SCHOOL OF
ECONOMICS AND
MANAGEMENT

Master's Program in Finance

RISK MEASUREMENT OF CRYPTOCURRENCIES USING VALUE AT RISK AND EXPECTED SHORTFALL

Author: Van Cao Thi Hong

Supervisor: Andreas Johansson

May 2022

ABSTRACT

Cryptocurrencies are highly volatile and risky assets, therefore, it is of vital importance to find an appropriate model for risk measurement. This thesis compares three parametric and three non-parametric estimation methods to estimate the value at risk and the expected shortfall of five cryptocurrencies, namely Bitcoin (BTC), Ethereum (ETH), Binance coin (BNB), Ripple coin (XRP), and Cardano (ADA). We estimate the value at risk and expected shortfall using these methods at the confidence level of 95% and 99%. We then perform five backtesting procedures and use these test results to compare the performance of these estimation methods. Consequently, we can conclude that the volatility-weighted historical simulation (VWHS) method using the exponential weighted moving average (EWMA) model and GARCH-type models to rescale cryptocurrency loss for VaR and ES estimation perform the best in most cases. The basic historical simulation (BHS) method and the peak over threshold (POT) method also show positive performance in several cases. Meanwhile, the age-weighted historical simulation (AWHS) has a poor performance in almost all cases.

ACKNOWLEDGEMENT

I would like to express my sincere thanks to my supervisor, Andreas Johansson, for all of his valuable comments and advice, which enormously help me define my thesis purpose and approach. His clear and concise guidance helps me to build a logical thesis structure and improve the coherence of my writing. I also want to send a special thanks to the Swedish Institute and the Swedish Ministry of Foreign Affairs for sponsoring my study at Lund University.

TABLE OF CONTENT

1. Introduction

The explosive growth of cryptocurrencies has become a major matter of concern for the financial market. The market capitalization of cryptocurrencies reached its all-time-high peak of 3 trillion dollars in 2021. As of 30 April 2022, there are more than 19 thousand types of cryptocurrencies in the market (coinmarketcap.com). Since its whitepaper was first published in 2008 (Nakamoto, 2008), Bitcoin has remained the largest cryptocurrency based on the market capitalization. It is followed by Ethereum, the cryptocurrency of a platform built to write smart contracts and decentralized applications (Buterin, 2014).

Being driven by not only their intrinsic value but also market news and speculations, cryptocurrencies are highly volatile and risky assets. Risk management of cryptocurrencies is a primary concern of most investors, policymakers, governments, institutions, researchers, and economists. Finding a practical and prominent method to measure the risk of cryptocurrencies is therefore absolutely imperative, laying a concrete foundation for the development of a cryptocurrency policy framework, the enforcement of necessary laws and regulations, or the implementation of a more effective investment strategy.

In this thesis, we propose the application of value at risk and expected shortfall to quantify the risk of cryptocurrencies. As there are several methods to estimate these risk measures, our research purpose is then to find out "which method performs the best in measuring the value at risk and expected shortfall of cryptocurrencies?". Our general approach is to estimate the value at risk and expected shortfall of cryptocurrencies using three parametric and three non-parametric estimation methods, then perform several backtesting procedures to compare the performance of these methods in measuring the risk of cryptocurrencies. The conclusion about the best risk estimation method is beneficial to the cryptocurrency risk control process.

1.1. Overview of cryptocurrencies

Cryptocurrencies are decentralized digital assets that apply blockchain technology. Such technology allows the online transfer of value in a decentralized network without the participation of traditional financial intermediaries like banks. The cryptocurrency blockchain is similar to a general ledger that records every transaction in this network. However, unlike the traditional ledger, it is enforced by a large number of computers by the consensus algorithm known as “proof-of-work”, in which the computers have to mine or solve the increasingly hard mathematic problems to verify a new transaction in the blockchain. Accordingly, all transactions in this network cannot be manipulated or reversed, except for the situation when more than 50% of computing power is controlled. It can also be enforced by an alternative mechanism named “proof-of-stake” that does not heavily rely on computer resources.

1.2. Overview of literature on cryptocurrencies

Due to the fast-growing market of cryptocurrencies, many research papers have been carried out. Literature on cryptocurrencies can be divided into three main research areas. The first strand of literature focuses on cryptocurrency characteristics, its main drivers, and its classification as a real currency or a speculative investment (Ariefianto, 2020; Baldan & Zen, 2020; Baur, Hong & Lee, 2018; Chen Y. Wu & Pandey, 2014; Janson & Karoubi, 2021; Ratajczak-Mrozek & Marszałek, 2022; Yermack, 2015). The conclusions from these research works remain ambiguous, but most empirical results indicate that cryptocurrency is more similar to a speculative investment than a medium of exchange.

The second strand of research investigates the use of cryptocurrencies as a portfolio hedge or a safe haven, and its relationship with other financial assets in a portfolio such as commodities, equities, and bonds (Bouri, Hussain Shahzad & Roubaud, 2020; Choi & Shin, 2022; Corbet et al., 2018; Dyhrberg, 2016a; Guesmi et al., 2019; Hasan et al., 2022; Mroua, Bahloul & Naifar, 2022; Platanakis & Urquhart, 2020; Ustaoglu, 2022; Wang, Ma & Wu, 2020; Wang et al., 2019). These literature works produce different results related to the use of cryptocurrencies in investment strategies in turbulent periods.

The third and most practical research area examines the risky aspects of cryptocurrencies and the application of different risk measures as a tool to improve the risk forecasting process and investment decisions. This research area on cryptocurrency risk measurement is also the focal point of this thesis. Overall, many research papers focus on analyzing the volatility dynamics as a risk measure of cryptocurrencies and comparing the performance of different GARCH-type models in estimating volatility (for details of the model specification, see Section 3.2.1.3). In contrast, the number of research papers on measuring value at risk and expected shortfall of cryptocurrencies is relatively limited. These papers will be discussed in more details in Section 2.

Upon reviewing the existing research papers, we note that there are several limitations in the sample selection process. First, due to the constraint of the research period, most research papers (all papers in the literature review section with date before 2020) do not capture the recent volatility behavior of cryptocurrencies in their risk estimates. The price of cryptocurrencies fluctuates significantly from 2020 onwards due to the covid-19 pandemic and the explosive technological development. Therefore, their conclusions are based on the study of limited historical data, which may not be applicable to the current period. Second, many papers (Liu et al., 2020; Platanakis & Urquhart, 2020; Van der Auwera et al., 2020) use return or loss data from a specific cryptocurrency spot exchange (for example, Bitstamp, Kraken, Bitfinex…), which may not be representative of the whole market because cryptocurrencies may have different listing dates and trading volumes on different spot exchanges. Third, several research papers only focus on Bitcoin rather than other major types of cryptocurrencies, while some other papers still include cryptocurrencies that do not have any intrinsic value in the research. In contrast, we will improve the sample selection process by choosing the more representative, informative, and updated set of data, as discussed in more details in Section 4.

1.3. Value at risk and expected shortfall as risk measures of cryptocurrencies

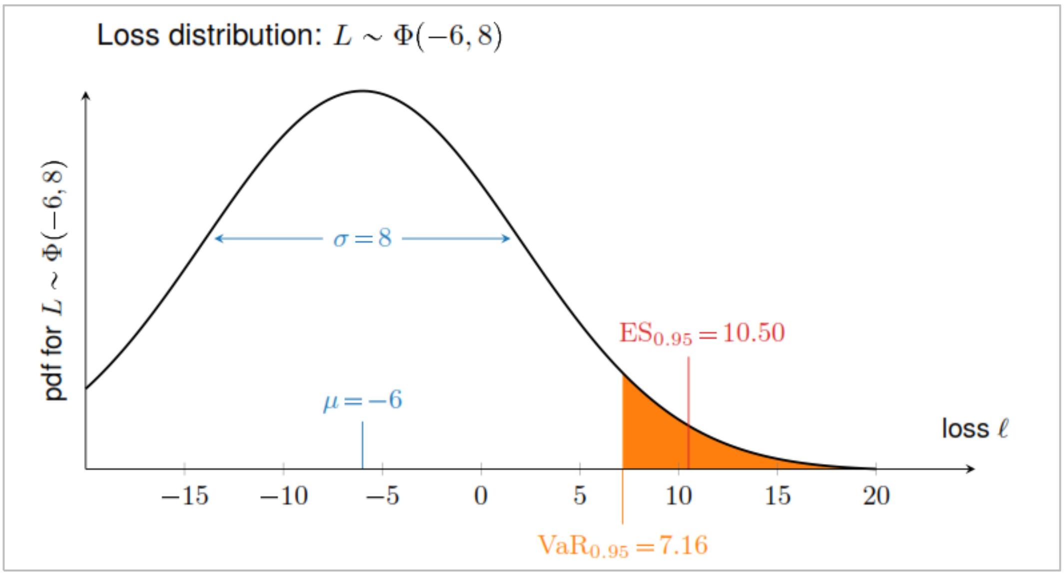
Value at risk and expected shortfall are two commonly used risk measures for traditional financial assets, but they have not been widely applied to cryptocurrencies.

Value at risk is a useful tool in both risk reporting and other stages of the risk management process, including setting the regulatory capital requirements (Basel III, Solvency II), defining the acceptable risk levels, establishing the risk budgets, or performing sensitivity analysis and stress testing.

Value at risk measures the maximum loss at the given confidence level, calculated by a quantile of the loss distribution $\Pr(L > \mathrm{VaR}_\alpha) = 1 - \alpha$ (1). In other words, we can interpret that the possibility of losing more than VaR_α is equal to $1 - \alpha$.

Despite its simplicity in interpretation, value at risk is not a coherent risk measure as it does not always satisfy the subadditivity requirement. Therefore, expected shortfall is also used in support of value at risk. Expected shortfall is a coherent risk measure that quantifies the average magnitude of the loss in case the tail events happen. However, backtesting procedures for expected shortfall is more difficult to implement than value at risk, thus these two risk measures are used simultaneously for risk management.

Graph 1.1. *Example of* $VaR_{0.95}$ *and* $ES_{0.95}$



Graph 1.1 illustrates an example of value at risk and expected shortfall at the 95% confidence level. As $\Pr(L > 7.16) = 1 - 0.95 = 0.5$, which satisfies the definition of VaR so that $\mathrm{VaR}_\alpha = 7.16$. ES_α is then equal to the average of all losses larger than VaR.

An advantage of value at risk and expected shortfall is that they take all underlying risk factors into consideration by directly studying the loss distribution. Given that value at risk and expected shortfall can be applied to all asset classes, it is beneficial to carry out more research on the application of these risk measures to cryptocurrencies. Some recent papers started to introduce specific methods to estimate value at risk and expected shortfall of cryptocurrencies, which will be discussed in Section 2.2.

2. Literature review

2.1. Volatility analysis of cryptocurrencies

As cryptocurrencies are consistently more volatile than other traditional securities (Van der Auwera et al., 2020), risk management based on volatility modelling is of crucial importance. Many research works have carried out volatility analysis of cryptocurrencies, notably Bitcoin.

The most frequently used model in literature for modeling volatility and estimating value at risk and expected shortfall is the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. Some studies analyzed the volatility dynamics of Bitcoin with different GARCH-type models, including autoregressive jump-intensity GARCH (Gronwald, 2014), threshold GARCH (TGARCH) (Bouoiyour & Selmi, 2015; Dyhrberg, 2016a), exponential GARCH (EGARCH) (Dyhrberg, 2016b), component GARCH (CGARCH) (Katsiampa, 2017), fractionally integrated GARCH (FIGARCH) (Lahmiri, Bekiros & Salvi, 2018; Ulmer & Chen, 2021), mixed data sampling GARCH (GARCH-MIDAS) (Fang et al., 2019), Markov-switching GARCH (MSGARCH) (Ardia, Bluteau & Rüede, 2019; Caporale & Zekokh, 2019; Maciel, 2021).

Several authors have measured and compared relative performance of different GARCH-type models. Bouoiyour & Selmi (2016) carried out many extensions of GARCH models to estimate Bitcoin price volatility and concluded that Bitcoin remains more reactive to negative news than positive news, consequently proposing the use of asymmetry TGARCH model. Chu et al. (2017) fitted twelve GARCH-type models to seven most popular cryptocurrencies, using their log return data from 2014 to 2017, and chose the best fitting models based on five model selection criteria. The results indicate that IGARCH (1,1) may be a good fit for five among seven selected cryptocurrencies, while GARCH (1,1) may fit Ripple coin better and GJR-GARCH (1,1) is the best fitting model for Dogecoin. Katsiampa (2017) compared six GARCH-type models in explaining Bitcoin volatility and concluded that asymmetric component GARCH (AR-CGARCH) is the optimal model to fit the data. This result implies the importance of including both the long-term and short-term components of the conditional variance. Gyamerah (2019) evaluated Bitcoin volatility returns using three GARCH-type models and suggested the use of TGARCH models. Obeng (2021) compared SGARCH and

EGARCH model and concluded that EGARCH model perform better in measuring the volatility of 30 cryptocurrencies during the period from 2017 to 2020, due to the larger impact of bad news on volatility than positive news.

There is no consensus about which is the best GARCH model to use. According to Jiménez, Mora-Valencia & Perote (2022), the inconsistent conclusions result from the fact that analyzing the high volatility of cryptocurrencies requires not only the conditional heteroskedastic processes, but also the full shape modeling of the density of the underlying stochastic process. Due to the existence of outliers, such density should reflect skewness, kurtosis, and extreme values at the tail ends. Following the idea of Cerqueti, Giacalone & Mattera (2020) to employ skewed non-Gaussian GARCH models for cryptocurrency volatility analysis, Jiménez, Mora-Valencia & Perote (2022) then propose the use of semi-nonparametric approach (SNP) based on Gram-Charlier series to approximate the distribution tails.

Several research papers point out that the volatility of cryptocurrency returns has some characteristics similar to traditional financial assets, including heteroscedasticity (Gkillas & Katsiampa, 2018; Van der Auwera et al., 2020), leptokurtosis (Chan et al., 2017), and long memory (Assaf, Alberiko Gil-Alana & Mokni, 2021; Lahmiri, Bekiros & Salvi, 2018; Phillip, Chan & Peiris, 2019). For example, Lahmiri, Bekiros & Salvi (2018) identify a fractional long-range memory and inherent stochasticity in Bitcoin volatility under normal, student-t, t-skewed, and generalized error distribution.

According to Van der Auwera et al. (2020), the log returns of cryptocurrencies observe heavier tails than normal distribution. Therefore, to find the most appropriate distribution for the log returns of cryptocurrencies, these authors choose to fit different types of distributions that are more suitable for fat tails, namely lognormal, student t, Pareto, Cauchy, Burr, Weibull, and Frechet distribution. Using Kolmogorov-Smirnov test (KS test), Akaike's Information Criteria (AIC), and Bayesian Information Criteria (BIC) to select the most fitted distribution with respect to maximum likelihood, the results show that student t-distribution may be a better fit than others for most selected cryptocurrencies. However, this conclusion is drawn from the limited observation of historical returns of only four cryptocurrencies between Jan 2016 and Feb 2019. The volatility behavior of cryptocurrencies has changed significantly from that time onwards, especially in 2021.

More recently, by investigating the return behavior of 254 cryptocurrencies from March 2019 to March 2021, Fung, Jeong & Pereira (2021) confirm the conclusion of previous studies that common characteristics of cryptocurrencies' return behavior include heavy tails, long memory, volatility persistence, and negative leverage effects. The research suggests that it is of vital importance to incorporate a heavy-tailed distribution into the GARCH-type models to capture the volatility clusters and large kurtosis of cryptocurrencies' returns. In addition, this study also indicates that student t-distribution is a good fit for most cryptocurrencies.

In our thesis, we will use the popular GARCH-type models, including GARCH, EGARCH and GJR-GARCH models, to incorporate the volatility clustering in our VaR and ES estimates. More details about the model specification and the motivation for model selection are stated in Section 3.2.1.3. Using a more updated and representative sample, we can investigate whether there is any significant change in the volatility behavior compared to the previous research. We also fit normal distribution and student t-distribution to our loss sample (following the idea of Van der Auwera et al. (2020) and Fung, Jeong & Pereira (2021) that student-t distribution may be a good fit for cryptocurrencies). Moreover, we will employ the generalized Pareto distribution in our VaR and ES estimates, which have not been examined by the existing research papers.

2.2. Risk measures of cryptocurrencies

Value at risk (VaR) and expected shortfall (ES) have been popularly employed for risk management of different asset classes, however, research papers on the application of these risk measures for cryptocurrencies are relatively limited. Recent literature started to investigate the use of value at risk and expected shortfall for risk measurement of cryptocurrencies.

Likitratcharoen et al. (2018) used historical simulation VaR and Gaussian parametric VaR to measure the risk of cryptocurrencies. A similar method is employed in a more recent paper to test the accuracy of Bitcoin's risk measures during Covid-19 pandemic (Likitratcharoen et al., 2021). Gkillas & Katsiampa (2018) applied the extreme value theory to study the tail behavior for VaR and ES estimates.

Many authors considered including the time-varying volatility of cryptocurrencies in their risk estimation. An early paper by Stavroyiannis (2017) showed some preliminary calculations of value at risk and expected shortfall for several major cryptocurrencies at that time, using GARCH modeling. His later paper then focused on the VaR and ES estimation of Bitcoin, implementing a GJR-GARCH model with residuals following the standardized Pearson type-IV distribution (Stavroyiannis, 2018). The results show that Bitcoin violates VaR measures more than other assets, consequently being subject to higher capital requirements.

Troster et al. (2019) concluded that heavy-tailed generalized autoregressive score (GAS) models yielded a better result than GARCH-type models when estimating VaR of Bitcoin. Meanwhile, Trucíos (2019) proved that the robust-based residual bootstrap method outperformed GARCH-type and GAS models for Bitcoin's VaR forecast.

In a later research paper, Trucíos, Tiwari & Alqahtani (2020) proposed a vine copula-based approach to estimate VaR and ES of seven cryptocurrencies with daily return series from 1 January 2015 to 14 June 2019. Such approach based on robust volatility models to some extent solved the existence of extreme values and the correlation between cryptocurrencies.

Using high-frequency data of Bitcoin, Pele & Mazurencu-Marinescu-Pele (2019) suggested a new VaR estimation procedure based on the entropy of the intraday log-return distribution. The VaR backtesting results show that this method produces better VaR forecasts than traditional GARCH models.

Liu et al. (2020) estimated VaR of three cryptocurrencies (Bitcoin, Ethereum, Litecoin) with exponentially weighted moving average (EWMA) model. Using different VaR backtesting methods, they concluded that EWMA model can be used to forecast VaR of cryptocurrencies. Similar to the conclusion of Trucíos (2019), they also observed that GAS models showed a good performance at most levels.

Görgen, Meirer & Schienle (2022) proposed to use the quantile version of generalized random forests (GRF) for out-of-sample VaR prediction of four popular cryptocurrencies. The results indicate that GRF method outperformed other methods in measuring the tail risk of cryptocurrencies, especially in unstable periods. In addition, Li (2022) suggested using a forecast combination of value at risk and expected shortfall for the risk management of cryptocurrencies. This combination, under some certain conditions, is more useful compared to individual forecasts.

Most research papers above employed similar methods for VaR and ES backtesting, among which the most frequently used methods are Kupiec test (1995), Christoffersen test (1998), and a joint conditional coverage test named Christoffersen and Pelletier (2004). Some literature papers also used Engle & Manganelli test (2004), and model confidence set procedure (MCS) of Hansen et al. (2011).

To the best of our knowledge, none of the above papers conduct a full comparison between the performance of different non-parametric and parametric VaR and ES estimation methods which we intend to use in our thesis. More specifically, no existing research compare the performance between the basic historical simulation method, the age-weighted historical simulation method, the generalized Pareto distribution with other popular methods to estimate VaR and ES. Therefore, we aim to contribute to the current research papers by comparing the ability of six estimation methods to measure VaR and ES of cryptocurrencies (more details in Section 3) and finding the best model among them.

3. Methodology

3.1. General approach

This thesis evaluates the performance of several non-parametric and parametric methods in estimating VaR and ES. A general approach is to divide the total loss sample into a rolling estimation window and a testing window. The one-day ahead VaR and ES are estimated along with the testing window by rolling the fixed-size estimation window forward. For example, with a rolling estimation window of size 1000, the first 1000 loss observations ($\ell_1 - \ell_{1000}$) are used to estimate the first VaR and ES in the testing window at day 1001. These estimates are then compared with the corresponding actual data at day 1001. For the second VaR and ES estimates in the testing window, we remove the first loss observation ℓ_1 and add the latest one ℓ_{1001}, meaning that we keep the estimation window size unchanged, and use the next 1000 loss observations ($\ell_2 - \ell_{1001}$) to estimate VaR and ES at day 1002. Similar steps are repeated until the end of the testing window. Graph 3.1 below illustrates this approach.

Graph 3.1. *Rolling estimation window for VaR estimates and testing window*

	First estimation window				*Testing window*			
		Second estimation window						
Days	*1*	*2*	*(…)*	*1000*	*1001*	*1002*	*(…)*	*M*
VaR estimates					VaR_{1001}	VaR_{1002}	*(…)*	$\mathrm{VaR}_{\mathrm{M}}$
Actual losses	ℓ_1	ℓ_2	*(…)*	ℓ_{1000}	ℓ_{1001}	ℓ_{1002}	*(…)*	ℓ_{M}

We implement different parametric and non-parametric methods (more details in Section 3.2) to estimate 95% and 99% one-day ahead VaR and ES using observations in the rolling estimation window. A 95% one-day ahead value at risk of X million dollars indicates that the probability of losing more than X million dollars in the next day is 5%. The 95% expected shortfall of Y million dollars means that the average of VaR at all confidence levels larger than 95% in the next day is Y million dollars. We then use several backtesting procedures (more details in Section 3.3) to measure the accuracy of VaR and ES estimates in the testing window for the purpose of comparing these methods.

3.2. Estimating value at risk and expected shortfall

We choose several non-parametric and parametric methods to estimate VaR and ES. Non-parametric methods are based on empirical or historical loss distribution, meaning that historical data is used as a guide for the future (Hull, 2018). Meanwhile, parametric methods assume a parametric distribution for losses. Three popularly used non-parametric methods to estimate VaR and ES include basic historical simulation, age-weighted historical simulation, and volatility-weighted historical simulation. To estimate volatility of cryptocurrencies for the VaR and ES calculations using the volatility weighted historical simulation method, we use several conditional variance models (GARCH-type models). We also use three different parametric methods, assuming normal distribution, student t-distribution, and generalized Pareto distribution to estimate parameters of the distribution, which are then used to calculate VaR and ES.

3.2.1. Non-parametric methods

3.2.1.1. Basic historical simulation (BHS)

The basic historical simulation method estimates VaR and ES directly by sorting the sample of historical losses from the greatest to the smallest loss, then choosing the loss that satisfies VaR and ES definition. For example, given a sample of T historical loss observations, the one-day ahead VaR estimate at the α confidence level is the $(1-\alpha)T+1$ largest loss, and correspondingly, the ES estimate is the average of $(1-\alpha)T$ losses that are greater than VaR.

The intuition behind this method can be explained by considering the probability of losing more than the $(1-\alpha)T+1$ largest loss as follows.

$$\Pr\left(L > \ell_{(1-\alpha)T+1}^{sorted}\right) = \frac{(1-\alpha)T}{T} = 1-\alpha \Rightarrow VaR_\alpha = \ell_{(1-\alpha)T+1}^{sorted} \quad (2)$$

As proved by equation (2), the $(1-\alpha)T+1$ largest loss satisfies the α-quantile definition of VaR that $\Pr(L > VaR_\alpha) = 1-\alpha$, therefore it is an estimate of VaR at the α confidence level.

The basic historical simulation method is simple to implement and interpret, however, one problem of this method is that the uncertainty of VaR and ES estimates remains unknown. Therefore, we also use the bootstrapping method to deal with the estimation uncertainty. We simulate a large number of new bootstrapping samples from the original loss sample, then estimate VaR and ES for each new sample. By re-sampling, the uncertainty in VaR and ES estimates can be measured by the standard error of these new VaR and ES estimates.

Another major drawback of this method is that it does not take time variation into account. As such, all historical loss observations are equally weighted when estimating the one-day ahead VaR and ES. This problem can be solved by the next two non-parametric methods.

3.2.1.2. Age-weighted historical simulation (AWHS)

The underlying motivation for the age-weighted historical simulation method is that the more recent loss observation is more relevant for VaR and ES estimates than those observed in longer time ago, as it may contain more updated and similar information for the near future estimate. Therefore, higher weights should be allocated to the newer loss observations, and lower weights should be assigned to the more remote observations in the past. Table 3.2 illustrates the weights of each loss observation from the beginning point of estimation period t = 1 to the most recent time t = T. In our VaR and ES estimates, we use the decay factor $\lambda = 0.94$, which is a standard value proposed in the RiskMetrics™ technical document (J.P. Morgan & Reuters, 1996).

Table 3.2. *Allocated weights to each loss distribution (AWHS method)*

Loss	Weight (probability)	Note
ℓ_T	$w_T = \dfrac{1-\lambda}{1-\lambda^T}$	Highest weight/probability for the most recent loss observation
ℓ_{T-1}	$w_{T-1} = \lambda w_T$	
…	…	As $0 < \lambda < 1$, the weight is decreasing
ℓ_2	$w_2 = \lambda w_3 = \lambda^{T-2} w_T$	
ℓ_1	$w_1 = \lambda w_2 = \lambda^{T-1} w_T$	Lowest weight for the furthest loss observation

It is noted that the allocated weights can be interpreted as probabilities, because all the weights sum up to one.

$$\sum_{t=1}^{t=T} w_t = \sum_{t=1}^{t=T} \lambda^{T-t} w_T = w_T(1 + \lambda + \cdots + \lambda^{T-1}) = w_T \left(\frac{1-\lambda^T}{1-\lambda} \right) = 1 \tag{3}$$

Due to this interpretation, we can estimate VaR by sorting losses in the descending order (the greatest loss is the first), then accumulating the corresponding probabilities until $\Pr\left(L > \ell_{k+1}^{sorted}\right) > 1 - \alpha$. In this case, VaR estimate is equal to ℓ_k^{sorted} because it is the smallest loss that meet the VaR definition that $\Pr\left(L > \ell_k^{sorted}\right) \leq 1 - \alpha$. ES is the average of losses that larger than ℓ_k^{sorted}.

3.2.1.3. Volatility-weighted historical simulation (VWHS)

Volatility-weighted historical simulation method aims to include volatility clustering effect and time-varying volatility into VaR and ES estimates. The motivation for this method can be explained by a common empirical finding that the volatility of financial assets' returns or losses is not always constant over time. Normally unstable days happen in clusters, leading to the higher market risk during the turbulent period, which should be reflected in the estimates of VaR and ES.

The general approach is that the original losses ℓ_t (t ranges from time 1 to time T) should be scaled up or down according to the forecast volatility σ_{T+1} on the day after the estimation period. The volatility scaling formula is as follows, where ℓ_t^* is the rescaled loss at time t (with t = 1, 2, …, T).

$$\ell_t^* = \frac{\sigma_{t+1}}{\sigma_t} \ell_t \tag{4}$$

After that, we can implement similar steps as the basic historical simulation method to the rescaled losses (not the original losses) to estimate VaR and ES. VaR is then equal to the smallest rescaled loss that satisfies $\Pr(L > \ell_k^*) \leq 1 - \alpha$, and ES is the average of all rescaled losses greater than VaR.

The volatility in the scaling formula can be estimated by several GARCH-type models. The GARCH-type models have the proven capability to capture clustering pattern of volatility. As discussed in the literature review section, there is no consistent conclusion on the best GARCH model to estimate volatility of cryptocurrencies. In this thesis, we use four conditional variance models, including (i) EWMA model as suggested in the RiskMetrics™ technical document by J.P. Morgan & Reuters (1996) to forecast variances, (ii) GARCH model which is a popular standard model, (iii) EGARCH model to capture the additional leverage effect, and (iv) GJR-GARCH model that is similar to TGARCH model, proposed as the most suitable model to model the volatility of cryptocurrencies by Bouoiyour & Selmi (2016) and Gyamerah (2019).

We consider the time series of losses $\ell_t = \mu + \varepsilon_t$, where μ is the expected loss and $\varepsilon_t = \sigma_t z_t$ is the stochastic error, which can be interpreted as the unexpected loss. The term z_t reflects the shock to the market, which can be negative or positive. The motivation behind the equation $\varepsilon_t = \sigma_t z_t$ is that the unexpected loss is influenced by the shock to the market, where the size of the influence is equal to the volatility. In other words, the volatility impacts the magnitude of unexpected loss, and consequently impacts the total loss.

The conditional variance is calculated differently by four GARCH-type models, but generally, the variance of the unexpected loss at time t+1 is conditional on the set of historical information H_t available at time t.

$$\sigma_{t+1}^2 = Var(\varepsilon_{t+1}|H_t) \tag{5}$$

The historical information includes past variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_t^2$ and past innovations $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_t$. With four GARCH-type models in this thesis, two common characteristics of loss series, namely volatility clustering and leverage effect, can be incorporated into the estimates. While all four models can address the volatility clustering, the EGARCH and GJR-GARCH models can reflect the leverage effect, which will be discussed in more details in the following subsections (i) to (iv).

i) EWMA model

The original form of the exponential weighted moving average (EWMA) model forecasts the conditional variance σ_{T+1}^2 by taking the weighted average of past loss innovations $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_t$. The interpretation of the decay factor λ and the weights are similar to the age-weighted historical simulation method.

$$\sigma_{T+1}^2 = \frac{1-\lambda}{1-\lambda^T} \sum_{t=1}^{T} \varepsilon_t^2 \tag{6}$$

When the number of loss observations is reasonably large, the EWMA model can also be interpreted as the simplified GARCH model without the constant term.

$$\sigma_{T+1}^2 = \lambda\sigma_T^2 + (1-\lambda)\varepsilon_T^2 \tag{7}$$

The first variance is set equal to the variance of 1000 first loss observations. We also use the standard decay parameter $\lambda = 0.94$, which is recommended for EWMA model when dealing with the daily data set (J.P. Morgan & Reuters, 1996). There is a trade-off between the convenience of using a fixed parameter and the possibility that this parameter may not always be optimal.

ii) GARCH (1,1) model

The generalized autoregressive conditional heteroscedastic (GARCH) model (Bollerslev, 1986) was developed from the ARCH model (Engle, 1982). Empirical studies find that GARCH (1,1) model with one lag of past variance and one lag of past innovation is normally sufficient to capture the dynamics of volatility (Asgharian, 2021). In GARCH (1,1) model, the conditional variance is calculated as:

$$\sigma_{T+1}^2 = \kappa + \gamma\sigma_T^2 + \alpha\varepsilon_T^2 \tag{8}$$

The parameters in this equation are estimated by maximum likelihood, in which we choose the values of parameters that maximize the probability that the fitted distribution generates the observed innovations. We fit two types of distribution to the innovation process, the Gaussian-normal distribution (denoted as GARCHn) and the student t-distribution (denoted as GARCHt), to estimate three parameters of the GARCH (1,1) model.

To ensure the stationarity and positivity, the following constraints on the parameters should be satisfied: $\kappa > 0, \gamma \geq 0, \alpha \geq 0$, $\gamma + \alpha < 1$. The first variance and innovation follow the default set-up in the Matlab programming language. Accordingly, the first variance σ_1^2 is equal to the long-term unconditional variance, calculated by taking the sample average of the squared disturbances of input data. The first innovation ε_1 is by default the squared root of the average squared value of the loss data set. These settings minimize the initial temporary effects.

iii) EGARCH (1,1) model

The motivation for the use of the exponential GARCH (EGARCH) is that it can model the leverage effect, which is not captured by EWMA or GARCH model. The leverage effect indicates that negative shocks to the market may have a more significant impact on volatility than positive news. The EGARCH model aims to reflect this asymmetric characteristic by calculating the logarithm of variance as follows.

$$\log \sigma_{T+1}^2 = \kappa + \gamma \log \sigma_T^2 + \alpha \left[\frac{|\varepsilon_T|}{\sigma_T} - E\left\{\frac{|\varepsilon_T|}{\sigma_T}\right\}\right] + \xi \left(\frac{\varepsilon_T}{\sigma_T}\right) \tag{9}$$

The four parameters $\kappa, \gamma, \alpha,$ and ξ in this model are also estimated by the maximum likelihood method. As this model measures the logarithm of variance, the constraint for positivity is eliminated. The constraint for stationarity still holds. The first variance is similar to GARCH (1,1) model, but the first innovation is equal to zero by Matlab default set-up.

We also estimate EGARCH (1,1) model in two cases of normal (denoted as EGARCHn) and student-t innovation distribution (denoted as EGARCHt). If the innovation follows Gaussian-normal distribution, then:

$$E\left\{\frac{|\varepsilon_T|}{\sigma_T}\right\} = E\{|z_T|\} = \sqrt{\frac{2}{\pi}} \tag{10}$$

If the innovation follows student's t distribution and v > 2, then:

$$E\left\{\frac{|\varepsilon_T|}{\sigma_T}\right\} = E\{|z_T|\} = \sqrt{\frac{v-2}{\pi}} \frac{\Gamma\left(\frac{v-1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \tag{11}$$

iv) GJR-GARCH (1,1) model

Similar to the EGARCH (1,1) model, the GJR-GARCH (1,1) model also incorporates the leverage effect into the variance estimation as follows.

$$\sigma_{T+1}^2 = \kappa + \gamma \sigma_T^2 + \alpha \varepsilon_T^2 + \xi \boldsymbol{I}[\varepsilon_T < 0] \varepsilon_T^2 \tag{12}$$

$$\text{with } \boldsymbol{I}[\varepsilon_T < 0] = \begin{cases} 1 \text{ if } \varepsilon_T < 0 \\ 0 \text{ if } \varepsilon_T \geq 0 \end{cases}$$

In this model, the indicator function $\boldsymbol{I}[\varepsilon_T < 0] = 1$ implies that additional weight of innovations (ξ) should be added to the variance estimate when there is a negative shock ($\varepsilon_T < 0$). Otherwise, in case of positive information ($\varepsilon_T \geq 0$), $\boldsymbol{I} = 0$ then the conditional variance is calculated in the same way as the standard GARCH model.

Similar to GARCH (1,1) and EGARCH (1,1) models, we also use the maximum likelihood estimation method to obtain the four parameters of GJR-GARCH (1,1) model. For stationarity and positivity, these constraints on parameters should be applied:

$\kappa > 0, \gamma \geq 0, \alpha \geq 0, \alpha + \xi \geq 0, \gamma + \alpha + 1/2\xi < 1$. The default set-up in Matlab for the first variance and the first innovation in the GJR-GARCH model is exactly the same as the GARCH model.

This model has a close similarity with the below threshold GARCH (TGARCH) model, with the same interpretation of the indicator function.

$$\sigma_{T+1} = \kappa + \gamma\sigma_T + \alpha|\varepsilon_T| + \xi|\varepsilon_T|\boldsymbol{I}[\varepsilon_T < 0] \tag{13}$$

The TGARCH model has been proven to perform properly in modeling volatility of some types of cryptocurrencies as discussed in the review of existing research papers. Given that the GJR-GARCH (1,1) and the TGARCH (1,1) have a similar interpretation, we choose the GJR-GARCH (1,1) model for more convenient implementation using Matlab programming language. We also consider two types of innovation distribution, including normal (denoted as GJRGARCHn) and student-t (denoted as GJRGARCHt), in our estimates.

3.2.2. Parametric methods

The parametric method fits a specific type of distribution to the losses. To obtain the parameters for the distribution, typically we apply the maximum likelihood, which is similar to the method to estimate parameters of GARCH-type models.

3.2.2.1. Normal distribution

The normal distribution is a symmetric distribution around the mean, which has the following probability distribution function (pdf).

$$f(\ell) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ell - \mu}{\sigma}\right)^2\right] \tag{14}$$

In which, the two parameters are the mean μ (or location) and the standard deviation σ (or scale). First, we fit the normal distribution to loss data, using maximum likelihood method to estimate these parameters. Then, we use these parameters to calculate the VaR and ES at the α confidence level as follows.

$$VaR_\alpha = \mu + \sigma z_\alpha \text{ where } z_\alpha = \frac{VaR_\alpha - \mu}{\sigma} \tag{15}$$

$$ES_\alpha = \mu + \sigma \frac{f_{std}(z_\alpha)}{1-\alpha} \text{ where } f_{std}(z_\alpha) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_\alpha^2\right) \tag{16}$$

3.2.2.2. Student t-distribution

The student t-distribution has heavier tails than the normal distribution, therefore it is suitable for modeling distribution that is more sensitive to outliers. Some research papers show that the student t-distribution is the most fitted distribution for several types of cryptocurrencies (Fung, Jeong & Pereira, 2021; Van der Auwera et al., 2020). The student t-distribution has three parameters, including the location parameter μ, the scale parameter σ^*, and the shape parameter v, which can be interpreted as the mean, the standard deviation, and the degree of freedom respectively.

The pdf of student t-distribution is as follows.

$$f^*(\ell) = \frac{\Gamma[(v+1)/2]}{\sigma^*\sqrt{v\pi}\,\Gamma(v/2)}\left[1+\frac{1}{v}\left(\frac{\ell-\mu}{\sigma^*}\right)^2\right]^{-(v+1)/2} \tag{17}$$

Given the relationship between σ and σ^* in case the degree of freedom is higher than 2, the pdf can be rewritten as below.

$$f(\ell) = \frac{\Gamma[(v+1)/2]}{\sigma\sqrt{(v-2)\pi}\,\Gamma(v/2)}\left[1+\frac{1}{v-2}\left(\frac{\ell-\mu}{\sigma}\right)^2\right]^{-(v+1)/2} \tag{18}$$

The steps to estimate VaR and ES are similar to those of normal distribution. We fit the student t-distribution to the losses to find the value of three parameters μ, σ^*, and v with the maximum likelihood method, then calculate the α confidence level VaR and ES using these parameters.

$$VaR_\alpha = \mu + \sqrt{\frac{v-2}{v}}\sigma t_{\alpha,v} \text{ where } t_{\alpha,v} = \frac{VaR_\alpha - \mu}{\sigma^*} \tag{19}$$

$$ES_\alpha = \mu + \sqrt{\frac{v-2}{v}}\sigma \frac{f^*_{std}(t_{\alpha,v})}{1-\alpha}\left(\frac{v+t^2_{\alpha,v}}{v-1}\right) \text{ where } f^*_{std}(t_{\alpha,v}) = \frac{\Gamma[v+1)/2]}{\sqrt{v\pi}\Gamma(v/2)}\left[1+\frac{1}{v}t^2_{\alpha,v}\right]^{-\frac{v+1}{2}} \tag{20}$$

3.2.2.3. Generalized Pareto distribution (GPD-POT)

The general approach is to fit the generalized Pareto distribution to a subsample of the largest losses. This employs the idea of the peaks over threshold (POT) method, which is a popular version of the extreme value theory (EVT).

The POT method provides a solution to estimate VaR and ES with an unspecified loss distribution f. VaR is by definition the α-quantile of f. Given the corresponding cumulative distribution F, we can estimate VaR by solving the equation $F(VaR) = \alpha$.

According to the extreme value theory (Balkema & de Haan, 1974; Gnedenko, 1943; Pickands, 1975), the unknown equation $F(VaR) = \alpha$ can be approximated by the generalized Pareto distribution (GDP) as below, if the threshold level u is large enough.

$$F(VaR) \approx GDP(\xi, \beta) = 1 - (1 + \xi \frac{\ell - u}{\beta})^{-\frac{1}{\xi}} \tag{21}$$

In which, the parameters ξ and β are related to the right tail and the volatility respectively in the unspecified loss distribution f.

Overall, the POT method is conducted in three steps. First, we draw a subsample of the largest losses from the original loss sample. Second, we estimate the parameters of the generalized Pareto distribution, using maximum likelihood method. Finally, these parameters are used to estimate VaR and ES.

In the first step, the largest losses are defined as all losses in the sample that are beyond the threshold level u. A practical choice is using roughly 4% - 6% largest losses in the sample (Nilsson, 2021), thus we choose to use 5% of the largest losses in this thesis. As our main focus is the largest losses, the POT method is mainly used to estimate VaR and ES at a high confidence level. In this thesis, we choose to use the POT method at the 99% confidence level.

For the estimates of parameters ξ and β by maximum likelihood in the second step, the generalized Pareto probability density function $g(\xi, \beta)$ is required. It can be calculated by taking the derivative of GDP with respect to loss ℓ as follows.

$$g(\xi, \beta) = \frac{dGDP(\xi, \beta)}{d\ell} = \frac{1}{\beta}(1 + \xi \frac{\ell - u}{\beta})^{-\frac{1}{\xi}-1} \tag{22}$$

Then the value of parameters ξ and β can be estimated by maximizing the following log-likelihood function.

$$logL(\xi, \beta) = \sum_{i=1}^{N_u} ln \left[\frac{1}{\beta}\left(1 + \xi \frac{\ell_i - u}{\beta}\right)^{-1/\xi - 1}\right] \tag{23}$$

As a final step, we can use these parameters to estimate VaR and ES by the following formulas, in which N is the size of the original sample, and N_u is the size of the subsample of largest losses.

$$VaR_\alpha = u + \frac{\beta}{\xi}\left[(\frac{N}{N_u}(1 - \alpha))^{-\xi} - 1\right] \tag{24}$$

$$ES_\alpha = \frac{VaR_\alpha + \beta - \xi u}{1 - \xi} \tag{25}$$

3.3. Backtesting procedures

The purpose of VaR and ES backtesting procedures is to assess the performance of VaR and ES estimation methods. Normally, VaR and ES are estimated daily, using information available at time T to forecast one-day ahead VaR and ES at time T+1. Then when actual loss is known at the end of day T+1, it can be compared with the estimate of VaR one day before. Similarly, we can compare ES estimated on the previous day with the actual number when the loss data is available at day T+1. Therefore, actual daily loss data can be employed to measure the accuracy of VaR and ES models.

A violation is constituted when the actual loss ℓ is larger than the estimated VaR for the same day. If $\ell_t > VaR_t$, then we record a violation as "one", otherwise we record a non-violation as "zero". Consequently, the stochastic variable that represents the number of violations is a Bernoulli variable, following a binomial distribution.

To measure the performance of VaR models, we use several popular backtesting procedures, including the unconditional coverage test of Kupiec (1995), the traffic light test proposed by the Basel Committee (1996), the conditional coverage test and the independence test of Christoffersen (1998). For the backtesting of ES, we employ the unconditional test of Acerbi & Szekely (2014). For all the tests, we choose the pragmatic test confidence level of 95%.

3.3.1. Backtesting value at risk

3.3.1.1. Kupiec test (POF)

The Kupiec test is also known as the proportion of failures (POF) test. The purpose of this test is to assess whether the proportion of actual VaR violations is in line with the probability of failures indicated by the VaR confidence level.

The POF test follows the idea of the binomial test that the actual number and the expected number of VaR violations should be consistent if the VaR estimation method is correct. The actual number of violations is counted directly from the testing sample of size N. According to VaR definition, the expected number of violations is then equal to N(1 - VaR level). Following this idea, Kupiec (1995) develops the test by using the likelihood ratio for the comparison of probabilities. The POF's likelihood ratio or test statistic is calculated as follows.

$$LR_{POF} = -2\log\left(\frac{(1-p)^{N-x}p^{x}}{\left(1-\frac{x}{N}\right)^{N-x}\left(\frac{x}{N}\right)^{x}}\right) \sim \chi^{2}(1) \quad (26)$$

In which, x represents the number of observed violations, N is the size of testing sample and p = 1 - VaR level is the probability of violations inferred from the VaR confidence level.

If the test statistic is lower than the critical value, which depends on the test confidence level, the test result is to accept. Otherwise, the VaR estimation model fails the POF test.

3.3.1.2. Christoffersen tests (CC and CCI)

Two VaR backtesting procedures proposed by Christoffersen (1998) include the conditional coverage independence (CCI) test and the conditional coverage (CC) test. While the former tests for the independence of VaR violations and non-violations on the consecutive days, the latter tests for both the independence and the frequency of the VaR violations.

(i) The CCI test

The CCI test examines if the probability of violation on the following day depends on the violation today or not. Christoffersen (1998) states that the series of violations and non-violations (the string of "ones" and "zeros") follows the Markov chain with two states (violation state s_1 and non-violation state s_0). We can interpret that π_{01} is the probability of transition to a violation day on the next day, given that there is no violation today. Similarly, the matrix of conditional transition probabilities between two states are as follows.

$$\begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix} = \begin{pmatrix} \Pr\,(s_0|s_0) & \Pr\,(s_0|s_1) \\ \Pr\,(s_1|s_0) & \Pr\,(s_1|s_1) \end{pmatrix} \quad (27)$$

Under the null hypothesis, violations and non-violations are independent over time ($\pi_{11} = \pi_{01}$). In other words, the next day is a violation day, regardless today is a violation day or a non-violation day.

The test statistic of the CCI test is calculated as:

$$LR_{CCI} = -2\log\left(\frac{{\pi_0}^{n_0}{\pi_1}^{n_1}}{{\pi_{00}}^{n_{00}}{\pi_{01}}^{n_{01}}{\pi_{10}}^{n_{10}}{\pi_{11}}^{n_{11}}}\right) \sim \chi^2(1) \tag{28}$$

In which, n_1 is the number of ones, n_0 is the number of zeros, n_{ij} is the number of transitions from state s_i to state s_j (i and j can be 0 or 1). If the test statistic is higher than the critical value at the given significant level, the null hypothesis of independence is rejected.

(ii) The CC test

The CC test is a combination of the POF test and the CCI test, thus its test statistic is measured as follows.

$$LR_{CC} = LR_{POF} + LR_{CCI} \sim \chi^2(2) \tag{29}$$

Similar to POF test and CCI test, we compare the CC test statistic with the critical value at a statistical significant level. If it is higher than the critical value, then the VaR estimation method is rejected.

3.3.1.3. Traffic light test (TF)

The traffic light test, also known as the three zones test, is developed by the Basel Committee (1996) as a proxy to set up the bank's capital requirement. This test focuses on the underestimation of VaR, as it is an incentive for a lower capital reserve. It is originally used for bank regulation, but we can employ this test for cryptocurrencies to see the probabilities of observing up to a specific number of VaR violations.

There are three zones in the traffic light tests, including red zone, yellow zone, and green zone. The red zone starts when the cumulative probability of VaR violations is equal to or higher than 99.99%. The yellow zone covers the number of VaR violations where the probability is lower than 99.99% and higher than 95%. Otherwise, it enters the green zone.

The red zone implies an issue with the estimation method. If the VaR model is correct, it is not likely that there are too many violations. It is noted that only a too high number of VaR violations result in the rejection, as this test only covers the risk of VaR underestimation.

3.3.2. Backtesting expected shortfall

We use the unconditional test proposed by Acerbi and Szekely (2014) for the backtesting of ES. While most ES backtesting procedures require information about the loss distribution or the tail distribution, this unconditional test allows us to produce approximate test results without specifying the distribution.

The test statistic of Acerbi and Szekely test is derived from the ES definition, in which the indicator function I = 1 if loss is larger than VaR and I = 0 otherwise.

$$ES_{\alpha,t} = \frac{E[L_t I_t]}{1-\alpha} \Leftrightarrow -\frac{E[L_t I_t]/1-\alpha}{ES_{\alpha,t}} + 1 = 0 \tag{30}$$

The unconditional test statistic is then defined as follows.

$$Z_{uncond} = -\frac{1}{N(1-\alpha)} \sum_{t=1}^{t=N} \frac{L_t I_t}{ES_{\alpha,t}} + 1 \tag{31}$$

Under the null hypothesis, we expect that the test statistic is equal to zero, as a correct estimate of ES indicates that $ES_{\alpha,t} = E[L_t I_t]/(1-\alpha)$ for all days t. A negative value of the test statistic implies that $ES_{\alpha,t} < E[L_t I_t]/(1-\alpha)$ for at least one day t. Therefore, the test result is rejection if the test statistic is largely negative. In this test, the critical values for this test are precomputed and stable. The unconditional test is conducted in two cases that the critical values follow the normal distribution or student t-distribution.

4. Data

4.1. Data selection

As discussed in the introduction section, we aim to obtain a more representative, informative, and updated set of data than the existing research papers.

First, we collect the daily closing cryptocurrency prices from coinmarketcap.com, which is more representative of the market than the price obtained from an individual spot exchange. As stated in the metric methodology paper developed by CoinMarketCap (2019), the prices of cryptocurrencies are calculated as a volume-weighted average of all pair prices available on the market. The higher volume has a greater impact on the average price. The underlying motivation for this calculation is that markets with larger transaction volumes have higher liquidity and are possibly less susceptible to price fluctuation than those with lower volumes.

Second, the data is obtained for the entire period that the cryptocurrencies are listed on the market, from the first available date to 30 April 2022. It is more informative, containing both information further in the past and more updated information.

Third, we select cryptocurrencies that satisfy the requirements of (i) being in the top 20 cryptocurrencies by market capitalization as of 30 April 2022, (ii) having more than 1500 observations to obtain more robust results, and (iii) excluding stable coins that attempt to maintain a relatively stable price, and meme coins that are inspired by humorous ideas and jokes. Consequently, our sample of cryptocurrencies includes Bitcoin (BTC), Ethereum (ETH), Binance coin (BNB), Ripple coin (XRP), and Cardano (ADA). The main characteristics of these cryptocurrencies are stated in Appendix 1. These five cryptocurrencies account for roughly 67% of total market cap as of 30 April 2022, and mostly maintain dominance in the market (See Appendix 2). Therefore, this sample of cryptocurrencies is fairly representative of the market.

4.2. Descriptive statistics

Based on historical price series of all selected cryptocurrencies, we calculate the losses ℓ_t at time t as follows, where P_t and P_{t-1} are the daily closing prices of cryptocurrencies at time t and the day before respectively.

$$\ell_t = -100[\ln (P_t) - \ln (P_{t-1})] \tag{32}$$

It is notable that the loss here is defined as a positive number, so that it is more convenient to perform the tests and examine the test results. Table 4.1 below describes the descriptive statistics of the cryptocurrency loss sample. We also test for the normality assumption of the loss distribution using the Jarque-Bera test, and the stationarity of the loss time series using the Augmented Dickey-Fuller (ADF) test. More details about the methodology, the test statistic, and the expected results of these tests are stated in Appendix 3 and Appendix 4.

Table 4.1. *Descriptive statistics of cryptocurrency losses*

	BTC	**ETH**	**BNB**	**XRP**	**ADA**
Mean	-0.171	-0.280	-0.470	-0.144	-0.204
Median	-0.185	-0.081	-0.120	0.182	0.000
Minimum	-35.745	-41.241	-67.506	-102.746	-86.123
Maximum	46.473	130.214	54.281	61.638	50.371
Variance	17.657	42.837	49.184	50.603	49.754
Standard deviation	4.202	6.545	7.013	7.114	7.054
Skewness	0.514	3.201	-0.938	-1.591	-1.891
Kurtosis	13.916	72.009	18.074	29.823	25.545
Jarque-Bera statistic	16,475	491,933	16,729	97,007	36,406
Jarque-Bera result	reject	reject	reject	reject	reject
Jarque-Bera p-value	0.000	0.000	0.000	0.000	0.000
ADF test statistic	-58.254	-52.806	-38.784	-54.676	-41.439
ADF test result	reject	reject	reject	reject	reject
ADF p-value	0.001	0.001	0.001	0.001	0.001
Starting date	29/4/2013	8/8/2015	26/7/2017	5/8/2013	2/10/2017
Ending date	30/4/2022	30/4/2022	30/4/2022	30/4/2022	30/4/2022
Total observations	3,289	2,458	1,740	3,191	1,672

As can be seen from Table 4.1 and Graph 4.3, the mean and median value of losses for each cryptocurrency are relatively similar (approximately zero), however, its minimum and maximum value of losses range broadly, indicating the presence of outliers. Among five cryptocurrencies, ETH observed the largest loss of 130.214 due to a substantial drop in its price at the beginning of the period (See Graph 4.3). After that, the greatest losses of all cryptocurrencies are around 50 to 60. On the contrary, XRP and ADA achieved the exceptionally high returns of 102.746 and 86.123 respectively in 2017, while the highest return of BTC is only 35.745.

Overall, BTC losses are less volatile than other cryptocurrencies with the lowest standard deviation of roughly 4.2. Meanwhile, XRP, ADA, and BNB experience the highest volatility in their loss series, with the standard deviation of around 7. Graph 4.3 also illustrates the same conclusion. It can be explained by the fact that BTC is more well-known with the larger trading volume and market capitalization, so it is less prone to price fluctuation than other cryptocurrencies.

The Jarque-Bera test statistics of all cryptocurrencies in Table 4.1 are significantly higher than zero, clearly showing the rejection of normal distribution. Graph 4.4 presents the shape of each cryptocurrency loss distribution, with a closer display to the right and the left of the distribution. BTC and ETH losses witness the positive skewness, meaning the loss data spreads out more to the left than to the right of the mean, while the other cryptocurrencies observe the negative skewness. The kurtosis values of all five cryptocurrencies are greater than the standard value of three (See Appendix 3), indicating that the distribution is peaked with heavier tails. The high kurtosis results from the existence of extreme values. This observation is similar to the existing research papers (Stephen Chan et al., 2017). Among all, ETH is the most skewed and leptokurtic cryptocurrency, with the highest skewness of 3.2 and an extremely large kurtosis of 72.

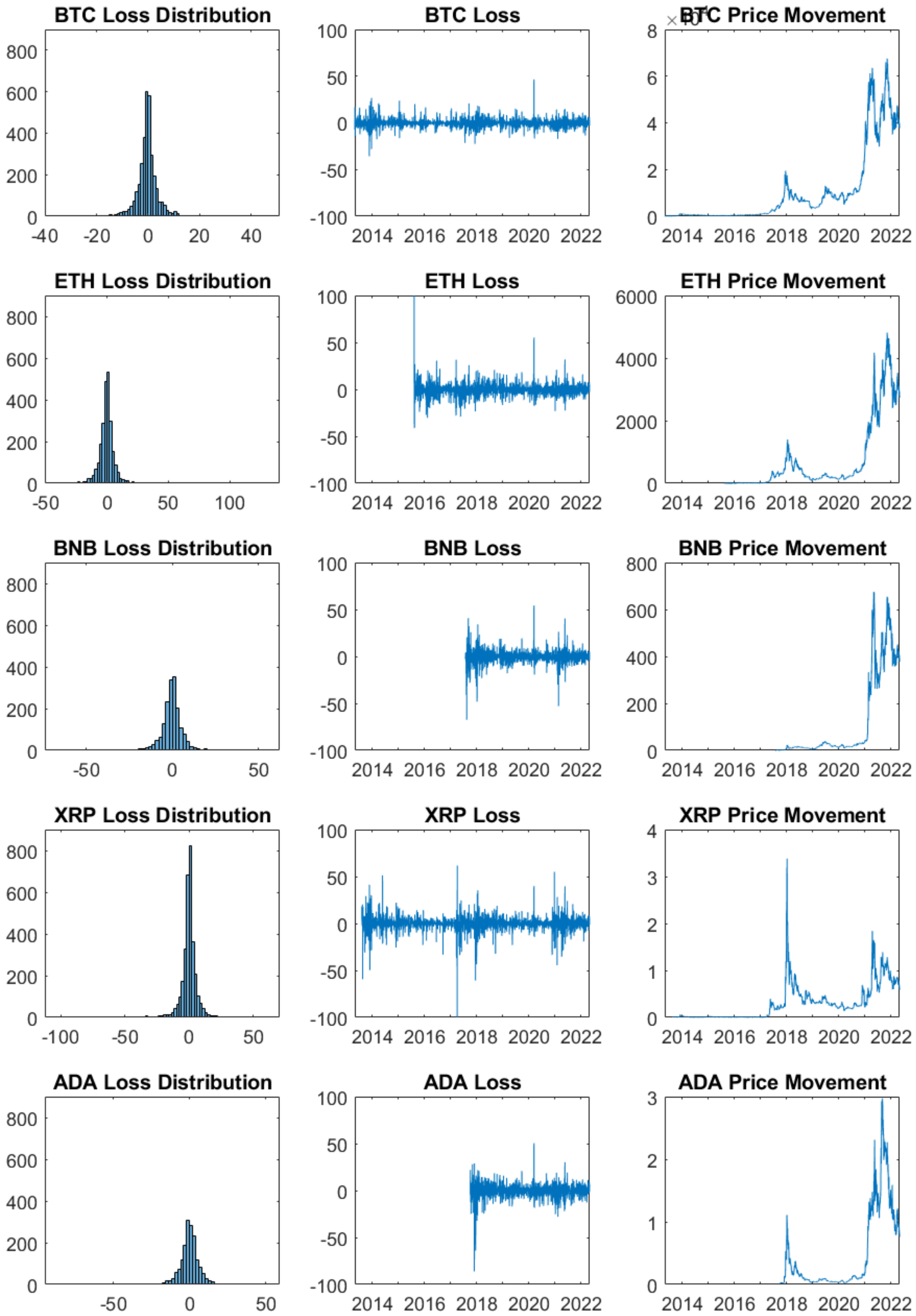
The ADF test statistics of all five cryptocurrencies are negative and far from zero, indicating that we can reject the null hypothesis of having unit roots in the time series of cryptocurrency losses. In other words, we can reject the possibility of a non-stationary process. Graph 4.3 also shows that the time series of cryptocurrency losses are to some extent stationary, while the cryptocurrency prices do not follow a stationary process.

Following the approach in section 3.1, Table 4.2 divides the total loss observations into a rolling estimation window of size 1000 and a testing window as below.

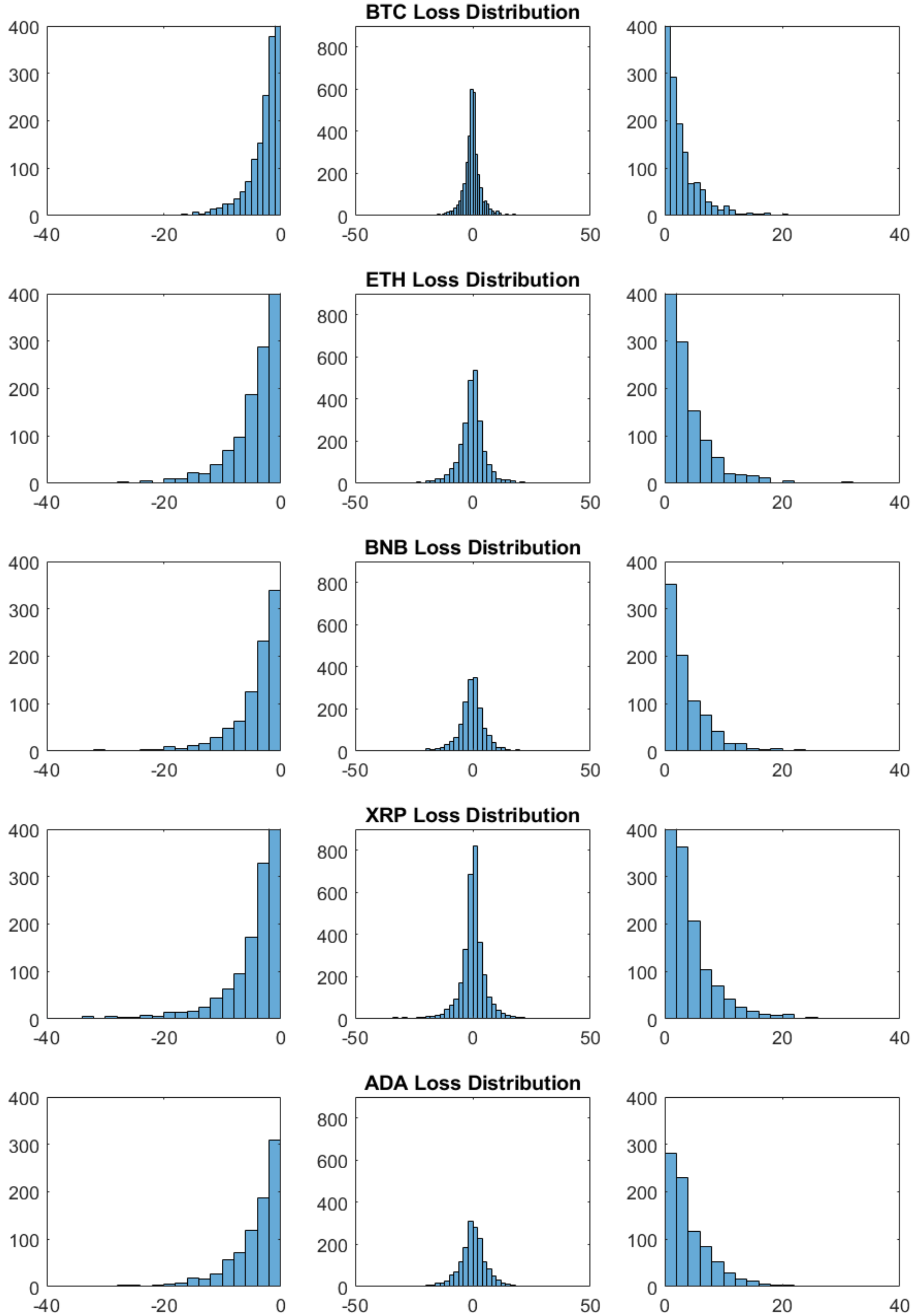
Table 4.2. *First estimation window and testing window*

	Total losses	**First estimation window**		**Testing window**	
		Period	**Days**	**Period**	**Days**
BTC	3,289	29/04/2013 - 23/01/2016	1,000	24/01/2016 - 30/04/2022	2,289
ETH	2,458	08/08/2015 - 03/05/2018	1,000	04/05/2018 - 30/04/2022	1,458
BNB	1,740	26/07/2017 - 20/04/2020	1,000	21/04/2020 - 30/04/2022	740
XRP	3,191	05/08/2013 - 30/04/2016	1,000	01/05/2016 - 30/04/2022	2,191
ADA	1,672	02/10/2017 - 27/06/2020	1,000	28/06/2020 - 30/04/2022	672

Graph 4.3. *Cryptocurrency loss, loss distribution, and price movement*



Graph 4.4. *Cryptocurrency loss distribution*



5. Results and analysis

5.1. Estimation results

Graphs 5.1 to 5.5 presents the results of one-day ahead VaR estimates for five cryptocurrencies (BTC, ETH, BNB, XRP, ADA) by different non-parametric and parametric VaR estimation methods (in Section 3.2), at 95% and 99% confident levels.

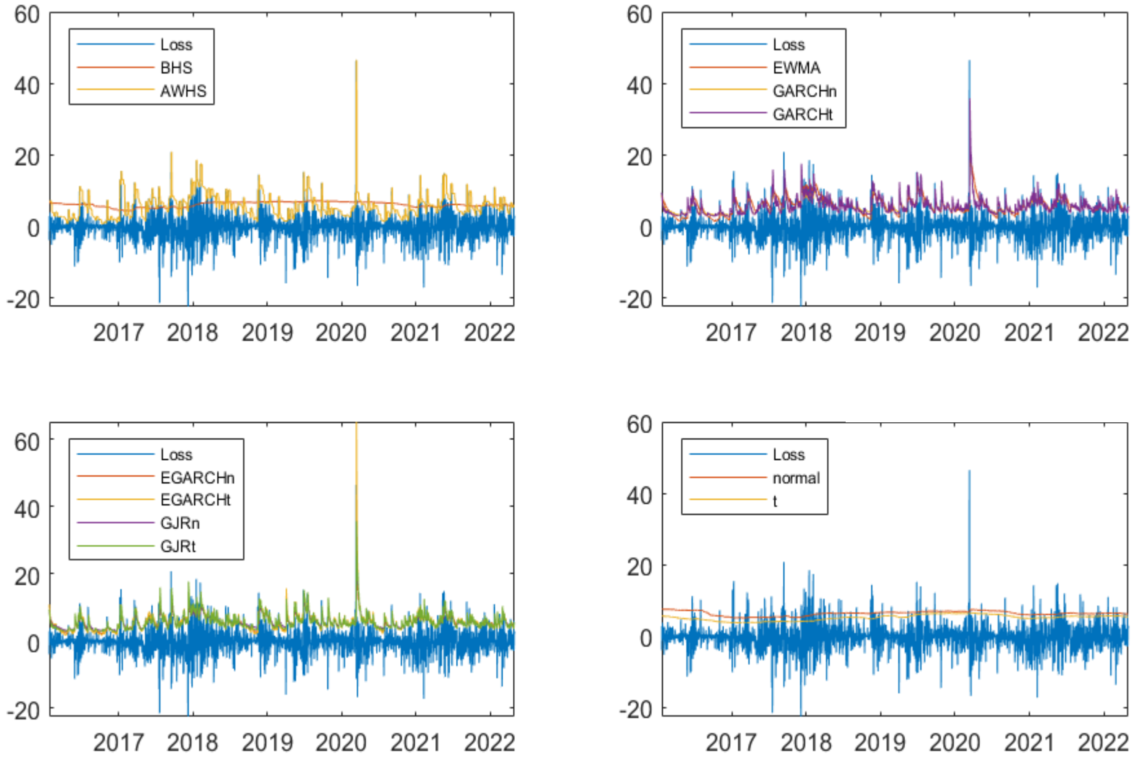
By visual inspection, the VWHS method using the EWMA model and all GARCH-type models reacts more swiftly to the market movement than other estimation models at all VaR levels. It is followed by the AWHS model that also has a fairly quick reaction to the loss fluctuation. On the contrary, the BHS method and all three parametric methods seem to react more slowly. It can be explained by the fact that these methods do not incorporate the time variation in VaR estimation.

Under most circumstances, the POT method generates higher VaR estimates than the two other parametric methods, since this method fits a distribution to a subsample of highest losses, not all loss observations, for the parameter estimation. As expected, the VaR estimates at 99% confidence level are higher than those at 95% VaR level, as it investigates further to the tails of the distribution which is more sensitive to outliers.

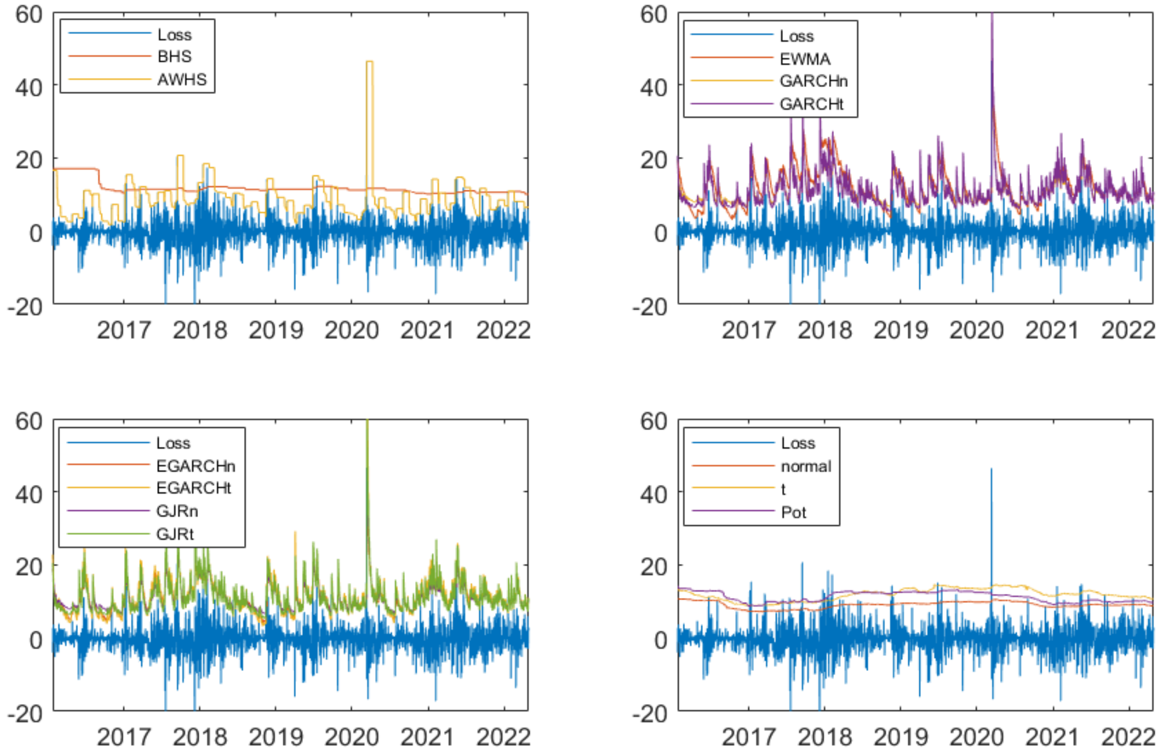
The VWHS method using EGARCH models tends to overreact to the market, meaning that it generates a substantially higher VaR estimate than the actual loss. The overreaction of EGARCH models can be clearly seen in Graph 5.1, 5.2, and 5.4. The GJR-GARCH models have the same tendency in several cases, especially at 99% level of confidence, for example in case of BTC and XRP (illustrated in Graph 5.1 (b) and Graph 5.4 (b) respectively).

Regarding BHS method, as noted in section 3.2.1.1, VaR and ES estimated by this method are subject to the uncertainty in estimation, hence we use the bootstrapping method to solve this problem. By resampling 1000 times with replacement, we are 95% certain that our loss ranges within the confidence interval 95% of the time (for VaR level of 95%) and 99% of the time (for VaR level of 99%). Further information about the VaR and ES confidence intervals for each cryptocurrency under the BHS method are stated in Appendix 5. Due to the greater uncertainty in ES estimates, it is observable that the confidence intervals for ES-BHS are wider than VaR-BHS in all cases. The distribution of VaR-BHS and ES-BHS is illustrated in Appendix 6.

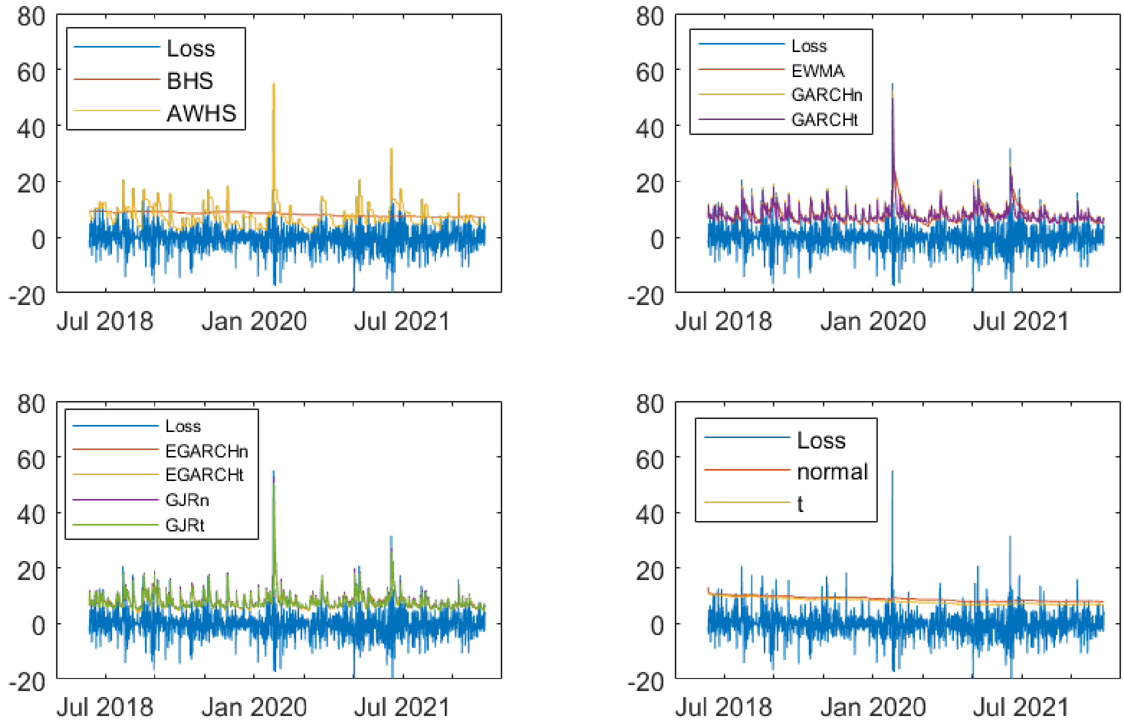
Graph 5.1. *(a) 95% one-day ahead VaR estimates (BTC)*



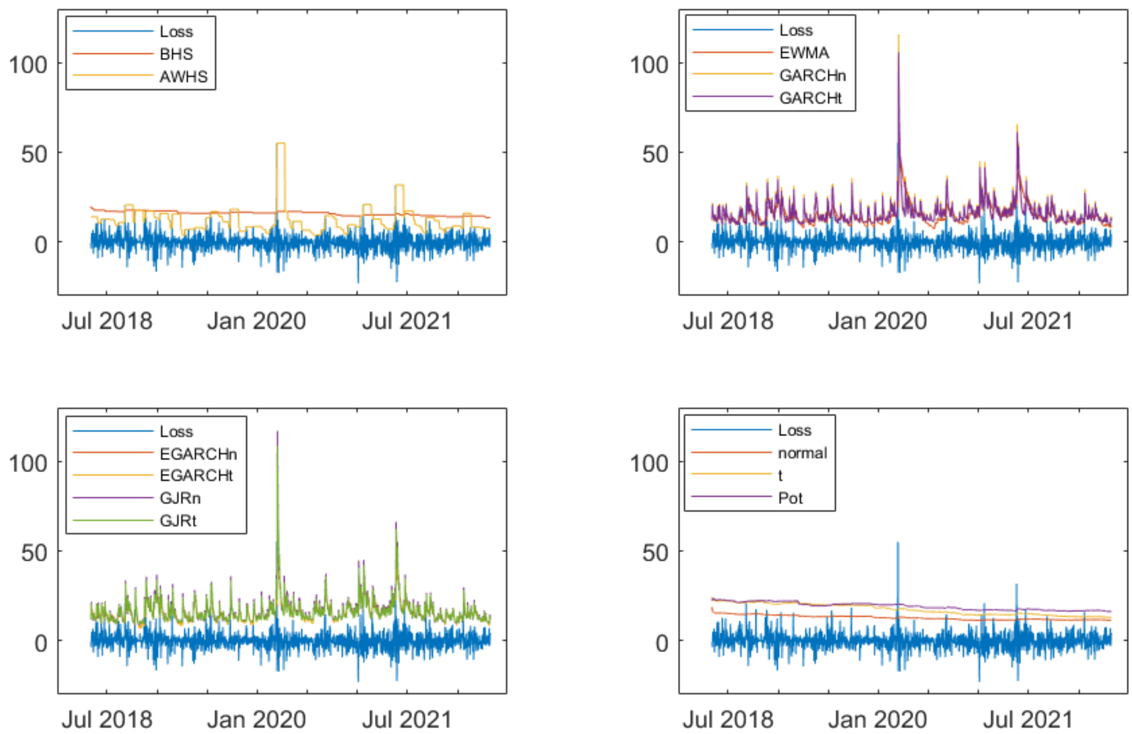
(b) 99% one-day ahead VaR estimates (BTC)



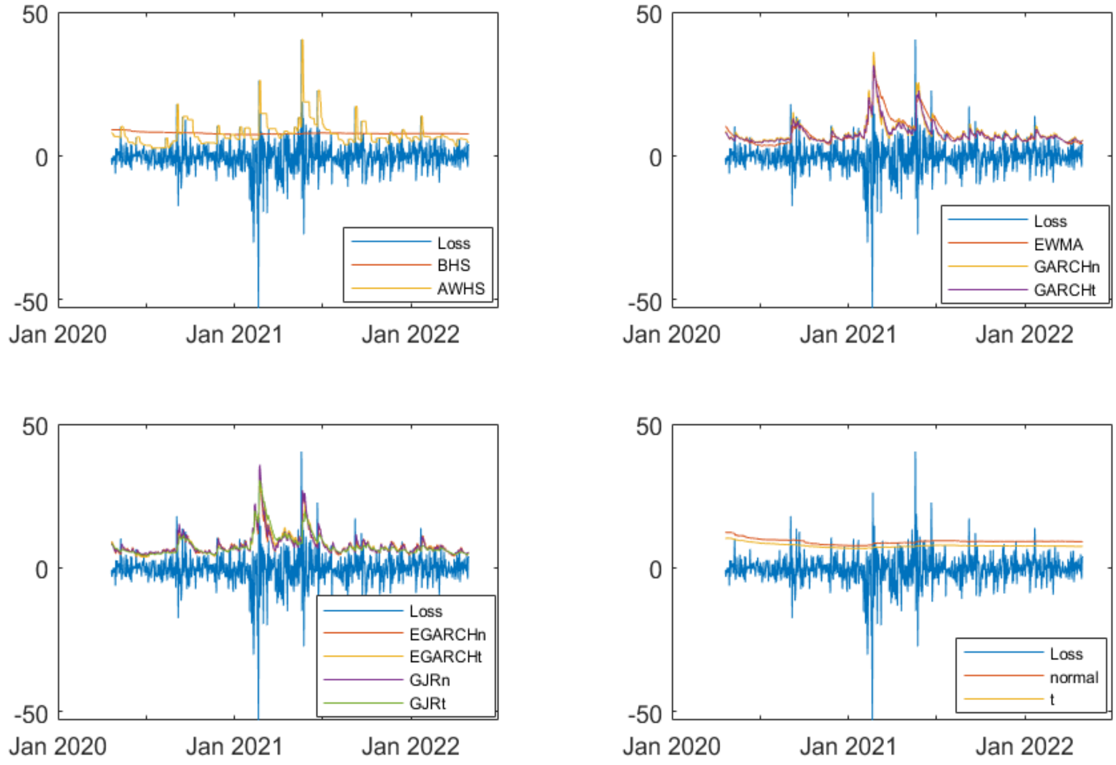
Graph 5.2. *(a) 95% one-day ahead VaR estimates (ETH)*



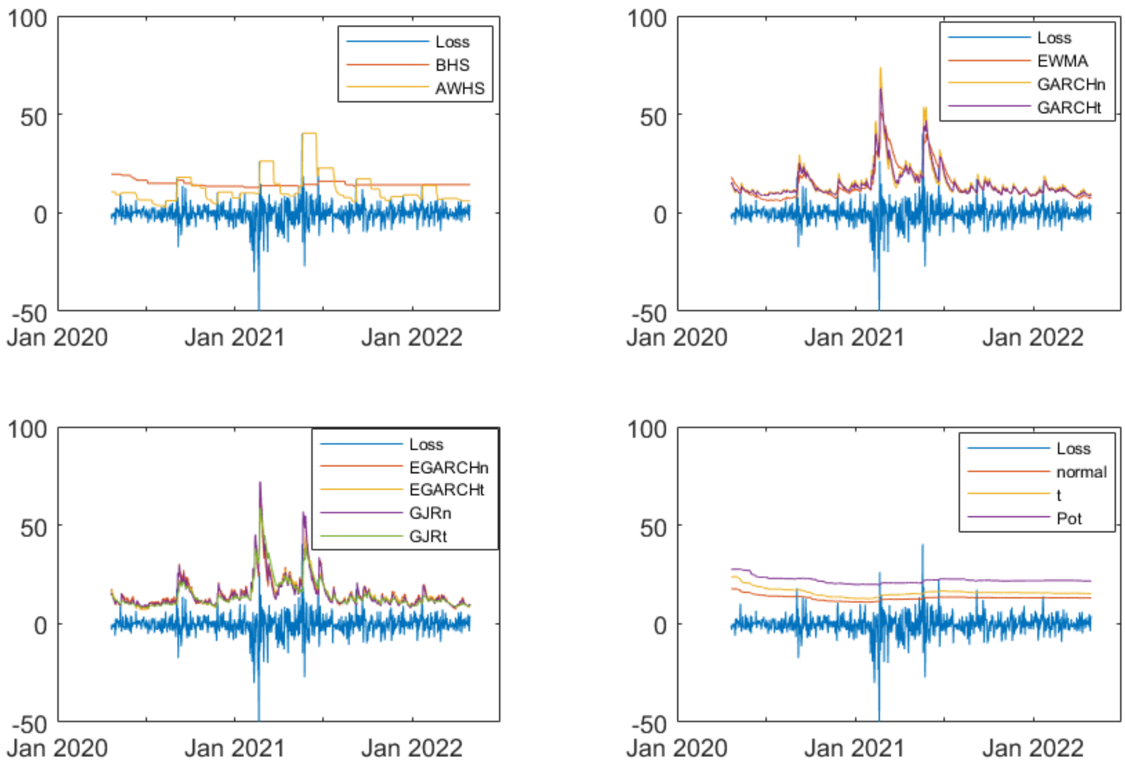
(b) 99% one-day ahead VaR estimates (ETH)



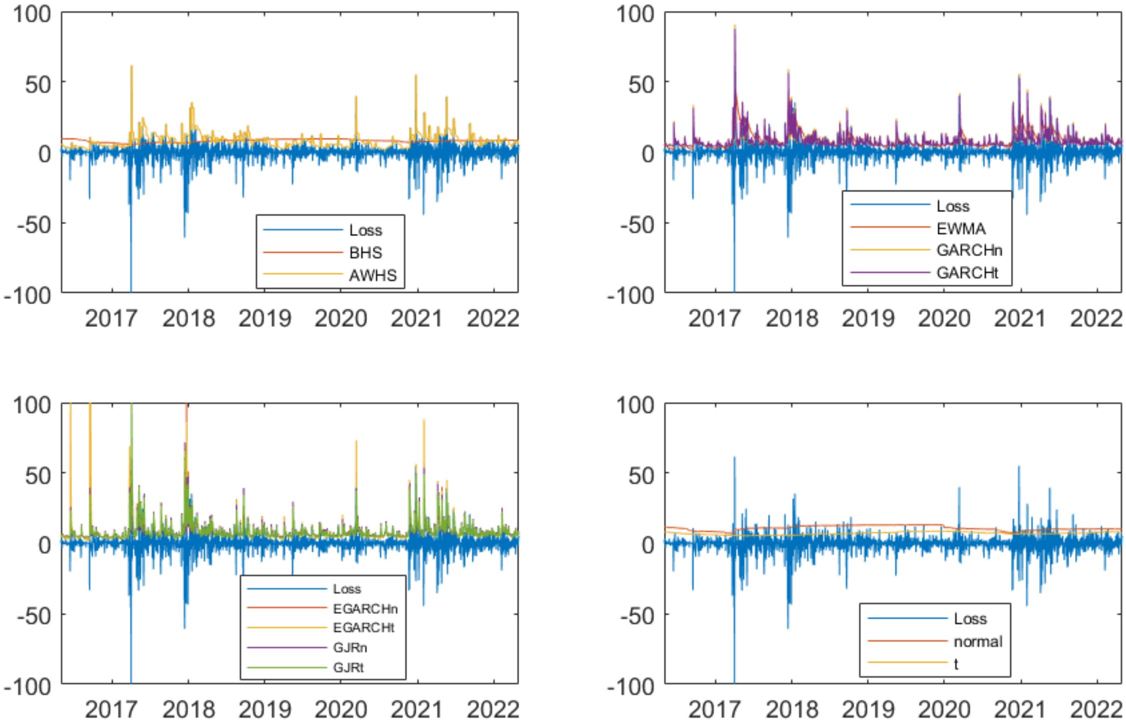
Graph 5.3. *(a) 95% one-day ahead VaR estimates (BNB)*



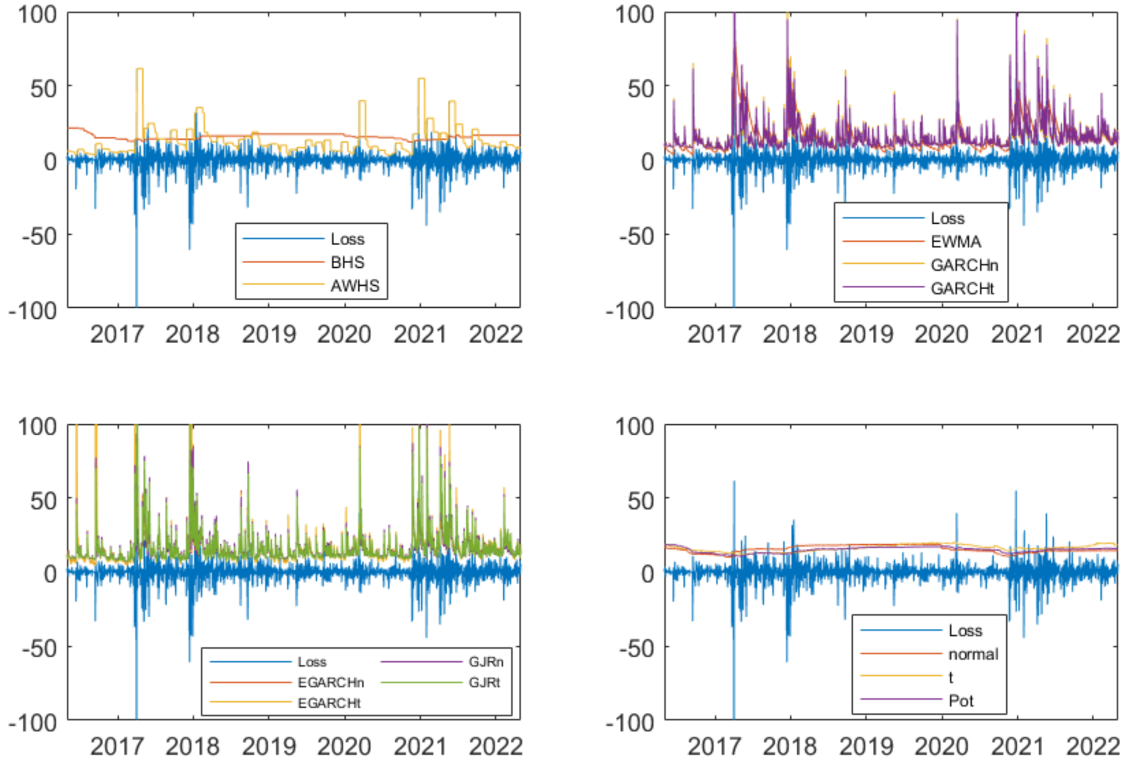
(b) 99% one-day ahead VaR estimates (BNB)



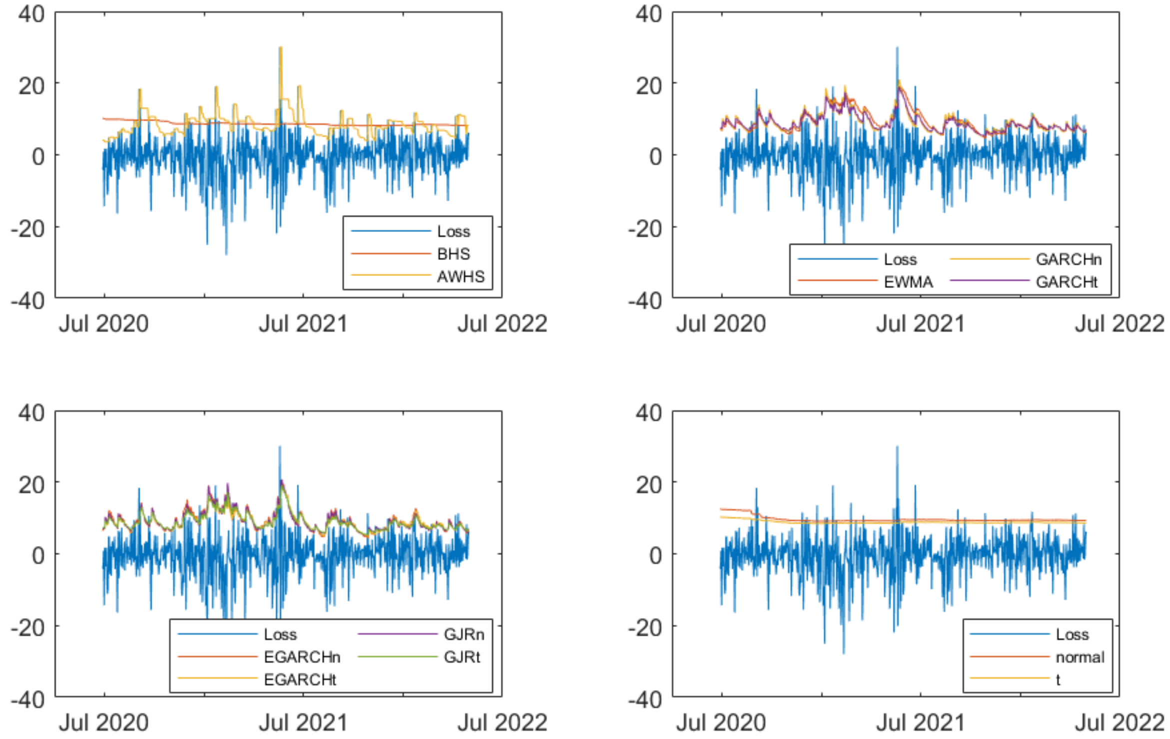
Graph 5.4. *(a) 95% one-day ahead VaR estimates (XRP)*



(b) 99% one-day ahead VaR estimates (XRP)



Graph 5.5. *(a) 95% one-day ahead VaR estimates (ADA)*



(b) 99% one-day ahead VaR estimates (ADA)

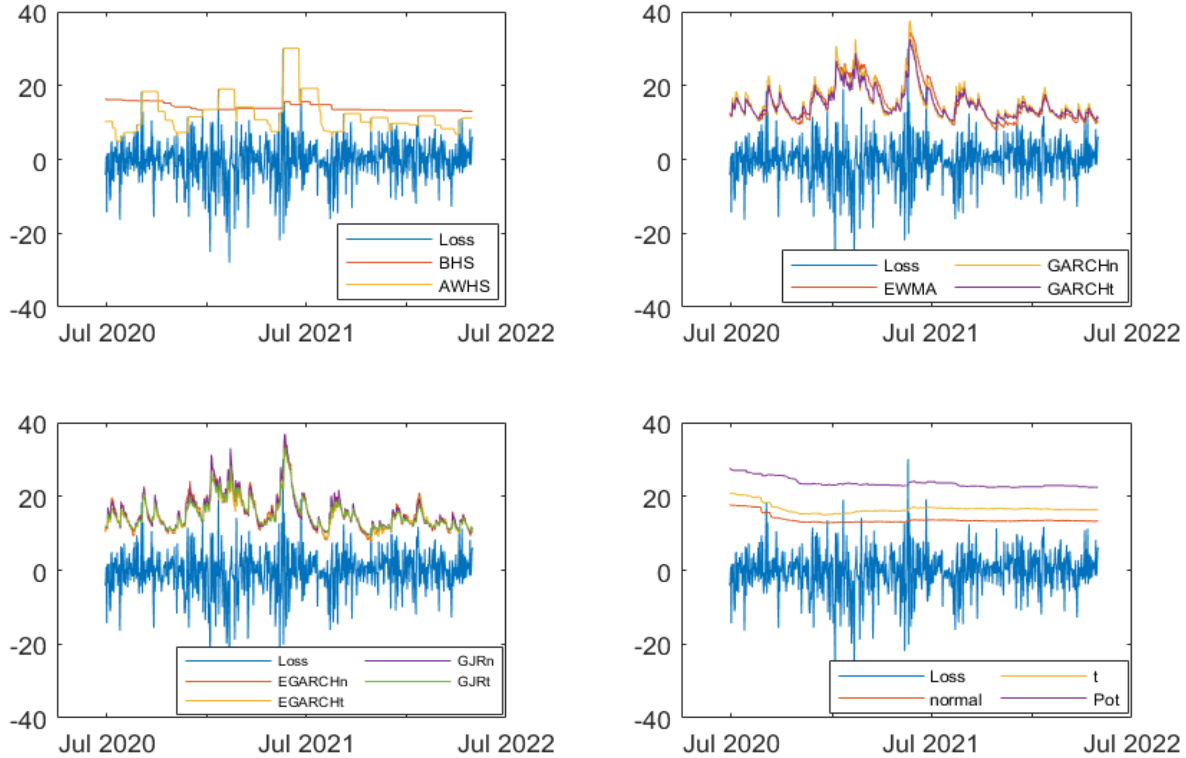


Table 5.6 below summarized all parameters of the GARCH-type models estimated by the maximum likelihood method and the standard parameter values of the EWMA model as proposed by the Basel Committee. For most cryptocurrencies excluding XRP, the GARCH (1) term is typically higher than the ARCH (1) term, indicating that the past variance has a more considerable influence on the current conditional variance than the past innovation. This is somewhat similar to the standard parameters ($\gamma = 0.94$) of the EWMA model. The parameters of the GARCH-type models are statistically significant. As for XRP, the impacts of the historical variance and past innovation on the conditional variance are relatively equal.

For all cryptocurrencies, it is observable that the EGARCH models have the highest value of the GARCH (1) term and the lowest value of the constant term. The differences in parameter estimation between the EGARCH models and other GARCH-type models are particularly striking in cases of ETH and XRP. It is possibly the reason for the overreaction of EGARCH models to the market as observed before.

The leverage terms of the EGARCH and the GJR-GARCH models are negative in several cases (see Table 5.6). Notably, ETH experiences negative leverage terms in all EGARCH and GJR-GARCH models. Similarly, the leverage terms of the EGARCH models for XRP are also smaller than zero. If the leverage term is positive, then the bad news has a greater effect on volatility than good news. The reverse is true for the negative leverage term, meaning that the volatility of cryptocurrencies is likely to be higher after positive news. This observation is to some extent different from the conclusion of Obeng (2021) that bad news has a more significant impact on the market. In our cases, we observe both positive and negative leverage terms. Such difference may result from the fact that Obeng (2021) observed the price index of all 30 cryptocurrencies for a period of only four years from 2017 to 2020, rather than examining a particular cryptocurrency for a longer period.

Table 5.6. *Maximum likelihood estimation of GARCH-type model parameters and standard EWMA parameters*

BTC		**Constant** κ	**GARCH(1)** γ	**ARCH(1)** α	**Leverage** ξ
EWMA (standard)	value	N/A	0.94	0.06	N/A
GARCHn	value	0.9913	0.8321	0.1242	N/A
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	
GARCHt	value	1.0584	0.7641	0.2359	N/A
	p-value	*(0.0021)*	*(0.0000)*	*(0.0000)*	
EGARCHn	value	0.1715	0.9507	0.2420	0.0224
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0025)*
EGARCHt	value	0.1904	0.9411	0.4955	-0.0035
	p-value	*(0.0023)*	*(0.0000)*	*(0.0000)*	*(0.9361)*
GJRGARCHn	value	1.0008	0.8311	0.1219	0.0059
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.7155)*
GJRGARCHt	value	1.0600	0.7636	0.2328	0.0072
	p-value	*(0.0021)*	*(0.0000)*	*(0.0011)*	*(0.9175)*
ETH		κ	γ	α	ξ
EWMA (standard)	value	N/A	0.94	0.06	N/A
GARCHn	value	3.9964	0.6386	0.3615	N/A
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	
GARCHt	value	3.6216	0.6810	0.3190	N/A
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	
EGARCHn	value	0.3505	0.9146	0.3915	-0.0360
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0035)*
EGARCHt	value	0.3223	0.9148	0.4283	-0.0213
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.5135)*
GJRGARCHn	value	4.0247	0.6375	0.3686	-0.0122
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.7918)*
GJRGARCHt	value	3.6697	0.6785	0.3415	-0.0400
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.6447)*
BNB		κ	γ	α	ξ
EWMA (standard)	value	N/A	0.94	0.06	N/A
GARCHn	value	1.7292	0.8003	0.1879	N/A
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	
GARCHt	value	1.1245	0.8635	0.1365	N/A
	p-value	*(0.0080)*	*(0.0000)*	*(0.0000)*	

Table 5.6. *(continued)*

***BNB** (continued)*		κ	γ	α	ξ
EGARCHn	value	0.2130	0.9523	0.3595	0.0242
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0085)*
EGARCHt	value	0.0770	0.9801	0.2532	-0.0177
	p-value	*(0.0145)*	*(0.0000)*	*(0.0000)*	*(0.4955)*
GJRGARCHn	value	1.6885	0.7993	0.2078	-0.0307
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0553)*
GJRGARCHt	value	0.6943	0.9080	0.0661	0.0519
	p-value	*(0.0156)*	*(0.0000)*	*(0.0112)*	*(0.0000)*

XRP		κ	γ	α	ξ
EWMA (standard)	value	N/A	0.94	0.06	N/A
GARCHn	value	5.0469	0.4879	0.5121	N/A
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	
GARCHt	value	3.5141	0.5651	0.4349	N/A
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	
EGARCHn	value	0.7266	0.8195	0.7222	-0.1044
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
EGARCHt	value	0.5404	0.8592	0.8698	-0.0919
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0991)*
GJRGARCHn	value	5.7782	0.4406	0.4003	0.3183
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
GJRGARCHt	value	3.8295	0.5403	0.3639	0.1916
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0954)*

ADA		κ	γ	α	ξ
EWMA (standard)	value	N/A	0.94	0.06	N/A
GARCHn	*value*	*1.6894*	*0.8539*	*0.1221*	*N/A*
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	
GARCHt	*value*	*1.3539*	*0.8945*	*0.0868*	*N/A*
	p-value	*(0.0124)*	*(0.0000)*	*(0.0008)*	
EGARCHn	*value*	*0.1719*	*0.9616*	*0.2706*	*-0.0133*
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0830)*
EGARCHt	*value*	*0.1107*	*0.9704*	*0.2068*	*0.0090*
	p-value	*(0.0160)*	*(0.0000)*	*(0.0000)*	*(0.7081)*
GJRGARCHn	*value*	*1.7621*	*0.8526*	*0.1152*	*0.0121*
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.2874)*
GJRGARCHt	*value*	*1.3966*	*0.8914*	*0.0953*	*-0.0103*
	p-value	*(0.0136)*	*(0.0000)*	*(0.0053)*	*(0.7565)*

5.2. Backtesting results

5.2.1. Summary report on violations

Table 5.7 below summarizes the number of actual VaR violations for all cryptocurrencies, in comparison with the expected number of violations at 95% and 99% confidence levels, given that the VaR estimation method is correct. In this table, the red-colored number illustrates the highest number of VaR violations at a certain level of confidence for each cryptocurrency, while the blue number indicates the lowest one.

Table 5.7. *Summary report on number of observations and failures*

		95% confidence level			99% confidence level		
VaR methods	***Obs.***	***failures***	***expected***	***1st failure***	***failures***	***expected***	***1st failure***
BTC							
BHS	2289	124	114.45	150	20	22.89	348
AWHS	2289	144	114.45	31	56	22.89	95
EWMA	2289	116	114.45	95	19	22.89	95
GARCHn	2289	121	114.45	95	24	22.89	192
GARCHt	2289	124	114.45	95	24	22.89	192
EGARCHn	2289	117	114.45	95	23	22.89	192
EGARCHt	2289	118	114.45	95	22	22.89	192
GJRGARCHn	2289	121	114.45	95	24	22.89	192
GJRGARCHt	2289	124	114.45	95	23	22.89	192
Normal-fit	2289	112	114.45	150	51	22.89	151
t-fit	2289	167	114.45	150	23	22.89	348
POT	2289				23	22.89	348
ETH							
BHS	1458	63	72.9	20	10	14.58	125
AWHS	1458	88	72.9	8	43	14.58	38
EWMA	1458	74	72.9	19	17	14.58	38
GARCHn	1458	72	72.9	8	16	14.58	38
GARCHt	1458	74	72.9	8	16	14.58	38
EGARCHn	1458	72	72.9	19	15	14.58	38
EGARCHt	1458	74	72.9	19	16	14.58	38
GJRGARCHn	1458	73	72.9	8	16	14.58	38
GJRGARCHt	1458	73	72.9	8	16	14.58	38
Normal-fit	1458	49	72.9	38	23	14.58	125
t-fit	1458	65	72.9	20	7	14.58	675
POT	1458				3	14.58	679

Table 5.7. (continued)

BNB							
BHS	740	31	37	20	7	7.4	136
AWHS	740	42	37	20	17	7.4	20
EWMA	740	37	37	20	5	7.4	136
GARCHn	740	36	37	20	7	7.4	20
GARCHt	740	36	37	20	7	7.4	20
EGARCHn	740	33	37	20	6	7.4	20
EGARCHt	740	35	37	20	6	7.4	20
GJRGARCHn	740	36	37	20	7	7.4	20
GJRGARCHt	740	37	37	20	7	7.4	20
Normal-fit	740	22	37	136	10	7.4	136
t-fit	740	38	37	20	7	7.4	136
POT	740				3	7.4	306
XRP							
BHS	2191	125	109.55	139	26	21.91	329
AWHS	2191	129	109.55	25	56	21.91	52
EWMA	2191	119	109.55	25	24	21.91	139
GARCHn	2191	126	109.55	139	30	21.91	139
GARCHt	2191	121	109.55	139	32	21.91	139
EGARCHn	2191	121	109.55	87	29	21.91	139
EGARCHt	2191	121	109.55	25	30	21.91	139
GJRGARCHn	2191	116	109.55	139	29	21.91	139
GJRGARCHt	2191	116	109.55	139	30	21.91	139
Normal-fit	2191	68	109.55	139	25	21.91	329
t-fit	2191	162	109.55	139	32	21.91	329
POT	2191				31	21.91	329
ADA							
BHS	672	34	33.6	68	7	6.72	68
AWHS	672	50	33.6	12	23	6.72	30
EWMA	672	31	33.6	55	5	6.72	68
GARCHn	672	26	33.6	55	6	6.72	68
GARCHt	672	29	33.6	55	5	6.72	68
EGARCHn	672	28	33.6	53	5	6.72	68
EGARCHt	672	29	33.6	53	5	6.72	68
GJRGARCHn	672	26	33.6	55	6	6.72	68
GJRGARCHt	672	29	33.6	55	5	6.72	68
Normal-fit	672	27	33.6	68	7	6.72	68
t-fit	672	33	33.6	68	4	6.72	68
POT	672				1	6.72	326

Overall, the first violation happens earlier when using AWHS and VWHS estimation methods than BHS or parametric methods. Especially, it occurs very late in case we fit student t-distribution to loss observations or use POT methods at 99% confidence level. While BTC, ETH, and XRP seem to underestimate VaR in most cases at both 95% and 99% confidence level, BNB and ADA mostly experience VaR overestimation.

At 95% confidence level, the normal distribution fitting method witnesses the lowest number of VaR violations for four cryptocurrencies apart from ADA. The first violation when using this method appears relatively late in most cases. Meanwhile, the lowest number of violations for ADA is observed when we use GARCHn and GJRGARCHn method.

The student t-distribution fitting method and AWHS seem to have the worst performance in terms of violations at 95% confidence level. Using AWHS method results in the highest number of failures for ETH, BNB, ADA, and also the second-highest number for BTC and XRP at the confidence level of 95%. The largest number of violations is witnessed when we use the student t-distribution fitting method for BTC and XRP at this confidence level.

At 99% confidence level, the POT method seems to perform the best for ETH, BNB, and ADA regarding the number of VaR violations, while using EWMA method leads to the smallest number of failures for BTC and XRP. In contrast, AWHS has the highest 99% VaR violations for all cryptocurrencies. Also, the first violation appears very early when using this method in all cases.

Compared to the expected number of violations, the VWHS estimation method using GARCH-type models for BTC and ETH mostly generates a closer number of actual VaR violations to the expectation than other estimation methods, at both confidence level. However, this observation is not applicable to other cryptocurrencies. As for the number of VaR violations, all GARCH models and the EWMA model have a relatively similar performance.

In conclusion, considering the number of violations and the first failure, the AWHS method seems to perform the worst with the early failure occurrence and high probability of VaR violation at both confidence levels. The POT method has an outstanding performance at 99% confidence level. The number of failures using this method is low, and the first failure happens relatively late compared to other estimation methods. Meanwhile, at 95% confidence level, the normal distribution fitting method appears to have a good performance in most circumstances. It is notable that the above conclusions are merely drawn from the number of observed violations, not the extent to which this number of violations is acceptable at the given statistical test level (it may be too high or too low compared to the expected value). This matter will be discussed in the next section.

5.2.2. VaR backtesting results

Table 5.8 summarizes all VaR backtesting results for five cryptocurrencies at 95% and 99% VaR levels. This table shows the results of the traffic light test (TL), the Kupiec test (POF), and the Christoffersen tests (CCI and CC) as introduced in the methodology.

The AWHS method consistently delivers a poor performance in most cases. It fails most of the tests at both 95% and 99% confidence levels, except for the independence test (CCI test) of all cryptocurrencies at 95% VaR level. The exceptional case is BNB where the AWHS method does not fail any test, however, it still has the highest number of violations among all methods. Therefore, it seems to perform poorly overall.

Despite having the lowest number of violations for four cryptocurrencies (apart from ADA) at the confidence level of 95%, the normal distribution fitting method fails at least two tests for each cryptocurrency. The only exception for this method at 95% VaR level is ADA, where it passes all tests and shows a relatively good performance. At 99% confidence level, this method does not perform well for BTC and ETH, while it still gives a satisfying performance for the other three cryptocurrencies.

As for BTC, all estimation methods fail the CCI test at 95% confidence level except for AWHS. However, the AWHS method gets rejected in both frequency tests (TL test and POF test) and the conditional coverage test (CC). Also, almost all methods excluding EWMA do not pass the CC test. Therefore, EWMA proves to the best estimation model for BTC at 95% VaR level. It also passes all the backtesting procedures for BTC and generates the least number of violations at 99% VaR level (19 violations). All GARCH-type models seem to perform well for BTC. However, these methods produce more violations than EWMA model at both levels, thus they are not possibly the best.

Regarding ETH, the EWMA method and the BHS method pass all tests at both confidence levels, while other methods get rejections at the minimum of two tests. Considering the number of actual violations, the BHS method performs better than the EWMA method. All GARCH-type models fail the conditional coverage test and independence test at 95% VaR level. This result suggests that the VaR violations may not be independent, and as a consequence, it is likely that a violation may follow another one in the previous day. These models seem to perform better at 99% VaR level without failing any test, but the number of failures produced by these models are fairly high.

Considering BNB, the POT method is obviously the best method at 99% confidence level, which passes all the tests and produces the lowest number of failures. Except for the AWHS, normal and t-fitted methods, the remaining methods pass all of the tests. Among them, the BHS method stands out with the lowest number of violations at 95% VaR level, and EWMA generates the second-lowest number of failures at 99% VaR level (only higher than POT method).

All estimation methods get rejected at least one test in the case of XRP. The EWMA method continues to perform fairly better than others, given that the number of tests it fails is smallest. In addition, this method gives the lowest number of failures at 99% confidence level, and relatively low number of violations at 95% confidence level.

Although the POT method produces only one failure for risk estimates of ADA, it does not pass the frequency test and the conditional coverage test. Apart from the AWHS method and the POT method, none of the other methods fail any test at both confidence levels. The number of violations generated by each method is roughly similar. At 95% VaR level, the GARCHn and the GJR-GARCHn have the lowest failure number of 26, followed by the EGARCHn with 28 failures. At 99% VaR level, the t-fitted method observes 4 failures, and other GARCH-type methods witness around 5 to 6 violations.

To sum up, for all cryptocurrencies, the EWMA method and the BHS method prove to have a remarkable performance. All GARCH-type models perform relatively well, but not outperform other methods due to the higher number of violations. The 99% POT method is the best method for BNB.

Table 5.8. *VaR backtesting results*

	METHOD	TL	POF			CC			CCI		
		Result	Result	LR ratio	P-value	Result	LR ratio	P-value	Result	LR ratio	P-value
BTC 95%	BHS	green	accept	0.818	0.366	reject	7.757	0.021	reject	6.939	0.008
	AWHS	yellow	reject	7.450	0.006	reject	7.895	0.019	accept	0.445	0.505
	EWMA	green	accept	0.022	0.882	accept	5.611	0.060	reject	5.589	0.018
	GARCHn	green	accept	0.388	0.534	reject	8.167	0.017	reject	7.779	0.005
	GARCHt	green	accept	0.818	0.366	reject	11.568	0.003	reject	10.750	0.001
	EGARCHn	green	accept	0.059	0.807	reject	7.132	0.028	reject	7.073	0.008
	EGARCHt	green	accept	0.115	0.735	reject	18.027	0.000	reject	17.912	0.000
	GJRn	green	accept	0.388	0.534	reject	8.167	0.017	reject	7.779	0.005
	GJRt	green	accept	0.818	0.366	reject	11.568	0.003	reject	10.750	0.001
	Normal-fit	green	accept	0.056	0.814	reject	8.583	0.014	reject	8.527	0.003
	t-fit	red	reject	22.384	0.000	reject	34.943	0.000	reject	12.559	0.000
BTC 99%	BHS	green	accept	0.385	0.535	accept	2.293	0.318	accept	1.908	0.167
	AWHS	red	reject	34.467	0.000	reject	38.037	0.000	accept	3.570	0.059
	EWMA	green	accept	0.709	0.400	accept	2.794	0.247	accept	2.085	0.149
	GARCHn	green	accept	0.054	0.817	accept	0.562	0.755	accept	0.509	0.476
	GARCHt	green	accept	0.054	0.817	accept	0.562	0.755	accept	0.509	0.476
	EGARCHn	green	accept	0.001	0.982	accept	1.445	0.486	accept	1.444	0.229
	EGARCHt	green	accept	0.035	0.851	accept	0.463	0.794	accept	0.427	0.513
	GJRn	green	accept	0.054	0.817	accept	0.562	0.755	accept	0.509	0.476
	GJRt	green	accept	0.001	0.982	accept	0.468	0.792	accept	0.467	0.494
	Normal-fit	red	reject	25.845	0.000	reject	28.084	0.000	accept	2.239	0.135
	t-fit	green	accept	0.001	0.982	accept	5.377	0.068	reject	5.376	0.020
	POT	green	accept	0.001	0.982	accept	1.445	0.486	accept	1.444	0.229

Table 5.8. *VaR backtesting results (continued)*

	METHOD	TL	POF			CC			CCI		
		Result	Result	LR ratio	P-value	Result	LR ratio	P-value	Result	LR ratio	P-value
ETH 95%	*BHS*	*green*	*accept*	*1.480*	*0.224*	*accept*	*1.510*	*0.470*	*accept*	*0.030*	*0.863*
	AWHS	*yellow*	*accept*	*3.097*	*0.078*	*accept*	*3.193*	*0.203*	*accept*	*0.096*	*0.756*
	EWMA	*green*	*accept*	*0.017*	*0.895*	*accept*	*2.553*	*0.279*	*accept*	*2.536*	*0.111*
	GARCHn	*green*	*accept*	*0.012*	*0.914*	*reject*	*6.744*	*0.034*	*reject*	*6.732*	*0.009*
	GARCHt	*green*	*accept*	*0.017*	*0.895*	*reject*	*6.076*	*0.048*	*reject*	*6.058*	*0.014*
	EGARCHn	*green*	*accept*	*0.012*	*0.914*	*reject*	*9.080*	*0.011*	*reject*	*9.068*	*0.003*
	EGARCHt	*green*	*accept*	*0.017*	*0.895*	*reject*	*8.279*	*0.016*	*reject*	*8.262*	*0.004*
	GJRn	*green*	*accept*	*0.000*	*0.990*	*reject*	*6.390*	*0.041*	*reject*	*6.390*	*0.011*
	GJRt	*green*	*accept*	*0.000*	*0.990*	*reject*	*6.390*	*0.041*	*reject*	*6.390*	*0.011*
	Normal-fit	*green*	*reject*	*9.278*	*0.002*	*reject*	*9.353*	*0.009*	*accept*	*0.076*	*0.783*
	t-fit	*green*	*accept*	*0.934*	*0.334*	*accept*	*0.938*	*0.626*	*accept*	*0.004*	*0.951*
ETH 99%	*BHS*	*green*	*accept*	*1.633*	*0.201*	*accept*	*1.771*	*0.412*	*accept*	*0.138*	*0.710*
	AWHS	*red*	*reject*	*36.737*	*0.000*	*reject*	*37.120*	*0.000*	*accept*	*0.384*	*0.536*
	EWMA	*green*	*accept*	*0.385*	*0.535*	*accept*	*2.096*	*0.351*	*accept*	*1.710*	*0.191*
	GARCHn	*green*	*accept*	*0.135*	*0.713*	*accept*	*0.491*	*0.782*	*accept*	*0.355*	*0.551*
	GARCHt	*green*	*accept*	*0.135*	*0.713*	*accept*	*0.491*	*0.782*	*accept*	*0.355*	*0.551*
	EGARCHn	*green*	*accept*	*0.012*	*0.912*	*accept*	*2.156*	*0.340*	*accept*	*2.144*	*0.143*
	EGARCHt	*green*	*accept*	*0.135*	*0.713*	*accept*	*2.053*	*0.358*	*accept*	*1.917*	*0.166*
	GJRn	*green*	*accept*	*0.135*	*0.713*	*accept*	*0.491*	*0.782*	*accept*	*0.355*	*0.551*
	GJRt	*green*	*accept*	*0.135*	*0.713*	*accept*	*0.491*	*0.782*	*accept*	*0.355*	*0.551*
	Normal-fit	*yellow*	*reject*	*4.178*	*0.041*	*accept*	*4.916*	*0.086*	*accept*	*0.738*	*0.390*
	t-fit	*green*	*reject*	*4.927*	*0.026*	*accept*	*4.995*	*0.082*	*accept*	*0.068*	*0.795*
	POT	*green*	*reject*	*13.766*	*0.000*	*reject*	*13.779*	*0.001*	*accept*	*0.012*	*0.911*

Table 5.8. *VaR backtesting results (continued)*

	METHOD	TL	POF			CC			CCI		
		Result	Result	LR ratio	P-value	Result	LR ratio	P-value	Result	LR ratio	P-value
BNB 95%	*BHS*	*green*	*accept*	*1.081*	*0.298*	*accept*	*2.900*	*0.235*	*accept*	*1.819*	*0.177*
	AWHS	*green*	*accept*	*0.683*	*0.409*	*accept*	*1.816*	*0.403*	*accept*	*1.133*	*0.287*
	EWMA	*green*	*accept*	*-0.000*	*1.000*	*accept*	*0.013*	*0.994*	*accept*	*0.013*	*0.910*
	GARCHn	*green*	*accept*	*0.029*	*0.865*	*accept*	*0.852*	*0.653*	*accept*	*0.823*	*0.364*
	GARCHt	*green*	*accept*	*0.029*	*0.865*	*accept*	*0.852*	*0.653*	*accept*	*0.823*	*0.364*
	EGARCHn	*green*	*accept*	*0.472*	*0.492*	*accept*	*0.659*	*0.719*	*accept*	*0.187*	*0.665*
	EGARCHt	*green*	*accept*	*0.116*	*0.734*	*accept*	*1.103*	*0.576*	*accept*	*0.987*	*0.321*
	GJRn	*green*	*accept*	*0.029*	*0.865*	*accept*	*0.852*	*0.653*	*accept*	*0.823*	*0.364*
	GJRt	*green*	*accept*	*-0.000*	*1.000*	*accept*	*0.013*	*0.994*	*accept*	*0.013*	*0.910*
	Normal-fit	*green*	*reject*	*7.443*	*0.006*	*reject*	*7.611*	*0.022*	*accept*	*0.168*	*0.682*
	t-fit	*green*	*accept*	*0.028*	*0.867*	*accept*	*3.876*	*0.144*	*reject*	*3.847*	*0.050*
BNB 99%	*BHS*	*green*	*accept*	*0.022*	*0.881*	*accept*	*0.156*	*0.925*	*accept*	*0.134*	*0.714*
	AWHS	*yellow*	*reject*	*9.205*	*0.002*	*reject*	*9.911*	*0.007*	*accept*	*0.706*	*0.401*
	EWMA	*green*	*accept*	*0.887*	*0.346*	*accept*	*0.956*	*0.620*	*accept*	*0.068*	*0.794*
	GARCHn	*green*	*accept*	*0.022*	*0.881*	*accept*	*0.156*	*0.925*	*accept*	*0.134*	*0.714*
	GARCHt	*green*	*accept*	*0.022*	*0.881*	*accept*	*0.156*	*0.925*	*accept*	*0.134*	*0.714*
	EGARCHn	*green*	*accept*	*0.286*	*0.593*	*accept*	*0.384*	*0.825*	*accept*	*0.098*	*0.754*
	EGARCHt	*green*	*accept*	*0.286*	*0.593*	*accept*	*0.384*	*0.825*	*accept*	*0.098*	*0.754*
	GJRn	*green*	*accept*	*0.022*	*0.881*	*accept*	*0.156*	*0.925*	*accept*	*0.134*	*0.714*
	GJRt	*green*	*accept*	*0.022*	*0.881*	*accept*	*0.156*	*0.925*	*accept*	*0.134*	*0.714*
	Normal-fit	*green*	*accept*	*0.831*	*0.362*	*accept*	*1.106*	*0.575*	*accept*	*0.274*	*0.600*
	t-fit	*green*	*accept*	*0.022*	*0.881*	*accept*	*0.156*	*0.925*	*accept*	*0.134*	*0.714*
	POT	*green*	*accept*	*3.409*	*0.065*	*accept*	*3.434*	*0.180*	*accept*	*0.024*	*0.876*

Table 5.8. *VaR backtesting results (continued)*

	METHOD	TL	POF			CC			CCI		
		Result	Result	LR ratio	P-value	Result	LR ratio	P-value	Result	LR ratio	P-value
XRP 95%	BHS	green	accept	2.198	0.138	reject	13.737	0.001	reject	11.539	0.001
	AWHS	yellow	accept	3.448	0.063	accept	4.236	0.120	accept	0.788	0.375
	EWMA	green	accept	0.836	0.361	accept	5.177	0.075	reject	4.341	0.037
	GARCHn	green	accept	2.485	0.115	reject	11.680	0.003	reject	9.194	0.002
	GARCHt	green	accept	1.220	0.269	reject	14.297	0.001	reject	13.077	0.000
	EGARCHn	green	accept	1.220	0.269	reject	12.121	0.002	reject	10.901	0.001
	EGARCHt	green	accept	1.220	0.269	reject	10.110	0.006	reject	8.890	0.003
	GJRn	green	accept	0.393	0.531	reject	13.180	0.001	reject	12.788	0.000
	GJRt	green	accept	0.393	0.531	reject	13.180	0.001	reject	12.788	0.000
	Normal-fit	green	reject	19.069	0.000	reject	32.906	0.000	reject	13.837	0.000
	t-fit	red	reject	23.187	0.000	reject	32.867	0.000	reject	9.681	0.002
XRP 99%	BHS	green	accept	0.728	0.394	accept	5.049	0.080	reject	4.321	0.038
	AWHS	red	reject	37.460	0.000	reject	43.311	0.000	reject	5.851	0.016
	EWMA	green	accept	0.195	0.659	accept	1.439	0.487	accept	1.244	0.265
	GARCHn	yellow	accept	2.705	0.100	accept	3.330	0.189	accept	0.624	0.429
	GARCHt	yellow	reject	4.110	0.043	accept	4.583	0.101	accept	0.474	0.491
	EGARCHn	green	accept	2.104	0.147	accept	2.813	0.245	accept	0.709	0.400
	EGARCHt	yellow	accept	2.705	0.100	accept	3.330	0.189	accept	0.624	0.429
	GJRn	green	accept	2.104	0.147	accept	2.813	0.245	accept	0.709	0.400
	GJRt	yellow	accept	2.705	0.100	accept	3.330	0.189	accept	0.624	0.429
	Normal-fit	green	accept	0.421	0.516	accept	1.542	0.463	accept	1.121	0.290
	t-fit	yellow	reject	4.110	0.043	accept	4.583	0.101	accept	0.474	0.491
	POT	yellow	accept	3.375	0.066	reject	6.483	0.039	accept	3.108	0.078

Table 5.8. *VaR backtesting results (continued)*

	METHOD	TL	POF			CC			CCI		
		Result	Result	LR ratio	P-value	Result	LR ratio	P-value	Result	LR ratio	P-value
ADA 95%	BHS	green	accept	0.005	0.944	accept	0.052	0.974	accept	0.047	0.828
	AWHS	yellow	reject	7.375	0.007	reject	8.780	0.012	accept	1.406	0.236
	EWMA	green	accept	0.217	0.641	accept	0.440	0.803	accept	0.223	0.637
	GARCHn	green	accept	1.956	0.162	accept	1.956	0.376	accept	0.000	0.994
	GARCHt	green	accept	0.694	0.405	accept	0.753	0.686	accept	0.060	0.807
	EGARCHn	green	accept	1.039	0.308	accept	1.055	0.590	accept	0.016	0.899
	EGARCHt	green	accept	0.694	0.405	accept	0.753	0.686	accept	0.060	0.807
	GJRn	green	accept	1.956	0.162	accept	1.956	0.376	accept	0.000	0.994
	GJRt	green	accept	0.694	0.405	accept	0.753	0.686	accept	0.060	0.807
	Normal-fit	green	accept	1.459	0.227	accept	1.466	0.480	accept	0.008	0.930
	t-fit	green	accept	0.011	0.915	accept	0.102	0.950	accept	0.091	0.763
ADA 99%	BHS	green	accept	0.012	0.914	accept	0.159	0.923	accept	0.148	0.701
	AWHS	red	reject	24.440	0.000	reject	26.073	0.000	accept	1.633	0.201
	EWMA	green	accept	0.488	0.485	accept	0.563	0.755	accept	0.075	0.784
	GARCHn	green	accept	0.081	0.776	accept	0.189	0.910	accept	0.108	0.742
	GARCHt	green	accept	0.488	0.485	accept	0.563	0.755	accept	0.075	0.784
	EGARCHn	green	accept	0.488	0.485	accept	0.563	0.755	accept	0.075	0.784
	EGARCHt	green	accept	0.488	0.485	accept	0.563	0.755	accept	0.075	0.784
	GJRn	green	accept	0.081	0.776	accept	0.189	0.910	accept	0.108	0.742
	GJRt	green	accept	0.488	0.485	accept	0.563	0.755	accept	0.075	0.784
	Normal-fit	green	accept	0.012	0.914	accept	0.159	0.923	accept	0.148	0.701
	t-fit	green	accept	1.301	0.254	accept	1.349	0.509	accept	0.048	0.827
	POT	green	reject	7.679	0.006	reject	7.682	0.021	accept	0.003	0.956

5.2.3. ES backtesting results

Table 5.9 presents the ES backtesting results for all cryptocurrencies, using the unconditional test by Acerbi and Szekely (2014). Overall, most estimation methods cannot be rejected at the 5% level, except for the AWHS method and the normal distribution fitting method.

At the VaR level of 99%, the AWHS and the normal distribution fitting method are rejected for most cryptocurrencies. Between these two methods, the AWHS has a worse performance with largely negative test statistics. The AWHS method is also rejected at 95% VaR level for BTC and ADA when the critical values are based on the normal distribution. At 95% confidence level, the normal distribution fitting method for BTC is rejected in both cases when we assume the critical values follow the normal distribution or the student-t distribution. Though the student t-distribution fitting method cannot be rejected in any case, its overall performance is not good due to the high value of test statistics.

The VWHS estimation method using the EWMA model and GARCH-type models seems to perform well in all cases, with the test statistics being approximately zero. The best performance among these models varies for each cryptocurrency. For BTC and XRP, the GARCH model that fits normal distribution to the innovation process outperforms the same type of GARCH model that fits student t-distribution (for instance, EGARCHn is better than EGARCHt). Meanwhile, ETH shows the opposite observation.

The best estimation methods for BTC at both VaR levels are GARCHn and GJR-GARCHn. For the risk estimation of XRP, GARCHn performs the best at the 95% VaR level and EGARCHn is the best model at the 99% VaR level. EWMA and EGARCHt prove to be the best estimation models for ETH at 95% and 99% respectively.

Despite its simplicity in implementation, the BHS method shows an outstanding performance for BNB and ADA, with the test statistics very close to zero at both VaR levels. This method also performs well for BTC at the 95% confidence level. This result is really impressive as the BHS method does not take volatility clustering into the risk estimates. However, as we use the rolling window, this method seems to work well. Currently, not many existing research papers examine the application of this method to VaR and ES estimates for cryptocurrencies.

Table 5.9. *ES backtesting results*

		95% CONFIDENCE LEVEL						99% CONFIDENCE LEVEL					
		Unconditional Normal			Unconditional t			Unconditional Normal			Unconditional t		
		Result	Test statistic	Crit value	Result	Test statistic	Crit value	Result	Test statistic	Crit value	Result	Test statistic	Crit value
BTC	BHS	accept	-0.0144	-0.1556	accept	-0.0144	-0.1849	accept	0.1467	-0.3586	accept	0.1467	-0.4183
	AWHS	reject	-0.1707	-0.1557	accept	-0.1707	-0.1850	reject	-1.4740	-0.3613	reject	-1.4740	-0.4214
	EWMA	accept	0.1369	-0.1556	accept	0.1369	-0.1849	accept	0.2721	-0.3586	accept	0.2721	-0.4183
	GARCHn	accept	0.0250	-0.1556	accept	0.0250	-0.1849	accept	0.0632	-0.3586	accept	0.0632	-0.4183
	GARCHt	accept	0.0511	-0.1556	accept	0.0511	-0.1849	accept	0.0838	-0.3586	accept	0.0838	-0.4183
	EGARCHn	accept	0.0648	-0.1556	accept	0.0648	-0.1849	accept	0.1044	-0.3586	accept	0.1044	-0.4183
	EGARCHt	accept	0.1165	-0.1556	accept	0.1165	-0.1849	accept	0.1805	-0.3586	accept	0.1805	-0.4183
	GJRn	accept	0.0251	-0.1556	accept	0.0251	-0.1849	accept	0.0601	-0.3586	accept	0.0601	-0.4183
	GJRt	accept	0.0519	-0.1556	accept	0.0519	-0.1849	accept	0.1103	-0.3586	accept	0.1103	-0.4183
	normal	reject	-0.2175	-0.1556	reject	-0.2175	-0.1849	reject	-1.7413	-0.3586	reject	-1.7413	-0.4183
	t	accept	0.6937	-0.1556	accept	0.6937	-0.1849	accept	0.8263	-0.3586	accept	0.8263	-0.4183
	POT							accept	0.1415	-0.3586	accept	0.1415	-0.4183
ETH	BHS	accept	0.2267	-0.1950	accept	0.2267	-0.2327	accept	0.3897	-0.4512	accept	0.3897	-0.5267
	AWHS	accept	-0.1230	-0.1952	accept	-0.1230	-0.2330	reject	-1.8090	-0.4565	reject	-1.8090	-0.5324
	EWMA	accept	0.0731	-0.1950	accept	0.0731	-0.2327	accept	-0.0702	-0.4512	accept	-0.0702	-0.5267
	GARCHn	accept	0.1187	-0.1950	accept	0.1187	-0.2327	accept	0.0545	-0.4512	accept	0.0545	-0.5267
	GARCHt	accept	0.0955	-0.1950	accept	0.0955	-0.2327	accept	0.0492	-0.4512	accept	0.0492	-0.5267
	EGARCHn	accept	0.1078	-0.1950	accept	0.1078	-0.2327	accept	0.0603	-0.4512	accept	0.0603	-0.5267
	EGARCHt	accept	0.0934	-0.1950	accept	0.0934	-0.2327	accept	0.0213	-0.4512	accept	0.0213	-0.5267
	GJRn	accept	0.1101	-0.1950	accept	0.1101	-0.2327	accept	0.0553	-0.4512	accept	0.0553	-0.5267
	GJRt	accept	0.1021	-0.1950	accept	0.1021	-0.2327	accept	0.0512	-0.4512	accept	0.0512	-0.5267
	normal	accept	0.1618	-0.1950	accept	0.1618	-0.2327	reject	-0.9723	-0.4512	reject	-0.9723	-0.5267
	t	accept	0.6457	-0.1950	accept	0.6457	-0.2327	accept	0.7708	-0.4512	accept	0.7708	-0.5267
	POT							accept	0.7273	-0.4512	accept	0.7273	-0.5267

Table 5.9. *ES backtesting results (continued)*

		95% CONFIDENCE LEVEL						99% CONFIDENCE LEVEL					
		Unconditional Normal			Unconditional t			Unconditional Normal			Unconditional t		
		Result	Test statistic	Crit value	Result	Test statistic	Crit value	Result	Test statistic	Crit value	Result	Test statistic	Crit value
BNB	BHS	accept	0.1696	-0.2759	accept	0.1696	-0.3288	accept	0.0475	-0.6445	accept	0.0475	-0.7477
	AWHS	accept	-0.0522	-0.2759	accept	-0.0522	-0.3288	reject	-1.3537	-0.6445	reject	-1.3537	-0.7477
	EWMA	accept	0.2374	-0.2759	accept	0.2374	-0.3288	accept	0.3577	-0.6445	accept	0.3577	-0.7477
	GARCHn	accept	0.2171	-0.2759	accept	0.2171	-0.3288	accept	0.1760	-0.6445	accept	0.1760	-0.7477
	GARCHt	accept	0.2024	-0.2759	accept	0.2024	-0.3288	accept	0.1802	-0.6445	accept	0.1802	-0.7477
	EGARCHn	accept	0.2581	-0.2759	accept	0.2581	-0.3288	accept	0.2896	-0.6445	accept	0.2896	-0.7477
	EGARCHt	accept	0.2544	-0.2759	accept	0.2544	-0.3288	accept	0.3040	-0.6445	accept	0.3040	-0.7477
	GJRn	accept	0.2156	-0.2759	accept	0.2156	-0.3288	accept	0.1793	-0.6445	accept	0.1793	-0.7477
	GJRt	accept	0.2053	-0.2759	accept	0.2053	-0.3288	accept	0.1936	-0.6445	accept	0.1936	-0.7477
	normal	accept	0.2123	-0.2759	accept	0.2123	-0.3288	reject	-0.8256	-0.6445	reject	-0.8256	-0.7477
	t	accept	0.5150	-0.2759	accept	0.5150	-0.3288	accept	0.5656	-0.6445	accept	0.5656	-0.7477
	POT							accept	0.6038	-0.6445	accept	0.6038	-0.7477
XRP	BHS	accept	-0.1542	-0.1591	accept	-0.1542	-0.1890	accept	-0.3089	-0.3668	accept	-0.3089	-0.4279
	AWHS	accept	-0.1457	-0.1594	accept	-0.1457	-0.1894	reject	-1.6060	-0.3740	reject	-1.6060	-0.4364
	EWMA	accept	0.1573	-0.1591	accept	0.1573	-0.1890	accept	0.0378	-0.3668	accept	0.0378	-0.4279
	GARCHn	accept	0.0271	-0.1591	accept	0.0271	-0.1890	accept	-0.1926	-0.3668	accept	-0.1926	-0.4279
	GARCHt	accept	0.0647	-0.1591	accept	0.0647	-0.1890	accept	-0.2371	-0.3668	accept	-0.2371	-0.4279
	EGARCHn	accept	0.1643	-0.1591	accept	0.1643	-0.1890	accept	0.0006	-0.3668	accept	0.0006	-0.4279
	EGARCHt	accept	0.2489	-0.1591	accept	0.2489	-0.1890	accept	0.1362	-0.3668	accept	0.1362	-0.4279
	GJRn	accept	0.1016	-0.1591	accept	0.1016	-0.1890	accept	-0.1314	-0.3668	accept	-0.1314	-0.4279
	GJRt	accept	0.1106	-0.1591	accept	0.1106	-0.1890	accept	-0.1508	-0.3668	accept	-0.1508	-0.4279
	normal	accept	0.1767	-0.1591	accept	0.1767	-0.1890	reject	-0.7885	-0.3668	reject	-0.7885	-0.4279
	t	accept	0.7090	-0.1591	accept	0.7090	-0.1890	accept	0.7397	-0.3668	accept	0.7397	-0.4279
	POT							reject	-0.3774	-0.3668	accept	-0.3774	-0.4279

Table 5.9. *ES backtesting results (continued)*

		95% CONFIDENCE LEVEL						99% CONFIDENCE LEVEL					
		Unconditional Normal			Unconditional t			Unconditional Normal			Unconditional t		
		Result	Test statistic	Crit value	Result	Test statistic	Crit value	Result	Test statistic	Crit value	Result	Test statistic	Crit value
ADA	BHS	accept	0.0587	-0.2934	accept	0.0587	-0.3491	accept	0.0303	-0.6862	accept	0.0303	-0.7949
	AWHS	reject	-0.2966	-0.2934	accept	-0.2966	-0.3491	reject	-2.0291	-0.6862	reject	-2.0291	-0.7949
	EWMA	accept	0.2484	-0.2934	accept	0.2484	-0.3491	accept	0.4041	-0.6862	accept	0.4041	-0.7949
	GARCHn	accept	0.3284	-0.2934	accept	0.3284	-0.3491	accept	0.3026	-0.6862	accept	0.3026	-0.7949
	GARCHt	accept	0.2580	-0.2934	accept	0.2580	-0.3491	accept	0.3950	-0.6862	accept	0.3950	-0.7949
	EGARCHn	accept	0.3077	-0.2934	accept	0.3077	-0.3491	accept	0.3969	-0.6862	accept	0.3969	-0.7949
	EGARCHt	accept	0.2677	-0.2934	accept	0.2677	-0.3491	accept	0.3846	-0.6862	accept	0.3846	-0.7949
	GJRn	accept	0.3303	-0.2934	accept	0.3303	-0.3491	accept	0.3030	-0.6862	accept	0.3030	-0.7949
	GJRt	accept	0.2548	-0.2934	accept	0.2548	-0.3491	accept	0.3923	-0.6862	accept	0.3923	-0.7949
	normal	accept	0.1286	-0.2934	accept	0.1286	-0.3491	accept	-0.2529	-0.6862	accept	-0.2529	-0.7949
	t	accept	0.4603	-0.2934	accept	0.4603	-0.3491	accept	0.6801	-0.6862	accept	0.6801	-0.7949
	POT							accept	0.8614	-0.6862	accept	0.8614	-0.7949

6. Conclusion

This thesis contributes to the existing research papers by thoroughly examining a more informative, representative, and updated data set of five cryptocurrencies (BTC, ETH, BNB, XRP, ADA) obtained from the first available day to 30 April 2022 on coinmarketcap.com. In addition, by implementing three parametric and three non-parametric VaR and ES estimation methods, fitting both normal distribution and student t-distribution to the innovation process and loss process, we examine a wide range of possibilities (11 models at 95% VaR level and 12 models at 99% VaR level) to obtain a more diversified and broader set of data for comparison. Totally, we carry out 575 tests and use these test results to compare the performance of the given models in estimating VaR and ES. Our major findings are summarized as follows.

First, the volatility-weighted historical simulation (VWHS) method shows a good performance in measuring VaR and ES. Among them, the exponential weighted moving average (EWMA) model consistently proves its outstanding capability in adequately estimating the volatility, and consequently the value at risk. Despite the use of fixed parameters, this model still passes almost all VaR backtesting procedures at both 95% and 99% VaR level. VWHS-EWMA is among the best VaR estimation models for all cryptocurrencies. Meanwhile, all GARCH-type models do not perform well in estimating VaR of BTC, ETH, and XRP at 95% VaR level. However, they are still among the best models for ES estimation.

Second, the basic historical simulation (BHS) method shows an impressive performance in both VaR and ES estimation. Despite the simple calculation, it still proves to be the best model for the ES estimate of ADA, and surprisingly the best model for the VaR estimates of BNB and ETH. It also performs relatively well in the VaR and ES estimation of other cryptocurrencies. Given these positive results, it can be implemented practically thanks to the simple calculation and good performance.

Third, the peak over threshold (POT) method that fits the Generalized Pareto distribution (GPD) to a subsample of the largest loss displays it good performance at the 99% level of confidence. It generates the lowest number of violations in most cases, and the first failure occurs relatively late. This method is the best model to estimate VaR of BNB, however, it does not perform well at ES estimates of most cryptocurrencies.

Finally, the age-weighted historical simulation (AWHS) seems to perform the worst in almost all cases, though it is expected to perform better due to the incorporation of time variation in the VaR and ES estimates. Therefore, this model should not be applied to measure VaR and ES of these five cryptocurrencies.

These conclusions satisfy our initial research question of finding the model with the best performance. The findings from this thesis are of crucial importance to the risk managers or policymakers to choose a convenient and appropriate method to estimate VaR and ES for cryptocurrencies, and subsequently establish the proper level of capital requirement. It is also valuable for investors to build a more adequate trading and investment strategy.

REFERENCES

Acerbi, C. & Szekely, B. (2014). Backtesting Expected Shortfall, [e-journal], Available Online: https://www.msci.com/documents/10199/22aa9922-f874-4060-b77a-0f0e267a489b.

Ardia, D., Bluteau, K. & Rüede, M. (2019). Regime Changes in Bitcoin GARCH Volatility Dynamics, *Finance Research Letters*, vol. 29, pp.266–271.

Ariefianto, Moch. D. (2020). Assessing Qualification of Crypto Currency as A Financial Assets: A Case Study on Bitcoin, *2020 International Conference on Information Management and Technology (ICIMTech), Information Management and Technology (ICIMTech), 2020 International Conference on*, pp.934–939.

Asgharian, H. (2021). Empirical Finance Lecture Notes Fall 2021.

Assaf, A., Alberiko Gil-Alana, L. & Mokni, K. (2021). True or Spurious Long Memory in the Cryptocurrency Markets: Evidence from a Multivariate Test and Other Whittle Estimation Methods, *EMPIRICAL ECONOMICS*.

Baldan, C. & Zen, F. (2020). The Bitcoin as a Virtual Commodity: Empirical Evidence and Implications, *Frontiers in Artificial Intelligence*, vol. 3, no. 3, p.21.

Balkema, A. A. & de Haan, L. (1974). Residual Life Time at Great Age, *The Annals of Probability*, vol. 2, no. 5, pp.792–804.

Basel Committee on Banking Supervision. (1996). Supervisory Framework for the Use of ‘backtesting’ in Conjunction with the Internal Models Approach to Market Risk Capital Requirements, [e-journal], Available Online: https://www.bis.org/publ/bcbs22.htm [Accessed 25 May 2022].

Baur, D. G., Hong, K. & Lee, A. D. (2018). Bitcoin: Medium of Exchange or Speculative Assets, *Journal of International Financial Markets, Institutions & Money*, vol. 54, pp.177–189.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, [e-journal] vol. 31, no. 3, pp.307–327, Available Online: https://www.sciencedirect.com/science/article/pii/0304407686900631 [Accessed 20 May 2022].

Bouoiyour, J. & Selmi, R. (2015). Bitcoin Price: Is It Really That New Round of Volatility Can Be on Way?, 65580, *MPRA Paper*, University Library of Munich, Germany, Available Online: https://ideas.repec.org/p/pra/mprapa/65580.html [Accessed 17 April 2022].

Bouoiyour, J. & Selmi, R. (2016). Bitcoin: A Beginning of a New Phase?, *Economics Bulletin*, vol. 36, no. 3, pp.1430–1440.

Bouri, E., Hussain Shahzad, S. J. & Roubaud, D. (2020). Cryptocurrencies as Hedges and Safe-Havens for US Equity Sectors, *Quarterly Review of Economics and Finance*, vol. 75, pp.294–307.

Buterin, V. (2014). Ethereum Whitepaper, *Ethereum.Org*, Available Online: https://ethereum.org/en/whitepaper/ [Accessed 19 May 2022].

Caporale, G. M. & Zekokh, T. (2019). Modelling Volatility of Cryptocurrencies Using Markov-Switching GARCH Models, *Research in International Business and Finance*, [e-journal] vol. 48, pp.143–155, Available Online: https://www.sciencedirect.com/science/article/pii/S027553191830669X [Accessed 15 April 2022].

Cerqueti, R., Giacalone, M. & Mattera, R. (2020). Skewed Non-Gaussian GARCH Models for Cryptocurrencies Volatility Modelling, *Information Sciences*, vol. 527, pp.1–26.

Chan, S., Chu, J., Nadarajah, S. & Osterrieder, J. (2017). A Statistical Analysis of Cryptocurrencies, *Journal of Risk and Financial Management*, vol. 10, no. 2, pp.12–12.

Chen Y. Wu & Pandey, V. K. (2014). The Value of Bitcoin in Enhancing the Efficiency of an Investor’s Portfolio, *Journal of Financial Planning*, [e-journal] vol. 27, no. 9, pp.44–52, Available Online: http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=bth&AN=98013371&site=eds-live&scope=site [Accessed 16 April 2022].

Choi, S. & Shin, J. (2022). Bitcoin: An Inflation Hedge but Not a Safe Haven, *Finance Research Letters*, vol. 46, no. Part B.

Christoffersen, P. F. (1998). Evaluating Interval Forecasts, *International Economic Review*, [e-journal] vol. 39, no. 4, pp.841–862, Available Online: http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=edsjsr&AN=edsjsr.10.2307.2527341&site=eds-live&scope=site [Accessed 24 May 2022].

Chu, J., Chan, S., Nadarajah, S. & Osterrieder, J. (2017). GARCH Modelling of Cryptocurrencies, *Journal of Risk and Financial Management*, vol. 10, no. 4, pp.1–15.

CoinMarketCap. (2019). How Are Prices Calculated on CoinMarketCap?, *CoinMarketCap*, Available Online: https://support.coinmarketcap.com/hc/en-us/articles/360015968632-How-are-prices-calculated-on-CoinMarketCap- [Accessed 25 May 2022].

Corbet, S. (1), Meegan, A. (1), Larkin, C. (2), Lucey, B. (2) & Yarovaya, L. (3). (2018). Exploring the Dynamic Relationships between Cryptocurrencies and Other Financial Assets, *Economics Letters*, vol. 165, pp.28–34.

Dyhrberg, A. H. (2016a). Hedging Capabilities of Bitcoin. Is It the Virtual Gold?, *Finance Research Letters*, vol. 16, pp.139–144.

Dyhrberg, A. H. (2016b). Bitcoin, Gold and the Dollar – A GARCH Volatility Analysis, *Finance Research Letters*, vol. 16, pp.85–92.

Fang, L., Bouri, E., Gupta, R. & Roubaud, D. (2019). Does Global Economic Uncertainty Matter for the Volatility and Hedging Effectiveness of Bitcoin?, *International Review of Financial Analysis*, vol. 61, pp.29–36.

Fung, K., Jeong, J. & Pereira, J. (2021). More to Cryptos than Bitcoin: A GARCH Modelling of Heterogeneous Cryptocurrencies, *Finance Research Letters*.

Gkillas, K. (1) & Katsiampa, P. (2). (2018). An Application of Extreme Value Theory to Cryptocurrencies, *Economics Letters*, vol. 164, pp.109–111.

Gnedenko, B. (1943). Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire, *Annals of Mathematics*, [e-journal] vol. 44, no. 3, pp.423–453, Available Online: http://www.jstor.org/stable/1968974 [Accessed 23 May 2022].

Görgen, K., Meirer, J. & Schienle, M. (2022). Predicting Value at Risk for Cryptocurrencies Using Generalized Random Forests.

Gronwald, M. (2014). The Economics of Bitcoins - Market Characteristics and Price Jumps, 5121, *CESifo Working Paper Series*, CESifo, Available Online: https://ideas.repec.org/p/ces/ceswps/_5121.html [Accessed 17 April 2022].

Guesmi, K., Saadi, S., Abid, I. & Ftiti, Z. (2019). Portfolio Diversification with Virtual Currency: Evidence from Bitcoin, *International Review of Financial Analysis*, vol. 63, pp.431–437.

Gyamerah, S. A. (2019). Modelling the Volatility of Bitcoin Returns Using GARCH Models, *Quantitative Finance and Economics*, vol. 3, no. 4, pp.739–753.

Hasan, Md. B., Hassan, M. K., Karim, Z. A. & Rashid, Md. M. (2022). Exploring the Hedge and Safe Haven Properties of Cryptocurrency in Policy Uncertainty, *Finance Research Letters*, vol. 46, no. Part A.

Hull, J. (2018). Risk Management and Financial Institutions, Fifth edition., Wiley.

Janson, N. & Karoubi, B. (2021). The Bitcoin: To Be or Not to Be a Real Currency?, *Quarterly Review of Economics and Finance,* vol. 82, pp.312–319.

Jiménez, I., Perote, J. & Mora-Valencia, A. (2022). Semi-Nonparametric Risk Assessment with Cryptocurrencies, *Research in International Business and Finance*, [e-journal] vol. 59, p.101567, Available Online: https://www.sciencedirect.com/science/article/pii/S0275531921001884 [Accessed 15 April 2022].

J.P. Morgan & Reuters. (1996). RiskMetrics Technical Document - Fourth Edition, [e-journal] p.296, Available Online: https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a.

Katsiampa, P. (2017). Volatility Estimation for Bitcoin: A Comparison of GARCH Models, *Economics Letters*, vol. 158, pp.3–6.

Kupiec, P. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models, SSRN Scholarly Paper, 6697, Rochester, NY: Social Science Research Network, Available Online: https://papers.ssrn.com/abstract=6697 [Accessed 24 May 2022].

Lahmiri, S., Bekiros, S. & Salvi, A. (2018). Long-Range Memory, Distributional Variation and Randomness of Bitcoin Volatility, *Chaos, Solitons and Fractals: the interdisciplinary journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena*, vol. 107, pp.43–48.

Li, Z. (2022). Honour Thesis: A Joint Value at Risk and Expected Shortfall Combination Framework and Its Applications in the Cryptocurrency Market.

Likitratcharoen, D., Kronprasert, N., Wiwattanalamphong, K. & Pinmanee, C. (2021). The Accuracy of Risk Measurement Models on Bitcoin Market during COVID-19 Pandemic, *Risks*, vol. 9, no. 12, pp.222–222.

Likitratcharoen, D., Ranong, T. N., Chuengsuksomboon, R., Sritanee, N. & Pansriwong, A. (2018). Value at Risk Performance in Cryptocurrencies, 1, *The Journal of Risk Management and Insurance*, [e-journal] vol. 22, no. 1, pp.11–28, Available Online: https://jrmi.au.edu/index.php/jrmi/article/view/171 [Accessed 18 April 2022].

Liu, W., Semeyutin, A., Lau, C. K. M. & Gozgor, G. (2020). Forecasting Value-at-Risk of Cryptocurrencies with RiskMetrics Type Models, *Research in International Business and Finance*, vol. 54.

Maciel, L. (2021). Cryptocurrencies Value-at-Risk and Expected Shortfall: Do Regime-Switching Volatility Models Improve Forecasting?, *International Journal of Finance & Economics*, [e-journal] vol. 26, no. 3, pp.4840–4855, Available Online: http://onlinelibrary.wiley.com/doi/abs/10.1002/ijfe.2043 [Accessed 15 April 2022].

Mroua, M., Bahloul, S. & Naifar, N. (2022). Should Investors Include Bitcoin in Their Portfolio? New Evidence from a Bootstrap-Based Stochastic Dominance Approach, *Applied Economics Letters*, vol. 29, no. 1, pp.53–62.

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System, [e-journal], Available Online: https://bitcoin.org/bitcoin.pdf.

Nilsson, B. (2021). Parametric Estimation Methods.

Obeng, C. (2021). Measuring Value at Risk Using Garch Model - Evidence from the Cryptocurrency Market, *International Journal of Entrepreneurial Knowledge*, [e-journal] vol. 9, no. 2, pp.63–84, Available Online: http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=bth&AN=154289193&site=eds-live&scope=site [Accessed 13 April 2022].

Pele, D. t. & Mazurencu-Marinescu-Pele, M. (2019). Using High-Frequency Entropy to Forecast Bitcoin's Daily Value at Risk, *Entropy*, [e-journal] vol. 21, no. 2, Available Online: http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=edselc&AN=edselc.2-52.0-85061994122&site=eds-live&scope=site [Accessed 18 April 2022].

Phillip, A., Chan, J. & Peiris, S. (2019). On Long Memory Effects in the Volatility Measure of Cryptocurrencies, *Finance Research Letters*, vol. 28, pp.95–100.

Pickands, J. (1975). Statistical Inference Using Extreme Order Statistics, *The Annals of Statistics*, vol. 3, no. 1, pp.119–131.

Platanakis, E. & Urquhart, A. (2020). Should Investors Include Bitcoin in Their Portfolios? A Portfolio Theory Approach, *The British Accounting Review*, vol. 52, no. 4.

Ratajczak-Mrozek, M. & Marszałek, P. (eds). (2022). Digitalization and Firm Performance: Examining the Strategic Impact, *Springer Nature eBook*.

Stavroyiannis, S. (2017). Value-at-Risk and Expected Shortfall for the Major Digital Currencies, *arXiv:1708.09343 [q-fin]*, [e-journal], Available Online: http://arxiv.org/abs/1708.09343 [Accessed 15 April 2022].

Stavroyiannis, S. (2018). Value-at-Risk and Related Measures for the Bitcoin, *The Journal of Risk Finance*, vol. 19, no. 2, pp.127–136.

Troster, V., Tiwari, A. K., Shahbaz, M. & Macedo, D. N. (2019). Bitcoin Returns and Risk: A General GARCH and GAS Analysis, *Finance Research Letters*, [e-journal] vol. 30, pp.187–193, Available Online: https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S1544612318304859 [Accessed 19 April 2022].

Trucíos, C. (2019). Forecasting Bitcoin Risk Measures: A Robust Approach, *International Journal of Forecasting*, vol. 35, no. 3, pp.836–847.

Trucíos, C., Tiwari, A. K. & Alqahtani, F. (2020). Value-at-Risk and Expected Shortfall in Cryptocurrencies' Portfolio: A Vine Copula–Based Approach, *Applied Economics*, [e-journal] vol. 52, no. 24, pp.2580–2593, Available Online: http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=bth&AN=143196357&site=eds-live&scope=site [Accessed 14 April 2022].

Ulmer, J. M. & Chen, L. (2021). Volatility and Risk – FIGARCH Modelling of Cryptocurrencies, [e-journal], Available Online: http://lup.lub.lu.se/student-papers/record/9050072 [Accessed 17 April 2022].

Ustaoglu, E. (2022). Safe-Haven Properties and Portfolio Applications of Cryptocurrencies: Evidence from the Emerging Markets, *Finance Research Letters*.

Van der Auwera, E., Schoutens, W., Petracco Giudici, M. & Alessi, L. (2020). Financial Risk Management for Cryptocurrencies, [e-book] Cham: Springer International Publishing, Available Online: http://link.springer.com/10.1007/978-3-030-51093-0 [Accessed 14 April 2022].

Wang, G.-J., Ma, X. & Wu, H. (2020). Are Stablecoins Truly Diversifiers, Hedges, or Safe Havens against Traditional Cryptocurrencies as Their Name Suggests?, *Research in International Business and Finance*, vol. 54.

Wang, P., Zhang, W., Li, X. & Shen, D. (2019). Is Cryptocurrency a Hedge or a Safe Haven for International Indices? A Comprehensive and Dynamic Perspective, *Finance Research Letters*, vol. 31, pp.1–18.

Yermack, D. (1, 2). (2015). Is Bitcoin a Real Currency? An Economic Appraisal, *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data*, Elsevier Inc.

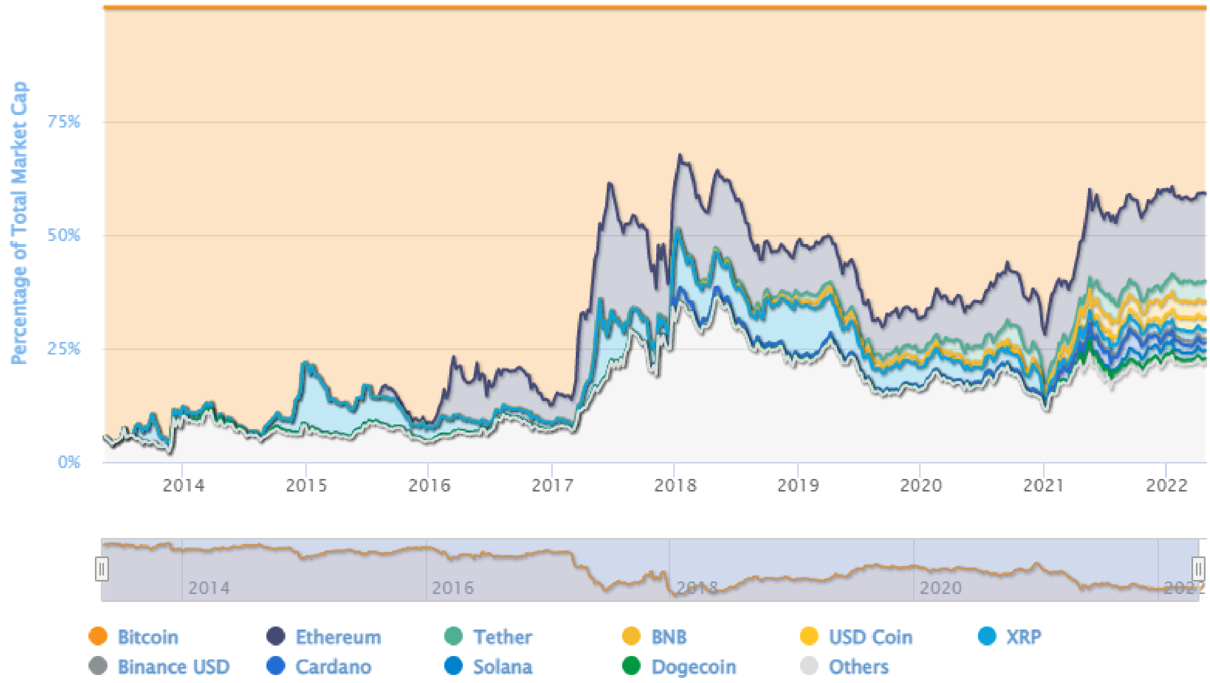
APPENDIX

Appendix 1: Characteristics of selected cryptocurrencies

Cryptocurrencies		Main characteristics
Bitcoin	BTC	- A decentralized cryptocurrency, being the very first cryptocurrency on the market - First described in a whitepaper by Satoshi Nakamoto in 2008, then official launched in January 2009 - Total supply in circulation is limited within 21,000,000 coins; new coins are added through a mining process based on "proof of work" consensus, which is operated using a large number of computers - An application of blockchain technology
Ethereum	ETH	- A "proof of work" decentralized blockchain platform which allows the execution of smart contracts and applications - First introduced by Vitalik Buterin in 2013 - Smart contracts are internet-based computer programs that automatically carry out the steps required to complete a contract between many parties, which was designed to reduce the intermediary organizations between parties
Binance coin	BNB	- Binance is the world's largest cryptocurrency exchange, which created a whole ecosystem of services for its users. - First launched by Changpeng Zhao in 2017 - An application of blockchain technology
Ripple coin	XRP	- A open-source ledger using decentralized technology, offering many payment-related applications and use cases, including micropayments, DeFi, and NFTs - Founded by David, Jed, and Arthur in 2012
Cardano	ADA	- A "proof-of-stake" blockchain platform, aiming to consume less energy than the "proof-of-work" - Founded by Charles Hoskinson in 2017

(Source: coinmarketcap.com)

Appendix 2: Cryptocurrency dominance chart from 5 May 2013 to 30 April 2022



(Source: coinmarketcap.com)

Appendix 3: Jacque-Bera test to validate the normality assumption

Normal distribution (also named Gaussian distribution) is a symmetric probability distribution around the mean. The normality assumption can be tested using the third moment (skewness) and the fourth moment (kurtosis) of a distribution, or a joint test of both moments known as the Jacque-Bera test.

Skewness measures the shape of a distribution. In other words, it illustrates how asymmetric the distribution is around the sample mean. Kurtosis measures the fatness of the distribution tails. Skewness and kurtosis are defined as below, in which μ is the mean and σ is the standard deviation.

$$skew = \frac{E(x-\mu)^3}{\sigma^3};\ kurt = \frac{E(x-\mu)^4}{\sigma^4}$$

The Jacque-Bera test statistic is calculated as follows, where n is the sample size.

$$JB = \frac{n}{6}(skew^2 + \frac{(kurt-3)^2}{4}$$

Due to the symmetry of normal distribution, the skewness is expected to be zero, and the kurtosis is expected to be three. Accordingly, Jacque-Bera test statistic is expected to be zero.

Appendix 4: Augmented Dickey Fuller (ADF) test for stationarity

A process is (weakly) stationary if it has finite mean and variance, the mean and the autocovariance is constant. The purpose of Augmented Dickey Fuller (ADF) test is to examine whether there is a unit root in the time series y_t. In case of a unit root, the time series is not stationary and the estimated parameters may be biased. We consider the following time series y_t:

$$y_t = \mu + \delta t + \emptyset y_{t-1} + \beta_1 \Delta y_{t-1} + \cdots + \beta_p \Delta y_{t-p} + \varepsilon_t$$

In which, $\Delta y_t = y_t - y_{t-1}$, μ is the mean or the drift term, δ is the trend coefficient, and ε_t is the innovation or error term. The time series has a unit root when the parameter $\emptyset = 1$, which is also the null hypothesis of the ADF test. The test statistic is then computed as follows, where SE is the standard error.

$$t = \frac{\widehat{\emptyset} - 1}{SE(\widehat{\emptyset})}$$

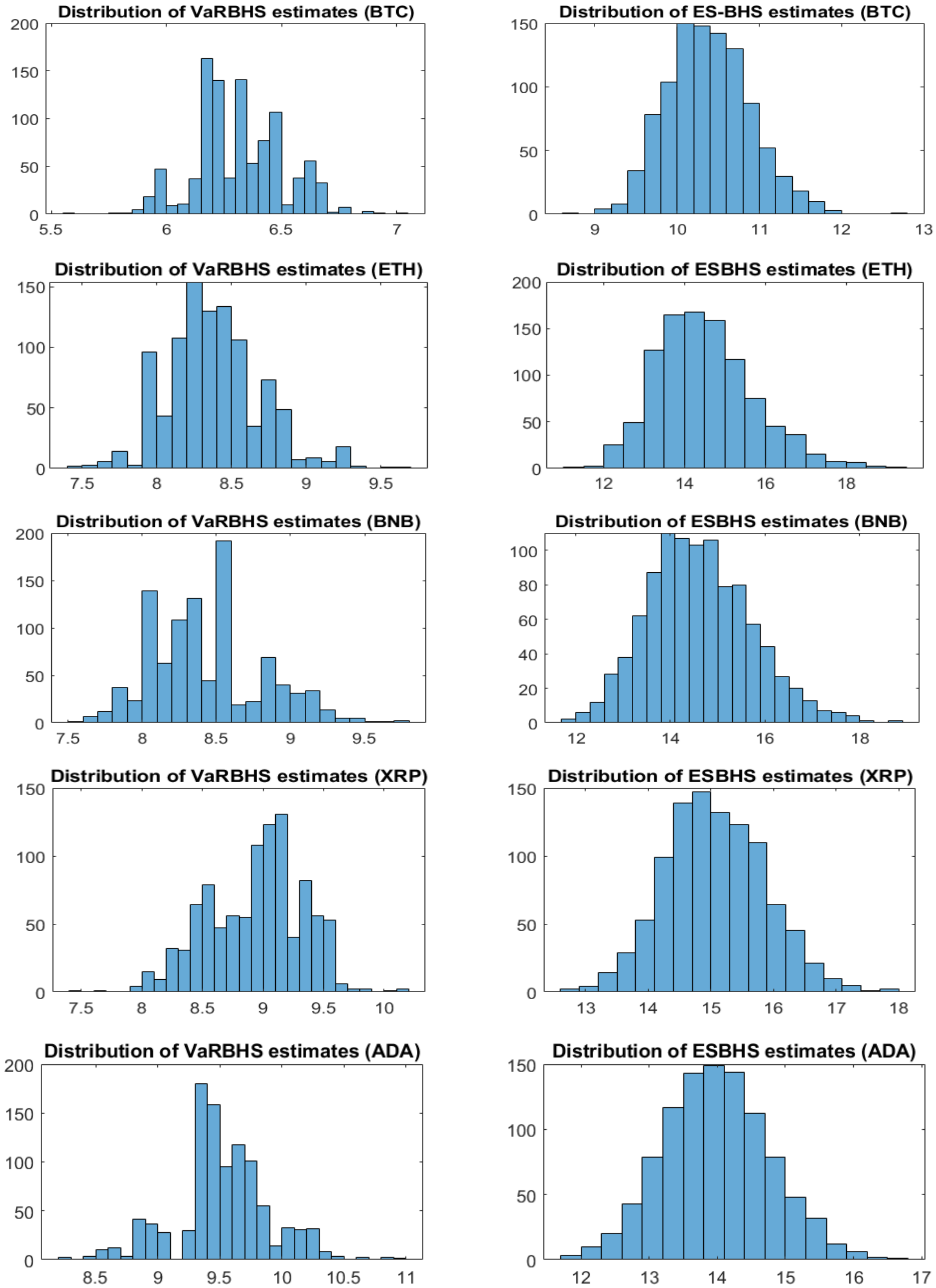
Under the null hypothesis, the test statistic is expected to be zero. The alternative hypothesis is $\emptyset < 1$, therefore, the test statistic is negative.

Appendix 5: VaR-BHS and ES-BHS confidence interval (CI) using bootstrapping

		VaR-BHS		ES-BHS	
	VaR level	Lower CI	Upper CI	Lower CI	Upper CI
BTC	95%	5.9385	6.6797	9.4791	11.4140
	99%	11.1035	14.7977	15.5512	21.4007
ETH	95%	7.8280	9.1762	12.4415	17.1328
	99%	14.6581	17.8901	20.0334	38.6334
BNB	95%	7.8253	9.2097	12.6871	16.9512
	99%	14.0667	22.7106	21.6942	34.6823
XRP	95%	8.1028	9.5368	13.5830	16.7299
	99%	15.6842	20.7614	23.2348	33.1942
ADA	95%	8.6899	10.2522	12.5280	15.5564
	99%	14.1668	18.7754	18.1008	27.9774

Appendix 6: VaR-BHS and ES-BHS distribution using bootstrapping method

a) At 95% VaR confidence level



Appendix 6 *(continued)*

b) At 99% VaR confidence level