Volatility forecasting for cryptocurrencies under a heavy-tailed distribution

by

Alina Bylkova

Diego Mauricio Vargas Pico

NEKN02 Master Essay – Finance Programme

Supervisor: Associate Professor Anders Vilhelmsson

Lund University School of Economics and Management

Lund, Sweden

June 2019

VOLATILITY FORECASTING FOR CRYPTOCURRENCIES UNDER A HEAVY-TAILED DISTRIBUTION

Alina Bylkova
Diego Mauricio Vargas Pico[1]

Abstract

In the recent years, cryptocurrencies have gained popularity and have experienced high price volatility. This essay pretends to examine how the multivariate GARCH models predict the volatility of these digital currencies and what implications exist if we consider the correlations among them to forecast their volatility. Using a bivariate Diagonal VECH and bivariate Diagonal BEKK this thesis checks the covariance on the returns of the two largest cryptocurrencies in terms of market capitalization, Bitcoin and Ethereum. Since the price returns of financial assets tend to have a non-normal behavior, we consider a distribution that allows the data to have heavier tails, assuming that this could be more realistic as it is observed in other financial markets. After employing the different models, we find that the best-fit distribution for estimating the conditional (co)variance is the Student's t, which allows the data to have fatter tails. We find that GARCH(1,1) is the best specification for the conditional (co)variance of Bitcoin and Ethereum and when performing the forecast we find that the best model is estimated with the Diagonal VECH approach.

Keywords: Cryptocurrency, Bivariate Diagonal BEKK, Bivariate Diagonal VECH, MGARCH, Volatility

[1] The authors would like to thank the orientation, valuable comments and time dedicated by their supervisor, **Anders Vilhelmsson**

Contents

List of Tables

List of Figures

Terminology

ACGARCH Asymmetric Component Generalized Autoregressive Conditional Heteroscedasticity

AIC Akaike's Information Criteria

APARCH Asymmetric Power Autoregressive Conditional Heteroscedasticity

AR Autoregressive model

BEKK The acronym is based on authors' last names Baba, Engle, Kraft and Kroner

BTC Bitcoin

CGARCH Component Generalized Autoregressive Conditional Heteroscedasticity

EGARCH Exponential Generalized Autoregressive Conditional Heteroscedasticity

ETH Ethereum

GARCH Generalized Autoregressive Conditional Heteroscedasticity

HMAE Heteroscedasticity-adjusted Mean absolute error

HQIC Hannan-Quinn's Information Criteria

IC Information Criteria

MAE Mean absolute error

MSE Mean squared error

RMSE Root-mean squared error

SBIC Schwarz Bayesian's Information Criteria

TGARCH Threshold Generalized Autoregressive Conditional Heteroscedasticity

VECH Half Vectorization

1. Introduction

During the recent years, cryptocurrencies and blockchain have gained great popularity worldwide, and the number of cryptocurrencies has increased accordingly. As of April 2019, there were over 2000 different ones, with a rising trend and a total market capitalization of 175 billion US dollars (CoinMarketCap, 2019). Nonetheless, it is a relatively new instrument in the financial markets, since its creation dates back to 2008 with the paper "Bitcoin: A peer-to-peer Electronic Cash System" published by Satoshi Nakamoto (Antonopoulos, 2014), and has had an exponential growth during its 10 years of existence. From the vast majority of current digital currencies, two of them (Bitcoin and Ethereum) account for 62% of the total market capitalization (CoinMarketCap, 2019), implying that it is a very concentrated market despite of the large amount of different assets. Moreover, the prices of these two digital currencies rose rapidly and fluctuated significantly in recent years and, therefore, many cryptocurrency investors increased their investments in Bitcoin and Ethereum (Katsiampa, 2018b). That is the reason why we focus on these both cryptocurrencies for our research. The paper could be easily extended by including other cryptocurrencies, but we consider that investigating these two is the best choice, since the following in market capitalization have a small market share and will not make a big contribution to the main results.

As it is well known, variance (volatility), just after the expected return, is considered the second most important statistical property of financial securities for portfolio construction, risk management and pricing. Hence, it becomes important to estimate this value as it is not possible to observe it given its latent variable nature. For this purpose, and considering that it is a security traded in the financial markets and, therefore, it is subject to volatility clustering, the estimation is conducted using Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models. Specifically, we use multivariate GARCH models such as bivariate Diagonal VECH and bivariate Diagonal BEKK models to estimate the variance of the cryptocurrencies mentioned above, and to check the covariance between them and predict future returns. In our study, we opted for diagonal type models over ordinary VECH and BEKK models because the number of parameters to be estimated is lower, and the results obtained are very similar. It is important to remark that in this paper we allow the returns to present fatter tails in their distribution and investigate whether this well-known feature of "common" financial instruments applies to cryptocurrencies as well. Katsiampa (2018b) addresses this topic by estimating Diagonal BEKK models under normal and Student's t distribution, but does not employ Diagonal VECH parametrization nor utilizes loss functions to decide which model to use; the decision relies upon the information given by the Information Criteria.

The main purpose of the present study is to investigate and find the most accurate model for forecasting the variance of digital currencies using diagonal VECH and diagonal BEKK models and their performance for Bitcoin and Ethereum accounting for the correlation between them. Comprehending the interrelation of these two cryptocurrencies and forecasting their volatility is essential for suitable investment strategy developing and portfolio construction purposes as mentioned before. This research can help cryptocurrency investors to become more aware of their decisions and at the same time adds to the existing hitherto scarce literature in the area.

The structure of the thesis is organized as follows: in section 2 we review the literature that has been written about cryptocurrencies, different implied models and works that were made considering their volatility. In section 3 we provide a short description of cryptocurrencies. In section 4 we describe the data used for this thesis. In section 5, methodology and statistical models to be considered are presented. In section 6, forecast evaluation procedure is described. In section 7, the main results of the thesis are presented after performing the corresponding models on the data and finding the best Multivariate-GARCH model for the sample. Finally, in section 8, the main conclusions of our research are presented.

2. Literature review

Recently, cryptocurrencies have become a popular topic in academic research. For example, Ciaian, Rajcaniova and Kancs, d'Artis (2016) investigated the Bitcoin price formation process and the determinants of its returns and found evidence that supply and demand of Bitcoin have considerable impact on its price and it can be explained in a standard economic model. Some other researchers studied the correlation between the prices of digital currency and trading volumes, for instance, Blau (2017) found evidence that the level of speculative trading is not directly related to the level of Bitcoin volatility, whereas Balcilar, Roubaud, Gupta and Bouri (2017) showed that the Bitcoin price returns and their relationship with trading volume develop in a nonlinear way. Also, the linkages of cryptocurrencies to mainstream assets was covered in a more recent study of Corbet, Meegan, Larkin, Lucey and Yarovaya (2018) where the authors found a very low correlation between them and therefore argued that cryptocurrencies establish a new investment asset class. Other researchers as Phillip, Chan and Peiris (2018) used the stochastic volatility model for 224 different digital currencies in order to explore their properties; they revealed that cryptocurrencies have specific characteristics such as leverage effect, volatility clustering, fat tails and long memory.

We have seen many recommendations in different papers how to model the cryptocurrencies price returns volatility and the GARCH model which was introduced by Bollerslev (1986) is the most widely used one. In this model the conditional variance in the current period depends on the squared errors and conditional variances of the previous period. GARCH models, unlike linear estimation models, assume that the variance of the returns is not constant over time, as returns of the assets show periods of weak and strong fluctuations on the financial market, i.e. volatility clustering (which is common for cryptocurrencies data). Further, Bollerslev, Engle and Wooldridge (1988) proposed a multivariate GARCH model named Diagonal VECH considering a possible correlation between several assets. In order to reduce the number of parameters in the model, they assumed that both parameter matrices A and B should be diagonal, which means that every element of conditional covariance matrix depends only on its own lag and on the previous value of error term (Bollerslev, Engle &Wooldridge, 1988). However, this model requires some restrictions to ensure the positivity of conditional covariance matrix, which is quite complicated in practice. Therefore, Engle and Kroner (1995) presented another type of multivariate GARCH model named Diagonal BEKK in which they ensured the positive definiteness of conditional covariance matrix by proposing a new parametrization of covariance matrix without any restrictions. These two models are discussed in detail in section 5.

Many authors employed different GARCH-type models in order to investigate the cryptocurrencies volatility. For instance, Katsiampa (2017) examined different conditional heteroscedasticity models performance to estimate Bitcoin price volatility. She employed GARCH, EGARCH, TGARCH, APARCH, CGARCH and ACGARCH models and based on information criteria she found evidence that the Component GARCH model is the optimal one. Chu, Chan, Nadarajah and Osterrieder (2017) investigated the price returns volatility of seven digital currencies such as Bitcoin, Dash, Dogecoin, Litecoin, Maidsafecoin, Monero and Ripple and by employing twelve GARCH-type models they concluded that IGARCH and GJRGARCH are the best fitting for volatility modelling purposes. However, aforementioned authors used only univariate GARCH models that overlook possible relationships over a set of cryptocurrencies (cross-sectional correlation) and only considers one variable. Therefore, multivariate GARCH models are more appropriate for estimating cryptocurrencies price returns volatility accounting for possible correlation between different digital currencies.

However, the literature on volatility dynamics and forecasting of cryptocurrencies using multivariate GARCH models remains limited. The most important related literature consists of the papers of Hultman (2018) and Katsiampa (2018a; 2018b). More specifically, Katsiampa (2018b) explored the volatility dynamics of Bitcoin and Ethereum over the period from August 2015 to January 2018 and by employing a bivariate Diagonal BEKK model, she revealed the dependence between two digital currencies and concluded that Ethereum is an effective hedging instrument against Bitcoin. In another study, Katsiampa (2018a) investigated the volatility dynamics of five different cryptocurrencies taking into account asymmetric effects of good and bad news in the conditional volatility of digital currencies. She used an asymmetric Diagonal BEKK model for the returns, covering the period from August 2015 to February 2018 and found evidence that asymmetric effects of positive and negative shocks have a considerable impact on conditional variance of Bitcoin, Ethereum, Ripple and Litecoin, whereas conditional volatility of Stellar Lumen does not capture asymmetric past shocks (Katsiampa, 2018a). Hultman (2018) investigated the effectiveness of the bivariate BEKK(1,1), GARCH(1,1) and the Standard stochastic volatility models to predict Bitcoin price returns volatility and concluded that GARCH(1,1) produces the most accurate forecast. In his study, he used the data of daily closing prices for Bitcoin and Ethereum (in case of bivariate BEKK) during the period from August 2015 to June 2018.

We intend to extend the existing literature not only by considering a larger dataset, including one more year (until April 2019 capturing the price growth and decline experienced during 2017-2018), but also by employing two multivariate GARCH models (such as a bivariate Diagonal VECH and bivariate Diagonal BEKK) for forecasting and accounting for possible correlation between Bitcoin and Ethereum. Moreover, we aim to estimate the models under multivariate normal and multivariate Student's t distributions to examine which distribution fits better to the price returns of cryptocurrencies.

3. Cryptocurrencies description

Nowadays, there are 2165 digital currencies present in 17747 markets, however, Bitcoin and Ethereum are the largest ones and together represent 62.3% of total market capitalization (CoinMarketCap, 2019). Besides, prices of both cryptocurrencies have experienced a rapid growth and decline during recent years, and therefore many investors started investing actively in Bitcoin and Ethereum (Katsiampa, 2018b). Thus, we pick these two cryptocurrencies for our research.

3.1 Bitcoin

Bitcoin was created in 2008 by an unknown individual or group of people under the pseudonym Satoshi Nakamoto as a peer-to-peer electronic cash system which was absolutely decentralized and free of any authority (Antonopoulos, 2014). Bitcoin is based on cryptography and uses encryption and digital signatures in order to make it protected and secure (Antonopoulos, 2014). The main difference between Bitcoin and traditional currencies is that Bitcoin is an entirely digital currency. It only exists virtually and does not have any physical equivalents like coins or banknotes.

Although current cryptocurrency markets are full of different analogues, Bitcoin is still the most demanded digital currency in terms of market capitalization. It is estimated at 89.5 billion US dollars which is 52.1% of the total market capitalization (CoinMarketCap, 2019). The current number of coins in existence available to public is about 17.6 million, with a fixed total limit in circulation of 21 million bitcoins (CoinMarketCap, 2019).

Trading operations using Bitcoin as a payment rapidly grew in recent years due to increasing public and media interest (Katsiampa, 2018a). Bitcoin can be used for different purposes like selling and buying goods and services, transferring money to people or companies or exchanging for other currencies on specialized markets (Hultman, 2018). On the other hand, Bitcoin users could also purchase illegal goods (Baur, Hong & Lee, 2018). According to Foley, Karlsen and Putniņš (2019) around 26% of all Bitcoin users and 46% of Bitcoin transactions are connected with prohibited activities. Despite the opportunity to finance illegal activities or use it just for regular trading purposes, Bitcoin is mainly used for speculation and investment purposes (Katsiampa, 2017; Dyhrberg, 2016; Baek & Elbeck, 2015; Blau, 2017).

Besides, there is no common opinion among researchers regarding the position of Bitcoin on the market, whether it is an asset or a currency. Dyhrberg (2016) argues that Bitcoin is very similar to both the US dollar and gold, therefore it can be classified somewhere in between a currency and a commodity. However, Corbet et al. (2018) investigated the connectedness of cryptocurrencies to traditional assets and concluded that there is very low correlation between them and therefore they argued that cryptocurrencies establish new investment asset class. According to Katsiampa (2018a) Bitcoin has more similarities with financial assets because of its specific properties such as volatility (Katsiampa, 2017; Chu et al., 2017), susceptibility to speculative bubbles (Cheah & Fry, 2015), heavy-tailed distribution (Chan, Chu, Nadarajah & Osterrieder, 2017; Gkillas & Katsiampa, 2018; Phillip, Chan & Peiris, 2018) and leverage effects (Phillip, Chan & Peiris, 2018). More descriptions of Bitcoin can be found in papers of Böhme, Christin, Edelman and Moore (2015), Dwyer (2015) and others.

3.2 Ethereum

Ethereum is a quite new digital currency, which was launched in July 2015, and as well as Bitcoin it is based on Blockchain platform (Katsiampa, 2018a). Despite the fact that Ethereum was released a few years ago, it has already become the second largest digital coin. It is estimated currently at 17.3 billion US dollars which represents 10.2% of the total market capitalization (CoinMarketCap, 2019). There are more than 105.6 million coins in circulation which is approximately six times higher than the circulating supply of Bitcoin (CoinMarketCap, 2019). Moreover, large industrial flagships recently supported and promoted Ethereum through the Enterprise Ethereum Alliance foundation (Katsiampa, 2018a).

4. Data

The dataset for this research consists of daily closing prices in US dollars for Bitcoin and Ethereum over the period from 7th August 2015 (as the starting date of trading Ethereum) to 8th April 2019. Summarizing, the sample consists of 1341 observations for each cryptocurrency. We split our sample in two intervals as shown in Figure 1. The first interval from 7th August 2015 to 8th April 2018 is used for in-sample estimation. The second interval from 9th April 2018 to 8th April 2019 (365 days) is used for out-of-sample forecast evaluation. The data are available online at the website CoinMarketCap (2019).

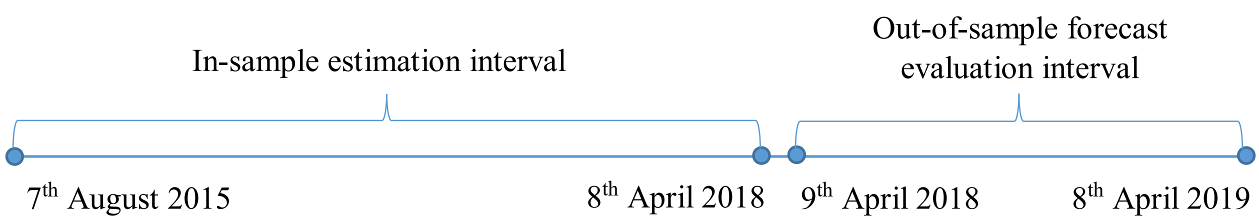

Figure 1. In-sample and out-of-samples intervals. Source: Brooks (2014)

Figure 2 shows the Bitcoin price evolution during the period of study. We can see that the prices of Bitcoin rose slowly until the beginning of 2017. Starting from the second quarter of 2017 there was a rapid growth, which reached the maximum point of 19497.40 US dollars in December 2017. This fact attracted public attention and increased investment and speculation. In 2018, the prices started to decline significantly and nowadays vary around 3000-5000 US dollars.

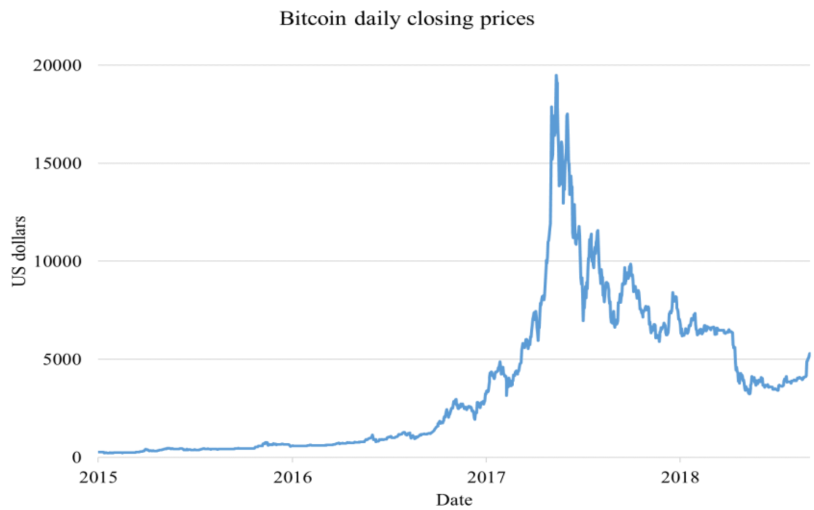

Figure 2. Bitcoin daily closing prices (in US Dollars) 07/08/2015 – 08/04/2019

Figure 3 represents the Ethereum price evolution over the period from 7th August 2015 to 8th April 2019. It can be seen that the price of Ethereum were stable until the beginning of 2017. After that, they gradually increased over the period from the second quarter of 2017 until the end of 2017. Then, from the beginning of 2018 there was a considerable price decrease.

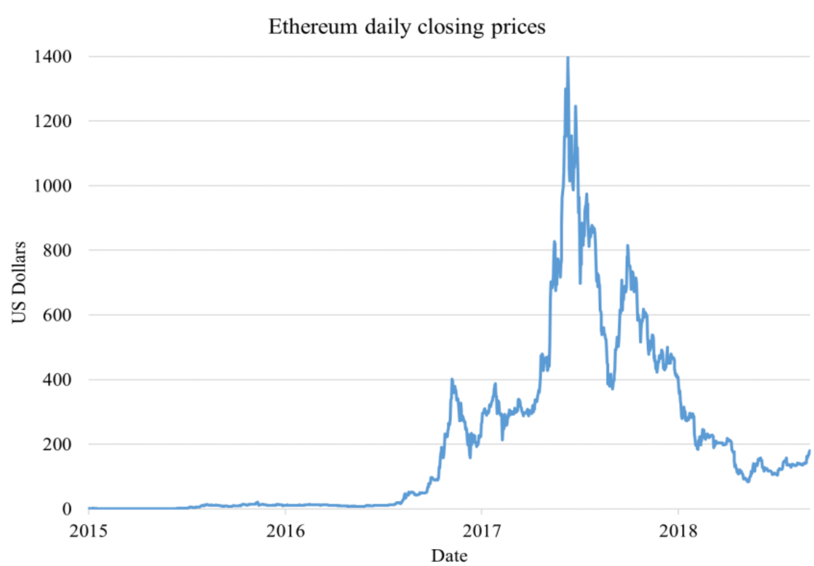

Figure 3. Ethereum daily closing prices (in US Dollars) 07/08/2015 – 08/04/2019

To sum up, the prices of both cryptocurrencies tend to have the same behavior and therefore could be correlated. This assumption is confirmed by the positive Pearson correlation coefficient with a value of 0.8998.

In order to make the data appropriate for the further analysis we make them stationary by transforming prices of cryptocurrencies into returns. The returns of Bitcoin $(i = 1)$ and Ethereum $(i = 2)$ are calculated on daily basis as follows:

$$r_{i,t} = lnP_{i,t} - lnP_{i,t-1} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right), \quad (1)$$

where $r_{i,t}$ is the logarithmic price return of cryptocurrency i at time t, $P_{i,t}$ is the observed price of cryptocurrency i at time t and $P_{i,t-1}$ is the observed price of cryptocurrency i at time $t-1$.

Table 1 illustrates the descriptive statistics and the unit root tests for the price returns of the two cryptocurrencies. The average daily price returns are positive (0.22% and 0.31%) for both Bitcoin and Ethereum respectively with a standard deviation of 3.93% and 7.60%.

Table 1. Descriptive statistics and unit root tests for the Bitcoin and Ethereum daily price returns

Panel I: Descriptive statistics		
N = 1340	Bitcoin	Ethereum
Mean	0.0022	0.0031
Median	0.0023	-0.0009
Maximum	0.2251	0.4123
Minimum	-0.2075	-1.3021
Std. Dev.	0.0393	0.0760
Skewness	-0.2240	-3.4058
Kurtosis	7.9887	69.6112
Jarque-Bera	1400.721***	250325.7***
Panel II: Unit root test statistics		
Augmented Dickey-Fuller	-36.4210***	-39.1649***
Phillips-Perron	-36.4399***	-38.8461***

*** *1% level of significance*

Note: This table in Panel I represents the descriptive statistics for the Bitcoin and Ethereum daily price returns and Jarque-Bera test with the null hypothesis that returns are distributed normally. Panel II shows the results of Augmented Dickey-Fuller and Phillips-Perron tests with null hypothesis of non-stationary data. Both panels cover sample period from 7th August 2015 to 8th April 2019.

The mean and median of Bitcoin price returns are close from each other (0.22% and 0.23% respectively). This may suggest that the data have symmetrical distribution, which is consistent with the histogram of Bitcoin returns presented in Figure 4. The range is small. Skewness of Bitcoin is negative and equals to -0.22 which is close to zero, but kurtosis is higher than 3 and equals to 7.99. Therefore, we can assume that the distribution of Bitcoin price returns is not normal and has heavy tails. This assumption is proven by the Jarque-Bera test, since p-value equals to 0.00 which is lower than any significance level. It means that the null hypothesis that skewness = 0 and kurtosis = 3 is rejected.

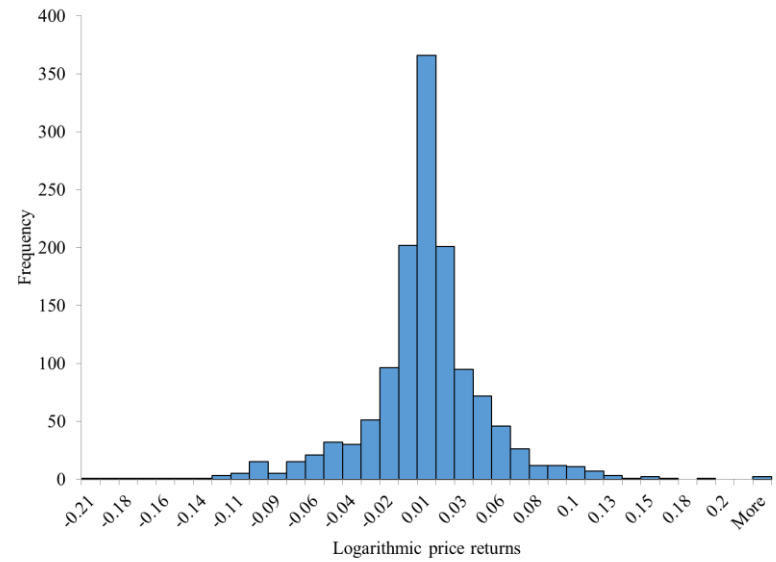

Figure 4. Bitcoin logarithmic price returns distribution histogram

Ethereum has a mean and median that are not close to each other (0.31% and -0.09% respectively) which suggest that the data have asymmetrical distribution. Besides, there is a big difference between the maximum and minimum value. The distribution of Ethereum price returns is right skewed, which is indicated by its skewness = -3.41 < 0. Moreover, the distribution is leptokurtic since kurtosis = 69.61 > 3. We can guess that Ethereum price returns are not normally distributed and have high kurtosis, which is confirmed by the results of the Jarque-Bera test – null hypothesis is rejected as p-value is 0%, lower than any significance level. This is also consistent with Ethereum histogram showed in figure 5.

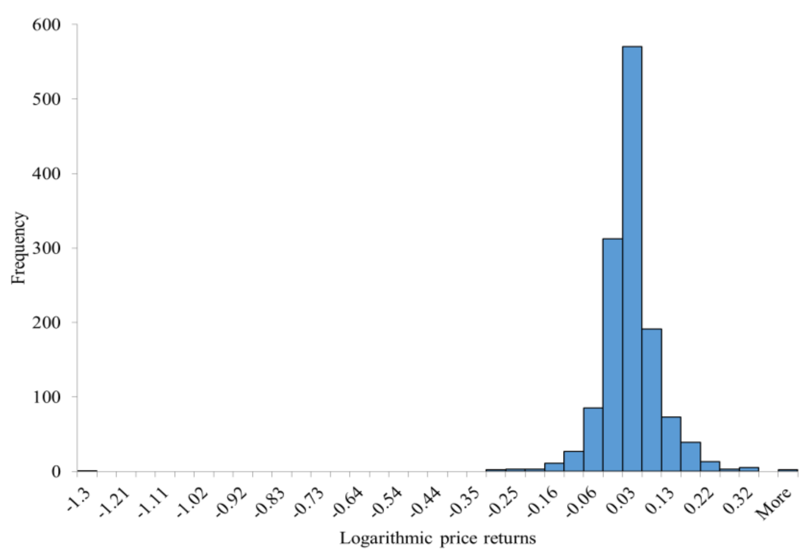

Figure 5. Ethereum logarithmic price returns distribution histogram

To conclude, the price returns of the two cryptocurrencies are leptokurtic because of significant excess kurtosis. However, Bitcoin has smaller kurtosis than Etherium. According to negatively skewed price returns of both digital currencies, we can make an inference that it is more likely to observe large negative returns. Moreover, the Jarque-Bera test results proved the deviation from normality. The results of the Augmented Dickey-Fuller and Phillips-Perron unit root tests proved that the price returns of both cryptocurrencies are stationary. Thus, the data are appropriate for further analysis.

5. Methodology

The starting point for our analysis is characterization of the price returns according to their descriptive statistics. This step was described in details in section 4. Then, in order to check the data for stationarity we perform Augmented Dickey-Fuller and Phillips-Perron unit root test. Thus, we test price returns of both cryptocurrencies under the null hypothesis of the unit root opposite to alternative hypothesis of stationary data. According to obtained results reported in the end of section 4, we can see that the Bitcoin and Ethereum price returns are stationary.

Hence, as the main research methods, we employ multivariate GARCH models such as bivariate Diagonal VECH and bivariate Diagonal BEKK models to estimate the volatility of Bitcoin and Ethereum price returns accounting for correlation between them. We prefer diagonal type of models instead of the ordinary VECH and BEKK, because the number of parameters to be estimated is lower in diagonal models, this produces better results in the numerical optimization. For example, already in the bivariate case we face problems regarding the optimization because there were several local optimums, depending on the starting point of the iteration. It becomes more severe the more parameters we have. Among other reasons, diagonal models are preferred due to the number of parameters.

Both, diagonal VECH and diagonal BEKK models are regarded as two of the most used models for volatility estimations. The main difference between them is that diagonal BEKK ensures the positivity of conditional covariance matrix by construction, without any restriction, whereas diagonal VECH requires imposing some limitations to parameters.

In addition, we estimate these models under both the multivariate normal and multivariate Student's t distributions in order to investigate whether models under a distribution that accounts for heavy-tailed behavior fits better to the price returns of cryptocurrencies than the models under the normal distribution. Additionally, we try two different equations for the conditional mean, the standard constant equation and other which follows an AR(1) process, as a first attempt to dig further into this “new” financial instrument.

Thus, we have two conditional mean equations:

$$r_{i,t} = \mu_i + \eta_{i,t}, i = 1,2, \eta_{i,t}|\Omega_{t-1} \sim D(0, \mathrm{H}_t) \quad (2)$$

$$r_{i,t} = \mu_i + \gamma_i r_{i,t-1} + \eta_{i,t}, i = 1,2, \eta_{i,t}|\Omega_{t-1} \sim D(0, \mathrm{H}_t) \quad (3)$$

Where $r_{i,t}$ is the vector of the logarithmic price return of cryptocurrency i at time t, μ_i is a vector of parameters that estimates the mean of the price return of cryptocurrency i, $\eta_{i,t}$ is the vector of error terms for i at time t, with a positive definite conditional covariance matrix H_t given the available information set Ω_{t-1}. Thus, the size of all components of the conditional mean equation is (Tx1). The third equation is similar to the second one, but it contains an additional variable, $r_{i,t-1}$, which is the vector of the lagged logarithmic price return of cryptocurrency i, γ_i is its corresponding estimator. $D(0, \mathrm{H}_t)$, can be either a Student’s t or a Normal distribution.

For the matter of this paper, the sub index 1 refers to Bitcoin, while sub index 2 refers to Ethereum. As it is known, when we model the conditional variance of $\eta_{i,t}$, we are estimating the conditional variance of the returns.

5.1 Bivariate Diagonal VECH

One of the most popular multivariate GARCH model is Diagonal VECH introduced by Bollerslev, Engle and Wooldridge (1988), it is defined in matrix notation as follows:

$$h_t = C + A\Pi_{t-1} + Bh_{t-1} \quad (4)$$

where

$$h_t = vech(H_t),$$

$$\Pi_t = vech(\eta_t \eta_t'),$$

C is an $\frac{(N+1)N}{2} \times 1$ parameter vector and A and B are paremeter matrices of order $\frac{(N+1)N}{2} \times \frac{(N+1)N}{2}$. According to Bollerslev, Engle and Wooldridge (1988) in Diagonal VECH model both parameter matrices A and B should be diagonal, which means that every element of $\sigma_{ij,t}$ depends only on its own lag and on the previous value of $\eta_{i,t-1}$ and $\eta_{j,t-1}$. Hence, in bivariate case the model is defined as:

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{22,t} \\ \sigma_{12,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} + \begin{bmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{22} & 0 \\ 0 & 0 & \alpha_{33} \end{bmatrix} \begin{bmatrix} \eta_{1,t-1}^2 \\ \eta_{2,t-1}^2 \\ \eta_{1,t-1}\eta_{2,t-1} \end{bmatrix} + \begin{bmatrix} \beta_{11} & 0 & 0 \\ 0 & \beta_{22} & 0 \\ 0 & 0 & \beta_{33} \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{22,t-1} \\ \sigma_{12,t-1} \end{bmatrix} \quad (5)$$

This model can also be expressed as a system of equations:

$$\begin{cases} \sigma_{11,t} = c_1 + \alpha_{11}\eta_{1,t-1}^2 + \beta_{11}\sigma_{11,t-1} \\ \sigma_{22,t} = c_2 + \alpha_{22}\eta_{2,t-1}^2 + \beta_{22}\sigma_{22,t-1} \\ \sigma_{12,t} = c_3 + \alpha_{33}\eta_{1,t-1}\eta_{2,t-1} + \beta_{33}\sigma_{12,t-1} \end{cases} \quad (6)$$

However, Diagonal VECH model needs imposing some limitations to ensure the positivity of conditional covariance matrix H_t which sometimes can be very complicated in practice. In order to solve this inconvenient, we can easily impose positivity on H_t by employing the Diagonal BEKK model.

5.2 Bivariate Diagonal BEKK

Another widely used multivariate GARCH model is the Diagonal BEKK (Engle & Kroner, 1995), in which the conditional covariance matrix is defined in the following way:

$$\mathrm{H}_t = \mathrm{C}'\mathrm{C} + \mathrm{A}'\eta_{t-1}\eta_{t-1}'\mathrm{A} + \mathrm{B}'\mathrm{H}_{t-1}\mathrm{B} \quad (7)$$

Where H_t is a conditional covariance matrix of all the price returns, A and B are square $\mathrm{N} \times \mathrm{N}$ matrices and C is an upper triangular matrix of constants. In the Diagonal BEKK model, both parameter matrixes A and B are diagonal, and all their off-diagonal elements equal to zero. Thus, in the bivariate case we have the following model:

$$\begin{aligned} H_t &= \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} \\ &= \begin{bmatrix} c_{11} & 0 \\ c_{12} & c_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{bmatrix} \\ &+ \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} \eta_{1,t-1} \\ \eta_{2,t-1} \end{bmatrix} \begin{bmatrix} \eta_{1,t-1}\eta_{2,t-1} \end{bmatrix} \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \\ &+ \begin{bmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} & \sigma_{12,t-1} \\ \sigma_{21,t-1} & \sigma_{22,t-1} \end{bmatrix} \begin{bmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{bmatrix} \end{aligned} \quad (8)$$

This model can also be written as a system of equations:

$$\begin{cases} \sigma_{11,t} = c_{11}^2 + \alpha_{11}^2\eta_{1,t-1}^2 + \beta_{11}^2\sigma_{11,t-1} \\ \sigma_{22,t} = c_{12}^2 + c_{22}^2 + \alpha_{22}^2\eta_{2,t-1}^2\beta_{22}^2\sigma_{22,t-1} \\ \sigma_{12,t} = c_{11}c_{12} + \alpha_{11}\alpha_{22}\eta_{1,t-1}\eta_{2,t-1} + \beta_{11}\beta_{22}\sigma_{12,t-1} \end{cases} \quad (9)$$

As we can clearly see from the equation, the elements of the covariance matrix are positive without imposing any restrictions (because all the relevant coefficients are squared values). However, the solutions are not unique, i.e. we can replace positive value with a negative one and vice versa. Therefore, Engle and Kroner (1995) restrict α_{11}, β_{11} and diagonal elements of C matrix to be positive. In addition, we keep the restrictions of the stationarity condition.

Under the normal distribution, the vector θ of the parameters of the model is estimated by maximizing the log-likelihood function:

$$lnL(\theta) = \sum_{t=1}^{T}(-\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(|\mathrm{H}_t|) - \frac{1}{2}\eta_t'\mathrm{H}_t^{-1}\eta_t) \tag{10}$$

Under Student's t distribution with $v > 2$ degrees of freedom, the log-likelihood function is defined as follows:

$$lnL(\theta) = \sum_{t=1}^{T}(N\ln\left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi(v-2)}\Gamma\left(\frac{v}{2}\right)}\right] - \frac{1}{2}\ln(|\mathrm{H}_t|) - \frac{v+1}{2}ln[1 + \eta_t'[(v-2)\mathrm{H}_t]^{-1}\eta_t] \tag{11}$$

The estimation under the multivariate normal and multivariate Student's t distributions is performed using the statistical package Econometric Views (EViews10).

The time-varying correlation between Bitcoin and Ethereum is defined as:

$$\rho_t = \frac{\sigma_{12,t}}{\sqrt{\sigma_{11,t}\sigma_{22,t}}} \tag{12}$$

Where $\sigma_{11,t}$ is the conditional variance of Bitcoin, $\sigma_{22,t}$ is the conditional variance of Ethereum, $\sigma_{12,t}$ is the conditional covariance of both cryptocurrencies.

We can summarize both models as follows:

$$\begin{cases} \sigma_{11,t} = c_1 + \alpha_1\eta_{11,t-1} + \beta_1\sigma_{11,t-1} \\ \sigma_{22,t} = c_2 + \alpha_2\eta_{22,t-1} + \beta_2\sigma_{22,t-1} \\ \sigma_{12,t} = c_3 + \alpha_3\eta_{12,t-1} + \beta_3\sigma_{12,t-1} \end{cases} \tag{13}$$

Where the first term of the equations refers to a constant, the second refers to the error term of the conditional mean equation and the third one to the previous value of the variance/covariance.

5.3 Forecast evaluation

As mentioned before in this paper, forecasting cryptocurrencies volatility is essential for suitable investment strategy developing and portfolio construction purposes, therefore, it can help cryptocurrency investors to be more aware of their decisions. For forecasting evaluation, we perform an out-of-sample approach meaning that we split our sample in two parts. The first part covers the period from 7th August 2015 to 8th April 2018 and is used for in-sample estimation. In the second part, we keep 365 days (from 9th April 2018 to 8th April 2019) for out-of-sample forecast. In addition, we intend to forecast considering 3 different time horizons (1-day, 5-days and 20-days) to evaluate the accuracy of the models.

5.3.1 Volatility proxy

In the forecasting stage, we face the problem that the variable of interest σ_t^2 (conditional variance) is a latent variable, meaning that it is unobservable. To evaluate forecast accuracy, we have to compare the forecasted variance $\widehat{\sigma_t}^2$ to the latent variance σ_t^2. However, this is impossible and we should use an unbiased proxy for the latent volatility. A common approach among researchers is to use squared daily returns r_t^2 as the proxy for the true variance (Poon & Granger, 2003). Although squared returns are unbiased, they are an extremely noisy estimator of the latent volatility according to Andersen and Bollerslev (1998). To solve this problem, they recommend using cumulative squared intraday returns (also known as a realized volatility) as an alternative proxy. However, the realized volatility cannot be used as a proxy in our thesis, since we do not have access to high frequency data. Thus, the latent variance is proxied by using squared daily returns despite of its known flaws, and the covariance by multiplying the returns of each of the currencies:

$$\sigma_{11,t} = r_{1,t}^2 \quad ; \sigma_{22,t} = r_{2,t}^2 \quad ; \sigma_{12,t} = r_{1,t} r_{2,t} \tag{14}$$

5.3.2 Loss functions

There are different approaches to measure forecast accuracy, and the most popular one is the loss function. It evaluates the forecast error. The error in is defined as the difference between the latent variance and the forecasted variance.

Thus, let's define the forecast error at time t with e_t and the loss function with $f(e_t)$. Granger (1999) claims that a loss function should satisfy three properties:

- $f(0) = 0$ which means that if there is no error, there will be no loss
- $\min f(e_t) = 0$ and thus $f(e_t) \geq 0$
- $f(e_t)$ is monotonic non-decreasing as the forecast error moves away from zero.

According to Poon and Granger (2003) the most widely used loss functions include Mean Square Error (MSE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Hence, to measure the forecast accuracy in this research we employ MSE, MAE and RMSE.

$$MSE: L(\sigma_t, \hat{\sigma}_t) = \frac{1}{T}\sum_{t=1}^{T}(\sigma_t - \hat{\sigma}_t)^2 \tag{15}$$

$$MAE: L(\sigma_t, \hat{\sigma}_t) = \frac{1}{T}\sum_{t=1}^{T}|\sigma_t - \hat{\sigma}_t| \tag{16}$$

$$RMSE: L(\sigma_t, \hat{\sigma}_t) = \sqrt{\frac{\sum_{t=1}^{T}(\sigma_t - \hat{\sigma}_t)^2}{T}} \tag{17}$$

Patton (2011) argues that Mean Square Error is a robust loss function regardless the choice of proxy for the latent variance. However, Wilhelmsson (2006) claims that Mean Square Error is very susceptible to outliers whereas Mean Absolute Error is more robust to outlying values.

6. Empirical results

As we mentioned before, we want to compare the results obtained assuming a normal distribution on the data and assuming another distribution that allows the sample to have fatter tails, in this case Student's t distribution. Besides, we will check if the conditional mean for each one of the cryptocurrencies follows a AR(1) process or not. According to this, we estimate 8 different models: 4 for Diagonal VECH, and 4 for Diagonal BEKK. Each one of these groups split into Normal distribution and Student's t distribution. Finally, we estimate one following an AR(1) process and other that does not. We name them from I to VIII and they are described now:

Table 2. Description of the different models to estimate

Model	MGARCH		Distribution		Cond. Mean Process	
	D. BEKK	D. VECH	Student-t	Normal	Constant	AR(1)
I	X		X		X	
II	X		X			X
III	X			X	X	
IV	X			X		X
V		X	X		X	
VI		X	X			X
VII		X		X	X	
VIII		X		X		X

Note: Table 2 presents the convention of each one of the models i.e. Model I is estimated using Diagonal BEKK approach, under Student-s distribution and constant conditional mean. Model VII is estimated using Diagonal VECH, under normal distribution and with autoregressive conditional mean.

6.1 Parameters estimations

After running the regressions in Eviews we obtain the following results shown in Table 3:

Table 3. Values of the estimated parameters of the employed models

	I	II	III	IV	V	VI	VII	VIII
μ_1	0.0027***	0.0027***	0.0028***	0.0027***	0.0026***	0.0027***	0.0027*	0.0026***
γ_1		-0.0191		0.0687**		-0.0202		0.0709**
μ_2	0.0005	0.0006	0.0026	0.0025	0.0005	0.0005	0.0027	0.0017
γ_2		-0.0146		0.0733**		-0.0131		0.0699**
c_1	0.0000**	0.0000**	0.0000***	0.0000***	0.0000**	0.0000**	0.0000***	0.0000***
α_1	0.2936***	0.2836***	0.1661***	0.1653***	0.2917***	0.2799***	0.1686***	0.1600***
β_1	0.8316***	0.8374***	0.8404***	0.8388***	0.8314***	0.8362***	0.8385***	0.8496***
c_2	0.0003***	0.0003***	0.0003***	0.0002***	0.0004***	0.0003***	0.0003***	0.0002***
α_2	0.3946***	0.3409***	0.2911***	0.2100***	0.4407***	0.3939***	0.3347***	0.2324***
β_2	0.7437***	0.7773***	0.6990***	0.7834***	0.7150***	0.7346***	0.6682***	0.7465***
c_3	0.0000	0.0000	0.0000***	0.0000***	0.0000	0.0000	0.0000***	0.0000***
α_3	0.3404***	0.3109***	0.2199***	0.1863***	0.3343***	0.3064***	0.2193***	0.1517***
β_3	0.7864***	0.806***	0.7664***	0.8107***	0.7903***	0.8050***	0.7676***	0.8408***
$Student - t$	2.7523***	2.7577***			2.7593***	2.7799***		
$L. Likelihoood$	3513.12	3517.33	3282.38	3283.05	3514.73	3518.96	3284.66	3292.96

*Level of significance: * = 10%, ** = 5%, *** = 1%*

Note: Table 3 presents the estimated coefficients following the notation of equations 2,3 and 13 for each of the models in Table 2 and their corresponding statistical significance.

We can observe from the conditional mean equations coefficients (see equations 2 and 3), that using either the Student's t distribution or the normal distribution, Bitcoin (μ_1) seems to have a long term expected return around 0.27%, statistically significant at 1%, except for model VII when it is significant at a 10% of confidence. For Ethereum (μ_2), the constant in the conditional mean is not significant for any of the 8 different models proposed, and that the coefficients of autocorrelation (γ_1 and γ_2) have a negative sign with the Student's t distribution around -1.4% (not significant). On the other hand, we found that using a normal distribution the autocorrelation coefficient of 1 period is around +7% and significant at 5%.

For the conditional variance equations, in all the 8 models the coefficients turned out to be significant, most of them with a confidence level of 1%; just a few with a different behavior, c_1 resulted significant at a 5% when assuming a Student's t distribution (models I, II, V and VI), and c_3

not significant for the same models. The interpretation of these coefficients is rather not of much interest for the purpose of this paper. However, it is important to remark that all the constant terms in $\sigma_{11}, \sigma_{22}, \sigma_{12}$, despite of their statistical (un)significance, are very close to 0, so the evolution of the Variance/Covariance of the currencies is mainly determined by the error term of the conditional mean and the previous value of itself.

Additionally, we would like to remark that all the coefficients of the variance have the expected positive values, which was one the reasons to include the Diagonal BEKK model, given that it ensures the positivity of the coefficients by construction.

At the end of the table, there are two additional rows, named Student-t distribution and Log-likelihood. This first contains the coefficient estimated for the degrees of freedom of the distribution. As we can see, it only exists in the 4 models where we consider a Student's t distribution, and the coefficients are statistically significant at 1% and are between 2 and 3. It is important to highlight that for having a defined variance the degrees of freedom have to be greater than 2, so, in our models we can assure that the variance is well defined. Additionally, as we know, the Student's t distribution tends towards the normal distribution when we consider infinite degrees of freedom. Having so small figures (still well defined as mentioned before) leads us to think and confirm that our data follows, most likely, a heavy-tailed distribution, but we must show if these models are statistically better than the ones when we assume a normal distribution.

For this comparison, we use the likelihood ratio test. For the implementation of the test, we must have an unrestricted and a restricted model, where we have imposed restrictions to the estimators. The null hypothesis of the test states that the restricted model is sufficient to explain the dependent variable. It follows asymptotically a Chi-squared distribution with *m* degrees of freedom, where *m* is the number of restrictions. It is described as follows:

$$LR = -2 * (L_r - L_u) \tag{18}$$

Where L_r is the log-likelihood of the restricted model and L_u is the log-lkelihood of the unrestricted model.

In this paper, the restricted models are the ones that we assume to follow a normal distribution, and the unrestricted would be assuming a Student's t distribution. Likewise, the restriction that we impose is that the coefficient of the degrees of freedom is zero, so m=1. We compare the models that are identical in all the estimation but using different distribution on the sample data. This means that we compare I vs III, II vs IV, V vs VII and VI vs VIII. After the calculations, we find that the statistics are 461, 468, 460, and 452 respectively. According to what we mentioned before, the critical value of the Chi-squared distribution with 1 degree of freedom and

95% of probability is 3.84, so we can reject the null hypothesis and therefore conclude that the models following a t-distribution are statistically better fit for explaining the conditional variance with our data.

6.2 Model Fitting

As it is widely known, the most used GARCH(p,q) model for standard financial instruments is the GARCH (1,1) specification (Brooks, 2014), but according to what we mentioned before, the academic literature regarding the cryptocurrencies is not that broad and extensive. So, considering the aim of this paper where we want to find the most accurate model for forecasting the conditional variance of Bitcoin and Ethereum we ran the models I, II, V and VI with a higher order GARCH: GARCH (1,2), GARCH (2,1) and GARCH (2,2) in order to compare them with the GARCH(1,1). It is important to highlight that these 4 models are the ones that follow a Student's t distribution, because as we discussed in the previous section, the heavy-tailed distributions models outperformed the models with a normal distribution.

For comparing the different models, we use the Log-Likelihood ratio used before and the Information Criteria (IC) to see how adding new variables to the models is punished through the overfitting and how the parsimony is affected. We use the Akaike's (AIC), Schwarz Bayesian's (SBIC) and Hannan-Quinn's (HQIC).

First of all, just by checking the IC for each model, we find that we could disregard the GARCH(2,1) and GARCH(2,2) specifications as those have a larger coefficient for each one of IC considered. Regarding the GARCH(1,2), we find that this could be a better specification for model VI (as the 3 IC turned out to be the lower than our reference model), and for model II is not completely conclusive, because with 1 (AIC) out of 3 we find a larger figure. For the other two models (II and V) GARCH(1,1) has better results.

Table 4. Log-Likelihood Ratio test for fitting models

Log-Likelihood Ratio Test				
Restricted: GARCH (1,1)	I	II	V	VI
GARCH (1,2)	3.81	4.24	3.81	4.24
GARCH (2,1)	0.00	0.00	0.00	0.00
GARCH (2,2)	0.00	0.00	0.00	0.00

Note: Table 4 presents the value of the Log-Likelihood Ratio Test, comparing the GARCH (1,1) specification with the other 3.

So far, we do not have statistical evidence that GARCH(1,1) is better than the others, so as a second step, we perform the Log-Likelihood ratio test. As expected, the test reflects that we should not use GARCH(2,1) and GARCH(2,2) specifications as the Log-Likehood turned to be exactly the same than in the GARCH(1,1), therefore we prefer to use the restricted model. At the same time, when comparing the GARCH(1,2) for models II and VI we find a statistic of 4.24 and the critical value is 3.84 which means that with a probability of 95% we should use this model instead of the "restricted" one. Nevertheless, if we start increasing the probability, after 96.06%, we get that we should use the restricted model. For the sake of being conservatives, we keep using the GARCH(1,1) for all the models.

We could have contemplated GARCH(p,q) models of higher order than GARCH(2,2); but considering that: this is a first approach to the check the specification of the conditional variance, and that we preferred the GARCH(1,1) over the others in most of the cases, it is not likely that we could improve the forecasting by introducing more lags on the model.

6.3 Forecast evaluation

For the forecasting part of our paper, it is important to remark that we are forecasting 3 different variables, $\sigma_{11}, \sigma_{22}, \sigma_{12}$ (conditional volatility of Bitcoin, conditional volatility of Ethereum and the conditional covariance respectively). Moreover, we forecast considering 3 different time horizons (1 day, 5 days and 20 days) and we are comparing 4 different models (I, II, V, VI). Additionally, we are considering 3 different loss functions (MSE, MAE and RMSE). Thus, we can obtain diverse results, where different models might give the best forecast for each one of our variables.

Table 5. The best forecasting model on different loss functions

MAE			
	σ_{11}	σ_{22}	σ_{12}
1 day	VI	II	VI
5 days	VI	II	VI
20 days	VI	II	VI

MSE			
	σ_{11}	σ_{22}	σ_{12}
1 day	VI	II	VI
5 days	VI	VI	VI
20 days	VI	II	VI

RMSE			
	σ_{11}	σ_{22}	σ_{12}
1 day	VI	II	VI
5 days	VI	VI	VI
20 days	VI	II	VI

Note: Table 5 shows the model with the lower loss for each one of the variables, loss function, and horizon.

As we can see from Table 5, the model VI (using diagonal VECH and following an AR(1) process) turned to be the best model when we want to model the conditional variance of Bitcoin and the conditional covariance. On the other hand, for the conditional variance of Ethereum we find that the most accurate model is the model II (using diagonal BEKK and following an AR(1) process) in 7 of the different 9 scenarios that we forecasted; in the other two scenarios we obtain better results with model VI.

Considering these results, we would like to point that according to our research the best forecasts are obtained when we allow the conditional mean of the returns to be dependent on the lagged value of the returns, as these two models outperformed the forecast ability of the constant conditional mean equations.

Table 6. Relative performance

	MAE			MSE			RMSE		
1 day	σ_{11}	σ_{22}	σ_{12}	σ_{11}	σ_{22}	σ_{12}	σ_{11}	σ_{22}	σ_{12}
I	0.0038233	0.0095539	0.0044004	0.0000233	0.0002103	0.0000497	0.0037541	0.0161057	0.0065631
II	0.0029201	**0.0000000**	0.0039396	0.0000147	**0.0000000**	0.0000225	0.0023655	**0.0000000**	0.0029889
V	0.0028386	0.0179626	0.0041158	0.0000176	0.0003979	0.0000429	0.0028363	0.0301538	0.0056715
VI	**0.0000000**	0.0057437	**0.0000000**	**0.0000000**	0.0001454	**0.0000000**	**0.0000000**	0.0111800	**0.0000000**
5 days	σ_{11}	σ_{22}	σ_{12}	σ_{11}	σ_{22}	σ_{12}	σ_{11}	σ_{22}	σ_{12}
I	0.0156794	0.0312976	0.0044749	0.0003479	0.0013840	0.0003134	0.0232196	0.0410880	0.0188323
II	0.0145824	**0.0000000**	0.0185247	0.0002890	0.0000025	0.0004115	0.0193411	0.0000740	0.0246443
V	0.0109054	0.0659527	0.0062607	0.0002444	0.0026478	0.0003107	0.0163906	0.0777614	0.0186755
VI	**0.0000000**	0.0056870	**0.0000000**	**0.0000000**	**0.0000000**	**0.0000000**	**0.0000000**	**0.0000000**	**0.0000000**
20 days	σ_{11}	σ_{22}	σ_{12}	σ_{11}	σ_{22}	σ_{12}	σ_{11}	σ_{22}	σ_{12}
I	0.0859012	0.2028501	0.0752323	0.0046331	0.0208523	0.0038219	0.1016806	0.1902270	0.0842336
II	0.0738848	**0.0000000**	0.1069620	0.0040666	**0.0000000**	0.0055674	0.0894865	**0.0000000**	0.1217002
V	0.0614430	0.3931709	0.0748629	0.0032052	0.0408883	0.0038229	0.0708214	0.3670846	0.0842564
VI	**0.0000000**	0.0987652	**0.0000000**	**0.0000000**	0.0051712	**0.0000000**	**0.0000000**	0.0477958	**0.0000000**

Note: Table 6 presents relative performance of the 4 models estimated under Student-t distribution. The loss has been scaled by 100, in order to present easier figures to the read. The best model is normalized to be 0, as the minimum value is subtracted in each part.

The results of the loss functions are somehow relative to each case, and they have their own drawbacks. For instance, when we use the MSE we elevate the returns to the power of 4, so it could be very noisy with the outliers. Additionally, MAE is a linear loss function that weighs all the differences equally, whereas the RSME is a quadratic function and weighs heavily the higher differences. In Table 6, the relative performance of the models is presented. It is obtained by subtracting the loss value of the best model to the other models; thus, the best model has a value of 0. It is possible to see that in most of cases the difference among the models is not very large, at least when using the loss functions described in section 6.2, which can be seen as a positive result. We tried to use the Heteroscedasticity adjusted-MAE (HMAE) to control for the heteroscedasticity (Andersen, Bollerslev & Lange, 1999), but as we do not count with high frequency data for the

Realized Volatility the values of these calculations tend to be really close to zero, and then the loss function calculation explodes and we got unnormal results. The HMAE is defined as follows:

$$HMAE: L(\sigma_t, \hat{\sigma}_t) = \frac{1}{T}\sum_{t=1}^{T}\left|1 - \frac{\hat{\sigma}_t}{\sigma_t}\right| \tag{19}$$

We depict the forecast of the conditional variables and the realized value of each variable in the Figures 6, 7, and 8. It can be seen that the forecast of the 1-day horizon has a poor performance, since the estimated value does not capture the big changes in the realized variable. On the 5-days horizon, we get better results as the calculated variable follows the real value in a more accurate way. Finally, for the 20-days horizon the forecast takes larger values than the proxied value.

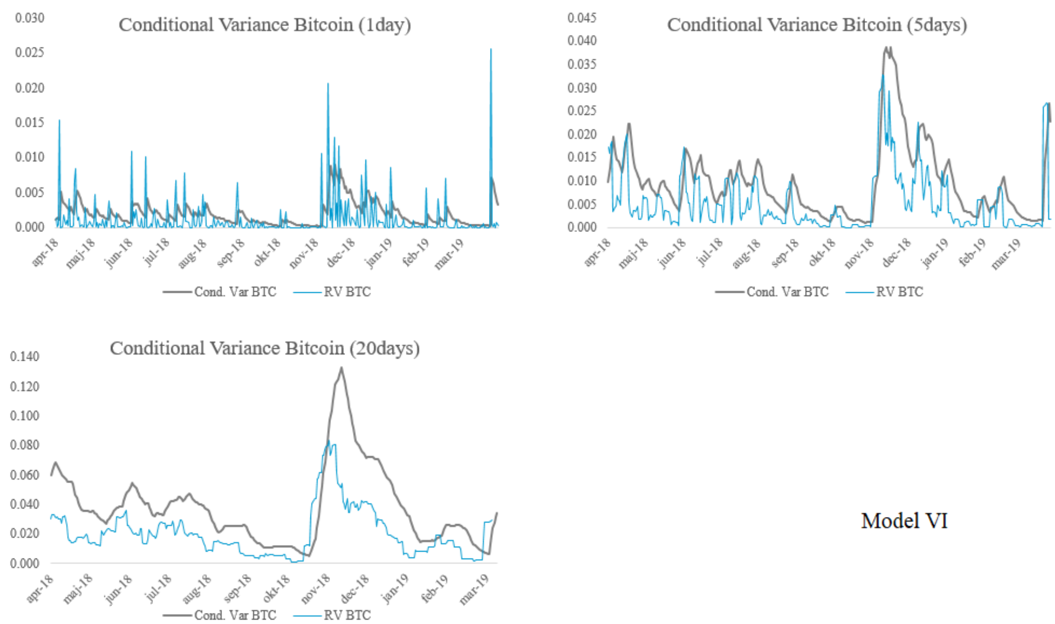

Figure 6. Realized Volatility and Forecasted Conditional Variance of Bitcoin for 1, 5, and 20 days using the specification of model VI

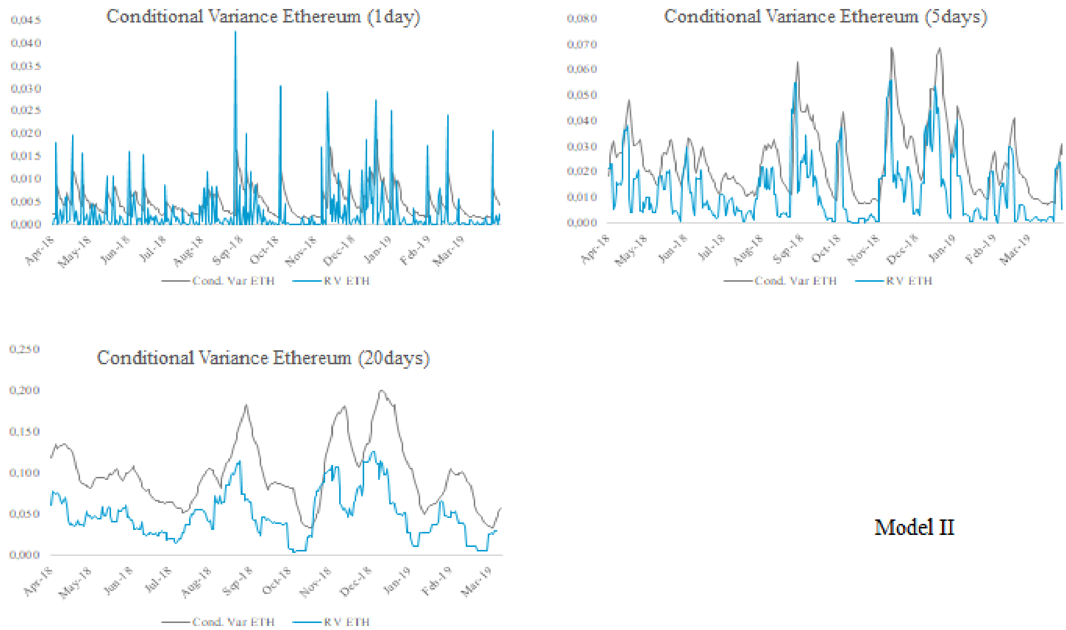

Figure 7. Realized Volatility and Forecasted Conditional Variance of Ethereum for 1, 5, and 20 days using the specification of model II

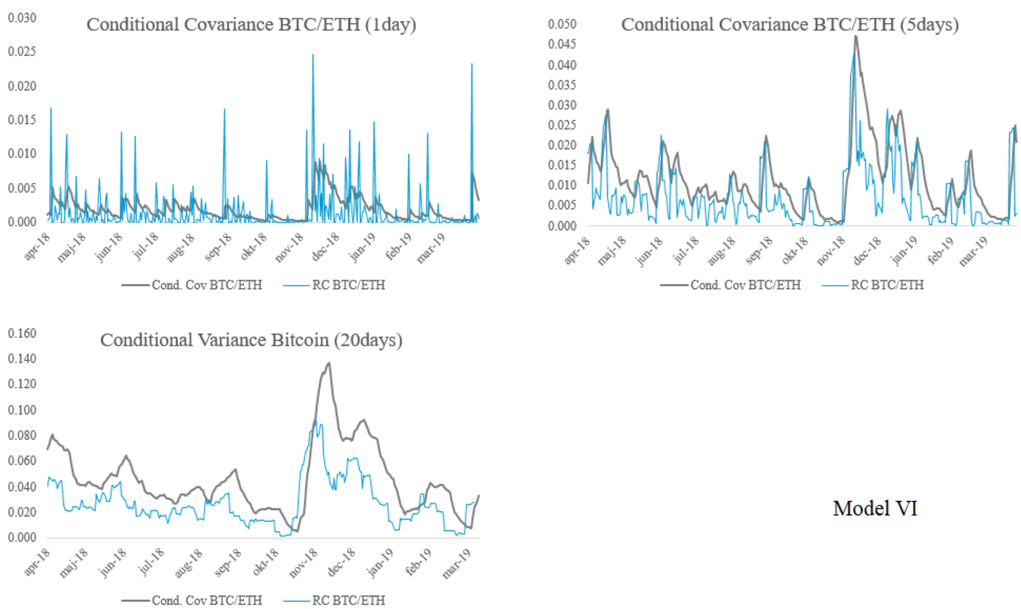

Figure 8. Realized Covariance and Forecasted Conditional Covariance of Bitcoin and Ethereum for 1, 5, and 20 days using the specification of model VI

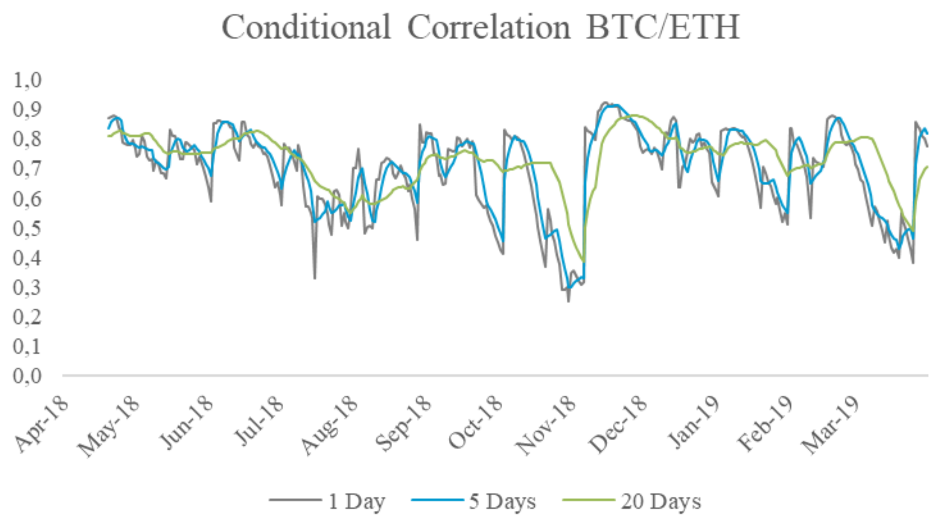

Figure 9. Conditional Correlation by time horizon

The Conditional Correlation is positive for the forecasted period, and we can confirm our hypothesis that the relation between the two variables is time varying, so it is better to treat them with this approach, than just assuming a constant value for them. Additionally, the correlation declines during high volatility periods, as in November of 2018; despite the fact that the covariance increases, the volatility grows in a larger magnitude.

7. Conclusions

One of the main aims of this paper was to test if the cryptocurrencies follow the known behavior of financial instruments on the distribution of their returns, i.e. if they have the feature of fatter tails. For this, we ran and compared the models assuming a normal distribution and a Student's t distribution, and identified heavy tails in the data as the models had between 2 and 3 degrees of freedom. Using the Log-Likelihood ratio test, we determined that we should use this approach since we got extreme values for the statistic. In this way, our findings are consistent with the outcome of Katsiampa (2018b), even though she employs only a Diagonal BEKK model, and we found a better estimate with model VI. Nevertheless, our daily conditional correlation ranges from 0,24 to 0,92; therefore, the latter results differ from the findings of Katsiampa (2018b) since she concludes that Ethereum can be used as a hedging instrument for the Bitcoin.

Overall, we find that the best model for forecasting the conditional variance using our two multivariate models is the Model VI, where we performed the estimation allowing our data to have fatter tails (since we used the Student's t distribution), and used the Diagonal VECH approach. It has the larger estimator of the degrees of freedom (2.77). We would like to highlight that it is the second-best model in the cases where it was not the best. Additionally, the shocks of the conditional forecast and covariance turned out to be statistically significant and persistent over time.

As the latter figures exhibit, the horizon of 5 days manages to produce the more similar forecast compared to the Realized Variables; as the 1-day horizon can be seen as high frequency data for the 5-days horizon. The daily forecast fails to reproduce the "real" value in times of high volatility, as we only count on close prices.

We are aware that one potential drawback in our paper is the lack of high frequency data for cryptocurrencies prices. As it was stated in section 6.1 for estimating the Realized Volatility, intradaily data should be collected and compared with the forecast of the conditional variance that we estimated to get more accurate results. This was the main reason why we could not use the HMAE in as a loss function for comparing models. A further research on the topic could be performed using high frequency data, as it fits better the real volatility of instruments.

References

Andersen, T. G. & Bollerslev, T. (1998). Answering the Skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review*, vol. 39, no. 4, pp. 885–905

Andersen, T. G., Bollerslev, T., & Lange, S. (1999). Forecasting Financial Market Volatility: Sample frequency vis-à-vis forecast horizon, *Journal of Empirical Finance*, vol. 6, no.5, pp. 457–477

Antonopoulos, A. M. (2014). Mastering Bitcoin: Unlocking digital cryptocurrencies, Sebastopol, California: O'Reilly Media Inc.

Baek, C. & Elbeck, M. (2015). Bitcoins as an Investment or Speculative Vehicle? A first look, *Applied Economics Letters*, vol. 22, no. 1, pp. 30–34, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Balcilar, M., Roubaud, D., Gupta, R., & Bouri, E. (2017). Can Volume Predict Bitcoin Returns and Volatility? A quantiles-based approach, *Economic Modelling*, vol. 64, pp. 74–81, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Baur, D. G., Hong, K., & Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets?, *Journal of International Financial Markets, Institutions and Money*, vol. 54, pp. 177–189, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Blau, B. M. (2017). Price Dynamics and Speculative Trading in Bitcoin, *Research in International Business and Finance*, vol. 41, pp. 493–499, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance, *Journal of Economic Perspectives*, vol. 29, no. 2, pp. 213–238, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, vol. 31, pp. 307-327

Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A Capital Asset Pricing Model with Time-Varying Covariances, *Journal of Political Economy*, vol. 96, no. 1, pp. 116-131

Brooks, C. (2014). Introductory Econometrics for Finance, New York: Cambridge University Press

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. 1–23, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Cheah, E. T. & Fry, J. (2015). Speculative Bubbles in Bitcoin Markets? An empirical investigation into the fundamental value of Bitcoin, *Economics Letters*, vol. 130, pp. 32–36, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Chu, J., Chan, S., Nadarajah, S., & Osterrieder, J. (2017). GARCH Modelling of Cryptocurrencies, *Journal of Risk and Financial Management*, vol. 10, no. 4, p. 17, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Ciaian, P., Rajcaniova, M., & Kancs, d'Artis (2016). The Economics of Bitcoin Price Formation, *Applied Economics,* vol. 48, no. 19, pp. 1799–1815, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

CoinMarketCap (2019). Top 100 Cryptocurrencies by Market Capitalization. Available Online: https://coinmarketcap.com [Accessed 11 April 2019]

Corbet, S., Meegan, A., Larkin, C., Lucey, B., & Yarovaya, L. (2018). Exploring the Dynamic Relationships Between Cryptocurrencies and Other Financial Assets, *Economics Letters*, vol. 165, pp. 28–34, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Dwyer, G. P. (2015). The Economics of Bitcoin and Similar Private Digital Currencies, *Journal of Financial Stability*, vol. 17, pp. 81–91, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Dyhrberg, A. H. (2016). Bitcoin, Gold and the Dollar - A GARCH Volatility Analysis, *Finance Research Letters*, vol. 16, pp. 85–92, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 8 March 2019]

Engle, R. F. & Kroner, K. F. (1995). Multivariate Simultaneous Generalized ARCH, *Econometric Theory*, vol. 11, no. 1, pp. 122–150

Foley, S., Karlsen, J. R., & Putniņš, T. J. (2019). Sex, Drugs, and Bitcoin: How much illegal activity is financed through cryptocurrencies?, *Review of Financial Studies*, vol. 32, no. 5, pp. 1798–1853, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Gkillas, K. & Katsiampa, P. (2018). An Application of Extreme Value Theory to Cryptocurrencies, *Economics Letters*, vol. 164, pp. 109–111, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Granger, C. W. J. (1999). Outline of Forecast Theory Using Generalized Cost Functions, *Spanish Economic Review*, vol. 1, no. 2, pp. 161–173

Hultman, H. (2018). Volatility Forecasting: An empirical study on bitcoin using GARCH and stochastic volatility models, Master thesis, Department of Economics, Lund

University, Available Online: http://lup.lub.lu.se/student-papers/record/8958504 [Accessed 08 March 2019]

Katsiampa, P. (2017). Volatility Estimation for Bitcoin: A comparison of GARCH models, *Economics Letters*, vol. 158, pp. 3–6, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 8 March 2019]

Katsiampa, P. (2018a). An Empirical Investigation of Volatility Dynamics in the Cryptocurrency, *SSRN Electronic Journal*, Available Online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3202317 [Accessed 8 March 2019]

Katsiampa, P. (2018b). Volatility Co-Movement between Bitcoin and Ether, *Finance Research Letters*, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 8 March 2019]

Patton, A. J. (2011). Volatility Forecast Comparison Using Imperfect Volatility Proxies, *Journal of Econometrics*, vol. 160, no. 1, pp. 246–256, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 14 May 2019]

Phillip, A., Chan, J., & Peiris, S. (2018). A New Look at Cryptocurrencies, *Economics Letters*, vol. 163, pp. 6–9, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 16 April 2019]

Poon, S.-H. & Granger, C. W. J. (2003). Forecasting Volatility in Financial Markets: A review, *Journal of Economic Literature*, vol. 41, no. 2, pp. 478–539, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 14 May 2019]

Wilhelmsson, A. (2006). Garch Forecasting Performance Under Different Distribution Assumptions, *Journal of Forecasting*, vol. 25, no. 8, pp. 561–578, Available through: LUSEM Library website http://www.lusem.lu.se/library [Accessed 14 May 2019]