation or exception will occur on a day $t + 1$ when the observed losses (L_{t+1}) exceed the predicted VaR for a given probability α and, consequently, we define an indicator function that takes the value 1 in this case and 0 otherwise, i.e.

$$I_{t+1}(\alpha) = \begin{cases} 1 & \text{if } L_{t+1} \geq VaR_{t+1}(\alpha) \\ 0 & \text{if } L_{t+1} < VaR_{t+1}(\alpha) \end{cases} \quad (32)$$

If the number of violations significantly exceed the $1 - \alpha$ %, then the exceptions of the model are overestimated, and its risk underestimated. Therefore, model performance can be tested with different procedures, among them we choose the Conditional Coverage test, which test jointly Bernoulli and Independence test, and Dynamic Conditional Quantile (DQ) test proposed by Engle and Manganelli (2004).

2.4.1.1. Bernoulli coverage test. It is also referred to Unconditional Coverage test of Kupiec (1995) and is based on the number of violations occurred every day. Under correct specification, the $I_{t+1}(\alpha)$ sequence should be identically and independently distributed as a Bernoulli distribution, which is exactly the null hypothesis of the test:

$$H_0 : I_{t+1} \sim i.i.d. \text{Bernoulli}(p) \quad (33)$$

The statistic is based on the likelihood ratio (LR),

$$LR_{uc} = -2 \ln \left(\frac{L(p)}{L(\hat{\Pi})} \right) = -2 \ln \left(\frac{(1 - \alpha)^{T_0} \alpha^{T_1}}{(1 - \hat{\pi})^{T_0} \hat{\pi}^{T_1}} \right) \xrightarrow{T \rightarrow \infty} \chi_1^2, \quad (34)$$

where T_0 (T_1) is the number of zeros (ones) for a given probability α and $\hat{\pi}$ is violation ratio computed as $T_0 / (T_0 + T_1)$. This LR follows a chi-square distribution with one degree of freedom and its p-value is computed as

$$p - \text{value} = 1 - F_{\chi_1^2}(LR_{uc}) \quad (35)$$

where $F_{\chi_1^2}(\cdot)$ is the cumulative distribution function of the χ_1^2 .

2.4.1.2. Independence test. This is a LR test built for examining the independence of the violations which considers that, under correct specification (null hypothesis), violation sequence holds this property. The test statistic is given in Eq. (36) and its p-value computed as in Eq. (35).

$$LR_{ind} = -2 \ln \left(\frac{L(\hat{\Pi})}{L(\hat{\Pi}_1)} \right) = -2 \ln \left(\frac{(1 - \hat{\pi})^{T_0} \hat{\pi}^{T_1}}{(1 - \hat{\pi}_{01})^{T_{00}} \hat{\pi}_{01}^{T_{01}} (1 - \hat{\pi}_{11})^{T_{10}} \hat{\pi}_{11}^{T_{11}}} \right) \xrightarrow{T \rightarrow \infty} \chi_1^2 \quad (36)$$

where violation ratios are based on $T_{ij} = \begin{cases} I_t(\alpha) = j \\ I_{t-1}(\alpha) = i \end{cases}$ and $\begin{cases} T_0 = T_{00} + T_{10} \\ T_1 = T_{11} + T_{01} \end{cases}$.

2.4.1.3. Dynamic Conditional Quantile test. This test, developed by Engle and Manganelli (2004), is based on a regression model that links the current and the former violations, using lagged values (the standard is four lags) and explanatory variables, transforming VaR series into a hit sequence, $Hit_t(\alpha)$, defined by:

$$Hit_t(\alpha) = \begin{cases} 1 - \alpha & \text{if } r_{t+1} < VaR_{t+1}(\alpha) \\ -\alpha & \text{if } r_{t+1} \geq VaR_{t+1}(\alpha) \end{cases} \quad (37)$$

Following the regression model:

$$Hit_t(\alpha) = \delta_0 + \sum_{i=1}^p \delta_i Hit_{t-i}(\alpha) + \sum_{j=p+1}^p \delta_j M_j + \epsilon_t \quad (38)$$

where $M_j = VaR(\alpha)$ Under the null hypothesis of correct model specification the statistic DQ is distributed as chi-square distribution with q degree of freedom:

$$DQ = \left(\text{Hit}_t' X_t [X_t' X_t]^{-1} X_t' \text{Hit}_t \right) / (T\alpha(1 - \alpha)) \sim \chi_q^2 \quad (39)$$

2.4.2. ES backtesting

Kratz et al. (2018) propose a simple approach to backtesting ES based on a multinomial test for different levels of VaR instead of a binomial test for each VaR level. Therefore, this approach is an implicit procedure to test ES, since ES can be approximated by a weighted sum of VaRs. Every single VaR will have (or not) an exception in each day t and the sum of the number of exceptions for the N considered VaR levels at every t day is gathered in variable X_t , with a minimum of zero and a maximum of the N . The test is based on the observed cell counts, O_j , i.e. the number of times that $X_t = 0, X_t = 1, \dots, X_t = N$ along the backtesting, i.e

$$O_j = \sum_{t=1}^{W_T} I_{\{X_t=j\}}, \quad j = 0, 1, \dots, N, \quad (40)$$

where W_T is the size of testing window, and $X_t = \sum_{j=1}^N I_{t,j}$, where $I_{t,j} = I_{\{L_t > \text{VaR}_{\alpha_j,t}\}}$, is, as before, an indicator function that takes value 1 if losses exceed the $\text{VaR}_{\alpha_j,t}$ level. Thus, the vector $(O_0, \dots, O_N)$ follows a multinomial distribution,

$$(O_0, \dots, O_N) \sim MN(W_T, \alpha_1 - \alpha_0, \dots, \alpha_{N+1} - \alpha_N) \quad (41)$$

According to this distribution, Pearson's and Nass's Chi-square tests can be used to determine adequate performance.

Pearson test (Pearson, 1900):

$$S_N = \sum_{j=0}^N \frac{(O_{j+1} - W_T[\alpha_{j+1} - \alpha_j])^2}{W_T(\alpha_{j+1} - \alpha_j)} \sim \chi_N^2 \quad (42)$$

where O_j and $W_T(\alpha_{j+1} - \alpha_j)$ are the observed cell counts and the expected number of exceptions, respectively.

Nass test (Nass, 1959):

$$hS_N \sim \chi_m^2 \quad (43)$$

where $h = \frac{2E(S_N)}{\text{var}(S_N)}$ and $m = hE(S_N)$. This is an enhanced version of the previous test when the cell probabilities are small (Kratz et al., 2018).

For the sake of comparison, we are using both tests and $N = 8$ VaR levels to approximate, which, in general, guarantees a robust performance of the tests. Apart from this novel ES test, it has been considered three brand-new tests in order to have a more comprehensive model performance assessment: Exceedance Residual test (ER) (McNeil and Frey, 2000), Conditional Calibration test (CCa) (Nolde and Ziegel, 2017) and Regression test (ESR) (Bayer and Dimitriadis, 2020) which encompass three different specifications: 'Strict' which include an intercept and slope in the regression, 'Auxiliary' with an auxiliary regression to jointly backtest VaR and ES and 'Intercept' where the slope parameter of the Strict specification is equal to one. One-sided and two-sided alternatives are considered in the results except for Strict and Auxiliary specifications.

2.4.3. MS backtesting

As MS can be directly computed as a VaR – see Eq. (31) – then its backtesting procedures do not differ from VaR backtesting methods described in Section 2.4.1.

2.4.4. Comparison for risk measures

The model's performance for VaR and ES is compared through Diebold-Mariano (DM) tests for pairwise comparisons and Equal Predictive Ability (EPA) tests for MCS.

2.4.4.1. Relative comparison for VaR. For this purpose, scoring function $S = (x, y)$ is considered, which is based on the forecasted (x) and the realized observations (y) and corresponds to a prediction error function that must be minimized. A scoring function for VaR is defined as Gneiting (2011); Fissler et al. (2015); Fissler and Ziegel (2016)

$$S_v(v, x) = (I_{\{x \leq v\}} - \alpha)(\delta(v) - \delta(x)) \quad (44)$$

where δ is a strictly increasing function, v the VaR forecast and x the real return.

DM test (Diebold and Mariano, 1995) is implemented to compare a pair of VaR models depending on each Scoring Function and determines whether the forecasts are significantly different. This comparison considers the absolute difference between VaRs and real observations,

$$d_{ij,t} = S_v^{(i)}(v_t^{(i)}, x_t) - S_v^{(j)}(v_t^{(j)}, x_t) \quad (45)$$

where $d_{ij,t}$ is a loss differential series, $S_v^{(i)}$ and $S_v^{(j)}$ denote the scoring function for each pair (i, j) of VaR models.

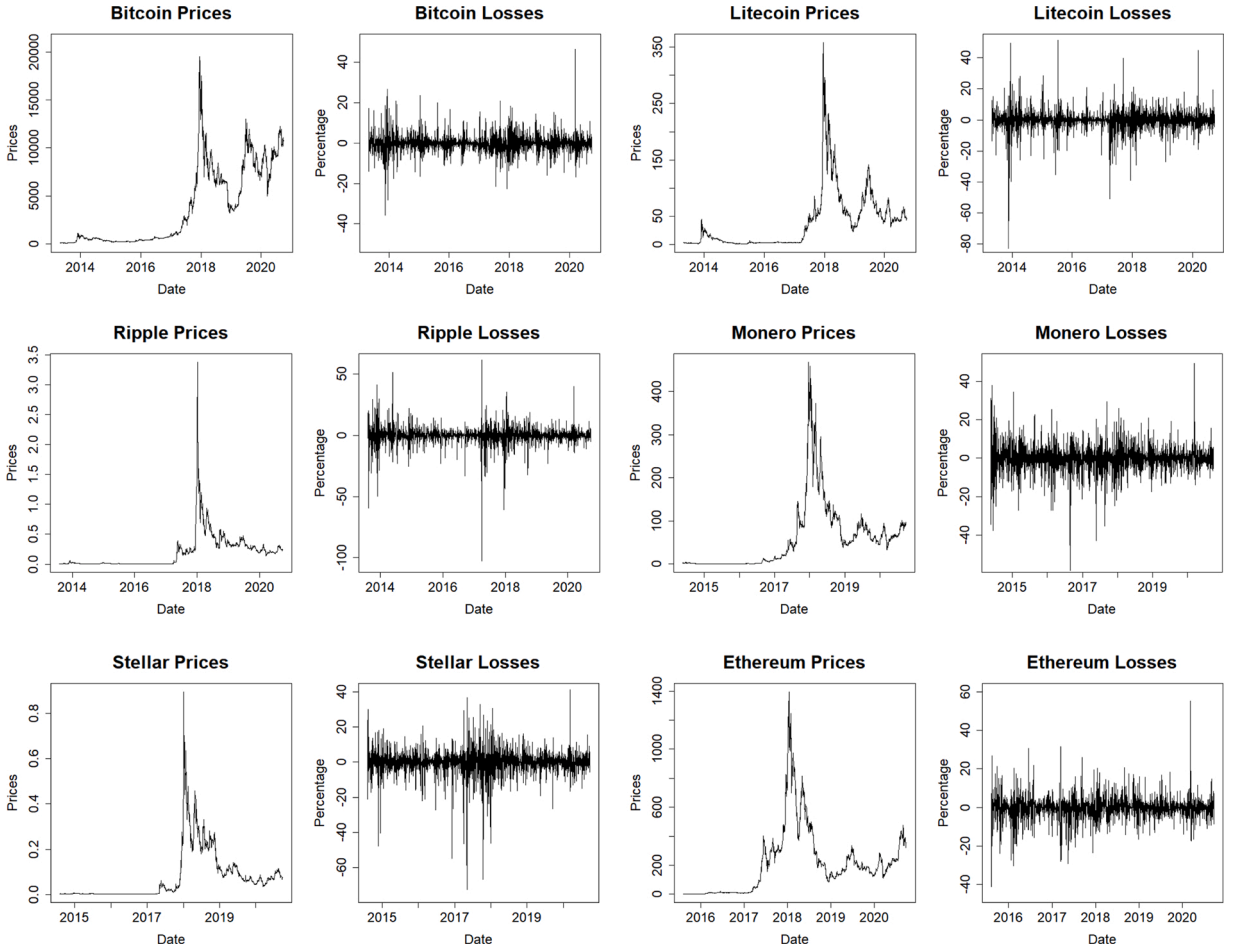


Fig. 2. Cryptocurrencies Prices and Losses.

Cryptocurrency prices and losses for Bitcoin from April 29th, 2013 to September 27th, 2020, Litecoin from April 29th, 2013 to September 27th, 2020, Ripple from August 4th, 2013 to September 27th, 2020, Monero from May 21st, 2014 to September 27th, 2020, Stellar from August 5th, 2014 to September 27th, 2020 and Ethereum from August 8th, 2015 to September 27th, 2020.

The DM test statistic is computed as:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{\text{var}}(\bar{d}_{ij})}} \sim N(0, 1) \quad (46)$$

where $\bar{d}_{ij}$ is the mean of $d_{ij,t}$ and $\widehat{\text{var}}(\bar{d}_{ij})$ its the heteroskedasticity and autocorrelation consistent (HAC) variance. When the null hypothesis, $E[d_{ij}] = 0$ is accepted both the models present statistically the same performance.

2.4.4.2. Relative comparison for ES. For ES, the scoring function takes into account that VaR and ES are jointly elicitable (Fissler et al., 2015):

$$S_{V,E}(v, e, x) = (I_{x \leq v} - \alpha)(\delta_1(v) - \delta_1(x)) + \frac{1}{\alpha} \delta_2(e) I_{x \leq v}(v - x) + \delta_2(e)(e - v) - \varphi_2(e) \quad (47)$$

where x is the observed return, v y e are the forecasted VaR and ES respectively, δ_1 and δ_2 are strictly increasing functions and $\varphi_2(e) = \delta_2(e)$. In this paper, $\delta_1(x) = x$ and $\delta_2(x) = \exp(x)$. To assess the relative performance of the models DM test in Eq. (46) is implemented.

2.4.4.3. Model Confidence Set. Hansen et al. (2011) developed a procedure that is used to compare different models considering scoring functions. Model Confident Set (MCS) consists of a sequence of statistic tests based on an initial set of models defined as $M_{0,\alpha}$ (with m number of models) for a given confidence level α . The best ones, M_α^* , form the Superior Set of Models (SSM) where $M_\alpha^* \leq M_{0,\alpha}$, which is obtained by an iterative elimination rule where the worst model is removed according to the results on the testing of the null

Table 1
Descriptive statistics for cryptocurrency losses.

	Bitcoin	Litecoin	Ripple	Monero	Stellar	Ethereum
No. Obs.	2708	2708	2611	2320	2245	1877
Minimum	-35.745	-82.897	-102.736	-58.464	-72.310	-41.233
Median	-0.176	0.023	0.227	0.000	0.203	0.005
Mean	-0.159	-0.087	-0.143	-0.177	-0.152	-0.328
Maximum	46.473	51.393	61.627	49.421	40.995	55.071
Standard deviation	4.271	6.332	7.028	6.873	7.283	6.261
Variance	18.237	40.096	49.388	47.231	53.037	39.196
Skewness	0.585	-1.498	-1.994	-0.438	-1.884	-0.078
Excess kurtosis	12.119	25.693	31.219	7.599	17.310	8.027

Series of losses for Bitcoin from April 30th, 2013 to September 27th, 2020, Litecoin from April 30th, 2013 to September 27th, 2020, Ripple from August 5th, 2013 to September 27th, 2020, Monero from May 22nd, 2014 to September 27th, 2020, Stellar from August 6th, 2014 to September 27th, 2020 and Ethereum from August 9th, 2015 to September 27th, 2020.

hypotheses of the EPA:

$$\begin{aligned} H_{0,M} &= \mathbb{E}(d_{ij,t}) = 0 \quad \forall i, j = 1, \dots, m \\ H_{1,M} &= \mathbb{E}(d_{ij,t}) \neq 0 \text{ for some } i, j = 1, \dots, m. \end{aligned} \quad (48)$$

where $d_{ij,t} = \ell_{i,t} - \ell_{j,t}$ is the loss differential between models, $i, j = 1, \dots, m$, assumed to have finite and time independent mean. The test is performed in terms of a DM statistic in Eq. (46) but in this case $\bar{d}_{ij} = \frac{1}{m} \sum_{t=1}^m d_{ij,t}$ is the average loss between models, $\widehat{\text{var}}(\bar{d}_{ij})$ is the bootstrapped estimate of its variance (Bernardi and Catania, 2016; Le, 2020). Finally, according to Hansen et al. (2011) the null hypothesis in Eq. (48) fits in with the following statistic:

$$T_M = \max_{i,j \in M} |t_{ij}| \quad (49)$$

3. Results

3.1. Descriptive and in-sample results

The ‘date of birth’ of every crypto is different and most of them are very heterogeneous in terms of features such as market capitalization or volatility. The criteria used to select the six cryptocurrencies in this paper took into consideration two main aspects: Firstly, they were among the top twenty cryptos by market capitalization according to the reference website www.coinmarketcap.com on September 27th, 2020. Secondly, we only considered those with a minimum range of one thousand and eight hundred observations to obtain robust results.

For such series we compute the negative daily returns or the losses as

$$L_t = -100[\ln(P_t) - \ln(P_{t-1})] \quad (50)$$

where L_t is referring to losses, P_t and P_{t-1} are the prices in one day and the day before.

Fig. 2 shows prices in levels and losses in percentage for the six cryptocurrencies for the whole samples.

Descriptive statistics are displayed in Table 1 for the losses – Eq. (50). The data were downloaded from www.coinmarketcap.com where 24/7 daily prices are traded, gathering the following series: Bitcoin from April 29th, 2013 to September 27th, 2020 ($T = 2,708$), Litecoin from April 29th, 2013 to September 27th, 2020 ($T = 2,708$), Ripple from August 4th, 2013 to September 27th, 2020 ($T = 2,611$), Monero from May 21 st, 2014 to September 27th, 2020 ($T = 2,320$), Stellar from August 5th, 2014 to September 27th, 2020 ($T = 2,245$) and Ethereum from August 7th, 2020 to September 27th, 2020 ($T = 1,877$).

High volatility is observed in all cases, with a range of 4.271–7.283, being the lowest for Bitcoin and the highest for Stellar. For comparative purposes, some statistics for the losses of S&P500 (from April 30th, 2013 to September 25th, 2020, ($T = 1,867$), downloaded from www.investing.com) are provided. This asset exhibits a volatility of 1.096, which means around four (six) times less than Bitcoin (Stellar). It must be also highlighted that the minimum and maximum value for losses in each crypto are very high with values from -102.736 to 61.627 being the values of S&P500 smaller, from -8.968 to 12.765. This remarkable feature may be explained by the conceptual definition of what a cryptocurrency is: a decentralized currency where no institution or central bank can control supply, thus its price may vary a high percentage in just one day depending on its demand and market conditions. As a consequence, extreme values are more frequent than in other assets causing a high excess of kurtosis and heavy (and wavy) tails. These issues are defined by the high excess kurtosis and skewness of the series. Regarding excess kurtosis, Ripple and Litecoin are the most prominent with values of 31.219 and 25.693, respectively and Monero and Ethereum have the smallest values with 7.599 and 8.027, respectively. To conclude, the only series which presents positive skewness for losses is Bitcoin. For the rest of cryptocurrencies, the highest values correspond to Litecoin, Ripple and Stellar (-1.498, -1.994 and -1.884).

These salient differences justify the analysis of the density of cryptocurrency losses by means of flexible methodologies capable of capturing the differences between them.

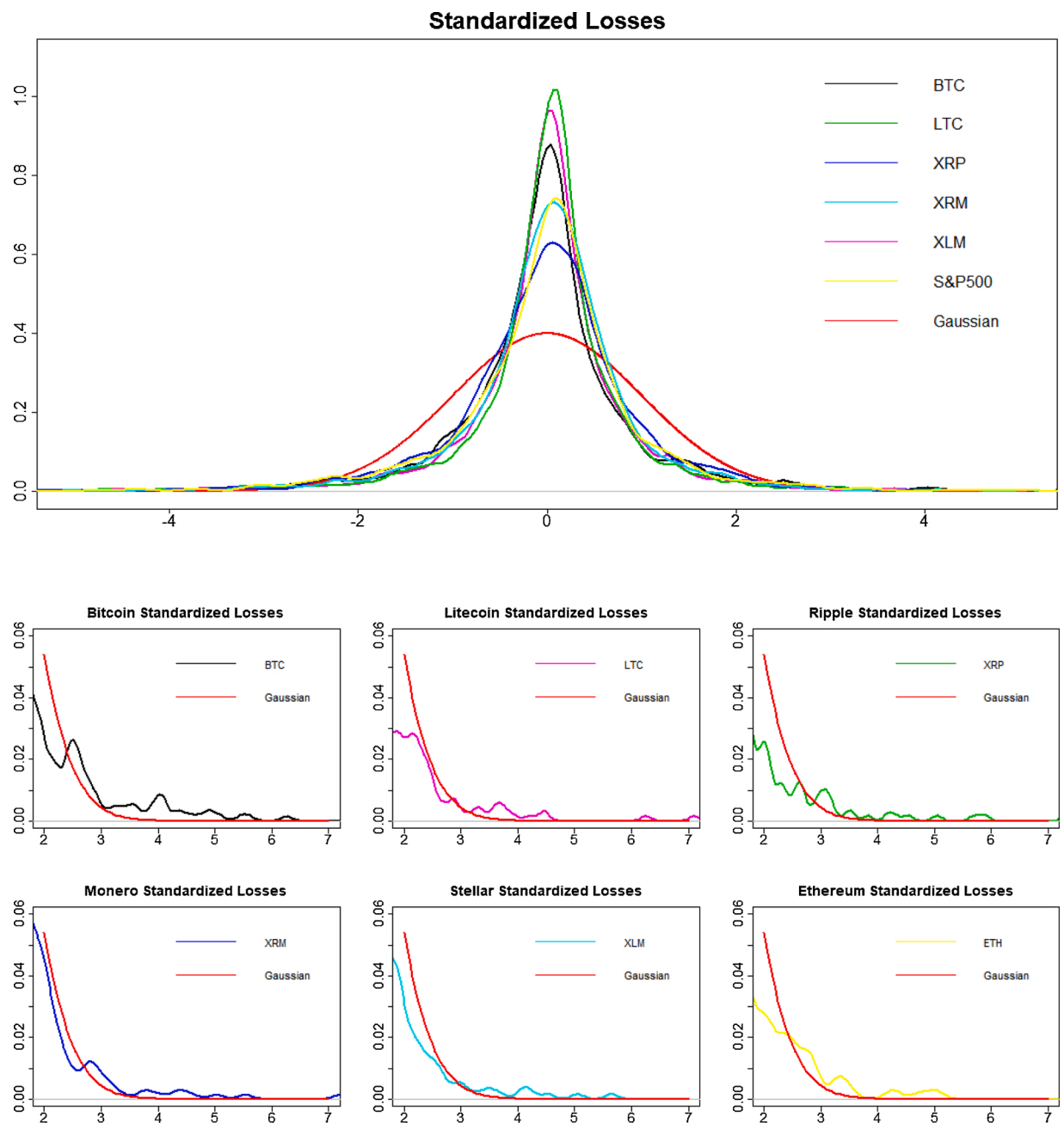


Fig. 3. Standardized Losses for cryptocurrencies.

BTC, LTC, XRP, MON, XLM and ETH refer to Bitcoin, Litecoin, Ripple, Monero, Stellar and Ethereum, respectively. Standardized Losses for all cryptocurrencies with Standard Gaussian distribution as a benchmark for all data and for the right tail of every cryptocurrency.

The non-Gaussian distribution of all these series is illustrated in Fig. 3, which depicts the empirical distributions for standardized losses of the six cryptocurrencies compared to the standard Gaussian distribution. A detail on the distribution right tail is also shown for every series for the entire period. It is noteworthy that the traditional Gaussian assumption clearly underestimates the risk but also that the empirical tail is wavy, which cannot be captured by traditional parametric distributions characterized by a smoothly decreasing pattern.

Table 2 records the Maximum Likelihood (ML) estimates for the parameters of the AR(1)-GJR-GARCH(1,1) model and those of the density specifications (Gaussian, Student's t, Skewed-t, and SNP densities) for each crypto asset. The estimates are computed for the first in-sample window of the backtesting. We set a constant size of 1,000 observations, in order to have the same window size for all cryptocurrencies, since each of them has a different sample length. Although some of the AR parameters are not significant (reflecting

Table 2
Fitted densities for cryptocurrencies.

	BTC	LTC	XRP	XMR	XTM	ETH
<i>Panel A: Gaussian</i>						
ϕ	0.030 (0.802)	0.295 (0.050)	0.316 (0.050)	−0.128 (0.547)	−0.027 (0.870)	−0.369 (0.049)
φ	0.028 (0.467)	−0.101 (0.007)	0.115 (0.010)	0.019 (0.587)	−0.062 (0.123)	0.070 (0.057)
ω	0.971 (0.001)	1.377 (0.000)	5.376 (0.000)	2.981 (0.001)	12.012 (0.002)	2.767 (0.000)
α	0.121 (0.000)	0.067 (0.000)	0.420 (0.000)	0.052 (0.031)	0.071 (0.074)	0.252 (0.000)
β	0.835 (0.000)	0.880 (0.000)	0.459 (0.000)	0.844 (0.000)	0.439 (0.003)	0.720 (0.000)
γ	−0.001 (0.966)	0.064 (0.002)	0.242 (0.020)	0.097 (0.001)	0.389 (0.000)	0.009 (0.855)
<i>Panel B: Student's t</i>						
ϕ	−0.044 (0.544)	0.200 (0.015)	0.257 (0.022)	0.070 (0.670)	0.384 (0.000)	−0.088 (0.529)
φ	−0.023 (0.457)	−0.047 (0.089)	0.086 (0.010)	−0.021 (0.521)	−0.187 (0.000)	−0.004 (0.913)
ω	1.042 (0.019)	0.869 (0.018)	3.529 (0.001)	3.218 (0.009)	11.881 (0.025)	3.084 (0.008)
α	0.224 (0.000)	0.137 (0.000)	0.368 (0.000)	0.138 (0.008)	0.412 (0.029)	0.321 (0.000)
β	0.769 (0.000)	0.861 (0.000)	0.560 (0.000)	0.788 (0.000)	0.421 (0.006)	0.698 (0.000)
γ	0.012 (0.852)	0.001 (0.989)	0.143 (0.204)	0.082 (0.143)	0.331 (0.146)	−0.039 (0.601)
ν	2.885 (0.000)	2.634 (0.000)	3.221 (0.000)	3.870 (0.000)	2.615 (0.000)	3.654 (0.000)
<i>Panel C: Skewed t</i>						
ϕ	0.058 (0.540)	0.244 (0.025)	0.323 (0.017)	−0.164 (0.390)	0.045 (0.725)	−0.376 (0.027)
φ	−0.026 (0.392)	−0.048 (0.087)	0.084 (0.011)	−0.022 (0.502)	−0.183 (0.000)	0.007 (0.833)
ω	1.008 (0.017)	0.872 (0.018)	3.497 (0.001)	2.931 (0.014)	8.597 (0.039)	2.838 (0.008)
α	0.226 (0.000)	0.138 (0.000)	0.370 (0.000)	0.128 (0.011)	0.287 (0.050)	0.307 (0.000)
β	0.771 (0.000)	0.861 (0.000)	0.560 (0.000)	0.803 (0.000)	0.562 (0.000)	0.715 (0.000)
γ	0.005 (0.940)	−0.001 (0.984)	0.135 (0.223)	0.082 (0.127)	0.305 (0.099)	−0.050 (0.484)
ξ	1.068 (0.000)	1.023 (0.000)	1.033 (0.000)	0.900 (0.000)	0.848 (0.000)	0.886 (0.000)
ν	2.911 (0.000)	2.635 (0.000)	3.229 (0.000)	3.880 (0.000)	2.641 (0.000)	3.667 (0.000)
<i>Panel D: SNP c_3, c_4</i>						
c_3	0.035 (0.044)	0.012 (0.516)	0.017 (0.342)	−0.032 (0.058)	−0.076 (0.000)	−0.038 (0.026)
c_4	0.105 (0.000)	0.129 (0.000)	0.096 (0.000)	0.072 (0.000)	0.104 (0.000)	0.071 (0.000)
<i>Panel E: SNP c_3, c_4, c_6, c_8</i>						
c_3	0.037 (0.007)	0.004 (0.765)	0.006 (0.706)	−0.030 (0.050)	−0.060 (0.000)	−0.030 (0.041)
c_4	0.168 (0.000)	0.238 (0.000)	0.144 (0.000)	0.107 (0.000)	0.160 (0.000)	0.100 (0.000)
c_6	0.029 (0.000)	0.045 (0.000)	0.021 (0.000)	0.015 (0.005)	0.025 (0.000)	0.013 (0.004)
c_8	0.004 (0.000)	0.005 (0.000)	0.003 (0.000)	0.002 (0.002)	0.003 (0.000)	0.002 (0.000)

ϕ_0, φ_1 are the parameters of the AR (1) model; $\omega, \alpha, \beta, \gamma$ are the parameters of the GJR-GARCH (1,1) model; ν and ξ are the parameters of the Student's-t and Skewed-t, respectively; c_3, c_4, c_6 and c_8 are the parameters for the semi-nonparametric (SNP) obtained in the second step of the model estimation, where the odd parameter captures skewness and even parameters excess kurtosis and fat tails. P-values in parentheses.

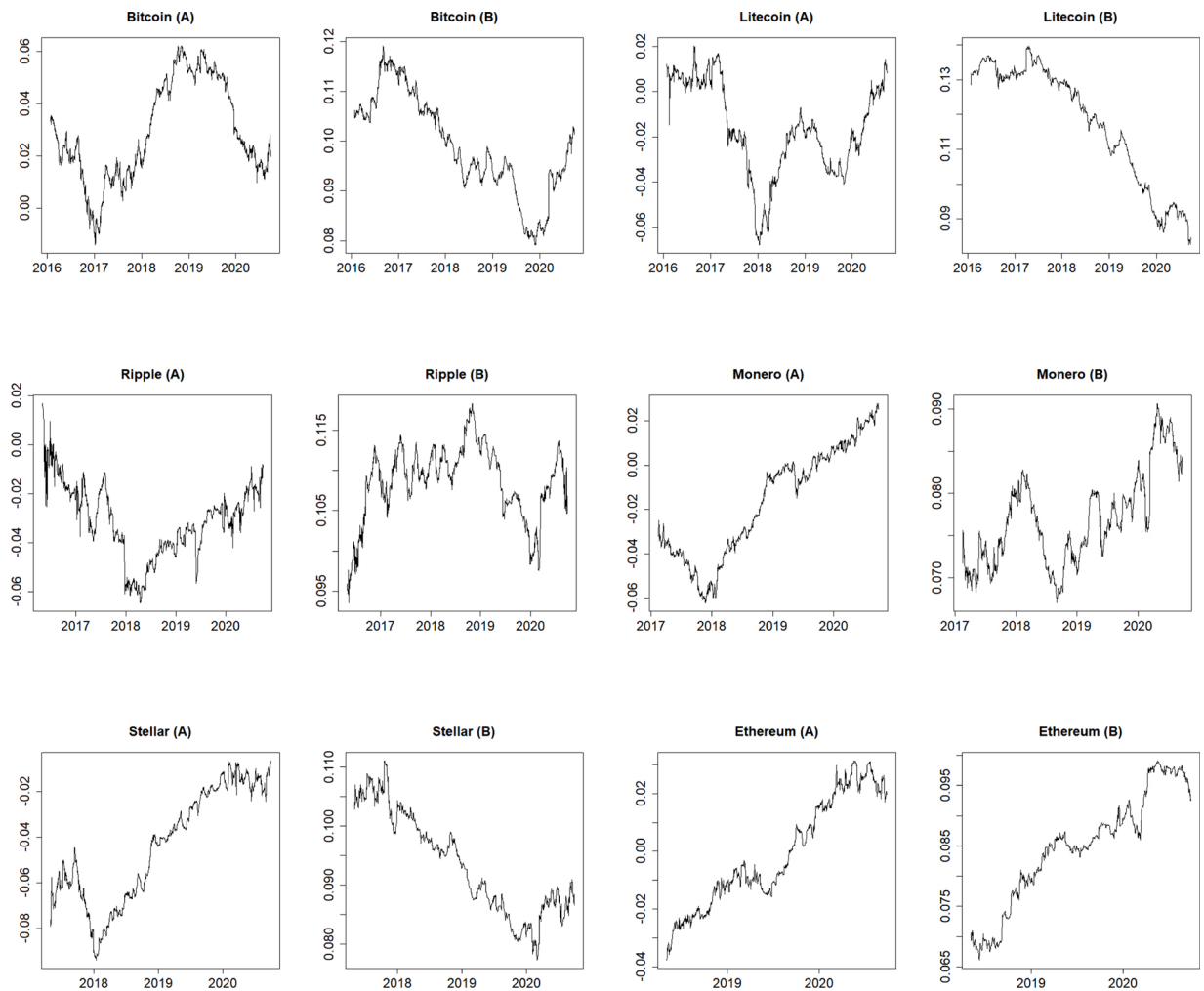


Fig. 4. Dynamics of parameters c_3 and c_4 of the SNP distribution.

Dynamics of parameters c_3 and c_4 of the SNP distribution, accounting for skewness and excess kurtosis, respectively. (A) corresponds to c_3 parameter and (B) corresponds to c_4 parameter.

the small predictable component of the mean of losses) we keep the AR structure since it might be relevant in some estimation window of the backtesting. The GARCH parameters are highly significant. The estimated parameters show the typical volatility persistence behavior in conditional variances and heavy tails in the standardized residuals. Moreover, it seems that the leverage effect is only exhibited when the Gaussian density is fitted to the standardized residuals of the GJR-GARCH specification.

The estimates for the df of Student's t distributions (ν) are significant and quite very low in all the cases, below four and even below three in the case of Bitcoin, Litecoin and Stellar. This is another disadvantage of these parametric distributions since moments above the df seem not to be properly defined, whilst for the SNP density all moments are finite. Note that the low value of this parameter not only captures the thick tails but also the sharp peak at the mean, as illustrated in Fig. 3. The skewness parameter of the Skewed- t (ξ) is significant for all cryptocurrencies and close to one in Bitcoin, Litecoin and Ripple, except in Monero, Stellar and Ethereum with values of 0.900, 0.848, and 0.886 respectively, which means they have a slight negative skewness.

We consider two different specifications of SNP densities to show the relevance of the expansion length, especially to highlight the extreme part of the tails (the main source of risk). The AR-GJR-GARCH and SNP parameters are estimated in two stages, as mentioned above in Section 2.1.1. All the even parameters of the expansion (c_4, c_6, c_8) are positive and significant, reflecting the leptokurtic shape (sharply peaked and with abundant extreme events) and the odd parameters (c_3) are all significant for the shorter expansion evidencing both positive and negative skewness, and for the larger SNP c_3 is significant for all series except for Litecoin and Ripple.

The evolution of the parameters of the two-parameter SNP distribution is depicted in Fig. 4. These dynamics corroborate the fact that the ML estimation of SNP distributions does not require the implementation of positivity constraints.

Table 3
Backtesting 99 %-VaR and 97.5 %-VaR.

		Bitcoin	Litecoin	Ripple	Monero	Stellar	Ethereum
99 %-VaR Panel A							
Gaussian	E.E.	17	17	16	13	12	9
	Exc.	32	25	25	21	24	17
		11.696	4.020	5.030	4.627	9.457	6.795
	CC Test	(0.003)	(0.134)	(0.081)	(0.099)	(0.009)	(0.033)
	DQ Test	21.821	9.534	12.612	11.383	28.476	18.756
Student's t		(0.003)	(0.217)	(0.082)	(0.123)	(0.000)	(0.009)
	Exc.	27	21	19	14	15	12
		5.814	1.370	0.949	0.348	0.861	1.411
	CC Test	(0.055)	(0.504)	(0.622)	(0.840)	(0.650)	(0.494)
	DQ Test	19.274	14.033	5.989	9.534	14.329	12.677
Skewed t		(0.007)	(0.050)	(0.541)	(0.217)	(0.046)	(0.080)
	Exc.	25	29	26	18	23	14
		3.989	9.702	5.774	2.081	8.091	3.123
	CC Test	(0.136)	(0.008)	(0.056)	(0.353)	(0.018)	(0.210)
	DQ Test	16.278	35.011	14.043	11.136	23.213	15.104
SNP c_3, c_4		(0.023)	(0.000)	(0.050)	(0.133)	(0.002)	(0.035)
	Exc.	18	14	15	11	12	12
		0.433	0.829	0.361	0.578	0.250	1.411
	CC Test	(0.805)	(0.661)	(0.835)	(0.749)	(0.882)	(0.494)
	DQ Test	9.310	0.862	7.206	10.094	21.266	13.061
SNP c_3, c_4, c_6, c_8		(0.231)	(0.281)	(0.408)	(0.183)	(0.003)	(0.071)
	Exc.	17	9	11	13	14	13
		0.342	4.762	1.994	0.262	0.506	2.186
	CC Test	(0.843)	(0.092)	(0.369)	(0.874)	(0.776)	(0.335)
	DQ Test	9.391	14.906	11.166	8.774	17.598	13.931
PGC c_2, c_4, c_6		(0.226)	(0.037)	(0.132)	(0.269)	(0.014)	(0.052)
	Exc.	22	17	15	13	15	12
		1.887	0.342	0.361	0.262	0.861	1.411
	CC Test	(0.389)	(0.843)	(0.835)	(0.877)	(0.650)	(0.494)
	DQ Test	9.170	5.903	7.045	9.264	15.871	12.837
		(0.241)	(0.551)	(0.424)	(0.234)	(0.026)	(0.076)
97.5 %-VaR Panel B							
Gaussian	E.E.	43	43	40	33	31	22
	Exc.	54	38	39	35	35	22
		2.884	7.229	0.045	0.128	0.476	1.134
	CC Test	(0.236)	(0.027)	(0.978)	(0.938)	(0.788)	(0.567)
	DQ Test	8.056	20.928	4.054	4.999	6.963	8.091
Student's t		(0.328)	(0.004)	(0.774)	(0.660)	(0.433)	(0.325)
	Exc.	68	59	45	31	39	23
		20.712	13.965	2.422	1.619	1.940	0.284
	CC Test	(0.000)	(0.001)	(0.298)	(0.445)	(0.379)	(0.868)
	DQ Test	43.052	38.640	7.686	6.271	10.105	8.025
Skewed t		(0.000)	(0.000)	(0.361)	(0.508)	(0.183)	(0.330)
	Exc.	64	66	49	43	45	28
		13.708	22.382	3.090	2.977	5.887	2.689
	CC Test	(0.001)	(0.000)	(0.213)	(0.226)	(0.053)	(0.261)
	DQ Test	28.592	61.978	7.564	9.244	15.604	13.638
SNP c_3, c_4		(0.000)	(0.000)	(0.373)	(0.236)	(0.029)	(0.058)
	Exc.	28	27	30	34	35	19
		6.832	7.377	3.240	0.048	2.503	1.261
	CC Test	(0.033)	(0.025)	(0.198)	(0.976)	(0.286)	(0.532)
	DQ Test	8.106	16.201	6.213	4.378	9.718	5.405
SNP c_3, c_4, c_6, c_8		(0.323)	(0.023)	(0.515)	(0.735)	(0.205)	(0.611)
	Exc.	35	31	32	36	33	20
		2.981	3.917	2.059	0.272	1.912	1.113
	CC Test	(0.225)	(0.141)	(0.357)	(0.873)	(0.384)	(0.573)
	DQ Test	4.783	10.807	5.032	3.673	9.895	5.245
PGC c_2, c_4, c_6		(0.686)	(0.147)	(0.656)	(0.817)	(0.195)	(0.630)
	Exc.	44	33	31	28	33	20
		2.369	2.632	2.613	1.060	1.911	1.113
	CC Test	(0.306)	(0.268)	(0.271)	(0.589)	(0.384)	(0.573)
	DQ Test	7.778	13.218	5.580	7.140	7.117	4.986
		(0.353)	(0.067)	(0.590)	(0.414)	(0.417)	(0.662)

E.E. stands for the expected exceptions for each crypto and sample size. The number of observed exceptions (Exc.) per each crypto and model is reported in the table. Conditional Coverage (CC) test and Dynamic Quantile (DQ) (with 4 lags) test are reported in the table with p-value for both tests in parentheses (significance level is 10 %). The table includes the results for two different semi-nonparametric (SNP) expansions and a Positive Gram-Charlier version (PGC).

Table 4
Diebold-Mariano (DM) test for 99 %-VaR.

Model A → Model B ↓	Student's t		Skewed t		SNP c ₃ , c ₄		SNP c ₃ , c ₄ , c ₆ , c ₈		PGC c ₂ , c ₄ , c ₆	
Panel A. Bitcoin										
Gaussian	1.097	(0.864)	0.315	(0.623)	−2.258	(0.012)	−2.301	(0.011)	−2.252	(0.012)
Student's t	−	−	−2.514	(0.006)	−2.046	(0.020)	−2.076	(0.019)	−1.981	(0.024)
Skewed t	−	−	−	−	−1.989	(0.023)	−2.021	(0.022)	−1.899	(0.029)
SNP c ₃ , c ₄	−	−	−	−	−	−	0.495	(0.690)	2.228	(0.987)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	2.350	(0.991)
Panel B. Litecoin										
Gaussian	−1.274	(0.101)	−0.716	(0.237)	−2.308	(0.011)	−2.069	(0.019)	−2.386	(0.009)
Student's t	−	−	2.287	(0.989)	−1.795	(0.036)	−1.711	(0.044)	−0.583	(0.280)
Skewed t	−	−	−	−	−2.578	(0.005)	−2.128	(0.017)	−2.332	(0.010)
SNP c ₃ , c ₄	−	−	−	−	−	−	−1.157	(0.124)	1.996	(0.977)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	1.569	(0.942)
Panel C. Ripple										
Gaussian	−1.957	(0.025)	−1.657	(0.049)	−2.462	(0.007)	−2.310	(0.010)	−2.335	(0.010)
Student's t	−	−	2.458	(0.993)	0.391	(0.652)	−1.242	(0.107)	0.619	(0.732)
Skewed t	−	−	−	−	−0.976	(0.165)	−2.250	(0.012)	−0.675	(0.250)
SNP c ₃ , c ₄	−	−	−	−	−	−	−1.795	(0.036)	1.096	(0.863)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	1.940	(0.974)
Panel D. Monero										
Gaussian	−1.895	(0.029)	−1.399	(0.081)	−2.086	(0.019)	−2.076	(0.019)	−2.288	(0.011)
Student's t	−	−	2.254	(0.988)	1.054	(0.854)	1.417	(0.922)	1.170	(0.879)
Skewed t	−	−	−	−	−1.371	(0.085)	−0.582	(0.280)	−0.365	(0.357)
SNP c ₃ , c ₄	−	−	−	−	−	−	1.617	(0.947)	1.115	(0.868)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	0.073	(0.529)
Panel E. Stellar										
Gaussian	−2.347	(0.010)	−1.572	(0.058)	−2.671	(0.004)	−2.709	(0.003)	−2.908	(0.002)
Student's t	−	−	2.584	(0.995)	1.034	(0.849)	1.121	(0.869)	1.684	(0.954)
Skewed t	−	−	−	−	−1.404	(0.080)	−1.155	(0.124)	−0.256	(0.399)
SNP c ₃ , c ₄	−	−	−	−	−	−	0.751	(0.774)	2.049	(0.980)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	2.044	(0.979)
Panel F. Ethereum										
Gaussian	−1.486	(0.069)	−1.321	(0.093)	−1.717	(0.043)	−1.615	(0.053)	−1.704	(0.044)
Student's t	−	−	2.143	(0.984)	0.406	(0.658)	1.024	(0.847)	0.973	(0.835)
Skewed t	−	−	−	−	−0.288	(0.387)	0.361	(0.641)	0.462	(0.678)
SNP c ₃ , c ₄	−	−	−	−	−	−	2.705	(0.997)	1.778	(0.962)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	0.776	(0.781)

The table shows the Diebold-Mariano statistics and p-values (in parentheses). Bold figures indicate that Model A is preferred than Model B, underlined figures indicate the opposite. Otherwise, both models present the same performance. The null hypothesis is rejected when $|DM| > 1.96$ at 5% significance level.

3.2. Value-at-Risk backtesting results

Our research is focused on both backtesting 99 %-VaR and 97.5 %-VaR as the most widely used measures for risk assessment. For the sake of comparability, we consider the same rolling estimation window estimation size, $W_E = 1,000$ observations, although different testing windows, $W_T = T - W_E$, depending on the total number of observations, T .

Table 3 exhibits two panels, one for each confidence level, 99 %-VaR and 97.5 %-VaR. Backtesting performance for the series is assessed according to the Conditional Coverage test (CC Test), Dynamic Quantil (DQ Test) and the number of observed exceptions (Exc).

Expected Exceptions (E.E) depend on the confidence level (97.5-% and 99 %) and the size of W_T and thus is reported for each crypto in Table 3. The closer the observed exceptions to the E.E. the better the model. It seems that, unlike the parametric specifications that tend to overestimate the exceptions, the SNP specifications tend to underestimate them in the majority of cases. This point to the fact that SNP seems to provide more conservative risk measures.

For most cryptocurrencies, both CC and DQ tests exhibit similar results thus we mainly focus the explanations on the former. For 99 %-VaR (Panel A). Gaussian distribution is rejected for all cryptos except Litecoin whereas Student's t cannot be rejected with the exception of Bitcoin. Skewed-t is rejected for Litecoin, Ripple and Stellar. Both SNP expansions and the PGC have good performances

Table 5
Diebold-Mariano (DM) test for 97.5 %-VaR.

Model A → Model B ↓	Student' t		Skewed t		SNP c ₃ , c ₄		SNP c ₃ , c ₄ , c ₆ , c ₈		PGC c ₂ , c ₄ , c ₆	
Panel A. Bitcoin										
Gaussian	<u>2.662</u>	(0.996)	<u>2.442</u>	(0.993)	−3.010	(0.001)	−3.335	(0.000)	−3.161	(0.001)
Student's t	−	−	−3.863	−	−2.896	(0.002)	−2.977	(0.001)	−2.852	(0.002)
Skewed t	−	−	−	−	−2.808	(0.003)	−2.872	(0.002)	−2.718	(0.003)
SNP c ₃ , c ₄	−	−	−	−	−	−	2.409	(0.992)	<u>2.887</u>	(0.998)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	<u>3.475</u>	(0.999)
Panel B. Litecoin										
Gaussian	<u>2.534</u>	(0.994)	<u>3.859</u>	(0.999)	−2.622	(0.004)	−2.989	(0.001)	−2.894	(0.002)
Student's t	−	−	<u>3.647</u>	(0.999)	−3.224	(0.001)	−3.475	(0.000)	−3.664	(0.000)
Skewed t	−	−	−	−	−3.557	(0.000)	−4.005	(0.000)	−4.073	(0.000)
SNP c ₃ , c ₄	−	−	−	−	−	−	<u>2.216</u>	(0.987)	<u>2.094</u>	(0.982)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−1.046	(0.148)
Panel C. Ripple										
Gaussian	−0.191	(0.424)	1.321	(0.907)	−2.440	(0.007)	−2.545	(0.006)	−2.687	(0.004)
Student's t	−	−	<u>3.112</u>	(0.999)	−1.196	(0.116)	−0.539	(0.295)	−1.543	(0.061)
Skewed t	−	−	−	−	−2.247	(0.012)	−1.860	(0.032)	−2.418	(0.008)
SNP c ₃ , c ₄	−	−	−	−	−	−	<u>2.272</u>	(0.988)	−0.880	(0.190)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−2.456	(0.007)
Panel D. Monero										
Gaussian	−1.279	(0.101)	1.011	(0.844)	−1.581	(0.057)	−0.810	(0.209)	−2.789	(0.003)
Student's t	−	−	<u>3.067</u>	(0.999)	−0.496	(0.310)	0.736	(0.769)	−1.080	(0.140)
Skewed t	−	−	−	−	−3.178	(0.001)	−2.405	(0.008)	−2.887	(0.002)
SNP c ₃ , c ₄	−	−	−	−	−	−	<u>3.115</u>	(0.999)	−0.477	(0.317)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−1.921	(0.027)
Panel E. Stellar										
Gaussian	−1.047	(0.148)	0.897	(0.815)	0.009	(0.503)	−0.606	(0.272)	−3.271	(0.001)
Student's t	−	−	<u>3.455</u>	(0.999)	1.067	(0.857)	0.885	(0.812)	−0.170	(0.432)
Skewed t	−	−	−	−	−0.940	(0.174)	−1.076	(0.141)	−1.940	(0.026)
SNP c ₃ , c ₄	−	−	−	−	−	−	−0.958	(0.169)	−2.605	(0.005)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−2.876	(0.002)
Panel F. Ethereum										
Gaussian	−0.660	(0.255)	0.154	(0.561)	−1.659	(0.049)	−1.783	(0.037)	−2.246	(0.012)
Student's t	−	−	<u>2.918</u>	(0.998)	−2.197	(0.014)	−2.699	(0.004)	−3.163	(0.001)
Skewed t	−	−	−	−	−2.671	(0.004)	−3.223	(0.001)	−3.305	(0.000)
SNP c ₃ , c ₄	−	−	−	−	−	−	1.376	(0.915)	1.023	(0.847)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	0.468	(0.680)

The table shows the Diebold-Mariano statistics and p-values (in parentheses). Bold figures indicate that Model A is preferred than Model B, underlined figures indicate the opposite. Otherwise, both models present the same performance. The null hypothesis is rejected when $|DM| > 1.96$ at 5% significance level.

except for the larger expansion for Litecoin that is almost accepted at 10 % confidence level (0.092). It seems that regardless the election GC distribution (Type A or Positive), both perform adequately to model 99 %-VaR. This feature also points to the fact that with controlled optimization implementing positive constraints is not mandatory when estimating SNP models, even in backtesting procedures.

Nevertheless, this evidence is not that clear for 97.5 %-VaR (Panel B) since all cryptocurrencies, except Litecoin, the normality assumption cannot be rejected. This means that the distribution seems to have higher impact on the VaR accuracy measures the higher confidence levels, particularly for those levels recommended by the regulators. Student's t, Skewed t and the shorter expansion tend to underestimate risk in Bitcoin and Litecoin. However, larger SNP and PGC still having a good performance for 97.5 %-VaR for all cryptos.

In general, results in Table 3 support our conjecture on the outperformance of the flexible distributions, in particular, all SNP versions for 99 %-VaR and the large SNP and PGC for 97.5 %-VaR. This evidence is consistent with the findings in Jiménez et al. (2020) for Bitcoin. Nonetheless, the behavior within the cryptocurrencies seems to be heterogeneous and Bitcoin exhibits some peculiarities with respect to the rest of cryptocurrencies or altcoins. This observation is more evident after considering MS results in Section 3.4.

Relative performance for pairwise comparisons through the DM test is displayed in Table 4. For all cryptocurrencies except for Ethereum both SNP models outperform the other parametric alternatives. It seems that Bitcoin have different results from the rest of

Table 6

Comparison for the SSM for 99 %-VaR and 97.5 %-VaR.

	99 %-VaR					97.5 %-VaR			
	Rank _M	t _{ij}	P-value _M	Loss		Rank _M	t _{ij}	P-value _M	Loss
<i>Bitcoin: 0 eliminations</i>					<i>Bitcoin: 0 eliminations</i>				
Student's t	5	1.986	(0.098)	0.168	6	1.308	(0.329)		0.305
Normal	4	0.514	(0.857)	0.164	2	-1.851	(1.000)		0.296
Skewed t	6	2.018	(0.092)	0.167	5	1.218	(0.379)		0.304
SNP c ₃ , c ₄	2	-1.555	(1.000)	0.158	4	0.465	(0.855)		0.301
SNP c ₃ , c ₄ , c ₆ , c ₈	3	-1.041	(1.000)	0.159	3	-1.287	(1.000)		0.295
PGC c ₂ , c ₄ , c ₆	1	-3.291	(1.000)	0.158	1	-3.272	(1.000)		0.293
<i>Litecoin: 0 eliminations</i>					<i>Litecoin: 0 eliminations</i>				
Student's t	4	-0.067	(1.000)	0.208	4	-0.134	(1.000)		0.389
Normal	2	-0.652	(1.000)	0.205	1	-4.785	(1.000)		0.379
Skewed t	5	0.533	(0.872)	0.211	5	0.701	(0.698)		0.396
SNP c ₃ , c ₄	3	-0.379	(1.000)	0.206	6	1.688	(0.159)		0.401
SNP c ₃ , c ₄ , c ₆ , c ₈	6	1.144	(0.462)	0.215	3	-0.378	(1.000)		0.389
PGC c ₂ , c ₄ , c ₆	1	-2.784	(1.000)	0.203	2	-0.993	(1.000)		0.387
<i>Ripple: 0 eliminations</i>					<i>Ripple: 3 eliminations</i>				
Student's t	1	-3.447	(1.000)	0.209	1	-3.138	(1.000)		0.392
Normal	3	0.081	(1.000)	0.219	3	2.344	(0.029)		0.406
Skewed t	2	-1.811	(1.000)	0.212	2	-1.352	(1.000)		0.393
SNP c ₃ , c ₄	6	2.423	(0.036)	0.224					
SNP c ₃ , c ₄ , c ₆ , c ₈	5	1.597	(0.234)	0.226					
PGC c ₂ , c ₄ , c ₆	4	0.95	(0.626)	0.22					
<i>Monero: 0 eliminations</i>					<i>Monero: 0 eliminations</i>				
Student's t	4	0.131	(1.000)	0.229	2	-0.502	(1.000)		0.425
Normal	3	0.05	(1.000)	0.229	1	-2.407	(1.000)		0.423
Skewed t	6	0.458	(0.950)	0.23	4	-0.078	(1.000)		0.427
SNP c ₃ , c ₄	5	0.29	(0.990)	0.23	6	1.399	(0.383)		0.432
SNP c ₃ , c ₄ , c ₆ , c ₈	1	-1.093	(1.000)	0.227	5	1.22	(0.497)		0.431
PGC c ₂ , c ₄ , c ₆	2	-0.567	(1.000)	0.228	3	-0.401	(1.000)		0.426
<i>Stellar: 0 eliminations</i>					<i>Stellar: 0 eliminations</i>				
Student's t	3	-0.178	(1.000)	0.266	3	-0.388	(1.000)		0.509
Normal	6	0.633	(0.848)	0.273	1	-0.689	(1.000)		0.51
Skewed t	5	0.44	(0.940)	0.271	4	0.306	(0.956)		0.517
SNP c ₃ , c ₄	4	0.126	(1.000)	0.268	6	0.784	(0.669)		0.518
SNP c ₃ , c ₄ , c ₆ , c ₈	2	-1.031	(1.000)	0.264	5	0.472	(0.876)		0.516
PGC c ₂ , c ₄ , c ₆	1	-1.355	(1.000)	0.263	2	-0.6	(1.000)		0.51
<i>Ethereum: 0 eliminations</i>					<i>Ethereum: 0 eliminations</i>				
Student's t	3	-0.741	(1.000)	0.242	2	-0.393	(1.000)		0.422
Normal	6	0.805	(0.716)	0.249	1	-1.983	(1.000)		0.418
Skewed t	5	0.456	(0.931)	0.245	4	0.031	(1.000)		0.424
SNP c ₃ , c ₄	2	-0.973	(1.000)	0.241	6	1.457	(0.267)		0.429
SNP c ₃ , c ₄ , c ₆ , c ₈	4	0.434	(0.940)	0.244	5	0.448	(0.903)		0.425
PGC c ₂ , c ₄ , c ₆	1	-2.056	(1.000)	0.241	3	-0.368	(1.000)		0.423

Rank_M refers to the ranking of the models, t_{ij} stands for the Diebold-Mariano test statistic for two models and p-value are in parenthesis. P-value_M is the minimum p-value for belonging to the Superior Set Model (SSM). The number of eliminations corresponds to the models that do not belong to the SSM. Loss is the average loss for each model.

cryptocurrencies for 99 %-VaR, since the SNP expansions have good performance for Bitcoin in comparison to other distributions whereas for altcoins this result is not that clear, only being superior to Student's t and the Skewed t for Litecoin and Ripple.

Table 5 presents the DM test for 97.5 %-VaR where Bitcoin and Litecoin seems to have a similar performance, with the SNP expansion being superior than the parametric alternatives. In contrast to 99 %-VaR, the PGC outperforms in all cryptocurrencies not only for the Gaussian but also for the Student's t and Skewed t. Mostly, SNP models are better than Skewed t.

The results of applying the MCS, for both levels of VaR, in order to obtain the Superior Set Models (SSM) are displayed in Table 6. It is noticeable that there are hardly any distributions eliminated, only in Ethereum for 97.5 %-VaR for all the SNP distribution. The explanation of the scarcity of distributions eliminated indicates that the models are statistically equivalent in terms of their VaR forecasting ability, in other words, SSM for each cryptocurrency shows homogeneity concerning the conditional distributions. P-values

Table 7

Multinomial test for 97.5 %-ES Backtesting N = 8.

	O0	O1	O2	O3	O4	O5	O6	O7	O8	Pearson	Nass
<i>Panel A: Gaussian</i>											
Bitcoin	1654	4	4	3	7	6	1	7	22	58.383	53.529
Litecoin	1670	2	3	2	3	3	4	3	18	38.644	35.449
Ripple	1572	4	3	2	2	3	2	8	15	28.818	26.304
Monero	1285	1	5	5	1	4	4	3	12	20.455	18.318
Stellar	1210	2	0	3	4	4	5	4	13	26.667	23.733
Ethereum	855	1	1	1	1	1	3	1	13	45.063	38.334
<i>Panel B: Student's t</i>											
Bitcoin	1640	8	11	6	12	5	8	9	9	22.110	20.282
Litecoin	1649	7	7	7	9	10	5	5	9	10.695	9.811
Ripple	1566	4	6	6	3	7	6	9	4	5.694	5.197
Monero	1289	6	1	4	3	6	3	5	3	5.182	4.641
Stellar	1206	5	3	5	5	6	6	5	4	3.760	3.346
Ethereum	854	3	1	2	2	4	2	6	3	6.210	5.283
<i>Panel C: Skewed t</i>											
Bitcoin	1644	10	7	9	8	6	8	7	9	12.873	11.809
Litecoin	1642	5	7	7	9	12	10	5	11	21.988	20.170
Ripple	1562	4	4	7	6	4	9	9	6	8.023	7.323
Monero	1277	7	6	6	3	3	7	6	5	7.364	6.595
Stellar	1200	2	2	5	9	5	7	9	6	19.519	17.371
Ethereum	849	5	3	2	2	3	3	4	6	6.792	5.777
<i>Panel D: SNP c_3, c_4</i>											
Bitcoin	1680	3	2	1	2	2	4	5	9	13.679	12.548
Litecoin	1681	3	5	1	2	2	4	4	6	9.496	8.711
Ripple	1581	3	4	6	2	0	3	6	6	9.276	8.467
Monero	1286	10	4	4	2	3	1	3	7	14.455	12.945
Stellar	1210	5	6	3	6	4	2	3	6	5.077	4.518
Ethereum	858	1	3	1	1	1	2	1	9	20.048	17.055
<i>Panel E: SNP c_3, c_4, c_6, c_8</i>											
Bitcoin	1673	5	2	3	3	7	5	3	7	6.236	5.721
Litecoin	1677	2	3	6	4	7	3	4	2	7.492	6.872
Ripple	1579	2	2	4	8	7	3	0	6	12.426	11.343
Monero	1284	2	6	7	2	6	4	6	3	7.061	6.323
Stellar	1212	2	3	4	4	7	5	3	5	4.450	3.961
Ethereum	857	1	2	1	3	1	2	6	4	8.196	6.973
<i>Panel F: PGC c_2, c_4, c_6</i>											
Bitcoin	1664	6	6	3	4	3	8	8	6	5.285	4.848
Litecoin	1675	4	3	1	3	5	5	10	2	12.110	11.109
Ripple	1580	1	6	1	7	1	9	1	5	17.009	15.525
Monero	1292	4	1	4	3	4	2	7	3	6.091	5.455
Stellar	1212	3	4	3	5	3	4	5	6	2.394	2.131
Ethereum	857	2	1	1	2	2	2	6	4	7.467	6.352

O_j shows the result of observed cells, which count how many times each VaR has j observed exceptions $j = 0, 1 \dots N$. Critical value for Pearson Statistic is 15.507 for all cryptocurrencies. Critical value for Nass test is 14.832 for Bitcoin and Litecoin, 14.806 for Ripple, 14.714 for Monero, 14.686 for Stellar and 14.515 for Ethereum.

are the minimum for belonging to the SSM, see more details about the interpretation of p-values (Hansen et al., 2011). The first column reports the ranking over the model where PGC is most of the times, the SNP alternatives being mostly in the first positions.

3.3. Expected Shortfall backtesting results

The sensitivity of the results on VaR backtesting shreds of evidence that a measure that averages the performance at the tail (once VaR has been exceeded), like ES, might capture more accurately the risk than just reporting the VaR for an arbitrary probability. Consequently, it is presented the 97.5 %-ES backtesting, which is the standard risk measure for replacing 99 %-VaR. In order to show a more comprehensive analysis, it has been considered several tests depending on the way of computing ES. Table 7 depicts the results for the approximation by Emmer et al. (2015) in terms of averaged VaR levels – see Eq. (30), which is the basis for the Kratz et al. (2018) proposal for a multinomial test to backtest ES, and Table 8 includes the results for the closed formulas for computing ES – e.g., Eq. (28) for SNP density, hereafter referred as ‘direct ES’.

For verification purposes, in Table 7, two multinomial tests are considered: Pearson’s and Nass’ tests – see Eqs. (42) and (43), which

Table 8

Test for 97.5 %-ES Backtesting.

	Model	ER		CCa		ESR			
		one	two	one	two	Str. two	Aux. two	Int.	
								one	two
Bitcoin	Gaussian	(0.001)	(0.002)	(0.016)	(0.037)	(0.008)	(0.008)	(0.009)	(0.018)
	Student's	(0.627)	(0.908)	(0.312)	(0.007)	(0.064)	(0.026)	(0.184)	(0.368)
	Skewed t	(0.685)	(0.772)	(0.540)	(0.022)	(0.043)	(0.037)	(0.269)	(0.537)
	SNP c_3, c_4	(0.018)	(0.026)	(0.008)	(0.001)	(0.320)	(0.304)	(0.418)	(0.837)
	SNP c_3, c_4, c_6, c_8	(0.075)	(0.190)	(0.283)	(0.255)	(0.339)	(0.351)	(0.491)	(0.982)
	PGC c_2, c_4, c_6	(0.117)	(0.258)	(0.740)	(0.779)	(0.266)	(0.276)	(0.286)	(0.573)
Litecoin	Gaussian	(0.000)	(0.000)	(0.060)	(0.005)	(0.057)	(0.055)	(0.041)	(0.082)
	Student's	(0.841)	(0.318)	(1.000)	(0.020)	(0.032)	(0.039)	(0.696)	(0.607)
	Skewed t	(0.767)	(0.513)	(0.623)	(0.007)	(0.045)	(0.031)	(0.440)	(0.880)
	SNP c_3, c_4	(0.019)	(0.073)	(0.004)	(0.001)	(0.309)	(0.385)	(0.790)	(0.419)
	SNP c_3, c_4, c_6, c_8	(0.496)	(0.829)	(0.051)	(0.106)	(0.170)	(0.172)	(0.918)	(0.165)
	PGC c_2, c_4, c_6	(0.178)	(0.335)	(0.133)	(0.167)	(0.376)	(0.350)	(0.720)	(0.560)
Ripple	Gaussian	(0.002)	(0.002)	(0.082)	(0.037)	(0.026)	(0.029)	(0.342)	(0.683)
	Student's	(0.738)	(0.571)	(1.000)	(0.119)	(0.753)	(0.787)	(0.704)	(0.591)
	Skewed t	(0.666)	(0.756)	(1.000)	(0.217)	(0.408)	(0.422)	(0.584)	(0.831)
	SNP c_3, c_4	(0.009)	(0.050)	(0.088)	(0.081)	(0.497)	(0.359)	(0.687)	(0.627)
	SNP c_3, c_4, c_6, c_8	(0.265)	(0.486)	(0.210)	(0.324)	(0.654)	(0.625)	(0.759)	(0.482)
	PGC c_2, c_4, c_6	(0.086)	(0.190)	(0.139)	(0.241)	(0.716)	(0.593)	(0.727)	(0.546)
Monero	Gaussian	(0.002)	(0.002)	(0.086)	(0.113)	(0.065)	(0.062)	(0.078)	(0.156)
	Student's	(0.559)	(0.920)	(1.000)	(0.815)	(0.478)	(0.506)	(0.782)	(0.436)
	Skewed t	(0.649)	(0.837)	(0.983)	(0.265)	(0.351)	(0.370)	(0.412)	(0.825)
	SNP c_3, c_4	(0.303)	(0.527)	(0.843)	(0.835)	(0.473)	(0.411)	(0.411)	(0.822)
	SNP c_3, c_4, c_6, c_8	(0.577)	(0.985)	(1.000)	(0.877)	(0.571)	(0.499)	(0.480)	(0.960)
	PGC c_2, c_4, c_6	(0.132)	(0.277)	(0.510)	(0.452)	(0.552)	(0.579)	(0.531)	(0.937)
Stellar	Gaussian	(0.000)	(0.000)	(0.033)	(0.040)	(0.024)	(0.024)	(0.217)	(0.433)
	Student's	(0.811)	(0.407)	(1.000)	(0.035)	(0.683)	(0.714)	(0.756)	(0.489)
	Skewed t	(0.414)	(0.772)	(0.594)	(0.081)	(0.245)	(0.234)	(0.554)	(0.893)
	SNP c_3, c_4	(0.183)	(0.365)	(0.701)	(0.745)	(0.386)	(0.314)	(0.508)	(0.984)
	SNP c_3, c_4, c_6, c_8	(0.288)	(0.527)	(0.944)	(0.922)	(0.499)	(0.439)	(0.566)	(0.869)
	PGC c_2, c_4, c_6	(0.115)	(0.248)	(0.838)	(0.842)	(0.397)	(0.372)	(0.513)	(0.974)
Ethereum	Gaussian	(0.012)	(0.012)	(0.067)	(0.045)	(0.061)	(0.060)	(0.060)	(0.119)
	Student's	(0.259)	(0.471)	(0.885)	(0.967)	(0.811)	(0.464)	(0.517)	(0.966)
	Skewed t	(0.344)	(0.596)	(0.911)	(0.499)	(0.631)	(0.658)	(0.433)	(0.865)
	SNP c_3, c_4	(0.016)	(0.024)	(0.373)	(0.130)	(0.487)	(0.482)	(0.289)	(0.579)
	SNP c_3, c_4, c_6, c_8	(0.015)	(0.039)	(0.493)	(0.284)	(0.508)	(0.503)	(0.301)	(0.602)
	PGC c_2, c_4, c_6	(0.014)	(0.036)	(0.481)	(0.282)	(0.480)	(0.482)	(0.286)	(0.572)

Exceedance Residuals backtesting for ES (ER) tests the null hypothesis that ER (i.e. residuals where an exception occurs) are *i.i.d.* with zero mean. In Conditional Calibration backtest (CCa) the sequences of ES predictions are conditionally calibrated comparing different models after the model is deemed as a correct specification. Regression based on ES (ESR) is displayed for three specifications: Strict ESR (Str.), Auxiliary ESR (Aux.) and Intercept ESR (Int.), the null hypothesis being correct specification of ES. Both one-sided (one) and two-sided (two) statistics are displayed. P-values in parentheses (significance level is 10 % in bold).

suggest that the model is adequate if the test results are below their critical values reported in Table 7. Apart from both tests, the cell counts O_j are shown in order to provide more comprehensive results –see explanations in Section 2.4.2. The table presents six panels, one for each distribution (Gaussian, Student's t, Skewed-t, SNP c_3, c_4 and SNP c_3, c_4, c_6, c_8 , PGC c_2, c_4, c_6). According to both tests, the most flexible distributions outperform Gaussian and Student's t for Bitcoin, but Student's t cannot be rejected either for the rest of the series. However, the SNP distributions, either the 'large' or 'short' version, as well as the PGC, cannot be rejected for most of the series.

Three different specifications are shown in Table 8 for 'direct 97.5 %-ES', as it was mentioned above. In those tests one-sided and two-sided alternatives are considered. Unlike multinomial test, the performance is quite similar, since the Gaussian distribution is rejected for all cryptocurrencies for the three tests with the exception of the Regression test (ESR) for the Intercept alternative in some cases. For Students' t and Skewed t the performance is similar to the multinomial test, following identical pattern, accepting (rejecting) the same tests. For Monero and Stellar both SNP and PGC are accepted for all tests and these three conditional distributions have a similar performance to the rest of cryptocurrencies but with some rejections depending on the test, but for all cases Regression test (ESR) is accepted.

The relative performance for 97.5 %-ES is displayed in Table 9. Pairwise model comparison exhibits the good performance for the PGC for all cryptos except for Bitcoin. The shorter and the larger SNP outperform the Gaussian for Litecoin and Ripple. It seems that SNP c_3, c_4 has better result than Skewed t for all cryptos with the exception of Stellar that only have good results with the PGC. Similar

Table 9
Diebold-Mariano (DM) test for 97.5 %-ES.

Model A → Model B ↓	Student' t		Skewed t		SNP c ₃ , c ₄		SNP c ₃ , c ₄ , c ₆ , c ₈		PGC c ₂ , c ₄ , c ₆	
Panel A. Bitcoin										
Gaussian	1.700	(0.955)	<u>2.080</u>	(0.981)	−1.943	(0.026)	−1.744	(0.041)	−1.587	(0.056)
Student's t	−	−	−1.314	(0.095)	−1.830	(0.034)	−1.731	(0.042)	−1.657	(0.049)
Skewed t	−	−	−	−	−2.034	(0.021)	−1.906	(0.028)	−1.810	(0.035)
SNP c ₃ , c ₄	−	−	−	−	−	−	<u>2.392</u>	(0.992)	<u>3.074</u>	(0.999)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	<u>3.046</u>	(0.999)
Panel B. Litecoin										
Gaussian	<u>2.598</u>	(0.995)	<u>3.883</u>	(0.999)	−2.638	(0.004)	−3.020	(0.001)	−2.921	(0.002)
Student's t	−	−	<u>3.809</u>	(0.999)	−3.275	(0.001)	−3.508	(0.000)	−3.689	(0.000)
Skewed t	−	−	−	−	−3.623	(0.000)	−4.048	(0.000)	−4.116	(0.000)
SNP c ₃ , c ₄	−	−	−	−	−	−	<u>2.216</u>	(0.987)	<u>2.096</u>	(0.982)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−1.031	(0.151)
Panel C. Ripple										
Gaussian	−0.240	(0.405)	1.292	(0.902)	−2.477	(0.007)	−2.612	(0.005)	−2.732	(0.003)
Student's t	−	−	<u>3.140</u>	(0.999)	−1.202	(0.115)	−0.547	(0.292)	−1.550	(0.061)
Skewed t	−	−	−	−	−2.261	(0.012)	−1.879	(0.030)	−2.432	(0.008)
SNP c ₃ , c ₄	−	−	−	−	−	−	<u>2.270</u>	(0.988)	−0.879	(0.190)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−2.452	(0.007)
Panel D. Monero										
Gaussian	−1.286	(0.099)	0.995	(0.840)	−1.586	(0.057)	−0.823	(0.205)	−2.785	(0.003)
Student's t	−	−	<u>3.070</u>	(0.999)	−0.498	(0.309)	0.733	(0.768)	−1.081	(0.140)
Skewed t	−	−	−	−	−3.178	(0.001)	−2.407	(0.008)	−2.887	(0.002)
SNP c ₃ , c ₄	−	−	−	−	−	−	<u>3.115</u>	(0.999)	−0.477	(0.317)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−1.920	(0.028)
Panel E. Stellar										
Gaussian	−1.049	(0.147)	0.899	(0.815)	−0.001	(0.500)	−0.616	(0.269)	−3.271	(0.001)
Student's t	−	−	<u>3.461</u>	(1.000)	1.066	(0.857)	0.884	(0.812)	−0.171	(0.432)
Skewed t	−	−	−	−	−0.945	(0.172)	−1.081	(0.140)	−1.944	(0.026)
SNP c ₃ , c ₄	−	−	−	−	−	−	−0.958	(0.169)	−2.605	(0.005)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−2.876	(0.002)
Panel F. Ethereum										
Gaussian	−0.721	(0.236)	0.112	(0.545)	−1.681	(0.047)	−1.814	(0.035)	−2.280	(0.011)
Student's t	−	−	<u>2.907</u>	(0.998)	−2.194	(0.014)	−2.695	(0.004)	−3.160	(0.001)
Skewed t	−	−	−	−	−2.675	(0.004)	−3.224	(0.001)	−3.301	(0.001)
SNP c ₃ , c ₄	−	−	−	−	−	−	1.375	(0.915)	1.020	(0.846)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	0.462	(0.678)

The table shows the Diebold-Mariano statistics and p-values (in parentheses). Bold figures indicate that Model A is preferred than Model B, underlined figures indicate the opposite. Otherwise, both models present the same performance. The null hypothesis is rejected when $|DM| > 1.96$ at 5% significance level.

results show approximation 97.5 %-ES but it is not reported to save space although they are available upon request.

Table 10 depicts the comparison for the SSM for 97.5 %-ES. The only cryptocurrency with eliminated models is Ripple, where SSM is formed by two models corresponding to Student's t and Skewed t. This evidences the high heterogeneity of the models in Ripple. For the rest cryptocurrencies is highlighted the homogeneity among conditional distributions for ES forecasting. As it mentioned above, p-values are the minimum for belonging to the SSM. Except for Bitcoin, in which the SNP alternatives are in the top three rankings, for the rest of the cryptocurrencies the order of the alternatives in the SSM is somehow mixed, the Gaussian model appearing to be in the first position.

3.4. Median Shortfall backtesting results

Previous results point to the fact that 97.5 %-ES might be a more adequate risk measure for cryptocurrencies than 99 %-VaR, but still is very sensitive to extreme events. For this reason, we conjecture that MS might solve this shortcoming, although it is not a priori evident which measure, either 97.5 %-MS or 99 %-MS, would be the best alternative to replace 99 %-VaR or 97.5 %-ES. The analyses below and the discussion in the next section shed some light on this issue.

Table 11 displays the expected and observed exceptions and the CC and DQ tests for 99 %-MS (Panel A) and 97.5 %-MS (Panel B),

Table 10
Comparison for the SSM for 97.5 %-ES.

	97.5 %-ES			
	Rank _M	t _{ij}	P-value _M	Loss
<i>Bitcoin: 0 eliminations</i>				
Student's t	6	1.616	(0.161)	0.322
Normal	4	0.277	(0.949)	0.309
Skewed t	5	1.523	(0.195)	0.317
SNP c ₃ , c ₄	3	−0.500	(1.000)	0.305
SNP c ₃ , c ₄ , c ₆ , c ₈	2	−1.644	(1.000)	0.298
PGC c ₂ , c ₄ , c ₆	1	−2.515	(1.000)	0.297
<i>Litecoin: 0 eliminations</i>				
Student's t	4	−0.090	(1.000)	0.390
Normal	1	−4.742	(1.000)	0.379
Skewed t	5	0.754	(0.657)	0.397
SNP c ₃ , c ₄	6	1.570	(0.188)	0.401
SNP c ₃ , c ₄ , c ₆ , c ₈	3	−0.432	(1.000)	0.389
PGC c ₂ , c ₄ , c ₆	2	−1.019	(1.000)	0.387
<i>Ripple: 4 eliminations</i>				
Student's t	1	−0.646	(1.000)	0.391
Skewed t	2	0.646	(0.525)	0.393
<i>Monero: 0 eliminations</i>				
Student's t	2	−0.505	(1.000)	0.425
Normal	1	−2.402	(1.000)	0.423
Skewed t	4	−0.076	(1.000)	0.427
SNP c ₃ , c ₄	6	1.380	(0.394)	0.432
SNP c ₃ , c ₄ , c ₆ , c ₈	5	1.224	(0.495)	0.431
PGC c ₂ , c ₄ , c ₆	3	−0.415	(1.000)	0.426
<i>Stellar: 0 eliminations</i>				
Student's t	3	−0.393	(1.000)	0.509
Normal	1	−0.693	(1.000)	0.510
Skewed t	4	0.310	(0.950)	0.517
SNP c ₃ , c ₄	6	0.785	(0.642)	0.518
SNP c ₃ , c ₄ , c ₆ , c ₈	5	0.472	(0.865)	0.516
PGC c ₂ , c ₄ , c ₆	2	−0.608	(1.000)	0.510
<i>Ethereum: 0 eliminations</i>				
Student's t	2	−0.419	(1.000)	0.422
Normal	1	−1.878	(1.000)	0.419
Skewed t	4	0.032	(1.000)	0.424
SNP c ₃ , c ₄	6	1.442	(0.281)	0.430
SNP c ₃ , c ₄ , c ₆ , c ₈	5	0.431	(0.912)	0.425
PGC c ₂ , c ₄ , c ₆	3	−0.386	(1.000)	0.423

Rank_M refers to the ranking of the models, t_{ij} stands for the Diebold-Mariano test statistic for two models and p-value are in parenthesis. P-value_M is the minimum p-value for belonging to the Superior Set Model (SSM). The number of eliminations corresponds to the models that do not belong to the SSM. Loss is the average loss for each model.

which, according to Eq. (31), is equivalent to calculate 99.5 %-VaR and 98.75 %-VaR, respectively. Results for 99 %-MS are consistent with those of 99 %-VaR. The Gaussian hypothesis is rejected in all cases. On the contrary, the more flexible specifications result in accurate risk measures. The three SNP conditional distributions have good results for all cryptos except for Ethereum, although for this latter series the expected number of exceptions is too low for 99 %-MS and results might be sensitive to this issue.

On the other hand, the results for 97.5 %-MS (Panel B) shows that Gaussian is only rejected model for Bitcoin and Stellar. The Students't and the Skewed-t are rejected in some cases. Nonetheless, SNP distributions are found to be accepted with good results for all cryptos, supporting the flexibility in those SNP alternatives.

Details about the relative performance test for 97.5 %-MS can be presented in Table 12 where, in general, the results are quite similar to 99 %-VaR. This is due to the fact, that 97.5-MS is equal to 98.75-VaR a similar confidence level to 99 %. PGC have good results for Bitcoin and Litecoin in comparison to the parametric distributions whereas for most altcoins differences are only significant compared to the Gaussian. For the case of Ethereum, Student's t and SNP c₃, c₄ seem to provide the best forecasting performance.

Relative performance (DM test) for 99 %-MS has a similar performance than 97.5 %-MS and the results are available upon request.

For 99 %-MS and 97.5 %-MS the comparison for the SSM is displayed in Table 13. The only cryptocurrency with eliminated models

Table 11
Backtesting 99 %-MS and 97.5 %-MS.

		Bitcoin	Litecoin	Ripple	Monero	Stellar	Ethereum
99 %-MS Panel A							
Gaussian	E.E.	9	9	8	7	6	4
	Exc.	25	20	20	14	16	14
		21.680	11.671	13.080	6.598	11.153	13.835
	CC Test	(0.000)	(0.003)	(0.001)	(0.037)	(0.004)	(0.001)
	DQ Test	47.881	25.680	30.935	26.188	45.103	42.435
Student's t		(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
	Exc.	14	14	8	6	5	8
		3.169	3.169	0.080	0.111	0.300	2.552
	CC Test	(0.205)	(0.205)	(0.961)	(0.946)	(0.861)	(0.279)
	DQ Test	25.637	9.409	0.800	0.859	1.296	35.376
Skewed t		(0.001)	(0.225)	(0.997)	(0.997)	(0.989)	(0.000)
	Exc.	14	14	12	9	9	9
		3.169	3.169	1.867	0.911	1.223	3.924
	CC Test	(0.205)	(0.205)	(0.393)	(0.634)	(0.543)	(0.141)
	DQ Test	26.105	9.938	17.540	24.871	3.415	34.324
SNP c_3, c_4		(0.000)	(0.192)	(0.014)	(0.001)	(0.844)	(0.000)
	Exc.	12	9	10	10	7	10
		1.420	0.120	0.563	1.672	0.172	5.525
	CC Test	(0.492)	(0.942)	(0.755)	(0.433)	(0.917)	(0.063)
	DQ Test	6.114	1.545	21.225	23.702	59.293	33.552
SNP c_3, c_4, c_6, c_8		(0.527)	(0.981)	(0.003)	(0.001)	(0.000)	(0.000)
	Exc.	8	5	6	5	7	10
		0.110	1.764	0.623	0.464	0.172	5.525
	CC Test	(0.946)	(0.414)	(0.732)	(0.793)	(0.917)	(0.063)
	DQ Test	4.041	1.613	1.143	0.841	58.799	34.496
PGC c_2, c_4, c_6		(0.775)	(0.978)	(0.992)	(0.997)	(0.000)	(0.000)
	Exc.	10	7	6	8	9	10
		0.356	0.355	0.623	0.377	1.223	5.525
	CC Test	(0.837)	(0.837)	(0.733)	(0.828)	(0.543)	(0.063)
	DQ Test	5.580	0.767	1.123	26.648	63.769	34.212
97.5 %-MS Panel B		(0.590)	(0.998)	(0.993)	(0.000)	(0.000)	(0.000)
	E.E.	21	21	20	16	16	11
	Exc.	36	28	28	23	26	18
		9.997	2.407	3.205	3.127	7.012	4.590
	CC Test	(0.007)	(0.300)	(0.201)	(0.209)	(0.030)	(0.101)
Gaussian		17.630	13.742	9.400	8.575	19.907	15.073
	DQ Test	(0.014)	(0.056)	(0.225)	(0.285)	(0.006)	(0.035)
	Exc.	31	29	26	17	21	15
		5.024	5.250	2.185	0.459	24.559	1.874
	CC Test	(0.081)	(0.072)	(0.335)	(0.795)	(0.293)	(0.392)
Student's t		14.977	17.906	7.640	6.560	12.374	10.064
	DQ Test	(0.036)	(0.012)	(0.365)	(0.476)	(0.089)	(0.185)
	Exc.	30	38	28	21	27	16
		4.227	21.117	3.205	1.824	7.238	2.649
	CC Test	(0.121)	(0.000)	(0.201)	(0.402)	(0.027)	(0.266)
Skewed t		14.195	69.930	10.548	8.619	22.029	12.033
	DQ Test	(0.048)	(0.000)	(0.160)	(0.281)	(0.003)	(0.099)
	Exc.	20	16	15	14	15	13
		0.563	1.789	1.738	0.705	0.387	0.754
	CC Test	(0.755)	(0.409)	(0.419)	(0.703)	(0.824)	(0.686)
SNP c_3, c_4		6.055	6.390	7.048	7.488	12.808	9.510
	DQ Test	(0.533)	(0.495)	(0.424)	(0.380)	(0.077)	(0.218)
	Exc.	22	16	16	19	20	13
		0.594	1.789	1.247	0.921	1.829	0.754
	CC Test	(0.743)	(0.409)	(0.536)	(0.631)	(0.401)	(0.686)
SNP c_3, c_4, c_6, c_8		6.508	7.723	6.489	5.869	20.712	9.773
	DQ Test	(0.482)	(0.358)	(0.484)	(0.555)	(0.004)	(0.202)
	Exc.	25	22	16	16	18	14
		1.342	1.156	1.247	0.408	0.897	1.239
	CC Test	(0.511)	(0.561)	(0.536)	(0.815)	(0.637)	(0.538)
PGC c_2, c_4, c_6		6.256	6.678	6.373	6.550	10.056	9.556
	DQ Test	(0.510)	(0.463)	(0.497)	(0.477)	(0.185)	(0.215)

E.E. stands for the expected exceptions for each crypto and sample size. The number of observed exceptions (Exc.) per each crypto and model is reported in the table. Conditional Coverage (CC) test and Dynamic Quantile (DQ) (with 4 lags) test are reported in the table with p-value for both tests in parentheses (significance level is 10 %). The table includes the results for two different semi-nonparametric (SNP) expansions and a Positive Gram-Charlier version (PGC).

Table 12
Diebold-Mariano (DM) test for 97.5 %-MS.

Model A → Model B ↓	Student's t		Skewed t		SNP c ₃ , c ₄		SNP c ₃ , c ₄ , c ₆ , c ₈		PGC c ₂ , c ₄ , c ₆	
Panel A. Bitcoin										
Gaussian	1.596	(0.945)	1.158	(0.876)	−2.401	(0.008)	−2.573	(0.005)	−2.460	(0.007)
Student's t	−	−	−2.780	(0.003)	−2.212	(0.014)	−2.297	(0.011)	−2.159	(0.015)
Skewed t	−	−	−	−	−2.150	(0.016)	−2.234	(0.013)	−2.064	(0.020)
SNP c ₃ , c ₄	−	−	−	−	−	−	1.575	(0.942)	2.298	(0.989)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	2.754	(0.997)
Panel B. Litecoin										
Gaussian	−0.947	(0.172)	0.070	(0.528)	−2.412	(0.008)	−2.132	(0.017)	−2.521	(0.006)
Student's t	−	−	2.545	(0.994)	−2.408	(0.008)	−1.932	(0.027)	−1.299	(0.097)
Skewed t	−	−	−	−	−2.837	(0.002)	−2.323	(0.010)	−2.832	(0.002)
SNP c ₃ , c ₄	−	−	−	−	−	−	−0.734	(0.232)	2.158	(0.984)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	1.613	(0.947)
Panel C. Ripple										
Gaussian	−1.844	(0.033)	−1.327	(0.092)	−2.510	(0.006)	−2.556	(0.005)	−2.444	(0.007)
Student's t	−	−	2.589	(0.995)	−0.427	(0.335)	0.121	(0.548)	0.314	(0.623)
Skewed t	−	−	−	−	−1.703	(0.044)	−1.305	(0.096)	−1.086	(0.139)
SNP c ₃ , c ₄	−	−	−	−	−	−	2.127	(0.983)	2.269	(0.988)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	1.017	(0.845)
Panel D. Monero										
Gaussian	−1.927	(0.027)	−1.198	(0.116)	−2.119	(0.017)	−1.870	(0.031)	−2.395	(0.008)
Student's t	−	−	2.459	(0.993)	0.384	(0.649)	1.667	(0.952)	1.084	(0.861)
Skewed t	−	−	−	−	−2.306	(0.011)	−0.675	(0.250)	−0.811	(0.209)
SNP c ₃ , c ₄	−	−	−	−	−	−	2.448	(0.993)	1.264	(0.897)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	−0.766	(0.222)
Panel E. Stellar										
Gaussian	−2.292	(0.011)	−1.100	(0.136)	−2.751	(0.003)	−2.661	(0.004)	−3.000	(0.001)
Student's t	−	−	2.756	(0.997)	0.428	(0.666)	1.367	(0.914)	1.484	(0.931)
Skewed t	−	−	−	−	−1.821	(0.034)	−0.991	(0.161)	−0.769	(0.221)
SNP c ₃ , c ₄	−	−	−	−	−	−	2.809	(0.997)	2.235	(0.987)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	0.764	(0.777)
Panel F. Ethereum										
Gaussian	−1.492	(0.068)	−1.258	(0.104)	−1.744	(0.041)	−1.665	(0.048)	−1.829	(0.034)
Student's t	−	−	2.408	(0.992)	−1.937	(0.027)	0.595	(0.724)	0.674	(0.750)
Skewed t	−	−	−	−	−2.355	(0.009)	−0.422	(0.337)	−0.002	(0.499)
SNP c ₃ , c ₄	−	−	−	−	−	−	2.106	(0.982)	1.539	(0.938)
SNP c ₃ , c ₄ , c ₆ , c ₈	−	−	−	−	−	−	−	−	0.736	(0.769)

The table shows the Diebold-Mariano statistics and p-values (in parentheses). Bold figures indicate that Model A is preferred than Model B, underlined figures indicate the opposite. Otherwise, both models present the same performance. The null hypothesis is rejected when $|DM| > 1.96$ at 5% significance level.

is Ripple where the shorter version of SNP dropped from the SSM. As well as in other risk measures, this altcoin shows evidences certain heterogeneity for 97.5 %-MS forecasting. However, there exists a great homogeneity for all the conditional distributions and for all the cryptocurrencies. Regarding to model ranking, the first positions correspond to the SNP approach.

4. Discussion

This paper compares the performance of different financial risk measures for capturing the risk of extremely volatile assets. For this purpose, we analyze different cryptocurrencies with an abundant presence of extreme values. In this context, we propose the use of MS, a risk measure robust to outliers, as a good alternative for risk management and decision-making under high degrees of uncertainty. The results on the best model, however, seem to depend on the cryptocurrency. On one hand, the use of flexible specifications should outperform the more rigid and non-Gaussian distributions. This result has been found for most financial assets, including Bitcoin, but, on the other hand, when market instability is too extreme the more complex structures do not necessarily provide the best forecasting performance (e.g., Litecoin), while for 'purely parametric' distribution (e.g., Student's t type) data fits tend to overestimate the exceptions. This might be a consequence of the smooth decreasing at the far end of the distribution tails (recall the low values of the df parameter), which is not exhibited by the SNP approach that allows wavy tails. Therefore, the SNP produces more conservative risk

Table 13

Comparison for the SSM for 99 %-MS and 97.5 %-MS.

	99 %-MS					97.5 %-MS			
	Rank _M	t _{ij}	P-value _M	Loss		Rank _M	t _{ij}	P-value _M	Loss
<i>Bitcoin: 0 eliminations</i>					<i>Bitcoin: 0 eliminations</i>				
Student's t	5	2.167	(0.056)	0.109	5	1.583	(0.213)		0.193
Normal	4	1.501	(0.225)	0.110	4	−0.007	(1.000)		0.188
Skewed t	6	2.622	(0.018)	0.108	6	1.731	(0.161)		0.193
SNP c ₃ , c ₄	1	−3.775	(1.000)	0.099	3	−0.839	(1.000)		0.185
SNP c ₃ , c ₄ , c ₆ , c ₈	3	−1.263	(1.000)	0.101	2	−1.019	(1.000)		0.186
PGC c ₂ , c ₄ , c ₆	2	−2.507	(1.000)	0.099	1	−2.855	(1.000)		0.184
<i>Litecoin: 0 eliminations</i>					<i>Litecoin: 0 eliminations</i>				
Student's t	5	0.643	(0.829)	0.130	3	−0.414	(1.000)		0.241
Normal	4	0.566	(0.875)	0.132	2	−1.383	(1.000)		0.237
Skewed t	6	1.286	(0.397)	0.132	5	0.447	(0.917)		0.246
SNP c ₃ , c ₄	1	−2.773	(1.000)	0.122	4	0.260	(0.985)		0.244
SNP c ₃ , c ₄ , c ₆ , c ₈	3	−0.112	(1.000)	0.128	6	1.511	(0.259)		0.252
PGC c ₂ , c ₄ , c ₆	2	−1.065	(1.000)	0.125	1	−2.599	(1.000)		0.238
<i>Ripple: 0 eliminations</i>					<i>Ripple: 3 eliminations</i>				
Student's t	2	−2.207	(1.000)	0.124	1	−2.724	(1.000)		0.245
Normal	4	0.897	(0.669)	0.141	3	0.315	(0.962)		0.254
Skewed t	1	−2.903	(1.000)	0.126	2	−1.363	(1.000)		0.248
SNP c ₃ , c ₄	3	0.056	(1.000)	0.134	–	–	–		–
SNP c ₃ , c ₄ , c ₆ , c ₈	6	2.120	(0.068)	0.140	5	2.417	(0.032)		0.260
PGC c ₂ , c ₄ , c ₆	5	1.745	(0.169)	0.138	4	1.261	(0.376)		0.256
<i>Monero: 0 eliminations</i>					<i>Monero: 0 eliminations</i>				
Student's t	4	−0.300	(1.000)	0.142	4	0.293	(0.991)		0.268
Normal	6	0.888	(0.682)	0.151	3	−0.374	(1.000)		0.265
Skewed t	3	−0.573	(1.000)	0.142	6	0.627	(0.867)		0.268
SNP c ₃ , c ₄	2	−0.880	(1.000)	0.142	5	0.475	(0.941)		0.268
SNP c ₃ , c ₄ , c ₆ , c ₈	5	−0.150	(1.000)	0.143	2	−0.685	(1.000)		0.265
PGC c ₂ , c ₄ , c ₆	1	−1.085	(1.000)	0.141	1	−0.855	(1.000)		0.265
<i>Stellar: 0 eliminations</i>					<i>Stellar: 0 eliminations</i>				
Student's t	3	−0.698	(1.000)	0.159	3	−0.260	(1.000)		0.312
Normal	6	1.050	(0.550)	0.175	5	0.428	(0.939)		0.317
Skewed t	2	−0.796	(1.000)	0.159	6	0.446	(0.931)		0.318
SNP c ₃ , c ₄	4	0.109	(0.999)	0.164	4	0.210	(0.995)		0.315
SNP c ₃ , c ₄ , c ₆ , c ₈	5	0.922	(0.639)	0.167	2	−0.455	(1.000)		0.312
PGC c ₂ , c ₄ , c ₆	1	−2.025	(1.000)	0.159	1	−1.377	(1.000)		0.309
<i>Ethereum: 0 eliminations</i>					<i>Ethereum: 0 eliminations</i>				
Student's t	4	−1.052	(1.000)	0.153	3	−0.415	(1.000)		0.278
Normal	6	1.597	(0.183)	0.176	5	0.413	(0.955)		0.281
Skewed t	5	−0.643	(1.000)	0.155	6	1.089	(0.556)		0.282
SNP c ₃ , c ₄	3	−1.215	(1.000)	0.156	2	−0.434	(1.000)		0.278
SNP c ₃ , c ₄ , c ₆ , c ₈	2	−1.703	(1.000)	0.154	4	0.222	(0.994)		0.279
PGC c ₂ , c ₄ , c ₆	1	−2.142	(1.000)	0.156	1	−2.094	(1.000)		0.276

Rank_M refers to the ranking of the models, t_{ij} stands for the Diebold-Mariano test statistic for two models and p-value are in parenthesis. P-value_M is the minimum p-value for belonging to the Superior Set Model (SSM). The number of eliminations corresponds to the models that do not belong to the SSM. Loss is the average loss for each model.

measures for the relevant high quantiles but with a faster convergence to the zero x-axis asymptote.

Furthermore, this issue is very sensitive to the risk measure employed and the probability level considered. These results should be taken into account when making portfolio decisions and managing risk, since replacing 99 %-VaR is linked to the methodology that we employ to compute it. In fact, the decision of the BCBS to shift 99 %-VaR by 97.5 %-ES is due to the fact that these measures are approximately the same under Gaussianity (i.e., the ratio 97.5 %-ES/99 %-VaR is approximately one). Nonetheless, this distribution does not seem to be appropriate for cryptocurrencies, neither for financial returns that have found to be closer to a Student's t distribution with low df (3 and 5 df seemed appropriate according to [Kumari and Tan \(2013\)](#)). Therefore, such a decision is not innocuous and implies increasing the capital buffer to mitigate potential market risks.

Along the same line with the BCBS, we propose the use of MS as a risk measure more robust to outliers, but replacing 99 %-VaR or

Table 14

Ratios of risk measures for alternative distributions.

	Gaussian	Student's-t (3)	Student's-t (5)	SNP c_3, c_4
97.5%ES/99%VaR	1.005	1.110	1.047	0.993
98.31%-MS/99%VaR	1.027	1.064	1.046	1.028
98.513%-MS/99%VaR	1.047	1.115	1.082	1.048
97.5%ES/98.31%-MS	0.979	1.043	1.000	0.966
97.5%ES/98.513%-MS	0.960	0.995	0.967	0.948

The table presents different ratios among ES (Expected shortfall), VaR (Value at Risk) and MS (Median Shortfall) under alternative distributions: Student's t distribution with 3 and 5 degrees of freedom and SNP with parameters $c_3 = 0.029$ and $c_4 = 0.099$.

Table 15

Average Ratios of risk measures for Bitcoin along the backtesting.

	Gaussian	Student's-t	SNP c_3, c_4
97.5%ES/99%VaR	1.005 (0.000)	1.099 (0.016)	0.993 (0.001)
98.31%-MS/99%VaR	1.027 (0.000)	1.062 (0.004)	1.028 (0.001)
98.513%-MS/99%VaR	1.048 (0.001)	1.111 (0.007)	1.049 (0.001)
97.5%ES/98.31%-MS	0.978 (0.000)	1.035 (0.011)	0.966 (0.002)
97.5%ES/98.513%-MS	0.959 (0.001)	0.989 (0.008)	0.947 (0.002)

Standard deviation in parenthesis.

97.5 %-ES by an 'equivalent' MS measure. This change requires to select the appropriate confidence level in order to have a similar ratio to the abovementioned comparison. To shed some light on this, [Table 14](#) exhibits several ratios among alternative risk measures (99 %-VaR, 97.5 %-ES and their MS counterpart measures), illustrating the sensitivity of the ratios to the assumed distribution. In particular, we compare the Gaussian, the Student's t with 3 and 5 df and the SNP distribution with parameters $c_3 = 0.029$ and $c_4 = 0.099$ (the average value of the fitted parameters along the backtesting). This table shows that the ratio between 98.31 %-MS/97.5 %-ES (1.000) and 98.513-MS/97.5 %-ES (0.995) are similar with a Student's-t distribution with 5 and 3 df, respectively, whereupon MS seems to be a good alternative for these new confidence levels.

Of course, [Table 14](#) provides just an example with fixed (although realistic) values of the parameters. In order to show the ratio performance along the backtesting, we replicate the same procedure implemented in Section 3 with Bitcoin series, although, in order to have exact results in the pairwise comparison, ES has been computed as Eq. (29), avoiding approximations. [Table 15](#) shows the average ratios between different risk measures under Gaussian, Student's t and SNP distributions (standard deviations of the ratios in parentheses). The results are very similar to those in [Table 14](#), thus emphasizing that regulatory changes between risk measures should be highly dependent on the assumed distribution for risk management, otherwise capital buffer might be significantly affected.

5. Conclusions

Due to the significant growth of Cryptocurrency industry and the salient differences between crypto assets, which exhibit extreme volatility, a specific analysis on risk assessment must be undertaken. This paper provides a comprehensive analysis on this issue for a wide variety of cryptocurrencies (including Bitcoin, Litecoin, Ripple, Monero, Stellar and Ethereum) focused on two main features: (i) Proposing the use of a SNP approach as a flexible technique to accurately approximating the distribution tails; (ii) Discussing on the different risk measures, particularly those recently accorded by the BCBS (99 %-VaR and 97.5 %-ES) and proposing the 99 %-MS and 97.5 %-MS as a more appropriate alternative to model cryptocurrency risk due to the abundant presence of extreme values.

The main conclusions of our empirical study are the following:

First, the SNP approach provides very accurate risk measures for cryptocurrencies, particularly when conservative risk measures are required. This is a consequence of the wavy tails of the distribution, which cannot be performed with traditional parametric alternatives, and the flexibility to improve the data fits with a variable number of parameters. The positive transformations preserve these properties, although generally speaking, their use is not necessary as long as a controlled optimization is implemented.

Second, the performance of the SNP approach in comparison to other leptokurtic and skewed alternatives might be sensitive to the cryptocurrency and confidence level. In particular, for high confidence levels and Bitcoin series it works remarkably well, but for lower confidence levels or certain volatile altcoins finding superior forecasting models seems to be a very difficult task. However, in the cases where discriminating between models is not plausible, the model ranking analyses show that the SNP models are mostly in the first positions in terms of risk forecasting ability.

Third, for modeling cryptocurrency risk, the MS results in more robust-to-outliers risk measures, but the appropriate probability levels must be chosen according to the assumed distribution, since an arbitrary choice might induce important implications for

regulatory capital requirements on investment decisions. In particular, the 98.31 %-MS and 98.51 %-MS seem to be similar to 99 %-VaR and 97.5 %-ES for the range of flexible distributions studied in this paper.

In a nutshell, the use of the MS with a proper probability level choice and computed under a flexible SNP approach seems to provide conservative but accurate, risk measures for cryptocurrencies. However, given the high heterogeneity among cryptocurrencies risk managers should choose the most accurate model for each when deciding capital buffer of hedging policies to deal with potential losses.

Authors statement

Inés Jiménez: Investigation, Methodology, Formal Analysis, Software, Data curation, Validation, and Writing-Original draft preparation.

Andrés Mora-Valencia: Conceptualization, Methodology, Formal Analysis, Software and Supervision Writing-Reviewing and Editing.

Javier Perote: Conceptualization, Methodology, Formal Analysis, Supervision Writing-Reviewing and Editing and Project Administration.

Acknowledgment

This work was supported by the Consejería de Educación, Junta de Castilla y León (Grant No. SA049G19), FAPA- Uniandes (Grant No. PR.3.2016.2807), Fundación Banco Santander (Grant No. Doctoral research scholarship). We also thank to participants to the Conference Research Conference 2020.

References

- Acerbi, C., Szekely, B., 2014. Backtesting expected shortfall. *Risk Mag.* 27, 76–81.
- Acereda, B., Leon, A., Mora, J., 2019. Estimating the expected shortfall of cryptocurrencies: an evaluation based on backtesting. *Finance Res. Lett.* (January), 1–6. <https://doi.org/10.1016/j.frl.2019.04.037>.
- Ardia, D., Bluteau, K., Rüede, M., 2019. Regime changes in Bitcoin GARCH volatility dynamics. *Finance Res. Lett.* 29, 266–271. <https://doi.org/10.1016/j.frl.2018.08.009>.
- Baek, C., Elbeck, M., 2015. Bitcoins as an investment or speculative vehicle? A first look. *Appl. Econ. Lett.* 22 (1), 30–34. <https://doi.org/10.1080/13504851.2014.916379>.
- Balcilar, M., Bouri, E., Gupta, R., Roubaud, D., 2017. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Econ. Model.* 64 (March), 74–81. <https://doi.org/10.1016/j.econmod.2017.03.019>.
- Barnard, R.W., Pearce, K., Trindade, A.A., 2018. When is tail mean estimation more efficient than tail median? Answers and implications for quantitative risk management. *Ann. Oper. Res.* 262 (1), 47–65. <https://doi.org/10.1007/s10479-017-2547-7>.
- Bayer, S., Dimitriadis, T., 2020. Regression-based expected shortfall backtesting. *J. Financ. Econom.* <https://doi.org/10.1093/jfinec/nbaa013>.
- Bernardi, R., Catania, L., 2016. Comparison of Value-at-Risk models using the MCS approach. *Comput. Stat.* 31, 579–608. <https://doi.org/10.1007/s00180-016-0646-6>.
- Blau, B.M., 2018. Price dynamics and speculative trading in Bitcoin. *Res. Int. Bus. Financ.* 43, 15–21. <https://doi.org/10.1016/j.ribaf.2017.07.183>.
- Borri, N., 2019. Conditional tail-risk in cryptocurrency markets. *J. Empir. Finance* 50, 1–19. <https://doi.org/10.1016/j.jempfin.2018.11.002>.
- Bouoiyour, J., Selmi, R., 2016. Bitcoin: a beginning of a new phase? Volume 36, Issue 3 Bitcoin: a beginning of a new phase? *Econ. Bull.* 36, 1430–1440.
- Bouri, E., Azzi, G., Dyhrberg, A.H., 2017. On the return-volatility relationship in the bitcoin market around the price crash of 2013. *Economics* 11 (Vix), 1–20. <https://doi.org/10.5018/economics-ejournal.ja.2017-2>.
- Cagli, E.C., 2019. Explosive behavior in the prices of Bitcoin and altcoins. *Finance Res. Lett.* 29 (August 2018), 398–403. <https://doi.org/10.1016/j.frl.2018.09.007>.
- Campbell, S., 2007. A review of backtesting and backtesting procedures. *J. Risk* 9 (2), 1–17. <https://doi.org/10.21314/jor.2007.146>.
- Cerqueti, R., Giacalone, M., Mattered, R., 2020. Skewed non-Gaussian GARCH models for cryptocurrencies volatility modelling. *Inf. Sci.* 527, 1–26. <https://doi.org/10.1016/j.ins.2020.03.075>.
- Chebyshev, P.L., 1855. Sur les Fractions Continues [in Russian]. *J. de Mathematiques et Appliquees* 3 (1958), 298–323 [in French].
- Chu, J., Chan, S., Nadarajah, S., Osterrieder, J., 2017. GARCH modelling of cryptocurrencies. *J. Risk Financ. Manag.* 10 (4), 17. <https://doi.org/10.3390/jrfm10040017>.
- Corbet, S., Meegan, A., Larkin, C., Lucey, B., Yarovaya, L., 2018. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Econ. Lett.* 165, 28–34. <https://doi.org/10.1016/j.econlet.2018.01.004>.
- Corrado, C.J., Su, T., 1996. Skewness and kurtosis in S&P500 index returns implied by option prices. *J. Financ. Res.* 19 (2), 175–192. <https://doi.org/10.1111/j.1475-6803.1996.tb00592.x>.
- Danielsson, J., 2011. *Financial Risk Forecasting*. Wiley Finances Series, New Jersey.
- De la Horra, L.P., de la Fuente, G., Perote, J., 2019. The drivers of Bitcoin demand: a short and long-run analysis. *Int. Rev. Financ. Anal.* 62, 21–34. <https://doi.org/10.1016/j.irfa.2019.01.006>.
- Del Brio, E.B., Níguez, T.M., Perote, J., 2011. Multivariate semi-nonparametric distributions with dynamic conditional correlations. *Int. J. Forecast.* 27 (2), 347–364. <https://doi.org/10.1016/j.ijforecast.2010.02.005>.
- Del Brio, E.B., Mora-Valencia, A., Perote, J., 2020. Risk quantification for commodity ETFs: backtesting value-at-risk and expected shortfall. *Int. Rev. Financ. Anal.* 70 (101163) <https://doi.org/10.1016/j.irfa.2017.11.007>.
- Díaz, A., García-Donato, G., Mora-Valencia, A., 2019. Quantifying risk in traditional energy and sustainable investments. *Sustainability* (Switzerland) 11 (3), 1–22. <https://doi.org/10.3390/su11030720>.
- Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. *J. Bus. Econ. Stat.* 13 (3), 253. <https://doi.org/10.2307/1392185>.
- Dwyer, G.P., 2015. The economics of Bitcoin and similar private digital currencies. *J. Financ. Stab.* 17, 81–91. <https://doi.org/10.1016/j.jfs.2014.11.006>.
- Dyhrberg, A.H., 2016a. Bitcoin, gold and the dollar – a GARCH volatility analysis. *Finance Res. Lett.* 16, 85–92. <https://doi.org/10.1016/j.frl.2015.10.008>.
- Dyhrberg, A.H., 2016b. Hedging capabilities of bitcoin. Is it the virtual gold? *Finance Res. Lett.* 16, 139–144. <https://doi.org/10.1016/j.frl.2015.10.025>.
- Emmer, S., Kratz, M., Tasche, D., 2015. What is the best risk measure in practice? A comparison of standard measures. *J. Risk* 18 (2), 31–60. <https://doi.org/10.21314/JOR.2015.318>.
- Engle, R.F., González-Rivera, G., 1991. Semi-parametric ARCH models. *J. Bus. Econ. Stat.* 9, 345–359.
- Engle, R.F., Manganelli, S., 2004. CAViaR: conditional autoregressive value at risk by regression quantiles. *J. Bus. Econ. Stat.* 22 (4), 367–381. <https://doi.org/10.1198/073500104000000370>.

- Fernández, C., Steel, M.F.J., 1998. On bayesian modeling of fat tails and skewness. *J. Am. Stat. Assoc.* 93 (441), 359–371. <https://doi.org/10.1080/01621459.1998.10474117>.
- Fischer, M., Moser, T., Pfeuffer, M., 2018. A discussion on recent risk measures with application to credit risk: calculating risk contributions and identifying risk concentrations. *Risks* 6 (4). <https://doi.org/10.3390/risks6040142>.
- Fissler, T., Ziegel, J.F., Gneiting, T., 2015. Expected Shortfall Is Jointly Elicitable With Value at Risk – Implications for Backtesting. Retrieved from: <http://Arxiv.Org/Abs/1507.00244>, pp. 1–7. <http://arxiv.org/abs/1507.00244>.
- Fissler, T., Ziegel, J.F., 2016. Higher order elicibility and osband's principle. *Ann. Stat.* 44 (4), 1680–1707. <https://doi.org/10.1214/16-AOS1439>.
- Gallant, A.R., Nychka, D.W., 1987. Semi-nonparametric maximum likelihood estimation. *Econometrica* 55 (2), 363. <https://doi.org/10.2307/1913241>.
- Gkillas, K., Katsiampa, P., 2018. An application of extreme value theory to cryptocurrencies. *Econ. Lett.* 164, 109–111. <https://doi.org/10.1016/j.econlet.2018.01.020>.
- Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M.C., Siering, M., 2014. Bitcoin – asset or currency? Revealing users' hidden intentions. *ECIS 2014 Proceedings – 22nd European Conference on Information Systems* 1–14.
- Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Finance* 48 (5), 1779–1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>.
- Gneiting, T., 2011. Making and evaluating point forecasts. *J. Am. Stat. Assoc.* 106 (494), 746–762. <https://doi.org/10.1198/jasa.2011.r10138>.
- Hald, A., 2000. The early history of the cumulants and the Gram-Charlier series. *Int. Stat. Rev.* 68 (2), 137–153. <https://doi.org/10.1111/j.1751-5823.2000.tb00318.x>.
- Hansen, P.R., Lunde, A., Nason, James M., 2011. The model confidence set. *Econometrica* 79 (2), 453–497. <https://doi.org/10.3982/ecta5771>.
- Hermite, C., 1859. Sur l'interpolation. In: *Comptes Rendus Des Séances de L'Académie Des Sciences Paris*, 48, pp. 62–67.
- Jiménez, I., Mora-Valencia, A., Perote, J., 2020. Risk quantification and validation for Bitcoin. *Oper. Res. Lett.* 48 (4), 534–541. <https://doi.org/10.1016/j.orl.2020.06.004>.
- Jondeau, E., Rockinger, M., 2001. Gram-Charlier densities. *J. Econ. Dyn. Control* 25 (10), 1457–1483. [https://doi.org/10.1016/S0165-1889\(99\)00082-2](https://doi.org/10.1016/S0165-1889(99)00082-2).
- Katsiampa, P., 2017. Volatility estimation for Bitcoin: a comparison of GARCH models. *Econ. Lett.* 158, 3–6. <https://doi.org/10.1016/j.econlet.2017.06.023>.
- Kou, S., Peng, X., 2014. Expected shortfall or median shortfall. *J. Financ. Eng.* 1 (1), 1450007. <https://doi.org/10.1142/S234576861450007x>.
- Kou, S., Peng, X., 2016. On the measurement of economic tail risk. *Oper. Res.* 64 (5), 1056–1072. <https://doi.org/10.1287/opre.2016.1539>.
- Kratz, M., Lok, Y.H., McNeil, A.J., 2018. Multinomial VaR backtests: a simple implicit approach to backtesting expected shortfall. *J. Bank. Finance* 88, 393–407. <https://doi.org/10.1016/j.jbankfin.2018.01.002>.
- Kumari, S.N., Tan, A., 2013. Characterization of student's T- distribution with some application to finance. *Math. Theory Model.* 3 (10), 1–10. <https://doi.org/10.7176/MTM>.
- Kupiec, P.H., 1995. Techniques for verifying the accuracy of risk measurement models. *J. Deriv.* 3 (2), 73–84. <https://doi.org/10.3905/jod.1995.407942>.
- Lahmiri, S., Bekiros, S., Salvi, A., 2018. Long-range memory, distributional variation and randomness of bitcoin volatility. *Chaos Solitons Fractals* 107, 43–48. <https://doi.org/10.1016/j.chaos.2017.12.018>.
- Le, T.H., 2020. Forecasting value at risk and expected shortfall with mixed data sampling. *Int. J. Forecast.* 36 (4), 1362–1379. <https://doi.org/10.1016/j.ijforecast.2020.01.008>.
- León, A., Mencía, J., Sentana, E., 2009. Parametric properties of semi-nonparametric distributions, with applications to option valuation. *J. Bus. Econ. Stat.* 27 (2), 176–192. <https://doi.org/10.1198/jbes.2009.0013>.
- Liu, W., Semeyutin, A., Lau, C.K.M., Gozgor, G., 2020. Forecasting value-at-Risk of cryptocurrencies with riskmetrics type models. *Res. Int. Bus. Finance* 54, 101259. <https://doi.org/10.1016/j.ribaf.2020.101259>.
- Mauleón, I., Perote, J., 2000. Testing densities with financial data: an empirical comparison of the Edgeworth-Sargan density to the student's t. *Eur. J. Finance* 6, 225–239.
- McNeil, A.J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *J. Empir. Finance* 7, 271–300.
- Nakamoto, S., 2008. Bitcoin: un sistema de dinero en efectivo electrónico peer-to-peer. https://bitcoin.org/files/bitcoin-paper/bitcoin_es.pdf%0Awww.bitcoin.org.
- Nass, C.A.G., 1959. The χ^2 test for small expectations in contingency tables, with special reference to accidents and absenteeism. *Biometrika*. <https://doi.org/10.2307/2333534>.
- Nešlehová, J., Embrechts, P., Chavez-Demoulin, V., 2006. Infinite-mean models and the LDA for operational risk. *J. Oper. Risk* 1 (1), 3–25. <https://doi.org/10.21314/jop.2006.001>.
- Nguyen, T.V.H., Nguyen, B.T., Nguyen, T.C., Nguyen, Q.Q., 2019. Bitcoin return: impacts from the introduction of new altcoins. *Res. Int. Bus. Finance* 48, 420–425. <https://doi.org/10.1016/j.ribaf.2019.02.001>.
- Ñíguez, T.M., Perote, J., 2012. Forecasting heavy-tailed densities with Positive Edgeworth and Gram-Charlier expansions. *Oxf. Bull. Econ. Stat.* 74 (4), 600–627. <https://doi.org/10.1111/j.1468-0084.2011.00663.x>.
- Ñíguez, T.M., Paya, I., Peel, D., Perote, J., 2012. On the stability of the constant relative risk aversion (CRRA) utility under high degrees of uncertainty. *Econ. Lett.* 115 (2), 244–248. <https://doi.org/10.1016/j.econlet.2011.12.049>.
- Nolde, N., Ziegel, J.F., 2017. Elicitability and backtesting: perspectives for banking regulation. *Ann. Appl. Stat.* 11 (4), 1901–1911. <https://doi.org/10.1214/17-AOAS1041F>.
- Pearson, K., 1900. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Lond. Edinburgh Dublin Philos. Mag. J. Sci.* <https://doi.org/10.1080/14786440009463897>.
- Rauchs, M., Blandin, A., Klein, K., Pieters, G.C., Recanatini, M., Zhang, B.Z., 2019. 2nd global cryptoasset benchmarking study. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.3306125>.
- Sargan, J.D., 1976. Econometric estimators and the Edgeworth approximation. *Econometrica* 44, 421–448. <https://doi.org/10.1016/j.jaci.2012.05.050>.
- So, M.K.P., Wong, C.M., 2012. Estimation of multiple period expected shortfall and median shortfall for risk management. *Quant. Finance* 12 (5), 739–754. <https://doi.org/10.1080/14697681003785967>.
- Stavroyiannis, S., 2018. Volatility modeling and risk assessment of the major digital currencies. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.3092173>.
- Trespalcacios, A., Cortés, L.M., Perote, J., 2020. Uncertainty in electricity markets from a semi-nonparametric approach. *Energy Policy* 137. <https://doi.org/10.1016/j.enpol.2019.111091>.
- Trucíos, C., 2019. Forecasting Bitcoin risk measures: a robust approach. *Int. J. Forecast.* 35 (3), 836–847. <https://doi.org/10.1016/j.ijforecast.2019.01.003>.
- Yermack, D., 2015. Is bitcoin a real currency? An economic appraisal. *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data*. Elsevier Inc, pp. 31–43. <https://doi.org/10.1016/B978-0-12-802117-0.00002-3>.