Chapter 4
Quantitative Risks

Eline Van der Auwera, Wim Schoutens,
Marco Petracco-Giudici & Lucia Alessi

Abstract Cryptocurrency stock prices are a rich source of information on cryptocurrency behaviour. Therefore, we will manipulate the stock prices of Bitcoin and other cryptocurrencies to pinpoint particular characteristics. Starting with a thorough descriptive analysis of the log returns of the most well-known cryptocurrencies. Afterwards, we will investigate the inter-correlation between cryptocurrencies and also the correlation between cryptocurrencies and other asset classes. Then, an ARMA-GARCH model is fitted to the log-returns and this model allows to make a value-at-risk prediction. To finish this chapter, we will look at the market efficiency of the cryptocurrency market.

4.1 Descriptive Analysis of Returns

Next, we perform some time-series analysis on a variety of cryptocurrencies and further also investigate their correlation with other more traditional assets. One must be cautious with cryptocurrency data since it varies considerable (Alexander and Dakos (2019)). We typically will work with the so-called log-returns over a single time unit (in most cases one day):

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{4.1}$$

where t denotes the time and P_t is the asset price at a certain time t. We will use the daily historical prices of Ripple, Monero, Bitcoin and Ethereum to sketch an overview of the distributional and volatile behaviour of cryptocurrencies. Note that cryptocurrency exchanges trade 24/7. We will mainly focus on the historical closing prices starting from the first of January 2016 onward.

Table 4.1 gives a descriptive analysis of the log returns of the different cryptocurrencies and regular assets. Note that in comparison to other more traditional assets, the difference between the maximal and minimal value of the log returns of cryptocurrencies is high, as is the yearly standard deviation (std.). The median is in each case for cryptocurrencies smaller than the mean, this indicates the presence of positive outliers. The value in between parenthesis represents the p-value of testing the null-hypothesis of zero average returns. For all the returns under examination, the null-hypothesis cannot be rejected.

All the log returns of cryptocurrencies also exhibit a higher kurtosis (kurt.) than the normal distribution, this means that the distribution is likely to have fat tails and will exhibit more peakedness. The kurtosis of

	Min.	Median	Mean	Max.	Yearly Std.	Kurt.	Skew.
BTC	-0.225	0.003	0.002 (0.117)	0.287	0.645	5.869	-0.218
ETH	-0.319	-0.000	0.004 (0.250)	0.312	1.027	3.466	0.058
XMR	-0.285	0.001	0.004 (0.433)	0.625	1.166	9.102	1.116
XMR*	-0.285	0.001	0.003 (0.143)	0.528	1.096	5.923	0.655
XRP	-0.496	-0.003	0.003 (0.179)	0.881	1.194	26.023	2.562
XRP*	-0.496	-0.003	0.002 (0.279)	0.593	1.085	14.013	1.471
XAU	-0.018	0.000	0.000 (0.410)	0.024	0.097	0.620	0.068
BCOM	-0.027	0.000	-0.000 (0.612)	0.027	0.104	1.492	-0.316
S&P500	-0.042	0.001	0.000 (0.286)	0.048	0.130	6.122	-0.621
EURUSD	-0.024	0.000	-0.000 (0.819)	0.030	0.098	2.311	0.101

The value between the parenthesis represents the p-value of the hypothesis of zero average returns. The starred (*) tickers represent the returns without the maximal return.

Table 4.1: Descriptive analysis of daily log returns

	ADF		KPSS		PP	
	Test statistic	P-value	Test statistic	P-value	Test statistic	P-value
BTC	-33.861	0.000	0.340	0.105	-31.120	0.000
ETH	-17.836	0.000	0.605	0.022	-34.790	0.000
XMR	-11.305	0.000	0.476	0.046	-36.319	0.000
XRP	-7.165	0.000	0.147	0.399	-35.356	0.000
XAU	-25.510	0.000	0.195	0.363	-29.201	0.000
BCOM	-25.395	0.000	0.172	0.330	-29.981	0.000
S&P500	-8.984	0.000	0.115	0.516	-23.842	0.000
EURUSD	-42.172	0.000	0.142	0.414	-42.417	0.000

Table 4.2: Result of different stationarity tests

the cryptocurrencies is comparable to those of emerging government bonds. The skewness (skew.) of the log returns is positive except for Bitcoin. This positive skewness is not present in typical equity markets. We also observe that commodities (gold and the commodity index) and the Euro Dollar exchange rate have a kurtosis lower than 3, the kurtosis of a normal distribution, while the S&P500 stock index has a higher kurtosis. The commodity index and the S&P500 stock index both have a negative skewness, while gold and the exchange rate have a positive skewness.

A comparison among the log returns of the cryptocurrencies shows that Ripple has simultaneously the smallest and the largest log returns on his name. Moreover, it has the highest yearly standard deviation among all the cryptocurrencies. From a historical time-series perspective, Ripple is hence the most volatile of all the cryptocurrencies under consideration. When we delete the maximal return, as depicted with the starred ticker, we can still conclude the same remarks. Bitcoin is the least volatile of all the cryptocurrencies which are under examination here. Ripple exhibits the highest kurtosis and skewness, which means that Ripple has heavier tails and more large outliers than the remaining cryptocurrencies. The median of the log returns of all the cryptocurrencies are similar and very close to zero.

The log returns of all the investigated assets from 01/01/2016 until 25/02/2019 are probably stationary according to the Augmented Dickey Fuller (ADF) and Phillips Perron (PP) test. The tested null-hypothesis is:

$$H_0 : \text{The time series is non-stationary (not necessarily with a trend)} .$$

Stationary timeseries are timeseries whose unconditional joint probability remains the same over time, the weak stationarity only requires the (co)variance and mean to be constant over time. On the other hand, non-stationary behaviour can have cycles, random walks and trends or a combination of these three. Table 4.2 shows that the null-hypothesis is rejected at 5% significance level because the p-value for the ADF- and PP-test is smaller than 0.05, which means that the process is stationary. The null-hypothesis test for the Kwiatkowski Phillips Schmidt and Shin (KPSS) test is

$$H_0 : \text{The time series is weakly stationary.}$$

The null-hypothesis is rejected for Ethereum and Monero at a 5% significance level because the corresponding value in Table 4.2 is smaller than 0.05. For all the other log returns the weakly stationary hypothesis is not rejected.

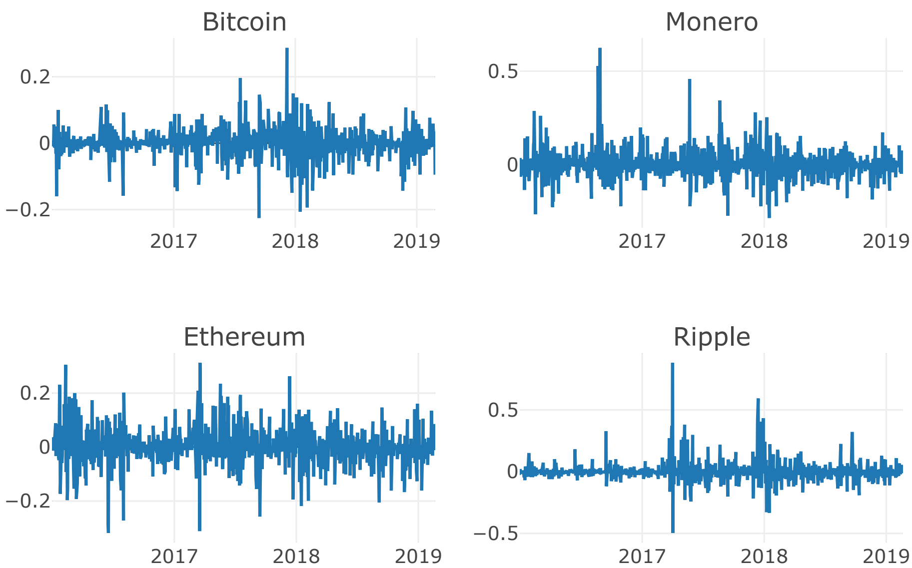
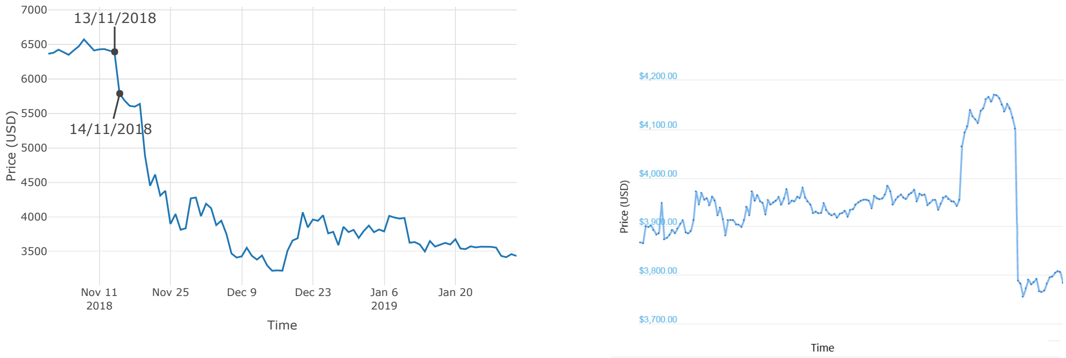


Fig. 4.1: Log returns of the different cryptocurrencies

Figure 4.1 shows one important remark, the log returns of cryptocurrencies do exhibit periods where high returns cluster together and periods where low returns reside, this could point to volatility clustering in the log returns. Another observation is that half of the time the returns are positive and half of the time the returns are negative, the daily log returns of cryptocurrencies seem to be mean-reverting to zero and Figure 4.2 provides some further examination of the log returns over time. The cumulative returns over time from mid 2015 until February 2019 is shown. The cryptocurrencies have had an upward movement until the end of 2017. From 2018 onward mainly negative returns were realized. In comparison to the traditional assets, cryptocurrencies have had a much higher cumulative return.

Fig. 4.2: Cumulative log returns

Cryptocurrencies often exhibit behaviour where large returns happen due to crashes in the price. In November 2018, the price of Bitcoin began to drop. The drop started with a crash of nearly 10% between the 13th of November and the 14th of November after nearly a month of having a stable price around 6200 USD, see Figure 4.3a. Moreover, the whole market dropped from a stable market price of 209 billion USD to 187 billion USD, which is a drop of 10.53% effectively. The mid November drop is the start of the price decline of Bitcoin, the level of 13 November 2018 will not be reached again in the near future. Between 24 February and 25 February the price of Bitcoin again suffered from a crash, as Figure 4.3b clearly shows. This time the price of Bitcoin dropped 9.7% over the course of only hours. Similar behaviour as in Figure 4.3b is found in Ethereum, Litecoin and many other cryptocurrencies. This could be a consequence of the leader (Bitcoin) follower behaviour in cryptocurrencies.

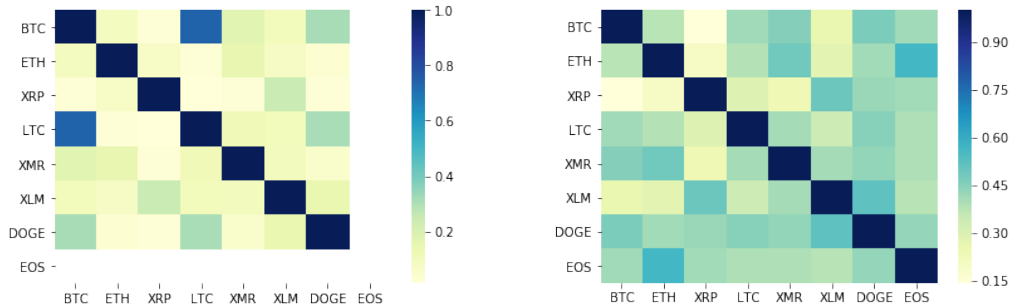


(a) November 2018 drop (b) February 2019 drop

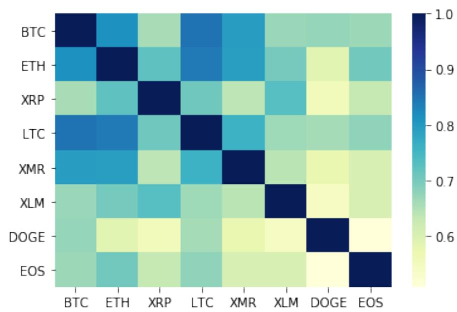
Fig. 4.3: Price drops in Bitcoin Figure 4.3a represents a price drop from 13/11/2018 until 14/11/2018 of nearly 10%. Figure 4.3b depicts a drop of 9.7% during the month February of that same year.

4.2 Correlation and Diversification

4.2.1 Correlation between Several Cryptocurrencies



(a) 2016 (b) 2017



(c) 2018

Fig. 4.4: Correlation heatmaps between different cryptocurrencies through the years.

Figures 4.4a, 4.4b and 4.4c represent heatmaps of the Pearson correlation between the daily log returns of several cryptocurrencies from 2016 until 2018. Bitcoin (BTC) and Litecoin (LTC) exhibit a high correlation in the year 2016, which is no surprise since both currencies have similar characteristics and prospects. The strong positive dependence between these two cryptocurrencies continues to exists for following years. As discussed, Ripple (XRP) and Stellar (XLM) also have similar characteristics, since they both act in the remittance field. The correlation between them also support this claim through the years. Note that EOS did not exist at in 2016.

Note that the correlation between several cryptocurrencies has significantly increased over the last three years. In other words, not only currencies with similar characteristics have high correlation, but the whole market seems to move in the same direction. One particular reason for the cryptomarket to be so correlated is because many cryptocurrencies are bought using Bitcoin or Ether. Hence, the dominant position of Bitcoin or Ethereum in the market is fortified.

4.2.2 Correlation between Cryptocurrencies and Other Stocks

Next, we investigate the correlation of cryptocurrencies with various other types of assets: commodities, equity and currencies. The commodity class will be represented by gold, the security class by S&P500 stock index and to model the currency behaviour we use the EUR/USD exchange rate.

First, we investigate the correlation between the log returns of Bitcoin and the S&P500 stock index. Figure 4.5a shows the rolling Pearson correlation over a window of a predefined number of days between the log returns of both assets. The assets appear to become less correlated if the number of data points to construct it becomes larger. Moreover, the correlation, when taking into account 180 data points per window, is nearly zero, but positive. Besides, short term correlations fluctuate around zero. The correlation between Bitcoin and the S&P500 based on a window of 180 days never exceeds 30%, which indicates that Bitcoin is not very correlated to it.

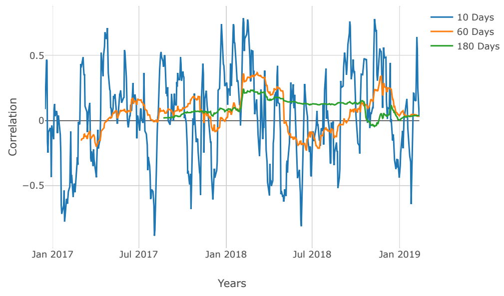
The correlation between Bitcoin and the price of gold in USD is shown in Figure 4.5b. A rolling window of respectively 10, 60 and 180 days is used to find the correlation over time. The 10-day correlation fluctuates around zero, like the correlation between S&P500 index and Bitcoin. The 180-day correlation, in this case is mainly negative, this means that the price movements in gold and Bitcoin act in opposite directions. Moreover, the 180-correlation is very close to zero but negative, which means that the price movements of gold and Bitcoin are more or less uncorrelated or move in opposite directions.

The Pearson correlation between the log returns of EUR/USD exchange rate and Bitcoin is plotted in Figure 4.5c. The 10-day correlation fluctuates around zero, however, longer time-periods lead to a more negative correlation. The 180-day correlation is even persistently below zero. Note that the 60-day correlation goes up in the end, hence it is likely that the 180-day correlation will follow this pattern.

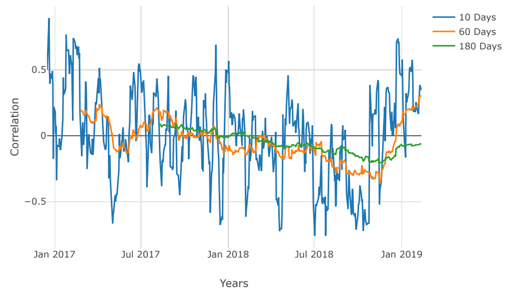
The one year correlation between Bitcoin and other major asset classes (bonds, oil, US real estate and emerging market currencies) stays within the boundaries of being a differentiated risk reducer (see Figure 4.5d) (Burniske and White (2017)) because the correlation remains under 0.3 in absolute value. The asset classes are quantified by

- Bond: Bloomberg Barclays US aggregate bond index (LBUSTRUU)
- Oil: crude oil futures (CL1 Comdty)
- Emerging markets currencies: J.P. Morgans emerging markets currency index (FXJPEMCS Index)
- Real estate: Vanguards US real estate ETF (VNQ).

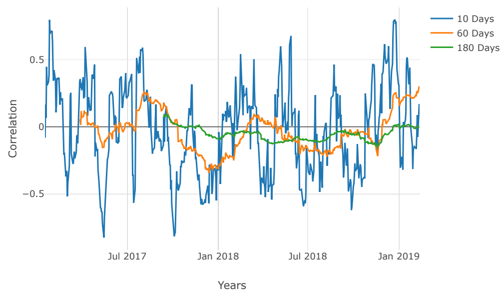
Moreover, the average correlation with these major asset classes is fluctuating around zero, which means that Bitcoin is not heavily dependent on other assets.



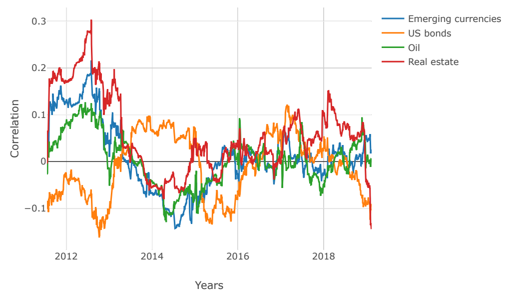
(a) Between BTC and S&P500



(b) Between BTC and Gold



(c) Between BTC and Euro Dollar exchange rate



(d) Between Bitcoin and other assets

Fig. 4.5: The rolling correlation with window sizes of 10, 60 and 180 days.

4.3 Flash Crash

A flash crash (Kenton (2018)) is a sudden steep decline in price of the underlying asset. Many people think flash crashes are a part of artificial market manipulation. There are several reasons why flash crashes can happen; human error, fraud, high frequency trading and computer/software glitches are all possible triggers. High frequency trading is an automated trading system, where computers are able to detect changing market conditions and make trades accordingly. These computers place enormous orders at light speed and as a result they can enlarge the negative price movements. Each crash is unique and it is difficult to pinpoint the exact cause. Many cryptocurrencies exchanges are open via an API system to all kinds of trading bots. Some bots try to play the bid ask spread, others do automated trading on the basis of technical analysis or other kinds of algorithms. The presence of such bots, often written by non-experienced participants increase the risk of flash-crashes.

In May 2019 the BTC-Canadian Dollar (CAD) price experienced a flash crash of 99.1% on the exchange Kraken (SFox (2019)), see Figure 4.6. The price for one Bitcoin dropped from 11 800 CAD to a low of 101.2 CAD. Only a few minutes after the flash crash, the price again reached its normal level. One of the possible reasons of the crash is the limited liquidity on Kraken to exchange Bitcoin for Canadian Dollar. Some traders were very lucky to obtain Bitcoins at the low price. However, this also means that the sellers of those particular Bitcoins made a huge loss. Stop-loss-orders can be the reason why there were Bitcoins available in the market at such a low price. A stop-loss-order is an order which is send to the market once a particular price-level is breached. If the level is breached, the stop-loss-order becomes an order. In this case, the order became a sell order at the market price during the crash.

Fig. 4.6: Flash crash of BTC/CAD in May 2019 from cryptowat.ch

4.4 Pump and Dump

Market manipulation is a crime in regulated markets, however many crypto-exchanges are unregulated and hence are under no real external supervision. This makes them vulnerable for all kinds of manipulation. Pump and dump is a form of such market abuse, in the regulated markets this scheme is illegal. However, since crypto-markets are mostly unregulated, it occurs much more frequently there. Some players drive the price of a stock they hold up (pump), by creating false interest in it. Once the price has risen sufficiently high, they sell their stocks on the market at an higher price than the original (true) market price (dump). The objective of a pump and dump scheme practitioner is distorting supply and demand in their favour. In general, pump and dump actions tend to work better on small illiquid stocks, because then a sharp increase in trade volume can send the price up drastically.

In cryptocurrencies there exists organized groups who perform pump and dump schemes. The groups post 'insider tips' regarding cryptocurrencies, which promise quick gains and wealth. The participants in a pump and dump scheme are told when and where to buy the specified coin via messages. Usually a pump and dump scheme drives the trading volume and price up simultaneously.

Figure 4.7 shows two obvious pump and dump operations, one on the 7^{th} of February 2019 and one on the 18^{th} of February 2019. From this figure, one can clearly see that the volume and price, on these days, has increased significantly. After the initial wave of buying, the participating investors need to act quickly to sell the asset, while at the same time they encourage non-suspecting investors to buy the asset. The pump and dump group tries to convince the non-suspecting investors to buy the asset by promoting the asset and as such try to find support for the upward price move. The wave of selling the assets, by the participants in the pump and dump, decreases the price often even more than its original level.

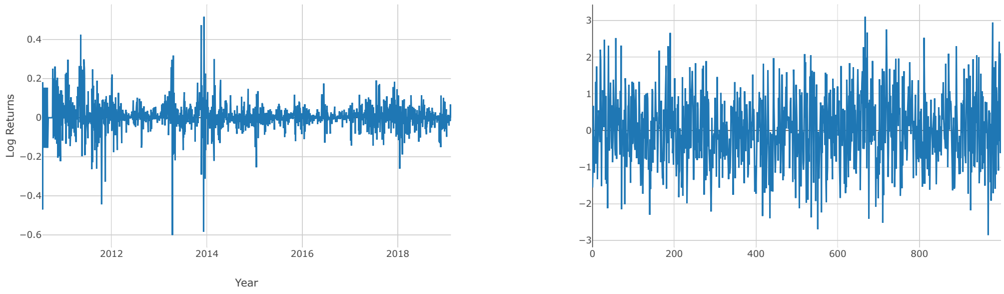
Fig. 4.7: The upper panel depicts the price history of Vertcoin and the lower panel the trading volume of Vertcoin.

Telegram channels advertise pump and dumps schemes nearly daily, sometimes even multiple times a day. Since the exchanges of cryptocurrencies are hardly regulated or even unregulated today.

4.5 Distributional Behaviour

The timeseries of the log returns of Bitcoin do not follow a white noise process. This can be seen by comparing Figure 4.8a, which represents the log returns of Bitcoin over time, and Figure 4.8b, which depicts a typical white noise process. The log returns of Bitcoin exhibit a more volatile pattern with so-called volatility-clusters (periods of high volatility). There are indications that the log returns are stationary, according to the tests in Table 4.2, but the price movement in different time periods vary quite a lot. As mentioned before, the log returns of cryptocurrencies exhibit volatility clustering, it is important to accurately model the time variability of the volatility.

Figure 4.9 shows the quantile plots of the different cryptocurrencies their log returns. In this case, the quantiles of the Bitcoin, Ethereum, Ripple and Monero their log returns are compared to quantiles



(a) Bitcoin (b) White noise

Fig. 4.8: Difference in log returns

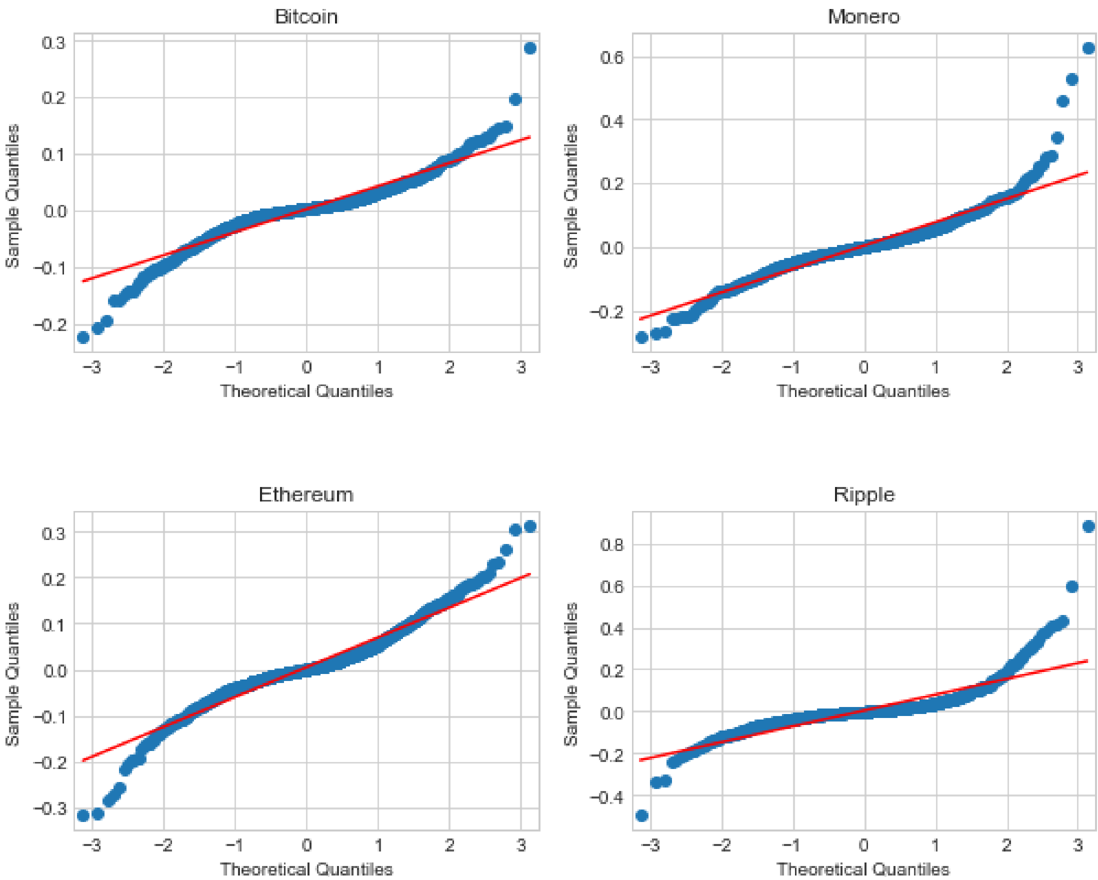


Fig. 4.9: Quantile-quantile plot of log returns of different cryptocurrencies

of the normal distribution. One can immediately see that the log returns have far heavier tails than the normal distribution suggests. Hence, the distributions of cryptocurrencies has to be chosen among the distributions which allow for fat tails. The most commonly used distributions with heavy tails are:

- Pareto distribution,
- Burr distribution,
- Cauchy distribution,
- Lognormal distribution,
- Weibull distribution,
- Frechet distribution,
- T distribution.

Next, these distributions are used to find an appropriate distribution for the log returns of cryptocurrencies. If $x_1, \ldots, x_n$ are independent observations of X, then the maximum-likelihood parameters of each distribution are the values maximizing the likelihood

$$L(\mathbf{\Theta}) = \prod_{i=1}^{n} f(x_i; \mathbf{\Theta})$$

where $\mathbf{\Theta} = (\theta_1, \ldots, \theta_s)$ is a vector of parameters specifying $f(\cdot)$. In order to check which distribution with the optimal parameters implemented gives the best fit, the Kolmogorov-Smirnov (KS) test statistic is used. KS-test performs a test to see if the distribution $G(x)$ of an observed random variable matches a given distribution $F(x)$. The null-hypothesis states that the two distributions are equal.

Using these optimal parameters, one can compare all of the distributions of the specified dataset. If more than one distribution offers a good fit according to the Kolmogorov-Smirnov test statistic, the distribution which offers simultaneously the lowest AIC and BIC value is chosen. AIC/BIC selects the best fit according to maximum likelihood and at the same time punishes for the number of parameters, as shown in Equations (4.2).

$$\begin{aligned} AIC &= 2k - 2\ln\left(L_{max}\right) \\ BIC &= \ln(n)k - 2\ln\left(L_{max}\right) \end{aligned} \tag{4.2}$$

In Equation (4.2) n denotes the number of observations, k is the number of estimated parameters and L_{max} represents the maximum of likelihood function.

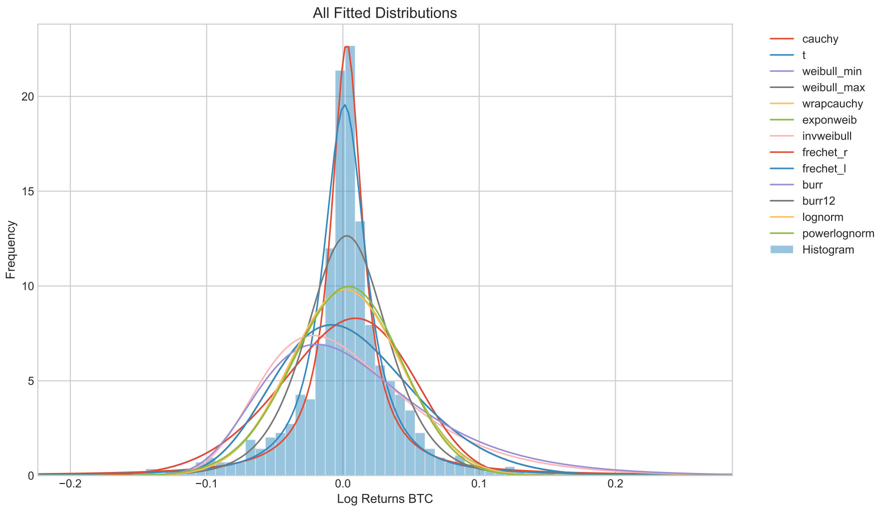
Table 4.3 represents the outcome of the KS-test statistic for all of the data samples. The Kolmogorov-Smirnov test tests if the data comes from a predetermined distribution. In other words, the null-hypothesis is

H_0: the data comes from the specified distribution.

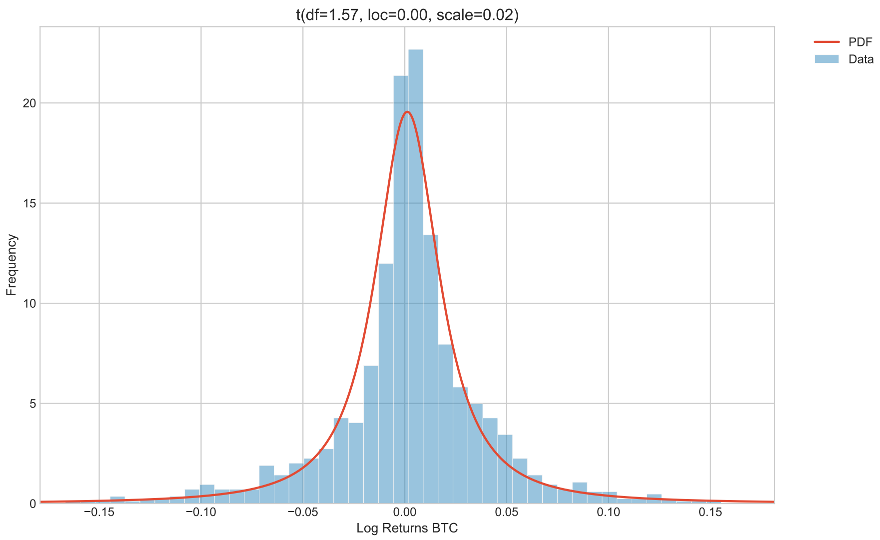
For the log returns of Bitcoin, the null-hypothesis for the t and Cauchy distribution cannot be rejected, since the p-value is larger than the significance level 0.05. Therefore, both distributions offer a good fit. However, the t distribution has a lower AIC and BIC value, therefore this distribution seems currently the most appropriate distribution in terms of fitting the historical time series. Figure 4.10a displays all of the fitted distributions to the empirical data of the log returns of Bitcoin. The distribution which fits the log returns of Bitcoin the best, according to AIC/BIC and the KS-test statistic, is the t distribution with location parameter 0, scale parameter 0.02 and the degrees of freedom are equal to 1.72. Moreover, we find similar results by doing the same analysis for Ethereum, Monero and Ripple their log returns. All the other cryptocurrencies also choose the t distribution as the best fitting distribution according to AIC/BIC and KS-test. These results can be seen visually in Figure 4.11.

	Cauchy distribution		T distribution	
	Test statistic	P-value	Test statistic	P-value
BTC	0.032	0.354	0.038	0.187
ETH	0.041	0.032	0.039	0.055
XMR	0.043	0.009	0.023	0.438
XRP	0.030	0.234	0.030	0.252

Table 4.3: Kolmogorov-Smirnov test on the log returns of different cryptocurrencies

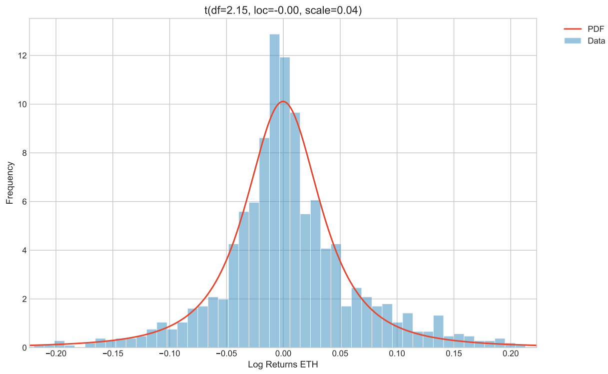


(a) Tests

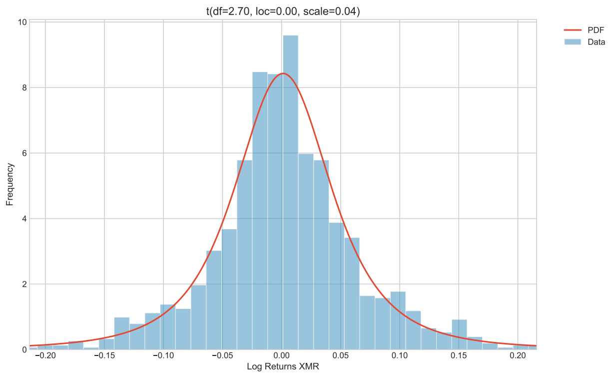


(b) Final

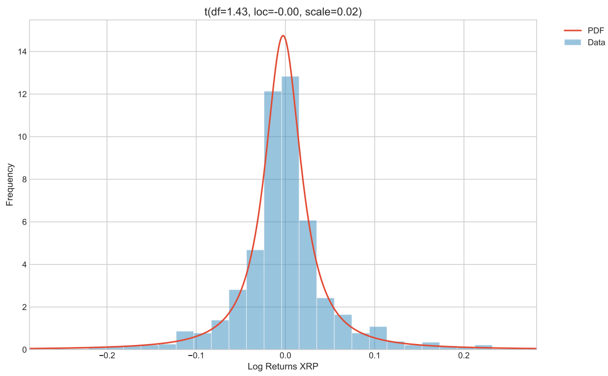
Fig. 4.10: Distribution of Bitcoin's Log Returns



(a) Ethereum



(b) Monero



(c) Ripple

Fig. 4.11: Distribution of log returns

4.6 Volatile Behaviour

Cryptocurrencies are highly volatile assets. Fiat currency, like Yen, Dollar or Euro, do not fluctuate much, while cryptocurrencies have seen their fair share of severe price movements. The highly volatile nature does not allow cryptocurrencies to accurately convey relative prices of goods and services in the economy and leads to uncertainty to its holders regarding its value. In the next chapter, we will have a closer look at the volatile behaviour of Bitcoin.

The analysis is restricted to a gold price index ("XAU Curncy"), a commodity index ("BCOM") and the Euro Dollar exchange rate ("EURUSD") for the non-crypto assets. Figure 4.12 shows the yearly volatility

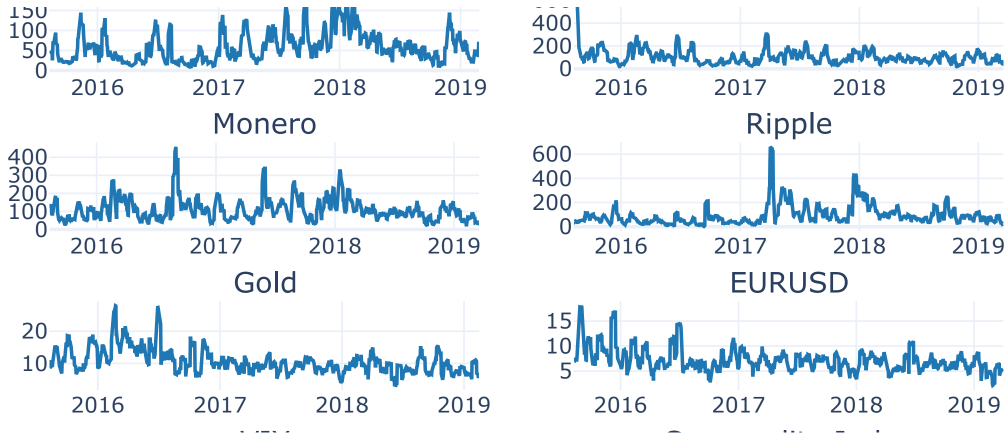


Fig. 4.12: Volatility of different 'securities' in percentage, quantified by yearly standard deviation of the log returns using a rolling window of 10 days.

in percentage of different 'securities' measured by standard deviation of the log returns using a rolling window of 10 days, we also include the VIX for comparison. VIX stands for the Chicago Board Options Exchange volatility index, which measures the implied volatility of S&P500 index options. The daily volatility at time t is calculated as :

$$\sigma_t = \sqrt{\frac{\sum_{s=1}^{w}(r_s - \bar{r})^2}{w}}$$

where $\bar{r}$ is the average return over the window (w), the yearly standard deviation is obtained by multiplying the daily standard deviation with $\sqrt{365}$. The data span a period from 10/08/2015 until 14/02/2019. One thing that is immediately clear from Figure 4.12 is that cryptocurrencies are persistently more volatile than regular securities, some days the yearly standard deviation is ten times higher.

The beginning of 2018 also seems to be the start of a new period volatility-wise for cryptocurrencies. This is around the same time that futures on Bitcoin were traded on exchanges, people now take short-positions on the Bitcoin market more easily and this apparently has calmed down the high volatility.

4.6.0.1 ARMA-GARCH Model

In this section, we will fit an ARMA-GARCH model to the daily log returns of Bitcoin from 01/01/2016 until 17/11/2018, this leaves a total of 1051 observations for the fitting. The data from 18/11/2018 until 25/02/2019 will be used as test data to check the goodness of fit of the ARMA-GARCH model. As mentioned before, the log returns of cryptocurrencies exhibit excess kurtosis and fat tails, these are typical evidences of heteroskedastic effects such as volatility clustering. Moreover, Figure 4.13 represents the squared log returns of Bitcoin, which fluctuate around a constant level. However, they also exhibit periods where large and small changes of log returns cluster together, this can indicate volatility clustering.

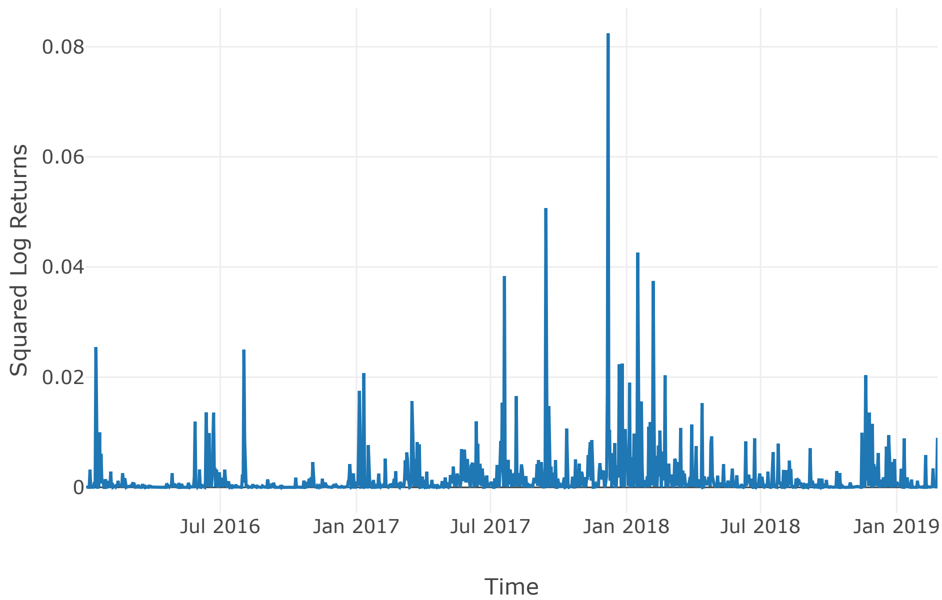
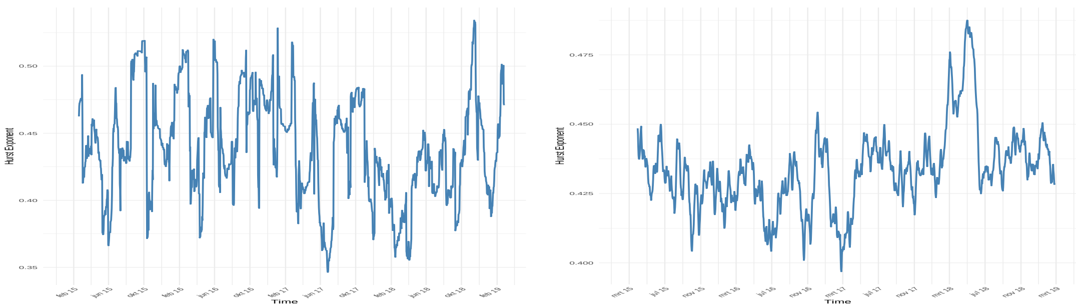


Fig. 4.13: Squared log returns of Bitcoin

We also test for long-run-dependency (LRD) of the returns and the volatility. We will calculate the LRD using the Hurst parameter (Bacon (2008)). Figure 4.14a and Figure 4.14b represent the rolling monthly Hurst parameter of respectively the returns and the volatility of Bitcoin. From these Figures, it can be concluded that the returns and the volatility of the returns have a Hurst exponent smaller than 0.5, this means that the these time-series are anti-persistent. In other words, they fluctuate violently but are mean-reverting.



(a) Returns (b) Volatility

Fig. 4.14: The long-run-dependency calculated over a monthly rolling window.

To further investigate this behaviour, the sample partial and regular autocorrelation function of the squared log returns are plotted in Figure 4.15. The sample autocorrelation function (ACF) and partial autocorrelation function (PACF) plots show significant autocorrelation in the squared log returns of Bitcoin. The Ljung-Box test formally supports this claim, the null-hypothesis of no-autocorrelation is rejected for the first 40 lags, see Table 4.4, hence, there is volatility clustering in the data.

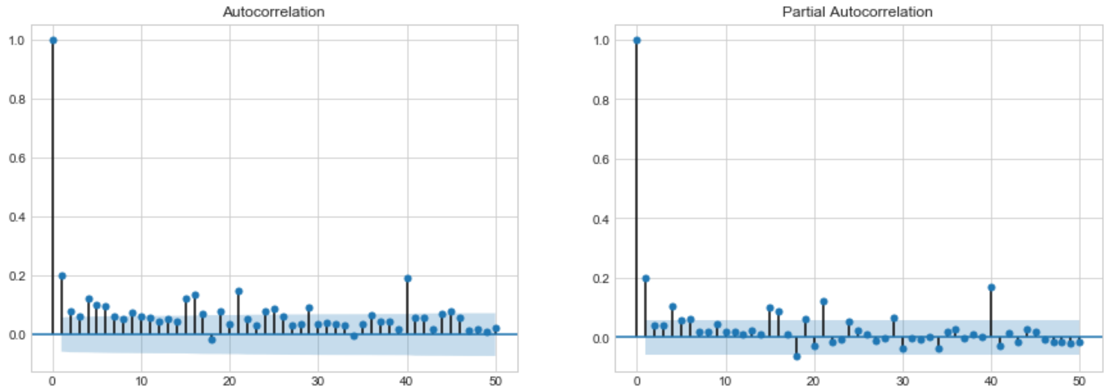


Fig. 4.15: Autocorrelation plots of the squared BTC log returns

The significance of the lags in the PACF and ACF plots shows that both the AR and MA part are needed to capture the behaviour of the mean-process of the log returns correctly. The volatility process is simulated using the GARCH model because the ARMA model assumes a constant variance given past information for the log returns, which is clearly not the case for Bitcoin. The GARCH model corrects the ARMA model for the volatility clustering and fat tails in the data, the fat tails are further corrected by using a t-distribution for the residuals. Therefore, in order to model the log returns accurately, we need both the ARMA and GARCH processes. In other words, the log returns (r_t) follow an ARMA(p_1, q_1)-GARCH(p_2, q_2) process when:

$$\begin{aligned} r_t &= c + \sum_{i=1}^{p_1} \phi_i r_{t-i} + \sum_{i=1}^{q_1} \theta_i \epsilon_{t-i} + \epsilon_t \\ \epsilon_t &= Z_t \sigma_t \\ Z_t &\sim t(\nu) \\ \sigma_t^2 &= \omega + \sum_{i=1}^{p_2} \gamma_i \sigma_{t-i}^2 + \sum_{i=1}^{q_2} \psi_i \epsilon_{t-i}^2 \end{aligned} \tag{4.3}$$

where ϵ_t denotes the residuals with zero mean of the log returns of Bitcoin at time t, it substitutes the part which cannot be predicted and is generated from the GARCH process with parameters ψ_i and γ_i. Z_t is a noise term of the t-distribution with ν degrees of freedom. The mean-model is predicted by an ARMA-

Lag	Test statistic	P-value	Lag	Test statistic	P-value
1	45.646	1.417e-11	21	208.671	7.071e-33
2	52.844	3.349e-12	22	211.740	5.646e-33
3	57.372	2.139e-12	23	213.018	1.001e-32
4	74.404	2.663e-15	24	220.160	1.246e-33
5	85.985	4.678e-17	25	229.144	6.771e-35
6	97.105	1.007e-18	26	233.710	2.682e-35
7	101.477	5.342e-19	27	235.012	4.557e-35
8	104.789	4.468e-19	28	236.667	6.526e-35
9	111.441	7.481e-20	29	246.790	2.134e-36
10	115.900	3.428e-20	30	248.146	3.454e-36
11	119.568	2.214e-20	31	249.899	4.633e-36
12	121.922	2.556e-20	32	251.557	6.417e-36
13	125.237	1.850e-20	33	252.606	1.149e-35
14	127.532	2.085e-20	34	252.615	3.205e-35
15	145.599	1.800e-23	35	254.057	4.717e-35
16	167.101	3.186e-27	36	259.481	1.219e-35
17	173.152	6.647e-28	37	261.729	1.249e-35
18	173.444	1.898e-27	38	263.909	1.310e-35
19	180.988	1.977e-28	39	264.241	3.025e-35
20	182.373	3.329e-28	40	308.823	3.136e-43

Table 4.4: Ljung-Box test for no-autocorrelation on the squared log returns of Bitcoin.

(p_1, q_1)	AIC	BIC
(2, 2)	5925.662	5950.450
(3, 3)	5927.039	5961.742
(3, 2)	5927.602	5957.347
(2, 3)	5927.611	5957.356
(4, 2)	5929.202	5963.905

Table 4.5: AIC/BIC values of the ARMA(p_1, q_1) fit on the log returns of Bitcoin.

model, where ϕ_i are the autoregressive (AR) coefficients, θ_i are the moving average (MA) coefficients and c represents the mean. All the parameters need to be estimated based on the data of the log returns.

First, we fit an ARMA(p_1, q_1) model, the best combination of (p_1, q_1) is chosen according to two information criteria, namely the Akaike (AIC) and Bayesian (BIC) information criteria. Table 4.5 depicts the different AIC- and BIC-values of several combinations of (p_1, q_1), it is clear that ARMA(2,2) offers simultaneously the lowest AIC- and BIC-value and therefore this model is chosen to model the mean process of the log returns.

Table 4.6 gives the specifications of the fitted ARMA(2,2) model; coef denotes the fitted coefficients, Z is the test statistic and the p-value gives an interpretation of the significance of the coefficient and the last column gives a 95% confidence interval of the coefficients. Note that all p-values for the coefficients are smaller than 0.05, hence the coefficients are deemed significant to predict the mean model. The standard error is 1% for the auto regressive part, this indicates that the Bitcoin log returns are strongly mean-reverting.

The squared residuals of the fitted ARMA(2,2) model still display autocorrelation, see Figure 4.16. In fact, this is what we expect because the GARCH-model needs to be used on the residuals. On first sight, the mean process is accurately modeled according to the test data as displayed in Figure 4.17. However, we can only be certain the ARMA-GARCH model is a good fit after the GARCH model is specified.

	Coef.	Std. err.	Z	P-value	95.0% Conf. Int.
ϕ_1	-1.694	0.009	-185.838	0.000	[-1.712, -1.676]
ϕ_2	-0.983	0.010	-102.622	0.000	[-1.002, -0.964]
θ_1	1.702	0.011	148.727	0.000	[1.680, 1.725]
θ_2	0.980	0.011	90.315	0.000	[0.959, 1.001]
c	0.0016	0.000	40.906	0.000	[0.002, 0.002]

Coef. denotes the fitted coefficients, Z is the test statistic and the p-value gives an interpretation of the significance of the coefficient and the last column gives a 95% confidence interval of the coefficients.

Table 4.6: Model results of an ARMA fit to the log returns of Bitcoin.

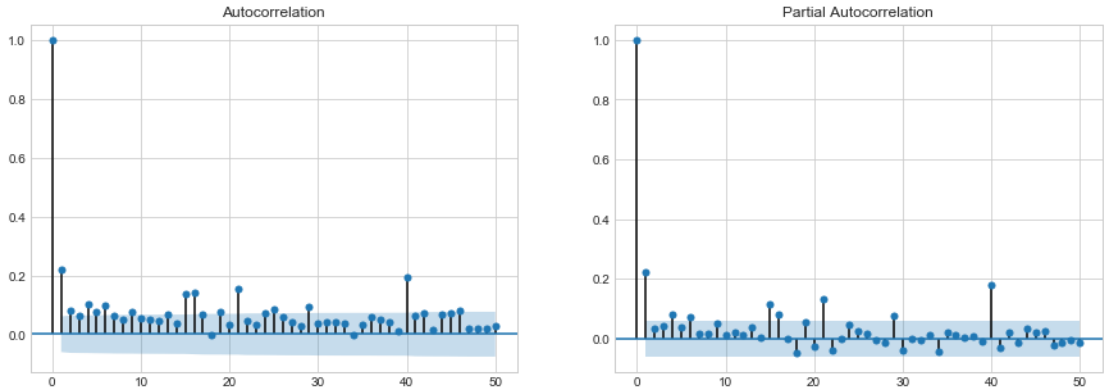


Fig. 4.16: Autocorrelation plot of the model residuals squared

We will now simultaneously fit the ARMA and GARCH models on the data, as explained in Equation (4.3). We will scale the returns by 100 before estimating the ARMA-GARCH model, this helps the optimizer to converge, since the scale of the volatility intercept is much closer to the scale of the other parameters in the model. Moreover, Z_t is given a student t-distribution due to the heavy tails of the distribution of Bitcoin log returns. The lowest AIC/BIC values across different combinations of (p_2, q_2) determine the best fitting GARCH model. Table 4.7 depicts the AIC/BIC values for these different combinations, as a result, (p_2, q_2)=(1,3) provides the best combined AIC/BIC score.

Table 4.8 provides the model specifications for the ARMA(2,2)-GARCH(1,3) model fitted to the data. The ARMA-GARCH model in full (with the scaled returns r_t^* for $t = 1 \ldots n$) is

$$\begin{aligned} r_t^* &= 0.217 - 0.005 r_{t-1}^* - 0.991 r_{t-2}^* - 0.001 \epsilon_{t-1} + 0.993 \epsilon_{t-2} + \epsilon_t \\ \epsilon_t &= Z_t \sigma_t \text{ where } Z_t \sim t(3.340) \\ \sigma_t^2 &= 0.225 + 0.267 \sigma_{t-1}^2 + 0.123 \epsilon_{t-1}^2 + 0.193 \epsilon_{t-2}^2 + 0.416 \epsilon_{t-3}^2. \end{aligned} \tag{4.4}$$

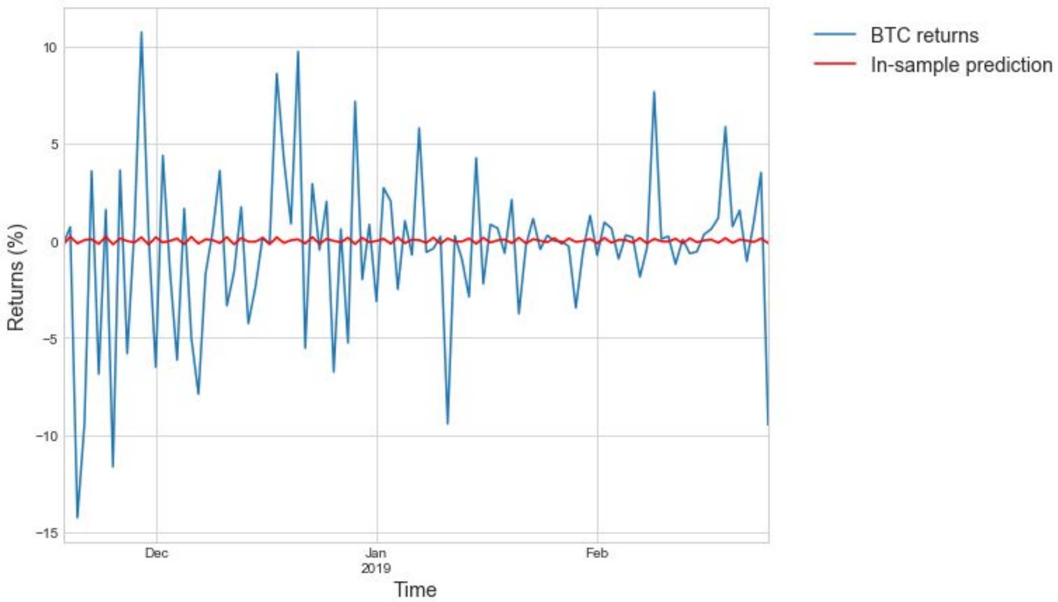


Fig. 4.17: Mean-process of the model

(p_2, q_2)	AIC	BIC
(1, 3)	5365.070	5394.815
(2, 3)	5366.980	5401.682
(3, 3)	5368.980	5408.640
(1, 2)	5369.999	5394.786
(2, 2)	5371.999	5401.744

Table 4.7: AIC/BIC values of the best fitting GARCH(p_2, q_2) model on the log returns of Bitcoin.

Notice that the parameters of the final ARMA(2,2) part are different from the previously specified ARMA(2,2) model because the final model is constructed by simultaneously fitting the ARMA and GARCH part. The sum of the α and ψ coefficients are close to 1, which indicates a high degree of volatility persistence. The estimated GARCH coefficients are all significant except for ψ_1 at a 5% significance level. ϕ_1 and θ_1 are both not-significant at a 5% significance level and both coefficients are also close to zero, this means that they do not differ significantly from zero. Note that θ_2 is highly significant and large, this confirms the hypothesis of the autocorrelation structure.

Table 4.9 represents the Ljunx-Box (LB) test for the squared residuals and the Arch LM (ARCH) test on the standardized residuals, with null-hypothesis':

	Coef.	Std. Err.	T-value	P-value
c	0.217	0.056	3.84794	0.000
ϕ_1	-0.005	0.003	-1.525	0.127
ϕ_2	-0.991	0.003	-329.674	0.000
θ_1	-0.001	0.003	-0.368	0.713
θ_2	0.993	0.000	5118.236	0.000
ω	0.225	0.113	1.989	0.047
γ_1	0.267	0.042	6.309	0.000
ψ_1	0.123	0.108	1.142	0.254
ψ_2	0.193	0.098	1.961	0.049
ψ_3	0.416	0.097	4.301	0.000
ν	3.340	0.259	12.884	0.000

Table 4.8: ARMA(2,2)-GARCH(1,3) model results on the log returns of Bitcoin.

	Test statistic	P-value
ARCH(5)	0.07081	0.7902
LB	0.03135	0.8595

Table 4.9: The Ljunx-Box (LB) test for the squared residuals and the Arch LM (ARCH) test on the standardized residuals, which respectively test the null-hypothesis of no-autocorrelation and no-arch effects.

$$H_0(LB): \text{ The residuals have no-autocorrelation}$$
$$H_0(ARCH): \text{ The residuals have no-arch effect.}$$

The p-value for the two tests is higher than 0.05, hence it is not possible to reject the null-hypothesis in both cases. The chosen model is probably appropriate for the Bitcoin log returns since there are no arch-effects left and the residuals do not exhibit volatility clustering.

The ARMA-GARCH model also allows for Value-at-Risk (VaR) prediction, which can be calculated as

$$VaR(\alpha) = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} t_\alpha(\nu),$$

where α denotes the confidence level, $\hat{\mu}_{t+1}$ and $\hat{\sigma}_{t+1}$ need to be estimated based on the ARMA-GARCH model and $t_\alpha(\nu)$ is the α-quantile of a t-distribution with ν degrees of freedom. The model is able to forecast the VaR, we use a one step ahead forecast with a rolling window of 100 days. Figure 4.18 provides an overview of the 95% VaR forecast of the test dataset, the actual returns and the breaches of the VaR are shown on the graph. There are only 3 breaches on 100 days. We use the conditional (Christoffersen) and the unconditional (Kupiec) test to check the accuracy of the VaR, with null-hypothesis:

$$H_0: \text{ Correct exceedances of the VaR limit.}$$

Table 4.10 shows the outcome of both the conditional and unconditional test for exceedances, both tests have a p-value bigger than the significance level (0.05), therefore, we can say that the VaR limit is accurate.

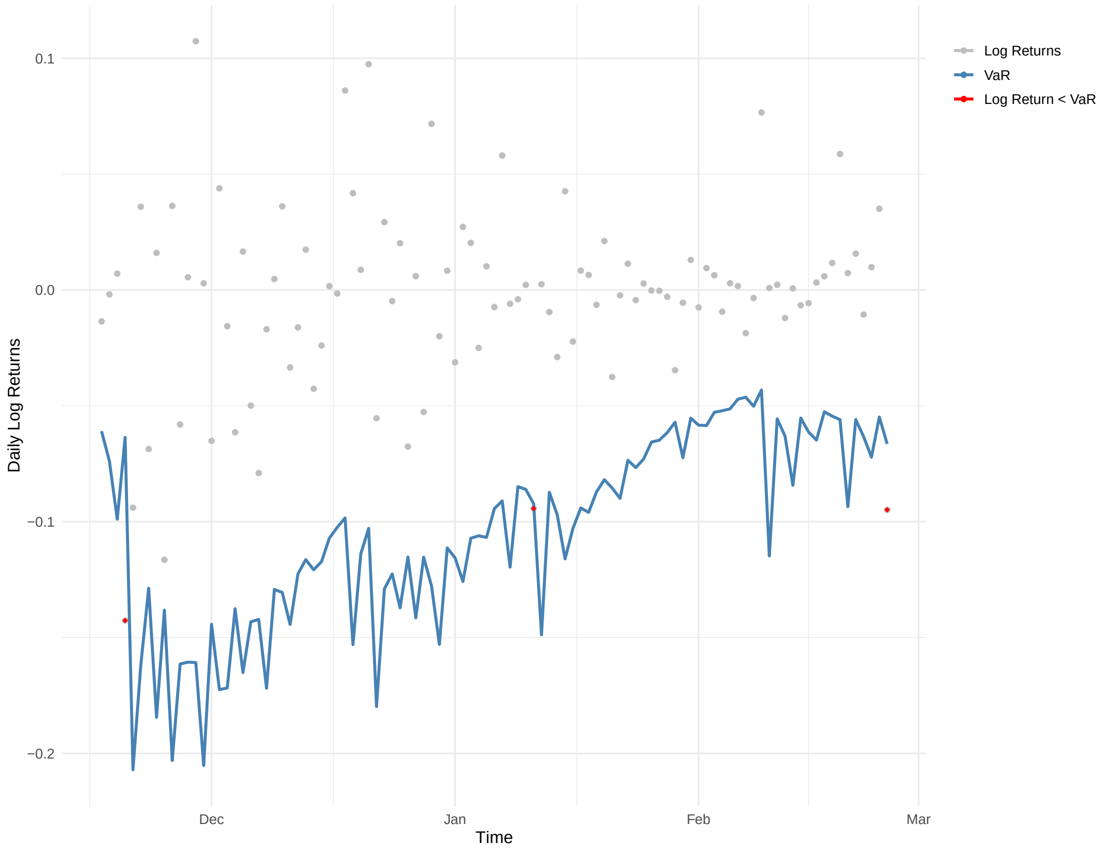


Fig. 4.18: 95% Value-at-Risk forecast using a rolling window of 100 days

	Test statistic	P-value
Kupiec	1.616	0.204
Christoffersen	2.841	0.242

Table 4.10: Results of the conditional and unconditional test for exceedances on the log returns of Bitcoin.

4.7 Market Efficiency

In this section, we will test the efficiency of the cryptocurrency market. The efficient market hypothesis developed by Malkiel and Fama (1970) and offers three types of market efficiency. The weak form of market efficiency says that past returns cannot be used to predict the future. Due to the erratic behaviour of Bitcoin, it is most likely that if cryptocurrency markets are efficient, they will uphold the weak form.

Urquhart (2016) finds that Bitcoin is not weakly efficient, when the hypothesis is tested on a sample from 2013 until mid 2016. We test for the weak form of market efficiency on the logreturns of Bitcoin from 2015 until September 2019. We will follow the approach of Urquhart to test if the Bitcoin market is weakly efficient. The Runs test (Wald and Wolfowitz (1940)) and the Bartels test (Bartels (1982)) check if the returns are independent as null hypothesis. The BDS test (Broock et al. (1996)) is a popular

Runs test	Bartels test	BDS test	AVR test	DL test-C_p	DL test-K_p
0.09	0.68	0.00	0.65	0.22	0.24

Table 4.11: P-values of the test results of the weak form of market efficiency.

non-parametric test for serial dependence where the null-hypothesis states that the returns are i.i.d. and the alternate hypothesis tells that the model is misspecified. The AVR test (Choi (1999)) determines if the returns are performing a random walk and the variance of price difference of order q is p times the variance of the price difference (p and q are determined based on the data). The Dominguez and Lobato test (DL test) (Dominguez and Lobato (2003)) has as null hypothesis that the returns follow a martingale difference process. The p-values of each test are depicted in Table 4.11.

From Table 4.11 one can see that the null-hypothesis of independence cannot be rejected by both the Bartels and Runs test on a 5% significance level. The BDS test rejects the fact that returns are i.i.d. on a 5% significance level. The null hypothesis of the DL test cannot be rejected, therefore the returns might follow a martingale difference process, also the AVR test cannot be rejected. The tests thus suggest that the returns of Bitcoin exhibit weak efficiency, which is also a conclusion of Wei (2018). Moreover, according to Wei (2018), liquidity plays a role in the market efficiency and return predictability.

References

Alexander, C. and Dakos, M. (2019). A critical investigation of cryptocurrency data and analysis. Technical report.

Bacon, C. (2008). *Practical Portfolio Performance Measurement and Attribution, 2nd Edition*, volume 546. John Wiley & Sons.

Bartels, R. (1982). The rank version of von neumann's ratio test for randomness. *Journal of the American Statistical Association*, 77(377):40–46.

Broock, W. A., Scheinkman, J. A., Dechert, W. D., and LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric reviews*, 15(3):197–235.

Burniske, C. and White, A. (2017). Bitcoin: Ringing the bell for a new asset class. Technical report.

Choi, I. (1999). Testing the random walk hypothesis for real exchange rates. *Journal of Applied Econometrics*, 14(3):293–308.

Dominguez, M. A. and Lobato, I. N. (2003). Testing the martingale difference hypothesis. *Econometric Reviews*, 22(4):351–377.

Kenton, W. (2018). Flash crash. https://www.investopedia.com/terms/f/flash-crash.asp.

Malkiel, B. G. and Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417.

SFox (2019). https://blog.sfox.com/bitcoin-flash-crash-short-history-cbcb0d6bc6a5.

Urquhart, A. (2016). The inefficiency of bitcoin. *Economics Letters*, 148:80–82.

Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are from the same population. *The Annals of Mathematical Statistics*, 11(2):147–162.

Wei, W. C. (2018). Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, 168:21–24.