University of Sussex

Empirical Research Project

978N1

Risk Metrics of Major Cryptocurrencies

Author:
Anthony Ede
212916

Supervisor:
Professor Carol Alexander

Word Count: 8246

UNIVERSITY OF SUSSEX

Abstract

With the growing interest in the world of cryptocurrencies from investors and researchers, there has been a need to investigate the risk measures to open the door to institutional investing in the market. Cryptocurrencies are currently unregulated by financial authorities and much of the investing in the markets is undertaken in a speculative manner. This unregulated and speculative manner of investing causes an extremely volatile nature to cyrptocurrency prices and even one of the least volatile cryptocurrencies, Bitcoin, has been known to be more volatile than the most volatile regulated financial assets available on standard markets. Previous literature has investigated into Bitcoin volatility dynamics and found that GARCH models perform the best at forecasting the volatility for both in-sample and out-of-sample data. We look to fill a gap in the literature through investigating which model performs the best over in-sample hourly data points using using a parity with bitcoin as the numeraire for our cyrptocurrency data, instead of the much investigated USD parity. We backtest both VaR estimates at the 95% and 99% confidence levels and ES at the 97.5% confidence level across four of the most liquid cryptocurrencies. Our findings are in line with current literature that GARCH models perform the strongest for VaR estimates, but we find that E-GARCH model outperforms the vanilla GARCH and GJR-GARCH model at the 99% confidence level for VaR and 97.5% confidence level for ES.

Acknowledgements

I would like to thank Professor Carol Alexander for presenting me with the opportunity to research this topic and for her support over the term, and Michael Dakos for helping with the code for the rolling window analysis.

Contents

1 Introduction

In recent years, there has been a growing interest in the world of cryptocurrencies from investors and researches alike, leading to a large growth in development in the market. Bitcoin was the first decentralized cyrptocurrency to be introduced using blockchain technology by Nakamoto 2008 . The purpose of the development of bitcoin was to make payments from peer to peer without using a centralized third party, hence the decentralization aspect of bitcoin.

Since the introduction of bitcoin, there has been much development in different forms of cryptocurrency, with some working on their own decentralized system such as Ethereum (Buterin 2013), Ripple (Schwartz, Youngs, and Britto 2014) and EOS (Larimer 2017), others as extensions of the original bitcoin like Litecoin. The constant development in the market has lead to a plethora of options in the cryptocurrency market. According to CoinMarketCap[1] there are over 6000 cryptocurrencies with a total market capitalisation of over $371bn and bitcoin making up just over 59% of the total capitalisation as of July 2020. The other most liquid coins like Ethereum, Litecoin, Ripple and EOS making up much of the rest. In fact, the top 23 cryptocurrencies listed on CoinMarketCap have a market capitalisation of over $1bn.

High returns and the fact that cryptocurrencies can be used for different purposes have attracted many new investors, with individual investors, banks and funds all looking at ways to utilise cryptocurrencies. They can be used as an alternative for foreign exchange as an instrument for international purchases and transactions of goods due to their low transaction costs and high liquidity. Cryptocurrencies can also be used an alternative asset for hedging purposes, as the they're extremely liquid and easily diversifiable. This property arises from the fact that cryptocurrencies, like bitcoin, aren't affected by market movements and have a low correlation with market traded assets and instruments. This would suggest that cryptocurrencies could be used to hedge market risk.

Two of the biggest exchanges, the Chicago Mercantile Exchange and Chicago Board Options Exchange have investigated futures trading on bit-

[1]https://coinmarketcap.com/all/views/all/

coin, just emphasising the huge interest in investment in the area of cryptocurrency. The current state of cryptocurrencies is that they are currently an unregulated market, but due to the development of interest it poses the question to central banks if and how they should be regulated. The unregulated and decentralized nature leads to cryptocurrencies being used more as a speculative instrument for investment at the current time. This speculative form of investment leads to bubbles and extreme volatility in most cryptocurrency prices. Hence the need to estimate the financial risk involved through appropriate risk models to prevent the potential sizeable losses one could face through exposure to the cyrptocurrency market.

Since the mid 1990's, regulators and banks have adopted value-at-risk measurement as the premier risk metric for estimating the capital risk of investing in an asset after it was made public to use after being created by JP Morgan. After the 2008 financial crisis, financial institutions have had to follow a set of rules under the Basel III regulatory framework, putting into place a framework of capital requirements and risk management procedures. Basel II, also introduced after the 2008 financial crisis, put into place a regulation for banks to use risk models to estimate the capital required from risk up to the 99th percentile of a one tailed confidence interval.

Value-at-risk (VaR) measures the highest amount of money that may be lost within a range of confidence (or put in other words, not exceed at certain confidence) on an investment in a financial asset or a portfolio of assets within a time interval. The attractiveness of VaR as a risk metric is its relative ease in computation and how it can be compared across different assets as a quantifier for risk with a single number.

Related to VaR is expected shortfall. Due to inadequacies in some VaR calculations, practitioners and researchers have started to use an alternative to the popular measure. Expected Shortfall, which may also be called expected tail loss or conditional VaR, is a spectral risk measure which is used for measurement when the the expected value becomes negative. In other words, this means expected shortfall looks to answer the question of when things get bad, how much are we expected to lose?

The focus of this paper is to examine the risk measures of four of the most liquid cryptocurrencies that are part of the MVIS CryptoCompare Digital

Assets 10 Index[2] and investigate which risk model performs the best through various backtesting procedures. Firstly, the returns are analysed using various GARCH models on a rolling window basis to investigate the econometric properties of the data. Secondly, various VaR methods are implemented on the data and the the performance of the methods are assessed using various backtesting criteria. Thirdly, an expected shortfall model is implemented on the data and again the performance of the model is asses using various backtesting criteria.

[2] https://www.mvis-indices.com/indices/digital-assets/mvis-cryptocompare-digital-assets-10

2 Literature Review

The research literature relevant to the work in this paper can be categorised into two sections. The first section outlines the important literature on cyrptocurrency in general, to provide the basis on why research into risk analysis on cyrptocurrency is important for financial research. The second subsection investigates the literature of risk measures on cyrptocurrency, this includes volatility analysis and the use of spectral risk measures. As such, this literature review shall follow these two subsections.

Much of the research into cryptocurrencies has been conducted mainly on BTC. Bitcoin is the most popular cryptocurrency in literature and in the financial world in general due to its liquidity and market capitalisation, which was stated previously in the introduction to this paper. Price formation in the BTC market has been researched by various groups and is an important topic to understand the motivation of conducting risk analysis on cryptocurrency markets. One of the first papers looking into BTC price dynamics was conducted by Buchholz et al. 2012, where they determined that BTC price is driven by supply and demand interaction. The work of Baek and Elbeck 2015 also agrees with the theory that Bitcoin prices are driven by supply and demand concluding that bitcoin is a highly speculative instrument. This simple supply and demand explanation may now be complicated as there is now a crowded field for cyrptocurrency competition. Recent research into the price dynamics of BTC by Kjaerland et al. 2018 suggest that Bitcoin prices are affected by the S&P500 index and Google searches which is line with previous literature in this topic. More recently, Blau 2017 conducted research into the area of bitcoin price dynamics arguing that Bitcoin's large volatility spikes and value are not determined by speculative trading. Research in price bubble formation in the cyrptocurrency market is also another important subject area for volatility analysis. Phillips, Shi, and Yu 2013 found two price bubbles of BTC/USD through their research which were formed in 2013. Their research was conducted using generalized ADF test. Cheung, Roca, and Su 2015 went on to extend the work of Phillips, Shi, and Yu 2013 through using the same generalized ADF test. The group used a data range of daily returns from 2010-2014, finding three large bubbles in the periods of 2010-2013 where these bubble lasted from 66-106 days. The largest of these bubbles was the demise of Mt. Gox, which was a bitcoin exchange that was hacked, resulting in huge financial losses for any participant keeping currency in the exchange. These findings have obvious financial risk implications for investors. An investor could find themselves with severe losses if caught by one of these large bubbles. It has has implications for financial risk modelling with multiple structural breaks appearing in the data, where standard forecasting models struggle to quickly account for these structural breaks.

Much of the current literature investigating risk measures of cryptocurrency focuses on Bitcoin and testing the performance of GARCH models. GARCH models can be applied in-sample or out-of-sample and most of the literature focuses on in-sample data. Katsiampa 2017 conducted research on in sample BTC data in a time period of 2010-2016, testing the performance on some of the most popular GARCH models. The GARCH models tested being; GARCH(1,1), E-GARCH, T-GARCH, A-PARCH, CGARCH and AR-CGARCH. Through the descriptive statistics a non-normal distribution was found in the BTC returns. Furthermore, when comparing the GARCH models through log-likelihood testing and information criteria, the AR-CGARCH model came out as being the most appropriate for modelling bitcoin volatility. Chu et al. 2017 extended the studies of Katsiampa 2017 by testing 12 different GARCH family models on 7 of the most popular cryptocurrencies available on exchanges; Bitcoin, Dash, Dogecoin, Litecoin, Maidsafecoin, Monero and Ripple. The first step of the group was to conduct a maximum likelihood estimation for the parameter estimation of each GARCH model then performing testing criteria to each of the models and finally conducting unconditional and conditional coverage value-at-risk exceedance tests to evaluate which models performed the best. Evidence of volatility clustering is found in the returns suggesting that GARCH family models are most appropriate for modelling cyrptocurrency volatility for in-sample data. From the testing it was concluded that both the I-GARCH and GJR-GARCH models performed best out of the GARCH family models for the most popular coins. The I-GARCH model outperforming the other models could be due to a structural change in the data that effects multiple cryptocurrencies, and the GJR-GARCH model is a vanilla GARCH model containing asymmetric effects which are common in financial return series which would be no surprise if cyrptocurrency returns contained these effects too.

Ardia, Bluteau, and Rüede 2019 also looked to extend the work of Katsiampa 2017 by testing whether Markov-switching GARCH (MS-GARCH) models outperform single regime GARCH models for bitcoin volatility dynamics and VaR forecasting. The volatility dynamics were performed on in-sample data and found that regime switching models performed better over this in-sample period as an inverted volatility effect is observed in all volatility regimes of bitcoin. The MS-GARCH models also outperformed all single regime models for 1 step ahead VaR forecasting. Employing a rolling window method for the 1 day step ahead VaR forecasting a two-regime MS-GARCH model performed more accurately than the no-regime GARCH model and the three-regime more accurately once again overall performing the best. The fact that regime switching models perform more accurately would suggest that structural breaks frequently within the data.

Trucios 2019 looked to fill a gap in the literature through using out-of-sample forecasting on BTC daily log returns. Firstly, the group set out to test the performance of several classical GARCH family models on an out-of-sample basis to evaluate which one is most adequate fore forecasting bitcoin volatility. Secondly they addressed the presence of outliers on the data to test whether volatility can be forecast more accurately with these effects. The group performed thorough testing on the forecasting procedures and found the AV-GARCH model to perform best out of the classical GARCH family models for forecasting of returns, but was outperformed in comparison to the robust GARCH model including outliers. When using VaR estimation with the non-robust approach, all models including the best performer AV-GARCH reported a large proportion of fails, whereas the robust approach did not and non of the test rejected the null-hypothesis. From the results it can be concluded that it is important to include outliers when forecasting bitcoin volatility. Although this paper only focused on bitcoin, it is likely that other major cryptocurrencies have the similar effects of outliers and should be forecast considering such effects.

Another section of the current literature focuses on comparing the spectral risk measures of cryptocurrencies against other popular assets. Two papers, Stavroyiannis 2017 and Stavroyiannis 2018, both published from the same author, focused on comparing the value-at-risk and expected shortfall metrics of different cryptocurrencies against the S$P500 index. The first paper, published as an electronic journal, compared the likes of Bitcoin, Ethereum, Litecoin and Ripple against the index using GARCH modelling followed by filtered historical simulation. The GARCH model used is the GJR-GARCH (1,1) which accounts for leverage effects. The filtered historical simulation is a method for calculating VaR and expected shortfall which combines monte carlo simulation and historical simulation via bootstrapping on standardized residuals. Testing at different confidence levels; 90%, 95%, 97.5% and 99% over a 10 day horizon it was found that each of the cryptocurrencies had a significantly higher measure of VaR and ES at each confidence level when compared to the index. This has obvious implications for an investor as one would need much a much higher capital requirement when investing in any of the major cryptocurrencies. The second of the papers (Stavroyiannis 2018 looks to compare bitcoin to three popular assets; the S&P500, Brent crude oil and gold. Again using the GJR-GARCH model and this time employing more rigorous back testing on the measurements when compared to the last paper. The author employs the Kupiec likelihood-ratio test, Christofferson unconditional coverage test and the Christoffersn and Pelletier conditional coverage test finding that bitcoin has more VaR failures. As with the previous VaR tests, bitcoin performed the worst when it came to the expected shortfall testing. Again this shows the obvious implications for investing in bitcoin as a significantly higher capital requirements would be needed under the current regulations for financial risk management.

Although the likes of bitcoin and other major cryptocurrencies are defined as digital currencies, they are mostly traded as speculative assets unlike the major tangible currencies of the world. If bitcoin and other major currencies are to be considered in the same light as other major currencies it is important to compare the risks involved between them. Uyar and Kahraman 2019 set out to compare the conventional currencies of the world to bitcoin through VaR methods. The major currencies the author looked to compare the bitcoin against are; Swiss Franc, Euro, UK Pound, Japanese Yen, Australia Dollar, Canadian Dollar and New Zealand Dollar with all, including bitcoin, including US Dollar as numeraire. Using both historical simulation and monte carlo simulation methods to calculate the VaR at 95% and 99% confidence levels it's found that bitcoin has a much higher level than the tangible currencies. This study again agrees with previous literature that bitcoin is a highly risky financial asset when compared with other mainstream financial assets.

One of the most extensive studies into the research of risk metrics in cryptocurrencies was conducted by Kopytin, Maslennikov, and Zhukov 2019. As has been stated, most of the current literature has been conducted on bitcoin and some branching into including other major cryptocurrencies. To fill the gap in the literature, Kopytin, Maslennikov, and Zhukov 2019 performed 1-day non-parametric VaR and expected shortfall measures on 283 cryptocurrencies at 99% confidence level and compared these metrics to 61 popular assets (including stocks, indices, commodities, bonds and currencies) in the financial markets. From their analysis, only Tether and Bitusd came out as less risky assets than bitcoin, with all other cryptocurrencies being riskier than the most popular bitcoin. Extending their analysis, the authors created portfolios of top-30, top-20 and top-5 cryptocurrencies by market capitalisation to test what additional risk an investor would take on by adding different cryptocurrencies to a portfolio including bitcoin. As is expected an equally weighted portfolio is riskier than one weighted by market capitalisation as bitcoin has the highest market cap. All the portfolios carry more financial risk than bitcoin alone, but less risky than the popular Ripple and Ethereum. When comparing the risk metrics of the most volatile assets to that of the cryptocurrencies, even the least risky cyrptocurrency bitcoin had a higher risk measure than the most volatile asset natural gas. Also conducting correlation clustering analysis on the cryptocurrencies it's found that an investor can diversify their risk through including Stellar and Monero in their cryptocurrency portfolios as these have the most prominent footprints after Bitcoin.

All of the current literature points to Bitcoin and cryptocurrencies in general being highly volatile assets and having obvious implications for financial risk management. Current literature also points to GARCH family models as the most capable for volatility modelling whether it be for in-sample modelling or out-of-sample forecasting, with the GARCH models that are more equipped to react to structural breaks in the data as the most accurate. There is an obvious gap in the literature to investigate the risk measures of more than just bitcoin and to look deeper into the data than daily time steps for spectral risk measurement. Many hedge funds and financial institutions will perform VaR analysis on hourly trading data due to high volume trading. Many cyrptocurrency speculators whom use algorithmic trading will also perform high volume trading so VaR at an hourly window is an important area to investigate and as such, hourly data points shall be analysed in this paper. The work in this paper is a natural extension to the work of

Katsiampa 2017 and Chu et al. 2017, with a look into which model is the most accurate at in-sample analysis of VaR and ES measures using hourly data points in the time period of 16/11/2017 to 20/07/2020 and using BTC as the numeraire, instead of the popular USD.

3 Data

The data of the cryptocurrencies included in this study was accessed from CryptoDataDownload[3] which has a database from all the main cryptocurrencies exchanges from around the world. All the data accessed from this site for this study was from the Binance exchange[4]. As cryptocurrencies are traded every time of every day, 7 days of data are used per week, with hourly data points. The data range of the sample is from 6/11/2017 to 20/07/2020, this is to include the tokens that were not trading on an exchange until the starting data of the sample range. We analyse the cryptocurrencies in parity of BTC as the numeraire as a lot of the cryptocurrency traded is exchange based, where the investor will trade often for cyrptocurrency to cyrptocurrency instead of a major national currency after their initial investment. There is also a gap in the literature to investigate this parity as much, if not all the current literature investigates cryptocurrency as USD as the numeraire.

The price data is converted into returns data for each cyrptocurrency through using the method of logarithmic returns as is the most popular method in financial literature in this area. The first step in our analysis of the data is to compute the descriptive statistics of the log returns of each cryptocurrency.

All of the cryptocurrencies exhibit excess kurtosis, suggesting the distribution of the returns are non-normal. This information is verified by the Jarque-Bera test for normality, where the test will return 0 for a normal distribution and 1 to reject the null hypothesis and indicate non-normality. For all the cryptocurrencies we reject the null hypothesis of normality in favour of the alternate hypothesis of non-normality. LTC and XRP in particular have very high skewness, which would indicate more extreme outcomes in the data. Both skewness and kurtosis are indicators of the extremeness of the data, this would suggest that XRP has the most extreme changes in hourly returns from glancing at the descriptive statistics alone.

[3]https://www.cryptodatadownload.com/data/

[4]https://www.binance.com/en

Statistics	ETH/BTC	EOS/BTC	LTC/BTC	XRP/BTC
Mean	-1.84e-05	4.28e-05	-3.79e-05	-9.80e-06
Median	0.0000	0.0000	0.0000	-1.91e-04
Maximum	0.0949	0.1712	0.1749	0.5916
Minimum	-0.0931	-0.1348	-0.0963	-0.4136
Std. Dev.	0.0069	0.0108	0.0081	0.0120
Skewness	0.5451	0.9678	1.9705	5.5431
Kurtosis	26.9059	28.5725	40.7661	372.2476
Jarque-Bera	1	1	1	1
Observations	23687	23687	23687	23687

Table 1: The descriptive statistics of our cyrptocurrency data. A 1 in the Jarque-Bera test indicates non-normality in the returns data.

From the figures of the price paths over time of each cyrptocurrency it's clear to see the volatile nature of each one. The hourly prices vary drastically over time. The figures of the returns data also shows a clear picture of how volatile each cyrptocurrency is, with each one exhibiting sharp spikes in volatility and volatility clustering on a regular basis. We can also observe similar spikes in the returns data across all four cryptocurrencies with a major volatility cluster at the start of 2018. Some of the volatility dynamics might be driven by the fact that we use Bitcoin as the numeraire, and this might explain the joint volatility characteristics that can be seen.

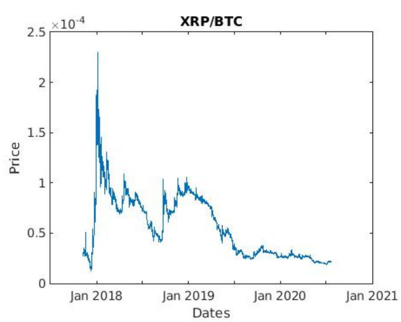

Figure 1: Price plot of XRP/BTC.

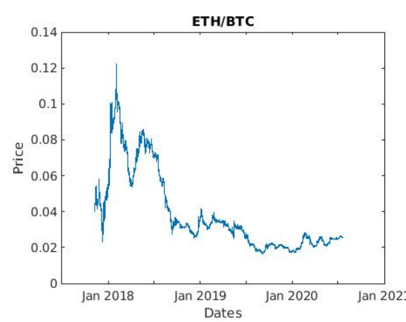

Figure 2: Price plot of ETH/BTC.

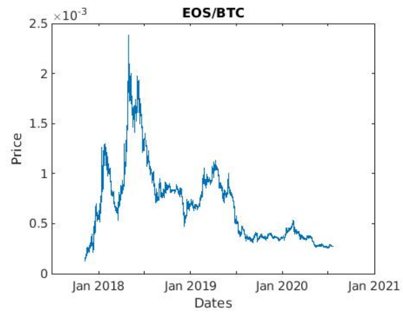

Figure 3: Price plot of EOS/BTC.

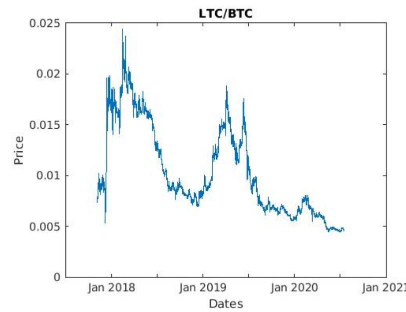

Figure 4: Price plot of XRP/BTC.

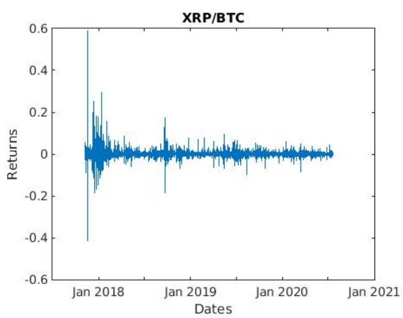

Figure 5: The returns plot of XRP/BTC.

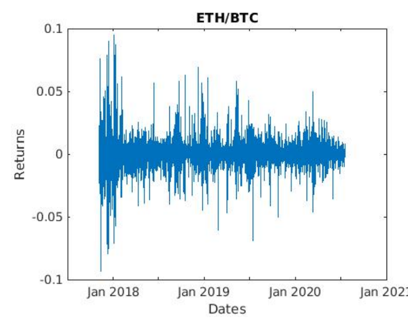

Figure 6: The returns plot of ETH/BTC.

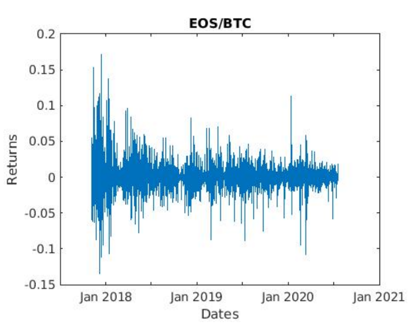

Figure 7: The returns plot of EOS/BTC.

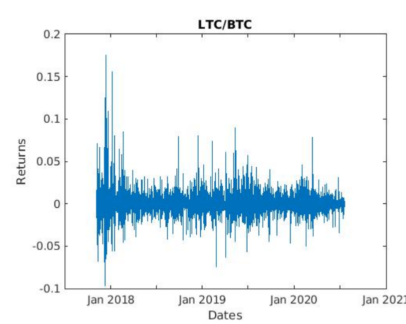

Figure 8: The returns plot of LTC/BTC.

4 Methodology

To investigate which VaR model performs the best over our data range of cryptocurrencies, we carry out various forms of popular VaR and ES methods and then backtest each method through various VaR and ES backtesting methods. These VaR and ES models are; Historical Simulation, Parametric Normal, Exponentially Weighted Moving Average, GARCH(1,1), GJR-GARCH(1,1) and E-GARCH(1,1). All programming has been performed in the MATLAB environment. A test window starting at January 1st 2018 up until the end of the data is used as our in-sample forecast range for the VaR calculations.

4.1 Econometric Models

Before the VaR calculations, a rolling window analysis is conducted on the GARCH models. Through the parameters estimated in the rolling window analysis, we can analyse what effects might be occurring in each of the cryptocurrency hourly return series.

Exponentially Weighted Moving Average (EWMA)

The EWMA model is one of the simplest models for predicting a security's (or portfolio's) volatility to then be used in VaR calculations, and has been popular ever since it was introduced by JP Morgan's RiskMetrics in 1996. The main concept of the model is that a security's past volatility can be used to predict its future volatility through a weighted average of its past returns, exponentially declining through history.

The following is the equation for calculating volatility at time t through the EWMA model:

$$\sigma^2 = \sum_{i=0}^{\infty}(1-\lambda)\lambda^i\left(r_{t-i-1}-\bar{r}\right)^2$$

Or more generally:

$$\sigma_t^2 = (1-\lambda)\epsilon_{t-1}^2 + \lambda\sigma_{t-1}^2$$

Where:

- λ = The smoothing factor ($0 < \lambda < 1$) which can be estimated from the data through maximum likelihood estimation. The lower the value of λ, the more reactive the volatility is to market shocks and events. The value for λ is commonly set to 0.94 as quoted by RiskMetrics (J.P.Morgan 1996) and for the purpose of this study this is the value that we use.
- ϵ_{t-1} = The unexpected return at time t-1.
- σ_{t-1} = The volatility at time t-1.

Past literature and tests have shown that the EWMA model is inaccurate when compared to the more sophisticated volatility models available to use. i.e GARCH family models. Due to it's simplicity in application, it is still one of the most widely used and popular models for VaR application within financial institutions. It is important to test whether the EWMA model can accurately cope with the volatile nature of cyrptocurrency data.

GARCH(p,q)

The GARCH family models are were specifically designed to model volatility clustering in financial returns, as this is something the commonly used EWMA model does not account for. Volatility clustering is the behaviour of volatility being exceptionally high or exceptionally low for periods of time, and this has obvious implications for risk management. The models also accounts for fat tails in the distribution of returns. Fat tails are when the the probability distribution moves 3 or more standard deviations more than a normal distribution, meaning that more extreme outcomes in the returns are likely. Fat tails are seen more in times of crisis, so again have obvious implications for risk management. The strength in GARCH models is the fact that they can provide short-medium term volatility forecasts based on an accurate econometric model. (Alexander 2008)

The 'vanilla' GARCH(1,1) model estimates the volatility of of returns through the following equation:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_i \sigma_{t-i}^2$$

Where:

- α = The error term of the equation and accounts for the reaction of conditional volatility to market shocks. This is also called the news term.
- β = The lag term of the equation and measures the persistence of conditional volatility. This is also called the memory term.
- ω = The mean reversion drift in the model and is a constant.
- When the sum of α and β is relatively large, the volatility term structure becomes relatively flat. This sum also tells us how fast the predictability of the model dies out. If $\alpha + \beta$ is close to zero then predictability of the process will die out quickly and if the sum is close to one then the process will die out slowly.

Since the model is for conditional variance, two conditions of the models are covariance stationarity and non-negativity ($\sigma_t^2 > 0$). These can be simply tested through the model estimation. In the GARCH model, the covariance stationary condition is satisfied when the summation of α and β is less than 1. The non-negativity of the model is satisfied when each parameter is greater than zero, i.e ω, β, $\alpha > 0$. If either of these constraints are violated, the the GARCH model being used is deemed inadequate for the modelling purpose.

GJR-GARCH(p,q)

The GJR-GARCH model was proposed by Glosten, Jagannathan, and Runkle 1993 as an extension to the vanilla GARCH model. The model extends the vanilla GARCH model through adding a leverage factor. The leverage effect is empirically observed in financial returns where a negative shock at time t-1 has more of an impact on volatility at time t than positive shocks do. The leverage effect is named as such as it was thought that the increased volatility was caused by the increased leverage from the negative shock which nowadays is known not to be true so is called the asymmetric effect.

The GJR-GARCH(1,1) model is expressed by the following equation:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma I_{t-1}\epsilon_{t-1}^2$$

The equation is similar to the equation for GARCH(1,1) with added terms to account for the leverage effect. If γ is positive (> 0) then we observe a leverage effect in the data series. The indicator term I_{t-1} will be equal to 1 when ϵ_{t-1} is negative and will equal 0 if positive. The volatility persistence is given by $\alpha + \beta + \gamma/2$ and must be less than 1 for covariance non stationary to be satisfied. As with vanilla GARCH, we require the parameters α, β and ω with the addition of $\alpha + \gamma$ to be positive for non negativity of the variance equation.

E-GARCH(p,q)

Another extension of the vanilla GARCH model is the Exponential-GARCH (E-GARCH) model proposed by Nelson 1991. It looks to address the volatility clustering in a time series, or heteroscedasticity. Much like the GJR-GARCH model before, it also tries to account for the leverage effect. The model also has the same parameters as the vanilla GARCH with the added leverage parameter as in GJR-GARCH. The conditional volatility is given by the following equation:

$$\ln\left(\sigma_t^2\right) = \omega + \beta \ln\left(\sigma_{t-1}^2\right) + \gamma \frac{\epsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[\frac{|\epsilon_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right]$$

As can be seen from the equation the conditional volatility will always provide a positive output from the model with he use of the natural logarithm in the left hand-side volatility term. For a leverage effect to be observed we would see $\gamma < 0$.

4.2 VaR Models and Calculations

Historical Simulation

One of the simplest methods of calculating VaR is through the method of historical simulation. This methods operates on the assumption that all future volatility events have happened in the previous data. To find the VaR through historical simulation method, the returns are sorted from lowest to highest which can be illustrated in a histogram. Taking 1-α multiplied by the sample size, e.g 1000 sample size and $\alpha = 99\%$, the VaR would then be the 10th (10th data point in the lower quantile of the tail) worst return from the sample. To calculate this for each day, this would be applied on a rolling window basis. The obvious implications with this method is that it relies heavily on the size of the sample being used in the rolling window, too small a sample size would not include enough data points in the lower quantile of the distribution to accurately forecast the VaR.

Parametric VaR

Often called the variance-covariance method, with this method an assumption on the distribution is made. With the assumption of normal distribution, the mean and standard deviation is found for a rolling sample size. That is for time t, the standard deviation is found for say the previous 250 days and is assumed to be the standard deviation at time t. To find the VaR, the standard deviation is multiplied by the z-score of the one-sided confidence interval. In this paper, to set up the parametric normal method, we use an estimation window size of 250 data points, as with the historical simulation. To find the positive value of VaR, we use the following calculation and multiply it by -1:

$$\text{VaR}_\alpha(X) = \Phi^{-1}(1-\alpha)\sigma + \mu$$

Where:

- α = The confidence level e.g $\alpha = 95\%$.
- Φ^{-1} = The inverse cumulative standard normal distribution.
- σ = The standard deviation.
- μ = The mean which we can assume to be zero.

For calculating the VaR using the EWMA and GARCH family models methods, the volatility is found through the model and using the same calculation as in the parametric normal method for VaR above, the value of VaR at each data point is found using a loop within the MATLAB environment.

Expected Shortfall

Expected Shortfall (ES for short) is the value that is expected to be lost on an investment if VaR is breached and is more sensitive to the tail shape of the distribution. For this reason, it's seen as a more robust measure than VaR itself. When finding the ES, it must be found together with the VaR as ES depends on VaR estimate being exceeded. The new Basel regulations propose using 97.5% ES for 99% VaR confidence estimates. [5] To find the ES under normal distribution and quote it as a positive value we multiply the value gained from using the following equation by -1:

[5] Basel Committee on Banking Supervision. "Minimum Capital Requirements for Market Risk." January 2016, https://www.bis.org/bcbs/publ/d352.htm

$$\mathrm{ES}_\alpha(X) = \mu - \sigma \cdot \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha}$$

Where:

- ϕ = The probability density function.

4.3 Backtesting Procedures

VaR Backtesting

A number of backtesting methods are used to test the performance of our VaR models:

- Traffic Lights Test[6] - One of the most basic ways of testing VaR performance. Based on a binomial distribution where N is the number of observations, p = 1 - VaR level and x is the number of failures that occur. $F(x \mid N, p)$ is the cumulative distribution of observing up to x failures.

$$\begin{aligned}&\text{Green: } F(x \mid N,p) \leq 0.95\\ &\text{Yellow: } 0.95 < F(x \mid N,p) \leq 0.9999\\ &\text{Red: } 0.9999 < F(x \mid N,p)\end{aligned}$$

$$\text{Probability} \;=\; \text{Probability}\;(X \leq x \mid N,p) = F(x \mid N,p)$$

- Bin Test[7] - A test to see whether the number of failures is consistent with the VaR confidence level. The test statistic for the Bin test is:

[6]Basel Committee on Banking Supervision, Supervisory Framework for the Use of 'Backtesting' in Conjunction with the Internal Models Approach to Market Risk Capital Requirements. January, 1996, https://www.bis.org/publ/bcbs22.htm.

[7]Jorion, P. Financial Risk Manager Handbook. 6th Edition. Wiley Finance, 2011.

$$ZScoreBin = \frac{(x - Np)}{\sqrt{Np(1-p)}}$$

The p-value is defined as two times the tail probability as this is not a one-sided tail test. The p value is:

$$\text{PValueBin } = 2 * \text{ TailProbability } < \text{ alpha}$$

Where:

$$\text{TailProbability } = 1 - F(\mid \text{ZScoreBin} \mid)$$

- Proportion of Failures (POF) Test - A likelihood ratio test proposed by Kupiec 1995. As with the Bin test, it test whether the proportion of failures is in line with the VaR confidence level but this time using a likelihood ratio test with an asymptotically distributed chi-square distribution with 1 degree of freedom test statistic:

  $$\text{LRatioPOF } = -2\log\left(\frac{(1-pVaR)^{N-x} pVaR^x}{\left(1-\frac{x}{N}\right)^{N-x}\left(\frac{x}{N}\right)^x}\right)$$

  $$= -2\left[(N-x)\log\left(\frac{N(1-pVaR)}{N-x}\right) + x\log\left(\frac{NpVaR}{x}\right)\right]$$

  And a p-value:

  $$\text{PValuePOF } = 1 - F(\text{LRatioPOF})$$

  Where F is an asymptotically distributed chi-square distribution. The test is accepted if $F(\text{LRatioPOF}) < F(\text{ TestLevel })$ and rejected if otherwise.

- Time Until First Failure (TUFF) Test - Another test proposed by Kupiec 1995, for testing if the periods until failure are consistent with the

VaR confidence level. The asymptotically distributed as a chi-square distribution with 1 degree of freedom test statistic is given by:

$$\text{LRatioTUFF} = -2\log\left(\frac{p\,\text{VaR}(1-p\,\text{VaR})^{n-1}}{\left(\frac{1}{n}\right)\left(1-\frac{1}{n}\right)^{n-1}}\right)$$

$$= -2(\log(p\,\text{VaR}) + (n-1)\log(1-p\,\text{VaR}) + n\log(n) - (n-1)\log(n-1))$$

And a p-value:

$$\text{PValueTUFF} \ = 1 - F(\text{LRatioTUFF})$$

Where F is an asymptotically distributed chi-square distribution. The test is accepted if $F(\text{LRatioTUFF}) < F(\text{TestLevel})$ and rejected if otherwise.

- Conditional Coverage Independence (CCI) Test - A likelihood ratio test proposed by Christoffersen 1998 for assessing the independence of failures on consecutive time periods. The test statistic is given by:

$$\text{LRatioCCI} \ = -2\log\left(\frac{(1-pUC)^{N00+N10}pUC^{N01+N11}}{(1-p01)^{N00}p01^{N01}(1-p11)^{N10}p11^{N11}}\right)$$

Where:

$$pUC = \frac{(\text{N01}+\text{N11})}{(\text{N00}+\text{N01}+\text{N10}+\text{N11})}$$

'N00': Number of periods with no failures followed by a period with no failures.

'N10': Number of periods with failures followed by a period with no failures.

'N01': Number of periods with no failures followed by a period with failures.

'N11': Number of periods with failures followed by a period with failures.

And a a p-value that is given by the probability that a chi square distribution of 1 degree of freedom exceeds the likelihood ratio.

$$\text{PValueCCI} = 1 - F(\text{LRatioCCI})$$

Where F is an asymptotically distributed chi-square distribution. The test is accepted if $F(\text{LRatioCCI}) < F(\text{TestLevel})$ and rejected if otherwise.

- Conditional Coverage (CC) Mixed Test - A test proposed by Christoffersen 1998, combining the CCI and POF tests to asses the independence of failures on consecutive time periods. The test statistic is given by the sum of likelihood ratios of the two tests:

$$\text{LRatioCC} = \text{LRatioPOF} + \text{LRatioCCI}$$

And a P value:

$$\text{PValueCC} = 1 - F(\text{LRatioCC})$$

Where F is an asymptotically distributed chi-square distribution with 2 degress of freedom. The test is accepted if $F(\text{LRatioCC}) < F(\text{TestLevel})$ and rejected if otherwise.

- Time Between Failures Independence (TBIF) Test - Proposed by Haas 2001 as a test for failure independence and is an extension of Kupiec's time until first failure test. It tests the time between all failures, not just the time until the first failure. Using the same test statistic as in the TUFF test, but with n_i instead of n, where n_i is the number of periods between failure i − 1 and failure i. We then use the sum of the individual likelihood for all times between failures:

$$\text{LRatioTBFI} = \sum_{i=1}^{x} \text{LRatioTBFI}_i$$

And a a p-value that is given by the probability that a chi square distribution of x (where x is the number of failures) degrees of freedom exceeds the likelihood ratio.

$$\text{PValueTBFI } = 1 - F(\text{LRatioTBFI})$$

Where F is an asymptotically distributed chi-square distribution with x degrees of freedom. The test is accepted if $F(\text{LRatioTBFI}) < F(\text{TestLevel})$ and rejected if otherwise.

- Time Between Failures (TBF) Mixed Test - A test that combines both the Kupiec 1995 POF test and Haas 2001 TBFI test. With test statistic that has an asymptotically distributed chi-squared distribution with x+1 (with x being the number of failures) degrees of freedom:

  $$\text{LRatioTBF } = \text{LRatioPOF } + \text{LRatioTBFI}$$

  And a a p-value that is given by the probability that a chi square distribution of x+1 (where x is the number of failures) degrees of freedom exceeds the likelihood ratio.

  $$\text{PValueTBF } = 1 - F(\text{LRatioTBF})$$

  Where F is an asymptotically distributed chi-square distribution with x degrees of freedom. The test is accepted if $F(\text{LRatioTBF}) < F(\text{TestLevel})$ and rejected if otherwise.

ES Backtesting

- Conditional Test - Proposed by Acerbi and Szekely 2014, a simulation based test and the test statistic is based on a conditional relationship where a distribution is assumed:

  $$ES_t = -E_t\left[X_t \mid X_t < -VaR_t\right]$$

  X_t represents the portfolio profit or loss at time t, VaR_t the VaR at time t and ES_t the expected shortfall at time t. The number of failures is given by NumFailures $= \sum_{t=1}^{N} I_t$, where I_t is the indicator function and N the number of periods in the test window. The test statistic is then given by:

$$Z_{\text{uncond}} = \frac{1}{Np_{\text{VaR}}} \sum_{t=1}^{N} \frac{X_t I_t}{ES_t} + 1$$

This test is only accepted if the both the VaR test passes and the conditional test ES passes. When the model underestimates the risk, a negative value will be produced and the test will fail and will also reject the test if the p-value is less than 1 minus the confidence level. The p-value for the ES tests is given by:

$$P_{\text{value}} = \frac{1}{M} \sum_{s=1}^{M} I\left(Z^s \leq Z^{\text{obs}}\right)$$

Where M is the amount of simulated scenarios (in our case 1000) and Z^s the test statistic. The critical value is the minimum simulated test statistic with P-Value greater than P_test.

- Unconditional Test - Proposed by Acerbi and Szekely 2014, a simulation based test and based on the unconditional relationship:

  $$ES_t = -E_t\left[\frac{X_t I_t}{p_{VaR}}\right]$$

  Where p_{VaR} is the probability of a VaR failure and is given by 1 minus the VaR level. The p-value for the unconditional test is the same as above in the conditional test and the test statistic is given by:

  $$Z_{\text{uncond}} = \frac{1}{Np_{\text{VaR}}} \sum_{t=1}^{N} \frac{X_t I_t}{ES_t} + 1$$

  The test can be run under a normal or t distribution assumption.

- Quantile Test - Proposed by Acerbi and Szekely 2014 utilises a sample estimator for the Expected shortfall and is another simulation based test, where the expected shortfall for sample y is given by:

$$\widehat{ES}(Y) = -\frac{1}{\lfloor Np_{VaR} \rfloor} \sum_{i=1}^{\lfloor Np_{VaR} \rfloor} Y_{[i]}$$

The process for computing the sample test statistic is to convert the portfolio outcome into ranks and invert them, compute the sample estimator and then to compute the expected value of the sample estimator. The test statistic is given by:

$$Z_{quantile} = -\frac{1}{N} \sum_{t=1}^{N} \frac{\widehat{ES}\left(P_t^{-1}(U)\right)}{E\left[\widehat{ES}\left(P_t^{-1}(V)\right)\right]} + 1$$

Where the denominator is given by:

$$-\frac{N}{\lfloor N_{pVaR} \rfloor} \int_0^1 I_{1-p}\left(N - \lfloor N_{pVaR} \rfloor, \lfloor N_{pVaR} \rfloor\right) P_t^{-1}(p) dp$$

The p-value is the same as for the other two simulation based tests above.

5 Empirical Results

5.1 GARCH Rolling Window Analysis

Using a rolling window analysis on each GARCH model on the different cyrptocurrency time series, we are able to gain an insight into how well the models cope with the data. The rolling window for each model is set up with an observation window of 2 months (1344 data points) and the window moves along weekly. The alpha, beta and omega parameters are saved a graphically shown for the rolling window to investigate these change throughout the time series of each cyrptocurrency.

In the figures showing the alpha estimations under rolling window analysis it is clear to see that ETH, LTC and XRP all experience a spike in the GARCH alpha value in the middle of data to as high as 0.9 in the case of XRP. In fact for much of the time series the alpha value is above 0.1, meaning that the volatility of each cyrptocurrency is incredibly sensitive to market events. This is in line with previous literature stating that cyrptocurrency markets are very sensitive to supply and demand and over market interactions. Each of the graphs show frequent spikes in the alpha value to above 0.2. The beta values for ETH and XRP especially are persistently high (above 0.9) for much of the analysis, meaning the volatility in sensitive to crisis in the market and indicates the presence of volatility clustering. Literature has stated the presence of bubbles in the market which can cause market crashes.

The rolling window analysis is also conducted using the GJR-GARCH model, plotting the leverage coefficient, gamma. Each of the cryptocurrencies experience a spike in the leverage coefficient at the midway point of the data, with ETH experiencing a positive spike, LTC, EOS and XRP experiencing a negative spike in the value of the leverage coefficient. This spike might be caused by an effect in the cyrptocurrency market effecting the major liquid cryptocurrencies at the same time. From the plotting of the leverage coefficient, it's clear to see an asymmetric effect happening in the data, where a positive shock effects the volatility more for XRP and LTC on average throughout the analysis and a negative shock effects the volatility of ETH and EOS more. The apparent presence of the asymmetric effects in the each of the time series would suggest that a GJR-GARCH or E-GARCH model would perform better at modelling the volatility of cyrptocurrency returns than vanilla GARCH.

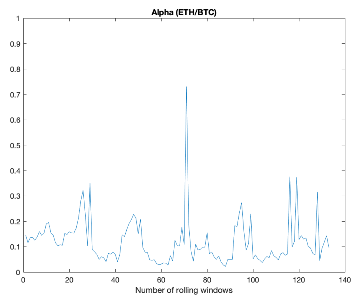

Figure 9: The values of alpha using GARCH(1,1) model rolling window analysis for ETH/BTC.

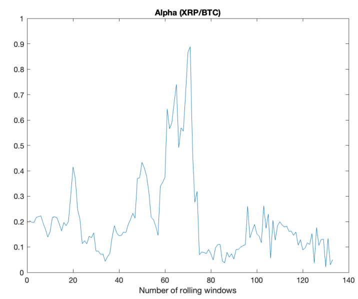

Figure 10: The values of alpha using GARCH(1,1) model rolling window analysis for XRP/BTC.

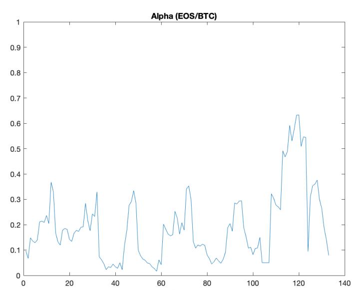

Figure 11: The values of alpha using GARCH(1,1) model rolling window analysis for EOS/BTC.

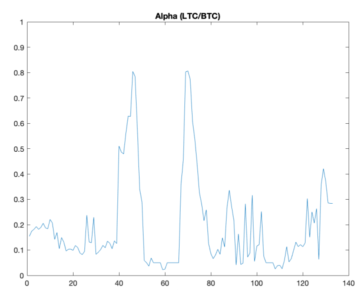

Figure 12: The values of alpha using GARCH(1,1) model rolling window analysis for LTC/BTC.

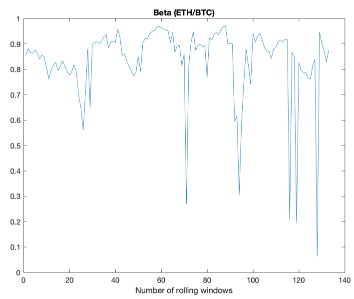

Figure 13: The values of beta using GARCH(1,1) model rolling window analysis for ETH/BTC.

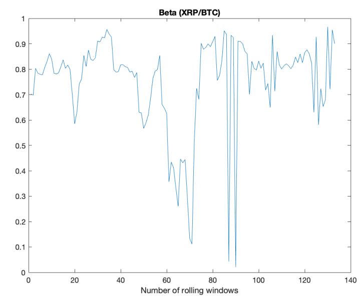

Figure 14: The values of beta using GARCH(1,1) model rolling window analysis for XRP/BTC.

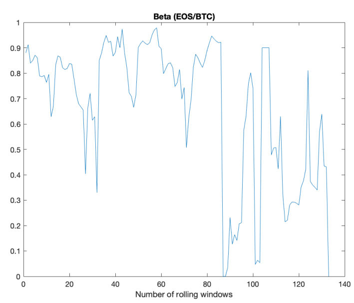

Figure 15: The values of beta using GARCH(1,1) model rolling window analysis for EOS/BTC.

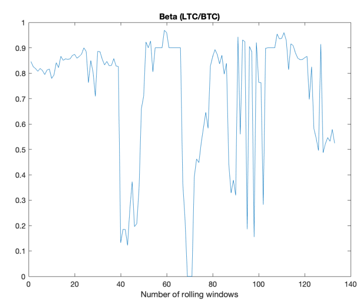

Figure 16: The values of beta using GARCH(1,1) model rolling window analysis for LTC/BTC.

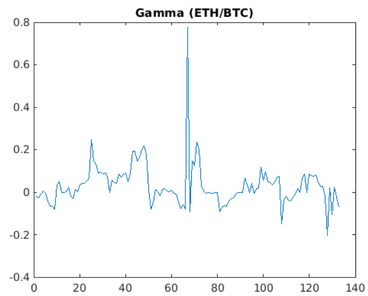

Figure 17: The leverage coefficient using GJR-GARCH(1,1) model rolling window analysis for ETH/BTC.

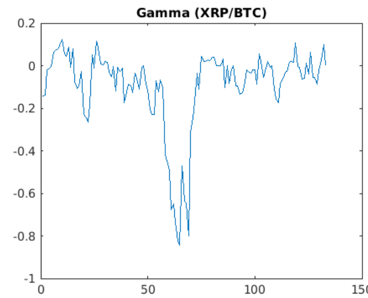

Figure 18: The leverage coefficient using GJR-GARCH(1,1) model rolling window analysis for XRP/BTC.

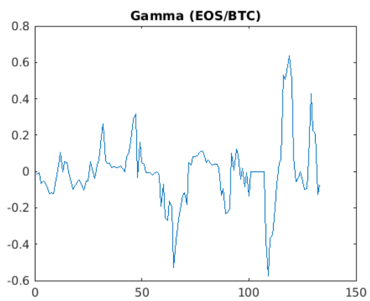

Figure 19: The leverage coefficient using GJR-GARCH(1,1) model rolling window analysis for EOS/BTC.

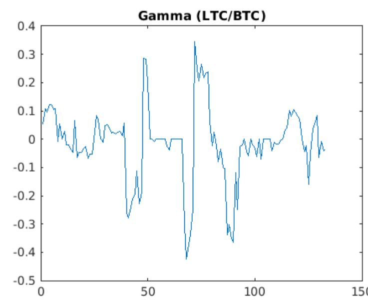

Figure 20: The leverage coefficient using GJR-GARCH(1,1) model rolling window analysis for LTC/BTC.

5.2 VaR Model Analysis

Six different models were used and backtested for the VaR estimation of four cryptocurrencies. The VaR models used are Parametric Normal, Historical Simulation, EWMA, GARCH, GJR-GARCH and E-GARCH. The VaR is tested at the 95% confidence level (often used by day traders) and 99% (the VaR level set by Basel accords). Using the various backtesting models, we see that the models struggled to cope with the volatility dynamics of the cryptocurrency data. From the models, the E-GARCH and Historical Simulation methods outperform the rest, with them having the lowest violations, as can bee seen in Table 2. The EWMA model, the simplest of the volatility forecasting models consistently performed the worst across the cryptocurrencies, with the most VaR violations. A VaR violation is when the an observation breaks the border specified by the VaR confidence level. This may be due to the persistent volatility clustering that appears within the returns times series. The performance of E-GARCH at modelling the volatility may be due to the absence of linearity restrictions on the model.

The traffic light test is a simplistic model that doesn't provide p-values for us to analyse, instead it provides the lights system. From the Table 3 we can see how well the models performed using the traffic lights test, with the E-GARCH model performing best overall.

The p-values for the test across the cryptocurrencies indicate that the models aren't performing accurately (which can be found in Tables 4-7), in fact we can see that the models performs very poorly with extremely low p-values for a lot of the tests. We can also see from the p-values for each of the tests that E-GARCH has a larger p-value than the other volatility models for 99% VaR across the cryptocurrencies under all the test apart from the CCI test. This indicates fewer VaR test violations by the E-GARCH model. The E-GARCH model with the historical simulation model both have the lowest violation ratios too. When we examine the p-values of the CCI test, we can see that the GJR-GARCH model had marginally higher p-values than the E-GARCH model for 3 out of the 4 cryptocurrencies. This indicates that the VaR violations didn't cluster together as much as the other models, which could be due to the leverage coefficient. Interestingly, the historical simulation method actually had the largest P-Values of all for the tests, but using an Historical Simulation method is not always advisable as it doesn't properly model the shape of the volatility dynamics - this can be seen by the plots of VaR model estimations found in the appendix.

	Failures (99%)	Ratio (99%)	Failures (95%)	Ratio (95%)
ETH/BTC				
Normal	426	1.91	901	0.81
Historical	275	1.23	1196	1.07
EWMA	466	2.09	1026	0.91
GARCH	337	1.51	780	0.69
GJR-GARCH	388	1.74	885	0.79
E-GARCH	310	1.39	746	0.67
LTC/ TC				
Normal	362	1.62	866	0.78
Historical	284	1.27	1206	1.08
EWMA	404	1.81	983	0.88
GARCH	309	1.38	715	0.64
GJR-GARCH	310	1.39	719	0.64
E-GARCH	297	1.33	691	0.62
EOS/BTC				
Normal	379	1.69	802	0.72
Historical	280	1.25	1182	1.06
EWMA	437	1.96	910	0.81
GARCH	312	1.39	656	0.59
GJR-GARCH	312	1.39	659	0.59
E-GARCH	298	1.33	645	0.58
XRP/BTC				
Normal	377	1.69	860	0.77
Historical	276	1.24	1148	1.03
EWMA	421	1.88	928	0.83
GARCH	263	1.18	604	0.54
GJR-GARCH	262	1.17	614	0.55
E-GARCH	203	0.91	529	0.47

Table 2: VaR failures and the ratio of VaR failures at the 99% and 95% confidence levels, where the ratio is the actual Var failures divided by the expected VaR failures. We can see that in this respect, the E-GARCH model outperforms the other parametric models.

	TL (99%)	TL (95%)
ETH/BTC		
Normal	red	green
Historical	yellow	yellow
EWMA	red	green
GARCH	red	green
GJR-GARCH	red	green
E-GARCH	red	green
LTC/ TC		
Normal	red	green
Historical	red	yellow
EWMA	red	green
GARCH	red	green
GJR-GARCH	red	green
E-GARCH	red	green
EOS/BTC		
Normal	red	green
Historical	yellow	yellow
EWMA	red	green
GARCH	red	green
GJR-GARCH	red	green
E-GARCH	red	green
XRP/BTC		
Normal	red	green
Historical	yellow	green
EWMA	red	green
GARCH	yellow	green
GJR-GARCH	yellow	green
E-GARCH	green	green

Table 3: The results of the traffic lights test conducted on the 99% VaR models and 95% VaR models.

The backtesting of the 95% VaR estimates was a little more inconclusive which just a glance at the p-values is would seem as though many of the models performed around the same level with mixed results as which one performed the most accurately across the tests. A closer look reveals that the GJR-GARCH model performed the best out of the GARCH models at this level, but was out-performed in the tests by the simpler models like EWMA and the Historical Simulation. For VaR estimation at 95% a lower level of accuracy is needed so this could be the case as to why the simpler models are performing better.

From looking at the currencies individually, we can see that the VaR for XRP/BTC was most accurately estimated. This might seem unusual as through out descriptive statistics we found that XRP exhibits the most extreme outcomes. In this case, the E-GARCH model out-performed the other models with a higher p-value across the tests.

99%	POF	CC	CCI
ETH/BTC			
Normal	9.927e-34	2.201e-39	2.005e-08
Historical	0.0008	3.182e-10	1.167e-08
EWMA	5.629e-46	2.77e-45	0.108
GARCH	8.039e-13	1.107e-12	0.051
GJR-GARCH	1.276e-23	7.914e-23	0.231
E-GARCH	3.846e-08	1.588e-08	0.017
LTC/ TC			
Normal	1.245e-17	8.106e-24	8.070e-09
Historical	9.318e-05	2.994e-09	9.717e-07
EWMA	1.094e-27	1.649e-29	0.0002
GARCH	5.405e-8	2.125e-08	0.016
GJR-GARCH	3.846e-08	1.588e-08	0.017
E-GARCH	2.494e-06	1.841e-08	0.0002
EOS/BTC			
Normal	1.851e-21	4.245e-25	3.124e-06
Historical	0.0003	6.288e-06	0.001
EWMA	6.128e-37	1.099e-38	0.0002
GARCH	2.729e-08	4.679e-09	0.006
GJR-GARCH	3.846e-08	6.242e-09	0.0059
E-GARCH	1.845e-06	1.023e-06	0.0279
XRP/BTC			
Normal	5.435e-21	1.377e-33	2.105e-15
Historical	0.0007	3.210e-17	1.049e-15
EWMA	2.598e-32	1.801e-35	7.926e-06
GARCH	0.0097	0.0008	0.0056
GJR-GARCH	0.0116	0.0025	0.0175
E-GARCH	0.1628	0.0186	0.0141

Table 4: P-values from the Proportion of Failures, Conditional Coverage and Conditional Coverage Independence tests conducted the 99% confidence VaR models.

99%	TUFF	BIN	TBFI	TBF
ETH/BTC				
Normal	0.45604	0	1.152e-59	1.158e-79
Historical	0.45604	0.0005	1.242e-31	7.403e-33
EWMA	0.45604	0	2e-23	1.781e-45
GARCH	0.45604	2.243e-14	2.424e-10	1.477e-14
GJR-GARCH	0.45604	0	9.11e-12	1.145e-20
E-GARCH	0.45604	5.887e-09	5.992e-08	3.329e-10
LTC/ TC				
Normal	0.44509	0	6.985e-44	2.894e-53
Historical	0.44509	4.665e-05	6.189e-23	1.621e-24
EWMA	0.44509	0	8.277e-17	1.389e-28
GARCH	0.44509	8.782e-09	1.535e-07	1.065e-09
GJR-GARCH	0.44509	5.887e-09	1.101e-07	6.548e-10
E-GARCH	0.44509	7.58e-07	3.534e-12	4.299e-14
EOS/BTC				
Normal	0.45604	0	3.319e-37	2.458e-48
Historical	0.45604	0.0001	2.908e-19	1.453e-20
EWMA	0.45604	0	3.862e-22	2.085e-39
GARCH	0.45604	2.609e-09	1.616e-08	6.103e-11
GJR-GARCH	0.45604	2.609e-09	1.616e-08	6.103e-11
E-GARCH	0.45604	5.356e-07	2.816e-08	5.572e-10
XRP/BTC				
Normal	0.98385	0	2.734e-59	1.950e-71
Historical	0.98385	0.0004	2.068e-37	9.657e-39
EWMA	0.45604	0	1.238e-19	3.825e-34
GARCH	0.45604	0.0078	0.0002	7.298e-05
GJR-GARCH	0.45604	0.0095	0.0048	0.0027
E-GARCH	0.45604	0.1694	0.0317	0.0291

Table 5: P-values from the Time Until First Failure, Bin Test, Time Between Failures Independence and Time Between Failure Mixed Test conducted the 99% confidence VaR models.

95%	POF	CC	CCI
ETH/BTC			
Normal	6.997e-12	2.049e-22	3.599e-13
Historical	0.0167	1.097e-17	1.777e-17
EWMA	0.0046	1.121e-09	8.444e-09
GARCH	9.275e-28	6.004e-30	9.030e-05
GJR-GARCH	1.529e-13	8.233e-16	0.00011
E-GARCH	1.115e-33	1.406e-35	0.00016
LTC/ TC			
Normal	1.108e-15	8.538e-24	9.094e-11
Historical	0.0070	3.695e-12	1.608e-11
EWMA	2.649e-05	1.203e-06	0.0019
GARCH	1.113e-39	9.446e-40	0.0147
GJR-GARCH	7.146e-39	4.355e-39	0.0102
E-GARCH	9.762e-45	5.977e-46	0.0007
EOS/BTC			
Normal	2.772e-24	8.283e-31	3.043e-09
Historical	0.0486	2.505e-10	2.146e-10
EWMA	5.287e-11	9.138e-15	3.386e-06
GARCH	8.783e-53	6.651e-53	0.0109
GJR-GARCH	4.627e-52	1.328e-52	0.0037
E-GARCH	1.757e-55	7.422e-57	0.005
XRP/BTC			
Normal	2.141e-16	1.549e-31	6.329e-18
Historical	0.3465	3.654e-20	4.771e-21
EWMA	2.299e-09	1.859e-13	1.684e-06
GARCH	2.735e-66	7.717e-70	2.108e-06
GJR-GARCH	1.529e-63	2.421e-65	0.00014
E-GARCH	2.472e-89	3.359e-91	0.00010

Table 6: P-values from the Proportion of Failures, Conditional Coverage and Conditional Coverage Independence tests conducted the 95% confidence VaR models.

95%	TUFF	BIN	TBFI	TBF
ETH/BTC				
Normal	0.7153	3.216e-11	1.381e-62	1.262e-67
Historical	0.7153	0.0156	8.447e-70	2.983e-70
EWMA	0.37075	0.0051	1.225e-17	4.342e-18
GARCH	0.37075	0	8.859e-20	1.619e-29
GJR-GARCH	0.37075	1.022e-12	1.433e-14	3.852e-18
E-GARCH	0.37075	0	1.891e-21	3.722e-34
LTC/ TC				
Normal	0.7153	1.243e-14	5.122e-53	9.919e-60
Historical	0.7153	0.00641	3.133e-54	8.559e-55
EWMA	1	3.8e-05	4.089e-10	5.485e-11
GARCH	1	0	3.73e-17	1.018e-31
GJR-GARCH	1	0	2.797e-17	1.6905e-31
E-GARCH	0.38769	0	7.875e-22	4.637e-40
EOS/BTC				
Normal	0.37075	0	5.453e-61	6.715e-50
Historical	0.50568	0.04669	6.033e-49	1.043e-48
EWMA	0.37075	2.017e-10	6.109e-19	4.379e-16
GARCH	0.37075	0	6.010e-50	2.748e-26
GJR-GARCH	0.37075	0	2.1979e-50	5.692e-27
E-GARCH	0.37075	0	7.8273e-57	1.297e-30
XRP/BTC				
Normal	0.37075	2.887e-15	9.677e-73	1.662e-80
Historical	0.37075	0.3445	5.790e-80	6.572e-80
EWMA	0.37075	6.339e-09	9.070e-17	3.959e-19
GARCH	0.37075	0	1.528e-38	3.355e-73
GJR-GARCH	0.37075	0	7.473e-31	5.113e-62
E-GARCH	0.37075	0	3.763e-44	4.681e-96

Table 7: P-values from the Time Until First Failure, Bin Test, Time Between Failures Independence and Time Between Failure Mixed Test conducted the 95% confidence VaR models.

5.3 Expected Shortfall Analysis

For the Expected Shortfall backtesting the cyrptocurrency data was split up into 6 month time periods starting from January 1st 2018 through until December 31st 2019 as the expected shortfall backtesting models can only accurately cope with up to 5000 data points of data. This analysis also allows a deeper looks into how the models perform as different periods in the data may be more volatile than others. The first two backtests, the Unconditional Normal and Unconditional T Distribution tests can be performed on all the models, but the simulated tests (Conditional, Unconditional and Quantile tests) can not be performed on the historical simulation model as a distribution is assumed about the data. The simulated tests are performed assuming a normal distribution and t distribution. The normal distribution tests are stricter than the t distribution tests which have fatter tails and allow for more extreme outcomes. The simulated unconditional test is the same test as the unconditional normal and t test, but using critical values instead based on a simulation with the mean and standard deviation provided from the data with either a normal distribution or a t distribution with 5 degrees of freedom. This simulation approach provides more precision than the non simulated unconditional tests.

When we examine the figures including the observed and expected severity, we can see the observed severity of the volatility models is much higher than that of the expected severity and this is the case across the cryptocurrencies and all time periods. The observed severity ratio is the average ratio of loss to VaR when VaR is violated. The expected severity ratio is the average ratio of ES to VaR for the violation periods.

ETH/BTC

From the figures Of the expected failures against observed failures, it's clear to see that the E-GARCH model performed consistently well with less failures than expected in three of the 4 data ranges, outperforming the other models in this aspect. The E-GARCH model also had the least test rejections from out backtesting procedures. Conversely, the EWMA model consistently had more failures observed than expected, with it being the worst performing model in this aspect. The worst time period for failures is the July2018-Dec2018 data range.

	UnconditionalNorm	UnconditionalT
Jan2018-June2018		
Historical	Reject	Accept
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Reject	Reject
GJR-GARCH	Reject	Reject
E-GARCH	Accept	Accept
July2018-Dec2018		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Reject	Reject
GJR-GARCH	Reject	Reject
E-GARCH	Reject	Reject
Jan2019-June2019		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Reject	Accept
E-GARCH	Accept	Accept
July2019-Dec2019		
Historical	Reject	Accept
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Reject	Reject
E-GARCH	Accept	Accept

Table 8: The results of the Unconditional Normal and Unconditional T tests for ETH/BTC in the four time periods we tested on.

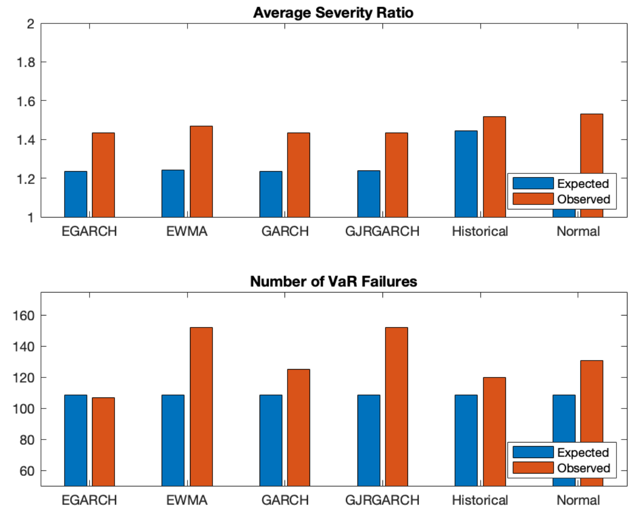

Figure 21: Histogram of the average severity ratio and number of VaR failures for Jan 2018 to June 2018 for ETH/BTC.

Figure 22: Histogram of the average severity ratio and number of VaR failures for July 2018 to Dec 2018 for ETH/BTC.

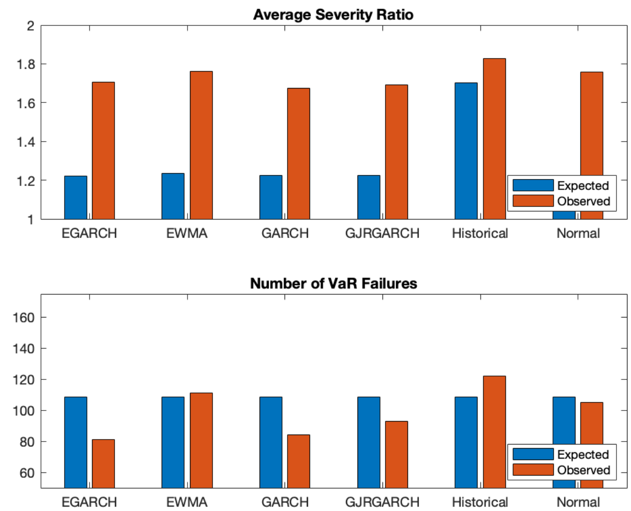

Figure 23: Histogram of the average severity ratio and number of VaR failures for Jan 2019 to June 2019 for ETH/BTC.

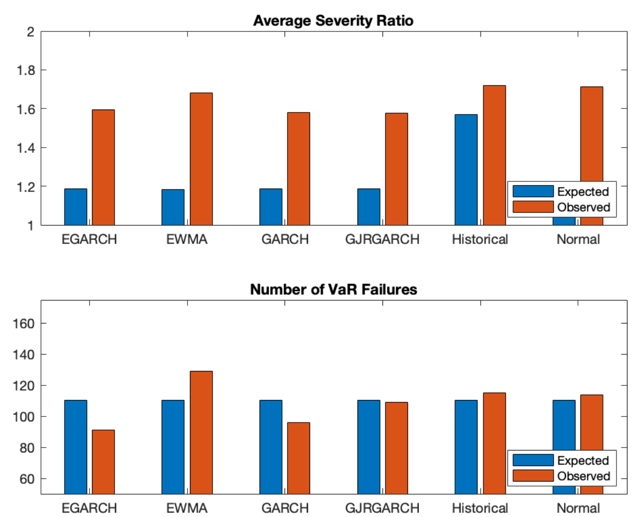

Figure 24: Histogram of the average severity ratio and number of VaR failures for July 2019 to Dec 2019 for ETH/BTC.

t Dist.	Conditional	Unconditional	Quantile
Jan2018-June2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Accept	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Accept	Accept	Reject
July2018-Dec2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Accept	Accept	Reject
GJR-GARCH	Accept	Accept	Reject
E-GARCH	Accept	Accept	Reject
Jan2019-June2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2019-Dec2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject

Table 9: The test results for the Conditional, Unconditional and Quantile ES backtests simulated under the assumption of t-distribution for ETH/BTC.

Normal Dist.	Conditional	Unconditional	Quantile
Jan2018-June2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Accept	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Accept	Accept	Reject
July2018-Dec2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
Jan2019-June2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2019-Dec2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject

Table 10: The test results for the Conditional, Unconditional and Quantile ES backtests simulated under the assumption of normal distribution for ETH/BTC.

LTC/BTC

The GARCH family models performed the best when tested with the Unconditional Normal and Unconditional T test, with each one passing at each time period, whilst the Historical, Normal and EWMA models performed much worse. As with the ETH/BTC data sets, the EGARCH model had the lowest violation ratio with EWMA having the highest violation ratio. When using the simulated tests, only the EWMA and GARCH model passed the conditional test in the first time period, with every model failing this test for the other time periods. In the first time period, every model passed the quantile test, and failing the quantile test for the other three time periods.

	UnconditionalNorm	UnconditionalT
Jan2018-June2018		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Accept
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept
July2018-Dec2018		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept
Jan2019-June2019		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept
July2019-Dec2019		
Historical	Accept	Accept
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept

Table 11: The results of the Unconditional Normal and Unconditional T tests for LTC/BTC in the four time periods we tested on.

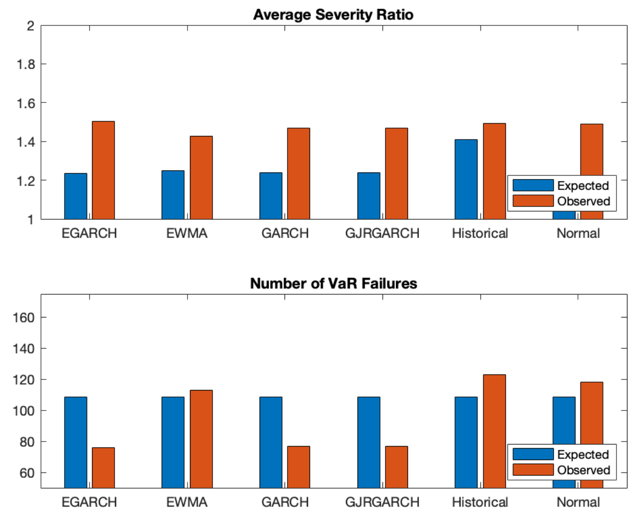

Figure 25: Histogram of the average severity ratio and number of VaR failures for Jan 2018 to June 2018 for LTC/BTC.

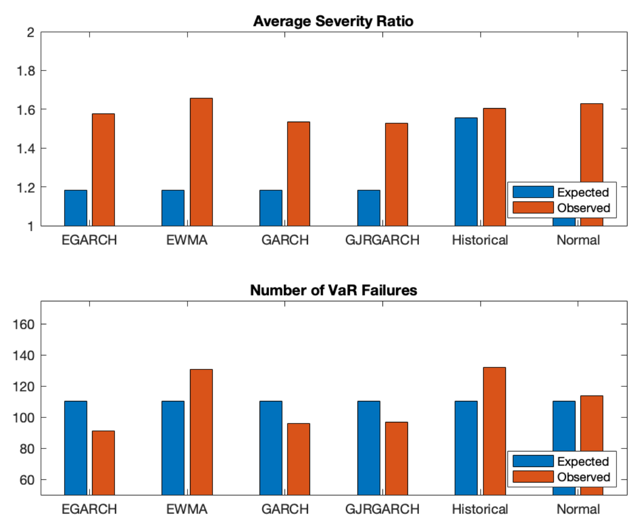

Figure 26: Histogram of the average severity ratio and number of VaR failures for July 2018 to Dec 2018 for LTC/BTC.

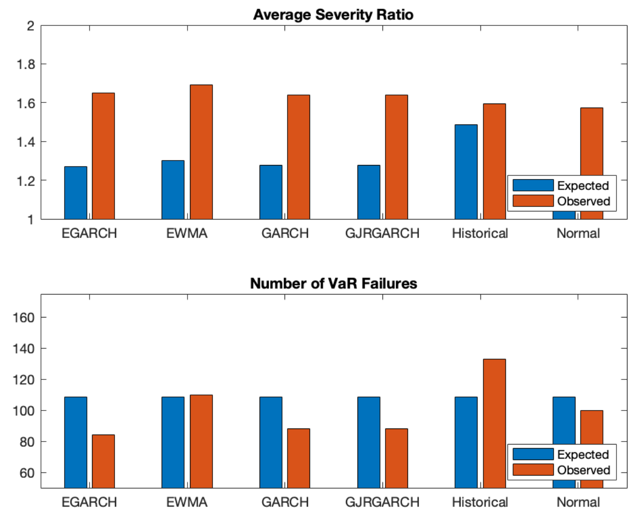

Figure 27: Histogram of the average severity ratio and number of VaR failures for Jan 2019 to June 2019 for LTC/BTC.

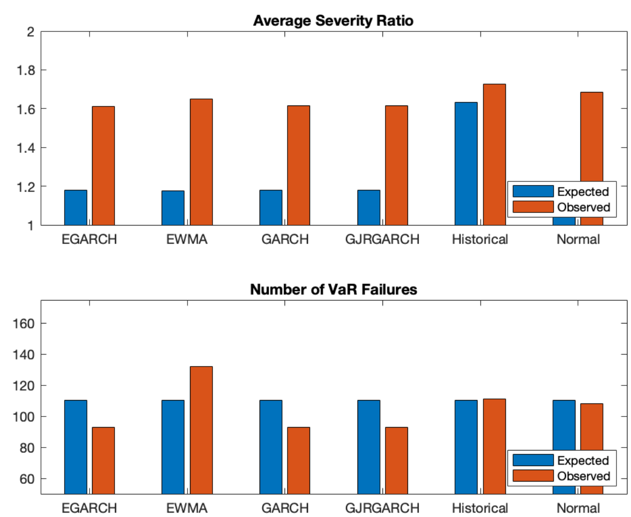

Figure 28: Histogram of the average severity ratio and number of VaR failures for July 2019 to Dec 2019 for LTC/BTC.

t Dist.	Conditional	Unconditional	Quantile
Jan2018-June2018			
Normal	Reject	Accept	Reject
EWMA	Accept	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2018-Dec2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
Jan2019-June2019			
Normal	Reject	Accept	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2019-Dec2019			
Normal	Reject	Accept	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject

Table 12: The test results for the Conditional, Unconditional and Quantile ES backtests simulated under the assumption of t-distribution for LTC/BTC.

Normal Dist.	Conditional	Unconditional	Quantile
Jan2018-June2018			
Normal	Reject	Accept	Reject
EWMA	Accept	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2018-Dec2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
Jan2019-June2019			
Normal	Reject	Accept	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2019-Dec2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject

Table 13: The test results for the Conditional, Unconditional and Quantile ES backtests simulated under the assumption of normal for LTC/BTC.

EOS/BTC

For EOS/BTC we can see that the GARCH family models performed best over first half of both 2018 and 2019, with the historical simulation performing the best of the second half of these years, when tested with the unconditional normal and t tests. When it comes to the failure ratios, the E-GARCH model marginally performs better than the GARCH and GJR-GARCH models. For the simulation based tests all models passed every test for the first time period under the assumption of a t distribution and only the normal model failing under the assumption of normal distribution. We could assume there is less volatility clustering exhibited in this time period.

	UnconditionalNorm	UnconditionalT
Jan2018-June2018		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept
July2018-Dec2018		
Historical	Accept	Accept
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Reject	Reject
GJR-GARCH	Reject	Reject
E-GARCH	Reject	Reject
Jan2019-June2019		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept
July2019-Dec2019		
Historical	Accept	Accept
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Reject	Reject
GJR-GARCH	Reject	Reject
E-GARCH	Reject	Reject

Table 14: The results of the Unconditional Normal and Unconditional T tests for EOS/BTC in the four time periods we tested on.

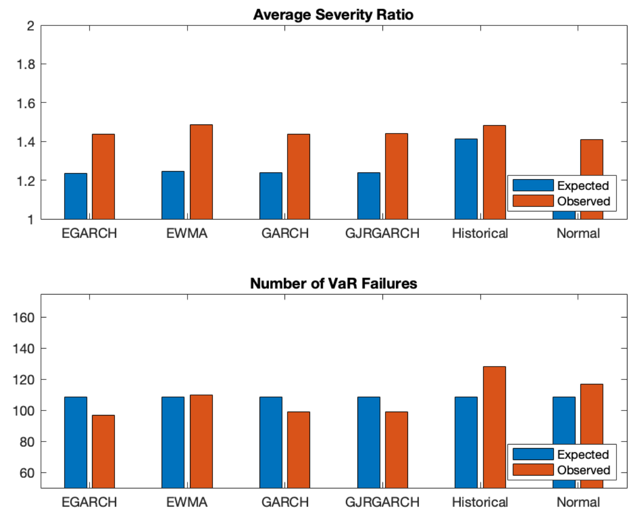

Figure 29: Histogram of the average severity ratio and number of VaR failures for Jan 2018 to June 2018 for EOS/BTC.

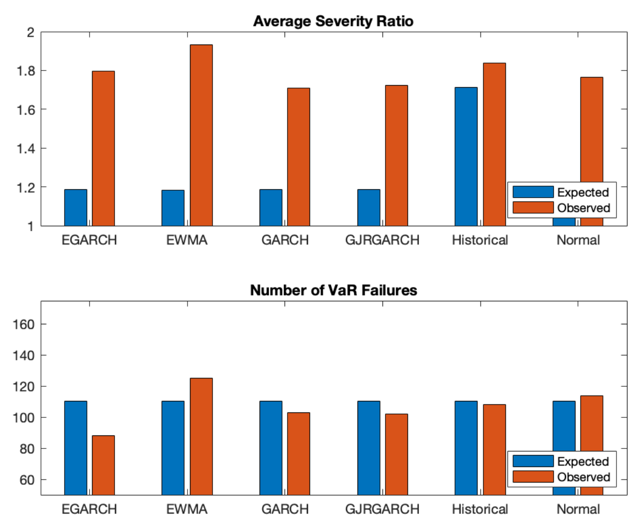

Figure 30: Histogram of the average severity ratio and number of VaR failures for July 2018 to Dec 2018 for EOS/BTC.

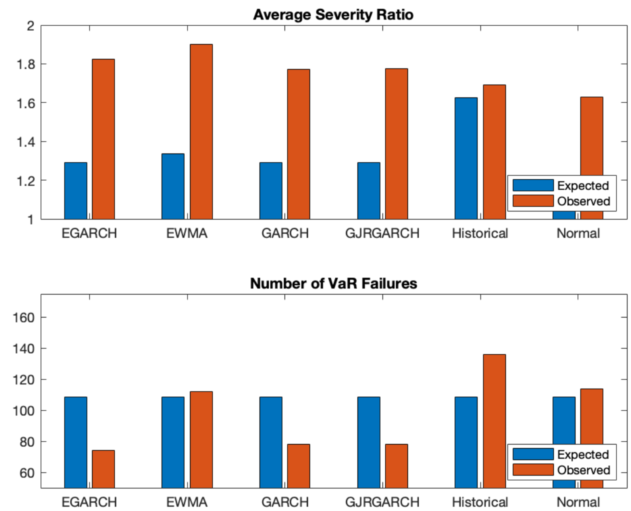

Figure 31: Histogram of the average severity ratio and number of VaR failures for Jan 2019 to June 2019 for EOS/BTC.

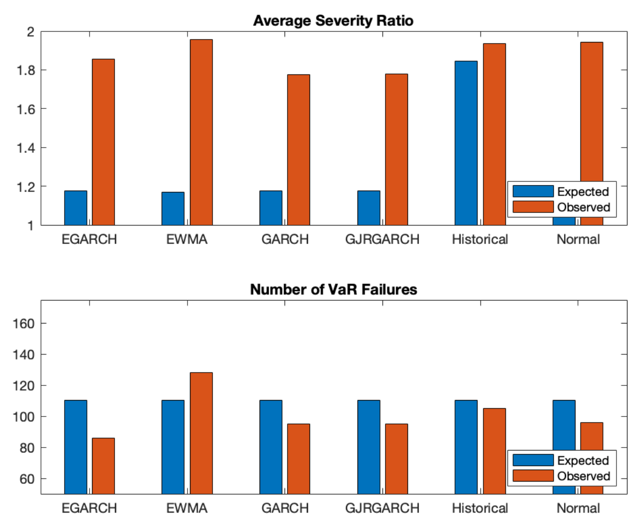

Figure 32: Histogram of the average severity ratio and number of VaR failures for July 2019 to Dec 2019 for EOS/BTC.

t Dist.	Conditional	Unconditional	Quantile
Jan2018-June2018			
Normal	Accept	Accept	Accept
EWMA	Accept	Accept	Accept
GARCH	Accept	Accept	Accept
GJR-GARCH	Accept	Accept	Accept
E-GARCH	Accept	Accept	Accept
July2018-Dec2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
Jan2019-June2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2019-Dec2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject

Table 15: The test results for the Conditional, Unconditional and Quantile ES backtests simulated under the assumption of t-distribution for EOS/BTC.

Normal Dist.	Conditional	Unconditional	Quantile
Jan2018-June2018			
Normal	Reject	Reject	Reject
EWMA	Accept	Accept	Reject
GARCH	Accept	Accept	Reject
GJR-GARCH	Accept	Accept	Reject
E-GARCH	Accept	Accept	Reject
July2018-Dec2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
Jan2019-June2019			
Normal	Reject	Accept	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2019-Dec2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject

Table 16: The test results for the Conditional, Unconditional and Quantile ES backtests simulated under the assumption of normal distribution for EOS/BTC.

XRP/BTC

It is much the same narrative for XRP as it is for the other cryptocurrencies with the GARCH family models passing the non simulated unconditional tests for each time period. E-GARCH was again had much less violations when compared to the other models, where it has consistently performed better in this aspect across all cryptocurrencies and time periods. When it comes to the simulated tests, interestingly the EWMA model was accepted for every time period under the t distribution assumption for the conditional test but this isn't the case under the stricter assumption of normal distribution.

It is clear to see from the backtesting results, that the GARCH family models are superior to the other models tested at modelling the volatility of hourly cryptocurrency returns and providing VaR and expected shortfall estimations. From these GARCH models, the E-GARCH model outperformed the rest, with passing the various back tests and providing less violations in the data than the other models across the 4 different cryptocurrencies and time periods.

	UnconditionalNorm	UnconditionalT
Jan2018-June2018		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept
July2018-Dec2018		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept
Jan2019-June2019		
Historical	Reject	Reject
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept
July2019-Dec2019		
Historical	Reject	Accept
Normal	Reject	Reject
EWMA	Reject	Reject
GARCH	Accept	Accept
GJR-GARCH	Accept	Accept
E-GARCH	Accept	Accept

Table 17: The results of the Unconditional Normal and Unconditional T tests for XRP/BTC in the four time periods we tested on.

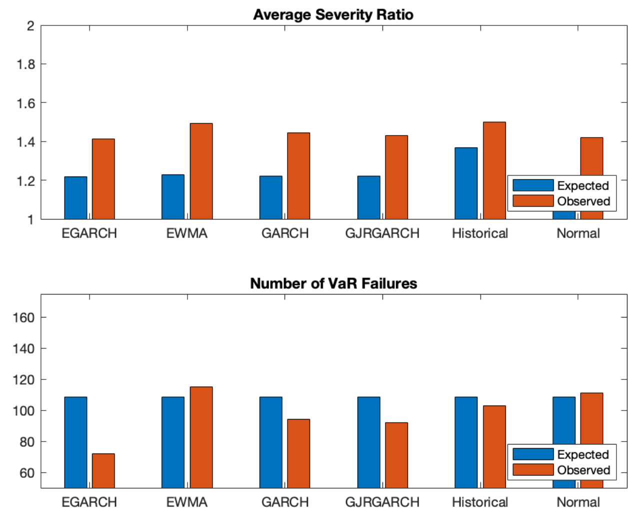

Figure 33: Histogram of the average severity ratio and number of VaR failures for Jan 2018 to June 2018 for XRP/BTC.

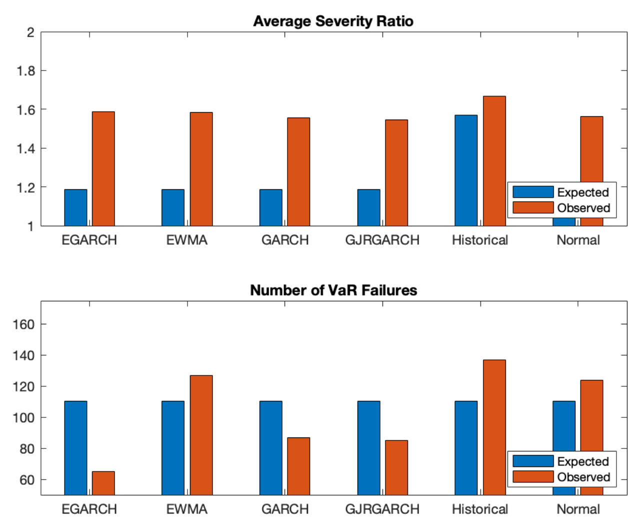

Figure 34: Histogram of the average severity ratio and number of VaR failures for July 2018 to Dec 2018 for XRP/BTC.

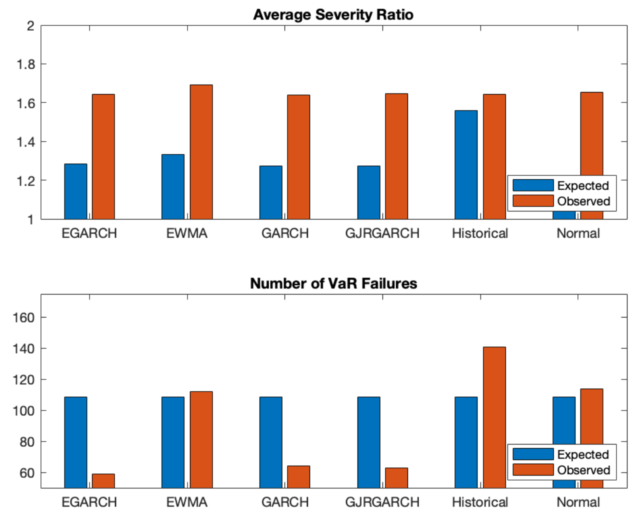

Figure 35: Histogram of the average severity ratio and number of VaR failures for Jan 2019 to Jun 2019 for XRP/BTC.

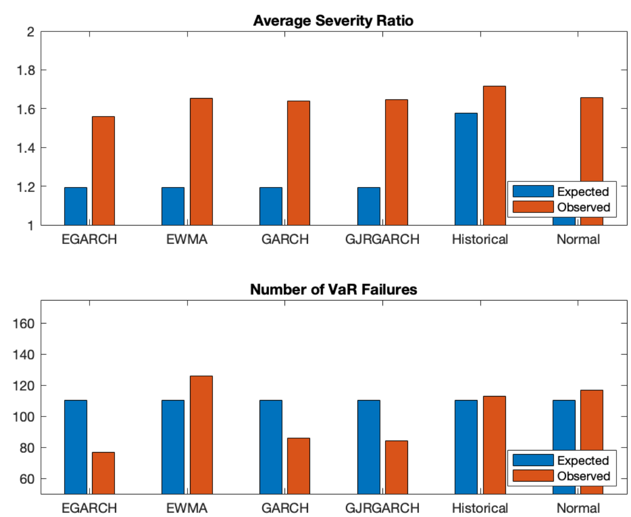

Figure 36: Histogram of the average severity ratio and number of VaR failures for July 2019 to Dec 2019 for XRP/BTC.

t Dist.	Conditional	Unconditional	Quantile
Jan2018-June2018			
Normal	Accept	Accept	Accept
EWMA	Accept	Accept	Accept
GARCH	Accept	Accept	Accept
GJR-GARCH	Accept	Accept	Accept
E-GARCH	Reject	Accept	Accept
July2018-Dec2018			
Normal	Reject	Reject	Reject
EWMA	Accept	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
Jan2019-June2019			
Normal	Reject	Reject	Reject
EWMA	Accept	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2019-Dec2019			
Normal	Reject	Reject	Reject
EWMA	Accept	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject

Table 18: The test results for the Conditional, Unconditional and Quantile ES backtests simulated under the assumption of t-distribution for XRP/BTC.

Normal Dist.	Conditional	Unconditional	Quantile
Jan2018-June2018			
Normal	Reject	Accept	Reject
EWMA	Accept	Accept	Reject
GARCH	Accept	Accept	Reject
GJR-GARCH	Accept	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2018-Dec2018			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
Jan2019-June2019			
Normal	Reject	Accept	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject
July2019-Dec2019			
Normal	Reject	Reject	Reject
EWMA	Reject	Accept	Reject
GARCH	Reject	Accept	Reject
GJR-GARCH	Reject	Accept	Reject
E-GARCH	Reject	Accept	Reject

Table 19: The test results for the Conditional, Unconditional and Quantile ES backtests simulated under the assumption of normal distribution for XRP/BTC.

6 Conclusions

We looked to extend the work of Katsiampa 2017 and Chu et al. 2017 with testing six different VaR and Expected Shortfall models (Historical Simulation, Normal Parametric, EWMA, GARCH(1,1), GJR-GARCH(1,1) and E-GARCH(1,1)) through using in-sample analysis on past returns data of four of the most liquid cryptocurrencies across a range of over two years. As much of cyrptocurrency trading is conducted 24/7 by speculative traders on cryptocurrency exchanges we chose to study hourly data with BTC as numeraire. Through statistical analysis on the data we found that the cryptocurrency prices are extremely volatile with heavy skew and kurtosis pointing to more extreme outcomes in the return which is in line with all previous literature into the area of cryptocurrencies.

Using a VAR confidence level of 99% (in line with the Basel accords for estimating value-at-risk values) and 95% (often the confidence level used by risk managers overseeing day trading activities) we conducted thorough backtesting on the models at these levels. Through the backtesting of the models we find that the E-GARCH model has the largest p-values for the tests of 99% VaR. The tests on 95% confidence VaR were more inconclusive with GJR-GARCH marginally outperforming the other GARCH models, but was outperformed by the simpler models such as EWMA and Historical Simulation when it came to the backtests. Despite this, the E-GARCH model consistently had the lowest VaR violations and lowest violation ratio across all cryptocurrencies.

For Expected Shortfall estimation, we used the 97.5% confidence level which is in line with the new regulation on capital risk management proposed. For testing the ES measurements from the models, we had to split the data up into 4 half yearly time periods ranging from the start of 2018 to the end of 2019. Over these time periods we found the GARCH family models to outperform the other models. Out of the GARCH models, the E-GARCH model once again has the lowest amount of violations, indicating it forecast the cryptocurrency volatility more accurately than the other models. Our findings are in line with previous literature, with being that GARCH models are superior at modelling VaR and ES estimates for cyrptocurrency data, but we find that the E-GARCH model is the most accurate out of the models we analysed. Although we omitted 5 other cryptocurrencies that make up the MVIS MVIS CryptoCompare Digital Assets 10 Index, our results would suggest that the E-GARCH model would too supply the most accurate results for the other cryptocurrency constituents.

The E-GARCH model is absent of non-linearity constraints unlike the other GARCH models, due to it's exponential term within the model. No matter the value of the estimated parameters, the volatility dynamics will return positive. This nature could be the reason why it outperforms the other models, as the constraints on the parameters could cause inaccuracies in highly dynamic data like cryptocurrencies.

This analysis could be extend further to use more complex GARCH models that contain regime switches that cope with the extreme volatility of cryptocurrencies which have been examined by Ardia, Bluteau, and Rüede 2019, and this would be a natural extension to further our work with these models.

Bibliography

[1] Carlo Acerbi and Balazs Szekely. "Backtesting Expected Shortfall". In: (2014). DOI: https://www.msci.com/documents/10199/22aa9922-f874-4060-b77a-0f0e267a489b.

[2] Carol Alexander. *Market risk analysis II: practical financial econometrics*. Chichester: John Wiley & Sons, Mar. 2008. URL: http://sro.sussex.ac.uk/id/eprint/40643/.

[3] D. Ardia, K. Bluteau, and M. Rüede. "Regime changes in Bitcoin GARCH volatility dynamics". In: *Finance Research Letters* 29 (2019), pp. 266–271. DOI: 10.1016/j.frl.2018.08.009.

[4] C. Baek and M. Elbeck. "Bitcoins as an investment or speculative vehicle? A first look". In: *Applied Economics Letters* 22.1 (2015), pp. 30–34. DOI: 10.1080/13504851.2014.916379. eprint: https://doi.org/10.1080/13504851.2014.916379. URL: https://doi.org/10.1080/13504851.2014.916379.

[5] Benjamin M. Blau. "Price dynamics and speculative trading in bitcoin". In: *Research in International Business and Finance* 41 (2017), pp. 493–499. DOI: https://doi.org/10.1016/j.ribaf.2017.05.010.

[6] M. Buchholz et al. "Bits and Bets, Information, Price Volatility, and Demand for BitCoin". In: *Economics* 312 (2012). DOI: https://www.reed.edu/economics/parker/s12/312/finalproj/Bitcoin.pdf.

[7] Vitalik Buterin. "A Next-Generation Smart Contract and Decentralized Application Platform". In: (2013). DOI: https://ethereum.org/en/whitepaper/.

[8] Adrian (Wai-Kong) Cheung, Eduardo Roca, and Jen-Je Su. "Cryptocurrency bubbles: an application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox bitcoin prices". In: *Applied Economics* 47.23 (2015), pp. 2348–2358. DOI: 10.1080/00036846.2015.1005827. eprint: https://doi.org/10.1080/00036846.2015.1005827. URL: https://doi.org/10.1080/00036846.2015.1005827.

[9] Peter F. Christoffersen. "Evaluating Interval Forecasts". In: *International Economic Review* 39.4 (1998), pp. 841–862. ISSN: 00206598, 14682354. URL: http://www.jstor.org/stable/2527341.

[10] J. Chu et al. "GARCH Modelling of Cryptocurrencies". In: *J. Risk Financial Manag* 10.17 (2017). DOI: https://doi.org/10.3390/jrfm10040017.

[11] LAWRENCE R. Glosten, RAVI Jagannathan, and DAVID E. Runkle. "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks". In: *The Journal of Finance* 48.5 (1993), pp. 1779–1801. DOI: 10.1111/j.1540-6261.1993.tb05128.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1993.tb05128.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1993.tb05128.x.

[12] Marcus Haas. "New Methods in Backtesting". In: *Financial Engineering, Research Center Caesar, Bonn.* (2001). DOI: https://www.ime.usp.br/~rvicente/risco/haas.pdf.

[13] J.P.Morgan. *RiskMetrics: Technical Document.* J. P. Morgan, 1996. URL: https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95aJ.

[14] P. Katsiampa. "Volatility estimation for Bitcoin: A comparison of GARCH models". In: *Economics Letters* 158 (2017), pp. 3–6. DOI: 10.1016/j.econlet.2017.06.023.

[15] Frode Kjaerland et al. "An Analysis of Bitcoin's Price Dynamics". In: *Journal of Risk and Financial Management* 11.4 (2018), p. 63. DOI: 10.3390/jrfm110400634.

[16] I. Kopytin, A. Maslennikov, and S. Zhukov. "What Value-at-Risk and Expected Shortfall Metrics Tell a Risk Averse Investor in Cryptocurrencies". In: *Smart Innovation, Systems and Technologies* 139 (2019), pp. 398–410. DOI: 10.1007/978-3-030-18553-4_50. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066911315&doi=10.1007%2f978-3-030-18553-4_50&partnerID=40&md5=4c53e714140c551883fab1bf0e9a9410.

[17] Paul H. Kupiec. "Techniques for Verifying the Accuracy of Risk Measurement Models". In: *The Journal of Derivatives* 3.2 (1995), pp. 73–84. ISSN: 1074-1240. DOI: 10.3905/jod.1995.407942. eprint: https://jod.pm-research.com/content/3/2/73.full.pdf. URL: https://jod.pm-research.com/content/3/2/73.

[18] Daniel Larimer. "EOS.IO Technical White Paper". In: (2017). DOI: https://eoscollective.org/papers.

[19] Satoshi Nakamoto. "Bitcoin: A Peer-to-Peer Electronic Cash System". In: (2008). DOI: https://bitcoin.org/bitcoin.pdf.

[20] Daniel B. Nelson. "Conditional Heteroskedasticity in Asset Returns: A New Approach". In: *Econometrica* 59.2 (1991), pp. 347–370. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/2938260.

[21] Peter Phillips, Shuping Shi, and Jun Yu. "Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the SP 500". In: *Cowles Foundation Discussion Paper* 1914 (2013). DOI: http://dx.doi.org/10.2139/ssrn.2327609.

[22] David Schwartz, Noah Youngs, and Arthur Britto. "The Ripple Protocol Consensus Algorithm". In: (2014). DOI: https://ripple.com/files/ripple_consensus_whitepaper.pdf.

[23] S. Stavroyiannis. "Value-at-Risk and Expected Shortfall for the Major Digital Currencies". In: *SSRN Electronic Journal* (Aug. 2017). DOI: 10.2139/ssrn.3028625.

[24] S. Stavroyiannis. "Value-at-risk and related measures for the Bitcoin". In: *Journal of Risk Finance* 19.2 (2018), pp. 127–136. DOI: 10.1108/JRF-07-2017-0115.

[25] C. Trucios. "Forecasting Bitcoin risk measures: A robust approach". In: *International Journal of Forecasting* 35.3 (2019), pp. 836–847. DOI: 10.1016/j.ijforecast.2019.01.003.

[26] U. Uyar and I.K. Kahraman. "The risk analysis of Bitcoin and major currencies: value at risk approach". In: *Journal of Money Laundering Control* 22.1 (2019), pp. 38–52. DOI: 10.1108/JMLC-01-2018-0005. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85064161466&doi=10.1108%2fJMLC-01-2018-0005&partnerID=40&md5=448346cf8a875b72bf4ee99c9bd5bad1.

Appendix

	alpha	beta	omega	gamma
ETH/BTC				
GARCH	0.08674	0.90843	4.8207e-07	N/A
GJR-GARCH	0.094435	0.87717	7.7629e-07	0.003395
E-GARCH	0.19602	0.97892	-0.19277	-0.013276
LTC/ TC				
GARCH	0.098913	0.8888	1.1321e-06	N/A
GJR-GARCH	0.098065	0.88945	1.1264e-06	-4.8431e-05
E-GARCH	0.19391	0.96842	-0.29173	0.010015
EOS/BTC				
GARCH	0.095608	0.89759	1.4018e-06	N/A
GJR-GARCH	0.097082	0.89906	1.3676e-06	-0.0054102
E-GARCH	0.20553	0.9707	-0.24907	0.0082805
XRP/BTC				
GARCH	0.12959	0.87041	1.2933e-06	N/A
GJR-GARCH	0.092714	0.87197	1.2906e-06	0.070623
E-GARCH	0.24181	0.98558	-0.093562	-0.02182

Table 20: GARCH model parameters from the estimation using the cyrptocurrency data.

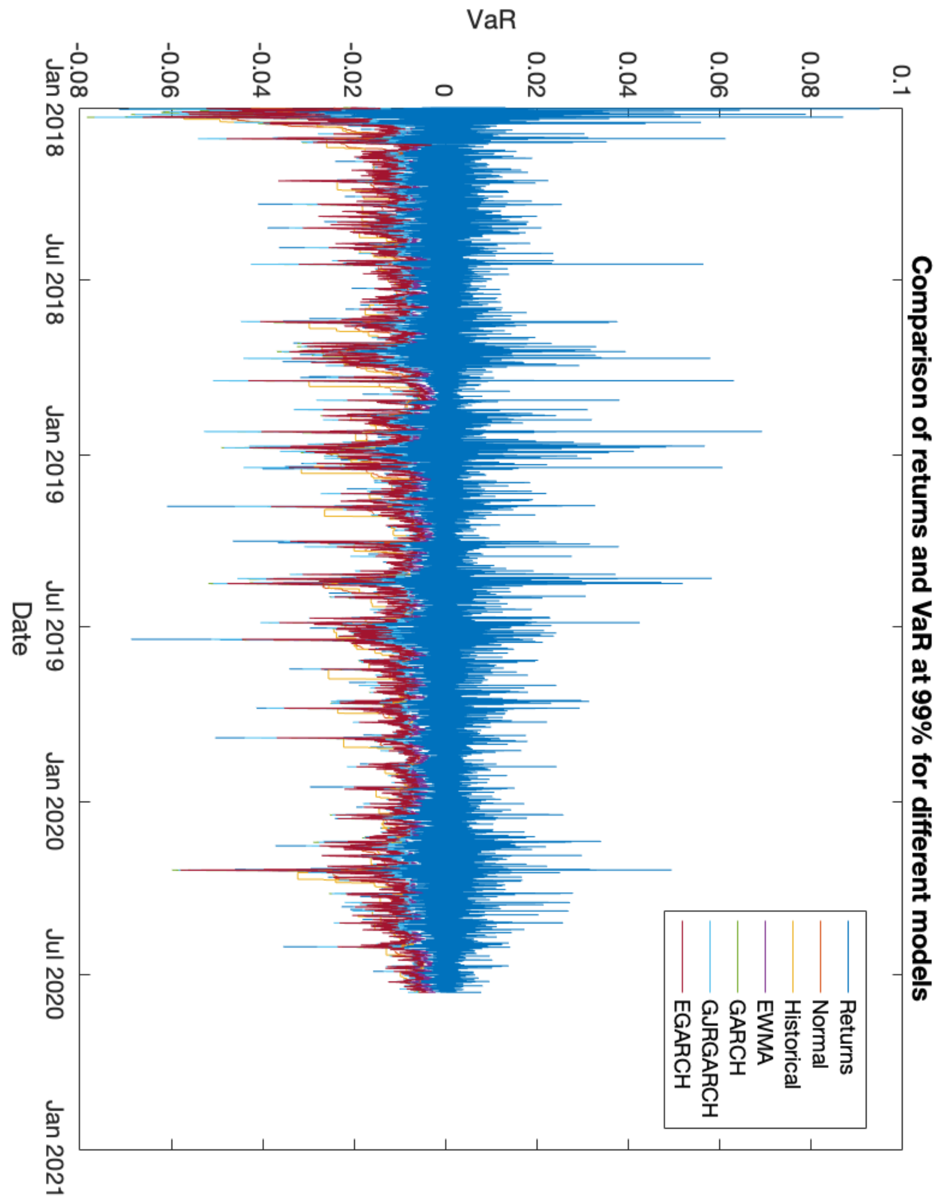

Figure 37: Comparison of VaR models to returns for ETH/BTC

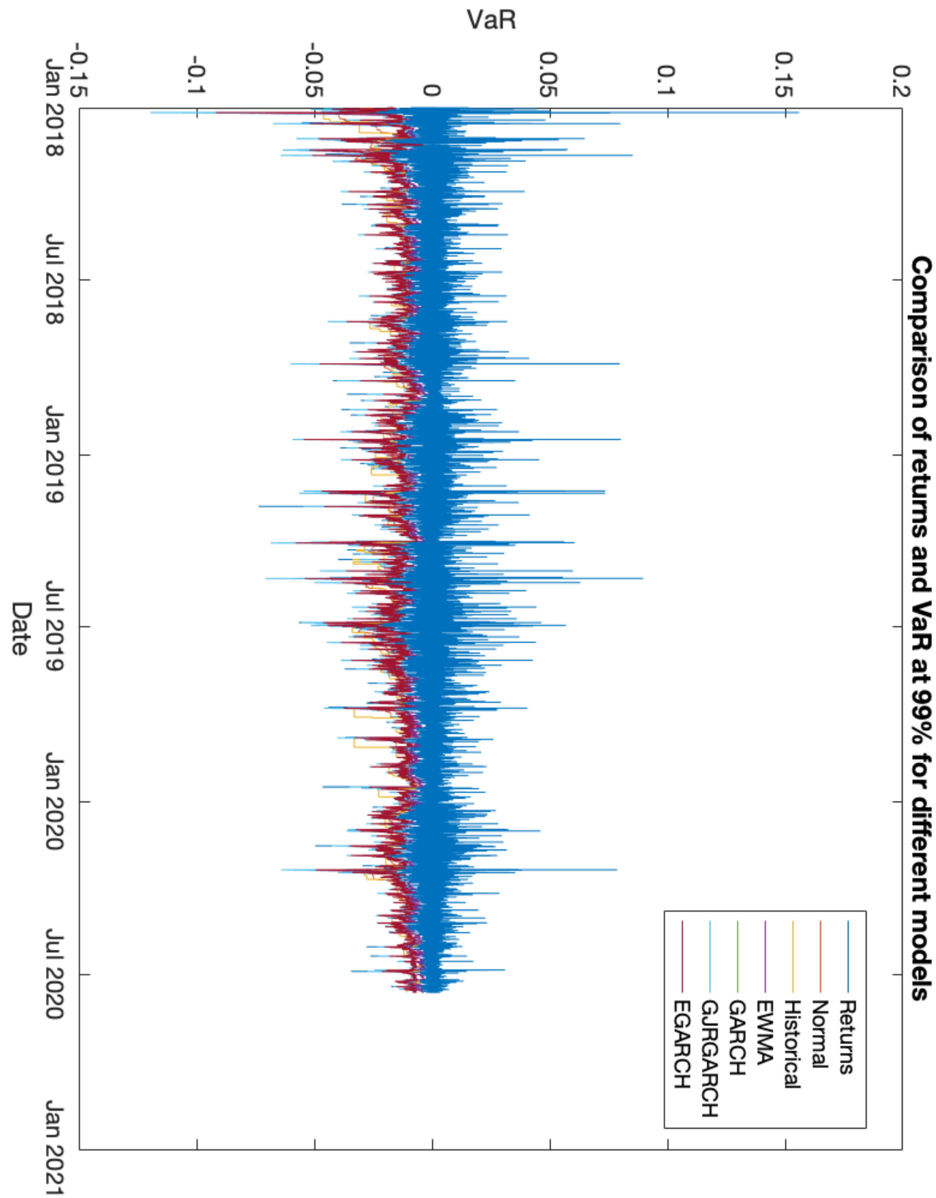

Figure 38: Comparison of VaR models to returns for LTC/BTC

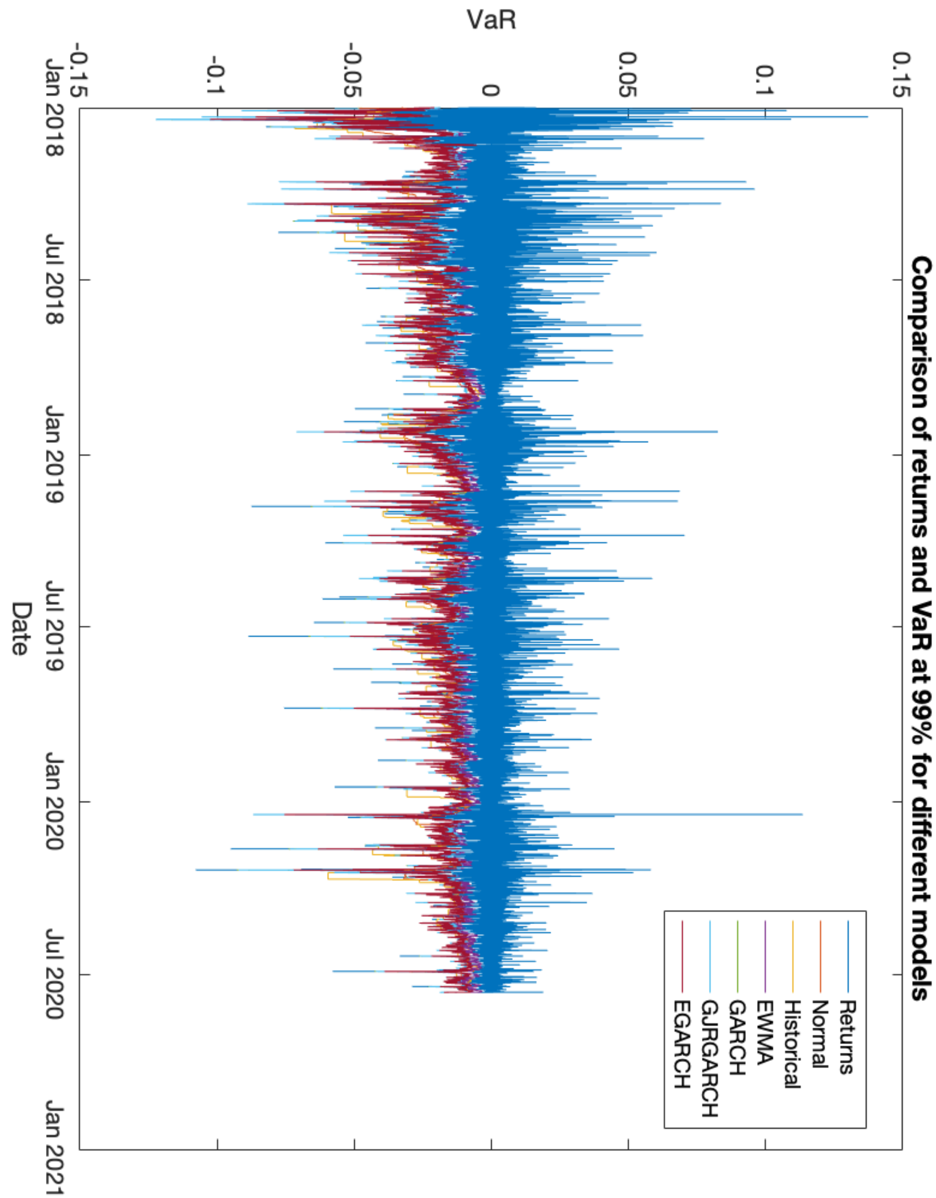

Figure 39: Comparison of VaR models to returns for EOS/BTC

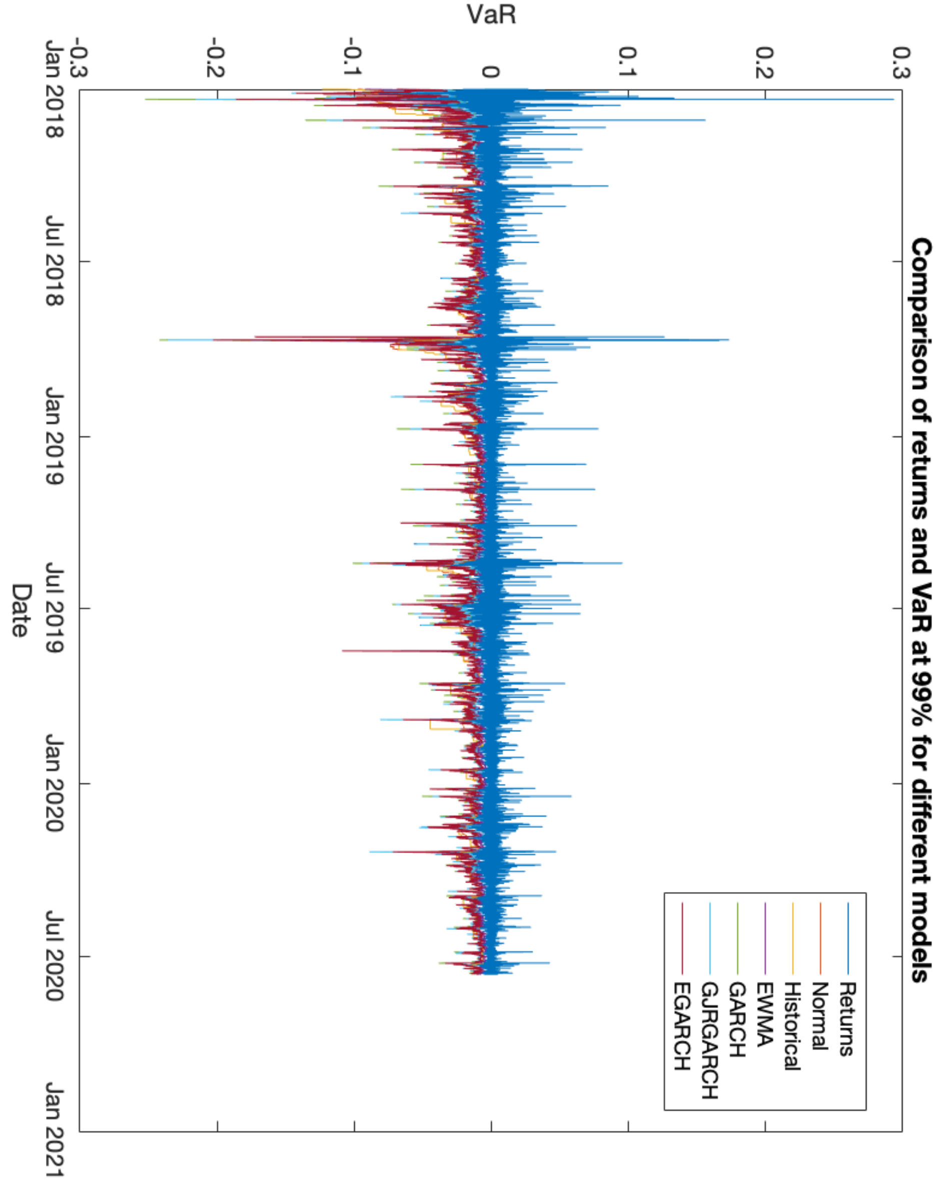

Figure 40: Comparison of VaR models to returns for XRP/BTC