LMU Munich
Faculty of Mathematics, Computer Science and Statistics
Chair of Financial Econometrics

Multivariate Volatility Modelling for Cryptocurrencies

Master Thesis

Stephanie Riedl

Full Professor:	Prof. Stefan Mittnik, PhD
Supervisor:	Malte Kurz, M. Sc.
Date of Submission:	26^{th} October 2018

Acknowledgement

I would like to acknowledge everyone who supported me during the creation of my thesis.

I first would like to thank my thesis advisor Malte Kurz. His door was always open for me to come with any problems or questions, and he always guided me in the right direction. His advice was on point and extremely helpful at any time, while he still allowed this thesis to be my own work.
Next, I want to thank Prof. Dr. Stefan Mittnik for his enthusiasm and input on the topic of this thesis.
I would also like to acknowledge Fernando Cuervo and my mother Christiane Riedl as second readers of this thesis, and I am grateful for their valuable comment.
Finally, I would like to thank my family and friends who provided my with continuous support and encouragement throughout my whole academic career. They never failed to believe in me, even in times when I doubted myself. This accomplishment would not have been possible without them.

Abstract

Cryptocurrencies as an investment have received increasing attention by media and international governments over the last years. However, little is known yet about the dynamics that drive these highly volatile alternative assets. This thesis studies the dynamic interdependencies between the volatility of Bitcoin, Litecoin, Ripple, Dogecoin and Feathercoin via the Dynamic Conditional Correlation model by Engle (2002) with the multivariate Student-t distribution. The main question is whether a multivariate approach improves the Value at Risk forecasting accuracy for the conditional heteroscedasticity in comparison to univariate GARCH-type models. Results show that there is a high interconnectedness between the volatility of the currencies. However, the Dynamic Conditional Correlation model can not deliver better forecasting results than the univariate GARCH-type models for the individual cryptocurrency return series.

Contents

List of Figures

List of Tables

1 Introduction

Over the last years, the public interest in cryptocurrencies has increased dramatically. The decentralized peer-to-peer system makes cryptocurrencies an attractive phenomenon on the internet and many enthusiasts see it as an independent alternative towards the traditional financial market which has experienced bad reputation ever since the market turmoil of 2008. Since mid 2017, many cryptocurrencies have experienced an extreme price increase which provided the early adopters with an immense return for their former low priced investments. The cryptocurrency market has been extremely volatile since then, which has lead to a lot of economic and scientific attempts to understand and predict the risk of investment. Previous analyses have focused on the determinants of cryptocurrency prices or their interconnectedness with traditional markets. Other authors studied the sociodemographic profile of cryptocurrency buyers and tried to find relations with social media and browsing activities. Yelowitz and Wilson (2015) analysed Google trends data and found that a high interest in Bitcoin is positively correlated with computer programming and illegal activity search terms online.
Many academic analyses focus on understanding the highly volatile behaviour of the price developments. Using GARCH-type models to determine the risk of Bitcoin and other cryptocurrencies has become a popular topic in academic research over the last two years. However, less is known yet about the interdependencies between the volatility behaviour of different coins.

The main attempt of this thesis is to explore whether a multivariate approach can improve the modelling and forecasting of volatility for the selected cryptocurrencies. It is to find out if there are any correlations and mutual dynamics in the altcoin market that help to determine the volatility of the single currencies or if the trends are very particular for every currency and they are modelled more appropriately via a univariate approach. Therefore, there are GARCH-type models applied to the univariate log return series first. Several variance models from the GARCH family are utilized to model the evolution of the conditional variance of the return series. For the distribution of the innovations, heavy-tailed and skewed alternatives are applied next to the Gaussian distribution since previous research has brought up findings in favour of heavy-tailed distributions for the returns of financial assets. The different models with different distribution assumptions are going to be evaluated via model diagnostics and information criteria, as well as their ability to forecast the downside portfolio risk (VaR). Next to the single return series of the

five analysed currencies, an aggregated portfolio is also modelled via a univariate approach.
The results of the univariate empirical analysis are going to be used to proceed to a multivariate volatility modelling approach. The residuals of the selected GARCH models are implemented to model the time-varying covariance of the volatility of the five currencies. The multivariate Dynamic Conditional Correlation model by Engle (2002) thus is applied to the multivariate log return series. The fit of the multivariate Normal distribution for the innovations is outperformed by the multivariate Student-t distribution. After fitting the DCC model, its ability to forecast Value at Risk for an aggregated portfolio is evaluated and eventually compared to the results for the univariate GARCH models.

The thesis is organized as follows: First, there is an introduction on the cryptocurrencies used in this paper. The currencies Bitcoin (BTC), Litecoin (LTC), Ripple (XRP), Dogecoin (DOGE) and Feathercoin (FTC) are selected because they have been operational for some time already, which is in favour of statistical time series analysis. There is a short description of the single currencies, pointing out the main technical differences between them. This is followed by a review of previous literature on volatility modelling for cryptocurrencies. The majority of authors focus on GARCH-type models for Bitcoin - the most popular currency on the altcoin market. The choice for a certain GARCH model differs, but most authors have found the Gaussian distribution to be inappropriate to model the innovations of the log return series and proceed to skewed alternatives. Other authors analyse the dynamics between cryptocurrencies and the traditional asset market. Fewer have focussed on the multivariate dynamics between the volatility of different crypto coins so far.

It follows a theoretical section with an explanation of the model equations of the different applied GARCH type models. Also there is an overview of the different distributions that the innovations are assumed to follow. Besides the Normal distribution, the heavy-tailed Student-t distribution, as well as its skewed modification and the skewed Generalized Error distribution are described. The Maximum-Likelihood estimation for the simple GARCH model by Bollerslev (1986) is outlined. Then, there is an explanation of the different model diagnostics that are appropriate for financial volatility analysis. Since the models are also evaluated by their ability to forecast the realized portfolio loss, the riskmetric Value at Risk and its application to GARCH processes is introduced.
Next, the results for the univariate empirical analysis of the five crypto return series are presented. First, the stylized facts are shown, which highlight the skewed

and heavy-tailed distribution of the log returns that is typical for financial time series. Also, an analysis of the autocorrelation of the returns reveals conditional heteroscedasticity in the data, which justifies the application of GARCH models. The different GARCH models, namely the simple GARCH, eGARCH, iGARCH, apARCH, cs-GARCH and gjr-GARCH with the different distribution assumptions for the innovations are applied to the series and compared via diagnostic checking, information criteria and Value at Risk forecasts. The results show that the Gaussian distribution is clearly outperformed by the heavy-tailed distributions when it comes to model fit and forecasting performance. The apARCH, eGARCH and csGARCH, combined with the skewed Student-t or skewed Generalized error distribution for the innovations, seem to be most appropriate to model and forecast the dynamics of the univariate volatility. Then, the out-of-sample rolling forecast is executed with a higher number of model refits for the aggregated portfolio. It shows that a more frequent update of model parameters does not improve the forecasting results.
The next section outlines the theoretical framework for the multivariate analysis. Since the estimation of Dynamic Conditional Correlation models requires higher computational effort, the presence of multivariate conditional heteroscedasticity in the data should be detected ex ante. Therefore, it is useful to apply the multivariate Ljung-Box test to find time-variant dependencies in the cross correlation matrix. Then, the Dynamic Conditional Correlation model by Engle (2002) is defined. The model estimates the dynamic correlation matrix of the standardized innovations from the univariate series and thus models dynamics between the data. The conditional distribution of the multivariate volatility can be modelled by a multivariate Gaussian distribution or the multivariate Student-t distribution. Next, the 2-stage estimation of the DCC via Quasi-Maximum-Likelihood is described. At the first stage, the volatility of the univariate return series is obtained and the standardized residuals are used to estimate the mutual conditional correlation matrix at the second stage. The last chapter for the theoretical framework gives an overview of additional model diagnostic techniques that can be applied to the multivariate residuals.
After that, the results for the multivariate part of the empirical analysis are presented. Multivariate conditional heteroscedasticity in the sample correlation matrix is detected, which justifies the application of the DCC. Several DCC orders are applied with the conditional multivariate Normal and multivariate Student-t distribution, using the results from the univariate GARCH analysis to estimate the correlation of volatilities. It shows that a simple DCC order of lags (1,1) with the multivariate Student-t distribution provides the most appropriate fit according to

information criteria. After fitting the model, a rolling out-of-sample forecast for the Value at Risk is performed. Therefore, an aggregated portfolio is created. Its distribution is determined by the forecasted mean and covariance matrix of the DCC, according to the approach by Bauwens and Laurent (2005). However, the Value at Risk forecast with the multivariate approach does not show more accurate results than the univariate approach for the aggregated portfolio. Moreover, Value at Risk forecasting of the univariate individual return series can not be outperformed by the aggregated univariate or multivariate approach.
As a final conclusion, the results and their limitations are discussed and implications for further research are given.

2 Background on Cryptocurrencies

A cryptocurrency can be defined as "a digital asset designed to work as a medium of exchange using cryptography to secure the transactions and to control the creation of additional units of the currency" (Chu et al., 2017, pg. 1).

In the following section, there is a brief overview on the five cryptocurrencies that are used in this study. Cryptocurrencies are digital currencies and mainly characterized by their decentralization and trading on an online peer-to-peer network. The digital coins are generated via cryptographic techniques and recorded and verified by the community. Bitcoin was the first decentralized currency emerging in 2008 and has entailed many successors. The website Coinmarketcap.com lists 1969 different operating cryptocurrencies in September 2018. The five currencies used in this thesis were chosen since they are operational for a long time and provide a high observation span, which is of benefit to statistical analysis.

Bitcoin (BTC) is the first cryptocurrency and has been operational since 2009. In 2008, a person with the pseudonym "Natoshi Sakomoto" created a document on the alternative currency, which has become the most traded coin on the market today. It was the first decentralized currency that runs on a peer-to-peer network with transactions between users happening without mediation by a third party (e.g. a financial institution). The transactions are recorded via a *Blockchain*, an extendible list of datasets that are connected by cryptographic procedures and verified by a network of individuals using the Bitcoin software. The individuals who offer their computer power to keep track of the transactions are called *miners*. Miners solve a cryptographic puzzle that uses the transaction data. This is how Bitcoins are created, which gives further incentive to miners on top of transaction fees (Chuen et al., 2017). When Sakomoto first published his document on the new cryptocurrency, there were still 50 BTC generated during the mining process of one block. The supply of Bitcoin is limited to 21 million and the reward drops by 50% after every 210.000 transaction blocks (Nakamoto, 2008). In September 2018, the reward for every mined block is 12.5 coins and only 18 percent of the Bitcoins are left to be mined (Blockchainhalf.com).

Due to the immense consumption of computation power - a regular PC would need several years to solve a puzzle - mining has become non profitable for individuals and mainly been commercialized over the last years (Chuen et al., 2017).

The usage of cryptographic techniques ensures that transactions of Bitcoins can only be executed by the Bitcoin owners and the currency can not be spent twice. The system is safe against transaction hackers since changing the transaction history would require redoing all puzzles of all blocks linked together in the chain. That again would take enormous computational power for the hacker (Nakamoto, 2008).

The source code for Bitcoin is publicly available online on Github.com, which has motivated many successors to create alternative cryptocurrencies with improved qualities.

Litecoin (LTC) released in October 2011, is one of the most important successors of Bitcoin and its operation is nearly identical to Bitcoin. It is a peer-to-peer decentralized currency based on an open source security protocol, published in October 2011 by Charles Lee (Chuen et al., 2017). The only difference to Bitcoin is that blocks are generated faster (2,5 minutes instead of 10), which is why transactions by users can be made faster. Therefore, higher trading volumes can be handled and the network is scheduled to produce 84 million coins eventually (Litecoin.org).

Ripple (XRP) was developed completely independent from Bitcoin. It is built on an open source decentralized consensus protocol, even though the deployment is provided by Ripple Labs who hold 25% of the currency, next to 20% held by the founders. It is operating since 2012. In general, Ripple is based on a public data bank where different account balances are registered. Additionally, the register contains options on goods and traditional currencies (Dollar, Yen, etc.) which can be bought in the system using the intern currency "XRP". XRP can either be traded itself or used to make payments for other goods and currencies.

In the Ripple system there are users who make payments, market makers who enable

the trading in the system, and validating servers that run the protocol in the system to check and validate transactions. All historic transactions and account balances are publicly available.
Just like Bitcoin, the transactions in the Ripple system are ensured by an Elliptic Curve Digital Signature Algorithm. But instead of using miners, Ripple relies on the consensus of the validating servers to vote for the correct transaction in the system. While Bitcoin transactions can only be confirmed after mining the blocks, which takes an hour on average, the consensus in Ripple system is reached within a few seconds. Therefore, faster payments are supported (Armknecht et al., 2015).

Dogecoin (DOGE) was usually intended as a parody of Bitcoin and is named after the Shiba Inu dog "Doge" which became a popular internet phenomenon in 2013. Dogecoin is based on the same operating system as Bitcoin and Litecoin, but the block generating algorithms work even faster than for Litecoin with one block being produced every minute. The ultimate number of Dogecoins is not limited (Dogecoin.com). The currency, which was originally intended as a joke, has experienced immense popularity after its creation in 2013. The trading of DOGE is processed in social networks like Reddit and Twitter (Chuen et al., 2017).

Feathercoin (FTC) is another cryptocurrency which is based on the Bitcoin operating system and was released in April 2013. Just like other Bitcoin successors it works with a faster average block time of one minute. The total number of Feathercoins is limited to 336 millions, with a block reward halving every 2.1 million blocks. A special feature of Feathercoin is the Neoscript Algorithm that is used for mining and requires less computer power than the Bitcoin algorithms (Feathercoin.com). Therefore, Feathercoin experienced a hype in late 2013/early 2014 - see figure (1) in section (5) - but was eventually outperformed by other emerging currencies.

3 Previous Research on Volatility of Cryptocurrencies

The interest in the analysis of the variation in cryptocurrency prices mainly arises since they are highly volatile compared to traditional currencies. When invested at the right time, they provided their owners with immense profits. In the past two years, researchers have made effort to study the variations of cryptocurrency returns via Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models. However, fewer studies have made effort to study the multivariate dynamics of the crypto market.

Most studies so far focus on the modelling of Bitcoin only, which is the first and most popular cryptocurrency. Its high market volume and long observation span makes it most attractive for statistical analysis. Since Bitcoin and other cryptocurrencies have experienced immense popularity over the last years, a lot of academic effort has been made to determine the development of prices. A lot of studies have focused on the interconnectedness between Bitcoin and Twitter or Google activities. Georgoula et al. (2015) have shown that the Twitter sentiment ratio is positively related with Bitcoin prices. Other authors like Ciaian et al. (2016) have focused on the impact of investors and macro-economics and found that market forces and Bitcoin attractiveness for investors have a big influence on the short run but does not seem to influence the overall long-term price development.

Concerning stylized facts, cryptocurrency time series are characterized similarly to other financial time series. They exhibit time-varying volatility, extreme observations and an asymmetry of the volatility process to the sign of past innovation (Catania et al., 2018). Just like other econometric time series, the log returns of the cryptocurrencies have been found to show major deviations from normality. Chu et al. (2017) find that seven of the most important currencies are positively skewed. They also show extreme volatility compared to traditional assets, especially with regard to inter-daily prices. Gkillas and Katsiampa (2018) study the heavy-tail behaviour of Bitcoin, Bitcoin Cash, Etherum, Ripple and Litecoin. They find that cryptocurrencies are more risky and volatile than traditional currencies and also show heavier tail behaviour. Bitcoin and Litecoin showed to be the least risky currencies of the five analysed coins. Phillip et al. (2018) study the stylized facts of 224 different cryptocurrencies and find that they show leverage effects and a negative correlation between one-day ahead volatility and returns. They also find that the returns follow a Student-t distribution rather than a Normal distribution. Besides stylized facts, many authors have studied the conditional variance of the individual cryptocurrency return series via Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models. Chen et al. (2016) give one of the first analysis of the dynamics in the CRIXindex family. The CRIX (CRypto currency IndeX) provides market information based on the 30 most important currencies and is modelled by a ARIMA(2,0,2)-t-GARCH(1,1) process to capture the volatility of the index return series. Chu et al. (2017) apply different univariate GARCH-type models to seven of the most popular currencies (Bitcoin, Dash, Litecoin, MaidSafeCoin, Monero, Dogecoin and Ripple) and mainly claim the iGARCH and gjrGARCH to be the best models to explain univariate volatility for those currencies. They also show that cryptocurrencies exhibit extremely high volatility when they look at inter-daily prices.

Katsiampa (2017) explores several conditional heteroscedasticity models to Bitcoin and finds the autoregressive compenent GARCH model to fit the data best, which has both a long-run and a short-run component of conditional variance. Katsampias study has been replicated by Charles and Darné (2018) who use robust QML estimators to fit the GARCH-type models instead of standard Maximum-Likelihood estimators to take the conditional non-normality of the returns into account. Improvements in the estimation method again result in the choice of an AR-csGARCH process to describe the conditional heteroscedasticity in the Bitcoin return series. However, both papers miss out to investigate whether an alternative distribution like the Student-t or skewed Student-t distribution might fit the log returns better. Several studies have found that fat-tailed, possibly skewed distributions are more adequate to describe financial data (Kuester et al., 2006).

Angelini and Emili (2018) attempt to forecast the volatility for six cryptocurrencies with GARCH-type models. They compare the forecasting performance of the simple GARCH, eGARCH, tGARCH, GARCH-M and apARCH with a training dataset of 700 daily prices, combined with a Student-t distribution for the innovations. Then, they perform $h = 1, ..., 7$ steps-ahead forecasts with a recursive window. The eGARCH seems to perform best for the higher forecast horizons overall, however, results differ between the different currencies and forecasting horizons.

Catania and Grassi (2017) claim that standard volatility models like the GARCH are outperformed by alternatives like the Score Driven model with conditional Generalized Hyperbolic Skew Student-t innovations (GHSKT). With the chosen model they legitimately react to the skewness of the distribution of the log returns, however, they miss out on comparing the models via Value at Risk - performance is evaluated by AIC and BIC.

Other studies also explore the dynamics between Bitcoin price volatility and the

US stock market. Conrad et al. (2018) use the GARCH-MIDAS model to extract long-term and short-term volatility components of the Bitcoin price series. They find that the S&P realized volatility has a negative and highly significant effect on long-term Bitcoin volatility. According to their results, Bitcoin volatility is negatively correlated to the US stock market volatility.

However, fewer literature has focussed on the dynamic correlation of volatility between the different cryptocurrencies yet, with the most papers on this issues being published in 2018. Corbet et al. (2018) study the relationship between Bitcoin, Ripple and Litecoin and other traditional financial indices by generalized variance decomposition methods and find that the price developments of cryptocurrencies are highly connected to each other while they are disconnected to mainstream assets on the long run. Spillovers between Bitcoin and traditional indices (e.g. SP500 and VIX) can only be observed on the short run.

Katsiampa (2018) employs an Asymmetric Diagonal BEKK multivariate GARCH model with a multivariate Student-t distribution for the error terms to the log returns of Bitcoin, Etherum, Ripple, Litecoin and Stellar to estimate the dynamic volatility of those currencies. He found that the conditional covariances were significantly affected by the past covariances of the innovations. The conditional correlation between the five currencies has shown to be mainly positive but changing over time.

The analysis of the evolution of different cryptocurrency volatilities has been a popular field of econometric research for the last two years. Many authors have already compared several univariate GARCH-type models to predict Value at Risk forecasts for the most popular cryptocurrencies and few focused on the dynamic interdependencies of conditional covariances. This thesis attempts to combine those two approaches and find out whether the performance of Value at Risk forecasting can be improved by employing a multivariate approach.

4 Univariate Volatility Modelling

This section deals with the theoretical framework for univariate volatility modelling and Value at Risk forecasting.
Volatility models are also referred to conditional heteroscedasticity models. Here, volatility means "the conditional standard deviation of the underlying asset return" (Tsay, 2005, pg. 109). Volatility modelling provides a simple approach to calculating Value at Risk of a financial position in risk management. It can improve the efficiency in parameter estimation and the accuracy in interval forecasting. Volatility is not directly observable, however, it has some characteristics that are common for asset return series. First, there are volatility clusters - volatility may be high for certain periods and low for others. Second, volatility evolves over time in a continuous manner. Third, volatility does not diverge to infinity, but varies within some fixed range. Fourth, volatility seems to react differently to a big price increase or a big price drop, which is referred to as "leverage effect". (Tsay, 2005).

4.1 Structure of Univariate Volatility Models

Let r_t be the log return of an asset at time index t. In volatility modelling the series r_t is serially uncorrelated but dependent. The conditional mean and variance of r_t given F_{t-1} can be described as

$$\mu_t = E(r_t|F_{t-1}), \quad \sigma_t^2 = Var(r_t|F_{t-1}) = \mathbb{E}[(r_t - \mu_t)^2|F_{t-1}], \tag{1}$$

where F_{t-1} denotes the information available at time $t-1$. If the serial dependence of r_t is weak, a simple time series model for μ_t can be entertained, such as a stationary Autoregressive Moving Average (ARMA-p, q)-process:

$$r_t = \mu_t + a_t, \tag{2}$$

$$\mu_t = \sum_{i=1}^{p} \phi_i \gamma_{t-1} - \sum_{i=1}^{q} \theta_i a_{t-i}, \quad \gamma_t = r_t - \phi_0 - \sum_{i=1}^{k} \beta_i x_{it}, \tag{3}$$

where k, p and q are nonnegative integers, and x_{it} are explanatory variables. γ_t represents the adjusted return series after removing the effect of explanatory variables. a_t is referred to as the *shock* or *innovation* of a time series. Combining equation

(1) and (3) gives:

$$\sigma_t^2 = Var(r_t|F_{t-1}) = Var(a_t|F_{t-1}). \tag{4}$$

Conditional heteroscedasticity models are concerned with the evolution of σ_t^2. The patterns under which σ_t^2 evolves over time distinguishes one volatility model from another. Equation (4) is referred to as the volatility equation of r_t (Tsay, 2005).
In volatility modelling, the first step is to test for conditional heteroscedasticity, also known as Autoregressive Conditional Heteroscedasticity (ARCH) effects. These can be detected by applying the univariate Ljung-Box Test (McLeod and Li, 1983). Let $a_t = r_t - \mu_t$ be the residuals to the mean equation, then the test statistic $Q(m)$ can be applied to the $[a_t^2]$ series. The null hypothesis is that the first m lags of the Auto Correlation Function of a_t^2 are zero. The test statistic is defined as:

$$Q(m) = T(T+2)\sum_{\ell=1}^{m}\frac{\rho_\ell^2}{T-\ell}, \tag{5}$$

where T is the length of the return series, ρ_ℓ is the estimated autocorrelation at lag ℓ and m is the maximum number of tested lags. The test is rejected if $Q(m) > \chi^2_{1-\alpha,d}$ for d degrees of freedom.

Variance Model. The most common approach to model conditional heteroscedasticity for univariate time series is a **simple GARCH** model (Generalized Autoregressive Conditional Heteroscedasticity) (Bollerslev, 1986)). Here, a_t follows a GARCH-(m, s) model if

$$a_t = \sigma_t \epsilon_t, \tag{6}$$

and the volatility of the innovations evolves according to:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m}\alpha_i a_{t-i}^2 + \sum_{j=1}^{s}\beta_j \sigma_{t-j}^2, \tag{7}$$

where $[\epsilon_t]$ is a sequence of iid random variables with mean zero and variance one. For $m = 0$ the process reduces to the ARCH(s)-process, while for $m = s = 0$ the innovations a_t are assumed to be white noise (Bollerslev, 1986). For the simplest version of a GARCH-$(1, 1)$ model, the equation can be reduced to

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad 0 < \alpha_0, 0 \le \alpha_1, \beta_1 \le 1, (\alpha_1 + \beta_1) < 1. \tag{8}$$

The latter constraint on $(\alpha_i+\beta_j)$ implies that the conditional variance of a_t is finite, whereas its conditional variance of σ_t^2 evolves over time. A large shock a_{t-1} or a large σ_{t-1} will give rise to a larger σ_t. The tail distribution of a GARCH process is heavier than that of the Normal distribution. The model provides a simple parametric function that can be used to describe the evolution of volatility (Tsay, 2005). This version of the GARCH model can be denoted as the **standard GARCH-**$(1,1)$ model.
Furthermore, there is the strictly stationary **integrated GARCH** model (iGARCH) (Engle and Bollerslev, 1986) for the particular case of the standard GARCH$(1,1)$ model where $\alpha_1+\beta_1=1$. This change of the constraint on the α and β parameters makes the model stationary, therefore, structural breaks in the data should be investigated ex ante (Ghalanos, 2018).
Some models take the asymmetry of positive and negative shocks into account. The market might react differently to a large negative shock in terms of the evolution of volatility than to a large positive shock. This is modelled by the **exponential GARCH** (eGARCH) model (Nelson, 1991), where the volatility equation can be written as

$$log(\sigma_t^2) = \alpha_0 + \alpha_1 a_{t-1} + \gamma_1[|a_{t-1}| - \mathbb{E}(|a_{t-1}|)] + \beta_1 log(\sigma_{t-1}^2), \tag{9}$$

for $0<\alpha_0, 0<\alpha_1, 0<\beta_1, 0<\gamma_1$. α_1 captures the sign effect and γ_1 captures the size effect of the past innovation.
Another asymmetric GARCH is denoted by the **GJR-GARCH** model due to Glosten et al. (1993):

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \gamma_1 I_{t-1}\alpha_{t-1}^2 + \beta_1\sigma_{t-1}^2, \tag{10}$$

for $0<\alpha_0, 0<\alpha_1, 0<\beta_1, 0<\gamma_1$, where $I_{t-1}=1$ if $a_{t-1}\leq 0$ and $I_{t-1}=0$ if $a_{t-1}>0$. Here, γ_1 represents the asymmetric parameter since a positive shock will affect σ_t^2 by α_1 and a negative shock by $\alpha_1+\gamma_1$.
The **asymmetric Power ARCH** (Ding et al., 1993) denoted by

$$\sigma_t^\delta = \alpha_0 + \alpha_1(|a_{t-1}^2| - \gamma_1 a_{t-1})^\delta + \beta_1\sigma_{t-1}^\delta, \tag{11}$$

for $0<\alpha_0, 0\leq\alpha_1, 0\leq\beta_1, 1<\gamma_1, 0<\delta$ models for both the leverage and the effect that the sample autocorrelation of absolute returns are usually larger than that of squared returns. The δ parameter of the apARCH is a parameter for the Box-Cox transformation and γ_1 is a leverage parameter (Chu et al., 2017). It is equivalent to the standard GARCH if $\delta = 2$ and $\gamma_1 = 0$ and to the gjrGARCH if $\delta = 2$ (Ghalanos, 2018).

The **component standard GARCH** model (csGARCH) (Engle and Lee, 1999) contains a time-varying intercept and decomposes the conditional variance into permanent and transitory components to investigate long- and short-run moments of volatility. The model is deployed as follows:

$$\sigma_t^2 = q_t + \alpha_1(a_{t-1}^2 - q_{t-1}) + \beta_1(\sigma_{t-1}^2 - q_{t-1}), \tag{12}$$

$$q_t = \alpha_0 + pq_{t-1} + \phi(a_{t-1}^2 - \sigma_{t-1}^2), \tag{13}$$

for $0 < \alpha_0, 0 \leq \alpha_1, 0 \leq \beta_1, 0 < \delta, 0 \leq \phi$. If $\alpha_1 + \beta_1 < 1$ and $p < 1$ weak stationarity holds. q_t represents the permanent component of the conditional variance. It can be seen as a time-varying intercept for the conditional heteroscedasticity.

The simple GARCH, iGARCH, eGARCH, apARCH, gjrGARCH and csGARCH are going to be applied to the five cryptocurrency time series.

Distribution model In the standard version of the GARCH-model, ϵ_t follows an independent identically Gaussian distribution. The simple GARCH-(1,1) model with Normal distribution assumption is then denoted by:

$$\begin{aligned} r_t &= \mu_t + \sigma_t\epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1\sigma_{t-1}^2, \\ \epsilon_t &\sim N(0,1). \end{aligned}$$

However, for financial time series analysis it has shown to be more appropriate to assume that ϵ_t follows a heavy-tailed distribution, such as a standardized **Student-t** distribution (see Kuester et al. 2006). Cryptocurrency log returns have already been found to follow a Student-t distribution by Phillip et al. (2018). Let x_ν be a Student-t distribution with ν degrees of freedom. Then, $Var(x_\nu) = \nu/(\nu-2)$ for $\nu > 2$, and we use $\epsilon_T = x_\nu/\sqrt{\nu/(\nu-2)}$. The probability function of ϵ_t is

$$f(\epsilon_t|\nu) = \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}}\left(1 + \frac{\epsilon_t^2}{\nu-2}\right)^{-(\nu+1)/2}, \quad \nu > 2, \tag{14}$$

where $\Gamma(x)$ is the usual Gamma function (Tsay, 2005). Besides fat tails, empirical distributions of fiancial asset returns may also be skewed. Skewed and heavy-tailed distributions have shown to provide better forecasting results than the Normal distribution (Kuester et al. 2006, Lee et al. 2008). For this purpose, the standardized Student-t distribution can been modified to be become a skew distribution. Fernández and Steel (1998) proposed introducing skewness into any unimodal and symmetric distribution. Applying their method to the Student-t distribution gives the resulting probability density function for the **skewed Student-t** distribution:

$$g(\epsilon_t|\xi\nu) = \begin{cases} \frac{2}{\xi+1/\xi}\varrho f[\xi(\varrho\epsilon_t + \varpi)|\nu] & \text{if } \epsilon_t < -\varpi/\varrho \\ \frac{2}{\xi+1/\xi}\varrho f[(\varrho\epsilon_t + \varpi)/\xi|\nu] & \text{if } \epsilon_t \geq -\varpi/\varrho \end{cases}, \tag{15}$$

where $f(\cdot)$ is the probability density function of the standardized Student-t distribution in equation 14 and $\xi \in \mathbb{R}^+$ is the skew parameter, implying symmetry for $\xi = 1$.

Another useful distribution for financial assets is the **Skewed Generalized Error** distribution which belongs to the exponential family and is a transformation of the Generalized Error distribution (Theodossiou, 2000). Its density function can be described as:

$$f(\epsilon_t) = \frac{k^{1-1/k}}{2\psi}\Gamma(\frac{1}{k})^{-1}exp(-\frac{1}{k}\frac{|\epsilon_t - m|^k}{(1 + sng(\epsilon_t - m)\lambda)^k\psi^k}, \tag{16}$$

where m is the mode of ϵ_t, ψ is a scaling constant derived from the standard deviation of ϵ_t, λ is a skewness parameter, k is a kurtosis parameter, $\Gamma(\cdot)$ is the gamma function and sgn is the sign function:

$$sng(\epsilon_t - m) = \begin{cases} -1 & \text{if } \epsilon_t - m < 0 \\ 1 & \text{if } \epsilon_t - m > 0 \end{cases} \tag{17}$$

k controls the heavy tails and peakness of the distribution, while λ controls the skewness (Theodossiou, 2000). As k increases, the density becomes flatter. For the original version of the Generalized Error distribution, it tends towards the Normal distribution for the case when $k = 2$ (Ghalanos, 2018). The skewed Generalized Error distribution has already shown to provide better Value at Risk forecasts in GARCH modelling for the traditional financial market than the more common distribution assumptions (Lee et al., 2008). Since it has been found that GARCH type models paired with Student-t and skewed distributions deliver better forecasting results for financial data, they are going to be utilized next to the Normal distribution.

4.2 Estimation

GARCH models are usually estimated via Maximum-Likelihood. According to Bollerslev (1986) the GARCH-(m, s) process can be written as a regression model with a_t being the innovations in a linear regression:

$$a_t = y_t - x_t' b, \tag{18}$$

where y_t is the dependent variable, x_t a vector of explanatory variables and b a vector of unknown parameters.
Then, if $z_t' = (1, a_{t-1}^2, ..., a_{t-m}^2, \sigma_{t-1}^2, ..., \sigma_{t-s}^2)$, $\omega' = (\alpha_0, \alpha_1, ..., \alpha_s, \beta_1, ..., \beta_s)$ and $\theta \in \Theta$ with $\theta = (b', \omega')$ and Θ being a subspace of the Euclidian space such that the second moments of a_t are finite. The true parameters are denoted $\theta_0 \in \Theta$. Bollerslev (1986) then rewrites the model as:

$$a_t | F_{t_1} \sim N(0, \sigma_t), \tag{19}$$

$$\sigma_t = z_t' \omega, \tag{20}$$

under normality assumption of the distribution for the innovations. The Log Likelihood can then be written as:

$$L_T(\theta) = T^{-1} \sum_{t=1}^{T} l_t(\theta), \tag{21}$$

$$l_t(\theta) = -\frac{1}{2} log(\sigma_t) - \frac{1}{2} a_t^2 \sigma_t^{-1}. \tag{22}$$

After differentiating with respect to ω, we receive:

$$\frac{\partial l_t}{\partial \omega} = \frac{1}{2} \sigma_t^{-1} \frac{\partial \sigma_t}{\partial \omega} \left(\frac{a_t^2}{\sigma_t} - 1 \right), \tag{23}$$

$$\frac{\partial^2 l_t}{\partial \omega \partial \omega'} = \left(\frac{a_t^2}{\sigma_t} - 1 \right) \frac{\partial}{\partial \omega'} \left[\frac{1}{2} \sigma_t^{-1} \frac{\partial \sigma_t}{\partial \omega} \right] - \frac{1}{2} \sigma_t^{-2} \frac{\partial \sigma_t}{\partial \omega'} \frac{a_t^2}{\sigma_t}, \tag{24}$$

with

$$\frac{\partial \sigma_t}{\partial \omega} = z_t + \sum_{j=1}^{s} \beta_i \frac{\partial \sigma_{t-i}}{\partial \omega}. \tag{25}$$

The Fisher's information matrix for ω can be estimated only by the sample analogue of the last term in equation (24) since the conditional expectation of the first term is zero.

Differentiating with respect to b parameters yields:

$$\frac{\partial l_t}{\partial b} = a_t x_t \sigma_t^{-1} + \frac{1}{2}\sigma_t \frac{\partial \sigma_t}{\partial b}\left(\frac{a_t^2}{\sigma_t} - 1\right), \tag{26}$$

$$\frac{\partial^2 l_t}{\partial b \partial b'} = -\sigma_t^{-1} x_t x_t' - \frac{1}{2}\sigma_t^{-2}\frac{\partial \sigma_t}{\partial b}\frac{\partial \sigma_t}{\partial b'}\left(\frac{a_t^2}{\sigma_t}\right) - 2\sigma_t^{-2} a_t x_t \frac{\partial \sigma_t}{\partial b} + \left(\frac{a_t^2}{\sigma_t} - 1\right)\frac{\partial}{\partial b'}\left[\frac{1}{2}\sigma_t^{-1}\frac{\partial \sigma_t}{\delta b}\right], \tag{27}$$

with

$$\frac{\partial \sigma_t}{\partial b} = -2\sum_{j=1}^{m} \alpha_i x_{t-i} a_{t-i} + \sum_{j=1}^{s} \beta_j \frac{\partial \sigma_{t-j}}{\partial b}. \tag{28}$$

Since there is no closed-form solution for the Maximum-Likelihood estimates, there is a need for an iterative procedure. Bollerslev (1986) names the algorithm by Berndt, Hall, Hall, and Hausman (1974). To find the true parameter θ_0, let $\theta^{(i)}$ denote the estimates after the ith iteration. Then $\theta^{(i+1)}$ for the $i+1$th iteration is calculated by:

$$\theta^{(i+1)} = \theta^{(i)} + \lambda_i \left(\sum_{t=1}^{t} \frac{\partial l_t}{\partial \theta}\frac{\partial l_t}{\partial \theta'}\right)^{-1} \sum_{t=1}^{T} \frac{\partial l_t}{\partial \theta}, \tag{29}$$

with λ being a pre-defined variable to maximize the likelihood function in the given direction. The Maximum-Likelihood estimator $\hat{\theta}$ is consistent for θ_0 and asymptotically normal with mean θ_0 and covariance matrix $\mathcal{F}^{-1} = \mathbb{E}(\partial^2 l_t / \partial\theta\partial\theta')^{-1}$.
For this thesis the package *rugarch* (Ghalanos, 2018) in R is used to estimate the different univariate GARCH models.

4.3 Diagnostic Checking

For a correctly specified GARCH model the standardized residuals

$$\epsilon_t = \frac{a_t}{\sigma_t}, \tag{30}$$

are randomly independent identically distributed with mean zero and variance one. To check whether the volatility of a return series has been modelled correctly, the Ljung Box-Test can be applied to the standardized squared residuals to detect remaining ARCH effects. The distribution assumption can be validated by quantile-to-quantile plots and parameters for skewness and kurtosis (Tsay, 2005).
Another method to detect for remaining serial dependence in the residuals is the ARCH LM test by Engle (1982). The alternative hypothesis assumes remaining autocorrelation for the standardized residuals:

$$H_1 : \epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + ... + \alpha_m \epsilon_{t-m}^2 + u_t, \tag{31}$$

with u_t as an error term for the autoregressive model of ϵ_t^2. Under the H_0 there are no remaining ARCH effects in the residuals, thus

$$H_0 : \alpha_0 = \alpha_1 = ... = \alpha_m = 0, \tag{32}$$

applies. The F test statistic is asymptotically distributed as χ^2 with m degrees of freedom under the null hypothesis.

4.4 Value at Risk

Regular model diagnostics are useful to detect for violations in the model assumptions and error terms. But the fit of GARCH models should also be evaluated by its forecasting performance. An important aspect is how well the model can determine potential portfolio losses.
Value at Risk is a popular concept of market downside risk. It was first introduced in 1994 by JP Morgan in a technical document which revealed their methodologies on financial risk measurement (Xu and Chen, 2012). Generally, the VaR covers the losses of a portfolio return distribution by stating that the portfolio loss will exceed a certain threshold with the small probability α. Technically, the Value at Risk for a certain period $t + h$ at the α-level can be described as the negative α-quantile of

the conditional return distribution:

$$VaR_{t+h}^{\alpha} := -Q_{\alpha}(r_{t+h}|F_t) = -\inf_x\{x \in \mathbb{R} : P(r_{t+h} \leq x|F_t) \geq \alpha\}, \quad 0 < \alpha < 1, \tag{33}$$

with $Q_{\alpha}(\cdot)$ denoting the quantile function and F_t being the past information available up to time t. There are several approaches to determine the distribution of r_t in equation (33). In the case of GARCH models, it is calculated by using $r_t = \mu_t + \sigma_t \epsilon_t$, with σ_t being modelled by a GARCH process (Kuester et al., 2006).
To evaluate the predictive performance of volatility models, Christoffersen (1998) has set up a framework to evaluate out-of-sample interval forecasts. Therefore, he defines the sequence of violations $H_t = I(r_t < -VaR_t)$ which has to be independent from any variable in the information set F_{t-1}. The VaR forecast is efficient with respect to F_{t-1} if $\mathbb{E}(H_t|F_{t-1}) = \lambda$. Assuming efficiency, H_t follows the Bernoulli distribution: $H_t|F_{t-1} \overset{iid}{\sim} B(\lambda)$, for $t = 1, ..., T$ (Kuester et al., 2006).
This leads to the first test of unconditional coverage:

$$H_0 : \mathbb{E}(H_t) = \lambda \quad vs. \quad H_0 : \mathbb{E}(H_t) \neq \lambda. \tag{34}$$

The likelihood ratio test statistic

$$LR_{uc} = 2\left[L(\hat{\lambda}, H_1, ..., H_T) - L(\lambda, H_1, ..., H_T)\right] \overset{H_0}{\sim} \chi_1^2, \tag{35}$$

tests for the correct number of unconditional violations, with $L(\cdot)$ denoting the log likelihood. The ML-estimator $\hat{\lambda}$ is the ratio of the number of violations to the total number of observations.
The test of independent violations checks for violation clusters in the VaR interval forecasts. Under the null hypothesis, a violation at t has no influence on the violation at $t+1$. The test statistic

$$LR_{cc} = 2\left[L(\hat{\Pi}, H_2, ..., H_T|H_1) - L(\Pi, H_2, ..., H_T|H_1)\right] \overset{H_0}{\sim} \chi_2^2, \tag{36}$$

tests for the conditional coverage of violations as well as the correct number of unconditional violations, with Π denoting a first-order Markov-Chain model corresponding to the independence of violations (Christoffersen, 1998). Both of the tests provide an evaluation on the performance of the Value at Risk forecasts.

Previous research has shown that one-day-ahead Value at Risk forecasts provide

better results of coverage than, for example, weekly forecasts (Kole et al., 2017). Therefore, the following analysis will focus on VaR forecasts for the next day.
The GARCH models described in section (4.1) are going to be utilized to forecast the return distribution function of the five cryptocurrency series. Also the aggregated portfolio of all five currencies is going to be modelled. Kole et al. (2017) show that lower levels of aggregation lead to better forecasting results. This might be the case, because an extreme development of one asset can lead to biased forecasting results for the whole portfolio. However, this method is going to be applied to make the univariate GARCH models comparable to a multivariate Dynamic Correlation model which forecasts the aggregated return series also by means of the dynamic conditional correlation between the innovations of the different return series.

5 Univariate Empirical Results

In this chapter, the historical price developments of the used cryptocurrencies will be described, next to the stylized facts of the logarithmic returns. It will be shown that the return series show typical characteristics of financial time series and, furthermore, are appropriate for GARCH-type modelling. Next, the model fit of the different conditional heteroscedasticity models will be compared due to their model fit and VaR forecasting performance.

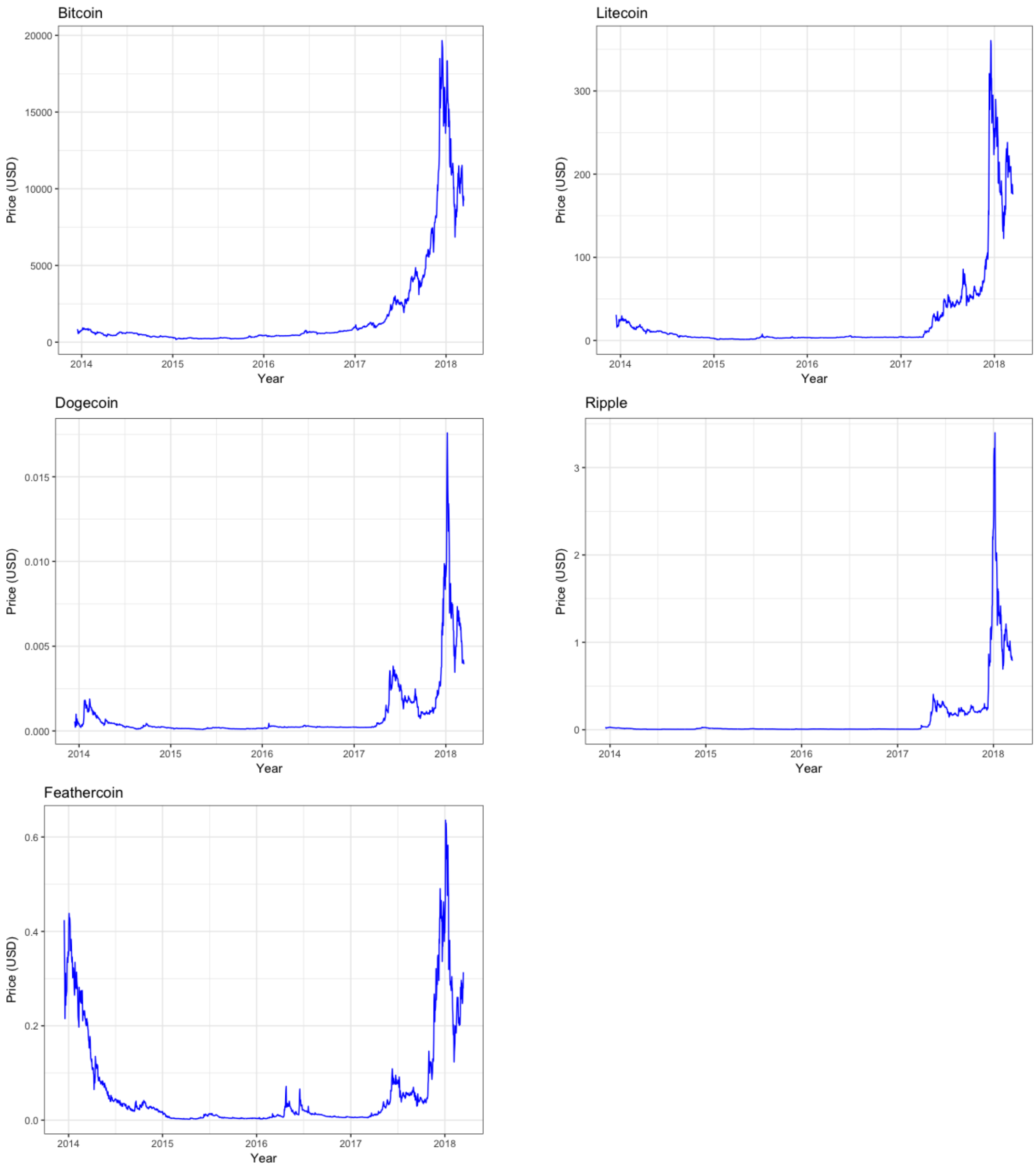


Figure 1: The evolution of daily closing prices in \$US for Bitcoin, Litecoin, Dogecoin, Ripple and Feathercoin from 17^{th} of December 2013 until 14^{th} of March 2018.

The most appropriate models are chosen to implement them for the first stage estimation of the Dynamic Conditional Correlation model in the next section.

The data used are daily closing prices for the five cryptocurrency series Bitcoin, Litecoin, Dogecoin, Feathercoin and Ripple. It is publicly available online at Coingecko.com. The evolution of prices since December 2013 until March 2018 is shown in figure (1). Dogecoin and Feathercoin have just emerged in 2013 and experience a price increase in late 2013 that fades out in 2014. The most noticeable development is the enormous growth in prices during the hype in mid 2017 which is followed by a short stagnation and another price increase in 2018. All crypto series experienced an extreme multiplication of their value, just before an immense price drop by the end of 2017. Figure (1) thus implicates that the price developments of the five different coins are driven by mutual determinants.

	Bitcoin	Litecoin	Dogecoin	Feathercoin	Ripple	Portfolio
Market Cap in USD	155,921,695,042	9,761,938,418	447,366,961	44,076,840	30,826,633,076	-
n. obs	1544	1544	1544	1544	1544	1544
Minimum	-25.18	-54.72	-94.04	-94.00	-91.34	-52.40
Maximum	28.71	51.44	84.33	72.76	88.13	36.83
Mean	0.15	0.11	0.13	-0.02	0.21	0.14
Median	0.19	-0.08	-0.36	-0.69	-0.17	-0.04
Variance	17.51	38.03	76.22	126.05	59.05	39.05
Stdev	4.18	6.17	8.73	11.23	7.68	6.24
Skewness	-0.42	0.34	0.72	0.51	0.84	-0.57
Kurtosis	6.91	14.66	29.88	9.09	34.08	9.09
Jarque-Bera	3128.9	13901	57742	5399.4	75116	5420.7

Table 1: Summary Statistics for daily log returns $\times$ 100 of cryptocurrencies.
Log returns are calculated using: $r_t = 100 \times ln(P_t/P_{t-1})$.
Returns are observed until 14^{th} of March 2018.
Market cap is captured at 14^{th} of March 2018.
Jarque-Bera-Test checks for deviation from normality (skewness S different from zero and kurtosis K different from 3): $JB = T(S/6 + (k-3)^2/24)$, is distributed as $\mathcal{X}^2(2)$ with 2 degrees of freedom. Its critical value at the five-percent level is 5.99 and at the one-percent it is 9.21.

Stylized Facts. Table (1) shows the summary statistics for the cryptocurrency log return series. The log returns are calculated by taking the natural logarithm of the ratio of two consecutive daily closing prices: $r_t = ln\frac{P_t}{P_{t-1}} \times 100$. For numerical stability in the statistical software R the returns are multiplied by 100. Next to the

individual series, there is also a return series for an aggregated portfolio calculated. The portfolio return is simply the sum of the individual returns, divided by the number of currencies: $r_{PF,t} = \frac{1}{5}\sum_{i=1}^{5} r_{i,t}$.
The data was downloaded on 14^{th} of March 2018. Since the main attempt is to analyse the dynamic correlation and multivariate volatility, the datasets are trimmed to the same length with Dogecoin being the youngest currency. This leads to an overall observation span of $T = 1544$ for the log returns. The market capitalization in US Dollar reflects the market value and trading volume of the coins. Bitcoin is the most popular currency with the highest market capitalization, followed by Ripple, Litecoin and Dogecoin. Feathercoins market capitalization is the lowest.
All currencies except Feathercoin show a positive mean log return. The median of all currencies except for Bitcoin is smaller than zero, which leads to a positive skewness in combination with a positive mean. The standard deviation of the returns is in all cases larger than the mean, which is a typical property of highly volatile financial data (Theodossiou, 1998).
It can be noted that Feathercoin has the lowest median return and minimum out of all currencies with the third highest maximum after Dogecoin and Ripple. Also it shows the highest variance. Bitcoin shows the lowest variance and is the only currency that is negatively skewed, which is congruent with the results by Catania and Grassi (2017) and Catania et al. (2018). Also negative skewness implies that more values bigger than the mean of a distribution were observed, specifically more positive returns. Gkillas and Katsiampa (2018) found Bitcoin to be the least risky coin among the five most popular cryptocurrencies. The other currencies are positively skewed, which has also been shown by Chu et al. (2017). That means that there is more mass on the left side of the density function, i.e. on the negative returns.
The kurtosis k is higher for all currencies than it should be expected for a Normal distribution ($k = 3$). This is a typical characteristic of financial asset returns since a lot of extreme values are observed and the distribution is highly peaked. The Jarque-Bera-Test for normality (Jarque and Bera, 1987) is rejected for all currencies, so the log returns deviate from the Gaussian distribution. For financial data the Central Limit Theorem - stating that the distribution of the sum of a random variable is going to converge to normality for big samples - does not apply in many cases. This is because daily log returns often show higher order moment dependencies like asymmetric volatility or conditional heteroscedasticity (Theodossiou, 1998). The Augmented-Dickey-Fuller test (Said and Dickey, 1984), testing the null hypothesis that the time series x has a unit root, is applied to all currencies and

shows that all log return series are stationary.
Figure (2) shows the histograms for the density of the log returns. To visualize the deviation from normality, the black line illustrates the theoretical Normal distribution, given the same mean and variance. It is visible that Feathercoin has the highest variance and many extreme observations, while Bitcoin has many observations around the return of zero. Dogecoin also shows more heavy-tail behaviour.
Figure (3) shows the evolution of the calculated log returns over the available observation span. It shows that Bitcoin has the least risky and volatile behaviour, while Feathercoin and Ripple have a bigger span of variation.
Since the returns for the aggregated portfolio are averaged over the five currencies, the span of the data is smaller, and standard deviation and kurtosis take values on an average level, see table (1). The histogram also shows that there are less extreme observations and more data is located around a return of zero.

Another typical characteristic of financial time series is the autocorrelation of returns. If the evolution of the price experiences an upward or downward dynamic, then the returns are positively correlated for a period of time. Figure (3) shows that for all currencies the returns exhibit periods of higher and lower volatility (volatility clusters). Volatility is extremely high in the period of the cryptocurrency hype during 2017. Figure (12) in the Appendix shows the Auto Correlation Function of the squared returns. For all currencies there is a significant correlation of returns with the preceeding days that fades out after a couple of days, with some peaks around lag $\ell = 20$ or $\ell = 30$. This structure is typical for ARCH-effects and justifies the application of GARCH-models (Tsay, 2005). Furthermore, the Ljung-Box-Test for serial correlation identifies a dependency within the first $ln(T) = ln(1544) = 7.34$ lags of the squared innovations $\hat{a}_t$ from equation (2), that were calculated by subtracting the mean return from the daily return: $\hat{a}_t = r_t - \overline{r}_t$ (Tsay, 2005).

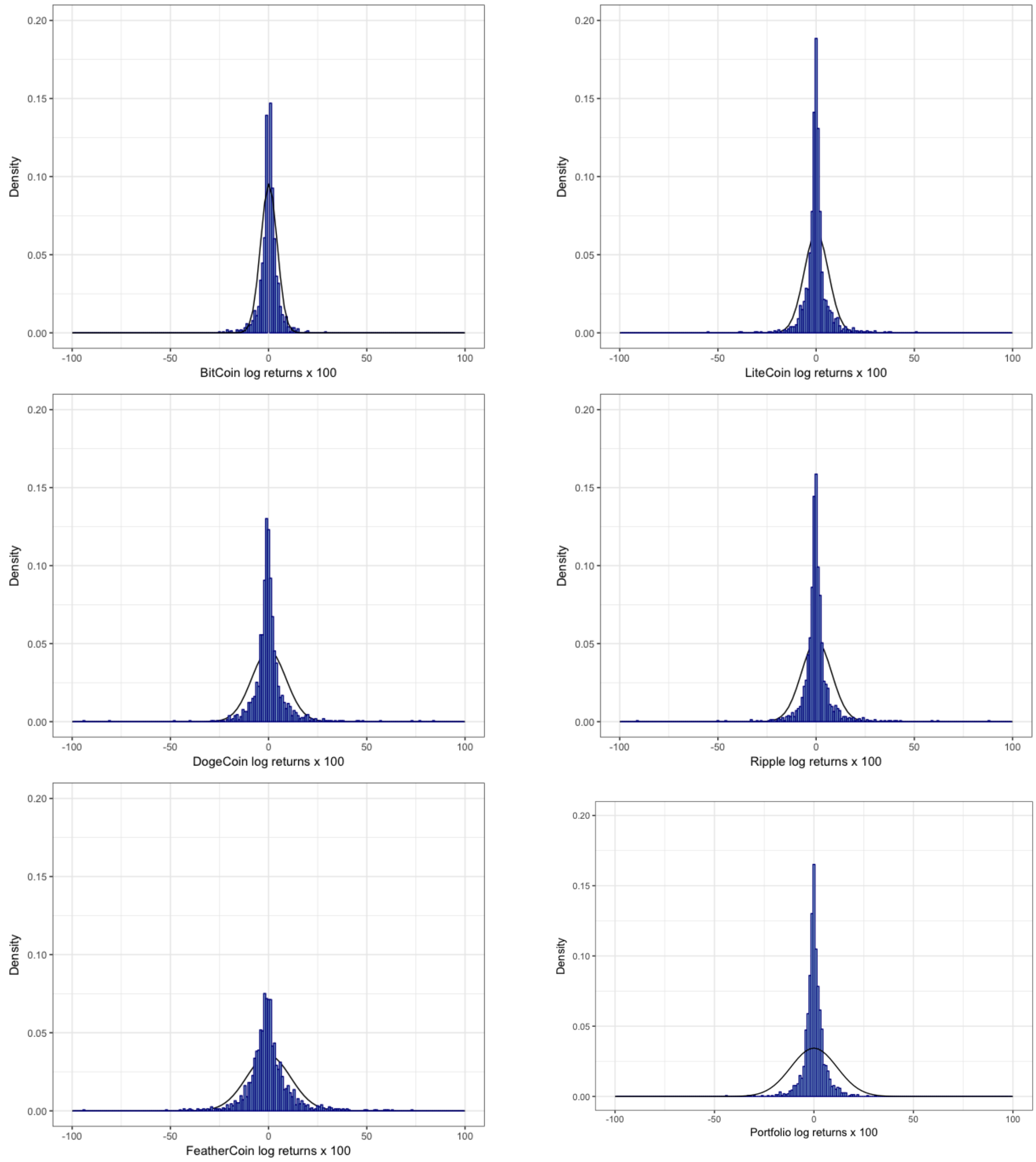


Figure 2: Histograms of the log returns of the five cryptocurrencies and the aggregated portfolio in blue.
Log returns are calculated using: $r_t = 100 \times ln(P_t/P_{t-1})$.
Black line shows the density of the Normal distribution that would occur for the empirical mean and standard deviation:
$x \sim N(\overline{r}_i, \sqrt{Var(r_i)})$.
T=1544.

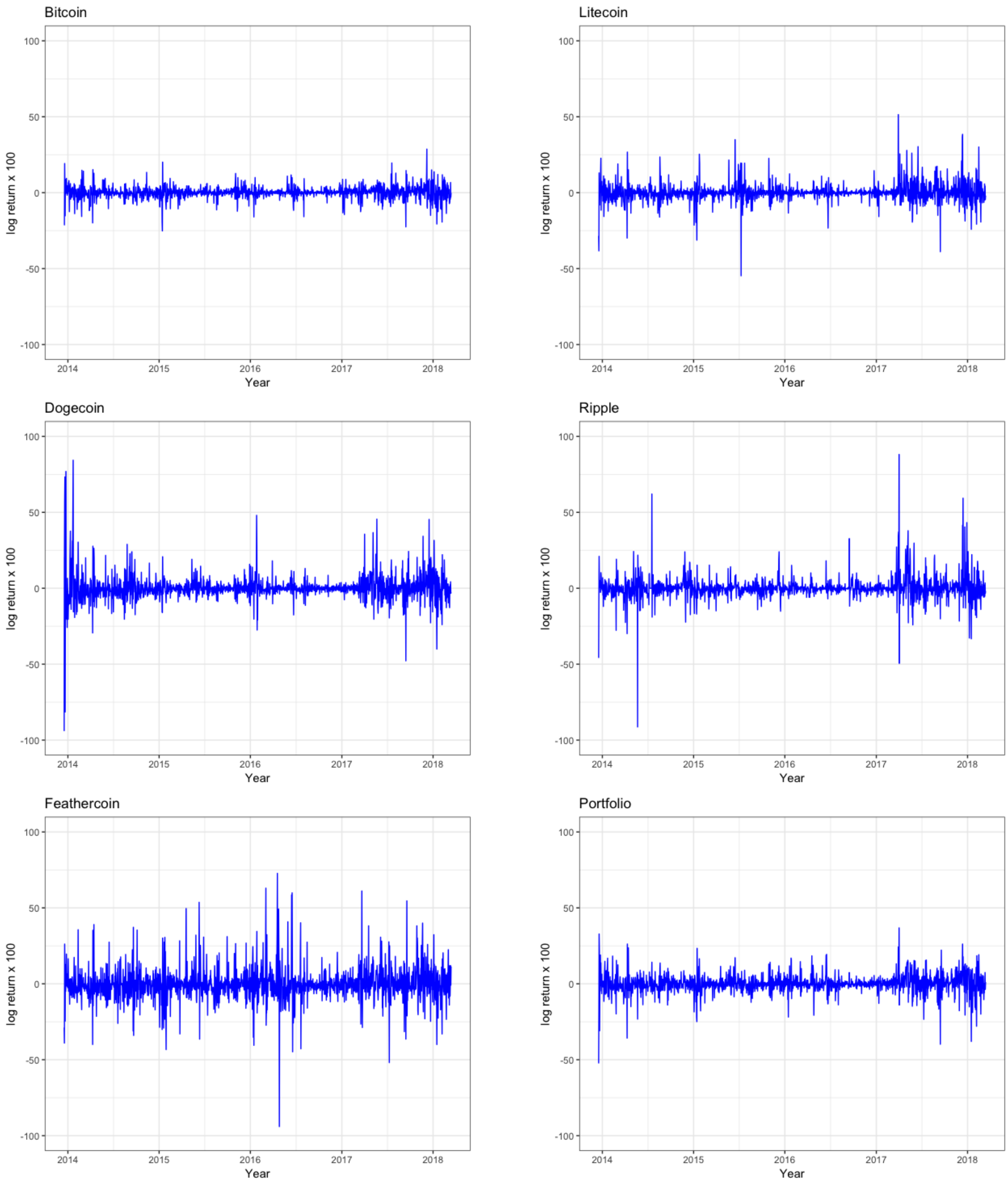


Figure 3: The evolution of log returns in \$US of Bitcoin, Litecoin, Dogecoin, Ripple, Feathercoin and the aggregated portfolio until 14^{th} of March 2018.

Univariate GARCH Models. After conditional heteroscedasticity has been detected for all five cryptocurrencies, the GARCH models discussed in chapter (4.1) are applied to the univariate log return series. Since the descriptive analysis of the returns has shown fat tails, not only the Gaussian distribution is going to be used for the residual terms but also the Student-t, skewed Student-t and skewed Generalized Error distribution.
All GARCH-type models with different distribution assumptions for the error terms are applied to the univariate time series. The fit is evaluated by diagnostic checking and information criteria. More interesting, however, is which model is able to provide the best Value at Risk forecasts. The results are going to be used in chapter (7) to model the DCC model. Additionally, the GARCH models are going to be applied to the combined portfolio time series consisting of the five currencies. The goal is to find out whether a multivariate approach incorporating the dynamic correlation improves the Value at Risk forecasts for an aggregated portfolio.

To model the mean of the time series, a simple ARMA-(1,1) process is defined for μ_t equivalent to equation (3). When defining the volatility equation (4), the first step is determine to the order of the ARCH effects. This can be done by looking at the PACF of the squared innovations a_t which can be estimated by the squared series of mean adjusted returns: $\hat{a}_t = r_t - \overline{r}_t$. For all five currencies the partial autocorrelation function reveals significances at higher order lags (around 10 to 100 days). In this case, it is more appropriate to choose the more parsimonious GARCH-model, instead of applying higher order ARCH-models. For the GARCH process usually a lower order model like the GARCH-(1,1) or GARCH-(2,1) is appropriate in most applications (Tsay, 2014).
A GARCH-(1,1) process is defined for the standard GARCH, iGARCH, iGARCH, gjr-GARCH, apARCH and csGARCH models with the Normal, Student-t, skewed Student-t, and skewed Generalized Error distribution for each logarithmic return series of the five currencies. The models are estimated via Maximum-Likelihood.

Bitcoin	norm		t		skewed t		sged	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
sGarch	5.3652	5.3860	5.1142	5.1384	5.1142	5.1418	5.0974	5.1251
iGarch	5.3639	5.3812	5.1127	5.1334	5.1127	5.1369	5.0960	**5.1202**
eGarch	5.3492	5.3734	5.0981	5.1257	5.0979	5.1290	not converged	
apArch	5.3664	5.3941	5.1012	5.1323	5.1008	5.1354	5.0960	5.1306
csGARCH	5.3388	5.36651	5.1055	5.1366	5.1058	5.1404	**5.0895**	5.1241
gjrGarch	5.3657	5.3899	5.1151	5.1428	5.1151	5.1462	5.0986	5.1297

Litecoin	norm		t		skewed t		sged	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
sGarch	6.1702	6.1910	5.5229	5.5472	5.5227	5.5504	5.5456	5.5732
iGarch	6.1736	6.1909	5.5213	5.5420	5.5210	5.5453	5.5440	5.5683
eGarch	6.1293	6.1535	5.4984	**5.5261**	**5.4973**	5.5285	not converged	
apArch	6.1650	6.1927	5.5003	5.5315	5.4985	5.5331	5.5399	5.5745
csGARCH	6.1635	6.1911	5.5075	5.5386	5.5058	5.5404	5.5376	5.5722
gjrGarch	6.1709	6.1951	5.5217	5.5494	5.5215	5.5526	5.5457	5.5768

Ripple	norm		t		skewed t		sged	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
sGarch	6.4823	6.5031	5.8276	5.8519	5.8289	5.8566	5.8442	5.8719
iGarch	6.4811	6.4984	5.8262	5.8470	5.8275	5.8517	5.8427	5.8669
eGarch	6.4740	6.4982	5.8259	5.8536	5.8266	5.8578	not converged	
apArch	6.4674	6.4951	5.7947	5.8258	5.7951	5.8297	5.8354	5.8700
csGARCH	6.3950	6.4227	**5.7937**	**5.8248**	5.7944	5.8290	5.8139	5.8485
gjrGarch	6.4622	6.4864	5.8286	5.8563	5.8298	5.8610	5.8446	5.8757

Dogecoin	norm		t		skewed t		sged	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
sGarch	6.3902	6.4109	6.0487	6.0729	6.0414	6.0691	6.0592	6.0869
iGarch	6.3885	6.4058	6.0472	6.0680	6.0399	6.0641	6.0577	6.0820
eGarch	6.3536	6.3779	6.0333	6.0610	6.0247	6.0559	not converged	
apArch	6.3907	6.4149	6.0484	6.0761	6.0413	6.0724	6.0588	6.0899
csGARCH	6.3029	6.3306	6.0276	6.0588	**6.0209**	**6.0555**	6.0360	6.0706
gjrGarch	6.3907	6.4149	6.0484	6.0761	6.0413	6.0724	6.0588	6.0899

Feathercoin	norm		t		skewed t		sged	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
sGarch	7.4188	7.4396	7.1496	7.1738	7.1426	7.1703	7.1356	7.1633
iGarch	7.4183	7.4356	7.1483	7.1690	7.1413	7.1655	7.1368	**7.1610**
eGarch	7.3963	7.4205	7.1387	7.1664	7.1314	7.1625	not converged	
apArch	7.3838	7.4115	7.1381	7.1692	7.1309	7.1655	**7.1268**	7.1614
csGARCH	7.4205	7.4481	7.1572	7.1884	7.1491	7.1837	7.1350	7.1696
gjrGarch	7.4122	7.4364	7.1486	7.1763	7.1414	7.1725	7.1340	7.1652

Portfolio	norm		t		skewed t		sged	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
sGarch	6.2541	6.2749	6.0194	6.0436	6.0208	6.0485	6.0222	6.0499
iGarch	6.2529	6.2702	6.0186	6.0394	6.0194	6.0437	6.0209	6.0451
eGarch	6.2332	6.2574	**6.0051**	**6.0327**	6.0053	6.0364	not converged	
apArch	6.2544	6.2821	6.0098	6.0409	6.0103	6.0449	6.0194	6.0540
csGARCH	6.1949	6.2226	6.0058	6.0369	6.0065	6.0411	6.0061	6.0407
gjrGarch	6.2563	6.2805	6.0181	6.0458	6.0189	6.0500	6.0211	6.0523

Table 2: AIC and BIC for the estimated GARCH-type models.
μ_t is modelled via an ARMA-(1,1) process. σ_t is modelled via a GARCH-type process of order (1,1). T=1544.
Lowest AICs and BICs per group are written in **bold** letters.

Table (2) shows the AIC and BIC for the fitted GARCH-type models with different distribution assumptions for the innovations. The models that performed best for every currency or the portfolio according to the Akaike or Bayes Information Criterion are written in bold letters. As a first result, it shows that the Gaussian distribution for the error terms is outperformed by its heavy-tailed alternatives. AIC and BIC indicate a lower fit here over all different currencies. The eGARCH in combination with the skewed Generalized Error distribution showed convergence problems for all currencies that remained even after several adjustments of solver options.

For Bitcoin and Feathercoin, the skewed Generalized Error distribution delivers the best fit for almost all models. AIC is lowest for the csGARCH, nearly followed by the iGARCH for Bitcoin and the apARCH for Feathercoin. This is supported by the findings by Katsiampa (2017), who also showed that the component-standard GARCH provides the best fit to model Bitcoins volatility according to information criteria. The quantile-to-quantile plot and the ACF of the squared residuals indicate a good fit. For Litecoin, the skewed Student-t distribution seems to be most appropriate and works best along with the eGARCH model. For Dogecoin, the eGARCH, csGARCH and apARCH deliver the best information criteria with the apARCH showing most reasonable plots in model diagnostics on the error terms. The apARCH with Student-t distribution works best for Ripple according to residual plots. However, the skew parameter for the equivalent model with the skewed Student-t distribution is significant. For the aggregated portfolio, the information criteria show that the return series is best modelled via the Student-t distribution with an eGARCH, however, the shape parameter for the corresponding model with a skewed Student-t distribution is also significant and reports good information criteria and residual plots.

In general, the univariate GARCH models show that there is a need for skewed distributions to model the volatility process of the currencies. The apARCH, csGARCH and eGARCH models seem to work best according to AIC and BIC, however, there is usually only a slight difference between the models if the right distribution for error terms is chosen. Bitcoin is most accurately modelled via a csGARCH, which is supported by the findings of Katsiampa (2017). Bitcoin and Feathercoin show a good model fit if the innovations are modelled via the skewed Generalized error distribution. It should be noted that the apARCH and eGARCH are more parsimonious than the csGARCH which contains a time-varying intercept for the conditional variance. However, the models should also be evaluated via their performance on Value at Risk forecasting. The choice of the right distribution assumption for the error terms is of great importance here.

Value at Risk. Now it is interesting to find out which of the models produce the best Value at Risk forecasts.

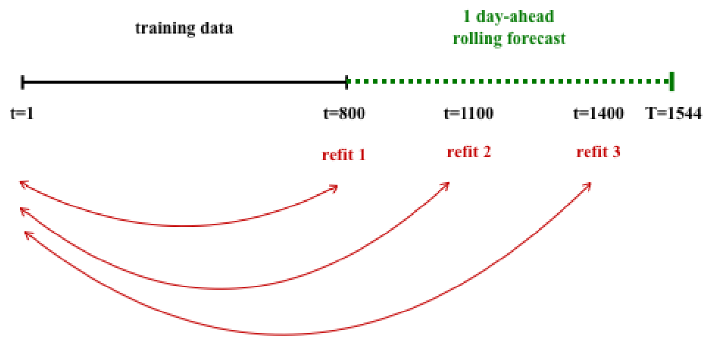


Figure 4: Outline for the 1-day-ahead forecast.
The time line shows the training data set of length T=800 in black and the rolling out-of-sample forecast of length T=744 in green.
Red arrows indicate the observations included in the three parameter re-estimations (recursive window).

Since the altcoin market is quite young, the length of the return series is short in comparison to other traditional currencies or indices. Therefore, it is more appropriate to estimate the rolling forecast based on a recursive window. That means all past observations are included in the estimation of the current parameters in contrast to a moving window where all the previous data is used for the first estimation and then moved by a pre-defined length for every forecast.
Figure (4) shows the outline for the forecast. The rolling forecast starts at $t = 800$, which leaves 744 one-day-ahead forecasts. The parameters are refitted every 300 days, thus the model parameters are refitted three times. The red arrows show the past observations that are included for the estimation of the model parameters. Within the refitting period, the parameters are fixed but data is updated for every trading day.

Table (3) reports the backtesting results for the Value at Risk forecast. VaR is calculated for every GARCH model and distribution at the 1%-, and 5%-level and evaluated via VaR backtesting described in section (4.4). The percent violations show how many times the returns dropped below the VaR for $\alpha = 0.01$ and $\alpha = 0.05$ predicted by the model in relation to the total number of forecasts. Additionally, the p-values for the tests of conditional and unconditional coverage are reported. The models where the tests for conditional or unconditional coverage were rejected at the 95% confidence level are written in bold.

VaR 1%

Bitcoin	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	2.3	**0.006**	**0.017**	1.3	0.37	0.584	0.9	0.87	0.923	0.8	0.583	0.819
iGarch	2.3	**0.003**	**0.007**	1.2	0.578	0.767	0.9	0.87	0.923	0.8	0.583	0.819
eGarch	2.3	**0.003**	**0.007**	0.7	0.339	0.612	0.5	0.165	0.373	not converged		
apArch	2.4	**0.001**	**0.003**	0.4	0.063	0.175	0.4	0.063	0.175	0.5	0.165	0.373
csGarch	2	**0.014**	**0.037**	1.3	0.37	0.585	1.5	0.221	0.41	0.8	0.853	0.819
gjrGarch	2.2	**0.006**	**0.017**	1.2	0.578	0.767	0.9	0.87	0.923	0.7	0.339	0.612

Litecoin	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	0.7	0.339	0.612	0.9	0.87	0.923	1.1	0.838	0.898	0.5	0.165	0.373
iGarch	0.8	0.535	0.819	0.9	0.87	0.923	1.1	0.838	0.898	0.5	0.165	0.373
eGarch	0.5	0.165	0.273	0.8	0.583	0.819	0.8	0.583	0.819	not converged		
apArch	0.7	0.339	0.612	0.5	0.165	0.373	0.7	0.339	0.612	0.5	0.165	0.373
csGarch	0.9	0.87	0.923	1.2	0.878	0.787	1.6	0.123	0.125	0.5	0.165	0.373
gjrGarch	0.7	0.339	0.612	1.1	0.838	0.898	1.1	0.383	0.895	0.5	0.165	0.373

Ripple	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	0.7	0.339	0.612	0.8	0.583	0.819	0.8	0.583	0.819	0.4	0.063	0.175
iGarch	0.7	0.339	0.612	0.9	0.87	0.923	0.8	0.583	0.819	0.5	0.165	0.373
eGarch	1.3	0.37	0.584	0.5	0.165	0.373	0.5	0.165	0.373	not converged		
apArch	3.6	**0.000**	**0.000**	0.7	0.339	0.612	0.7	0.339	0.612	0.5	0.165	0.373
csGarch	2.4	**0.001**	**0.003**	0.7	0.339	0.612	0.7	0.339	0.612	0.8	0.583	0.819
gjrGarch	1.2	0.578	0.767	0.8	0.583	0.819	0.8	0.583	0.819	0.5	0.165	0.373

Dogecoin	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	0.9	0.87	0.923	0.8	0.583	0.819	0.8	0.538	0.819	0.8	0.538	0.819
iGarch	0.9	0.87	0.923	0.8	0.583	0.819	0.8	0.538	0.819	0.8	0.538	0.819
eGarch	1.1	0.838	0.898	0.7	0.339	0.612	0.8	0.538	0.819	not converged		
apArch	0.8	0.583	0.819	0.5	0.165	0.273	0.8	0.538	0.819	0.8	0.538	0.819
csGarch	0.9	0.87	0.923	1.1	0.838	0.898	1.3	0.538	0.819	0.9	0.87	0.923
gjrGarch	1.1	0.838	0.898	0.8	0.583	0.819	0.8	0.538	0.819	0.8	0.538	0.819

Feathercoin	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	1.3	0.37	0.584	0.8	0.538	0.819	1.1	0.838	0.898	1.1	0.838	0.898
iGarch	1.2	0.578	0.767	0.8	0.538	0.819	1.1	0.838	0.898	0.9	0.87	0.932
eGarch	1.2	0.578	0.767	0.5	0.165	0.373	0.9	0.87	0.932	not converged		
apArch	1.3	0.37	0.584	0.7	0.339	0.612	0.8	0.538	0.819	1.1	0.838	0.898
csGarch	1.3	0.37	0.584	0.8	0.583	0.819	1.1	0.838	0.898	1.1	0.838	0.898
gjrGarch	1.5	0.221	0.400	0.8	0.583	0.819	1.1	0.838	0.898	1.1	0.838	0.898

Portfolio	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	1.7	0.064	0.143	0.7	0.339	0.612	1.1	0.838	0.898	0.8	0.583	0.819
iGarch	1.6	0.123	0.25	0.7	0.339	0.612	1.1	0.383	0.898	0.7	0.339	0.612
eGarch	1.5	0.221	0.4	0.5	0.165	0.373	0.9	0.87	0.923	not converged		
apArch	1.9	**0.031**	0.075	0.5	0.165	0.373	0.7	0.339	0.612	0.5	0.165	0.373
csGarch	1.7	0.064	0.143	1.5	0.221	0.400	1.9	0.031	0.075	1.5	0.221	0.400
gjrGarch	1.7	0.064	0.143	0.8	0.583	0.819	1.2	0.578	0.767	0.8	0.583	0.819

VaR 5%

Bitcoin	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	5.1	0.893	0.142	6.2	0.153	**0.028**	6	0.203	**0.027**	5	0.973	0.118
iGarch	4.8	0.839	0.289	6.2	0.153	**0.028**	6	0.203	**0.027**	4.8	0.839	0.095
eGarch	5.1	0.893	0.142	6.2	0.203	0.027	5.8	0.341	0.078	not converged		
apArch	4.8	0.839	0.289	5.6	0.428	0.074	5.6	0.428	0.074	4.7	0.709	0.073
csGarch	5.1	0.893	0.382	6.6	0.058	**0.023**	6.6	0.058	**0.023**	4.8	0.839	0.289
gjrGarch	5	0.973	0.339	6.3	0.113	**0.009**	6	0.203	**0.027**	4.8	0.839	0.095

Litecoin	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	2.3	**0.000**	**0.001**	4.4	0.472	0.142	4.6	0.585	0.187	3.2	**0.03**	**0.015**
iGarch	1.9	**0.000**	**0.000**	4.4	0.472	0.142	4.4	0.472	0.142	3.4	0.165	0.373

eGarch	1.6	**0.000**	**0.000**	4	0.211	**0.047**	4.3	0.371	0.103	not converged		
apArch	2.6	**0.001**	**0.003**	4	0.211	0.16	4.3	0.371	0.103	3.2	**0.018**	**0.029**
csGarch	2.7	**0.002**	**0.006**	5.6	0.428	0.074	6.6	0.058	**0.023**	3.2	**0.018**	**0.029**
gjrGarch	0.7	0.339	0.612	4.6	0.585	0.053	1.1	0.383	0.895	3.2	**0.018**	**0.029**

Ripple	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	2.6	**0.001**	**0.002**	4.3	0.371	0.585	4	0.211	0.361	2.8	**0.003**	**0.007**
iGarch	2.6	**0.002**	**0.007**	4.3	0.283	0.469	4	0.211	0.361	2.7	**0.002**	**0.004**
eGarch	2.8	**0.003**	**0.011**	4.7	0.709	0.897	4.3	0.371	0.629	not converged		
apArch	5.8	0.341	0.203	4.4	0.472	0.701	4.2	0.283	0.469	3	**0.006**	**0.02**
csGarch	4.4	0.472	0.166	5.1	0.893	0.99	4.8	0.839	0.96	3.9	0.152	0.11
gjrGarch	3.1	**0.01**	0.018	4.7	0.709	0.565	4.2	0.283	0.469	3.9	0.152	0.11

Dogecoin	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	2.7	**0.002**	**0.006**	3.9	0.152	0.109	4.7	0.709	0.073	3.5	**0.047**	0.138
iGarch	2.7	**0.002**	**0.006**	3.8	0.106	0.071	4.6	0.585	0.187	3.5	**0.047**	0.138
eGarch	2.3	**0.000**	**0.001**	3.2	**0.018**	0.059	4.6	0.585	0.808	not converged		
apArch	2.7	**0.002**	**0.006**	3.1	**0.01**	0.035	4	0.211	0.447	3.5	**0.047**	0.138
csGarch	3	**0.003**	**0.02**	4.3	0.371	0.301	5.4	0.642	0.059	4.3	0.371	0.301
gjrGarch	2.6	**0.001**	**0.003**	3.9	0.152	0.266	4.4	0.472	0.142	3.5	**0.047**	0.138

Feathercoin	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	2	**0.000**	**0.000**	3.4	**0.003**	**0.002**	4	0.211	0.16	3.1	**0.01**	**0.035**
iGarch	1.9	**0.000**	**0.000**	3.4	**0.03**	0.092	4	0.211	0.16	2.6	**0.001**	**0.002**
eGarch	2.2	**0.000**	**0.000**	2.7	**0.000**	**0.000**	3.4	**0.03**	0.092	not converged		
apArch	2	**0.000**	**0.000**	3	**0.006**	**0.02**	3.8	0.106	0.188	3.2	**0.018**	0.059
csGarch	2.3	**0.000**	**0.000**	3.5	**0.047**	0.138	4.6	0.585	0.187	3.2	0.018	0.059
gjrGarch	2.2	**0.000**	**0.000**	3.2	**0.018**	0.059	4.4	0.472	0.142	3.5	**0.047**	0.081

Portfolio	norm			t			skewed t			sged		
Model	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}	%Viol	L_{uc}	L_{cc}
sGarch	2.8	**0.003**	**0.007**	4.2	0.283	0.541	4.6	0.585	0.761	3.4	**0.030**	0.092
iGarch	2.7	**0.002**	**0.004**	4.2	0.283	0.146	4.6	0.585	0.761	3.4	**0.030**	**0.039**
eGarch	3.2	**0.018**	**0.027**	4.4	0.1472	0.705	4.6	0.585	0.761	not converged		
apArch	3.6	0.072	0.071	4.2	0.283	0.541	4.7	0.706	0.793	3.6	0.072	0.198
csGarch	3.8	0.106	0.271	5.4	0.642	0.892	6	0.203	0.439	4.6	0.585	0.761
gjrGarch	2.8	**0.003**	**0.007**	4.3	0.371	0.629	4.6	0.585	0.761	3.2	**0.018**	**0.027**

Table 3: 1%- and 5%-Value at Risk results for the univariate GARCH-type models.
1-day-ahead rolling forecast with recursive window, model parameters refitted every 300 observations. Model is built on a training data set of 800 observations, which leaves 744 out-of-sample forecasts.
% Viol: Percentage of VaR violations at $\alpha = 1\%$ and $\alpha = 5\%$.
L_{uc}: p-value for test of unconditional coverage; L_{cc}: p-value for test of conditional coverage. Values printed **bold** if $p < 0.05$.

The VaR forecasts prove some of the findings of the previous section right. The results for the Gaussian distribution of innovations are mainly outperformed by the heavy-tailed alternatives. At the 5%-VaR level, the tests for unconditional and conditional coverage are rejected for most models and currencies if the Normal distribution is applied. For all currencies, except for Bitcoin, also the Student-t and skewed Generalized Error distribution is outperformed by the skewed Student-t distribution according to the backtesting results. At the 5%-level, there are more incorrect violations found than at the 1%-level. The skewed Generalized Error distribution combined with an exponential GARCH has convergence problems that remain even after several adjustments of solver options.

The results for Bitcoin draw a different picture than for the other currencies. The Gaussian and the Student-t distribution for the error terms are outperformed by the skewed alternatives. The percentage of violations are highest for the Normal distribution of error terms at the 1% VaR level, and also both the tests for conditional and unconditional coverage are rejected for all six GARCH models at the 1% VaR level. For the Student-t distribution, L_{uc} and L_{cc} are non-significant, but the percentage of violations are still outperformed by the skew alternatives in general. At the 5-% VaR level, the test results indicate inaccurate VaR violations for both versions of the Student-t distribution in some cases. The skewed Generalized Error distribution produces the smallest number of violations in combination with a good model fit.
For Litecoin, the skewed Generalized Error distribution clearly produces the best VaR forecasts for all models with violations lower than 1%, but QQ-plots show that the SGED-distribution assumption for the error terms is not inconsiderably violated. At the 1%-VaR level, the backtests by Christoffersen (1998) are rejected in many cases. The skewed Student-t distribution produces the best results with the apARCH or eGARCH, wich both show reasonable residual plots next to a low AIC or BIC. Therefore, the more parsimonious eGARCH is chosen.
A similar picture can be seen for Ripple that seems to have a better overall forecasting performance at the 1%-VaR level for the SGED distribution, but models fail diagnostic checkings here and backtesting at the 5%-VaR. The csGARCH along with Student-t distribution reported the best AIC, however, the skewed Student-t distribution performs equally well during forecasting. The apARCH is chosen over the csGARCH since it shows fewer model violations in the residual plots and has an equal forecasting performance. It also models for a leverage effect but is more parsimonious.
For Dogecoin, the best backtesting results can be found for the skewed Student-t distribution. Again, the apARCH performs well during the forecast and shows low model violations in the residual plots, next to good values for the AIC and BIC. While table (2) indicates a good fit of the skewed Generalized Error distribution for Feathercoin, the 5%-VaR backtesting results draw a different picture. The tests for conditional and unconditional coverage is rejected in many cases with the SGED distribution. The apARCH in combination with the skewed Student-t distribution provides a better fit. For the combined portfolio return, the Student-t and skewed Student-t distribution have the best forecasting performance, with the eGARCH combined with Student-t distribution giving the best model fit. However, the shape

parameter of the skewed Student-t is significant for the eGARCH.
It can also be noted that the percentage of violations for the aggregated portfolio tends to be somewhat higher than for the individual return series on average. This finding is supported by previous research, where lower levels of portfolio aggregation provided better VaR forecasting results. This might be the case since more extreme developments in the volatility of one asset can lead to biased volatility forecasts for the whole portfolio (Kole et al., 2017).
It shows that also for the Value at risk forecasts, the Gaussian distribution model is outperformed by its skewed alternatives. This result is already very well documented in the literature, where the Normal distribution has shown to be inappropriate to model the forecasts for innovations of financial time series (Kuester et al., 2006). The analysis has shown that these findings are also valid for cryptocurrency return series. The skewed Generalized Error distribution does indeed in many cases produce a low number of VaR violations, which has already been shown by Lee et al. (2008). However, it produces worse backtesting results at the same time and shows convergence problems in combination with the eGARCH.
Since the Dynamic Conditional Correlation model requires a mutual distribution for the multivariate standardized error terms - see DCC estimation in section (6.3) - all univariate time series are going to be modelled with a skewed Student-t distribution model at the first stage estimation of the DCC.

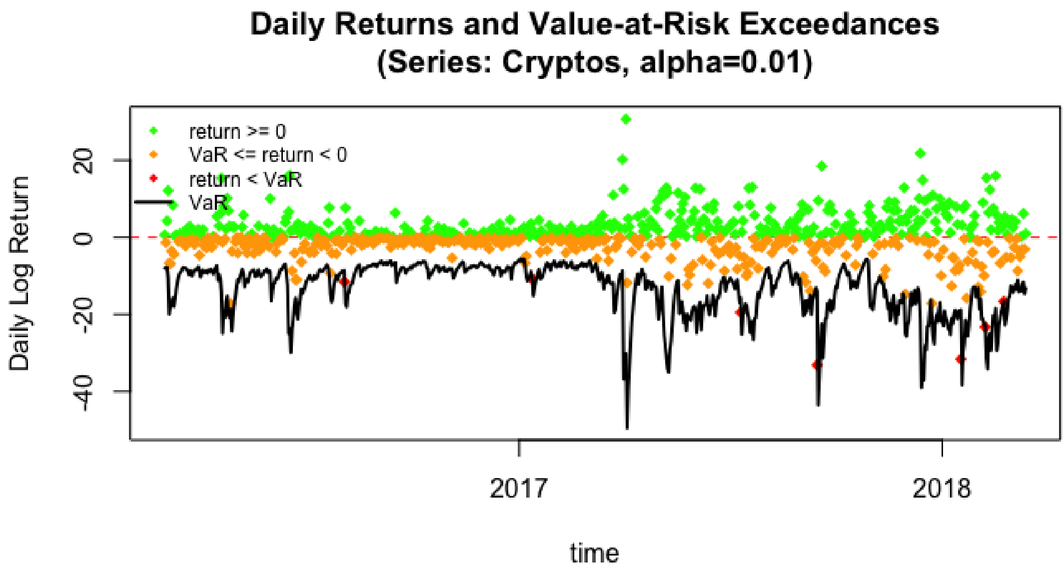
Table (4) shows the model parameters for the selected GARCH models for the DCC. The model equations have been described in section (4.1). Analysis of the Value at Risk performance and model diagnostics have shown that the Asymmetric Power ARCH and the Exponential GARCH fit the data best. Especially the apARCH that allows for leverage effects seems to perform best for the cypto currency return series. Both models take the asymmetry of positive and negative shocks into account. The market thus reacts differently to a positive shock in terms of volatility than to a negative shock. Both models are also more parsimonious than the csGARCH, while their forecasting performance is just as good. Almost all the model parameters are significant. Especially the significance of the skewness and shape parameters for the skew Student-t distribution show the need for a skew, heavy-tailed distribution model. The Ljung-Box test on the squared residuals at lag 10 and the Arch LM test at lag 5 are all non significant, therefore, the null hypothesis for no remaining ARCH effects and no autocorrelation in the residuals cannot be rejected. Figure (13) in the Appendix shows the quantile-to-quantile plots for the models in table (4). The auto correlation of the squared residuals can be seen in figure (14).

	Bitcoin	Litecoin	Ripple	Dogecoin	Feathercoin	Portfolio
Variance Model	apARCH	eGARCH	apARCH	apARCH	apARCH	eGARCH
Distribution Model	skew Student-t	skew Student-t	skew Student-t	skew Student-t	skew Student-t	Skew student-t
μ	0.0867	0.0415***	-0.2719***	-0.0945	-0.3379*	0.0136
AR(1)	0.3727***	0.4444***	0.4336***	-0.0628	0.3170***	0.3574***
MA(1)	-0.4236***	-0.5372***	-0.4784***	-0.0869	-0.4437***	-0.4060***
const α_0	0.1294*	0.0636***	0.2947***	0.1768*	1.3941	0.2528*
ARCH α_1	0.3067***	0.0428	0.4430***	0.2780***	0.3385***	0.0697*
GARCH β_1	0.8325***	0.9864***	0.6651***	0.8219***	0.6395***	0.9270***
E-GARCH γ		0.6221***				0.5134***
AP-ARCH γ	-0.0350		-0.0665***	0.0584	-0.2678**	
AP-ARCH δ	1.1477***		0.5621***	1.1527***	0.8607***	
skew	0.9543***	1.0505***	0.9902***	1.1212***	1.1174***	1.0410***
shape	2.4802***	2.0948***	2.2813***	2.7988***	2.6690***	3.0782***
Q(10)	0.4719	0.9984	1.000	1.000	0.5917	0.9993
ARCH(5)	0.5091	0.8707	0.9979	0.5093	0.4730	0.9952

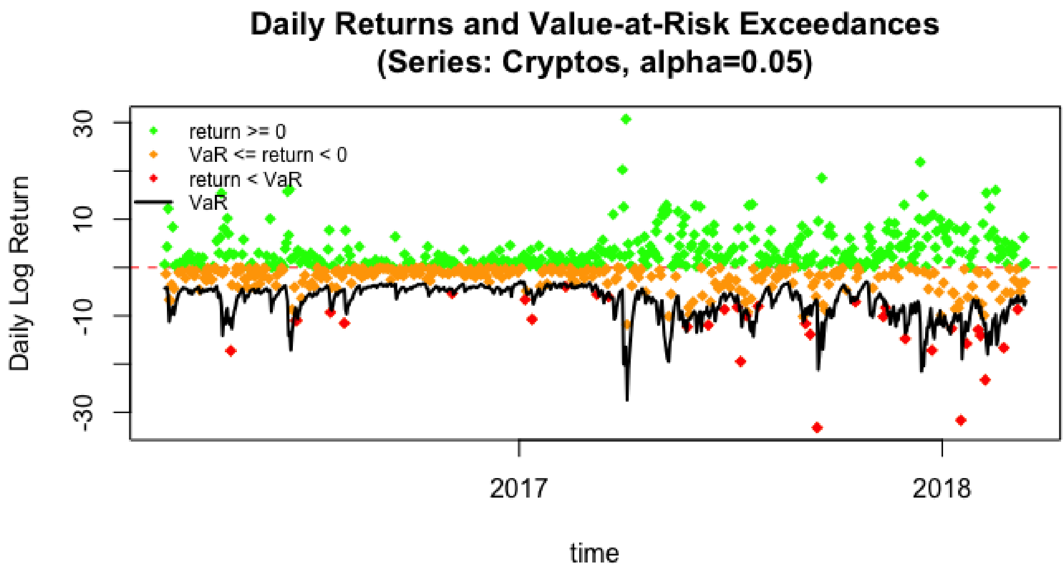
Table 4: Model parameters of the selected GARCH models.
μ_t is modelled via an ARMA-(1,1) process. T=1544.
*** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05.
Q(10): p-value of Ljung-Box test on squared standardized residuals for lag $\ell = 10$; ARCH(5): p-value for weighted ARCH LM test for lag $\ell = 5$.

The GARCH models selected in this chapter are going to be utilized to built a multivariate Dynamic Conditional Correlation model that takes account of volatility interdependencies and forecasts the downside risk of an aggregated cryptocurrency portfolio. The results are compared to the performance of the univariate approach for VaR estimation with the simple eGARCH for the aggregated portfolio.

Figure (5) shows the forecasted 1%- and 5%-Value at Risk for the combined portfolio return series of the selected eGARCH with skewed Student-t distribution with parameter re-estimation every 300 days. The red dots mark the seven days where the portfolio loss exceeds the forecasted 1%- and 5%-VaR limit. From 2017, after the general cryptocurrency hype, the expected portfolio loss becomes higher. In late 2016, the predicted volatility is more stable. The most violations are found during the cryptocurrency hype in late 2017 and early 2018. At the 5%-VaR level in figure (5b), there are more violations allowed than at the 1%-level. There are 34 violations found, especially during the hype in late 2017 and early 2018 the portfolio loss exceeds the forecasted Value at Risk.



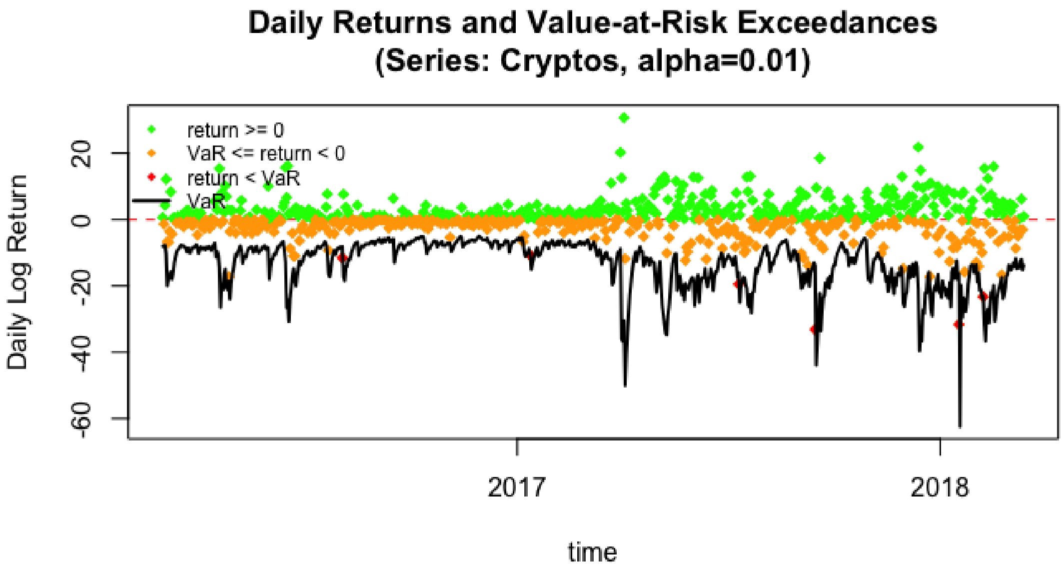
(a) 1%-Value at Risk. Parameters re-estimated every **300 days**. % Viol: 0.9, p-Values: LR_{uc} : 0.87, LR_{cc} : 0.923.



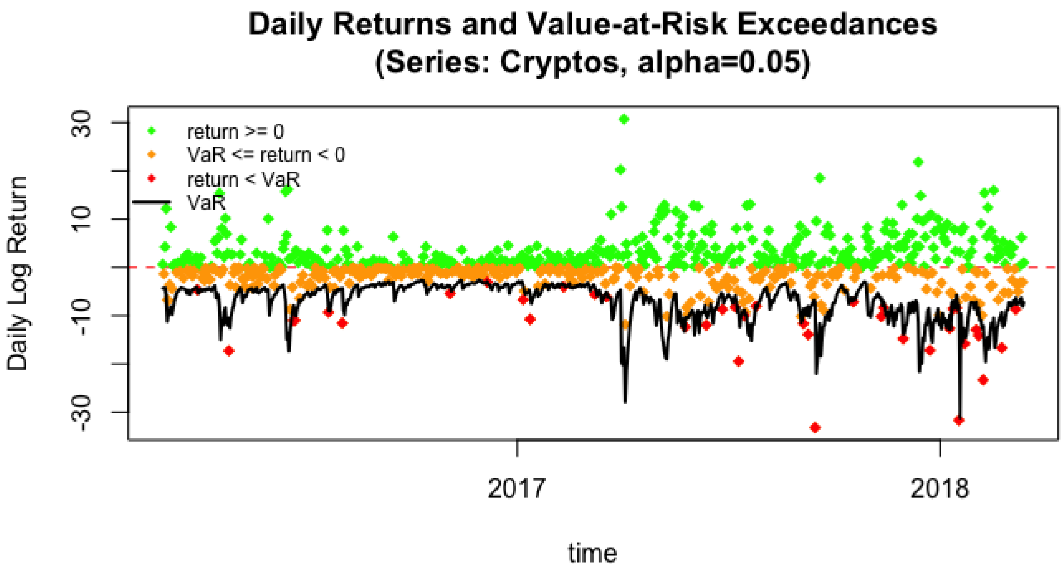
(b) 5%-Value at Risk. Parameters re-estimated every **300 days**. % Viol: 4.6, p-Values: LR_{uc} : 0.585, LR_{cc} : 0.761.

Figure 5: Value at Risk forecasts for the 1-day-ahead rolling forecast of the 1/k portfolio estimated with a univariate eGARCH with skewed Student-t distribution.
The model is built on a training data set of 800 observations, which leaves 744 out-of-sample forecasts.
Parameters are re-estimated every **300 days** with a recursive window, which leads to a total number of 3 refits.

Since the crypto market shows different patterns of volatility over time, it is worth considering to re-estimate the model parameters on a more frequent basis. For the combined portfolio, a VaR forecast with a re-estimated eGARCH every ten days is utilized. The 744 out-of-sample one-day-ahead rolling forecast with recursive window has now 75 parameter refits.
The results for Value at Risk forecasting do not improve with a more frequent refit. Figure (6a) shows the forecasted 1%-Value at Risk for the combined portfolio return series of the selected eGARCH with skewed Student-t distribution and parameter re-estimation every 10 days. There are six VaR exceedances observed at the 1%-VaR level instead of seven, which makes a violation rate of 0.8%. The tests for conditional and unconditional coverage are not rejected. A more frequent update, thus, improves the violation rate by 0.1% for the $\alpha = 1\%$ VaR level. At the 5%-VaR level, there are 38 violations found, which results in a violation rate of 5.1%. Compared to the forecast based on the refit every 300 days, this is not an improvement, see figure (5b). Both the tests for unconditional and conditional coverage are not rejected (pLR_{uc} : 0.839, pLR_{cc} : 0.799). On average, the more frequent update of model parameters has not delivered better Value at Risk forecasting results.



(a) 1%-Value at Risk. Parameters re-estimated every **10 days**. % Viol: 0.8, p-Values: LR_{uc} : 0.583, LR_{cc} : 0.819.



(b) 5%-Value at Risk. Parameters re-estimated every **10 days**. % Viol: 5.1, p-Values: LR_{uc} : 0.893, LR_{cc} : 0.731.

Figure 6: Value at Risk forecasts for the 1-day-ahead rolling forecast of the 1/k portfolio estimated with a univariate eGARCH with skewed Student-t distribution.
The model is built on a training data set of 800 observations, which leaves 744 out-of-sample forecasts.
Parameters are re-estimated every **10 days** with a recursive window, which leads to a total number of 75 refits.

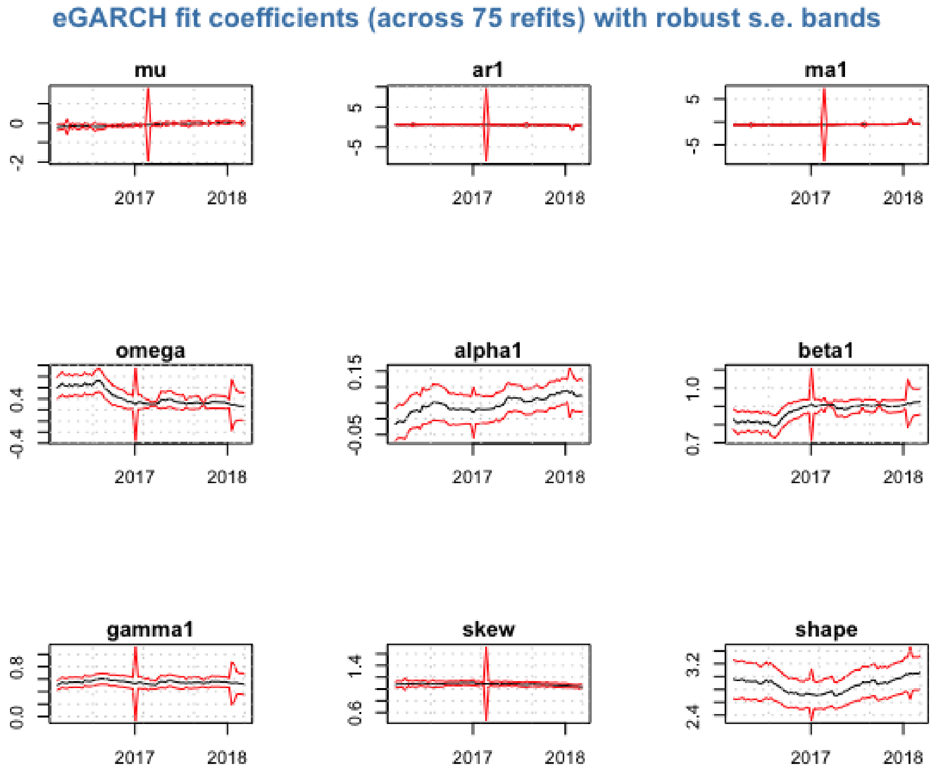


Figure 7: Model parameters for the rolling forecast of the aggregated portfolio.
eGARCH fit coefficients with robust standard error bands across 75 refits.
Parameters are re-estimated every 10 days with a recursive window on a 744 day out-of-sample rolling 1-day-ahead forecast.

Figure (7) shows the estimated eGARCH parameters across 75 refits for the out-of-sample forecasting period. Updating the model parameters more frequently during the rolling forecast has made no improvements to the VaR performance at the $\alpha = 1\%$-level. In fact, the model parameters show only slight variations over the forecasting horizon. The intercept (omega) for the conditional variance decreases over the rolling forecast, while β_1, the autoregressive coefficient for σ_t, increases over time. The shape parameter tends to increase as well, but an overlay of the robust standard error bands indicates statistical insignificance. Overall, the VaR forecasting performance does not improve by more frequent parameter updates.

One of the key findings of the univariate analysis is that (skewed) heavy-tailed distributions seem to provide a better model fit and VaR forecasts than the Normal distribution for innovations. This stands in contrast with the results by Chu et al. (2017) who found that the iGARCH with Normal innovations provides a good fit to model the volatility of seven of the top cryptocurrencies. The paper by Katsiampa

(2017) names the component standard GARCH as the most appropriate GARCH-type model for Bitcoin. This is supported by the findings of this chapter. The high performance of the apARCH and eGARCH, which allow for the leverage effect, is congruent with the results by Catania et al. (2018). They found that the volatility process of cryptocurrencies show an asymmetry to the sign of past innovations. As in the survey by Angelini and Emili (2018), it has been shown that the eGARCH also provides good forecasting results, especially for Litecoin and the aggregated portfolio. A more frequent update for the parameters does not lead to improvements regarding the accuracy of forecasting results and parameter plots show only a few changes over the whole forecasting horizon.

6 Multivariate Volatility Modelling

In the following section, the theoretical framework for multivariate volatility modelling is described. After discussing multivariate conditional heteroscedasticity in general, the Dynamic Conditional Correlation model is introduced. Finally, there is a discussion of diagnostic checkings for the DCC.

6.1 Multivariate Conditional Heteroscedasticity

Since multivariate GARCH models are time consuming to implement and estimate, it should be checked ex ante whether the data shows multivariate ARCH effects (Bauwens et al., 2006).

The previous results have shown that there is some source of conditional heteroscedasticity in the univariate time series of the cryptocurrency returns. However, the main goal here is to detect some kind of multivariate time-dependent variation and model it. The multivariate equivalent to the time-dependent σ_t of the univariate time series is the volatility matrix $\mathbf{\Sigma}_t$ of the k-dimensional time series $\boldsymbol{z}_t$. If $\boldsymbol{a}_t$ are the innovations of the multivariate time series $\boldsymbol{z}_t$, then, $\mathbf{\Sigma}_t = Cov(\boldsymbol{a}_t|F_{t-1})$ is the covariance matrix of the innovations, with F_{t-1} being the Σ-field generated by the past data (Tsay, 2013).

Similarly to the univariate case, the multivariate time series $\boldsymbol{z}_t$ can be decomposed as

$$\boldsymbol{z}_t = \boldsymbol{\mu}_t + \boldsymbol{a}_t, \tag{37}$$

where $\boldsymbol{\mu}_t = \mathbb{E}(\boldsymbol{z}_t|F_{t-1})$ is the conditional expectation of $\boldsymbol{z}_t$ given F_{t-1}. The innovation $\boldsymbol{a}_t$ is unpredictable because it is serially uncorrelated. The shock $\boldsymbol{a}_t$ can be written as

$$\boldsymbol{a}_t = \mathbf{\Sigma}_t^{1/2}\boldsymbol{\epsilon}_t, \tag{38}$$

where $[\boldsymbol{\epsilon}_t]$ is a sequence of independent and identically distributed random vectors, such that $\mathbb{E}(\boldsymbol{\epsilon}_t) = \mathbf{0}$ and $Cov(\boldsymbol{\epsilon}_t) = \boldsymbol{I}_k$ and $\mathbf{\Sigma}_t^{1/2}$ denotes the positive definite square-root matrix of $\mathbf{\Sigma}_t$.

Like in the univariate case, the conditional heteroscedasticity of a multidimensional time series can be tested. If $\boldsymbol{a}_t$ has no conditional heteroscedasticity then its conditional covariance matrix $\mathbf{\Sigma}_t$ is time-invariant. This implies that any shock at time t does not depend on the shock at $t - i$ for $i > 0$.

Dependencies could either way exist through autocorrelation within the individual return series, which has already been shown in chapter (5), or lagged correlation between the series. To test the null hypothesis $H_0 : \rho_1, = ... = \rho_m = \mathbf{0}$ for lag $\ell = 1, ..., m$ the multivariate Ljung-Box Test can be applied to the conditional correlation matrix $\boldsymbol{\rho}$. The test statistic is defined as

$$Q_k(m) = T^2 \sum_{\ell=1}^{m} \frac{1}{T-\ell} \quad tr\left(\hat{\boldsymbol{\rho}}_\ell' \hat{\boldsymbol{\rho}}_0^{-1} \hat{\boldsymbol{\rho}}_\ell' \hat{\boldsymbol{\rho}}_0^{-1}\right). \tag{39}$$

Under the null hypothesis, Q_k is asymptotically distributed as $\mathcal{X}^2_{mk^2}$ (Tsay, 2013). The test assumes the innovations $\boldsymbol{a}_t$ to be Gaussian. Therefore, Tsay proposes some robustness modifications for heavy tails in financial data to avoid misleading results. One simple procedure to reduce the effect of heavy tails is trimming away data in the upper 5% tail. Another approach is a rank based test of autocorrelation for the standardized series $e_t = \boldsymbol{a}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{a}_t - k$. Tsay (2013) combines all of those tests in his R-package MTS.

6.2 Dynamic Conditional Correlation Models

A simple class of models for multivariate volatility is the Dynamic Conditional Correlation model (DCC). It uses the covariance matrix $\boldsymbol{\Sigma}_t = [\sigma_{ij,t}]$ as the volatility matrix of the k-dimensional innovation $\boldsymbol{a}_t$ to the asset return series $\boldsymbol{z}_t$. The DCC model takes advantage of the fact that correlation matrices are easier to handle than covariance matrices. Therefore, the conditional correlation matrix $\boldsymbol{\rho}_t$ is used:

$$\boldsymbol{\rho}_t = \boldsymbol{D}_t \boldsymbol{\Sigma}_t \boldsymbol{D}_t, \tag{40}$$

where $\boldsymbol{D} = diag[\sigma_{11,t}^{1/2}, ..., \sigma_{kk,t}^{1/2}]$ is the $k \times k$ diagonal matrix of the time-varying standard variations at time t. The first step is to obtain the volatility series $\{\sigma_{ii,t}\}$ for $i = 1, ..., k$ assets. The second step is to model the dynamic dependence of the correlation matrix $\boldsymbol{\rho}_t$.

Engle (2002) introduces the first approach for DCC models. Let $\boldsymbol{\epsilon}_t = \boldsymbol{D}_t^{-1} \boldsymbol{a}_t$ be the vector of the standardized innovations and $\boldsymbol{\rho}_t$ the volatility matrix of $\boldsymbol{\epsilon}_t$. Then, the DCC by Engle is defined as:

$$\boldsymbol{Q}_t = \boldsymbol{S}(1-\alpha-\beta) + \alpha \boldsymbol{Q}_{t-1} + \beta(\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}'), \quad \alpha + \beta < 1, \tag{41}$$

$$\boldsymbol{\rho}_t = \boldsymbol{J}_t \boldsymbol{Q}_t \boldsymbol{J}_t, \tag{42}$$

with $\boldsymbol{S}$ being the unconditional correlation matrix of $\boldsymbol{\epsilon}_t$, $\boldsymbol{Q}_t$ being positive semi-definite and $\boldsymbol{J}_t = [q_{11,t}^{-1/2}, ..., q_{kk,t}^{-1/2}]$, where $q_{ii,t}$ denotes the (i,i)-th diagonal element of $\boldsymbol{Q}_t$:

$$\boldsymbol{J}_t = \begin{pmatrix} \sqrt{q_{11,t}} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{q_{22,t}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{q_{kk,t}} \end{pmatrix}. \tag{43}$$

The dynamic conditional correlation is modelled by the two parameters α and β. It is possible to implement additional lags $\ell = 1, .., m$ for $\boldsymbol{Q}_{t-m}$ and $\boldsymbol{\epsilon}_{t-m}$ and receive $\alpha_1, ..., \alpha_m$ or $\beta_1, ..., \beta_m$ (Ghalanos, 2012). However, in most applications only a DCC-(1,1)-process is estimated (Tsay 2013, Laurent et al. 2012).

An advantage of the DCC is that it can be estimated sequentially. The first step is to estimate the conditional variance of k assets and then model the conditional time-varying correlation between them, see section (6.3). This procedure is less efficient, but reduces the computational effort for the likelihood (Laurent et al., 2012).

The conditional distribution of $\boldsymbol{\epsilon}_t$ can be Multivariate Standard Normal with the probability density function

$$f(\boldsymbol{\epsilon}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{k/2}|\boldsymbol{\Sigma}|^{1/2}} exp\left[-\frac{1}{2}(\boldsymbol{\epsilon} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\epsilon} - \boldsymbol{\mu})\right], \tag{44}$$

with mean $\boldsymbol{\mu} = (\mu_1, ..., \mu_k)$ and positive-definite covariance matrix $\boldsymbol{\Sigma} = [\sigma_{ij}]$ (Tsay, 2013).

The standardized innovations $\boldsymbol{\epsilon}_t$ can also follow a multivariate Student-t distribution (Harvey et al., 1994) which is more appropriate for heavy-tailed financial data (Katsiampa, 2018). Then, the probability density function of $\boldsymbol{\epsilon}$ is

$$f(\boldsymbol{\epsilon}|\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\Gamma((\nu + k)/2)}{\Gamma(\nu/2)(\nu\pi)^{k/2}|\boldsymbol{\Sigma}|^{1/2}} \left[1 - \frac{1}{\nu}(\boldsymbol{\epsilon} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\epsilon} - \boldsymbol{\mu})\right]^{-(\nu+k)/2}, \tag{45}$$

where $\Gamma(\nu)$ denotes the usual Gamma function. Then, $\boldsymbol{\epsilon}$ follows a multivariate Student-t distribution with ν degrees of freedom and with location and scale parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ (Tsay, 2013). When ν tends to infinity, the Student-t distribution tends to the Gaussian distribution. When it tends to zero, the tails of the distribution become thicker (Bauwens et al., 2006).

6.3 Estimation

The Dynamic Conditional Correlation Model is estimated with a two-step approach. First, the univariate conditional variance is estimated for every series, described in section (4.2). In the second step, the conditional correlation matrix is estimated which is denoted by $\boldsymbol{\rho}_t = \boldsymbol{D}_t \boldsymbol{\Sigma}_t \boldsymbol{D}_t$. Engle and Sheppard (2001) propose a 2-stage Quasi-Likelihood estimation method to obtain the model parameters for the DCC. To obtain the univariate residual series for the first stage, $\boldsymbol{\Sigma}_t$ is replaced by $\boldsymbol{I}_k$, a $k \times k$ identity matrix. The resulting first-stage Quasi-Likelihood function for the parameter θ_1^* of the volatility then is:

$$QL_1(\theta_1^*) = -\frac{1}{2}\sum_{t=1}^{T}\left(k \cdot log(2\pi) + log(|I_k|) + 2 \cdot log(|D_t|) + r_t' D_t^{-1} I_k D_t^{-1} r_t\right) \quad (46)$$

$$= -\frac{1}{2}\sum_{t=1}^{T}\left(k \cdot log(2\pi) + 2 \cdot log(|D_t|) + r_t' D_t^{-2} r_t\right) \quad (47)$$

$$= -\frac{1}{2}\sum_{t=1}^{T}\left(k \cdot log(2\pi) + \sum_{n=1}^{k}\left(log(\sigma_{it} + \frac{r_{it}^2}{\sigma_{it}}\right)\right) \quad (48)$$

$$= -\frac{1}{2}\sum_{n=1}^{k}\left(T \cdot log(2\pi) + \sum_{t=1}^{T}\left(log(\sigma_{it} + \frac{r_{it}^2}{\sigma_{it}}\right)\right), \quad (49)$$

which can be seen as the sum of the log-likelihoods for the univariate GARCH processes. After the first stage is estimated, the parameters for the dynamic correlations can be found by using the standardized residuals of the first stage. The Quasi-Likelihood for the second stage, based on the estimates of θ_1^*, can be written as:

$$QL_2(\theta_2^*|\hat{\theta}_1^*) = -\frac{1}{2}\sum_{t=1}^{T}\left(k \cdot log(2\pi) + 2 \cdot log(|D_t|) + log(|\Sigma_t|) + r_t' D_t^{-1} \Sigma_t^{-1} D_t^{-1} r_t\right) \quad (50)$$

$$= -\frac{1}{2}\sum_{t=1}^{T}\left(k \cdot log(2\pi) + 2 \cdot log(|D_t|) + log(|\Sigma_t|) + \epsilon_t' \Sigma_t^{-1} \epsilon_t\right), \quad (51)$$

with $\epsilon_t \sim N(0, \Sigma_t)$ denoting the standardized residuals according to equation (30). Since the second stage Quasi-Likelihood is conditioned on $\hat{\theta}_1^*$, one can exclude the constant terms and simply maximize:

$$QL_2(\theta_2^*|\hat{\theta}_1^*) = -\frac{1}{2}\sum_{t=1}^{T}(log(|\Sigma_t|) + \epsilon_t'\Sigma_t^{-1}\epsilon_t). \quad (52)$$

The Quasi-ML estimators for θ_1^* and θ_2^* are assumed to be consistent and asymptotically normally distributed (Engle and Sheppard, 2001). For this thesis, the package *rmgarch* (Ghalanos, 2012) in R is used to estimate the dynamic conditional correlation model.

6.4 Diagnostic Checking

While there is a lot of literature on the diagnostics for univariate volatility models, fewer tests are devoted specifically to multivariate models. It is possible to distinguish between diagnostics that are applied to each univariate series and multivariate tests that are applied to the $k \times T$- dimensional time series. For the first mentioned tests, the diagnostics can be done analogously to section (4.3). Since the second stage of the DCC estimation is built on the residuals of the univariate series - see chapter (6.3) - violations on the first stage are important to detect. Although univariate tests can provide guidance, contemporaneous correlation of disturbances entails that statistics from individual equations are not independent. Therefore, there is a need for joint testing (Bauwens et al., 2006).
According to Ding and Engle (2001), one important moment of the model is that the standardized error terms should follow the condition:

$$Cov(\epsilon_{it}^2, \epsilon_{jt}^2) = 0, \quad \forall i \neq j, \quad (53)$$

if the conditional distribution is Gaussian. If the true distribution is the multivariate Student-t distribution, then, the covariance of the residuals of the different assets should follow:

$$Cov(\epsilon_{it}^2, \epsilon_{jt}^2) = \frac{2\nu^2}{(\nu-4)(\nu-2)^2}, \quad \forall i \neq j. \quad (54)$$

The term $2\nu^2/(\nu-4)(\nu-2)^2$ is different from zero if $1/\nu \neq 0$, which would be the case for a Gaussian distribution. However, testing for a multivariate Student-t distribution is a quite unexplored field in academic research yet. Bai and Chen (2008) have made an attempt to test distributional assumptions based on empirical residual distributions. Therefore, they transform the multivariate time series to a

univariate format and apply K-transformations to it to purge the distribution and make distribution free tests possible. Their method is only applicable for autoregressive and vector GARCH models so far.
If there are no remaining multivariate ARCH effects in the standardized residuals, then, ϵ_t should obey:

$$Cov(\epsilon_{it}^2, \epsilon_{j,t-m}^2) = 0, \quad \forall m > 0. \tag{55}$$

To test whether this is true, the multivariate Ljung-Box test described in equation (39) can be applied. Here, the Normal distribution of innovations is assumed. Again, the robust version of the test with upper tail trimming by Tsay (2013) can be applied here to overcome the non-normality of the data.
However, the state of research on model diagnostics for multivariate GARCH models is quite scarce yet, especially when it comes to the implementation in statistical software.

7 Multivariate Empirical Results

This section deals with the multivariate empirical results of interdependencies in the volatility of the five crypto coins. First, the sample correlation of the multivariate returns is analysed and a test for multivariate GARCH effects performed. After multivariate conditional heteroscedasticity has been detected, the Dynamic Conditional Correlation model by Engle (2002) is estimated based on the results for the univariate GARCH models. There are several orders for the DCC fitted and the two different multivariate distribution assumptions applied. After the DCC-(1,1) process with a multivariate Student-t distribution showed the most appropriate fit according to information criteria, the Value at Risk forecast for the weighted portfolio returns series is prsented, aligned with the approach by Bauwens and Laurent (2005). Also here, the forecasting performance of parameter refits every 300 vs. every 10 days is compared. It shows that, even though significant interdependencies in the multivariate volatility matrix exist, the employment of the multivariate approach does not improve the Value at Risk forecasting performance compared to the univariate results.

	BitCoin	LiteCoin	DogeCoin	FeatherCoin	Ripple
Bitcoin	1.00				
Litecoin	0.59	1.00			
Dogecoin	0.47	0.40	1.00		
Feathercoin	0.39	0.29	0.28	1.00	
Ripple	0.24	0.29	0.31	0.21	1.00

Table 5: Lag $\ell = 0$ sample correlation matrix $\hat{\rho}_0$ (Pearson) of the five crypto currency log return series. $T = 1554$.

Combining the five cryptocurrency log return series into a multidimensional data frame of equal length results in $T \times k = 1544 \times 5 = 7720$ observations. Table (5) shows the sample correlation at lag $\ell = 0$ and reveals a high connectedness of the multivariate log returns. In general, Bitcoin shows the highest correlations with the other currencies, the lowest correlation can be found between Feathercoin and Ripple. Ripple in general shows the lowest correlation with the other currencies.

The Henze and Zirkler test (1990) rejects the assumption of multivariate normality for the log returns of the cryptocurrencies. The Henze and Zirkler test is based

on a non-negative functional distance that measures the distance between two distribution functions. If the data is multivariate normal, the test statistic HZ is approximately log-normally distributed (Korkmaz et al., 2014).
The multivariate Ljung-Box Test from equation (39) is applied to the mean adjusted multivariate log return series. The highly significant test results imply multivariate GARCH effects in the data - this means that the conditional covariance matrix $\mathbf{\Sigma}_t$ is time-variant. Results are also highly significant for the robust version of the test which has been modified for heavy-tailed financial data.
The test result justifies the application of a multivariate GARCH model like the Dynamic Conditional Correlation model.

Dynamic Conditional Correlation Model. The estimation of the DCC model is done in two steps. First, the univariate GARCH models for the individual series are fitted. In chapter (5), we obtained the models for the conditional univariate variance from table (4), where we selected the apARCH for Bitcoin, Ripple, Dogecoin and Feathercoin and the eGARCH for Litecoin. To provide a mutual distribution for the multivariate standardized errors in the Dynamic conditional Correlation model, all univariate innovations in the GARCH models are assumed to follow a skewed Student-t distribution, even though the skewed Generalized Error distribution seemed to deliver better forecasting results for Bitcoin - see univariate VaR results in table (3).
The conditional variance then evolves according to:

$$\sigma^{1.148}_{BC,t} = 0.129 + 0.307 \cdot (|a^2_{BC,t-1}| + 0.035 \cdot a_{BC,t-1})^{1.148} + 0.832 \cdot \sigma^{1.148}_{BC,t-1},$$
$$\epsilon_{BC,t} \sim t_{0.954,2.480}.$$

$$\begin{aligned}\sigma^2_{LTC,t} = exp(&0.0636 - 0.0428 \cdot a_{Ltc,t-1} + 0.6221 \cdot [|a_{LTC,t-1}| \\ &-\mathbb{E}(|a_{LTC,t-1}|] + 0.9864 \cdot log(\sigma^2_{LTC,t-1})),\end{aligned}$$
$$\epsilon_{LTC,t} \sim t_{1.050,2.909}.$$

$$\sigma^{0.562}_{XRP,t} = 0.2957 + 0.4430 \cdot (|a^2_{XRP,t-1}| - 0.0665 \cdot a_{XRP,t-1})^{0.562} + 0.6651 \cdot \sigma^{0.562}_{XRP,t-1},$$
$$\epsilon_{XRP,t} \sim t_{0.990,2.281}.$$

$$\sigma^{1.152}_{DOGE,t} = 0.176 + 0.278 \cdot (|a^2_{DOGE,t-1}| - 0.0584 \cdot a_{DOGE,t-1})^{1.152} + 0.821 \cdot \sigma^{1.152}_{DOGE,t-1},$$
$$\epsilon_{DOGE,t} \sim t_{1.121,2.798}.$$

$$\sigma^{0.861}_{FTC,t} = 1.394 + 0.338 \cdot (|a^2_{FTC,t-1}| + 0.267 \cdot a_{FTC,t-1})^{0.861} + 0.639 \cdot \sigma^{0.821}_{FTC,t-1},$$
$$\epsilon_{FTC,t} \sim t_{1.117,2.669}.$$

Here, $t_{\xi,d}$ denotes the skew Student-t distribution with skewness parameter ξ and d

degrees of freedom.
As the second step, the Dynamic Conditional Correlation model is now to estimate the correlation matrix of the standardized innovations which were obtained with the volatility series of the univariate GARCH models.

$$\hat{\boldsymbol{\epsilon}}_t = (\hat{\epsilon}_{BC,t}, \hat{\epsilon}_{Ltc,t}, \hat{\epsilon}_{XRP,t}, \hat{\epsilon}_{DOGE,t}, \hat{\epsilon}_{FTC,t})', \quad \hat{\epsilon}_{it} = \hat{a}_{it}/\sqrt{\sigma_{ii,t}}. \tag{56}$$

The DCC model by Engle (2002) is applied to $\hat{\boldsymbol{\epsilon}}_t$. The model is implemented both with the multivariate Gaussian and the multivariate Student-t distribution assumption for $\boldsymbol{\epsilon}_t$. For both distribution models, a DCC-(1,1), -(1,2), -(2,1), and -(2,2) process is estimated.

	MV Normal				**MV Student-t**			
	DCC-(1,1)	DCC-(1,2)	DCC-(2,1)	DCC-(2,2)	DCC-(1,1)	DCC-(1,2)	DCC-(2,1)	DCC-(2,2)
α_1	0.0474***	0.0474	0.0198	0.0186	0.0627***	0.0559	0.0473*	0.0473*
α_2			0.0344	0.0668			0.0197	0.0197
β_1	0.9358***	0.9358	0.9269***	0.3231	0.9051***	0.8973	0.8984***	0.8984*
β_2		0.0000		0.5619		0.0000		0.0000
shape					4.0000***	4.0000***	4.0000***	4.0000***
AIC	32.202	32.204	32.196	32.200	29.754	30.267	29.934	29.755
BIC	32.413	32.418	32.410	32.418	29.969	30.485	29.754	29.976

Table 6: Model parameters for the estimated DCC models.
T=1544, k=5, *** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05.
Model parameters for univariate volatility series are listed in table (4).

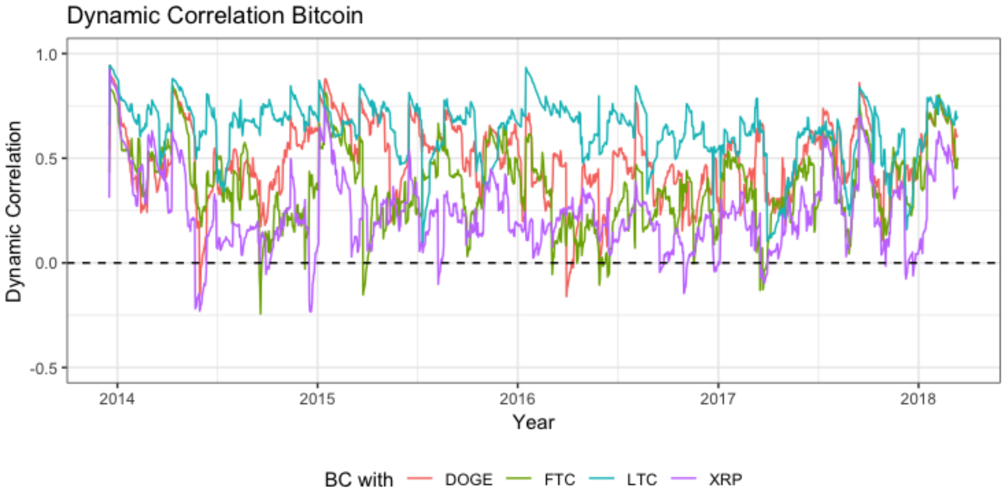
Table (6) shows the estimated parameters for the implemented DCC models. It shows that adding additional lags to the model is unnecessary and produces insignificant parameters. The DCC-(1,1) model performs best. The multivariate Student-t distribution with 4 degrees of freedom provides a better fit for the heavy-tailed data according to AIC and BIC. Thus, the resulting model equation for the estimated conditional correlation matrix is:

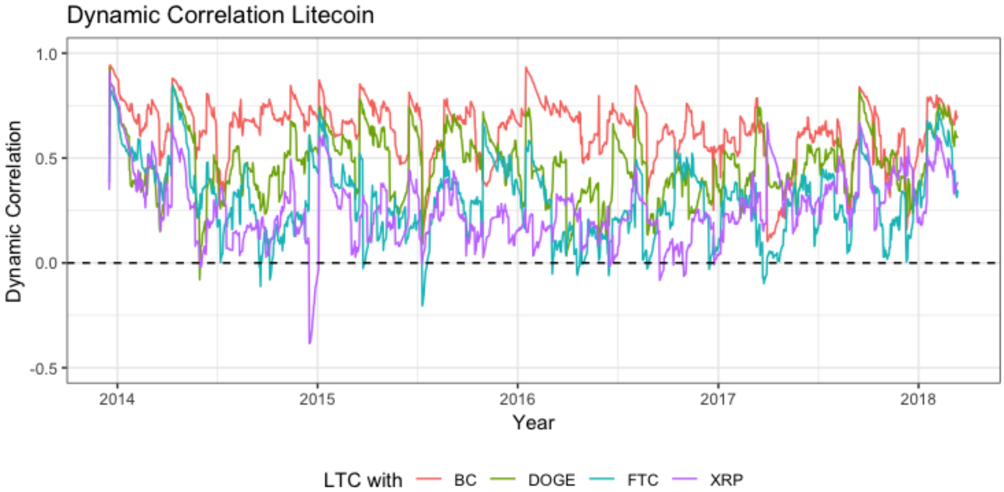
$$\begin{aligned} \boldsymbol{Q}_t = (1 - 0.0627 - 0.9051)\boldsymbol{S} + 0.0627\boldsymbol{Q}_{t-1} + 0.9051(\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1}), \\ \boldsymbol{\rho}_t = \boldsymbol{J}_t\boldsymbol{Q}_t\boldsymbol{J}_t, \end{aligned} \tag{57}$$

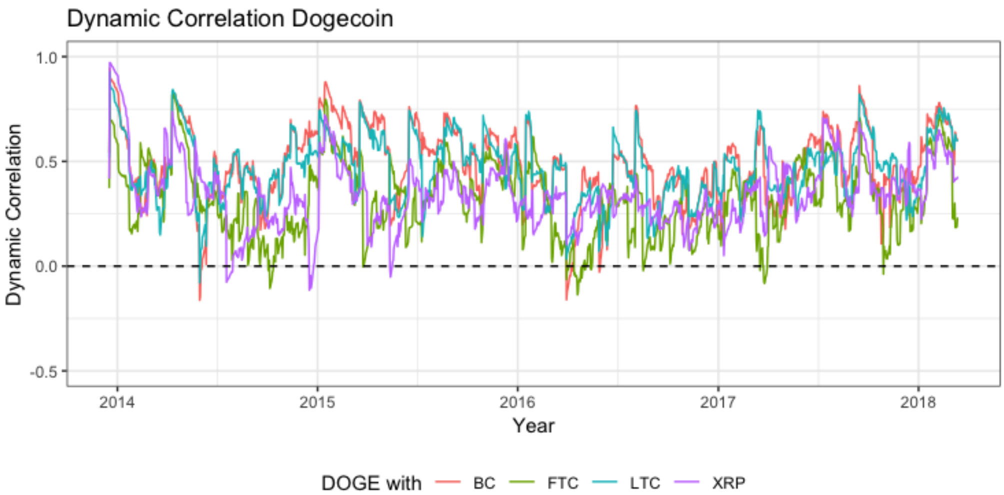
with $\boldsymbol{J}_t = [q_{11,t}^{-1/2}, ..., q_{kk,t}^{-1/2}]$, where $q_{ii,t}$ denotes the (i,i)-th element of $\boldsymbol{Q}$ and $\boldsymbol{S}$ is defined in equation (41). All parameters of the DCC-(1,1) with multivariate Student-t distribution are highly significant, which shows that the conditional variance of the

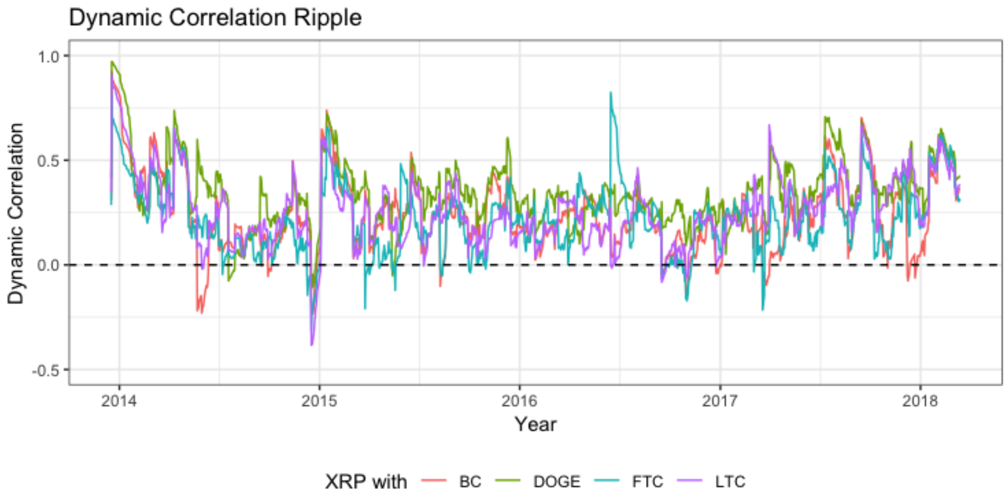
five cryptocurrencies is significantly affected by the preceding conditional correlation.
The Ljung-Box test at lag $\ell = 10$ for the univariate standardized residual series shows that there are no remaining ARCH effects in the squared residual terms. The Henze and Zirkler test shows that the multivariate residuals of the DCC do not follow a multivariate Normal distribution, which justifies the application of the heavy-tailed Student-t distribution. The multivariate Ljung-Box test from equation (39) is also applied to the squared multivariate standardized residual series. The test statistic $Q_k(m) = 323.472$ is highly significant and rejects the null hypothesis of no conditional heteroscedasticity. However, the standard version of the test is inappropriate if the innovations aren not assumed to follow a Gaussian distribution. Therefore, the robust version of $Q_k(m)$ with 5% upper tail trimming by Tsay (2013) is applied. $Q_k^r(m) = 323.3069$ also rejects the null hypothesis of no remaining ARCH effects with a p-value of 0.001. The DCC has often been shown to be rejected by model diagnostics (Tsay, 2013). However, some correlation between the volatility of the different currencies has been detected on a statistically significant level, which is now going to be further examined.

Figure (9) shows the dynamic correlations between the log returns series of the five crypto coins that were estimated by the DCC model in equation (57). Table (7) gives an additional summary of the mean and variance of the estimated dynamic correlation.
The correlation between Bitcoin and Litecoin as two of the most popular currencies is clearly the highest and also the most stable in comparison to the others since it shows the lowest standard deviation. Figure (9) also shows that it is the only correlation for Bitcoin with no peaks below zero. The correlation of Bitcoin with Dogecoin and Feathercoin is generally more unsteady but still shows a quite high mean correlation. Figure (9) reveals that the dynamic correlation of Dogecoin and Feathercoin with the other currencies evolves quite equally. A similar picture can be found for Ripple, where the correlation with the other currencies is evolving in a similar manner and relatively low. The relationship between Feathercoin and Ripple also appears to be unstable and shows swings in positive and negative directions.


Dynamic Correlation Bitcoin
Dynamic Correlation
1.0
0.5
0.0
-0.5
2014
2015
2016
2017
2018
Year
BC with
DOGE
FTC
LTC
XRP


Dynamic Correlation Litecoin
Dynamic Correlation
1.0
0.5
0.0
-0.5
2014
2015
2016
2017
2018
Year
LTC with
BC
DOGE
FTC
XRP


Dynamic Correlation Dogecoin
Dynamic Correlation
1.0
0.5
0.0
-0.5
2014
2015
2016
2017
2018
Year
DOGE with
BC
FTC
LTC
XRP



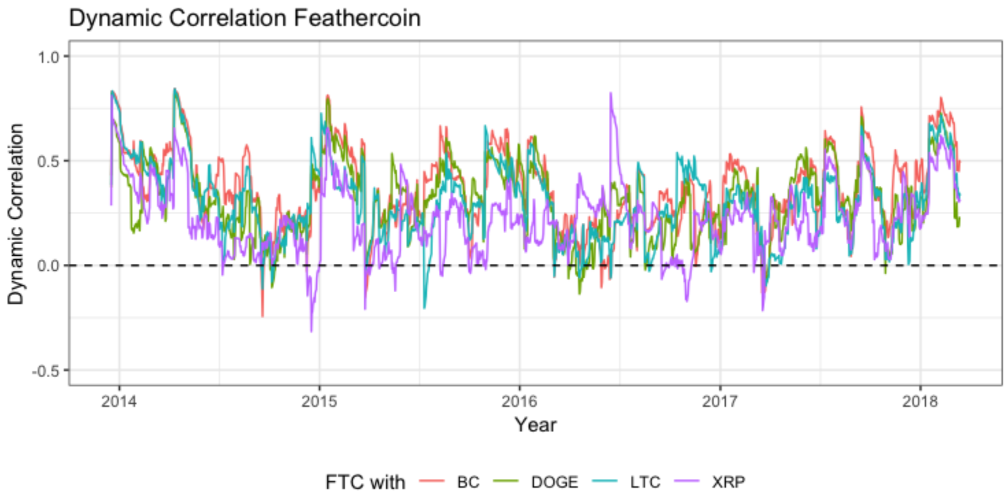


Figure 9: The estimated dynamic correlation between the currencies modelled by a DCC-(1,1) with multivariate Student-t distribution. T=1544.

	BitCoin	LiteCoin	DogeCoin	FeatherCoin	Ripple
BitCoin	1.00				
LiteCoin	0.6309 (0.1470)	1.00			
DogeCoin	0.4893 (0.1789)	0.4675 (0.1545)	1.00		
FeatherCoin	0.3661 (0.1899)	0.3027 (0.1847)	0.3190 (0.1795)	1.00	
Ripple	0.2396 (0.1815)	0.2612 (0.1668)	0.3453 (0.1607)	0.2227 (0.1693)	1.00

Table 7: Mean and (standard deviation) of the lag $\ell = 0$ correlations in the multivariate volatility of the currencies estimated by the DCC model in equation (57). T=1544.

Table (7) shows that the estimated dynamic conditional correlation of innovations between the currencies is positive on average. Bitcoin has the overall highest correlation with all other currencies, while Ripple shows the lowest correlation with other currencies. Dogecoins correlation with the other coins is quite high. The correlation between Bitcoin and Litecoin is the highest with the lowest standard deviation. Feathercoin has on average the highest standard deviation in the correlation with the other currencies. Figure (9) however, shows that the dynamic correlation is varying over time, thus the interdependency between the currencies behaves differently for several time periods.

Value at Risk Forecasting. To calculate the one-day-ahead Value at Risk forecast for the multivariate model, the time series has to be transformed into a univariate format. Equivalently to section (5) a rolling forecast with recursive window is performed, with re-estimated parameters for the DCC every 300 days and every 10 days separately. Subsequently, the forecasted mean portfolio return and volatility of all five currencies is calculated and the 1% and 5% quantile loss function generated. VaR forecasts are compared to the realized returns of the $1/k$ portfolio. To calculate the loss function for the multivariate portfolio, a weight vector is applied to the fitted mean and conditional covariance matrix of the DCC model, according to the approach by Bauwens and Laurent (2005):

$$\hat{\mu}_{PF,t+1} = \frac{1}{k}\sum_{i=1}^{k}\hat{\mu}_{i,t+1}, \quad i = 1, .., k, \quad k = 5. \tag{58}$$

In the 2-stage estimation of the DCC, $\hat{\mu}_{i,t}$ for the individual series is obtained by the univariate GARCH models. $\hat{\mu}_{PF,t}$ is then the average for the five fitted series. $\hat{\sigma}_{PF,t+1}$ is obtained from the $\hat{\mathbf{\Sigma}}_t$-matrix fitted at the second stage by applying a weight vector to it:

$$w = \begin{pmatrix} 1/k \\ \vdots \\ 1/k \end{pmatrix}_{k\times 1}, \tag{59}$$

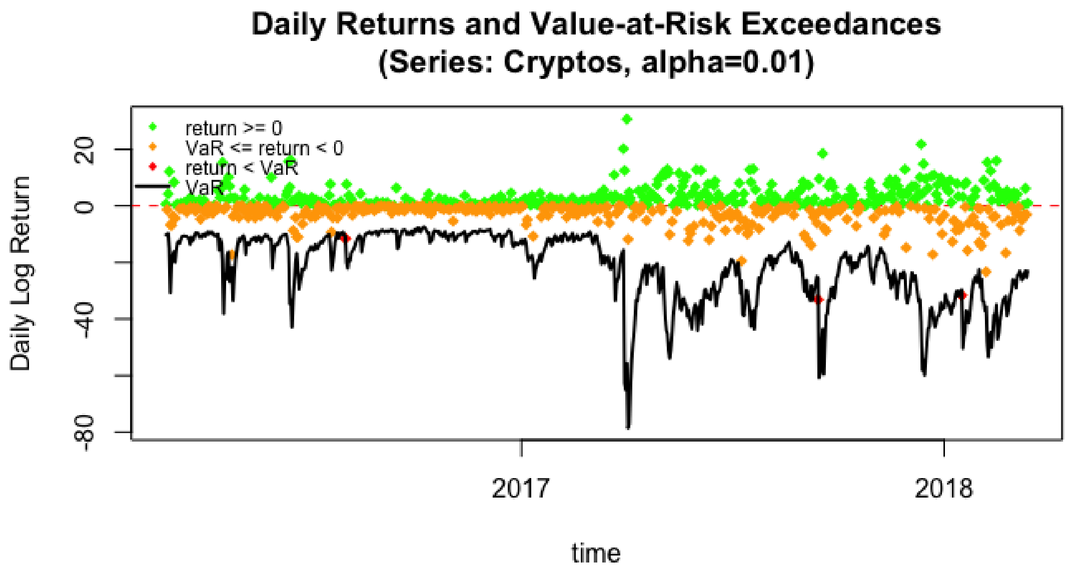
$$\hat{\sigma}_{PF,t+1} = \left[w\hat{\mathbf{\Sigma}}_{t+1}w'\right]^{1/2}. \tag{60}$$

Value at Risk is then the negative 1%- or 5%-quantile of the the conditional return distribution:

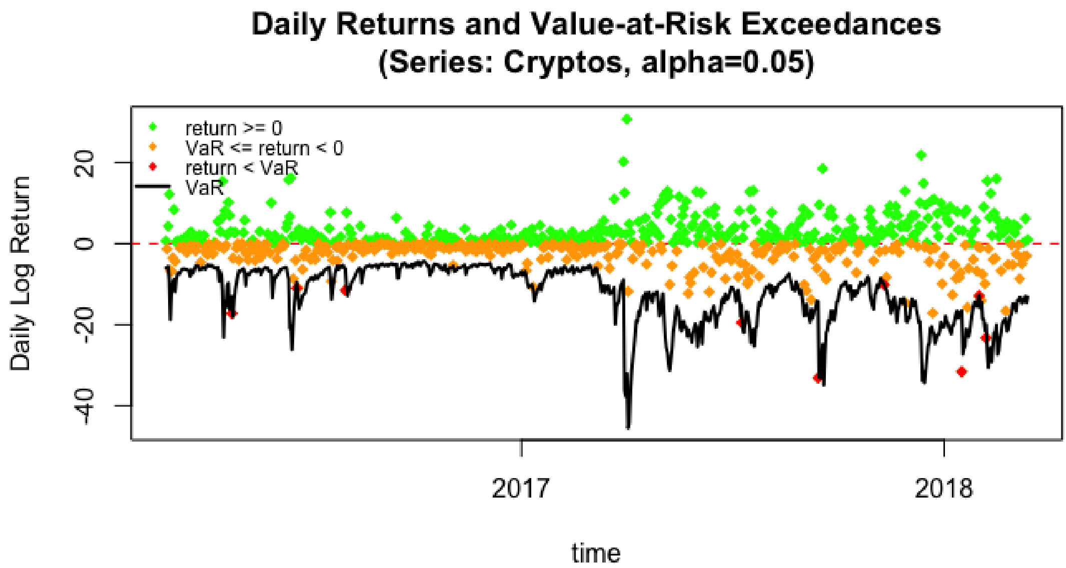
$$VaR^{\alpha}_{PF,t+1} := -Q_{\alpha}(r_{PF,t+1}|F_t). \tag{61}$$

which is determined by $r_{PF,t} = \mu_{PF,t} + \sigma_{PF,t}\epsilon_{PF,t}$, with $\epsilon_{PF,t}$ again being a random iid variable following the Student-t distribution with mean zero, standard deviation one and shape parameter obtained by the fitted DCC. The implementation of the VaR forecast via a DCC in *R* is aligned with the work by Ghalanos (2016), the author of the *rugarch* and *rmgarch* package.
The Value at Risk forecast is performed for parameter re-estimates every 300 days as well as every 10 days.

Figure (10a) and (10b) show the 1% and 5%-VaR forecasts of the DCC together with the daily returns of the $1/k$ portfolio with parameter re-estimates every 300 days. Compared to Figure (5), which shows the VaR results of the portfolio modelled via a univariate eGARCH, one can also identify a volatile period in mid 2016, followed by a stable period until early 2017 and a highly volatile period during the hype from early 2017 until end of the observation span. For the 1%-VaR limit, there are 3 data points where the realized portfolio loss exceeds the forecasted Value at Risk (0.4% violations). The LR test for unconditional coverage cannot reject the null hypothesis of correct exceedances for both α-levels. Also the test for conditional coverage does not reject the H_0 of independent exceedances. At the 5%-VaR level, there are 9 out of 37 violations found (1.2%). Even though this is a very low violation rate, both tests for unconditional and conditional coverage are rejected. Therefore, the violations in the forecast that was performed with the DCC are not correct and independent from each other.
Like in chapter (5), a second VaR forecast for the portfolio is performed, with parameters re-estimated every 10 days instead of every 300 days. The univariate analysis showed that a more frequent refit does not provide more accurate Value at Risk forecasts.

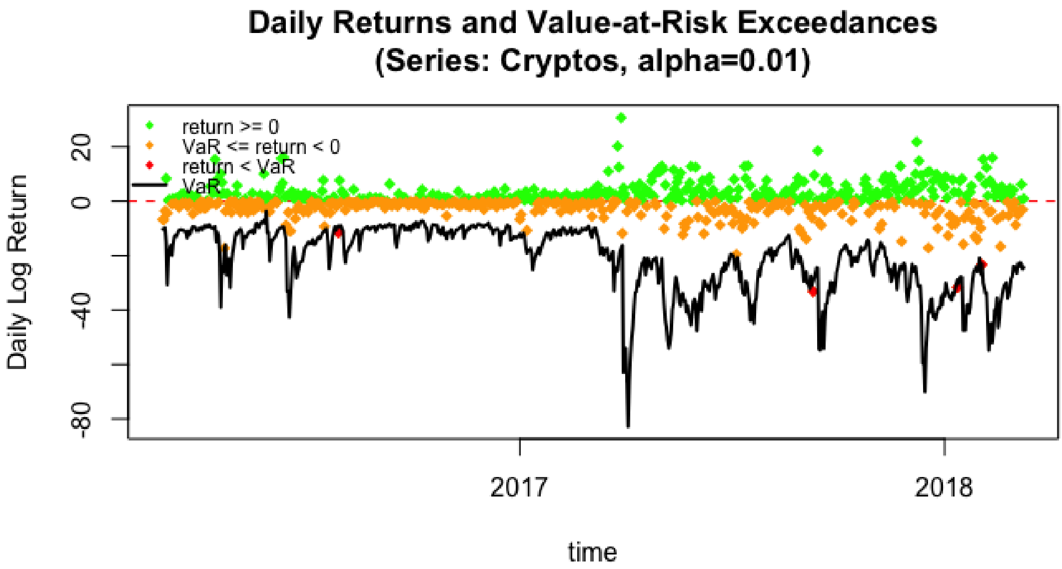


(a) 1%-Value at Risk. Parameters re-estimated every **300 days**. % Viol: 0.4, p-Values: LR_{uc} : 0.062, LR_{cc} : 0.175.

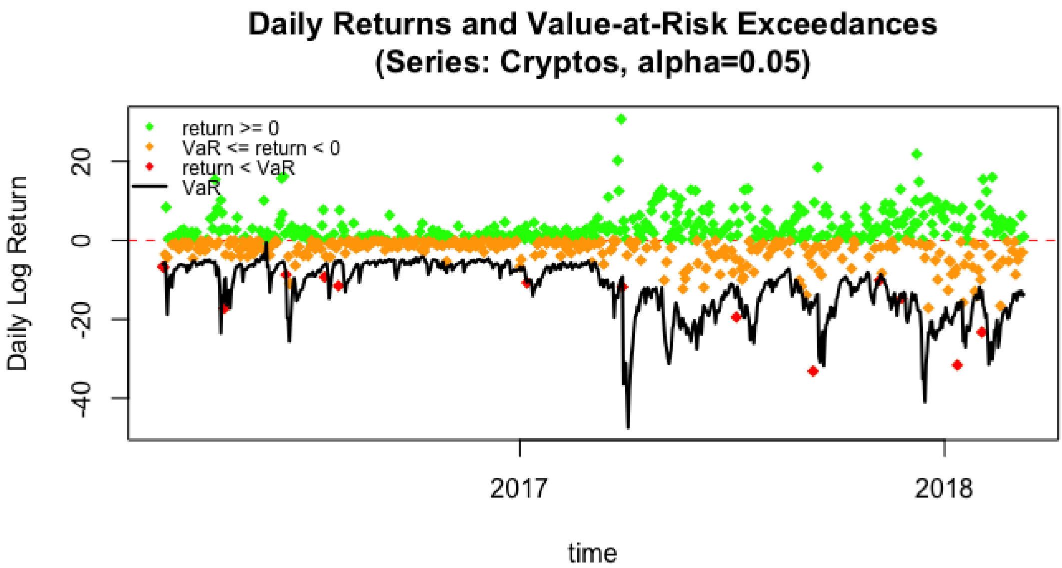


(b) 5%-Value at Risk. Parameters re-estimated every **300 days**. % Viol: 1.2, p-Values: LR_{uc} : 0.000, LR_{cc} : 0.000.

Figure 10: Value at Risk forecasts for the 1-day-ahead rolling forecast of the 1/k portfolio estimated with a multivariate DCC according to equation (57).
The model is built on a training data set of 800 observations, which leaves 744 out-of-sample forecasts.
Parameters are re-estimated every **300 days** with a recursive window, which leads to a total number of 3 refits.



(a) 1%-Value at Risk. Parameters re-estimated every **10 days**. % Viol: 0.6, p-Values: LR_{uc} : 0.168, LR_{cc} : 0.379.



(b) 5%-Value at Risk. Parameters re-estimated every **10 days**. % Viol: 1.9, p-Values: LR_{uc} : 0.000, LR_{cc} : 0.000.

Figure 11: Value at Risk forecasts for the 1-day-ahead rolling forecast of the 1/k portfolio estimated with a multivariate DCC according to equation (57).
The model is built on a training data set of 800 observations, which leaves 744 out-of-sample forecasts.
Parameters are re-estimated every **10 days** with a recursive window, which leads to a total number of 75 refits.

It shows that updating the Dynamic Conditional Correlation model parameters more often does not improve the multivariate forecasting results either.
Figure (11a) shows the 1% Value at Risk exceedances for the Dynamic Correlation model with parameters refitted every 10 days. In fact, the number of violations increases from 3 to 4 out of 7 expected violations, which implies a violation rate of 0.6%. The test for unconditional coverage shows a p-value of 0.168 and the test for conditional coverage can not be rejected with a p-value of 0.379 either. The VaR results at the $\alpha = 5\%$-level show a different picture in Figure (11b). Even though there is only a violation rate of 1.9%, with 14 out of 37 expected violations, the tests by Christoffersen (1998) show that the unconditional and conditional coverage is violated. Therefore, VaR violations in the forecast are not correct and independent from each other. In comparison to Figure (6) and (5) where the aggregated portfolio was forecasted with an eGARCH, the forecasted portfolio loss is also slightly higher, the DCC model thus draws a more pessimistic picture.
It can be concluded that the DCC does not outperform the results by the simple univariate eGARCH with skewed Student-t distribution for the aggregated portfolio which showed a good violation rate and did not fail the VaR backtesting. The most accurate VaR results are still found with a univariate forecast for each individual return series. Here, percentage violations between 0.4 and 0.7% can be found at the $\alpha = 1\%$-level, while both the tests for conditional and unconditional coverage are not rejected. Higher levels of portfolio aggregation might have led to information losses for the individual return processes. Since there was a weight vector applied to $\mathbf{\Sigma}$ for the calculated VaR quantile function in the multivariate approach - see equation (60) - the model might not be able to identify dependencies in the volatility of the individual currencies anymore. In the univariate analysis, the parameters for the negative quantile function can also be obtained more specifically, i.e. by the estimated skewness parameter.
Specifying the right distribution model is crucial when it comes to Value at Risk forecasting and the results of the DCC could be improved by introducing a skewness to negative quantile function of the aggregated portfolio. Dynamic conditional correlation models in combination with multivariate skew densities have shown to provide more accurate VaR results for financial data (Bauwens and Laurent, 2005).

The multivariate analysis has shown the following key findings: The multivariate distribution of innovations is more appropriately modelled with a multivariate Student-t distribution rather than the Gaussian. This is another finding that proves the need for a heavy-tailed distribution when modelling the volatility of cryptocurrencies and is already supported in the analysis of multivariate volatility interdependies by Katsiampa (2018). The conditional covariance of the five cryptocurrencies is significantly affected by the covariance of past error terms, which justifies the application of a multivariate approach to understand co-movements of the cryptocurrencies volatility. However, the Ljung-Box test rejects the assumption of no remaining multivariate ARCH effects in the squared residuals. The estimated conditional correlation between the currencies is time-varying but positive on average. Those results are supported by the findings of Katsiampa (2018) and Corbet et al. (2018), who also found that price and volatility developments of the currencies are highly connected to each other. Bitcoin shows the strongest correlation to the other currencies, while Ripples correlation to the other currencies appears to be more weak and unsteady.

The forecasting of Value at Risk has shown that consulting a multivariate approach does not make the results more accurate. Concering the forecasting of a portfolio downside risk for an aggregated portfolio, it is more efficient to apply a univariate GARCH model to the aggregated univariate portfolio return series rather than choosing the more complex Dynamic Conditional Correlation model. Furthermore, an aggregated portfolio shows less accurate forecasting performance than the forecasting of the individual return series. The most accurate results were produced with the apARCH, eGARCH and csGARCH in combination with a skewed Student-t or skewed Generalized error dsitribution for the individual volatility forecasts. This is a result that has already been found for the forecasting of traditional assets. Kole et al. (2017) have shown that higher levels of portfolio aggregation and the multivariate approach via a Dynamic Conditional Correlation model does not outperform the Value at Risk forecasting of univariate GARCH models in many cases.

8 Conclusion

The immense popularity of the altcoin market and the increasing interest in media and economics has lead to a lot of scientific attempts to explain what drives cryptocurrency price developments and their highly volatile behaviour. Still a lot is unknown about this emerging market. In this thesis, one of the first analyses of the interconnectedness between five different cryptocurrency volatilities and its impact on forecasting accuracy has been made. Its main attempt was to find out whether the forecasting accuracy of the conditional heteroscedasticity can be improved by taking the conditional covariance into account. It showed that there is a strong interconnectedness between the volatility of the five currencies, however, compared to a univariate forecast of the individual series, the implementation of a Dynamic Conditional Correlation model does not improve the Value at Risk forecasting accuracy.

A descriptive analysis of the stylized facts for Bitcoin, Litecoin, Dogecoin, Ripple and Feathercoin has been made. It showed that, except for Bitcoin, all currencies are positively skewed and all show heavy-tail behaviour, next to conditional heteroscedasticity in the evolution of their volatility. They also exhibit a high standard variation of returns, with Bitcoin being the least risky currency. Those results are supported by previous findings in academic research on the cryptocurrency market (Chu et al. 2017, Catania and Grassi 2017, Gkillas and Katsiampa 2018).

Several univariate GARCH models were fitted for the individual return series of the five coins. Results showed that the innovations of the log return series are most appropriately modelled with a skewed distribution like the skewed Student-t or the skewed Generalized Error distribution. It has been shown that the conditional heteroscedasticity is modelled best with an asymmetric power ARCH, a component-standard GARCH or an exponential GARCH. Predominately, the apARCH performed best, providing the best fit according to information criteria and residual plots. The high performance of the apARCH also remains during the rolling forecast of portfolio downside risk. It has already been identified in the literature as a powerful forecasting tool for conditional heteroscedasticity (Kuester et al., 2006). It also shows the need for a GARCH-model that takes the leverage effect into account since it has already been found that the volatility process of cryptocurrencies shows an asymmetry to the sign of past innovations (Catania et al. 2018, Phillip et al. 2018). The skewed Generalized Error distribution provides the overall lowest number of VaR violations but usually goes along with some violations of model assumptions and worse backtesting results. Furthermore, variations in the frequency

of parameter refits for the rolling 1-day ahead forecast show that the model parameters are stable over time and a more frequent update does not improve the performance of the forecast significantly.
After the apARCH and eGARCH with skewed Student-t distribution have been chosen to model the volatility of the univariate innovations, the Dynamic Conditional Correlation model with a multivariate Student-t distribution has been implemented to estimate the conditional correlation of the multivariate innovations. The parameters of the DCC-(1,1) process are highly significant, which shows that the conditional variance of the five cryptocurrency returns series is significantly affected by the preceding conditional correlation in the multivariate volatility. A graphical and descriptive analysis of the estimated dynamic correlation showed that the relation between the volatility of the currencies is positive on average but time-varying. In general, Bitcoin shows the highest connectedness with other currencies, while the conditional variance of Ripple is least correlated to the others. Even though a strong positive relationship between the currencies has been found, the multivariate approach does still not improve the Value at Risk forecasts. A more frequent update of the model parameters still does not outperform the portfolio forecast with a univariate GARCH model. The best forecasting results are found with the univariate GARCH-type models for the individual return series. The results are supported by the findings by Katsiampa (2018), who found a strong but time-varying relationship between the conditional heteroscedasticity of the main currency prices via a multivariate BEKK approach, and the results by Kole et al. (2017), who found that Dynamic Conditional Correlation models and aggregated portfolios usually can not outperform the Value at Risk accuracy of univariate GARCH models with lower portfolio aggregation.

The univariate results have shown that there is a need for skewed heavy-tailed distributions to model the innovations of the crypto returns. For the Dynamic Conditional Correlation model, forecasting results could have been improved by introducing a skewness into the multivariate Student-t distribution (Bauwens et al., 2006). Furthermore, the evaluation of a portfolio via Value at Risk has been found to be subject to a few limitations. Even though it is a popular tool that is also easily comprehensible to laymen, it does not differentiate between "good" and "bad" risk and reflects only the pessimistic view on the risk of a portfolio. The possibility of highly positive portfolio returns is not taken into account (Dembo and Freeman, 2001). Artzner et al. (1999) have also shown that VaR does not operate sub-additive, i.e. VaR can report a higher total portfolio risk than the sum of the individual positions. Further analysis of the performance of DCC models in the

crypto market could investigate alternative risk measures like the Conditional Value at Risk (CVaR) (Artzner et al., 1999)) or the Expected Shortfall (ES) (Acerbi and Tasche, 2002). Scientific research of the cryptocurrency volatility is an emerging topic of interest and there is going to be better understanding of dependencies in the altcoin market as soon as longer observation spans are available for statistical analysis. Right now, researchers need to choose coins that existed for a long time but might have lost their importance - like Feathercoin - to obtain a large sample size. Future analysis will hopefully benefit from a more stable market.
Regarding conditional heteroscedasticity, we have seen that the dynamics between the volatilities of the return series are not constant over time. The conditional correlation between the multivariate return series could be modelled more appropriately with a Regime Switching DCC (Pelletier, 2006). In this model, the correlation matrix of the volatilities is constant within a regime but varies across different regimes. This might be a tool to capture the different trends that the altcoin market experienced since it emerged in 2009.

The results have improved the understanding about how volatilities of cryptocurrencies are connected to each other. However, there is still much we don't know about this emerging market. A lot of academic work has already been spent on the effect of policy and social media activities on crypto prices, but there might be way more influencing variables that we are not aware of yet. An interesting upcoming event is the full utilisation of the Bitcoins that are left to mine. In fall 2018, only 18 percent are left until the limitation of 21 million Bitcoins available to mine is reached (Blockchainhalf.com). We have found that the development of all other cryptocurrencies is strongly correlated to Bitcoin. Therefore, the limit of the Bitcoin supply could have a huge impact on the altcoin market and the financial industry in general.

References

Carlo Acerbi and Dirk Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.

Giovanni Angelini and Silvia Emili. Forecasting cryptocurrencies: A comparison of garch models. 2018. URL https://dx.doi.org/10.2139/ssrn.3195704.

Frederik Armknecht, Ghassan O Karame, Avikarsha Mandal, Franck Youssef, and Erik Zenner. Ripple: Overview and outlook. In *International Conference on Trust and Trustworthy Computing*, pages 163–180. Springer, 2015.

Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

Jushan Bai and Zhihong Chen. Testing multivariate distributions in garch models. *Journal of Econometrics*, 143(1):19–36, 2008.

Luc Bauwens and Sébastien Laurent. A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics*, 23(3):346–354, 2005.

Luc Bauwens, Sébastien Laurent, and Jeroen VK Rombouts. Multivariate garch models: a survey. *Journal of applied econometrics*, 21(1):79–109, 2006.

Ernst R Berndt, Bronwyn H Hall, Robert E Hall, and Jerry A Hausman. Estimation and inference in nonlinear structural models. In *Annals of Economic and Social Measurement, Volume 3, number 4*, pages 653–665. NBER, 1974.

Blockchainhalf.com. accessed: 14.09.2018. URL https://www.bitcoinblockhalf.com.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

Leopoldo Catania and Stefano Grassi. Modelling crypto-currencies financial time-series. BI Norwegian Business School, Centre for Applied Macro-and Petroleum Economics, CAMP Working Paper Series, 2017.

Leopoldo Catania, Stefano Grassi, and Francesco Ravazzolo. Predicting the volatility of cryptocurrency time–series. BI Norwegian Business School, Centre for Applied Macro-and Petroleum Economics, CAMP Working Paper Series, 2018.

Amélie Charles and Olivier Darné. Volatility estimation for bitcoin: Replication and extension. *International Economics*, 2018. URL https://doi.org/10.1016/j.inteco.2018.06.004.

Shi Chen, Cathy Yi-Hsuan Chen, Wolfgang K Härdle, TM Lee, and Bobby Ong. A first econometric analysis of the crix family. 2016. URL https://dx.doi.org/10.2139/ssrn.2832099.

Peter F Christoffersen. Evaluating interval forecasts. *International economic review*, pages 841–862, 1998.

Jeffrey Chu, Stephen Chan, Saralees Nadarajah, and Joerg Osterrieder. Garch modelling of cryptocurrencies. *Journal of Risk and Financial Management*, 10 (4):17, 2017.

Kuo Chuen, Lee David, Li Guo, and Yu Wang. Cryptocurrency: A new investment opportunity? 2017. URL https://dx.doi.org/10.2139/ssrn.2994097.

Pavel Ciaian, Miroslava Rajcaniova, and d'Artis Kancs. The economics of bitcoin price formation. *Applied Economics*, 48(19):1799–1815, 2016.

Coingecko.com. accessed: 14.03.2018. URL https://www.coingecko.com/en.

Coinmarketcap.com. accessed: 18.09.2018. URL https://coinmarketcap.com/all/views/all/.

Christian Conrad, Anessa Custovic, and Eric Ghysels. Long-and short-term cryptocurrency volatility components: A garch-midas analysis. *Journal of Risk and Financial Management*, 11(2):23, 2018.

Shaen Corbet, Andrew Meegan, Charles Larkin, Brian Lucey, and Larisa Yarovaya. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, 165:28–34, 2018.

Ron S Dembo and Andrew Freeman. *Seeing tomorrow: Rewriting the rules of risk.* John Wiley & Sons, 2001.

Zhuanxin Ding and Robert F Engle. Large scale conditional covariance matrix modeling, estimation and testing. Technical report, NYU Working Paper No. FIN-01-029., 2001. URL https://ssrn.com/abstract=1294569.

Zhuanxin Ding, Clive WJ Granger, and Robert F Engle. A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1): 83–106, 1993.

Dogecoin.com. accessed: 20.09.2018. URL www.dogecoin.com.

Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

Robert F Engle and Tim Bollerslev. Modelling the persistence of conditional variances. *Econometric reviews*, 5(1):1–50, 1986.

Robert F Engle and Gary Lee. A long-run and short-run component model of stock return volatility. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, pages 475–497, 1999.

Robert F Engle and Kevin Sheppard. Theoretical and empirical properties of dynamic conditional correlation multivariate garch. *National Bureau of Economic Research*, 2001.

Feathercoin.com. accessed: 14.09.2018. URL https://feathercoin.com/about/.

Carmen Fernández and Mark FJ Steel. On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371, 1998.

Ifigeneia Georgoula, Demitrios Pournarakis, Christos Bilanakos, Dionisios Sotiropoulos, and George M Giaglis. Using time-series and sentiment analysis to detect the determinants of bitcoin prices. 2015. URL https://dx.doi.org/10.2139/ssrn.2607167.

Alexios Ghalanos. rmgarch: Multivariate garch models (r package version 0.98). 2012. URL https://cran.r-project.org/web/packages/rmgarch/rmgarch.pdf.

Alexios Ghalanos. How good are your var estimates? December 26 2012, unstarched.de, accessed: 12.07.2018, 2016. URL http://www.unstarched.net/2012/12/26/how-good-are-your-var-estimates/.

Alexios Ghalanos. Introduction to the rugarch package (version 1.3-1). 2018. URL http://cran.r-project.org/web/packages/rugarch.

Github.com. accessed: 14.09.2018. URL https://github.com/bitcoin/bitcoin.

Konstantinos Gkillas and Paraskevi Katsiampa. An application of extreme value theory to cryptocurrencies. *Economics Letters*, 164:109–111, 2018.

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. *The journal of finance*, 48(5):1779–1801, 1993.

Andrew Harvey, Esther Ruiz, and Neil Shephard. Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264, 1994.

N Henze and B Zirkler. A class of invariant consistent tests for multivariate normality. *Communications in Statistics-Theory and Methods*, 19(10):3595–3617, 1990.

Carlos M Jarque and Anil K Bera. A test for normality of observations and regression residuals. *International Statistical Review*, pages 163–172, 1987.

Paraskevi Katsiampa. Volatility estimation for bitcoin: A comparison of garch models. *Economics Letters*, 158:3–6, 2017.

Paraskevi Katsiampa. An empirical investigation of volatility dynamics in the cryptocurrency market. 2018. URL https://dx.doi.org/10.2139/ssrn.3202317.

Erik Kole, Thijs Markwat, Anne Opschoor, and Dick Van Dijk. Forecasting value-at-risk under temporal and portfolio aggregation. *Journal of Financial Econometrics*, 15(4):649–677, 2017.

Selcuk Korkmaz, Dincer Goksuluk, and Gokmen Zararsiz. Mvn: An r package for assessing multivariate normality. *The R Journal*, 6(2):151–162, 2014.

Keith Kuester, Stefan Mittnik, and Marc S Paolella. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1): 53–89, 2006.

Sébastien Laurent, Jeroen VK Rombouts, and Francesco Violante. On the forecasting accuracy of multivariate garch models. *Journal of Applied Econometrics*, 27 (6):934–955, 2012.

Ming-Chih Lee, Jung-Bin Su, and Hung-Chun Liu. Value-at-risk in us stock indices with skewed generalized error distribution. *Applied Financial Economics Letters*, 4(6):425–431, 2008.

Litecoin.org. accessed: 14.09.2018. URL https://litecoin.org.

Allan I McLeod and William K Li. Diagnostic checking arma time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, 4(4): 269–273, 1983.

Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. Technical report, 2008. URL https://bitcoin.org/bitcoin.pdf.

Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370, 1991.

Denis Pelletier. Regime switching for dynamic correlations. *Journal of econometrics*, 131(1-2):445–473, 2006.

Andrew Phillip, Jennifer Chan, and Shelton Peiris. A new look at cryptocurrencies. *Economics Letters*, 163:6–9, 2018.

Said E Said and David A Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.

Panayiotis Theodossiou. Financial data and the skewed generalized t distribution. *Management Science*, 44(12-part-1):1650–1661, 1998.

Panayiotis Theodossiou. Skewed generalized error distribution of financial assets and option pricing. 2000. URL https://dx.doi.org/10.2139/ssrn.219679.

Ruey S Tsay. *Analysis of financial time series*, volume 543. John Wiley & Sons, 2005.

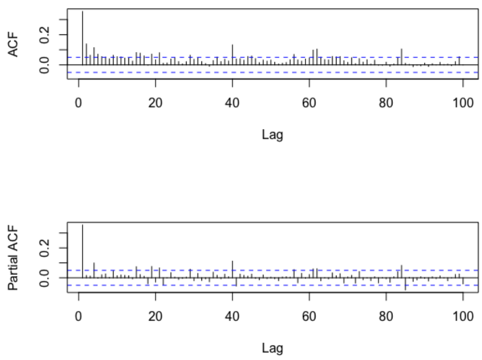
Ruey S Tsay. *Multivariate time series analysis: with R and financial applications*. John Wiley & Sons, 2013.

Ruey S Tsay. *An introduction to analysis of financial data with R*. John Wiley & Sons, 2014.

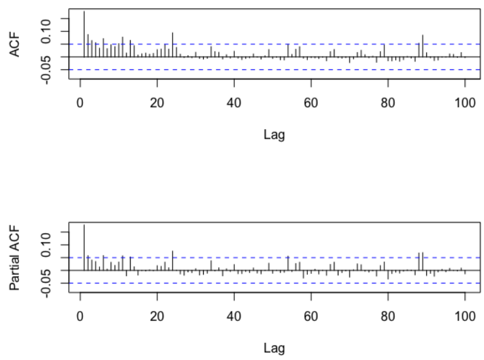
Chao Xu and Huigeng Chen. Measuring portfolio value at risk. 2012. URL http://lup.lub.lu.se/student-papers/record/2628584.

Aaron Yelowitz and Matthew Wilson. Characteristics of bitcoin users: an analysis of google search data. *Applied Economics Letters*, 22(13):1030–1036, 2015.

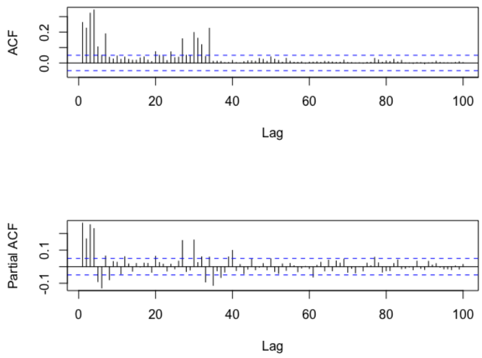
A Appendix



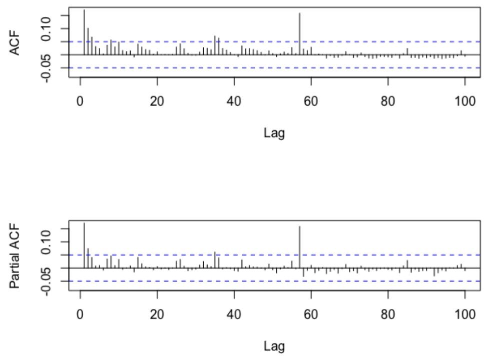
(a) Bitcoin



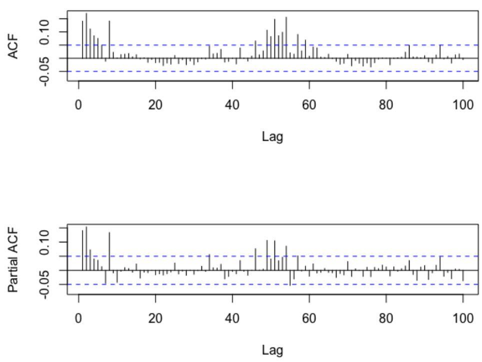
(b) Litecoin



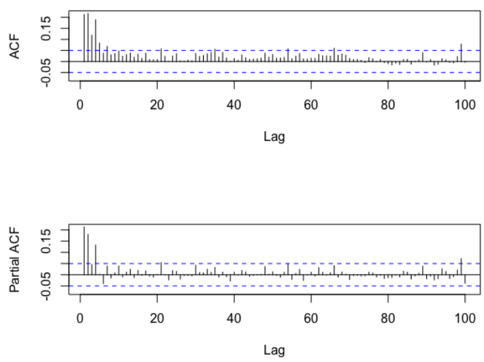
(c) Dogecoin



(d) Ripple

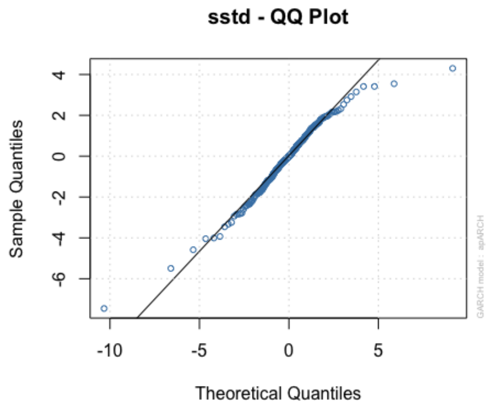


(e) Feathercoin

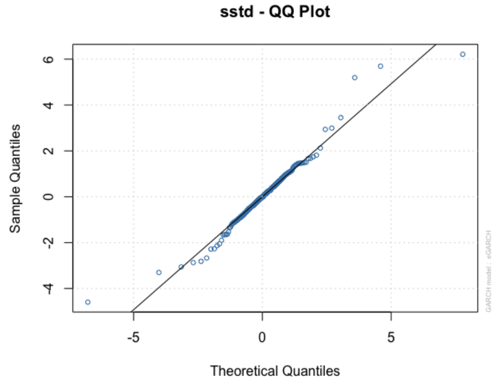


(f) Portfolio

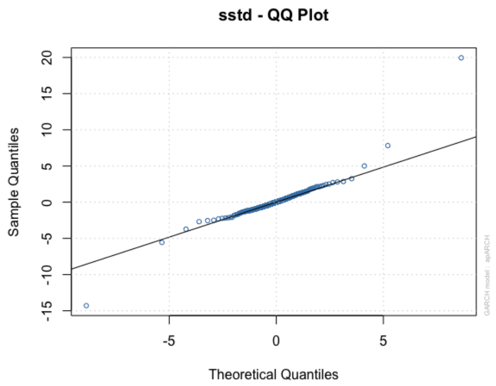
Figure 12: Plot for the autocorrelation and partial autocorrelation of the squared logarithmic returns of the individual cryptocurrency return series. Blue dashed line indicates the 5% significance level.



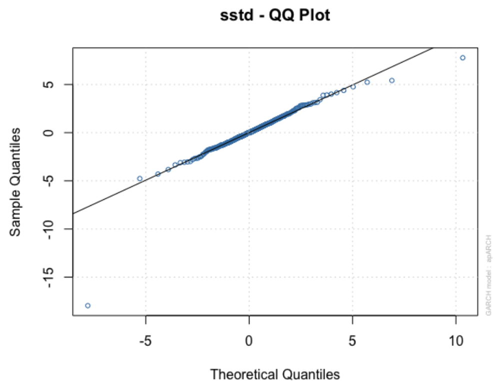
(a) Bitcoin: apARCH, skew Student-t distribution



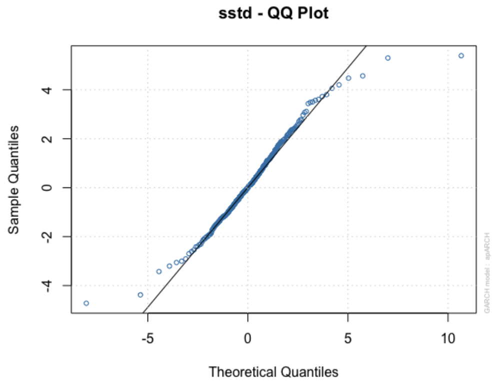
(b) Litecoin: eGARCH, skew Student-t distribution



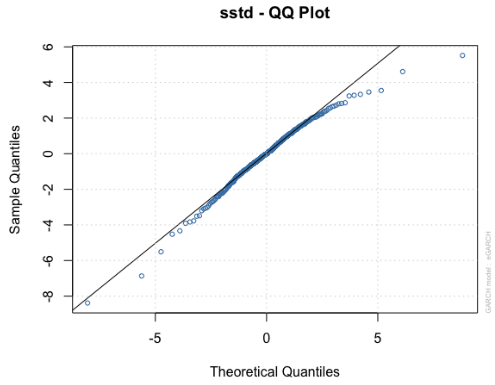
(c) Ripple: apARCH, skew Student-t distribution



(d) Dogecoin: apARCH, skew Student-t distribution

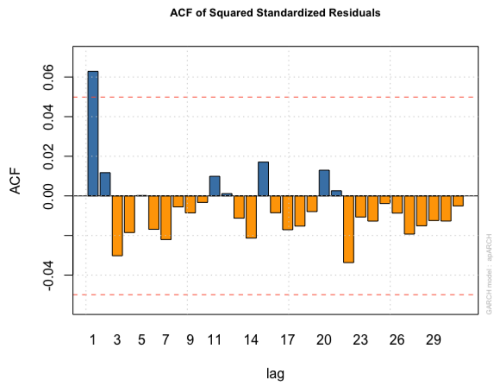


(e) FeatherCoin: apARCH, skew Student-t distribution

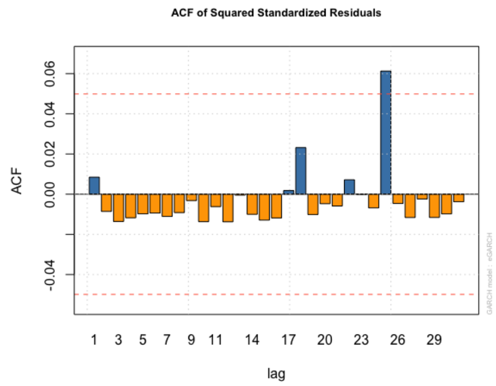


(f) Portfolio: eGARCH, skew Student-t distribution

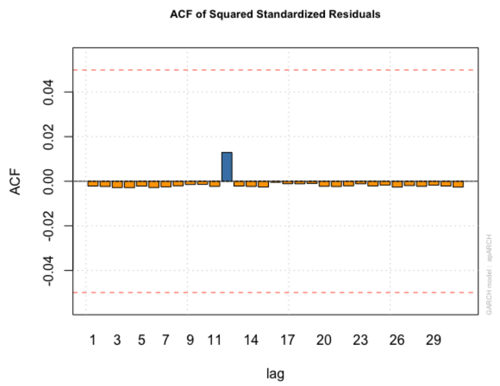
Figure 13: Quantile-to-quantile plots for the distribution of standardized errors for the selected univariate GARCH models



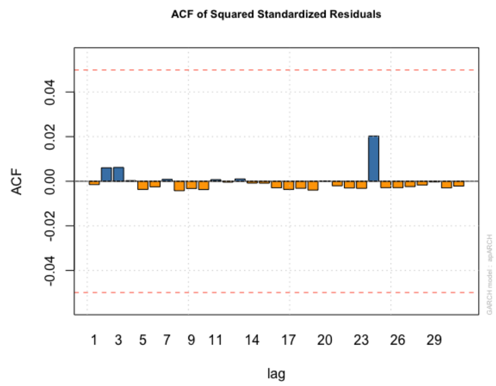
(a) Bitcoin: apARCH, skew Student-t distribution



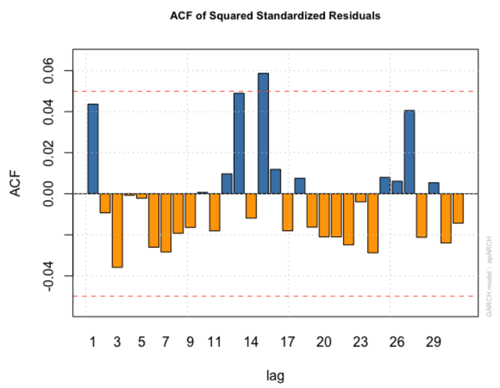
(b) Litecoin: eGARCH, skew Student-t distribution



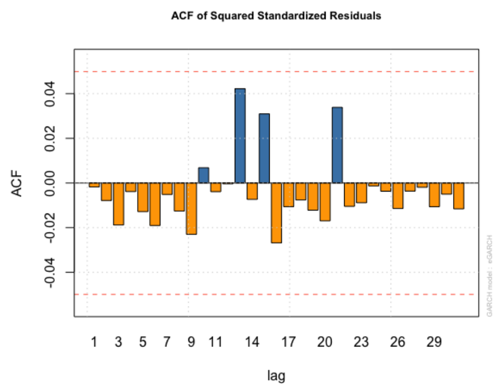
(c) Ripple: apARCH, skew Student-t distribution



(d) Dogecoin: apARCH, skew Student-t distribution



(e) FeatherCoin: apARCH, skew Student-t distribution



(f) Portfolio: eGARCH, skew Student-t distribution

Figure 14: Autocorrelation of squared standardized errors for the selected univariate GARCH models. Red dashed line indicates the 5% significance level.

B Electronic Attachments

The electronic attachment contains a PDF document "Guideline for R-Codes" and two folders "Code" and "Data":

- "Guideline for R-Codes" contains a listing of the different R-Codes and explains the content and data in more detail.
- "Code" contains 11 *R*-Codes that are to be executed in the consecutive order in which they are numbered.
- "Data" contains the original data sets of the five cryptocurrency daily closing price time series obtained from Coingecko.com, and the datasets built in the R-Codes.

C Declaration

I hereby declare that the thesis submitted is my own unaided work. All direct or indirect sources are acknowledged as references. Furthermore, no part of this thesis has been submitted elsewhere for any other degree or qualification. I am aware that this thesis in digital form can be examined for the use of unauthorized aid and in order to determine whether the thesis as a whole or parts incorporated in it may be deemed as plagiarism.

Place and date	Signature