Alexander Söllinger, BSc

Cryptoasset Return Modelling

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Mathematics

submitted to

Graz University of Technology

Supervisor

Gunther Leobacher, Univ.-Prof. Dr.
Karl-Franzens-Universität Graz

In cooperation with
Michael Kratochwil, Dr.
Dr. Nagler & Company

Graz, June 2024

Contents

List of Figures

Abstract

Cryptocurrencies have been more popular than ever during the COVID-19 pandemic, but shortly afterwards plummeted again in price. To mathematically capture the phenomenon of these rapidly changing prices, various approaches were carried out. The first part of this paper is dedicated to an in-depth literature review on the approaches that have been used and a brief summary of their findings. As a result of this analysis a certain model class called the GARCH model emerges as the most popular tool to describe model prices of cryptocurrencies. Subsequently, the most important models of this class are then derived and a popular subset of the class is introduced. For these, stationarity conditions are given, and their parameters are discussed. Additionally, an estimation method is given for the main models. Furthermore a short overview of the history of cryptocurrencies, especially Bitcoin, is given, demonstrating the high volatility of the cryptomarket. The last chapter uses the derived class of GARCH models to find the best univariate model for Bitcoin prices and the best multivariate model for six of the largest cryptocurrencies. The univariate models are used to analysze the price movements of Bitcoin and also to simulate European call option prices, which are, unsurprisingly, found to be exorbitant. The estimation of the mulivariate models gives insight into the correlation between the cryptocurrencies, suggesting that diversification within the cryptocurrency market is impossible.

1 Introduction

Cryptocurrencies had a bumpy ride during the last three years; Bitcoin alone reaching a market capitalization of over 1.2 trillion USD during November 2021 is impressive, but plummeting to below 400 billion USD, right after the COVID-19 crisis passed, was an equally impressive downswing. Predicting these heavy fluctuations in price is a challenging task, but if successful could yield a calculatable risk for investments in cryptocurrencies. Which in turn could increase their use to a wider variety of financial institutions. Of course, getting a handle on the risk of cryptocurrencies has already been attempted numerous times and with a variety of mathematical models. This is done in similar ways as for other assets, which is one of the reasons why they will be referred to as cryptoassets. With the other reasons being, that they are no stable value storage and that the Basel Committee refers to them as such [Committee, 2022]. Even though cryptoassets are similar to other assets, cryptoassets display unusually high volatility, which makes calculating a prudent but at the same time manageable risk reserve such a difficult task. First, an overview of previous attempts at modelling cryptoassets is given, which shows that not all models perform equally well over different time horizons. Most frequently, autoregressive models are used. These models not only have the advantage of their idea being straightforward, but also show good performance due to their flexibility. In particular, the GARCH model class, which exclusively models the volatility of a process, is a commonly used modelling tool, which can be used in combination with AR models, but is more frequently used without them. In shorter time periods the actual rise, which has mostly been the trend of successful cryptoassets, is less significant and favors a pure GARCH model even more. Moreover, this rising trend is not a given, since cryptoassets do sometimes fall from favour, but modelling is usually limited to those that are currently on top and therefore often gives the impression that all cryptoassets are rising in price. Therefore, a shorter time period will also be chosen here. Concerning the aforementioned flexibility, the class of GARCH models offers a number of different parametrizations, some of which are introduced in this thesis. First, a class of univariate models is introduced and an estimation method of the standard GARCH model is provided. Then multivariate GARCH models are presented, but only the CCC/DCC class is considered, since it features computational advantages over the costly VEC-GARCH model and additionally, is able to capture nonlinearities in the squared returns. After the introduction, a short detour into some history of the cryptoasset market is made. Here primarily Bitcoin is featured, which is a bias that is found throughout the whole paper. As Bitcoin is not only the cryptoasset with the highest market capitalization, but also the oldest one still active, it has a prominent place in the cryptoasset market and therefore deserves increased attention. In addition, for many who are not familiar with the topic, the words “cryptocurrency” and “Bitcoin” are somewhat of a synonym. This is due to the sustained reign of Bitcoin over the list of highest capped cryptoassets, while the ranks

below it have been subject to changes over the last years. Moreover, each asset has its own history and special features, but Bitcoin is a good candidate for what is commonly expected of a cryptoasset, since it utilizes the Blockchain technology and features the proof of work concept from [Nakamoto, 2008] which makes it decentralized. After the detour, the GARCH models are fitted to different cryptoassets. The univariate models are only fitted to the Bitcoin data. The best-performing univariate model is also used to calculate option prices, which reflect the volatility of the underlying asset. Risk management is briefly mentioned, but long-term investments with borrowed capital do not seem to be possible when adequately accounting for necessary reserves. Lastly the DCC-GARCH model is fitted to the data of the six cryptoassets Bitcoin, Ethereum, BNB, XRP, Cardano and Dogecoin. This model shows a strong positive correlation between the six assets, which reflects a behaviour similar to the volatility of the individual assets, i.e. both are GARCH models with high persistance. The thesis is structured in the following way: Section 2 contains the literature review, split in the two subcategories univariate modelling and multivariate modelling. Section 3 introduces the univariate GARCH models and possible estimation of the standard GARCH model. In Section 4, the multivariate GARCH models are defined and the CCC- and DCC-GARCH models are discussed with more depth. The short history of the cryptoasset market with the focus on Bitcoin is displayed in Section 5. In Section 6, the models that were introduced in Section 3 and 4 are fitted to available datasets and future returns are simulated. With those risk margin and option prices are derived.

2 Literature review

2.1 Univariate Analysis

Univariate models focus only on past values of the returns of the cryptoassets and use no further market related information. The advantages of univariate models come from their simplicity. Therefore, less assumptions have to be made and less computational power is required. This gives the opportunity to use more advanced models or to reduce the number of necessary parameters.

Source	Data Period	Methodology	Purpose and Findings
[Ardia et al., 2019]	18.08.2011-03.03.2018	Markov-switching GARCH models	Tests for regime switches in GARCH models for the volatility of Bitcoin. Therefore, two different GARCH models (GJR and normal) are tested with 1/2/3 regimes. The 2-regime models have the best in sample performance. Further it is shown that regime switching drastically improves Value-at-Risk forecasts.
[Phillip et al., 2018]	-31.07.2017	Generalized Long Memory SV model	In this paper some attributes of crypotassets are directly implemented in the model. These include long memory, leverage effect and heavy tails. This model is implemented for a total of 224 cryptoassets, which is the reason why the starting dates differ.
[Catania and Grassi, 2017]	2013-2019	Score-driven GHSKT model	Implements several different generalizations of a score driven model using a GHSKT distribution for the innovation. This is done for 606 different cryptoassets, which are selected based on time series length ($\geq$ 700 days). It is found that robust filters are necessary and time-varying skewness improves predictions. Further concern is expressed about the use of non-robust GARCH models.
[Chu et al., 2017]	22.06.2014-17.05.2017	various GARCH models	Twelve different GARCH models with eight different innovation distributions for six different cryptoassets are compared. The IGARCH and GJR-GARCH provide the best fit for the data, but concerns about structural changes influencing the results are expressed. Furthermore, Value at Risk is studied, and it is found that acceptable estimates of the measure can be obtained.

Source	Data Period	Methodology	Purpose and Findings
[Madan et al., 2019]	29.06.2018-31.08.2018	Black-Scholes model Laplace model Heston model Variance gamma models	Uses vanilla option prices to calibrate the models. Black-Scholes is found to not capture the dynamics very well, but the more advanced models produced a good fit. Lastly the liquidity of Bitcoin options is analysed. It is found that in-the-money long-term options are the most liquid.
[Hou et al., 2019]	31.07.2014-29.09.2017	SVCJ model	The SV, SVJ, SVCJ and a generalization of the SVCJ model are implemented for Bitcoin returns as well as for the CRIX index. For these models, the Bitcoin option prices are obtained via Monte Carlo Simulation. It is found that including jumps and co-jumps is significant for those prices. A comparison between the models yielded, that the SVCJ model mostly performs as well as its generalization. Another finding was that jumps in volatility and returns are anti-correlated.
[Guo, 2022]	18.12.2017-04.12.2020	GARCH model IGARCH model GJR-GARCH model EGARCH model TGARCH model	Considers Bitcoin futures, which are grouped in three categories (by maturity). Then several GARCH models are implemented with two different distributions, the normal distribution and the NIG distribution. Further Value at Risk forecasts are calculated. The results indicate that the heavy tailed GARCH models outperform the not heavy tailed ones and in case of the Value at Risk estimates the non-parametric estimates are outperformed as well.
[Troster et al., 2019]	19.07.2010-16.04.2018	various GARCH models GAS model	Implements eight types of GARCH models and the GAS model with five different distributions. In the comparison of them the heavy tailed GAS models outperform all GARCH models and the normal GAS model. Further Value at Risk forecasting is done, where again the heavy tailed GAS models perform the best.
[Chu et al., 2015]	13.09.2011-08.05.2014	fitting distributions	At first it is argued that the data is almost i.i.d., then 15 different distributions are fitted to the dataset via Maximum Likelihood Estimation. The generalized hyperbolic distribution is found to be the best fit to the data. Confidence intervals for future predictions are very large, which is attributed to Bitcoins high volatility.

Source	Data Period	Methodology	Purpose and Findings
[Katsiampa, 2017]	18.07.2010-01.10.2016	AR-GARCH model AR-EGARCH model AR-TGARCH model AR-APARCH model AR-CGARCH model AR-ACGARCH model	Studies several AR-GARCH models. The AR-CGARCH model has the best goodness of fit, which suggests that there is a short-run and a long-run component in the conditional variance.
[Cheikh et al., 2020]	28.04.2013-01.12.2018	ST-GARCH model	For four cryptoassets, namely Bitcoin, Ethereum, Ripple and Litecoin, a smooth transition GARCH model is implemented and compared with four other GARCH specifications. For most assets, the ST-GARCH model outperforms the others and strong asymmetry is found, where positive shocks increase the volatility more than negative shocks. Further the Ethereum time series does not start until 07.08.2015.
[Bouri et al., 2017a]	18.08.2011-29.04.2016	AGARCH model	Studies the impact of the price crash in 2013 on the save haven properties of Bitcoin. It is found that Bitcoin was only a safe haven before the crash.
[Matic et al., 2021]	01.01.2017-01.08.2020	SVCJ model KDE-GARCH	First, SVI-implied volatility surface is calibrated and used to price cryptoassets. Then the two models are used to generate Monte Carlo paths. These are in turn used to hedge options using a variety of hedging techniques. Results indicate, that a SVCJ model with low jump frequency is the best fit.
[Bouoiyour and Selmi, 2016]	01.12.2010-22.07.2016	various GARCH models	A total of nine GARCH models are implemented and compared. Then the time period is split in two at 31.12.2014. For the first period the Component with Multiple Threshold (CMT)-GARCH model is found to be the best performing model, while for the second period it is an asymmetric power GARCH model. It is found that from 2015 onward, the bitcoin returns become way less persistent, which argues against a long-memory property of the returns. Further the reaction to negative shocks is greater than to positive ones. In conclusion, it is suggested that while the volatility at the end of the researched time period is low, the bitcoin market was still far from being mature.
[Bouoiyour and Selmi, 2015a]			Prior version of the above using a smaller time horizon and a TGARCH model for the first and an EGARCH model for the second time period. It also leads to similar conclusions.

Source	Data Period	Methodology	Purpose and Findings
[Mensi et al., 2019]	01.07.2011-03.03.2018	ARFIMA-GARCH model ARFIMA-FIGARCH model ARFIMA-FIAPARCH model ARFIMA-HYGARCH model	Implements four types of GARCH models to explore the impact of long-term memory and structural breaks on the conditional volatility of Bitcoin and Ethereum (data starting at 09.08.2015). It is found, that without accounting for the two aforementioned properties volatility persistence is overestimated. Also forecasting is improved if structural breaks are taken into account. The best performing model was the FIGARCH model.
[Naimy and Hayek, 2018]	01.04.2013-31.03.2016	GARCH model EGARCH model	The EGARCH model outperforms the GARCH model.
[Venter et al., 2020]	01.01.2016-03.01.2020	GARCH model GJR-GARCH model	Both models are applied to Bitcoin returns and the CRIX index; these are used to create implied volatility surfaces for option prices. The results suggest that there is no asymmetry in the option prices, furthermore the obtained prices are consistent with the current market prices.

2.2 Multivariate Analysis

Within the multivariate analysis there is an additional differentiation. The first kind of multivariate papers examines the external factors influencing the cryptoasset market, they are called the price drivers of cryptoassets. The second kind attempts to understand the relations of cryptoassets among each other. For both kinds of papers, the most usual modelling approach are autoregressive time series models, usually some types of GARCH and ARDL models. Other approaches include SV models, wavelet analysis, regression analysis, Bayesian structural time series and machine learning.

2.2.1 Crypto price drivers

Many different time series were checked for a relation to cryptoassets. On the one hand internal factors, i.e. variables in direct relation to the cryptoassets themselves, and on the other hand factors only correlated through the economy.

Internal Factors

Hashrate: A hash is the evaluation of a hash function, which is used in proof of work networks such as Bitcoin. The hashrate describes how much computing power is currently invested in mining a cryptoasset.

Volume: The total amount of a cryptoasset that exists.

Trade Volume: The total amount of a cryptoasset that is traded within a certain time period.

Transaction Volume: The total amount of a cryptoasset that is used to purchase goods or services (not involving a crypto exchange) within a certain time period.

Google/Wikipedia searches: The amount of people that searched for a cryptoasset or a crypto related term.

Number of addresses: The total amount of addresses participating in a cryptoasset network.

Estimated output volume: The total amount of new cryptocurrency.

Difficulty: The difficulty of creating a new block, i.e. how much computing power is needed per new block.

Velocity: The number of times one coin is expected to be traded in a certain time interval.

Total number of transactions

Number of tweets

Twitter sentiment

Conformation time of transactions

External Factors

Commodities: Assets such as gold and oil are among the most frequently used factors for modelling cryptoassets.

Stock indices: Most often the S&P 500 is used, but many different market indices have been compared to cryptoassets.

Exchange rates: Most often USD/EUR or USD/GBP are used, but others appear as well.

Other financial indicators and assets: There are several other figures representing the market as well, like future prices or bond yields.

Source	Variables	Data Period	Methodology	Purpose and Findings
[Kjærland et al., 2018]	S&P 500 Hashrate Volume Gold Oil VIX Google searches	01.01.2013-20.02.2018	ARDL Model GARCH Model	Finds that S&P 500, BTC Trading Volume and the number of Google searches are relevant for predicting Bitcoin returns. The other variables were all insignificant.
[Dyhrberg, 2016a]	Federal funds rate USD/EUR USD/GBP FTSE index Gold futures Gold cash	19.07.2010-22.05.2015	GARCH model EGARCH Model	Conducts analysis if BTC behaves like a currency or a commodity. They find indications of both. The strong reaction to the federal funds rate is usually observed for currencies, but its response to exchange rates is similar to a commodity.
[Aalborg et al., 2019]	Volume Transaction volume Number of addresses VIX Google searches	01.03.2012-19.03.2017	linear Regression	Analysis on predicting returns, trading volume and volatility. Changes in unique addresses are positively correlated with the BTC return. Further they find that trading volume has only weak predictive power for the returns. Trading volume itself is influenced by transaction volume and the Google searches. Realized volatility is used to calculate the volatility from the returns and it is found that on a daily level it is correlated with the trading volume, which fails to extend to weekly data.
[Bouri et al., 2018b]	S&P GSCI Commodity Ounce of gold MSCI world PIMCO investment grade bond US dollar index	17.07.2010-02.02.2017	ARDL model NARDL model QARDL model QNARDL model	Finds asymmetric nonlinear relationships between BTC and two of the modelling variables, namely gold and aggregated commodity. This is not in line with several other studies and implies that BTC is not an isolated economy, leading to the conclusion that BTC is not a safe haven.
[Georgoula et al., 2015]	Total numer of transactions USD/EUR S&P 500 Hashrate Wikipedia searches Google searches Number of tweets Twitter sentiment ratio	27.10.2014-12.01.2015	Machine learning vector space machine Vector Error-Correction model	Finds positive short-term correlation between BTC media popularity and its price. Notably a positive correlation with the hashrate is found, in contrast to [Kjærland et al., 2018]. Furthermore a negative short time correlation to the USD/EUR exchange rate is found. The same is true for the S&P 500 index, which is attributed to people selling bad performing stocks to invest into BTC and vice versa. Lastly a positive correlation with the total number of BTC is found.

Source	Variables	Data Period	Methodology	Purpose and Findings
[Bouri et al., 2018a]	MSCI world MSCI Emerging markets MSCI China S&P GSCI commodity S&P GSCI energy Ounce of gold US Dollar index US 10 year treasury yields	19.07.2010-31.08.2017	smooth transition VAR - bivariate GARCH-in-mean	Studies spillovers in returns and volatility between BTC and other traditional assets and currencies. Findings indicate that the connection is stronger for returns than for volatility. Further asymmetric spillovers are found in both bull and bear markets and their effects vary conditional on the situation. Also, BTC usually is the recipient of these spillovers. So, it is possible to predict the volatility of BTC from the other volatilities, but not the other way around. For investors this means that not only is it necessary to observe the state of the market, but it is also necessary to differentiate between different markets in order to make an informed decision concerning BTC.
[Kristoufek, 2015]	Total number of BTC Number of transactions Estimated output volume Ratio trade volume/transaction volume Hashrate Difficulty BTC exchange rates Google searches Wikipedia searches Financial Stress index Gold	14.09.2011-28.02.2014	Wavelet coherent analysis	First it is found that trade, money supply and price level do influence the BTC price in the long-run. Secondly, a rising price should lead people towards starting to mine, but miners are found to be vanishing due to the increase in the difficulty and hashrate because of specialized hardware. The third finding is, that the interest in cryptocurrency is also driving the price of BTC, especially in the long-run. Fourth, Bitcoin is not a safe haven. Lastly no evidence is found that the Chinese and US markets are connected.
[Bouoiyour and Selmi, 2015b]	Google searches Trade volume/transaction volume Velocity Estimated output volume Hashrate Gold Shanghai market index	05.12.2010-14.06.2014	ARDL model Granger causality	They find that Google searches, Shanghai market index and trade/transaction ratio all have significant positive short-term impact (in descending order). But in the long run all but the trade/transaction ratio cease to have significant influence and even the trade/transaction ratio loses importance. Furthermore, the hashrate becomes significant. In conclusion, the long-term future is very difficult to predict and can be influenced by certain factors, whose impact are difficult to foretell (like the reaching of the maximum amount of BTC).

Source	Variables	Data Period	Methodology	Purpose and Findings
[Sovbetov, 2018]	Crypto 50 index Crypto 50 Volume Crypto 50 volatility S&P 500 Gold U.S. interest rates Google searches EUR/USD	2010-2018	ARDL model	The first contribution of this paper is the generation of a Crypto 50 index, that is the weighted average over 50 cryptoassets. Further, this paper focuses on five cryptoassets. It is found, that in the short- and long-run Bitcoin and Ethereum respond more to the crypto market than the other three cryptoassets. Trading volume is significant and positive for all five assets, but its significance is greater in the long-run. Volatility is also always significant and negative, but its short-run impact is larger. Google searches are only significant in the long-run and not significant at all for Dash. The S&P 500 shows a weak positive long-run relationship, but short-run it has only a minor negative influence on Bitcoin. Lastly the Error Correction Term was significant for all five cryptoassets and the strongest correction was displayed by Bitcoin.
[Bouri et al., 2017b]	S&P 500 FTSE 100 DAX 30 Nikkei 225 Shanghai A-share MSCI World MSCI Europe MSCI Pacific PIMCO Investment Grade Corporate Bond index US dollar index SPGS Commodity index Oil Gold	18.06.2011-22.12.2015	DCC-MGARCH	This paper applies the DCC-MGARCH model in order to find if BTC is a diversifier, hedge and/or a safe haven. They use weekly and daily data and show that the results on the safe haven and hedge properties differ between the different time horizons. In most cases, diversification is found to be possible.
[Guesmi et al., 2019]	MSCI Emerging Markets index MSCI Global Market index USD/EUR USD/CNY Gold Oil VIX	01.01.2012-05.01.2018	VARMA-DCC-GJR-GARCH model	First, several VARMA-GARCH models are compared and the beformentioned is chosen due to its AIC value. Further it is found that a short position in Bitcoin is a good hedge against the other variables. The hedge with Bitcoin is especially effective for a portfolio containing gold, oil and the MSCI Emerging Markets index.

Source	Variables	Data Period	Methodology	Purpose and Findings
[Nguyen et al., 2019]	Oil Gold S&P 500 LIBOR USD index	08.08.2014-07.06.2017	DCC-MGARCH model	Examines diversification options of seven cryptoassets against economic factors. Structural breaks and ARCH disturbances are found for each cryptoasset. Further, insignificant correlations with the economic factors are found, implying low hedge capabilities.
[González et al., 2021]	Gold	26.01.2015-30.06.2020	NARDL model	Studies the impact of the gold price on 12 different cryptoassets. Furthermore, two sub periods are analysed. One is the epicentre of the first COVID-19 wave (01.03.2020-30.06.2020) and the other one includes the build up as well (01.01.2020-30.06.2020). The first finding is that all cryptoassets except for Tether show significant positive dependencies on the gold price for all periods. This correlation increases particularly during the epicentre of the crisis. Second, cointegration increases in the COVID-19 periods. Third, the long run elasticities for the entire period are only significant for Bitcoin SV, while during the COVID-19 periods they become drastically more significant, and are generally positive, except for Tether. Fourth, long-run asymmetry can be observed only during the COVID-19 periods, but short-run asymmetry is not only present during the COVID-19 crisis, but also for ten of twelve observed assets in the entire period. Fifth, the cryptoassets show high persistence to changes in the gold price. Finally, the NARDL model explains a highly increased amount of cryptoasset returns during the COVID-19 period, which suggest that especially in times of economic turmoil, the connectedness of cryptoassets to the gold price increases.
[Jareño et al., 2020]	Gold Crude Oil S&P 500 VIX STLFSI index	08.2010-11.2018	Quantile Regression NARDL model	Analyses the sensitivity of Bitcoin returns to returns of other assets and indices. The reaction of the Bitcoin returns to the mentioned variables tends to be stronger in extreme market conditions. Further, a positive significant connectedness with gold is found.

Source	Variables	Data Period	Methodology	Purpose and Findings
[Klein et al., 2018]	Gold Silver Crude Oil S&P 500 MSCI World index MSCI Emerging Markets 50 index	01.07.2011-31.07.2017	APARCH model FIAPARCH model BEKK-GARCH model	Analyses the similarities between gold and Bitcoin (and CRIX). A t-distribution is chosen as the underlying distribution, which makes the FIAPARCH model the best performing one. Further, Bitcoin has the same asymmetric response to market shocks as precious metal, but it is very persistant to variance shocks. Also, Bitcoin was shown to decline as the market declines, which is contrary to gold. This implies that Bitcoin is not an effective hedge against the markets.
[Poyser, 2017]	USD trade volume on BTC exchanges Conformation time of transactions Hashrate Transactions per day Google searches (countrywise) S&P 500 CBOE VIX bearish sentiment (AAII survey) Gold EUR/USD USD/CNY	01.2013-05.2017	Bayesian Structural Time Series	Applies different takes of the Bayesian Structural Time Series approach to the topic. Finds negative correlation to the gold price and USD/CNY exchange rate, but a positive one for USD/EUR exchange rate. The trends in the different countries show varying signs, and their relevancy varies over time. The internal factors have no relevant impact.
[Symitsi and Chalvatzis, 2019]	US Dollar index Various Gold indices Various Oil indices Exchange rates Dow Jones Industrial Average S&P 500 Housing Sector index 30 year Treasury Bond index	20.09.2011-14.07.2017	DCC-GARCH model with univariate GJR-GARCH model	Uses three different models for gold (silver) and oil respectively, and seven exchange rates (all containing USD). Performs analysis on portfolios with and without Bitcoin in different states of the market. It is found, that portfolio risk is decreased by adding Bitcoin due to its low correlation with the other assets considered. However, these benefits decrease, if the portfolio already contains a wider range of assets.
[Charfeddine et al., 2020]	Gold Crude Oil S&P 500	18.07-2010-01.10.2018	ARFIMA-FIAPARCH model Time varying-copulas DCC-GARCH model	Investigates the relationship between conventional financial assets and two cryptoassets, namely Bitcoin and Ethereum. Time-varying dependence is found via the Copula approach. Further diversification benefits are found, but the optimal portfolio only includes a small weight in Bitcoin and Ethereum. Further hedging capabilities are found to be weak.

Source	Variables	Data Period	Methodology	Purpose and Findings
[Catania et al., 2019]	S&P 500 Nikkei 225 Stoxx Europe 600 Gold Silver 5 year Europe credit default swap US 1 month treasury yields US 10 year treasury yields VIX	08.08.2015-28.12.2017	TVP-VAR models KS models AR-EWMA model	Studies a variety of different univariate and multivariate models on four major cryptoassets. Further model combinations are made, where it is found that combinations of univariate models provide improved point forecasting, while density forecasting is improved by combining multivariate models.
[Dwita Mariana et al., 2021]	Gold S&P 500	01.07.2019-06.04.2020	DCC-GARCH model	Bitcoin and Ethereum are tested for safe haven properties in reference to the COVID-19 pandemic announcement. Findings indicate that both are short-term safe havens, with Ethereum being the better one, but this advantage comes at the price of higher return volatility.
[Stensås et al., 2019]	7 stock indices for developed markets 6 stock indices for developing markets 5 MSCI indices 4 commodity indices 6 commodity prices	13.09.2011-01.01.2018	DCC-GARCH model	Studies diversifier, hedge and safe haven properties of Bitcoin for developed and developing markets as well as for commodities. Bitcoin is found to be a hedge in most developing countries, for all other considered variables it acts as a diversifier. Further, save haven properties are only observed during periods of high uncertainty.
[Dyhrberg, 2016b]	USD/EUR USD/GBP FTSE index	19.07.2010-22.05.2015	AGARCH model TGARCH model	Researches Bitcoins similarities to gold by studying its responses to economic factors. Findings indicate, that Bitcoin can be used as a hedge against the FTSE index and the Dollar, though for the Dollar only in the short-term.

2.2.2 Inner relations of the cryptoasset market

The crypto market contains a huge battery of different assets, the listed papers consider between 3 and 17 of them for correlations. Bitcoin is the only asset considered in all papers; the next most popular ones are Ethereum and Litecoin.

Source	Cryptoassets	Data Period	Methodology	Purpose and Findings
[Ciaian et al., 2018]	Bitcoin Litecoin Dogecoin Monero Ripple DigitalCash NEM Peercoin Bitshares Nxt Namecoin Novacoin CounterParty Qora Mintcoin Feathercoin Primecoin	2013-2016	ARDL model	The paper analyses the interdependencies between Bitcoin and the altcoin market. To do so, Bitcoin, 16 Altcoins and two Altcoin indices are used in the model. It is found, that Bitcoin and the altcoin market are strongly interdependent in the short-run, but in the long-run the macro-financial indicators outweigh the impact of Bitcoin on the altcoin market (indicators used are Wikipedia searches, the total volume, the gold price, the NASDAQ index, US 10-year treasury yields, EUR/USD and USD/CNY exchange rates and the oil price). Further, the total volume has been found insignificant on the returns of the cryptoassets.
[Chaim and Laurini, 2019]	Bitcoin Ethereum Ripple Litecoin Stellar Dash Monero NEM Verge	16.08.2015-31.10.2018	Multivariate SV with jumps in mean and volatility	Describes the return and volatility dynamic between nine cryptoassets. Finds volatility periods in 2017 and early 2018, further jumps become larger and more frequent from 2017 on. Also, the long-term memory dependence is well represented in the chosen model.
[Hu et al., 2021]	Bitcoin Ethereum XRP Litecoin Bitcoin cash EOS Binance coin	25.07.2017-02.07.2019	Univariate ARMA-GARCH with multivariate innovations	Tests different marginal and joint distributions for the sample innovations from the GARCH part of the model. The best performing model is the one with gaussian marginals and joint t-distribution with five degrees of freedom. Further, optimal portfolios for different risk minimization objectives are constructed, with the result, that one of these optimal portfolios outperforms the S&P 500 in cumulative returns. Also, option prices on these optimal portfolios are calculated using Monte Carlo simulation.

Source	Cryptoassets	Data Period	Methodology	Purpose and Findings
[González et al., 2020]	Bitcoin Ethereum Ripple Bitcoin cash Tether Bitcoin SV Litecoin EOS Binance coin Tezos XRP	26.01.2015-07.03.2020	NARDL model	The connectedness of Bitcoin and ten other cryptoassets is studied via the NARDL model. Three different time horizons are used (daily, weekly and monthly). They find correlations for all assets and timeframes, but the newest assets in the monthly time horizons (probably due to lack of datapoints), but their connectedness differs. Further, asymmetric movement is found primarily in the short run.
[Demir et al., 2021]	Bitcoin Ethereum Ripple	07.2015-03.2019	NARDL model	This paper examines the asymmetric effects of BTC on Ethereum, Ripple and Litecoin. To do this they apply a NARDL model with several control variables, namely Wikipedia searches, oil price, gold price, 10-year US treasury yields, USD/EUR exchange rate and the NASDAQ Composite index. They find short-run asymmetries for all Altcoins considered and that a decrease in BTC price has greater impact than an increase. Further it is found that the period after the crash at the end of 2017 is the main contributor to the asymmetry.
[Omane-Adjepong and Alagidede, 2019]	Bitcoin BitShares Litecoin Stellar Ripple Monero Dash	08.05.2014-12.02.2018	Wavelet analysis GARCH model GJR-GARCH model	This paper studies the connectedness and volatility spillover between seven cryptoassets. Diversification among cryptoassets is limited. Co-movement and volatility spillovers are sensitive to both timescale and volatility measures.
[Shi et al., 2020]	Bitcoin Dash Ethereum Litecoin Ripple Stellar	08.08.2015-01.01.2020	Multivariate factor SV model	Analyses the correlation between six cryptoassets. It is found that Bitcoin is related to Litecoin and that Ethereum is related to Ripple, Dash and Stellar.
[Kumar and Suvvari, 2019]	Bitcoin Ethereum Ripple Litecoin	15.08.2015-18.01.2018	DCC-IGARCH model wavelet methods	This paper studies volatility spillovers between Bitcoin, Ethereum, Ripple and Litecoin. It is found that, there is significant short-run correlation over the whole time period. The results concerning volatility spillover are, that it is only evident after 2017. This is interpreted as a sign of the crypto market not being settled.

Source	Cryptoassets	Data Period	Methodology	Purpose and Findings
[Tu and Xue, 2019]	Bitcoin Litecoin	28.04.2013-31.07.2018	Granger causality test BEKK-MGARCH model	Studies the effect of the fork of Bitcoin into Bitcoin and Bitcoin Cash. This is done by comparing Bitcoin to Litecoin before and after the fork. While volatility spillovers in both periods go from Bitcoin to Litecoin, the direction of shock transmission changes after the fork (01.08.2017). This indicates a weakened market position of Bitcoin.
[Yousaf and Ali, 2020]	Bitcoin Ethereum Litecoin	03.10.2018-01.04.2020	VAR-AGARCH model	Examines the return and volatility spillover between Bitcoin, Ethereum and Litecoin in the COVID-19 period and before. It is found that diversification between the three cryptoassets is less effective during the COVID-19 period, but hedging effectiveness is increased.
[Canh et al., 2019]	Bitcoin Litecoin Ripple Stellar Monero Dash Bytecoin	05.08.2014-31.12.2018	DCC MGARCH model	Researches volatility spillovers between seven cryptoassets. It is found, that structural breaks occur first in the smaller capped cryptoassets and then spread to the larger ones. Further, significant positive correlations between the assets are found (over 0.6 in 6/7 cases), suggesting a lack of diversification in the crypto market.
[Katsiampa et al., 2019]	Bitcoin Litecoin Ethereum	07.08.2015-10.07.2018	BEKK-MGARCH model	Analyses Bitcoin, Litecoin and Ethereum pairwise for shock transmission and volatility spillover. There is no shock transmission from Litecoin to Ethereum, all other transmissions are found. Further volatility spillover is found between all three cryptoassets.
[Corbet et al., 2020]	Bitcoin Ethereum Ripple Bitcoin Cash Bitcion SV Litecoin Binance coin EOS Tezos Stellar Ethereum classic IOTA NEM	01.01.2019-31.03.2020	GARCH model	Analyses the relationship between 13 cryptoassets during the outbreak of COVID-19. To do so, the assets are modelled with three control variables namely, their traded volumes, the GBP/USD exchange rate and a sentiment time series, which was created by filtering twitter posts that mention both cryptocurrencies and COVID-19 and analysing their contents. It is found, that negative sentiment of COVID-19 does have an effect on the volatility of the cryptoasset returns. Growth in traded volumes and returns suggests, that cryptoassets became more attractive as value storage during the times of financial turmoil.

3 Univariate GARCH models

This Section aims to introduce the GARCH model and six possible generalizations.

3.1 Preliminary concepts

First, it is necessary to introduce a few basic concepts of time series analysis from [Francq and Zakoian, 2019].

Definition 3.1. A **time series** is a discrete stochastic process $(X_t)_{t\in\mathbb{Z}}$.

Remark. Often an observation of such a process is also called a time series, here such a single realization of a time series will be referred to as path.

Definition 3.2. A time series is called **weak white noise**, if it satisfies (i), (ii) and (iii), and **strong white noise**, if (iii) is replaced by (iii').

(i) $\mathbb{E}[X_t] = 0, \ \forall t \in \mathbb{Z}$

(ii) $\mathbb{E}[X_t^2] = \sigma^2 < \infty, \ \forall t \in \mathbb{Z}$

(iii) $\mathrm{Cov}[X_t, X_{t+h}] = 0, \ \forall t, h \in \mathbb{Z}, h \neq 0$

(iii') $(X_t)_{t\in\mathbb{Z}}$ is an independent and identically distributed sequence.

Definition 3.3. A time series $(X_t)_{t\in\mathbb{Z}}$ is called **strictly stationary**, if $(X_1, \ldots, X_k)'$ and $(X_{1+h}, \ldots, X_{k+h})'$ have the same joint distribution $\forall k \in \mathbb{N}$, $\forall h \in \mathbb{Z}$.

Definition 3.4. A time series $(X_t)_{t\in\mathbb{Z}}$ is **second order stationary**, if it satisfies the following

(i) $\mathbb{E}[X_t^2] < \infty, \ \forall t \in \mathbb{Z}$

(ii) $\mathbb{E}[X_t] = m, \ \forall t \in \mathbb{Z}$

(iii) $\mathrm{Cov}[X_t, X_{t+h}] = \gamma_X(h), \ \forall t, h \in \mathbb{Z}$

Remark. It is easy to see, that weak/strong white noise is a second order/strict stationary process, but not every second order/strict stationary process is a weak/strong white noise process.
Further, strict stationarity does not imply second order stationarity in all cases, because the notion of second order stationarity needs the first two moments to exist.

Remark. In finance, the time series that is modelled is the one of the logarithmic transformed data, i.e. $r_t = \log(p_t) - \log(p_{t-1})$ where p_t is the price of the asset at time t. With this transformation, financial time series usually become strict stationary. This can be checked with the augmented Dicky-Fuller test. [Kjærland et al., 2018], [Hou et al., 2020] and [Katsiampa, 2017] are examples for papers, that use this approach for cryptoassets.

Definition 3.5. The **autocorrelation function** of a second order stationary process $(X_t)_{t\in\mathbb{Z}}$ is defined as

$$\rho(h) := \frac{\gamma(h)}{\gamma(0)} \tag{1}$$

where

$$\gamma_X(h) := \mathrm{Cov}[X_t, X_{t+h}]$$

is called the **autocovariance function**.

Theorem 1. *Let $(X_k)_{k\in\mathbb{Z}}$ be a second order stationary time series, then $\forall k \in \mathbb{Z}$ and all $n, h \in \mathbb{N}$*

$$\hat{\gamma}(h) = \hat{\gamma}(h, n, k) := \frac{1}{n}\sum_{i=1}^{n-h}\left[\Big(X_i - \frac{1}{n}\sum_{j=1}^{n} X_{k+j}\Big)\Big(X_{i+h} - \frac{1}{n}\sum_{j=1}^{n} X_{k+j}\Big)\right] \tag{2}$$

is an asymptotically unbiased estimator for $\gamma(h)$, i.e.

$$\forall k \in \mathbb{Z}, \forall h \in \mathbb{N} : \lim_{n\to\infty} \hat{\gamma}(n, h, k) = \gamma(h) \ \ a.s.$$

Moreover, this provides an asymptotically unbiased estimator for $\rho(h)$ by

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

Proof.

$$\begin{aligned}\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n-h}\left[\Big(X_i - \frac{1}{n}\sum_{j=1}^{n} X_{k+j}\Big)\Big(X_{i+h} - \frac{1}{n}\sum_{j=1}^{n} X_{k+j}\Big)\right]\right] = \\ \frac{1}{n}\sum_{i=1}^{n-h}\mathbb{E}\left[\Big(X_i - \frac{1}{n}\sum_{j=1}^{n} X_{k+j}\Big)\Big(X_{i+h} - \frac{1}{n}\sum_{j=1}^{n} X_{k+j}\Big)\right]\end{aligned} \tag{3}$$

For second order stationary processes, it holds that $\mathbb{E}[X_t] = m \ \ \forall t \in \mathbb{Z}$, therefore for all $k \in \mathbb{Z}$

$$\frac{1}{n}\sum_{j=1}^{n} X_{k+j} \to m \ \ a.s. \text{ for } n \to \infty$$

by the strong law of large numbers. Since $(X_k)_{k\in\mathbb{Z}}$ is stationary, it follows that

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\left[\Big(X_i - \frac{1}{n}\sum_{j=1}^{n} X_{k+j}\Big)\Big(X_{i+h} - \frac{1}{n}\sum_{j=1}^{n} X_{k+j}\Big)\right] &= \mathbb{E}\left[\Big(X_i - \mathbb{E}[X_{i+h}]\Big)\Big(X_{i+h} - \mathbb{E}[X_{i+h}]\Big)\right] \\ &= \mathrm{Cov}[X_i, X_{i+h}] = \gamma(h)\end{aligned}$$

Using Dominated Convergence with the dominating function $4\mathbb{E}[X_i^2]$.
Taking the limit in equation (3) and using this yields the result. □

3.2 Purpose of the GARCH model

The goal of modelling a time series is to get the residuals of the chosen model to be white noise, i. e. to predict everything that follows a pattern. The well known autoregressive moving average model (ARMA model) predicts patterns in the values of the X_t, so what is the GARCH model needed for? The following example will illustrate the reason to implement GARCH:

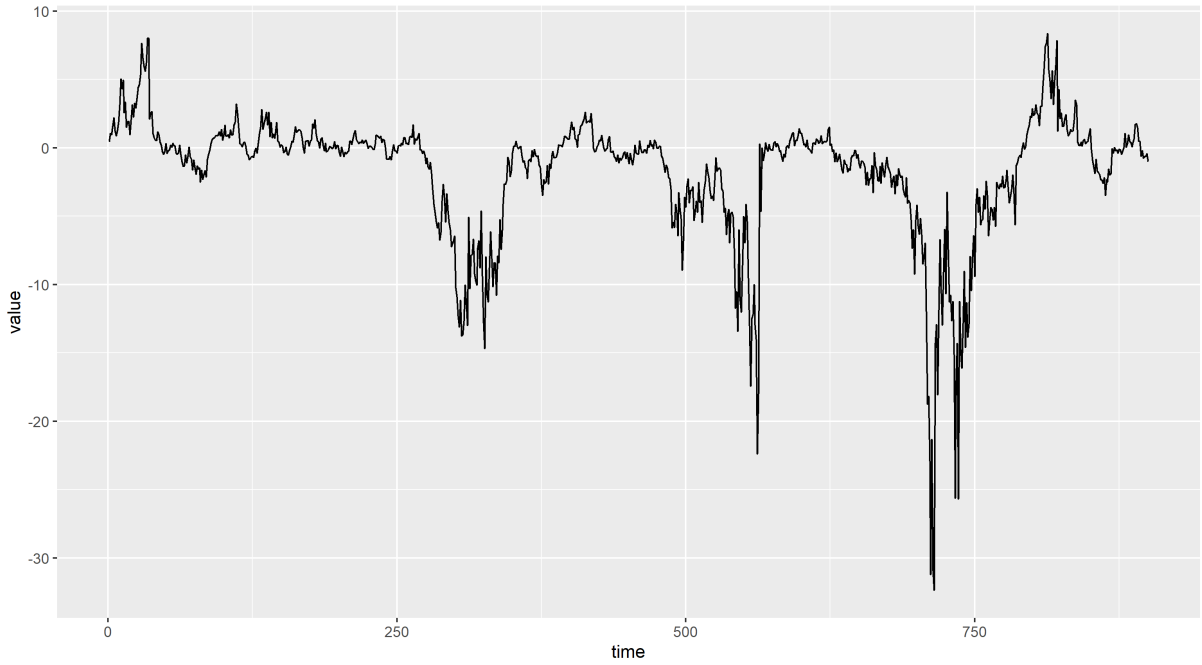


Figure 1: Sample time series path created with an ARMA(3,1) model $\alpha = [0.4, 0.3, 0.3]$, $\beta = 0.1$, with time varying volatility (created by GARCH(1,1) model).

In Figure 1 a path of a time series is displayed. Treating this as the input data e.g. the returns of a given asset, one wants to find a model to best describe its patterns. Just by looking at the graph, it is easy to observe, that high/low values are more likely to come after high/low ones. This contradicts (*iii*) of the properties of white noise. To change that one can fit an ARMA model to the observed data. In this case, since the data was created by such a model, the whole covariance structure is explained by the model.

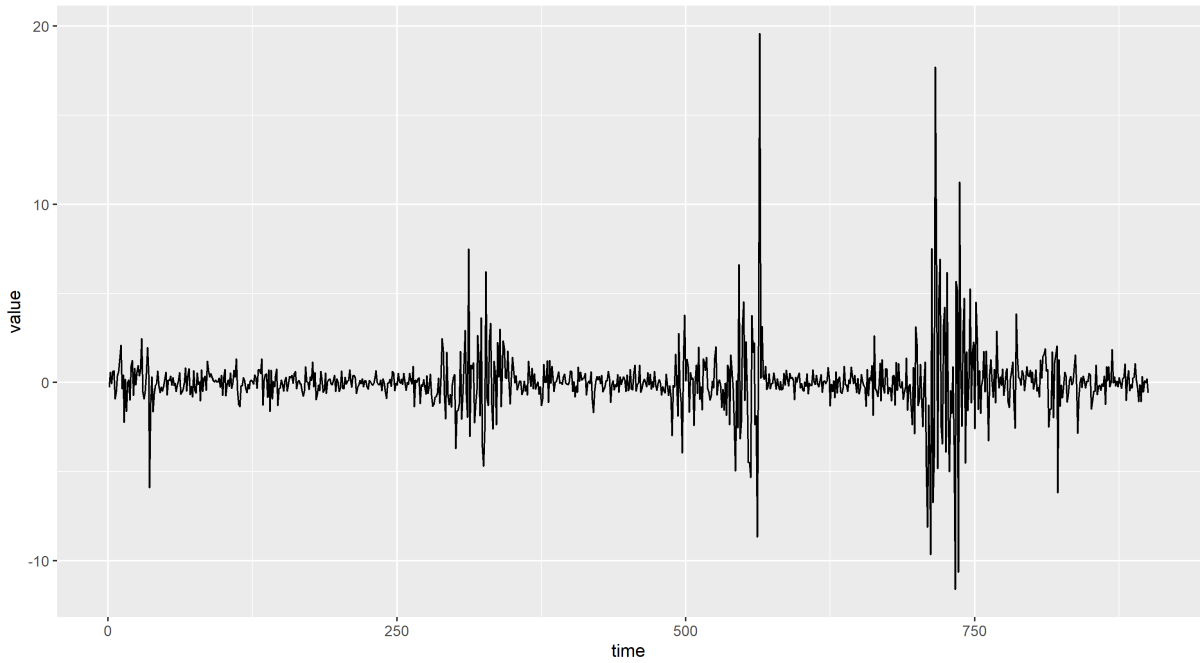


Figure 2: Errors of the ARMA(3,1) model

Figure 2 shows the error terms from the fitted ARMA model. Even though it does not look like it, the way this path was created did satisfy condition (*ii*) for white noise, which will be revisited later. So now the error terms are weak white noise, but it is easy to recognize that there still is some pattern left. The remaining structure, that can be observed, is that high and low variances alike form clusters.

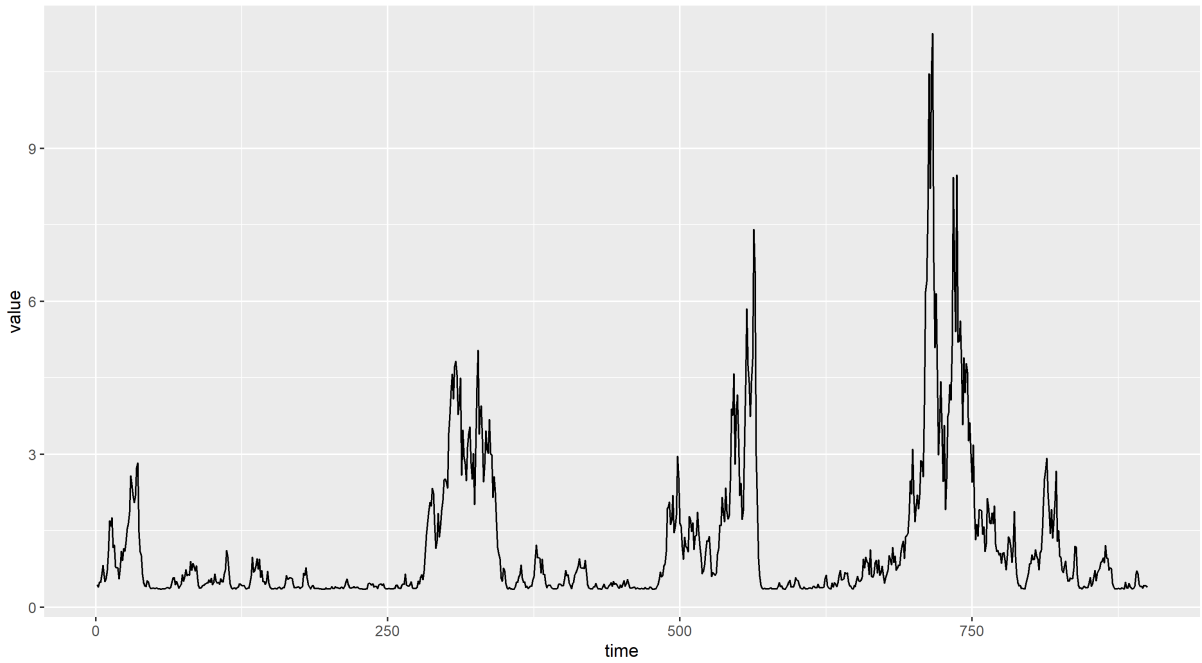


Figure 3: Volatility generated by GARCH(1,1) with $\omega = 0.1$, $\alpha = 0.1$ and $\beta = 0.2$

In Figure 3 the actual volatility of each error for the path of the chosen GARCH model is shown. Now the path of the volatility above has to be described by a model. If this were a real data set, one would need to estimate the volatility graph above. This is where the GARCH model comes into play, in a similar approach to the ARMA model it tries to predict the variance of future returns, thereby giving the errors independence.

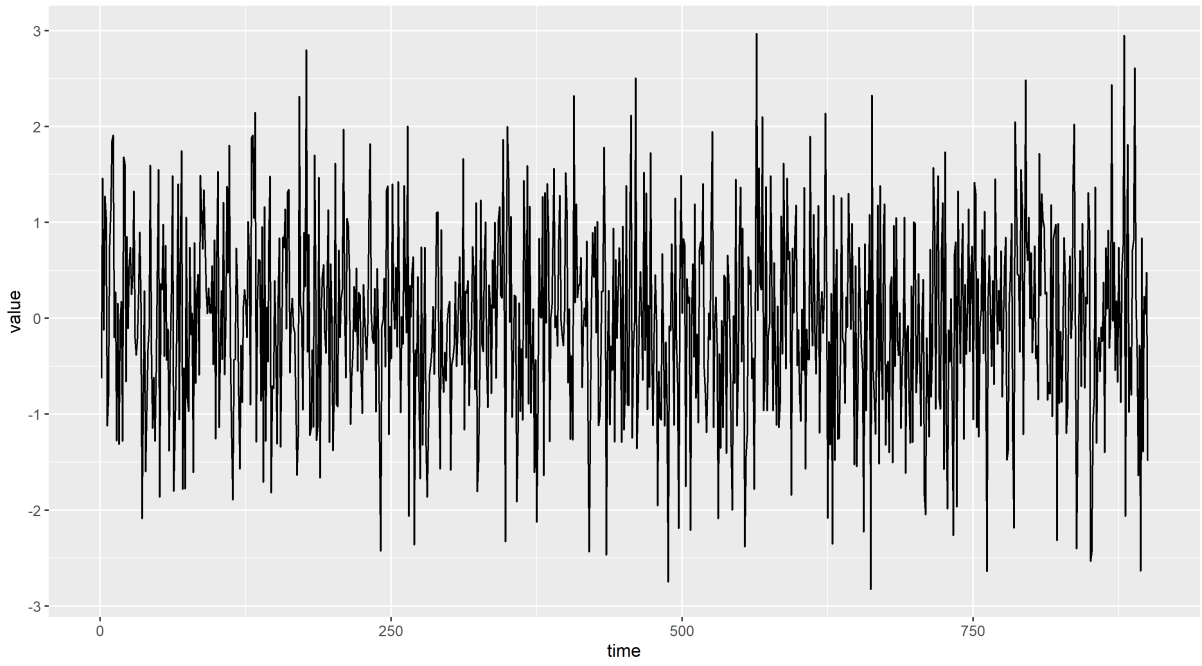


Figure 4: Errors after applying the GARCH model

In Figure 4 the final errors are shown. Now no patterns of first or second order can be observed and it holds that the variance is independent, i.e. $\mathbb{E}[X_t^2 \mid X_i \ \forall i \in \mathbb{Z} \setminus \{t\}] = \sigma^2$.

3.3 GARCH model

3.3.1 Definition of the model

Definition 3.6. Let $(X_t)_{t\in\mathbb{Z}}$ be a time series and $\mathcal{F}_t = \sigma(\{F_s \mid s \leq t\})$ the information set up to time t $\forall t \in \mathbb{Z}$. Further let $p \geq 0$, $q > 0$, $\omega > 0$, $\alpha_i \geq 0 \ \forall i = 1 \dots q$ and $\beta_i \geq 0 \ \forall i = 1 \dots p$. Then $(X_t)_{t\in\mathbb{Z}}$ is called a **(normal) GARCH(p,q) process**, if

$$X_t \mid \mathcal{F}_{t-1} \sim \mathcal{N}(0, \sigma_t) \tag{4}$$

with

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{5}$$

[Bollerslev, 1986]

GARCH stands for Gerneralised Autoregressive Conditional Heteroscedasticity, so it models the conditional variance of the process in an autoregressive way. If in the model above $qp = 0$, then it is the ARCH model by [Engle, 1982] (p=0 gives ARCH model and q=0 gives the ARCH regression). Further, it is easy to see that the variance is only depending on past observations,

meaning that the one step ahead variance forecast is deterministic, in contrast to Stochastic Volatility (SV) models.

Remark. The distribution can easily be generalised by replacing equation 4 by

$$X_t = \sigma_t \eta_t \tag{6}$$

and using any other distribution for an i.i.d. random series $(\eta_i)_{i\in\mathbb{Z}}$. Most models use normal distributions, so if no distribution is mentioned, the normal case is the default one. If all η_i have distribution $\mathcal{L}$, we call X an $\mathcal{L}$-GARCH(p,q) process. However in the literature, some papers have already shown that heavy tailed distributions outperform the normal distribution when studying the case of cryptoasset returns. Examples are [Guo, 2022], [Troster et al., 2019] or [Chu et al., 2017].

Equation (4) is called the mean equation and equation (5) is called the variance equation for obvious reasons.

3.3.2 Stationarity of the GARCH model

It is easy to see, that not for all admissible choices of α_i and β_i the corresponding GARCH process becomes stationary. For example, if one takes $\alpha_i = 0 \ \forall i \in \mathbb{N}$, $\beta_1 > 1$ and $\alpha_i = 0 \ \forall i \geq 1$, it holds that $\sigma_t^2 = h\omega + \sigma_{t-h}^2 \beta_1^h$ which would lead to an explosion of the variance. Stationarity analysis is quite cumbersome, therefore mostly results are presented here:

Theorem 2. *Let $(\eta_i)_{i\in\mathbb{Z}}$ be an i.i.d. random series with law $\mathcal{L}$. The $\mathcal{L}$-GARCH process is second order stationary, if and only if*

$$\sum_{i=1}^{q} \alpha_i \mathbb{V}[\eta_1] + \sum_{i=1}^{p} \beta_i < 1 \tag{7}$$

In this case it holds that

$$\mathbb{E}[\sigma_t^2] = \omega(1 - \sum_{i=1}^{q} \alpha_i \mathbb{V}[\eta_1] + \sum_{i=1}^{p} \beta_i)^{-1} \tag{8}$$

Proof. See [Bollerslev, 1986] for the proof of the normal case, which is simple to generalize. □

Theorem 3. *A GARCH(1,1) process is strictly stationary if*

$$\gamma := \mathbb{E}[\log(\alpha\eta_1^2 + \beta)] < 0 \tag{9}$$

holds.

Proof. Define

$$a(\eta_t) := \alpha\eta_t^2 + \beta$$

Equation (5) can be used recursively to get

$$\begin{aligned}\sigma_t^2 &= \omega + a(\eta_{t-1})\sigma_{t-1}^2 = \omega + a(\eta_{t-1})\omega + a(\eta_{t-1})a(\eta_{t-2})\sigma_{t-2}^2 \\ &= \omega\left(1 + \sum_{n=1}^{N} a(\eta_{t-1})\dots a(\eta_{t-n})\right) + a(\eta_{t-1})\dots a(\eta_{t-N-1})\sigma_{t-N-1}^2 \\ &=: h_t(N) + a(\eta_{t-1})\dots a(\eta_{t-N-1})\sigma_{t-N-1}^2\end{aligned} \tag{10}$$

The process $h_t := \lim_{N\to\infty} h_t(N)$ exists in $\mathbb{R}^+ \cup \{\infty\}$, since all summands are non-negative. Further h fulfills the same recursion as σ^2, in that

$$h_t = \omega + a(\eta_{t-1})h_{t-1}$$

Next it will be shown that h_t is a.s. finite if $\gamma < 0$.
By the strong law of large numbers on the i.i.d sequence $(\log(a(\eta_t)))_{t\in\mathbb{Z}}$ it holds that

$$\Big(a(\eta_{t-1})\dots a(\eta_{t-N})\Big)^{\frac{1}{N}} = \exp\left(\frac{1}{N}\sum_{n=1}^{N}\log(a(\eta_{t-n}))\right) \to \exp(\gamma) \text{ a.s. for } N\to\infty$$

The Cauchy rule for series with non-negative terms given in Lemma 1 can now be used to get the result for h_t. In the case $\gamma < 0$ it follows from equation (10) that $\sigma_t^2 = h_t$ a.s.. Therefore, X_t admits the following representation

$$X_t = \sqrt{h_t}\eta_t = \left(\omega + \sum_{n=1}^{\infty} a(\eta_{t-1})\dots a(\eta_{t-n})\omega\right)^{1/2}\eta_t$$

Lemma 2 bellow gives strict stationarity. □

Lemma 1. *Let $\sum_{n=1}^{\infty} a_n$ be a series with non-negaitve summands and $\lambda := \lim_{N\to\infty}\prod_{n=1}^{N} a_n^{1/N}$, then if $\lambda < 1$ the series converges and if $\lambda > 1$ it diverges.*

Lemma 2. *Let $(Z_t)_{t\in\mathbb{Z}}$ be a strict stationary sequence and $(Y_t)_{t\in\mathbb{Z}}$ be defined by*

$$Y_t := f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots) \tag{11}$$

then if f is a measurable function from $\mathbb{R}^\infty$ to $\mathbb{R}$, $(Y_t)_{t\in\mathbb{Z}}$ is also strict stationary.
[Billingsley, 1995] Theorem 36.4

Remark. By Jensen's inequality one can see that this is a relaxation to what was needed above. For a normal GARCH process this means α_1 can be as large as 3.65, but note that the second moment does not exist anymore. To state the general case, some addtitional work has to be done.

Definition 3.7. Let $(A_i)_{i\in\mathbb{N}}$ be a sequence of random $n\times n$ matrices, then the **top Lyapunov exponent** is defined as

$$\gamma(A) := \lim_{t\to\infty} \frac{1}{t} \log || A_t A_{t-1} \dots A_1 || \quad a.s. \tag{12}$$

Remark. The top Lyapunov exponent does not depend on the choice of the matrix norm.

For the GARCH model with parameters $\theta = (\omega, \alpha_1, \dots, \alpha q, \beta_1, \dots, \beta_p)$ define

$$A_t := \begin{bmatrix} \alpha_1\eta_t^2 & \cdots & \dots & \cdots & \alpha_q\eta_t^2 & \beta_1\eta_t^2 & \cdots & \dots & \cdots & \beta_p\eta_t^2 \\ 1 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots & \vdots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ \alpha_1 & \cdots & \dots & \cdots & \alpha_q & \beta_1 & \cdots & \dots & \cdots & \beta_p \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & \vdots & 0 & 1 & \ddots & & \vdots \\ \vdots & & & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{(p+q)\times(p+q)}$$

$$b_t := \begin{pmatrix} \omega\eta_t^2 \\ 0 \\ \vdots \\ 0 \\ \omega \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{p+q}, \qquad z_t := \begin{pmatrix} X_t^2 \\ \vdots \\ X_{t-q+1}^2 \\ \sigma_t^2 \\ \vdots \\ \sigma_{t-p+1}^2 \end{pmatrix} \in \mathbb{R}^{p+q}$$

With this the GARCH model admits the representation

$$z_t = b_t + A_t z_{t-1}$$

where the equations of the GARCH model are column 1 and q+1. Now let $A = (A_i)_{i\in\mathbb{N}}$, then $\gamma(A)$ is the top Lyapunov exponent of this GARCH model.

Theorem 4. *A GARCH(p,q) model is strictly stationary, if and only if its top Lyapunov exponent satisfies* $\gamma(A) < 0$.

Both results for strict stationarity stem from [Francq and Zakoian, 2019], where proofs can also be found.

Theorem 5. *A stationary GARCH process is weak white noise.*

Proof. (i) $\mathbb{E}[X_t] = \mathbb{E}[\sigma_t \eta_t] = \mathbb{E}[\mathbb{E}[\sigma_t \eta_t \mid \eta_{t-1}]] = \omega(1 - \sum_{i=1}^{q} \alpha_i \mathbb{V}[\eta] + \sum_{i=1}^{p} \beta_i)^{-1} \mathbb{E}[\eta_t] = 0$

(ii) Analogous to (i).

(iii) $\mathrm{Cov}[X_t, X_{t+h}] = \mathbb{E}[X_t X_{t+h}] = \mathbb{E}[\sigma_t \sigma_{t+h} \eta_t \eta_{t+h}] = \mathbb{V}[\sigma_t]^2 \mathbb{E}[\eta_t \eta_{t+h}] = 0$

□

3.3.3 Parameter interpretation

In the case of a weakly stationary GARCH process, ω is the parameter for the long-term variance. This can easily be seen in equation (8). From that equation one can also observe that the variance of the process is unaffected, as long as $\sum_{i=1}^{q} \alpha_i \mathbb{V}[\eta] + \sum_{i=1}^{p} \beta_i$ remains constant. In the most common case, it holds that $\mathbb{V}[\eta] = 1$, therefore, in the case of $p = q$, it is possible to swap the values of the α_i and β_i without changing the variance.

If the sum of the α_i is close to one, the reaction to volatility shocks in the process is a rapidly varying volatility. These rapid movements increase if p is smaller, since taking multiple past observations into account smoothens the process. Below, a process that admits these rapid changes is shown.

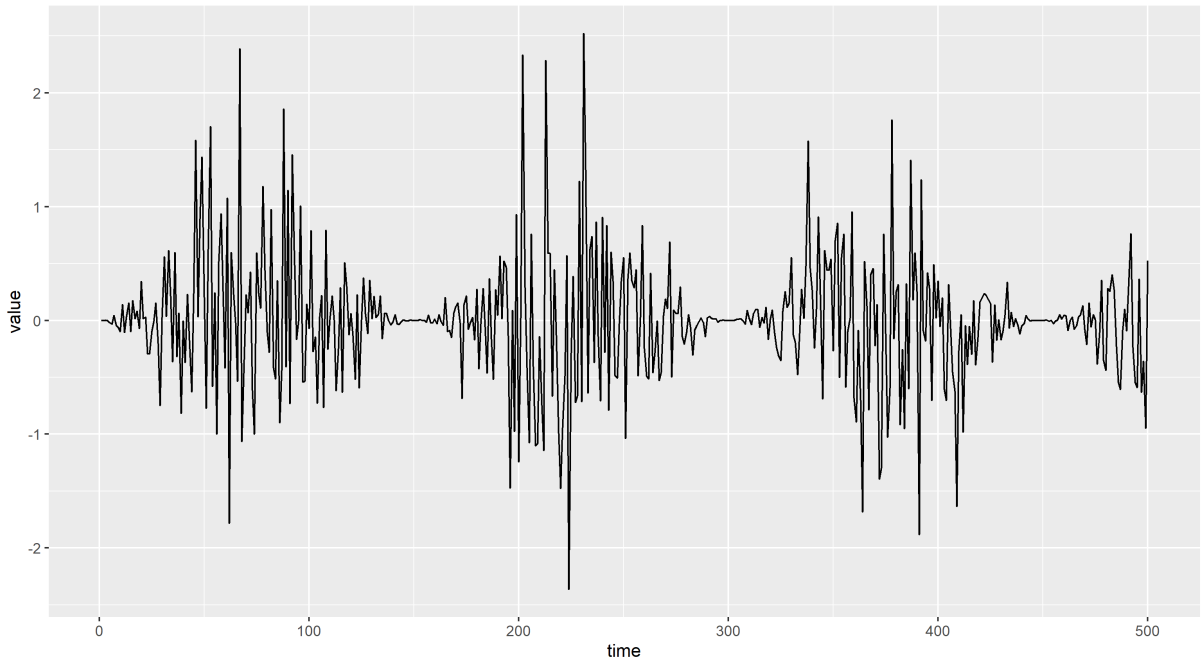


Figure 5: GARCH(1,1) with $\omega = 1$, $\alpha = 0.7$ and $\beta = 0.1$

On the other hand, when the sum of the β_i is high, the volatility shocks show high persistence. This displays in a constant high volatility, as the graph below demonstrates.

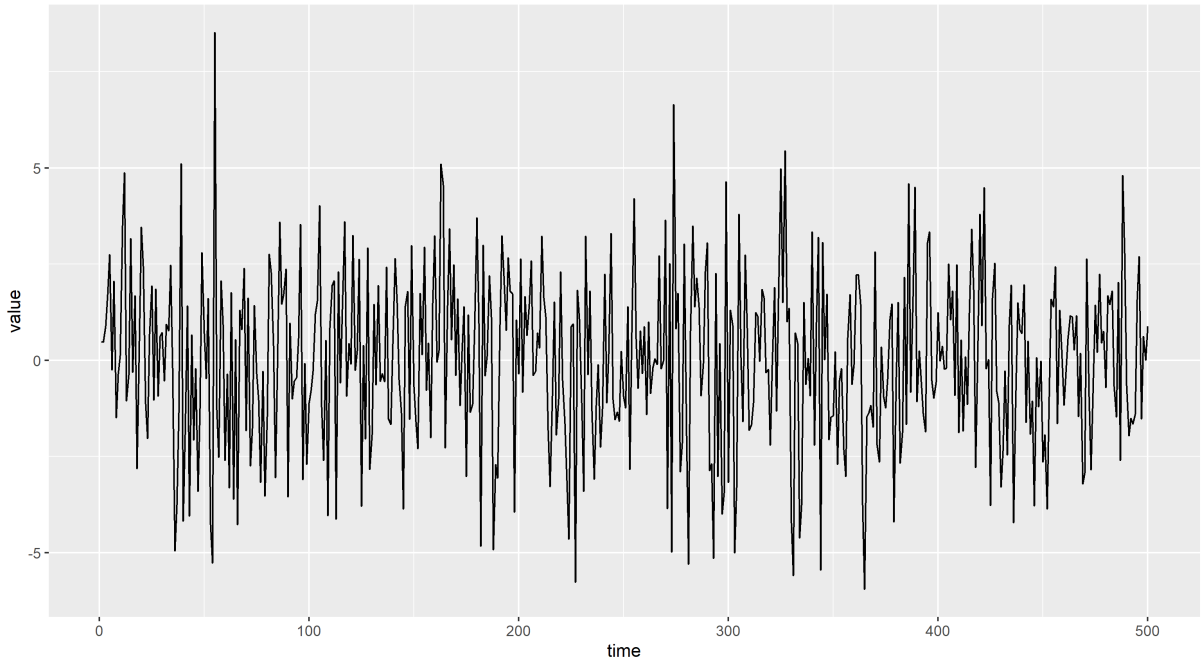


Figure 6: GARCH(1,1) with $\omega = 1$, $\alpha = 0.1$ and $\beta = 0.7$

3.3.4 Identifying GARCH orders

In this section the aim is to find plausible orders for the GARCH(p,q) model, which can be compared via an information criterion after estimation. The innovation of the process X_t^2 is defined by $\nu_t := X_t^2 - \sigma_t^2$. Substituting $X_t^2 - \nu_t$ for σ_t^2 in equation (5), the following representation of the process X^2 is obtained:

$$X_t^2 = \omega + \nu_t + \sum_{i=1}^{r}(\alpha_i + \beta_i)X_{t-i}^2 + \sum_{i=1}^{p}\beta_i\nu_{t-i} \tag{13}$$

with $\alpha_i := 0$, $\beta_j := 0$ for $i \geq q$, $j \geq p$, respectively, and $r = \max(p, q)$. Notice that this is a representation of an ARMA(r,p) model. The next step is to identify the orders of this ARMA model. For this the so-called corner method can be used.
Let ρ_{X^2} be the autocorrelation function of the process $(X_t^2)_{t\in\mathbb{Z}}$, then define the $j \times j$ Toeplitz matrix

$$D(i,j) := \begin{bmatrix} \rho_{X^2}(i) & \rho_{X^2}(i-1) & \cdots & \rho_{X^2}(i-j+1) \\ \rho_{X^2}(i+1) & \rho_{X^2}(i) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{X^2}(i-1) \\ \rho_{X^2}(i+j-1) & \cdots & \rho_{X^2}(i+1) & \rho_{X^2}(i) \end{bmatrix} \tag{14}$$

Theorem 6. *Let $\Delta(i,j) = \det(D(i,j))$, then P and Q are the minimal ARMA orders, if and only if*

(i) $\Delta(i,j) = 0 \ \forall i > Q$ *and* $j > P$

(ii) $\Delta(i,P) \neq 0 \ \forall i \geq Q$

(iii) $\Delta(Q,j) \neq 0 \ \forall j \geq P$

[De Gooijer and Heuts, 1981]

Proof. We give a short illustration of the proof idea. The ARMA(P,Q) model reads

$$X_t = \omega + \nu_t + \sum_{i=1}^{P}\alpha_i X_{t-i} + \sum_{i=1}^{Q}\beta_i\nu_{t-i} \tag{15}$$

For this model one can show that for $k > Q$ it holds that

$$\rho(k) = \sum_{i=1}^{P}\alpha_i\rho(k-i) \tag{16}$$

So, if $i > Q$ and $j > P$, the first column can easily be factorized into the following p columns. If $i = Q$, the above factorization ceases to be valid, and for any size of the matrix we have

a diagonal that prevents this factorization since $i \not> Q$. If $j = P$, the above does not yield a singular matrix, since the matrix has j columns and the factorization needs $j + 1$. The other direction follows by assuming there are smaller orders and using the factorization to contradict either (ii) or (iii). □

i \ j	1	2	$\cdots$	q	$q+1$	$\cdots$	K
1	$\Delta(1,1)$	$\Delta(1,2)$	$\cdots$	$\Delta(1,q)$	$\Delta(1,q+1)$	$\cdots$	$\Delta(1,K)$
2	$\Delta(2,1)$	$\Delta(2,2)$	$\cdots$	$\Delta(2,q)$	$\Delta(2,q+1)$	$\cdots$	$\Delta(2,K)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
p	$\Delta(p,1)$	$\Delta(p,2)$	$\cdots$	$\Delta(p,q)$	$\Delta(p,q+1)$	$\cdots$	$\Delta(p,K)$
$p+1$	$\Delta(p+1,1)$	$\Delta(p+1,2)$	$\cdots$	$\Delta(p+1,q)$	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	0	$\vdots$
K	$\Delta(K,1)$	$\Delta(K,2)$	$\cdots$	$\Delta(K,q)$	0	$\cdots$	0

Figure 7: Table for corner method

Figure 7 shows the matrix of the $\Delta(i,j)$ for an ARMA(p,q) model. At the bottom right is the name-giving corner of zeros in the matrix.

In practice, the exact values of ρ are not known. After replacing the exact values by the estimates $\hat{\rho}$, a test becomes necessary to evaluate the possible orders. Two possible ways to obtain these candidate orders will be illustrated below.

[De Gooijer and Heuts, 1981] used that if $\hat{\Delta}$ is a $1 \times n$ vector of determinants and H_0 is that they all have $i \geq q+1$ and $j \geq p+1$ it holds that

$$T(\hat{\Delta})(\hat{A}\hat{G}_H\hat{A}^t)^{-1}\hat{\Delta}^t \xrightarrow{d} \mathcal{X}^2_n, \qquad \text{under } H_0 \tag{17}$$

where $\hat{A}$ is an $n \times H$ matrix with

$$\hat{A}(r,s) = \frac{\partial\hat{\Delta}(r)}{\partial\hat{\rho}(s)}$$

T is the number of observed data points and $\hat{G}_H$ is the $H \times H$ covariance matrix with

$$\begin{aligned}\hat{G}_H(r,s) = \sum_{h=-\infty}^{\infty} \Big(&\hat{\rho}(h)\hat{\rho}(h-r+s) + \hat{\rho}(h+s)\hat{\rho}(h-r) \\ &- 2\hat{\rho}(h)\hat{\rho}(s)\hat{\rho}(h-r) - 2\hat{\rho}(r)\hat{\rho}(h)\hat{\rho}(h-s) + 2\hat{\rho}(r)\hat{\rho}(s)\hat{\rho}(h)^2 \Big)\end{aligned} \tag{18}$$

where H is the maximum index h such that $\hat{\rho}(h)$ is used in one of the determinants $\hat{\Delta}_1, \dots \hat{\Delta}_n$. The sum in (18) can in practice not be infinite. [Chan, 1999] suggests, that 50 should be a sufficient number for estimation. Another option to estimate this matrix is to use an estimator along the lines of equation (29). The inversion of the matrix in (17) can pose a problem, since it is not given that the estimated matrix is non-singular. [De Gooijer and Heuts, 1981] used a generalized inverse. They find disappointing performances for small samples, but large ones work well. Further, visual inspection was found to not give a clear answer.

Another approach is used by [Chan, 1999]. With similar notations as above, but with a shifted matrix $\tilde{\Delta} = (\hat{\Delta}(i+1, j+1))_{i,j}$ and $\tilde{A}_{i,j}(s) = \tilde{\Delta}(i+1, j+1)/\rho(s)$ for $s = 1 \dots h$, a new matrix is defined by

$$D^* = \begin{cases} O, & \text{if } \left| \dfrac{\tilde{\Delta}(i+1, j+1)}{SE(\tilde{\Delta}(i+1, j+1))} \right| < 2 \\ X, & \text{otherwise} \end{cases}$$

with

$$SE(\hat{\Delta}(i+1, j+1)) = \sqrt{\frac{\tilde{A}_{i,j} \hat{G}_H \tilde{A}_{i,j}^T}{T}}$$

In the new D^* matrix, the value O stands for a value that is not significantly different from zero, while X stands for a value that is significantly different form zero. In this case the problem of matrix inversion is gone and it also gives a good visual view which leads to some promising candidates for p and q, but it is not a formal significance test. Nevertheless, [Chan, 1999] finds good performance of the corner method in large samples, where it outperforms the other algorithms that are considered.

Now continuing with the model for X^2, it is possible, given enough sample data, to find promising candidates $(\hat{r}, \hat{p})$ for r and p. Since it holds, that $r = \max(p, q) \geq p$, [Francq and Zakoian, 2019] suggest using

$$(p, q) = \begin{cases} (\hat{p}, \hat{r}) & \text{if } \hat{r} \geq \hat{p} \\ (\hat{p}, i) & \text{for } i = 1 \dots \hat{r}, \text{ if } \hat{r} < \hat{p} \end{cases}$$

as candidates.

3.3.5 Testing for ARCH

Before estimating a GARCH model, first one needs to test, if the data displays conditional heteroskedasticity. This is generally done using a LM (Lagrange multiplier) test, which is used in [Dyhrberg, 2016a] or [Nguyen et al., 2019].

Definition 3.8. In a model with unknown parameter vector $\theta \in \mathbb{R}^d$ and observations $X = (X_1, \dots X_n)$, the hypothesis

$$H_0 : \quad R\theta_0 = r$$

with a matrix of full rank $R \in \mathbb{R}^{s \times d}$ and $r \in \mathbb{R}^s$ is tested with the **LM test statistic** or **score test statistic**

$$LM_n := \frac{1}{n} \frac{\partial l_n(\hat{\theta})}{\partial \theta^T} \hat{\mathcal{I}}^{-1}(\hat{\theta}) \frac{\partial l_n(\hat{\theta})}{\partial \theta} \tag{19}$$

where l_n is the log-likelihood function given the observations X, $\hat{\theta} = \arg\max_{R\theta=r} l_n(\theta)$ is the maximum of the constrained log-likelihood function and with $\hat{\mathcal{I}}$ being an estimator of the Fisher Information. For example

$$\hat{\mathcal{I}} = -\frac{1}{n} \frac{\partial^2 l_n(\hat{\theta})}{\partial \theta \partial \theta^T}$$

Further the term

$$\frac{1}{\sqrt{n}} \frac{\partial l_n(\theta)}{\partial \theta}$$

is called the score.

Theorem 7. *If the score exists almost everywhere and satisfies a CLT, it holds that under* H_0

$$LM_n \to \chi^2_s \tag{20}$$

Proof. A more precise formulation of the theorem and its proof can be found in [Aitchison and Silvey, 1958]. □

The following derivation of the LM test for the ARCH effect is following [Francq and Zakoian, 2019].
The ARCH model, that is tested for, can be stated as

$$X_t = \sigma_t \eta_t, \;\; \eta_t \sim i.i.d(0, 1) \tag{21}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 \tag{22}$$

with $\omega > 0$ and $\alpha_i \geq 0$.
For calculation purposes, the variance of the innovations is fixed, but it can easily be incorporated in the σ part, so nothing changes for the generality of the model. The hypothesis that is tested is

$$H_0 : \alpha_1 = \dots = \alpha_q = 0 \tag{23}$$

In the absence of a log-likelihood function, a gaussian quasi-log-likelihood function is used, which is given in the following form

$$ql_n(\theta) = -\frac{1}{2}\sum_{t=1}^{n}\left\{\frac{X_t^2}{\sigma_t^2(\theta)} + \log\sigma_t^2(\theta)\right\}, \quad \sigma_t^2(\theta) = \omega + \sum_{i=1}^{q}\alpha_i X_{t-i}^2 \tag{24}$$

For this, the convention of $X_t = 0$ for $t < 0$ is used. Conditions to the usability of the quasi-likelihood are briefly discussed in Section 3.3.7.
In the proof of the LM test, it is used that

$$\frac{1}{\sqrt{n}}\frac{\partial l_n(\theta_0)}{\partial\theta} \to \mathcal{N}(0, \mathcal{F})$$

This no longer holds true for the quasi-log-likelihood. Therefore, one has to calculate the limit. In this case $\theta_0 = (\omega_0, 0, \dots 0)$.

$$\frac{1}{\sqrt{n}}\frac{\partial ql_n(\theta_0)}{\partial\theta} = \frac{1}{2\sqrt{n}}\sum_{t=1}^{n}\frac{X_t^2 - \sigma_t^2(\theta_0)}{\sigma_t^4(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta} \tag{25}$$

Plugging in $X_t = \sigma_t\eta_t$ and using the form of $\sigma_t^2(\theta_0)$ represented in equation (24) one gets

$$\frac{1}{\sqrt{n}}\frac{\partial ql_n(\theta_0)}{\partial\theta} = \frac{1}{2\sqrt{n}}\sum_{t=1}^{n}\frac{\eta_t^2 - 1}{\omega_0}\begin{pmatrix}1\\ X_{t-1}^2\\ \vdots\\ X_{t-q}^2\end{pmatrix} \tag{26}$$

It is easy to see, that the expectation of this term is zero. Since η_t is independent of X_s for all $s < t$ it follows that we just need to check $\mathbb{E}[\eta_t^2 - 1]$, which is assumed to be zero by the variance and expectation conditions on η_t. To obtain the covariance matrix of the vector given in (26), more calculations are necessary. For this define $\kappa_\eta = \mathbb{E}[\eta_t^4]$. For the first entry, one gets

$$\mathbb{V}\left[\frac{1}{2\sqrt{n}}\sum_{t=1}^{n}\frac{\eta_t^2 - 1}{\omega_0}\right] = \frac{1}{4n\omega_0^2}\mathbb{E}\left[\sum_{t=1}^{n}(\eta_t^2 - 1)^2\right] = \frac{1}{4n\omega_0^2}\mathbb{E}\left[\sum_{t=1}^{n}\eta_t^4 - 1\right] = \frac{\kappa_\eta - 1}{4\omega_0^2}$$

The remaining diagonal elements of the covariance matrix are

$$\mathbb{V}\left[\frac{1}{2\sqrt{n}}\sum_{t=1}^{n}\frac{\eta_t^2 - 1}{\omega_0^2}X_{t-1}^2\right] = \frac{1}{4n\omega_0^2}\mathbb{E}\left[\sum_{t=1}^{n}(\eta_t^2-1)^2\sigma_t^4(\theta_0)\eta_{t-1}^4\right] = \frac{1}{4n}\mathbb{E}\left[\sum_{t=1}^{n}(\eta_t^2-1)^2\eta_{t-1}^4\right] = \frac{(\kappa_\eta - 1)\kappa_\eta}{4}$$

Let $i > j \geq 2$, then for those non-diagonal elements, which are not in the first column or row, one gets

$$\mathbb{E}\Big[\frac{1}{4n}\sum_{t=1}^{n}\frac{\eta_t^2-1}{\omega_0}X_{t-i}^2\sum_{t=1}^{n}\frac{\eta_t^2-1}{\omega_0}X_{t-j}^2\Big] = \frac{1}{4n\omega_0^2}\mathbb{E}\Big[\sum_{t=1}^{n}(\eta_t^2-1)^2X_{t-i}^2X_{t-j}^2\Big] = \frac{1}{4n}\mathbb{E}\Big[\sum_{t=1}^{n}(\eta_t^2-1)^2\Big] = \frac{\kappa_\eta-1}{4}$$

Lastly, for the non-diagonal elements of the first row or column, we almost analogously get

$$\mathbb{E}\Big[\frac{1}{4n}\sum_{t=1}^{n}\frac{\eta_t^2-1}{\omega_0}\sum_{t=1}^{n}\frac{\eta_t^2-1}{\omega_0}X_{t-j}^2\Big] = \frac{1}{4n\omega_0^2}\mathbb{E}\Big[\sum_{t=1}^{n}(\eta_t^2-1)^2X_{t-j}^2\Big] = \frac{1}{4n\omega_0}\mathbb{E}\Big[\sum_{t=1}^{n}(\eta_t^2-1)^2\Big] = \frac{\kappa_\eta-1}{4\omega_0}$$

So, the new CLT is

$$\frac{1}{\sqrt{n}}\frac{\partial ql_n(\theta_0)}{\partial\theta} \to \mathcal{N}(0,I) \tag{27}$$

where I is

$$I = \frac{\kappa_\eta-1}{4\omega_0^2}\begin{bmatrix} 1 & \omega_0 & \dots & \dots & \omega_0 \\ \omega_0 & \omega_0^2\kappa_\eta & \omega_0^2 & \dots & \omega_0^2 \\ \vdots & \omega_0^2 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \omega_0^2 \\ \omega_0 & \omega_0^2 & \dots & \omega_0^2 & \omega_0^2\kappa_\eta \end{bmatrix} =: \frac{\kappa_\eta-1}{4\omega_0^2}\begin{bmatrix} 1 & \boldsymbol{\omega_0} \\ \boldsymbol{\omega_0^t} & I_{pp} \end{bmatrix}$$

Now the last piece, that is missing is the score term for $\hat{\theta} = \arg\max_{\theta=(\omega,0\dots0)} l_n(\theta)$. For $\boldsymbol{\alpha} = (\alpha_1, \dots \alpha_q)$, it holds that

$$\frac{\partial ql_n(\hat{\theta})}{\partial\theta} = \begin{pmatrix} 0 \\ \frac{\partial ql_n(\hat{\theta})}{\partial\boldsymbol{\alpha}} \end{pmatrix}$$

since $\hat{\theta}$ is the maximiser. Further

$$\frac{\partial ql_n(\hat{\theta})}{\partial\boldsymbol{\alpha}} = -\frac{1}{2}\frac{\partial}{\partial\boldsymbol{\alpha}}\sum_{t=1}^{n}\left(\frac{X_t^2}{\sigma_t^2(\hat{\theta})} + \log\sigma_t^2(\hat{\theta})\right) = \frac{1}{2}\sum_{t=1}^{n}\left(\frac{X_t^2}{\hat{\omega}^2} - \frac{1}{\hat{\omega}}\right)\begin{pmatrix} X_{t-1}^2 \\ \vdots \\ X_{t-q}^2 \end{pmatrix}$$

So, to get LM_n, only the lower right-hand $q \times q$ matrix of I^{-1} needs to be calculated, this block will be referred to as I_{qq}^{-1}. This can simply be calculated by

$$I_{qq}^{-1} = \left(\frac{\kappa_\eta-1}{4\omega_0^2}(I_{qq} - \boldsymbol{\omega_0^t\omega_0})\right)^{-1} = \frac{4\omega_0^2}{\kappa_\eta-1}\Big(\omega_0^2(\kappa_\eta-1)I_q\Big)^{-1} = \frac{4}{(\kappa_\eta-1)^2}\mathrm{I}_q$$

where I_q is the identity matrix of size $q \times q$.

The closed-form expression of LM_n is

$$LM_n = \frac{1}{n}\frac{\partial ql_n(\hat{\theta})}{\partial \boldsymbol{\alpha}^T} I_{qq}^{-1} \frac{\partial ql_n(\hat{\theta})}{\partial \boldsymbol{\alpha}} = \frac{1}{n}\sum_{h=1}^{q}\left(\frac{1}{\kappa_\eta - 1}\sum_{t=1}^{n}\left(\frac{X_t^2}{\hat{\omega}^2} - \frac{1}{\hat{\omega}}\right)\begin{pmatrix} X_{t-1}^2 \\ \vdots \\ X_{t-p}^2 \end{pmatrix}\right)^2 \tag{28}$$

but I_{pp}^{-1} and κ_η are unknown and have to be estimated. There are several ways to do this. Here, only one example of a HAC (heteroscedastic and autocorrelation consistent) covariance matrix estimator will be presented. For this let $\text{Score}(n) := \frac{1}{\sqrt{n}}\frac{\partial ql_n(\theta_0)}{\partial \theta}$

$$\begin{aligned} \hat{I}(S_T) &:= \frac{T}{T-r}\sum_{j=-T+1}^{T-1} ker\left(\frac{j}{S_T}\right)\hat{\Gamma}(j), \text{ where} \\ \hat{\Gamma}(j) &= \begin{cases} \frac{1}{T}\sum_{n=j+1}^{T} \text{Score}(n)\,\text{Score}(n-j)^T & \text{for } j \geq 0, \\ \frac{1}{T}\sum_{n=-j+1}^{T} \text{Score}(n+j)\,\text{Score}(n)^T & \text{for } j < 0, \end{cases} \\ ker(x) &:= 1 - \mid x \mid, \qquad S_T \in o(T^{1/4}), \qquad \lim_{T\to\infty} S_T = \infty \end{aligned} \tag{29}$$

There are several choices for the kernel function and the S_T function. This estimator was proposed in [Newey and West, 1987], and is consistent under the conditions given in Theorem 2 of their paper. The notation, in which the estimator is stated, is from [Andrews, 1991], where the class of estimators is extended by on one hand giving several different kernel functions and on the other hand increasing the possible choices of S_T to $o(T)$.

3.3.6 Estimation of the GARCH model

Estimation of the GARCH model can be done via Maximum Likelihood Estimation. Here, again the gaussian quasi-log-likelihood will be used. This Section and the next one follow the steps presented in [Francq and Zakoïan, 2004], where the proofs for the results stated herein can be found.
For these models $q + p + 1$ parameters need to be estimated.

$$\theta = (\omega, \alpha_1, \dots \alpha_q, \beta_1, \dots, \beta_p)^t \in \Theta \subset]0, \infty[\times[0, \infty[^{p+q}$$

Let θ_0 be the true and unknown parameter values, and let $X_1, \dots X_n$ be a known realization from the GARCH model with parameters θ.

Conditional on starting values $X_0, \dots, X_{1-q}, \hat{\sigma}_0^2, \dots \hat{\sigma}_{1-p}^2$ the quasi-likelihood is given as

$$L_n(\theta) = L_n(\theta; X_1, \dots X_n) = \sum_{t=1}^{n} \frac{1}{\sqrt{2\pi\hat{\sigma}_t^2}} \exp\left(- \frac{X_t^2}{2\hat{\sigma}_t^2} \right)$$

where $\hat{\sigma}_t^2$ is recursively defined through

$$\hat{\sigma}_t^2 = \omega + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{i=1}^{p} \beta_i \hat{\sigma}_{t-i}^2, \quad \forall t \geq 0$$

The two different options of starting values suggested in [Francq and Zakoïan, 2004] are

$$X_0^2 = \dots = X_{1-q}^2 = \hat{\sigma}_0^2 = \dots = \hat{\sigma}_{1-p}^2 = \omega$$

and

$$X_0^2 = \dots = X_{1-q}^2 = \hat{\sigma}_0^2 = \dots = \hat{\sigma}_{1-p}^2 = X_1^2$$

In [Francq and Zakoian, 2019] additionally the long run variance from Theorem 2 is suggested

$$X_0^2 = \dots = X_{1-q}^2 = \hat{\sigma}_0^2 = \dots = \hat{\sigma}_0^2 = \omega(1 - \sum_{i=1}^{q} \alpha_i \mathbb{V}[\eta] + \sum_{i=1}^{p} \beta_i)^{-1}$$

for the case that second order stationarity is assumed.
Lastly, the estimator can be formulated via the quasi-log-likelihood as

$$\hat{\theta}_n := \underset{\theta \in \Theta}{\arg\min}\ l_n(\theta) = \underset{\theta \in \Theta}{\arg\min} \frac{1}{n} \sum_{t=1}^{n} \frac{X_t^2}{\hat{\sigma}_t^2} + \log \hat{\sigma}_t^2 \tag{30}$$

It is also possible to use a Maximum Likelihood Estimator with a predetermined density for η_t. This can sometimes yield an advantage, by either reducing the assumptions on the moments of η_t for asymptotic normality, that will be needed in Theorem 13, or by making the estimator efficient, if it is not already. In the case of a missspecified distribution, this equals a non-gaussian quasi-likelihood estimator, which can be consistent given a constraint on the distribution η, which is already fullfilled in the normal case for the choice of $\mathbb{E}[\eta] = 0$. A list of these constraints can be found in [Francq and Zakoian, 2019].

3.3.7 Consistency conditions of quasi-likelihood estimators

First define

$$\mathcal{A}_\theta(z) := \sum_{i=1}^{q} \alpha_i z^i, \;\; \mathcal{B}_\theta(z) := 1 - \sum_{i=1}^{p} \beta_i z^i$$

Further let $\gamma(A)$ be the top Lyapunov exponent as defined in Section 3.3.2 and θ_0 be the true value of θ.

Theorem 8. *For the quasi-likelihood estimator $\hat{\theta}_n$ in equation (30), with initial conditions as above, it holds that $\hat{\theta}_n \to \theta_0$, as $n \to \infty$, if the following conditions are met:*

1. *$\theta_0 \in \Theta$ and Θ is compact*
2. *$\gamma(A) < 0$ and $\forall \theta \in \Theta$, $\sum_{i=1}^{p} \beta_i < 1$*
3. *η_t^2 has a non-degenerate distribution and $\mathbb{E}[\eta_t^2] = 1$*
4. *If $p > 0$, $\mathcal{A}_\theta(z)$ and $\mathcal{B}_\theta(z)$ have no common root, $\mathcal{A}_\theta(1) \neq 0$ and $\alpha_q + \beta_p \neq 0$*

Theorem 9. *Under the conditions of Theorem 8 and the additional assumptions*

1. *$\theta_0 \in \mathring{\Theta}$ where $\mathring{\Theta}$ is the interior of Θ*
2. *$\kappa_\eta := \mathbb{E}[\eta^4] < \infty$*

it holds that

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \to \mathcal{N}(0, (\kappa_\eta - 1)J^{-1}) \tag{31}$$

where

$$J := \mathbb{E}_{\theta_0}\left[\frac{\partial^2 \left(\frac{X_t^2}{\sigma_t^2} + \log \sigma_t^2\right)}{\partial\theta\partial\theta^T}\right] = \mathbb{E}_{\theta_0}\left[\frac{1}{\sigma_t^4(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta^T}\right]$$

As mentioned above, these theorems originate from [Francq and Zakoïan, 2004], where their proofs can be found. Condition 1 of Theorem 9 implies that $\alpha_i > 0$ and $\beta_i > 0$, which is not in line with the model assumptions. This is a problem, making this asymptotic distribution not always valid. Although by imposing

$$\mathbb{E}[X_t^6] < \infty \tag{32}$$

as an additional assumption to ensure the existence of J from Theorem 13, one gets a similar asymptotic distribution, but constrained on a space $\Delta := \Delta_1 \times \Delta_2 \times \cdots \times \Delta_{p+q+1}$ where $\Delta_1 = \mathbb{R}$ and for $\imath = 1, \ldots, p+q+1$: $\Delta_i = \mathbb{R}$ if $\theta_i \neq 0$ and $\Delta_i = [o, \infty[$ if $\theta_i = 0$. Further the condition in equation (32) is sufficient, but not necessary. [Francq and Zakoian, 2019]

3.4 GARCH Extensions

There are several different types of univariate GARCH models that are being used to model cryptoassets. Usually, they include more parameters to model certain properties that the standard GARCH model misses. In this subsection, a few of these models will briefly be introduced, but their implementation will not be discussed in as much detail as for the GARCH model above.

3.4.1 IGARCH model

Definition 3.9. Let $(X_t)_{t\in\mathbb{Z}}$ be a time series, $(\eta_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence with law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_1] = 0$. Further let $p \geq 0$, $q > 0$, $\omega \geq 0$, $\alpha_i \geq 0 \ \forall i = 1 \dots q$ and $\beta_i \geq 0 \ \forall i = 1 \dots p$. Then $(X_t)_{t\in\mathbb{Z}}$ is called an **IGARCH(p,q) process**, if

$$X_t = \sigma_t \eta_t \tag{33}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{34}$$

$$\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i = 1 \tag{35}$$

IGARCH stands for integrated GARCH model and was introduced in [Engle and Bollerslev, 1986]. This model is more or less a special case of the GARCH model. The only extension is that ω is allowed to be zero, in this case the model will almost surely converge to zero, as shown in [Nelson, 1990]. For the following, it is assumed that $\omega > 0$.
The actual important attribute of this model is, that even though the second order stationarity is not given anymore (since the unconditional variance does not exist), it still is strictly stationary. For the IGARCH(1,1) model, the proof can be found in [Nelson, 1990]. The general stationarity results are the same as specified above for the GARCH model.
This the IGARCH model creates a volatility process that characterizes itself through shock persistence. Furthermore if the sample sizes for the estimation of a GARCH model increases, it often approaches an IGARCH model. This is typically attributed to model missspecification [Paolella, 2018].

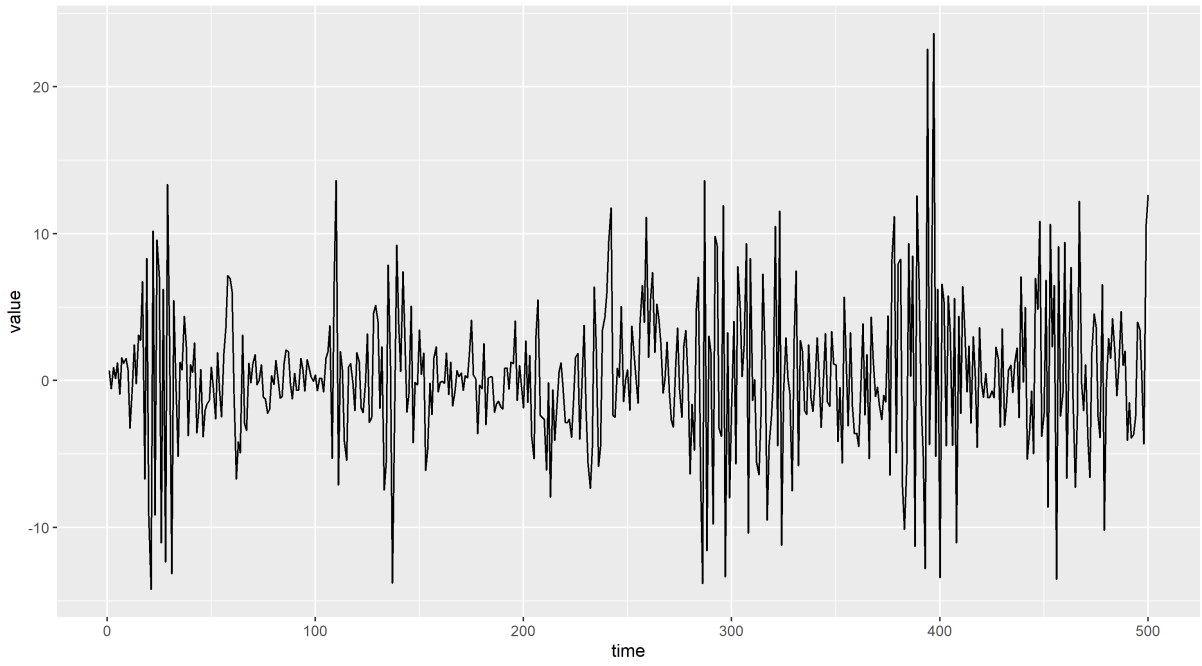


Figure 8: Sample path of an IGARCH(1,1) model with $\alpha = \beta = 0.5$ and $\omega = 1$

Figure 8 depicts a sample path of an IGARCH(1,1) model. Notable here is that even though the conditional variance does not exist, it is easy to see that the variance does not explode, which is in agreement with the stationarity result.

3.4.2 EGARCH model

Definition 3.10. Let $(X_t)_{t\in\mathbb{Z}}$ be a time series, $(\eta_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence with law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_1] = 0$ and $\mathbb{V}[\eta_1] = 1$. Further let $p \geq 0$, $q > 0$, $\omega, \alpha_i, \beta_j, \theta, \gamma \in \mathbb{R}$, for $i = 1, \ldots, q$ and $j = 1, \ldots, p$. Then $(X_t)_{t\in\mathbb{Z}}$ is called an **EGARCH(p,q) process**, if

$$X_t = \sigma_t \eta_t \tag{36}$$

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i g(\eta_{t-i}) + \sum_{i=1}^{p} \beta_i \ln \sigma_{t-i}^2 \tag{37}$$

$$g(\eta_{t-i}) = \theta \eta_{t-i} + \gamma(\mid \eta_{t-i} \mid - \mathbb{E}[\mid \eta_{t-i} \mid]) \tag{38}$$

EGARCH stands for exponential GARCH and was proposed in [Nelson, 1991] to address three problems. Firstly, that volatility tends to rise in response to "bad news", i.e. negative returns, and to fall in response to good news. This dynamic is named "leverage effect" and is not captured in the standard GARCH model, since the impact is weighted symmetrically. Secondly, the constraints of nonnegativity the aforementioned GARCH models impose on the

parameters. Third and lastly, shock persistence in the GARCH model is subject to the norms used to measure it.

The solution to the first two problems is visible in the model equations. Equation (38) has a symmetric component with parameter γ and an asymmetric one with parameter θ. In order to get the assymetric effect desired in [Nelson, 1991] in an EGARCH(1,1) model it must hold that $\theta\alpha_1 < 0$.

Further parameters for EGARCH can be chosen in $\mathbb{R}$, since the log of σ_t^2 can be negative, though it is suggested to choose parameters so that $-\gamma < \theta < \gamma$, guaranteeing that large absolute values of η_t will increase the variance, which is generally assumed to be true in financial time series.

Strict stationarity of the model is given under the fourth condition of Theorem 8 with the addition that no polynomial has a root in the unit disk [Francq and Zakoian, 2019].

Equation (37) can be rewritten as

$$\sigma_t^2 = e^{\omega} \prod_{i=1}^{q} \exp\left(\alpha_i g(\eta_{t-i})\right) \prod_{i=1}^{p} \sigma_{t-i}^{2\beta_i}$$

to illustrate that this model uses multiplicative dynamics. Further, by plugging in g one can see that in the exponent the gamma coefficient comes in front of

$$\frac{\exp(\mid \eta_{t-i} \mid)}{\exp(\mathbb{E}[\mid \eta_{t-i} \mid])}$$

showcasing that the ratio between the exponent of the expectation and the exponent of the realization is used for the symmetric part.

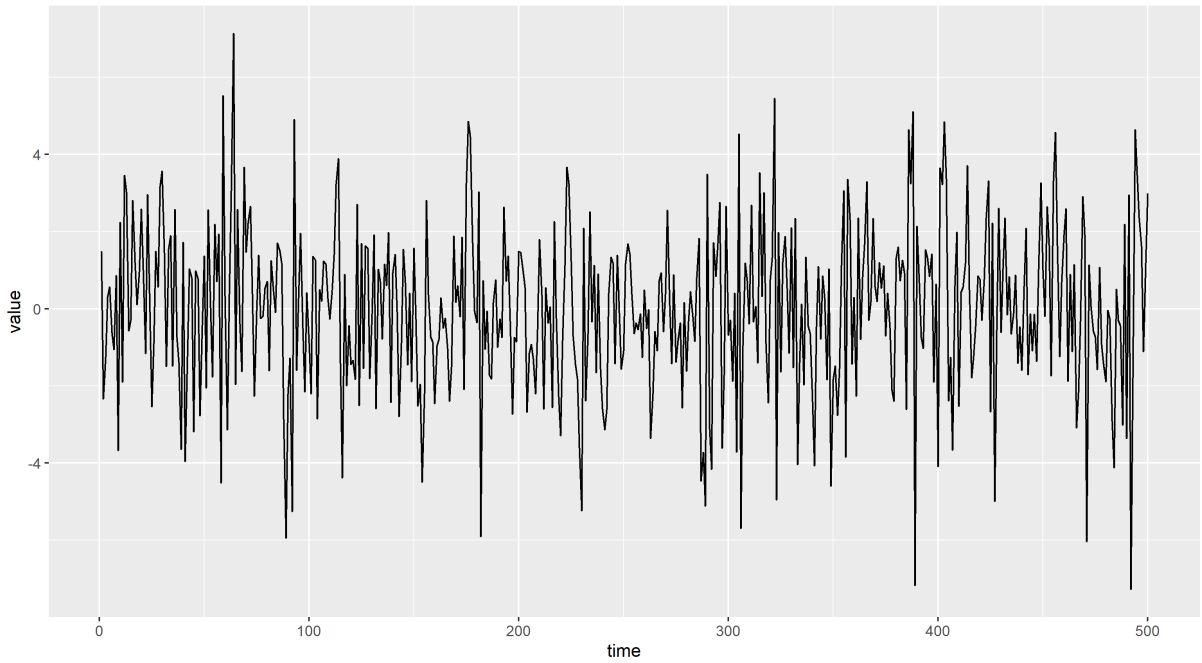


Figure 9: Sample path of an EGARCH(1,1) model with $\alpha = \beta = 0.3$, $\omega = 1$, $\theta = -0.2$ and $\gamma = 0.7$

Figure 9 shows a realization of an EGARCH(1,1) model, with the parameters chosen as above; one can see, that a positive η is multiplied by $\theta + \gamma = 0.7$, while a negative one is multiplied by $\theta - \gamma = -0.9$. The coefficient being negative results in a positive addition to the variance, which has more impact than that of a positive realization, implying presence of the leverage effect. However this can only be observed when the parameters are known and is hardly visible in the plot.

3.4.3 TGARCH model

Definition 3.11. Let $(X_t)_{t\in\mathbb{Z}}$ be a time series, $(\eta_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence with law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_1] = 0$ and $\mathbb{V}[\eta_1] = 1$. Further, let $p \geq 0$, $q > 0$, $\omega, \alpha_{i,+}, \alpha_{i,-}, \beta_j \in \mathbb{R}$ for $i = 1, \ldots, q$ and $j = 1, \ldots, p$. Then $(X_t)_{t\in\mathbb{Z}}$ is called a **TGARCH(p,q) process**, if

$$X_t = \sigma_t \eta_t \tag{39}$$

$$\sigma_t = \omega + \sum_{i=1}^{q} \left(\alpha_{i,+} X_{t-i}^{+} - \alpha_{i,-} X_{t-i}^{-} \right) + \sum_{i=1}^{p} \beta_i \sigma_{t-i} \tag{40}$$

where

$$X_t^{+} := \begin{cases} X_t, & \text{if } X_t > 0 \\ 0, & \text{else} \end{cases}, \quad X_t^{-} := \begin{cases} X_t, & \text{if } X_t < 0 \\ 0, & \text{else} \end{cases}$$

TGARCH stands for threshold GARCH and does not model the variance, but the standard deviation. It was proposed in [Zakoian, 1994] and also features asymmetry. The actual conditional standard deviation is $\mid \sigma_t \mid$ and not σ_t, but negative values are allowed for the equations above.
A stationarity condition for the TGARCH(1,1) model is given by

$$\omega > 0, \alpha_{1,+} \geq 0, \alpha_{1,-} \geq 0, \beta_1 \geq 0 \quad \text{and} \quad \mathbb{E}[\alpha_{1,+} + -\eta_1^+ \alpha_{1,-} \eta_1^- + \beta_1] < 0$$

Similar as for the GARCH model in Theorem 4 a strict stationarity condition can be found by regarding the top Lyapunov exponent for a matrix equation for the TGARCH model [Francq and Zakoian, 2019].

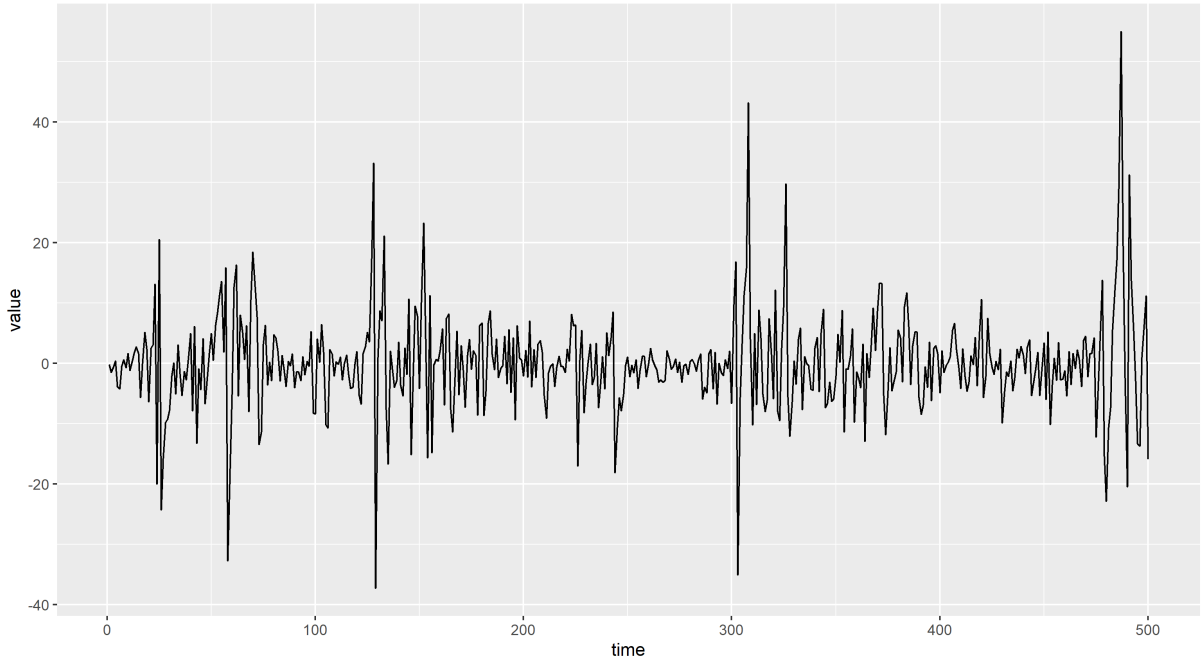


Figure 10: Sample path of a TGARCH(1,1) model with $\alpha_+ = 0.4$, $\alpha_- = -0.6$, $\beta = 0.5$ and $\omega = 1$

The sample path in Figure 10 belongs to a TGARCH(1,1) model; in contrast to the EGARCH model, the amount of asymmetry is directly visible in the choices of the α coefficients. Again, visibility of these asymmetries is not given in the plot.

3.4.4 GJR-GARCH

Definition 3.12. Let $(X_t)_{t\in\mathbb{Z}}$ be a time series, $(\eta_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence with law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_1] = 0$ and $\mathbb{V}[\eta_1] = 1$. Further let $p \geq 0$, $q > 0$, $\omega > 0$, $\alpha_i \geq 0$, $\alpha_i^* \geq 0$, $\beta_j \geq 0$ for

$i = 1, \dots q$ and $j = 1, \dots, p$. Then $(X_t)_{t\in\mathbb{Z}}$ is called a **GJR GARCH(p,q) process**, if

$$X_t = \sigma_t \eta_t \tag{41}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} (\alpha_i \mathbb{1}_{(X_{t-i}\geq 0)} + \alpha_i^* \mathbb{1}_{(X_{t-i}<0)}) X_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{42}$$

GJR-GARCH stands for Glosten Jagannathan and Runkle GARCH and was introduced in [Glosten et al., 1993]. It includes asymmetry in a framework that is even more similar to the original GARCH model than the previous two. The main selling point of the GJR-GARCH model is that it captures the leverage effect, but is way simpler than the EGARCH model and can be compared to a pure GARCH model more easily.
Stationarity of the GJR-GARCH model is studied in [Noori and Mohammad, 2021], where a necessary and sufficient condition for second order stationarity is found to be

$$\sum_{i=1}^{q} (\frac{\alpha}{2} + \frac{\alpha^*}{2}) + \sum_{i=1}^{p} \beta_i < 1$$

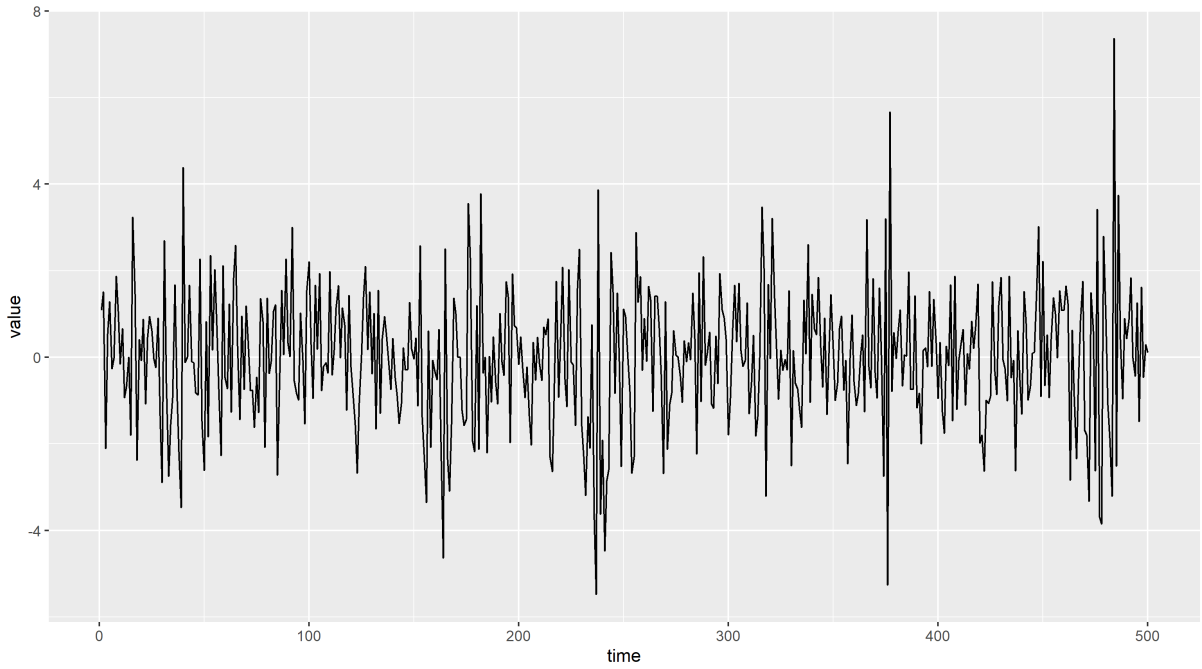


Figure 11: Sample path of a GJR-GARCH(1,1) model with $\alpha = 0.1$, $\alpha_* = 0.8$, $\beta = 0.2$ and $\omega = 1$

Figure 11 displays a realization of a GJR-GARCH(1,1) model, herein the asymmetry is vastly increased, but still it is not observable in the plot alone.

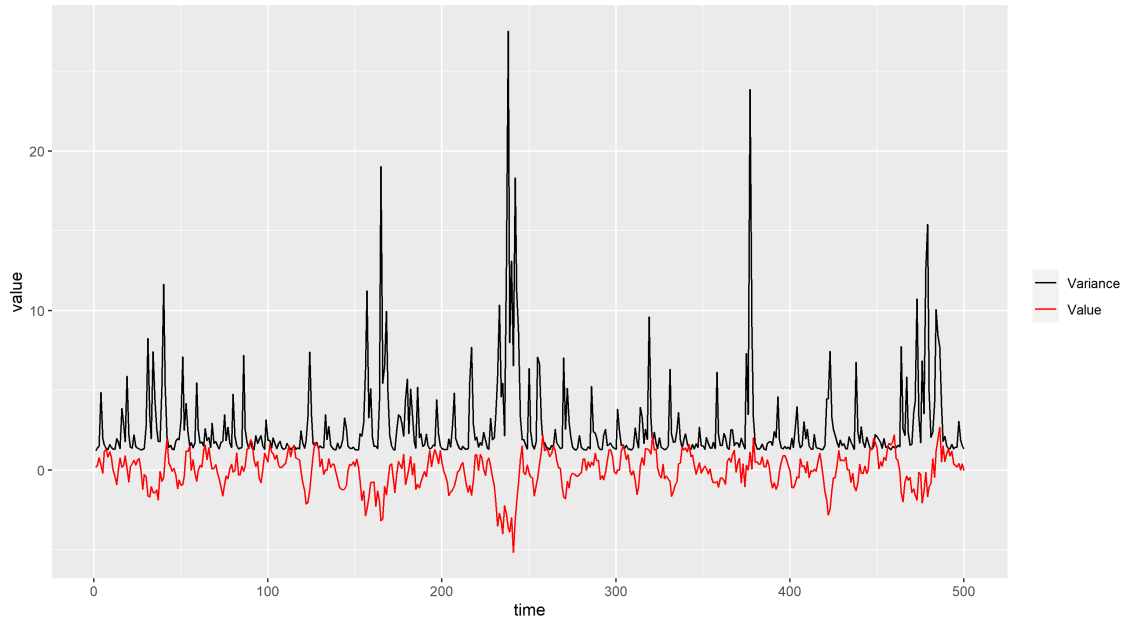


Figure 12: Shows the variance of the GJR-GARCH(1,1) realization from above (black) and a smoothed version of the returns (red) as comparison

In Figure 12, two paths are displayed. Shown in black is the variance of the GJR-GARCH(1,1) realization of Figure 11, and in red a smoothed version of the returns. The smoothing is done by taking an arithmetic mean over five observations $(i-2, i-1, i, i+1, i+2)$. The purpose is to depict the trend of the returns X_t around time t. With this smoothing, one can see, that in periods of time where X_t had a positive trend, the variance remains rather small, but as soon as a negative trend occurs, spikes are created, which grow until the trend becomes positive again. At this point, the variance starts to fall off again.

3.4.5 APGARCH

Definition 3.13. Let $(X_t)_{t\in\mathbb{Z}}$ be a time series, $(\eta_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence with law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_1] = 0$ and $\mathbb{V}[\eta_1] = 1$. Further, let $p \geq 0$, $q > 0$, $\omega > 0$, $\alpha_i \geq 0$, $|\ \zeta_i\ | \leq 1$, $\beta_j \geq 0$, for $i = 1, \ldots, q$, $j = 1, \ldots, p$ and $\delta > 0$. Then $(X_t)_{t\in\mathbb{Z}}$ is called an **APGARCH(p,q) process**, if

$$X_t = \sigma_t \eta_t \tag{43}$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^{q} \alpha_i (|\ X_{t-i}\ | - \zeta_i X_{t-i})^\delta + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^\delta \tag{44}$$

APGARCH stands for asymmetric power GARCH and was introduced in [Ding et al., 1993]. The purpose of δ is to improve the modelling of the "long memory" property. The δ_i are similar to those in the EGARCH model, just that the APGARCH model uses an additive variance

equation. The model can be reduced to the standard GARCH by setting $\delta = 2$ and $\zeta_i = 0$, or to the TGARCH model by setting $\delta = 1$. It is also a generalization of the GJR-GARCH model, by using $\delta = 2$ and $\alpha_i,\ \zeta_i$ that solve $\alpha_{GJR} = \alpha(1-\zeta)^2$ and $\alpha^*_{GJR} = \alpha(1+\zeta)^2$.
A strict stationarity condition for the APGARCH(1,1) model is

$$\mathbb{E}[\alpha_1(1-\zeta_1)^\delta \mid \eta_t \mid^\delta \mathbb{1}_{(\eta_t>0)} + \alpha_1(1+\zeta_1)^\delta \mid \eta_t \mid^\delta \mathbb{1}_{(\eta_t<0)} + \beta_1]$$

Since the APARCH model is a generalization of the above, for sample paths it is referred to those Figures, as they incorporate the more commonly used cases.

3.4.6 GARCHM model

Definition 3.14. Let $(X_t)_{t\in\mathbb{Z}}$ be a time series, $(\eta_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence with law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_1] = 0$ and $\mathbb{V}[\eta_1] = 1$. Further, let $p \geq 0$, $q > 0$, $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ for $i = 1, \ldots, q$, $j = 1, \ldots, p$ and $\lambda \in \mathbb{R}$. Then $(X_t)_{t\in\mathbb{Z}}$ is called a **GARCH(p,q)-in-mean process**, if

$$X_t = \sigma_t \eta_t + \lambda \sigma_t^2 \tag{45}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{46}$$

[Tsay, 2010]

GARCHM stands for GARCH-in-mean and is an add-on to the GARCH model, that can be used on any of the described models above. Here, the variance impacts the mean equation directly. This addition can be interpreted as a risk premium, by setting $\lambda < 0$ to reduce the returns.

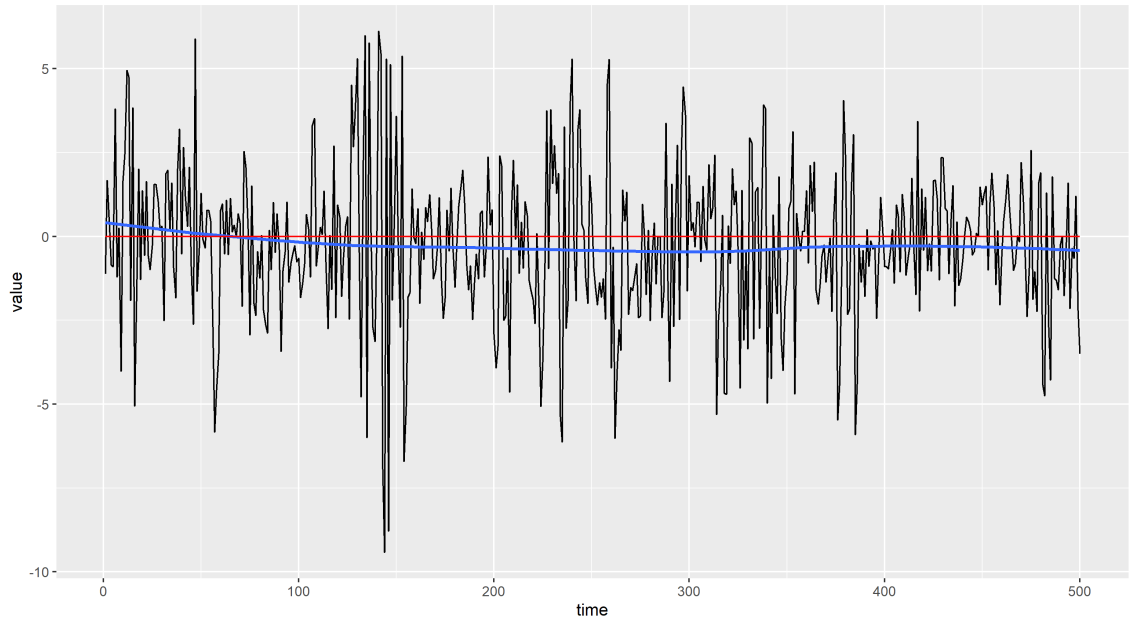


Figure 13: Sample path of GARCH-in-mean(1,1) model, with $\alpha = 0.4$, $\beta = 0.4$, $\lambda = 0.05$ and $\omega = 1$

In Figure 13, a sample path of a GARCH-in-mean(1,1) model is shown. In blue, a trendline is displayed, and in red, the x-axis is highlighted. From equation (45), one can see that it holds that $\mathbb{E}[X_t] = \mathbb{E}[\sigma_t \eta_t + \lambda \sigma_t^2] = \lambda \mathbb{E}[\sigma_t^2] < 0$ for $\lambda < 0$. This fact is visualised above by the comparison between the blue trendline and the red x-axis.

4 Multivariate GARCH models

4.1 Changes for multivariate models

While the concept of strict stationarity can easily be used for multivariate processes as well, second order stationarity needs to be generalized.

Definition 4.1. A process $X_t \in \mathbb{R}^m$ is **second order stationary** if

(i) $\mathbb{E}[X_{it}^2] < \infty, \ \forall t \in \mathbb{Z}, i = 1, \dots, m$

(ii) $\mathbb{E}[X_t] = \mu, \ \forall t \in \mathbb{Z}$

(iii) $\text{Cov}[X_t, X_{t+h}] = \Gamma(h), \ \forall t, h \in \mathbb{Z}$

[Francq and Zakoian, 2019]

Remark. While the input of Γ is still a scalar, the output now is an $m \times m$ matrix.

When looking at the first equation of the univariate GARCH model

$$X_t = \sigma_t \eta_t$$

one can observe, that, in the multivariate, case X_t and η_t become vectors, but σ_t becomes a matrix. So, the equation transforms to

$$X_t = H_t^{1/2} \eta_t$$

where $X_t \in \mathbb{R}^m$, $H_t \in \mathbb{R}^{m \times m}$, and $\eta_t = (\eta_t^1, \dots, \eta_t^m) \in \mathbb{R}^m$ are i.i.d. random vectors with distribution $\mathcal{L}$ on $\mathbb{R}^m$ with $\mathbb{E}[\eta_t^i] = 0, \ \forall i = 1, \dots, m$ and $\mathbb{V}[\eta] = \text{I}_m$. In the GARCH model, the variance only depends on the past so it holds that

$$\mathbb{V}[X_t \mid \mathcal{F}_{t-1}] = H_t \tag{47}$$

where $\mathcal{F}_{t-1}$ is the information set up to time $t-1$. The condition on the variance in the univariate case was $\sigma > 0$, which in the multivariate case generalizes to $H_t > 0$ and $h_{ij} = h_{ji}$. Ensuring positive definiteness and symmetry is more difficult than just ensuring positivity. However, if H_t is positive definite and symmetric, it is, by the Cholesky decomposition, possible to choose $H_t^{1/2}$ as a triangular matrix with positive diagonal elements.

Definition 4.2. Let $A = (a_{ij})_{i,j=1}^m \in \mathbb{R}^{m \times m}$ be a symmetric matrix. Then

$$\text{vech}(A) := (a_{11}, \dots, a_{1m}, a_{22}, \dots, a_{2m}, \dots, a_{mm})$$

is a bijective transformation from the class of symmetric $m \times m$ matrices to $m \times (m+1)/2$ vectors.

With this it is possible to define a very general multivariate GARCH (MGARCH) model

Definition 4.3. Let $(X_t)_{t\in\mathbb{Z}}$ be an m-dimensional time series vector, $(\eta_t)_{t\in\mathbb{Z}}$ be a vector valued i.i.d. sequence with law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_t] = 0$ and $\mathbb{V}[\eta_t] = \mathrm{I}_m$. Further, let $p \geq 0$, $q > 0$, $\omega \in \mathbb{R}^{\frac{m(m+1)}{2}}$ and $A_i, B_j \in \mathbb{R}^{\frac{m(m+1)}{2}\times\frac{m(m+1)}{2}}$ for $i = 1, \ldots, q$, $j = 1, \ldots, p$. Then $(X_t)_{t\in\mathbb{Z}}$ is called a **VEC-GARCH(p,q) process**, if

$$X_t = H_t^{1/2}\eta_t \tag{48}$$

where H_t is positive definite, such that

$$\mathrm{vech}(H_t) = \omega + \sum_{i=1}^{q} A_i \mathrm{vech}(X_{t-i}X_{t-i}^T) + \sum_{i=1}^{p} B_i \mathrm{vech}(H_{t-i}) \tag{49}$$

The ARCH version of this model was introduced in [Engle et al., 1984] and generalized in [Bollerslev et al., 1988]. The VEC-GARCH model uses all information given to predict the covariance matrix for the next step. Therefore, it is very general, which is not always an advantage. The model uses $(p+q)\big(\frac{m(m+1)}{2}\big)^2 + \frac{m(m+1)}{2}$ parameters, which is of order four in m and becomes large quickly. This can lead to several problems like, for example, overfitting or costly estimation.

There are several more parsimonious parametrisations for MGARCH models, in this paper the CCC- and DCC-GARCH models are discussed in more detail. Furthermore one other model has been frequently used to model cryptoassets, so for the sake of completeness it is briefly introduced here.

The BEKK (Baba Engel Kraft Kroner) GARCH model was introduced in [Engle and Kroner, 1995], but originally, four people worked on the creation, therefore the name. It is a direct restriction of the VEC-GARCH model above, by using $m\times m$ matrices and working on $X_{t-i}X_{t-i}^T$ directly instead of $\mathrm{vec}(X_{t-i}X_{t-i}^T)$, which drastically reduces the amount of parameters used.

Definition 4.4. Let $(X_t)_{t\in\mathbb{Z}}$ be an m-dimensional time series vector, $(\eta_t)_{t\in\mathbb{Z}}$ be a vector valued i.i.d. sequence with law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_1] = 0$ and $\mathbb{V}[\eta_1] = \mathrm{I}_m$. Further, let $p \geq 0$, $q > 0$, $K \in \mathbb{N}$, $\Omega \in \mathbb{R}^{m\times m}$, Ω positive definite and $A_i, B_j \in \mathbb{R}^{m\times m}$ for $i = 1, \ldots, q$, $j = 1, \ldots, p$. Then $(X_t)_{t\in\mathbb{Z}}$ is called a **BEKK-GARCH(p,q) process**, if

$$X_t = H_t^{1/2}\eta_t \tag{50}$$

with

$$H_t = \Omega + \sum_{i=1}^{q}\sum_{k=1}^{K} A_{ik}X_{t-i}X_{t-i}^T A_{ik}^T + \sum_{i=1}^{p}\sum_{k=1}^{K} B_{ik}H_{t-i}B_{ik}^T \tag{51}$$

4.2 CCC-GARCH model

4.2.1 Definition of the model

The Constant Conditional Correlations GARCH model was introduced by [Bollerslev, 1990]. The name refers to the restriction that is imposed on the data. The Pearson correlations of the X_t are assumed to be constant, i. e. if $X_t = H_t^{1/2}\eta_t$ and $H_t = (h_{ij})_{i,j=1}^m$ then it holds that

$$\frac{h_{ij,t}}{h_{ii,t}^{\frac{1}{2}} h_{jj,t}^{\frac{1}{2}}} = \rho_{ij} \tag{52}$$

This makes it possible to only update the $h_{ii},\ \ i = 1, \dots, m$ instead of the whole matrix. Define

$$R := (\rho_{ij})_{i,j=1}^m \tag{53}$$

Further, define for each m the function diagm, that returns a diagonal $m \times m$-matrix with diagonal equal to the input vector or the diagonal of the input matrix.

Definition 4.5. Let $(X_t)_{t\in\mathbb{Z}}$ be an m-dimensional time series vector, $(\eta_t)_{t\in\mathbb{Z}}$ be a vector valued i.i.d. sequence with independent components that follows a law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_1] = 0_m$ and $\mathbb{V}[\eta_1] = \mathrm{I}_m$. Further, let $h_t = (h_{11,t}, \dots, h_{mm,t})$, $X_t^2 = (X_{1t}^2, \dots, X_{mt}^2)$, $R = (\rho_{ij})_{i,j=1}^m$ with ρ_{ij} as in equation (52) and $D_t = \mathrm{diagm}\big((\sqrt{h_{11,t}}, \dots, \sqrt{h_{mm,t}})\big)$. Moreover let $p \geq 0$, $q > 0$, $\omega \in (\mathbb{R}^+)^m$ and $A_i, B_j \in (\mathbb{R}_0^+)^{m\times m}$ for $i = 1, \dots, q$, $j = 1, \dots, p$..
Then $(X_t)_{t\in\mathbb{Z}}$ is called a **CCC-GARCH(p,q) process**, if

$$X_t = H_t^{1/2}\eta_t \tag{54}$$

with

$$H_t = D_t R D_t \tag{55}$$

and

$$h_t = \omega + \sum_{i=1}^{q} A_i X_t^2 + \sum_{i=1}^{p} B_i h_{t-i} \tag{56}$$

Note, that the order of parameters in m is now only two. Ensuring positivity of H_t is quite straightforward. It suffices to choose R positive definite and symmetric, and that the coefficients of A_i and B_i are non-negative. Further, if both A_i and B_j are diagonal for $i = 1, \dots, q,\ j = 1, \dots p$, respectively, then the CCC-GARCH model consists of m univariate GARCH(p,q) models with correlation matrix R.

Example. Illustration of the difference to the VEC-GARCH model

Consider a bivariate CCC-ARCH(1) model, with

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad R = \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} \qquad \omega = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Then it holds that

$$h_{1,t} = 1 + 2X_{1,t-1}^2 + X_{2,t-1}^2$$

$$h_{2,t} = 1 + X_{1,t-1}^2 + 2X_{2,t-1}^2$$

From this follows for H_t

$$H_t = D_t R D_t = \begin{bmatrix} \sqrt{h_{1,t}} & 0 \\ 0 & \sqrt{h_{2,t}} \end{bmatrix} \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} \begin{bmatrix} \sqrt{h_{1,t}} & 0 \\ 0 & \sqrt{h_{2,t}} \end{bmatrix} = \begin{bmatrix} h_{1,t} & c\sqrt{h_{1,t}}\sqrt{h_{2,t}} \\ c\sqrt{h_{1,t}}\sqrt{h_{2,t}} & h_{2,t} \end{bmatrix}$$

Plugging in the values for $h_{1,t}$ and $h_{2,t}$, the off-diagonal elements of H_t become

$$c\sqrt{1 + 2X_{1,t-1}^2 + X_{2,t-1}^2}\sqrt{1 + X_{1,t-1}^2 + 2X_{2,t-1}^2}$$

This term is nonlinear in the squared returns, therefore there can not be a representation of this in form of a VEC-GARCH model as in Definition 4.3. If R is a diagonal matrix, i.e. an identity matrix, this becomes possible, though in this case one should not use a multivariate model to begin with.

4.2.2 Stationarity of the CCC-GARCH model

In a similar manner as in the steps leading to Theorem 4, a matrix for the CCC-GARCH model can be constructed.
Define $\hat{\eta}_t := R^{\frac{1}{2}}\eta_t$ and $\Gamma_t := \text{diagm}\big((\hat{\eta}_{1,t}^2, \ldots, \hat{\eta}_{m,t}^2)\big)$ with this and the convention that the squares are applied element wise, it holds that

$$X_t^2 = H_t\eta_t^2 = D_t R D_t \eta_t^2 = R D_t^2 \eta_t^2 = R \ \text{diagm}(\eta_t^2) h_t = \Gamma_t h_t$$

Multiplying equation (56) with Γ_t therefore yields

$$X_t^2 = \Gamma_t \omega + \sum_{i=1}^{q} \Gamma_t A_i X_t^2 + \sum_{i=1}^{p} \Gamma_t B_i h_{t-i} \tag{57}$$

which can be rewritten as

$$z_t = b_t + A_t z_{t-1}$$

with

$$b_t = \begin{pmatrix} \Gamma_t\omega \\ 0 \\ \vdots \\ 0 \\ \omega \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{m(p+q)} \qquad z_t = \begin{pmatrix} X_t \\ \vdots \\ X_{t-q+1} \\ h_t \\ \vdots \\ h_{t-p+1} \end{pmatrix} \in \mathbb{R}^{m(p+q)}$$

$$A_t = \begin{bmatrix} \Gamma_t A_1 & \dots & \dots & \dots & \Gamma_t A_q & \Gamma_t B_1 & \dots & \dots & \dots & \Gamma B_p \\ \mathrm{I}_m & \mathbf{0} & \dots & \dots & \mathbf{0} & \mathbf{0} & \dots & \dots & \dots & \mathbf{0} \\ \mathbf{0} & \mathrm{I}_m & \ddots & & \vdots & \vdots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & & & & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathrm{I}_m & \mathbf{0} & \mathbf{0} & \dots & \dots & \dots & \mathbf{0} \\ A_1 & \dots & \dots & \dots & A_q & B_1 & \dots & \dots & \dots & B_p \\ \mathbf{0} & \dots & \dots & \dots & \mathbf{0} & \mathrm{I}_m & \mathbf{0} & \dots & \dots & \mathbf{0} \\ \vdots & & & & \vdots & \mathbf{0} & \mathrm{I}_m & \ddots & & \vdots \\ \vdots & & & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \dots & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathrm{I}_m & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{m(p+q)\times m(p+q)}$$

Let $A = (A_t)_{t\in\mathbb{Z}}$, then $\gamma(A)$ is called the top Lyapunov coefficient for the CCC-GARCH model.

Theorem 10. *A CCC-GARCH(p,q) model is strictly stationary, if and only if its top Lyapunov exponent satisfies* $\gamma(A) < 0$.

Proof. The proof can be found in [Francq and Zakoian, 2019].

□

Second order stationarity is again very similar to the univariate case. With matrices, the spectral radius has to be constrained, but apart from that a very similar result to Theorem 2 holds.

Theorem 11. *The CCC-GARCH process is second order stationary, if and only if*

$$\rho\Big(\sum_{i=1}^{q} A_i + \sum_{i=1}^{p} B_i\Big) < 1 \tag{58}$$

where ρ *denotes the spectral radius. In this case it holds that*

$$\tilde{h} := \mathbb{E}[h_t] = \Big(I_m - \sum_{i=1}^{q} A_i - \sum_{i=1}^{p} B_i\Big)^{-1}\omega \tag{59}$$

4.2.3 Estimation of the CCC-GARCH model

Similar to chapter 3.3.6, a QML estimators can be derived. Again, all theorems for this Section come from [Francq and Zakoian, 2019], where proofs are provided.
Let $\theta = (\omega, \text{vec}(A_1), \dots, \text{vec}(A_q), \text{vec}(B_1), \dots, \text{vec}(B_p), \rho_{21}, \dots, \rho_{m1}, \rho_{32}, \dots \rho_{m2}, \dots, \rho_{m,m-1})$ be the vector of parameters to be estimated and $(X_1, \dots, X_n)$ be an observation. Again, let the initial values be

$$X_0^2 = \dots = X_{1-q}^2 = h_0 = \dots = h_{1-p} = 0$$

or

$$X_0^2 = \dots = X_{1-q}^2 = h_0 = \dots = h_{1-p} = \omega$$

The normal quasi-likelihood function can be stated as

$$L_n(\theta) = L_n(\theta; X_1, \dots, X_n) = \prod_{t=1}^{n} \frac{1}{(2\pi)^{m/2} \mid \hat{H}_t \mid^{1/2}} \exp\left(-\frac{1}{2} X_t^T \hat{H}_t^{-1} X_t \right) \tag{60}$$

where $|\cdot|$ denotes the determinant and $\hat{H}_t$ is recursively defined for $t \geq 1$ by

$$\hat{H}_t = \hat{D}_t R \hat{D}_t, \quad \hat{D}_t = \text{diagm}((\sqrt{\hat{h}_{11,t}}, \dots, \sqrt{\hat{h}_{mm,t}}))$$

$$\hat{h}_t = (\sqrt{\hat{h}_{11,t}}, \dots, \sqrt{\hat{h}_{mm,t}})^T = \omega + \sum_{i=1}^{q} A_i X_t^2 + \sum_{i=1}^{p} B_i \hat{h}_{t-i}$$

Thus, the QMLE is

$$\theta_n = \arg\max_{\theta \in \Theta} L_n(\theta) \tag{61}$$

with parameter space $\Theta \subset (0, \infty)^m \times (0, \infty)^{m^2(p+q)} \times (-1, 1)^{m(m-1)/2}$.

Regarding the consistency of the estimator in equation (61), let $\gamma(A)$ once again be the Lyapunov coefficient of the CCC-GARCH model as in Section 4.2.2 and θ_0 be the true value of θ. Further, define the polynomials

$$\mathcal{A}_\theta(z) := \sum_{i=1}^{q} A_i z^i, \;\; \mathcal{B}_\theta(z) := I_m - \sum_{i=1}^{p} B_i z^i$$

In the vector case the assumption of no common roots does not guarantee that $\mathcal{B}_{\theta_0}(z)^{-1}\mathcal{A}_{\theta_0}(z)$

is unique. Therefore, some preliminary notions are necessary. Define the $m \times 2m$ matrix

$$M_1(\mathcal{A}_{\theta_0}, \mathcal{B}_{\theta_0}) := [A_q, B_p]$$

Definition 4.6. Two polynomials $P(B)$ and $Q(B)$ are **left coprime** if, for every common factor $U(B)$, i.e., $P(B) = U(B)P_1(B)$ and $Q(B) = U(B)Q_1(B)$ for some polynomials $P_1(B)$, $Q_1(B)$, we have that $U(B)$ is constant.

Theorem 12. *For the quasi-likelihood estimator θ_n in equation* (61)*, with initial conditions as above, it holds that $\hat{\theta} \to \theta_0$, as $n \to \infty$, if the following statements hold:*

1. *$\theta_0 \in \Theta$ and Θ is compact*
2. *$\gamma(A) < 0$ and $\forall \theta \in \Theta$, $\mid \mathcal{B}(z) \mid = 0 \implies \mid z \mid > 1$*
3. *$\eta_{i,t}^2$ has a non-degenerate distribution for $i = 1, \ldots, m$*
4. *If $p > 0$, $\mathcal{A}_{\theta_0}(z)$ and $\mathcal{B}_{\theta_0}(z)$ are left coprime and $M_1(\mathcal{A}_{\theta_0}, \mathcal{B}_{\theta_0})$ has full rank m*
5. *R is a positive definite correlation matrix $\forall \theta \in \Theta$*

Theorem 13. *Under the conditions of Theorem 12 and the additional assumptions*

1. *$\theta_0 \in \mathring{\Theta}$ where $\mathring{\Theta}$ is the interior of Θ*
2. *$\mathbb{E}\big[\mid\mid \eta_t \eta_t^T \mid\mid^2 \big] < \infty$*

it holds that

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \to \mathcal{N}(0, J^{-1} I J^{-1}) \tag{62}$$

where J is a positive definite matrix and I is a positive semi-definite matrix, given by

$$J := \mathbb{E}\left[\frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta^T} \right], \quad I := \mathbb{E}\left[\frac{\partial l_t(\theta_0)}{\partial \theta} \frac{\partial l_t(\theta_0)}{\partial \theta^T} \right]$$

where $l_t(\theta) = X_t^T \hat{H}_t^{-1} X_t + \log(\mid H_t \mid)$ is the negative quasi-log-likelihood up to a multiplicative scalar and additive constants.

Under the conditions of Theorem 11, [Francq et al., 2014] suggest an approach to reduce the number of parameters estimated by the QML estimator by m. This is done using that

$$\mathbb{E}[X_t^2] = \mathbb{E}[h_t] = \tilde{h} = \left(I_m - \sum_{i=1}^{q} A_i - \sum_{i=1}^{p} B_i \right)^{-1} \omega$$

This leads to a new representation for equation (56), which is

$$h_t - \tilde{h} = \sum_{i=1}^{q} A_i(X_t^2 - \tilde{h}) + \sum_{i=1}^{p} B_i(h_{t-i} - \tilde{h}) \tag{63}$$

A simple estimator of $\tilde{h}$ is given by

$$\hat{\tilde{h}} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

The remaining parameters have to be estimated with the QMLE as above. In a very general setting, this has limited impact, but often just diagonal matrices are considered for A_i and B_i. Adding that (1,1) are the most popular choices for (p,q), this can become a significant reduction. [Francq and Zakoian, 2019] includes a table with a comparison of runtimes to the pure QML estimation.

4.3 DCC-GARCH model

4.3.1 Definition of the model

One weakness of the CCC-GARCH model is its primary assumption of constant conditional correlation. This is a big restriction and usually not true in practice. In [Engle, 2000], it was found that, in a bivariate case where many methods are feasible, the CCC model is outperformed by other multivariate GARCH specifications, such as the BEKK-GARCH model. To solve this inadequacy the Dynamic Conditional Correlation GARCH model was introduced, which performs best in the cases studied.

Definition 4.7. Let $(X_t)_{t\in\mathbb{Z}}$ be an m-dimensional time series vector, $(\eta_t)_{t\in\mathbb{Z}}$ be a vector valued i.i.d. sequence with independent components that follows a law $\mathcal{L}$ that satisfies $\mathbb{E}[\eta_t] = 0_m$ and $\mathbb{V}[\eta_t] = \mathrm{I}_m$. Further, let $h_t = (h_{11,t}, \ldots, h_{mm,t})$, $X_t^2 = (X_{1t}^2, \ldots, X_{mt}^2)$, $\epsilon_{i,t} = X_{i,t}/\sqrt{h_{ii,t}}$, with $\epsilon_t = (\epsilon_{1,t}, \ldots, \epsilon_{m,t})$ and $D_t = \mathrm{diagm}\big((\sqrt{h_{11,t}}, \ldots, \sqrt{h_{mm,t}})\big)$. Moreover let S be the unconditional correlation matrix $S = \mathbb{E}[\epsilon_t\epsilon_t] \in \mathbb{R}^{m\times m}$ and $p \geq 0$, $q > 0$, $\omega \in (\mathbb{R}^+)^m$, $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta < 1$ and $A_i, B_j \in (\mathbb{R}_0^+)^{m\times m}$ for $i = 1, \ldots, q$, $j = 1, \ldots, p$.
Then $(X_t)_{t\in\mathbb{Z}}$ is called a **DCC-GARCH(p,q) process**, if

$$X_t = H_t^{1/2}\eta_t \tag{64}$$

with

$$H_t = D_t R_t D_t \tag{65}$$

$$h_t = \omega + \sum_{i=1}^{q} A_i X_t^2 + \sum_{i=1}^{p} B_i h_{t-i} \tag{66}$$

$$R_t = \text{diagm}(Q_t)^{-1/2} \; Q_t \; \text{diagm}(Q_t)^{-1/2} \tag{67}$$

$$Q_t = (1-\alpha-\beta)S + \alpha\epsilon_{t-1}\epsilon_{t-1}^T + \beta Q_{t-1} \tag{68}$$

The correlation part of this model is presented as in [Aielli, 2013], in the original representation in [Engle, 2000], equation (68) used matrix coefficients. Let E be an $m \times m$ matrix with every entry being 1, then

$$Q_t = (E - A - B) \circ S + A \circ \epsilon_{t-1}\epsilon_{t-1}^T + B \circ Q_{t-1} \tag{69}$$

where $\circ$ denotes the Hadamard product (elementwise multiplication). Here A and B have to be positive semi-definite matrices and $E - A - B$ has to be positive definite.
The general idea here is to let the conditional correlation matrix follow a univariate GARCH(1,1) processes given by equation (68). Equation (67) ensures that R_t is a correlation matrix, i.e. $\text{diag}(R_t) = (1, \ldots, 1)$. Further, if $\alpha = \beta = 0$, then the CCC-GARCH model is obtained.
It is possible to model R_t differently, the only necessary condition here is that R_t is a correlation matrix. For example, [Engle, 2000] suggests an exponential smoothed empirical correlation matrix defined by

$$\rho_{ij,t} = \frac{\sum_{s=1}^{t-1} \lambda^s X_{i,t-s} X_{j,t-s}}{\sqrt{\sum_{s=1}^{t-1} \left(\lambda^s X_{i,t-s}^2\right) \sum_{s=1}^{t-1} \left(\lambda^s X_{j,t-s}^2\right)}} = [R_t]_{ij} \tag{70}$$

but in the end decides for the formulation as a GARCH process.

4.3.2 Corrected DCC-GARCH model

In the DCC-GARCH model, it is assumed that S is the unconditional correlation matrix, so it must hold that $S = \mathbb{E}[\epsilon_t \epsilon_t^T] = \mathbb{E}[R_t] = \mathbb{E}[\text{diagm}(Q_t)^{-1/2} \; Q_t \; \text{diagm}(Q_t)^{-1/2}]$.

Theorem 14. *Assume $\mathbb{E}[Q_t]$ and $\mathbb{E}[\epsilon_t \epsilon_t^T]$ are constant in t, then it holds for a DCC-GARCH model that*

$$S = \frac{1-\beta}{1-\alpha-\beta}\mathbb{E}[diagm(Q_t)^{1/2}\epsilon_t\epsilon_t^T diagm(Q_t)^{1/2}] - \frac{\alpha}{1-\alpha-\beta}\mathbb{E}[\epsilon_t\epsilon_t^T] \tag{71}$$

Proof. Taking the expectation of equation (68) and rewriting it yields

$$S = \frac{1-\beta}{1-\alpha-\beta}\mathbb{E}[Q_t] - \frac{\alpha}{1-\alpha-\beta}\mathbb{E}[\epsilon_t\epsilon_t^T]$$

From equation (4.3.4) follows

$$\mathbb{E}[Q_t] = \mathbb{E}[\text{diagm}(Q_t)^{1/2}\ R_t\ \text{diagm}(Q_t)^{1/2}] = \mathbb{E}[\text{diagm}(Q_t)^{1/2}\ \epsilon_t\epsilon_t^T\ \text{diagm}(Q_t)^{1/2}]$$

which proves the result above. [Aielli, 2013] □

Theorem 14 shows, that S being the unconditional correlation matrix is not generally true. This leads to inconsistency for the estimators provided in [Engle, 2000]. In [Aielli, 2013] instead, equation (68) is suggested to be

$$Q_t = (1-\alpha-\beta)S + \alpha\text{diagm}(Q_t)^{1/2}\epsilon_{t-1}\epsilon_{t-1}^T\text{diagm}(Q_t)^{1/2} + \beta Q_{t-1} \tag{72}$$

which yields

$$S = \mathbb{E}[\text{diagm}(Q_t)^{1/2}\epsilon_{t-1}\epsilon_{t-1}^T\text{diagm}(Q_t)^{1/2}]$$

This model is called the cDCC-GARCH model, for stationarity and estimation this model will be used, as results for stationarity of the DCC-GARCH model by [Engle, 2000] are non-explicit, and the estimation is very similar.

[Francq and Zakoian, 2019] points out that in equation (72) both sides can be multiplied by arbitrary positive definite matrices and still yields the same result for R_t, therefore the choice of S is not unique. Here S is chosen to be a correlation matrix, i.e. $\text{diag}(S) = (1, \dots, 1)$.

4.3.3 Stationarity of cDCC-GARCH

Stationarity of the cDCC-GARCH model is discussed in [Aielli, 2013]. For the stationarity analysis there they generalize equation (72) to be a BEKK-GARCH model. In the model here these conditions simplify to a not very surprising result.

Theorem 15. *Consider the following*

(A1) The law of η_t has absolutely continuous density, which is positive in a neighborhood of the origin

(A2) S is positive definite

(A3) The model fulfills the necessary strict stationarity conditions for a CCC-GARCH model, i.e. suppose $\alpha = \beta = 0$, then the condition $\gamma(A) < 0$ in Theorem 10 has to hold.

(A4) The model fulfills the necessary weakly stationarity conditions for a CCC-GARCH model

given in equation (58) *in Theorem 11, i.e.*

$$\rho\Big(\sum_{i=1}^{q} A_i + \sum_{i=1}^{p} B_i\Big) < 1$$

where ρ denotes the spectral radius.

(A5) $\mathbb{E}[X_i t^2] < \infty \ \forall i = 1, \dots, m.$

(i) Then if A1-A3 hold, then vech(R_t) and ϵ_t are strictly and second order stationary. In this case, the long-term correlation is

$$S = \mathbb{E}[diagm(Q_t)^{1/2} \epsilon_{t-1} \epsilon_{t-1}^T diagm(Q_t)^{1/2}] \tag{73}$$

(ii) Then if A1-A4 hold, then additionally vech(H_t) and X_t are strictly stationary.

(iii) Then if A1-A5 hold, then X_t is also second order stationary.

Remark. In [Aielli, 2013], the condition $\alpha+\beta < 1$ is also listed, which is here already assumed in the model definition. This ensures second order stationarity for the GARCH(1,1) processes that model the correlation matrix.

4.3.4 Estimation of cDCC-GARCH

Similar to Section 4.2.3, a normal quasi-likelihood can be defined. Again, the same starting values can be used, adding one for Q_0, for example I_m. The parameter vector enlarges to $\theta = (\phi, \psi)$, where $\phi = (\omega, \text{vec}(A_1), \dots, \text{vec}(A_q), \text{vec}(B_1), \dots, \text{vec}(B_p))$ are the parameters known from the CCC-GARCH model that do not model correlation and $\psi = (\alpha, \beta, \text{vec}(S)) \in \Psi \subset (0,1)^2 \times (-1,1)^{m(m-1)/2}$ is the vector of new parameters specifying the dynamic correlation.

$$L_n(\theta) = L_n(\theta; X_1, \dots, X_n) = \prod_{t=1}^{n} \frac{1}{(2\pi)^{m/2} \mid \hat{H}_t \mid^{1/2}} \exp\Big(-\frac{1}{2} X_t^T \hat{H}_t^{-1} X_t \Big) \tag{74}$$

where $\hat{H}_t, \hat{R}_t$ and $\hat{Q}_t$ are recursively defined for $t \geq 1$ by

$$\hat{H}_t = \hat{D}_t \hat{R}_t \hat{D}_t, \qquad \hat{D}_t = \text{diag}((\sqrt{\hat{h}_{11,t}}, \dots, \sqrt{\hat{h}_{mm,t}}))$$

$$\hat{h}_t = (\sqrt{\hat{h}_{11,t}}, \dots, \sqrt{\hat{h}_{mm,t}})^T = \omega + \sum_{i=1}^{q} A_i X_t^2 + \sum_{i=1}^{p} B_i \hat{h}_{t-i}$$

$$\hat{R}_t = \text{diagm}(\hat{Q}_t)^{-1/2} \ \hat{Q}_t \ \text{diagm}(\hat{Q}_t)^{-1/2}$$

$$\hat{Q}_t = (1 - \alpha - \beta) S + \alpha \epsilon_{t-1} \epsilon_{t-1}^T + \beta \hat{Q}_{t-1}$$

Similar to the CCC-GARCH model, a multi-step estimator can be introduced to shorten computation time. For this the quasi-log-likelihood function is used, which is given, up to additive constants, by

$$l_n(\theta) = -\frac{1}{2}\sum_{t=1}^{n}\left(\log \mid \hat{H}_t \mid + X_t^T \hat{H}_t^{-1} X_t\right) \tag{75}$$

Using $\hat{H}_t = \hat{D}_t \hat{R}_t \hat{D}_t$ and $\hat{D}^{-1} X_t = \epsilon_t$ this can be rewritten as

$$l_n(\theta) = -\frac{1}{2}\sum_{t=1}^{n}\left(2\log \mid \hat{D}_t \mid + \log \mid \hat{R}_t \mid + \hat{\epsilon}_t^T \hat{R}_t^{-1} \hat{\epsilon}_t\right)$$

In [Engle, 2000], this is split into two parts

$$l_n(\theta) = l_n(\phi, \psi) = l_{n,v}(\phi) + l_{n,c}(\phi, \psi)$$

First, $l_{n,v}$ is maximized w.r.t ϕ, which yields $\hat{\phi}$. The second part is done as in [Aielli, 2013] to ensure consistency for S.

Set

$$\hat{S}_{\phi,\alpha,\beta} = \frac{1}{n}\sum_{t=1}^{n} \text{diagm}(\hat{Q}_t)^{1/2} \hat{\epsilon}_{t-1} \hat{\epsilon}_{t-1}^T \text{diagm}(\hat{Q}_t)^{1/2}$$

where $\text{diag}(\hat{Q}_t) = (\hat{q}_{11,t}, \ldots, \hat{q}_{mm,t})$ is recursively defined for $t \geq 1$ by

$$\hat{q}_{ii,t} = (1 - \alpha - \beta) + \alpha \hat{\epsilon}_{i,t}^2 \hat{q}_{ii,t-1} + \beta \hat{q}_{ii,t-1}$$

Now ψ is obtained through a constrained optimization.

$$(\hat{\alpha}, \hat{\beta}) = \underset{\alpha,\beta \geq 0,\ \alpha+\beta<1}{\arg\max}\ l_n(\hat{\phi}, \alpha, \beta, S_{\hat{\phi},\alpha,\beta})$$

The constraint $\alpha + \beta < 1$ ensures stationarity of the GARCH processes that model Q_t. Lastly set $\hat{S} = S_{\hat{\phi},\hat{\alpha},\hat{\beta}}$. Thus, we never optimize over more parameters than given in ϕ at once.

5 Overview of the crypto market

This Section aims to give a short introduction to what cryptoassets are and what they are used for. For this Bitcoin will be used as a primary example. Furhtermore the Bitcoin data will be shortly discussed herein and some general figures about the size of the cryptoasset market will be given. The precise technical functionality of cryptocurrencies is omitted here, as just the financial aspects are of concern for this paper. All data used on cryptoasset prices is taken from the website [coinmarketcap.com, 2022] and daily closing prices are considered. The current value of the Bitcoin refers to the value at the 1st of September 2022, which is 20,127 USD.

5.1 Bitcoin

Bitcoin utilizes the Blockchain technology. This Blockchain contains all transactions ever made inside the Bitcoin network. In order to keep this chain up to date, so-called miners regularly add new blocks to it. Before a miner can add a block, they first must solve a puzzle, and if they are able to do so they are awarded some amount of Bitcoins. Furthermore, other miners check whether the block is correct. In this system, which is called proof of work, the miners check each other, and the system can just be corrupted if the computing power of the corruptors outweighs that of the uncorrupted miners.
In this chain the difficulty of the puzzles that need to be solved is adjusted in order to meet the current computing power of the network, i.e. the amount of Bitcoin given out to miners over a certain time period does not depend on the total computing power of the network. Also, the maximum amount of Bitcoin that can be issued has been set at 21 million, but the coins issued to a miner per block are halved in certain timeframes in order to push the date, on which this limit is reached, even further into the future. Currently, about 19 million coins are in circulation.

Important properties of Bitcoin

There are several important differences between Bitcoin and fiat money like the Euro or Dollar. Here the most important are outlined:

- **Divisible:** As there is no physical money, Bitcoin is in theory arbitrarily divisible. In practice, the smallest unit of Bitcoin is called a Satoshi. One Bitcoin equals 100 million Satoshis, which puts the current price of a Satoshi at about 0.0002 USD.
- **Decentralized:** Behind fiat currencies, there is always an organisation (e.g. government), but in the case of Bitcoin, transactions are saved in the blockchain and checked by the miners. Since the miners are spread across the globe, there is no single institution which holds the power in the Bitcoin network.

- **Anonymous:** While all transactions in the crypto network are visible and can be looked up at e.g. blockchain.com, the names of the people participating in these transactions are not shown.

These differences are associated with advantages and disadvantages alike. Decentralization prevents a direct regulation of the network by one central authority, but therefore it is also safe from having a corrupt authority and is not tied to the fate of such an authority, e.g., it is not tied to a single country that might have huge inflation because of a struggling economy. However, having a central authority is often also perceived as a safety behind an asset. The US government for example backs the USD and has reserves to support it, should need be. In the case of Bitcoin, there is no such thing, if no one wants to buy them they become worthless. Warren Buffet famously said that, if he could have all Bitcoins in existence for 25 Dollars, he would not buy them, because their only use would be to sell them back to the people he just bought them from. These lines of thought can often bring doubt into the minds of investors, leading to price fluctuations. In addition, there is of course the concern that somebody could possess 51% of all the computing power in the network, but this is highly unlikely. In the blog article [BRAIINS, 2021], an approximation is made on how much money the necessary hardware would cost, resulting in a figure of over five billion USD. While maybe theoretically payable, liquidity of the market for these products prevents this possibility. However one needs to keep in mind that mining always generates a certain payoff independent of the number of miners, which in return means that the number of profitable miners is dependent on the price of Bitcoin. This could lead to significantly less safety during periods of low prices.
Apart from the economic concerns, decentralization achieved through proof of work also has environmental concerns. According to the website [DIGICONOMIST, 2022], the Bitcoin network used up to 200 terawatts hours per year. The data from [WORLDDATAINFO, 2022] suggests, that this is more than three times the amount of power consumed in Austria. While power usage has dropped significantly due Bitcoins decline in value, it is still huge. In 2012, an alternative for proof of work was introduced in proof of stake in [King and Nadal, 2012]. This concept drastically decreases the environmental concerns and is already applied in Ethereum, which switched to it from proof of work.
While anonymity sounds good, it leads to some serious issues. Being nearly untraceable, Bitcoin is a perfect breeding ground for illegal activities. This in turn increases the desire to regulate it by centralized authorities.

5.2 History of cryptoassets

This subsection aims to review the early history of cryptoassets and continues with the history of Bitcoin up to present day. In more recent years, there are too many different assets on the market, leading to the history of the whole market becoming too wide, but [Ciaian et al.,

2018] found at least strong short-term dependencies between the assets, suggesting that major events probably have a similar impact on most of the popular cryptoassets.

Until 2007

The first cryptoasset was developed as early as 1990 by the company DigiCash. This cryptoasset was called eCash and lasted till 1998, when DigiCash went bankrupt. Several other cryptoassets such as E-Gold, Bit Gold (not to be confused with Bitcoin Gold) and B-Money were invented around that time. These predecessors of the modern cryptoassets had some of the desirable properties that make modern cryptoassets so popular, but ultimately, none of them were a success. Nevertheless, they laid the technological foundation for the modern cryptoasset market. It was not until 3 years later that the first cryptoasset which is still on the market was released. [Reiff, 2022]

2008 - 2010

In 2008, the paper [Nakamoto, 2008] proposed a new way of using the Blockchain technology. Satoshi Nakamoto is a pseudonym, and until today it isnot known who stands behind it. However, one year later, the same person or group released an implementation of their cryptoasset, which became Bitcoin. At first, Bitcoin was not used as actual cash, but more of a hobby for computer enthusiasts. It took until May 2010 before the first purchase with Bitcoins was made. Here, 10,000 of them were used to purchase two pizzas. With the current value, this would be just over 200 million USD.

2011

After that, Bitcoin exchanges emerged, where Bitcoins could be exchanged directly for fiat currencies and their price rose rapidly, reaching parity with the dollar in February [Scaillet et al., 2017]. After this, more and more vendors accepted payment in Bitcoins. Unfortunatly, this did not only include legal ones. The article [Chen, 2011] popularized a criminal website called Silk Road. This website used Bitcoin as a payment system for illegal substances. Other websites such as Wikileaks also started to accept Bitcoin in 2011, since normal donations were restricted by the US government. Of course there were plenty legal ones as well, but the illegal vendors gave Bitcoin a real use case.

Due to the code for Bitcoin being open sourced, it did not take long until other cryptoassets, called altcoins, entered the market. One example of an altcoin, that is still popular today, was introduced in 2011 as Litecoin.

2013 - 2014

The 29th of April 2013 is the begining of available price data at coinmarketcap.com. So far, Bitcoin was trending upwards and saw a steep rise in the end of 2013, where the price exceeded 1,000 USD for the first time. In February 2014, the largest crypto exchange called Mt.Gox was hacked. In this process, between 500,000 and 1,000,000 Bitcoins were lost, leading to the collapse of the exchange. This illustrated the danger of cryptoassets to investors and organizations alike. Leading not only to legal obstacles and an immediate downwards trend, but also to a long-term drop in value that continued until 2015. However, the database of Mt.Gox was leaked, giving unique data for a time period between February 2011 and November 2013. This made it possible to track the Bitcoin IDs in that time interval and sometimes even showed the location of the traders. The unique data set was used in [Scaillet et al., 2017] to do high frequency jump analysis for the Bitcoin returns. This paper is not included in the overview above, since its findings are sadly not reproducible for current returns.

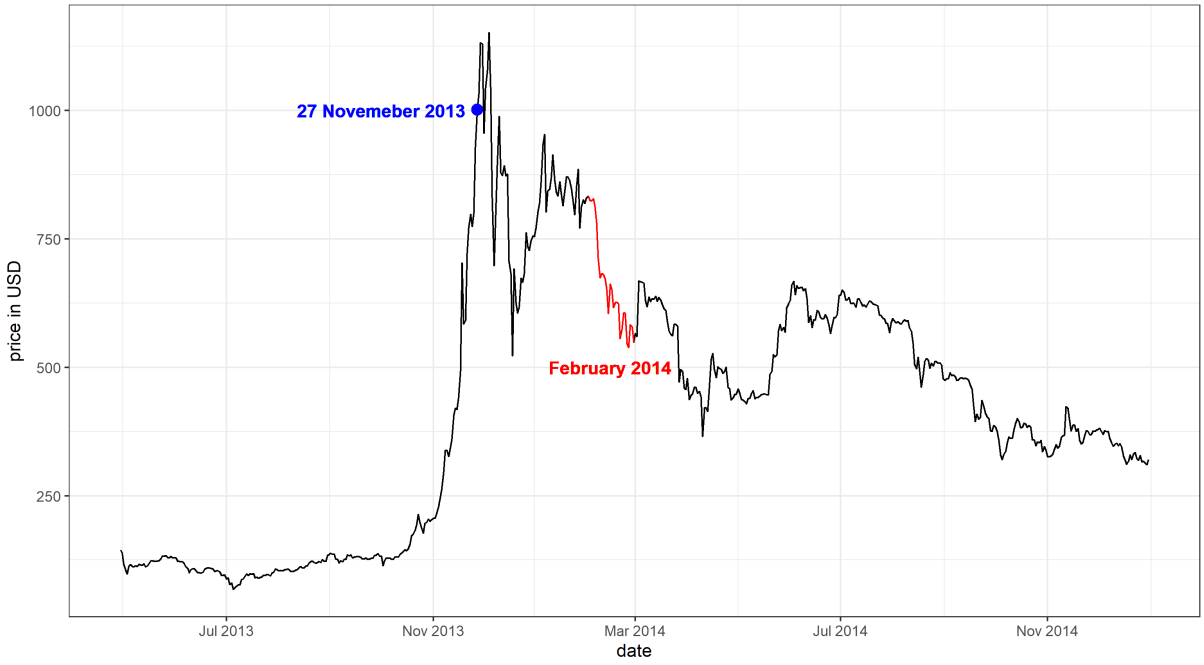


Figure 14: Bitcoin price history from 29.04.2013 until 31.12.2014. The first increase over 1,000 USD is marked in blue and the month with the Mt.Gox incident is highlighted red.

2015 - 2016

Still on the downtrend in the beginning of 2015, Bitcoin reached its minimum on the 14.01.2015 at a price of 178 USD. Then, during the rest of these two years, the price slowly crept back up to where it already had been in 2013. In 2015, Ethereum, currently the second largest cryptoasset, was created.

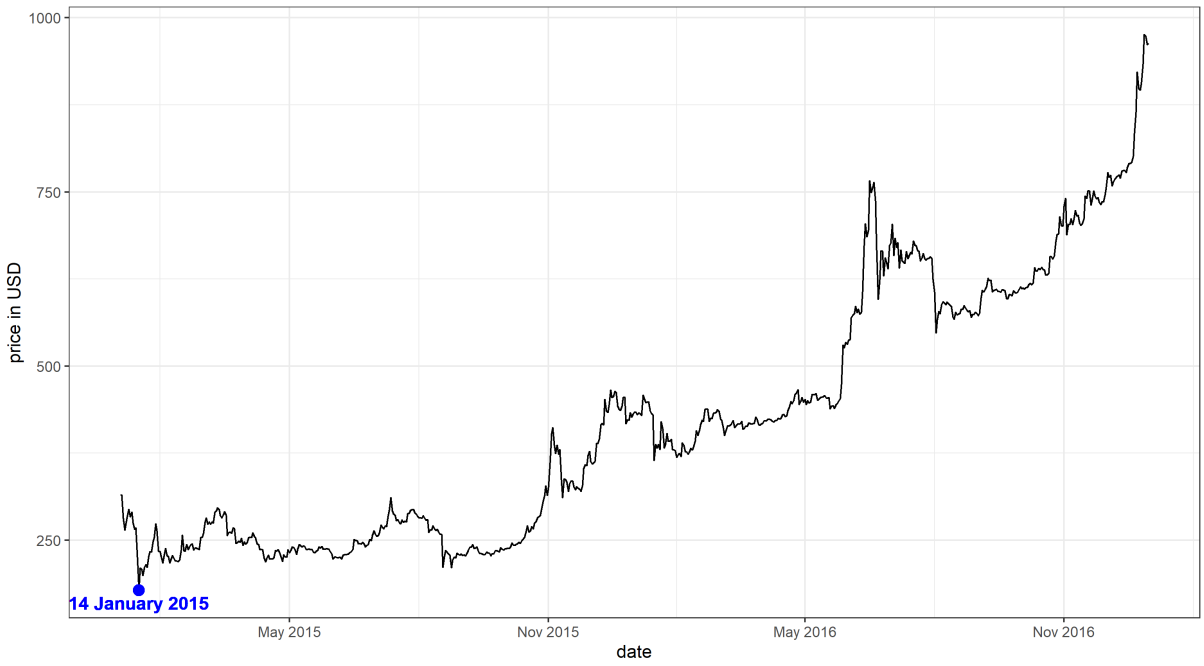


Figure 15: Bitcoin price history from 2015 - 2016. The lowest point 178 USD is marked in blue.

2017 - 2018

At the end of 2017, the Bitcoin price exploded for the second time. This rapid increase was attributed to the FOMO (fear of missing out) concept, i.e. people thinking it would just continue to increase and that if they do not invest, they would miss out on the easy money that was to be made there. The fastest growth occurred during the eight day stretch from 29.11.2017 to 07.12.2017 (marked red in Figure 16), where the price shot up over 8,000 USD. The price ultimately reached its climax at the 07.12.2017, where one Bitcoin was worth 19.497 USD; this record held for almost three years. The reversal in price occured due to the increasing believe that these high prices were a bubble. Since Bitcoin is backed mainly by the believe of its investors, the prophecy of a bubble made itself true in the sharp decline experienced shortly after the rise, leaving Bitcoin at around 4,000 USD at the end of 2018.

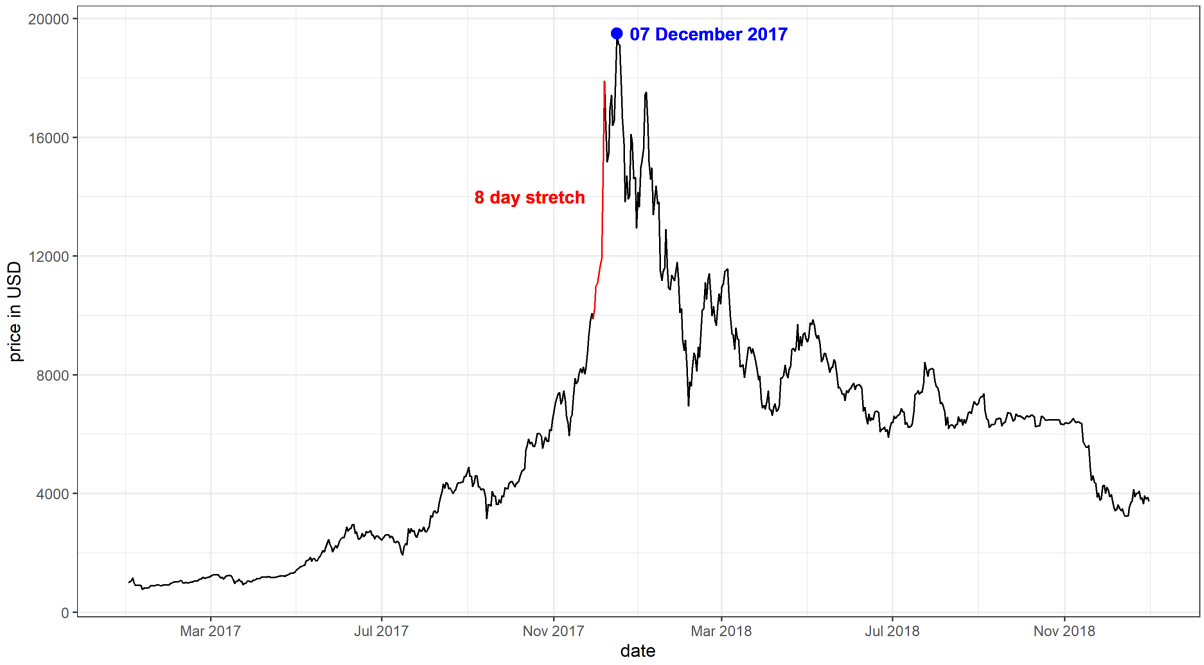


Figure 16: Bitcoin price history from 2017 - 2018. Shown in red is the price increase by more than 8,000 USD over the eight days from 29.11.2017 to 07.12.2017 and in blue the maximaum price of 19.497 USD.

2019 - June 2020

In the first half of 2019, the Bitcoin price doubled again to remain relatively calm until the end of February 2020. At 11.03.2020, the WHO declared COVID-19 a pandemic, which lead to the Bitcoin price falling by more than 37% within a single day (highlighted red in Figure 17). Until June, the price came back to its level previous to the drop.

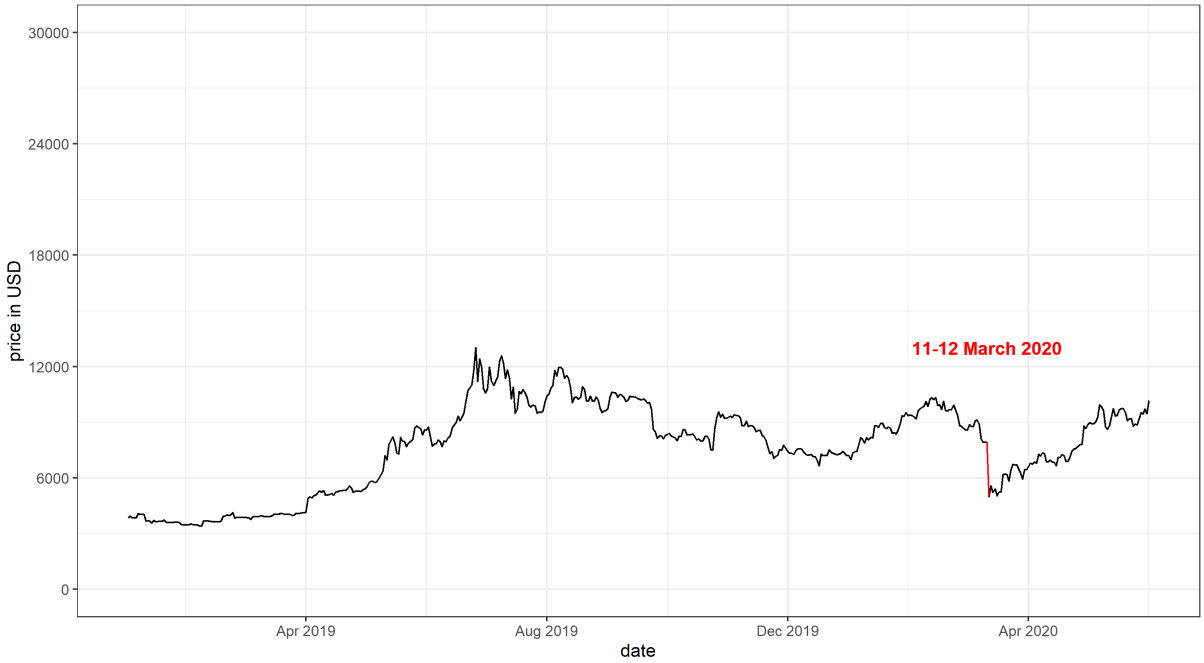


Figure 17: Bitcoin price history from 2019 - June 2020. SHown in red is the plummeting of the price after COVID-19 was announced a pandemic.

June 2020 - September 2022

The second wave of COVID-19 had a very different influence on the Bitcoin price from the first one. This time, it caused the Bitcoin price to rise. On the 8.11.2021, the record from 2017 was finally broken (marked green in Figure 18), but the price continued to skyrocket to over 60,000 USD, thereby increasing sixfold. The first peak was reached at 13.04.2021, where the price was 63,503 USD (marked red in Figure 18). After that, the price of Bitcoin more than halved as it went below 30,000 USD again, before reaching its all time high on 08.11.2021 with one Bitcoin equalling 67,566 USD. On this day, the market capitalization also reached its maximum at 1,274 billion USD. But it did not last for long, and shortly after, the Bitcoin price droped again and currently stands around 20,000 USD, less than a third of what it was worth once.

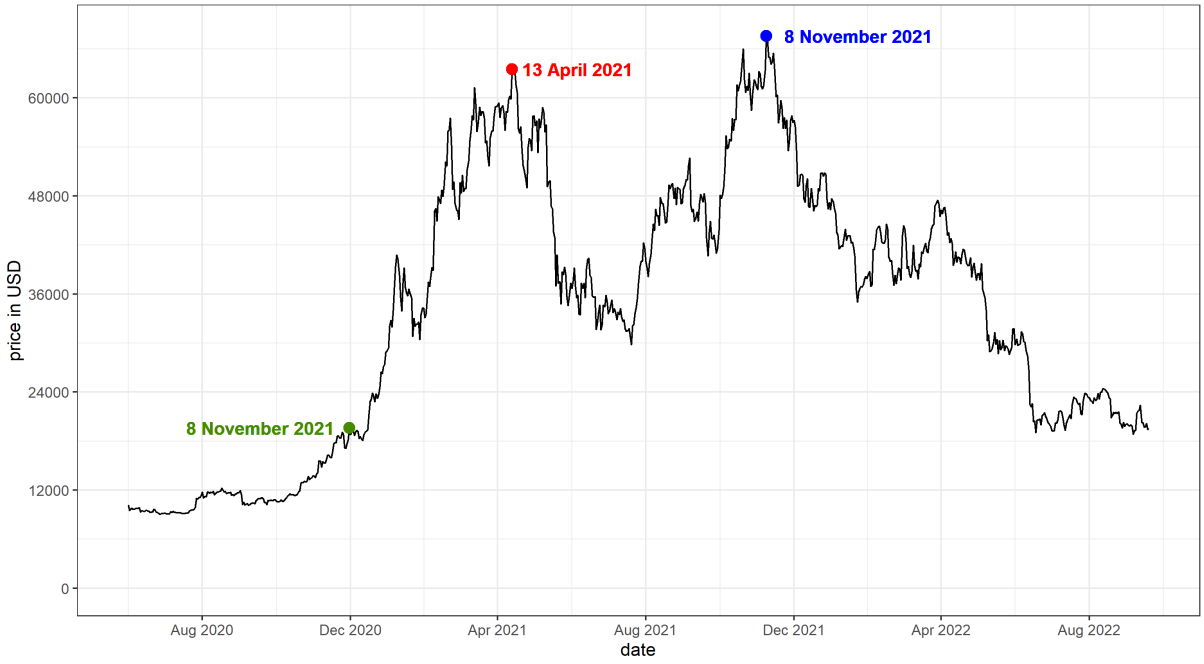


Figure 18: Bitcoin price history from June 2020 - September 2022. The three dots mark the breaking point of the mark set in 2017 (green), the first COVID-19 peak at 63,503 USD (red) and the all-time high at 67,566 USD (blue) in that order

Complete price graph

Figure 19: Complete chart of the Bitcoin price

The complete graph often misleads to assuming that the early days of Bitcoin saw fewer radical shifts than the current ones. However, it depicts, that even though it truly is a roller-coaster, it has, so far, an upwards trend.
This subsection shows, that Bitcoin is a true nightmare for any risk averse investor, since hardly any time passes until it doubles or halves.

5.3 Descriptive statistics of the cryptoasset market

Currently, coinmarketcap.com lists over 9,000 different cryptoassets, with almost 2,500 having more than 10,000 USD in market capitalization. In this Section, a few statistics of the top ten cryptoassets will be provided.
The data in the following table refers to the 01.09.2022.

Name	**Price in USD**	**market cap in million USD**	**supply**	**capped**	**starting year**
Bitcoin	20,127	385,208	19.13 million	21 million	2009
Ethereum	1,586	193,834	122.2 million	not capped	2015
Tether	1	67551	67.55 billion	not capped	2015
USD Coin	1	52,026	52.03 billion	not capped	2018
BNB	278	44,919	161.3 million	200 million	2017
Binance USD	1	19,427	19.41 billion	not capped	2019
XRP	0.33	16,542	49.64 billion	100 billion	2012
Cardano	0.45	15,638	34.18 billion	45 billion	2015
Solana	31	11,046	349.7 million	not capped	2019
Dogecoin	0.062	8,275	132.7 billion	not capped	2013

Notice that Tether, USD Coin and Binance coin are at parity with the USD. This is not a coincidence. These three assets are designed to stay at a constant price that is maintained by fiat reserves. They are actually better referred to as cryptocurrencies, as they do not have the characteristics of an asset, since they are just as volatile as the USD. While price stability can be an advantage, it also comes at the cost of being centralized again. However, for modelling purposes, these currencies are not of interest and will be omitted form here on.

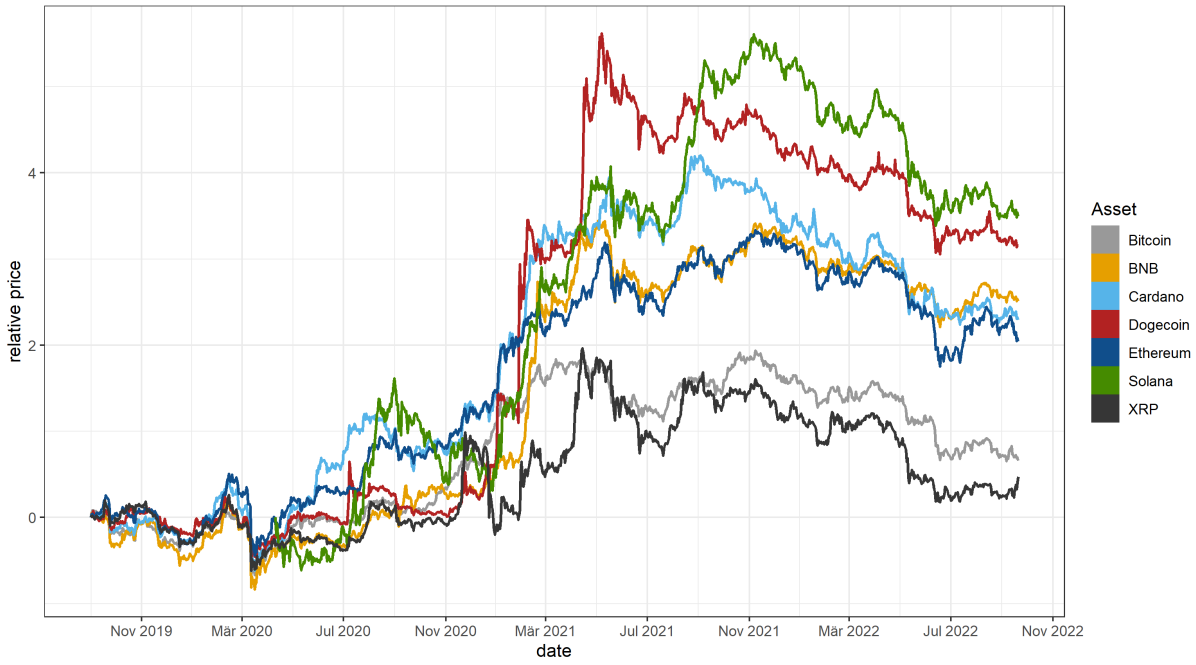


Figure 20: A normalized logarithmic price is displayed to highlight the growth rates of the assets from September 2019 to September 2022

Figure 20 aims to compare the seven remaining assets. Therefore, the data is transformed in the following way. Let $price(t)$ be the price of one of the cryptoassets t days after 01.09.2019, then the relative price is calculated by

$$\text{relative price}(t) = \log\left(\frac{\text{price}(t)}{\text{price}(0)}\right)$$

First, the price is normalized and then, the logarithm is taken. Therefore, every cryptoasset (but one) starts with the value 0 at 01.09.2019. For Solana, this transformation is done at the 10.04.2020, as this is the starting point for the data. This is why, the comparison with Solana lacks a little, but this way, the normalizing value is prior to the COVID-19 price explosion, which gives a better indicator of how this affected each of the cryptoassets. Furthermore, even when normalizing at the starting day for Solana, a new cryptoasset has more potential for multiplying itself, which makes it unsurprising to see Solana at the top of the chart from September 2021. Dogecoin toped the chart before that due to getting a lot of attention directed towards it by Elon Musk. Be aware that the general upwards trend all assets here have in common is biased by the choice of the assets. Since they are chosen due to their high market capitalization, these are assets that currently outperform the market and are therefore not indicative of the trend of the crypto market.

For modelling, not the actual price is used, but the returns of the assets or the log-returns. To get the log-returns, the transformation, that was already mentioned in Remark 3.1, is used.

$$\text{return}(t) = \frac{\text{price}(t)}{\text{price}(t-1)} - 1, \qquad \log-\text{return}(t) = \log\left(\frac{\text{price}(t)}{\text{price}(t-1)}\right)$$

The following table has statistics for the returns of the seven cryptoassets for the time period from 01.09.2019 to 01.09.2022.

Name	**Average size**	**Maximum**	**Minimum**	**Median**
Bitcoin	2.57%	18.7%	-37.2%	0.084%
Ethereum	3.50%	25.9%	-42.3%	0.267%
BNB	3.59%	69.8%	-41.9%	0.127%
XRP	3.73%	56.0%	-42.3%	-0.09%
Cardano	4.10%	32.2%	-39.6%	0.162%
Solana	5.99%	115.5%	-37.2%	0.116%
Dogecoin	4.47%	355.6%	-40.2%	-0.037%

In the table above, the first column gives the average size of the absolute returns, meaning the average daily change in these three years. This underlines the argument, that these assets are extremely volatile; Bitcoin even seems to be one of the less volatile among them. Also, the two assets leading the relative price chart both had days during the last three years on which their price more than doubled. Further, one can observe that all of the assets have a minimum return close to 40%, meaning that every single one of the top seven cryptoassets fell by more than a third of their value within a single day. This alone indicates, that necessary reserves for investments into the crypto market need to be relatively high. Also to add some contrast, the S&P 500 data included in the R-package [Hubbard, 2022], which ranges from October 2010 to October 2020, with 92 missing days, was analyzed. The data given there suggests, that the maximum return in these ten years was 9.38% and the minimum was -9.51%, though one has to keep in mind that the S&P 500 is a weighted average and therefore less volatile.

6 Modelling

For the modelling part, the time horizon from Figure 20 is used,. As the market is still relatively young, it might be influenced by structural breaks as mentioned in [Mensi et al., 2019]. Therefore, a shorter time period (three years) is chosen, which focuses on the COVID-19 period. In order to find a good observation window one could compare in- and out-of-sample performances for different time horizons, but this is not introduced and done here. Furthermore the log-returns are used for modelling, as these tend to be weakly stationary. Below, in Figure 21, the log-returns for Bitcoin are displayed. To get a better impression of whether there is a pattern or not, the red line is the smoothed absolute values of the log-returns.

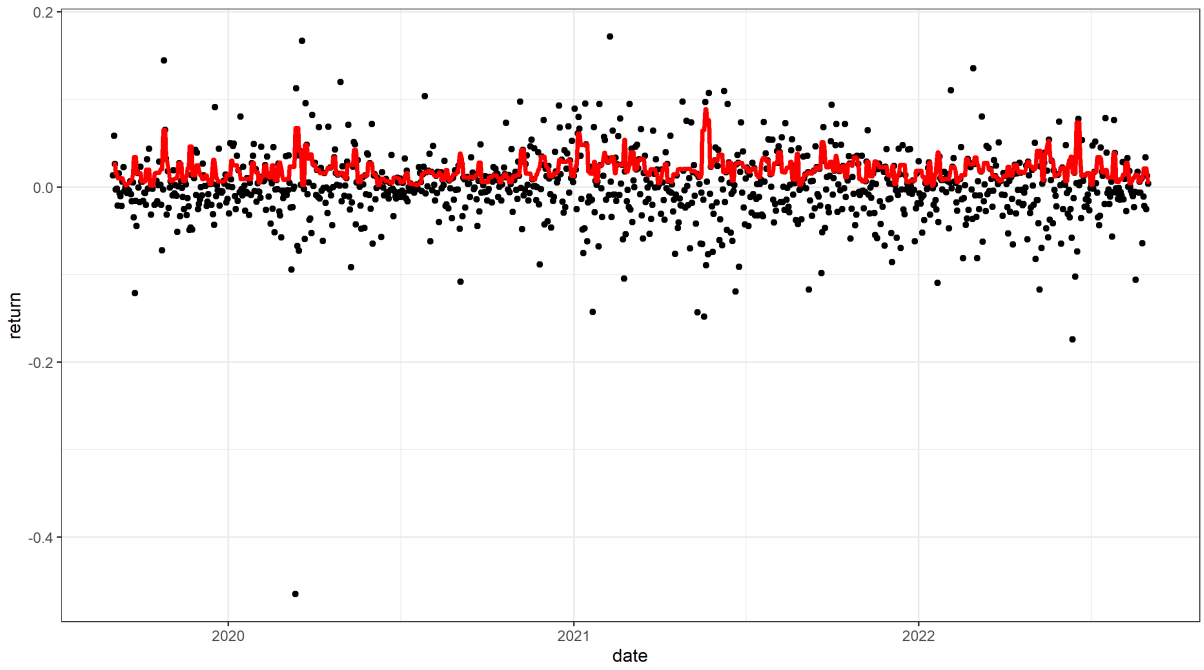


Figure 21: The log-returns of Bitcoin between the 01.09.2019 and 01.09.2022. The red line is the smoothed absolute values of the log-returns.

6.1 Univariate GARCH modelling

The univariate models are only considered for the Bitcoin data. However, before a model can be fitted to the data, some preliminaries have to be checked. These were introduced in the Chapters 3.3.4 and 3.3.5. Then, several models are checked against each other.

6.1.1 Preliminaries

Identifying the GARCH orders

Following the steps given in chapter 3.3.4, a 10×10 corner matrix is evaluated. In the following table, only the exponent is given such that the first significant digit is in front of the comma.

i \ j	1	2	3	4	5	6	7	8	9	10
1	-2	-3	-3	-2	-3	-4	-1	-2	-3	-3
2	-2	-5	-4	-3	-5	-4	-2	-4	-4	-4
3	-3	-4	-4	-4	-4	-4	-3	-5	-5	-6
4	-2	-3	-4	-6	-6	-6	-4	-5	-6	-6
5	-3	-5	-4	-6	-9	-7	-5	-6	-7	-8
6	-3	-4	-4	-6	-7	-7	-6	-8	-8	-9
7	-1	-3	-3	-4	-5	-6	-7	-8	-9	-10
8	-3	-4	-5	-5	-6	-8	-8	-9	-11	-11
9	-3	-5	-5	-5	-7	-8	-11	-11	-11	-12
10	-3	-4	-6	-6	-8	-0	-11	-12	-13	-13

Contrary to the statement in [De Gooijer and Heuts, 1981], here it is pretty straightforward to visually identify the GARCH orders. Marked above in red, the corner gives a jump of at least one order of magnitude in all but one cases, giving a very good hint at which orders to use. The approach suggested by [Chan, 1999] also yields a matrix containing only O values, confirming that $(p, q) = (1, 1)$ are indeed the suggested orders.

Stationarity- and LM-test

Stationarity is checked with an augmented Dicky-Fuller test available in the R-package [Trapletti et al., 2022], which yields -9.4, thus strongly implying weak stationarity. The test is not covered in this paper, but is very predictable, as its 1% critical value with the observation count present here would be a little more than -3.5 which is far away from the actual result.
For the LM-test, the package [Fisher, 2012] was used, for the hypothesis $\alpha_1 = 0$, the test returns a p value of 0.037, which suggests that ARCH disturbances are present in the data.

6.1.2 Choice of model

There are several ways to compare models, here the AIC (Akaike Information Criterion) will be used.

Definition 6.1. For a model with k parameters, let $\hat{l}$ be the maximum of the log-likelihood function. Then the **Akaike information criterion** is defined by

$$AIC := 2k - 2\hat{l} \tag{76}$$

This is used to prevent overfitting, which would occur, if just the likelihood would be considered. There are several other criteria, that can be considered, with the most popular being the BIC (Bayesian Information Criterion), but in this paper, only the AIC will be considered, since further analysis, that was not featured in this paper, would be necessary to decide between the best models for the AIC and BIC, should they be different.

The six different univariate GARCH models that were introduced in Section 3 will be compared here. For this the following four distributions will be used.

Definition 6.2.
Normal distribution:

$$f(x) := \frac{1}{\sqrt{2\pi}\sigma} \exp\bigg(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2\bigg), \qquad \mu \in \mathbb{R}, \sigma \in \mathbb{R}^+ \tag{77}$$

Student-t distribution:

$$f(x) := \frac{1}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\Gamma(\frac{n+1}{2})\bigg(1+\frac{x^2}{n}\bigg)^{-\frac{n+1}{2}}, \qquad n \in \mathbb{N}^+ \tag{78}$$

Normal inverse gaussian distribution:

$$f(x) := \frac{\alpha\delta K_1(\alpha\sqrt{\delta^2+(x-\mu)^2})}{\pi\sqrt{\delta^2+(x-\mu)^2}} \exp\bigg(\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)\bigg), \qquad \alpha,\beta,\delta,\mu \in \mathbb{R} \tag{79}$$

Generalized hyperbolic distribution:

$$f(x) := \frac{(\sqrt{\alpha^2-\beta^2})^\lambda \delta^{-\lambda}}{\sqrt{2\pi}K_\lambda(\delta\sqrt{\alpha^2-\beta^2})} \exp\bigg(\beta(x-\mu)\bigg) \times \frac{K_{\lambda-\frac{1}{2}}(\alpha\sqrt{\delta^2+(x-\mu)^2})}{(\frac{\sqrt{\delta^2+(x-\mu)^2}}{\alpha})^{\frac{1}{2-\lambda}}}, \quad \alpha,\beta,\gamma,\delta,\mu \in \mathbb{R} \tag{80}$$

With $K_\cdot$ being a modified Bessel function of the second kind.

So all in all, 24 different univariate models are considered for the Bitcoin data. The fitting of these models is done with the R-package [Galanos and Kley, 2022]. There, the AIC is defined slightly different, in the end it is divided by the number of observations, in our case 1,097. The table below lists the AIC value for every model considered.

AIC	GARCH	IGARCH	EGARCH	TGARCH	GJRGARCH	APGARCH
Normal	-3.692054	-3.670079	-3.727049	-3.729607	-3.725948	-3.727865
Student-t	-3.967987	-3.969791	-3.973637	-3.975366	-3.966203	-3.973549
Nig	-3.966720	-3.966014	-3.971416	-3.973364	-3.964996	-3.971543
Hyperbolic	-3.965566	-3.966415	-3.970788	-3.972764	-3.963762	-3.970943

The Student-t distribution has the best fit for all considered models, the same is true for the TGARCH model. The only distribution that clearly does not fit very well is the normal distribution. The heavy tailed ones lose out slightly, because they contain more parameters themself, which is penalized by the AIC. This is in contrast to the results of the study in [Guo, 2022], indicating that the dynamics between 2017 and 2020 differ from those between 2019 and 2022.
The choice of an asymmetric model is not really surprising, as several papers in the literature review already mentioned, that the impact of negative returns on the volatility was greater than that of positive returns, which is exactly what is found here as well. The following table gives the estimated model parameters for the TGARCH(1,1) model. Further the estimation determined that the optimal number of degrees of freedom for the student-t distribution was three.

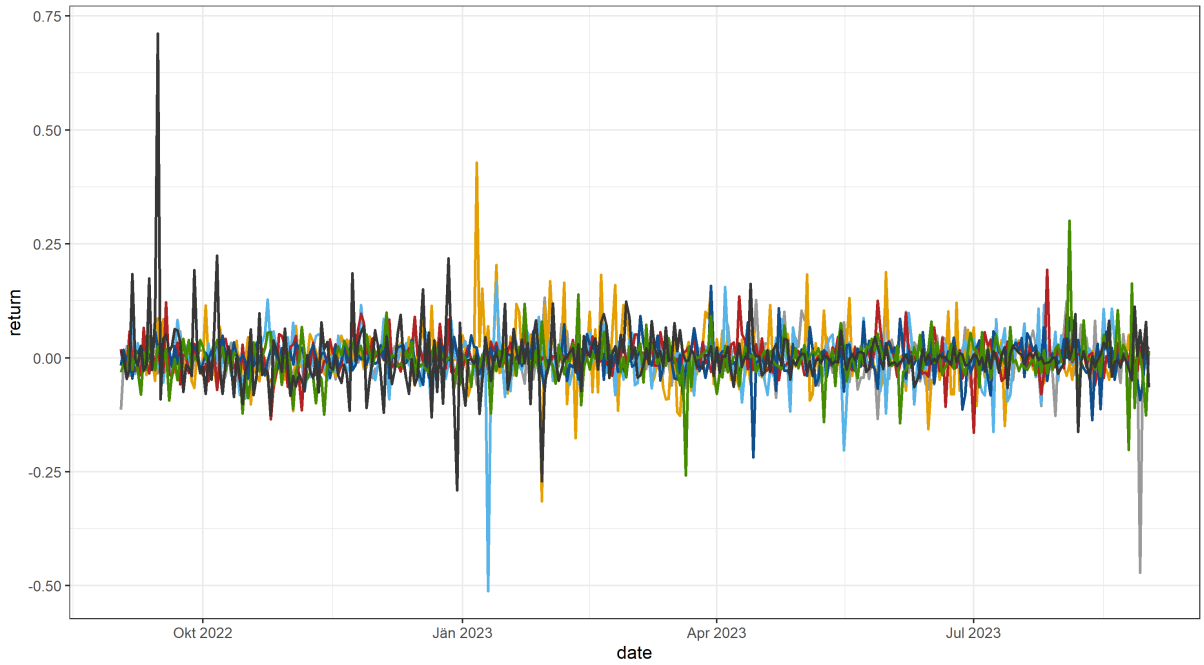
Parameter	Value
ω	0.000675
α_+	0.0648775
α_-	0.0816845
β	0.937365

The large value of β implies very high shock persistence. Furthermore, since $\alpha_+ + \beta > 1$ and $\alpha_- + \beta > 1$, the long-term variance does not exist.

6.1.3 Option prices

In order to determine option prices, the chosen TGARCH model from above is simulated. As starting point, two different choices are made. Firstly, the last point of the data will be chosen, so all prices are given as if the options were sold on the 01.09.2022, secondly, the model will

be given 100 days of burn-in and then the 101st day is used as starting point, this will give prices for an arbitrarily chosen day. Before the actual option prices, seven model paths, their sigma values and Bitcoin prices are displayed in Figure 22, 23 and 24, respectively.



re

Figure 22: Comparison of seven paths of simulated log-returns of the TGARCH(1,1) model

Figure 22 does not visualise each path clearly as most returns are very small and therefore not really possible to distinguish. The more interesting part of this plot is the comparison of the clearly visible peaks to Figures 23 and 24.

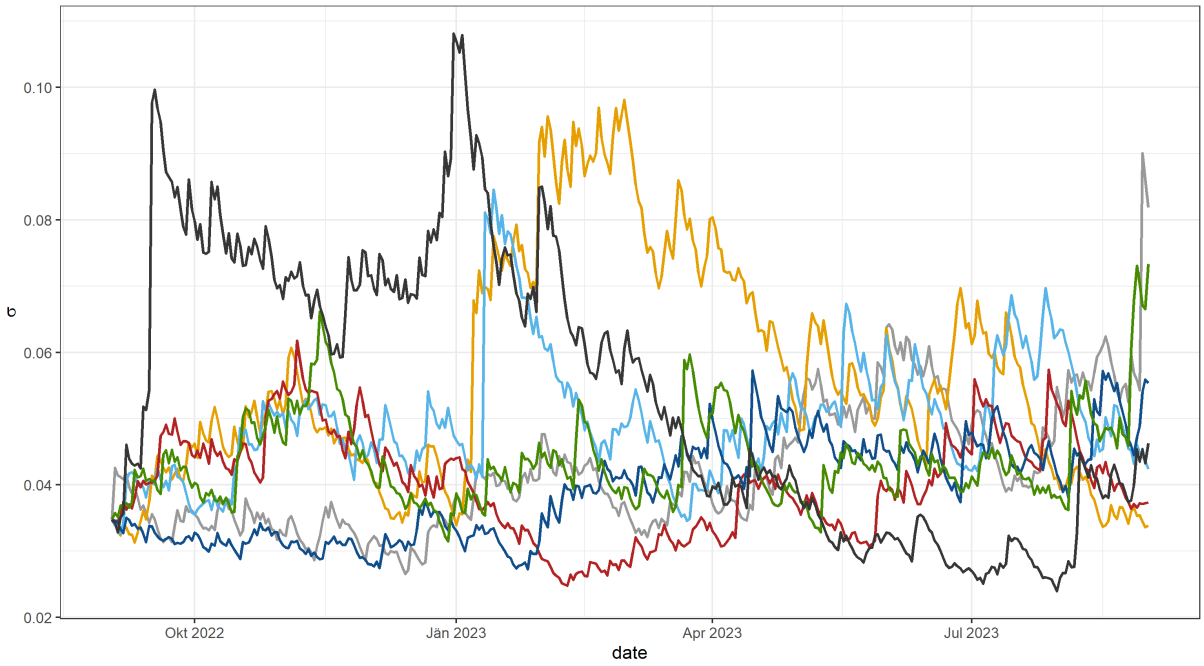


Figure 23: Comparison of the value of σ for the log-return paths in Figure 22

In Figure 23, the paths of the value σ from the log-return paths above is depicted. By comparison to Figure 22, one can see that extreme returns tend to be caused by high volatility, while not all periods of high volatility lead to extreme returns. Furthermore, the persistence of shocks is readily apparent as shocks cause returns to rise in an instant, but it always takes some time to settle again. Also, about half of the paths (blue, light blue and red) are quite calm over the entire period, while the others have at least one interval of high volatility.

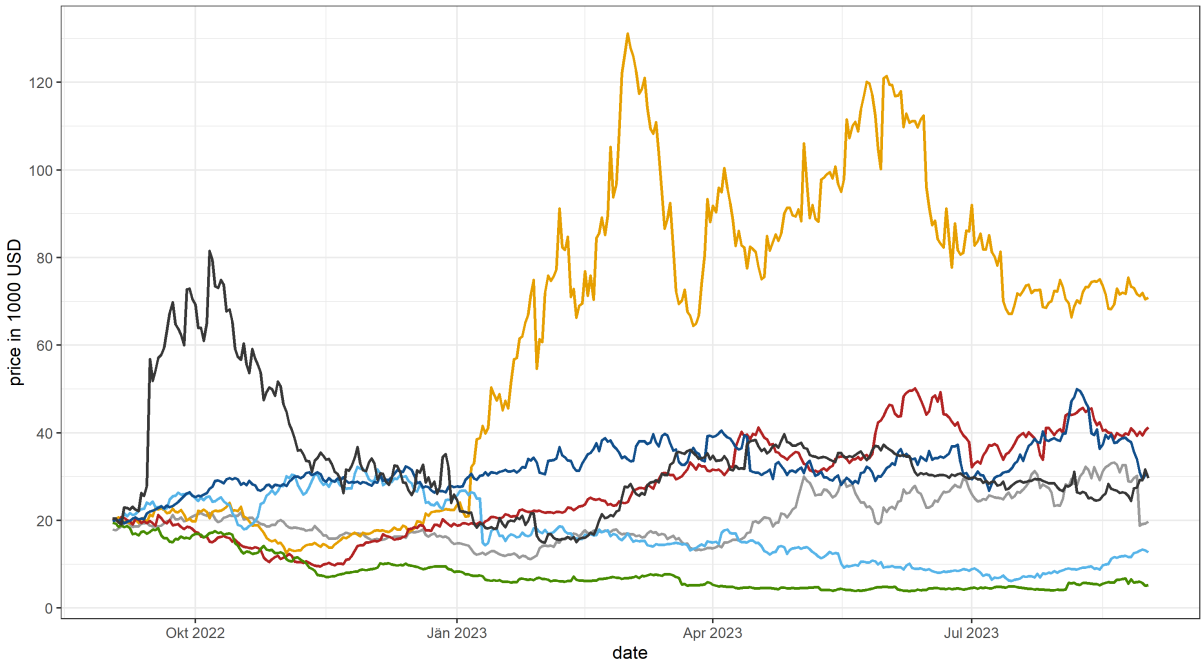


Figure 24: The actual Bitcoin prices that the log-return paths in Figure 22 display. The price is given in 1,000 USD.

Figure 24 shows the price process reflecting the log-returns from Figure 22. The high volatility of the paths becomes clearly visible here and while all but two paths appear as if they were calm, they surely are not. The green path more than halves in price, while the red path more than doubles. Only the grey path ends the year at a similar value as it began with.

One-day options

One-day options should obviously be the cheapest. Figure 25 shows the option price in relation to the strike price for the one-day options. The prices are for European call options and are calculated by simulating 1,000 paths. Strike prices are considered from 15,000 USD to 25,000 USD. Further, concerning the starting point, hardly any difference is visible for the one-day options, and in fact, at no strike price is there a difference greater than 5 USD between them. The blue line uses the starting value as given in the data and the red one uses the arbitrary chosen starting value, created via burn-in as mentioned above. The call with the current price as strike would cost 255.6 USD, which is 1.2% of the price of one Bitcoin. In the 1,000 considered paths, the maximum price after one day is 26,770 USD, while the minimum price is 15,820 USD.

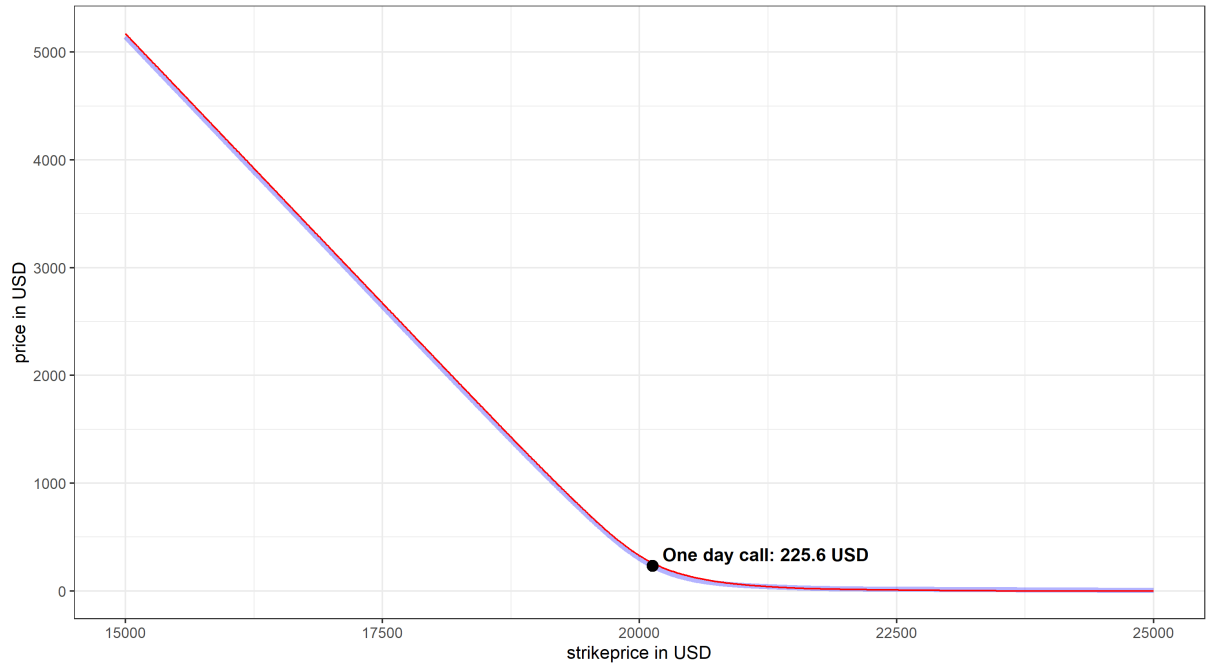


Figure 25: Shown above are the prices of one-day Bitcoin options by strike price. The blue line is the chosen starting value, the red line the arbitrary starting value. The black dot marks the point of an option with the price of the 01.09.2022 as strike price (20,127 USD) and marks its cost at 225.6 USD with starting day chosen as for the blue line.

One-week options

Figure 26 represents the same graph, but for a time horizon of one week. The interval of the considered strikes is increased to 12,500-27,500 USD. Some small separation between the blue and red line becomes visible, indicating that options at the 01.09.2022 are generally lower than on an average day. Here, the values for the red line are already always greater than those of the blue line, and the difference is 94 USD at most. The call with the current price as strike already tripled in price coming in at about 3.4% of the Bitcoin price. Furthermore, the maximum price after one week in the 1,000 paths is 29,235, while the minimum is 10,460 USD.

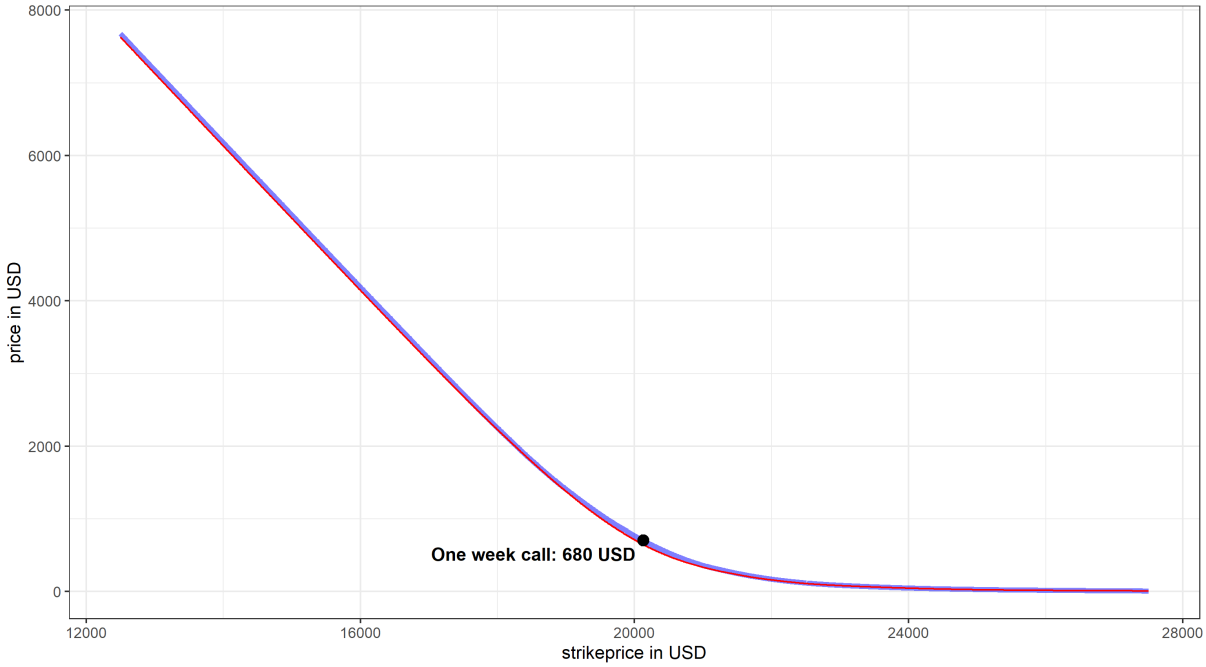


Figure 26: Shown above are the prices of one-week Bitcoin options by strike price. The blue line is the chosen starting value, the red line the arbitrary starting value. The black dot marks the point of an option with the price of the 01.09.2022 as strike price (20,127 USD) and marks its cost at 680 USD with starting day chosen as for the blue line.

One-month options

In Figure 27 the interval is increased again and now ranges from 10,000 USD to 30,000 USD. The difference between the blue and the red line increases again and now reaches up to 271 USD, while the minimal difference still is around 30 USD. The price of the one-month call is with 1,506 USD already about 7.5% of the price of a Bitcoin. The maximum and minimum values in the simulation are 42,531 USD and 10,102 USD, respectively.

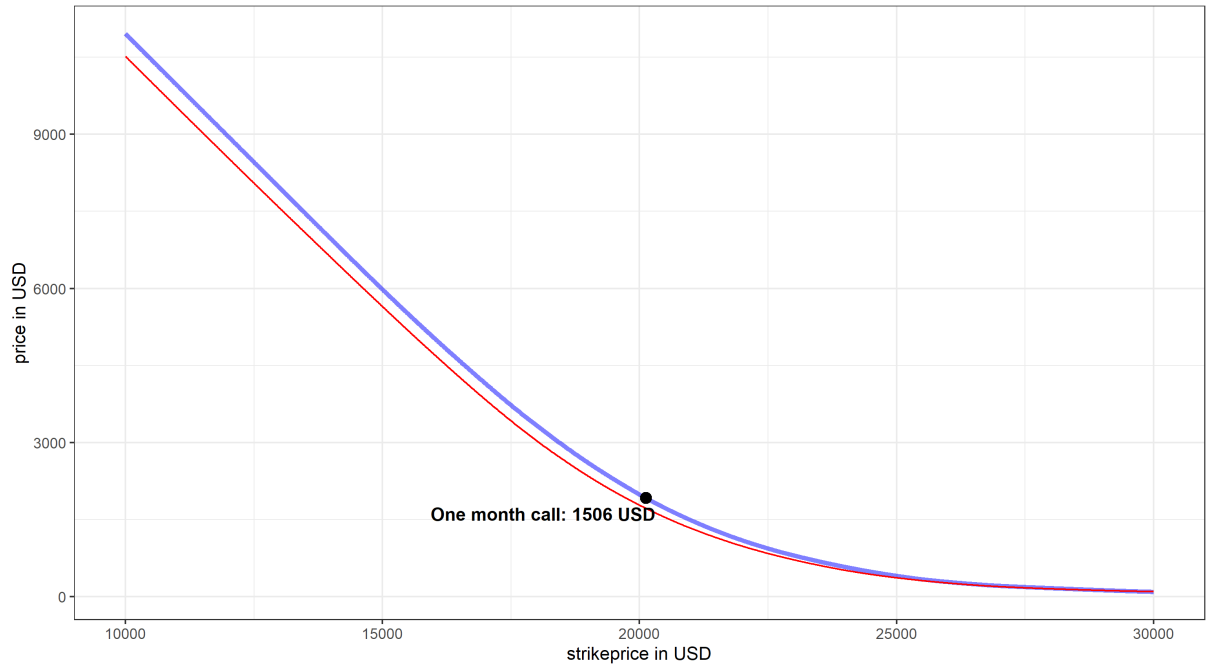


Figure 27: Shown above are the prices of one-month Bitcoin options by strike price. The blue line is the chosen starting value, the red line the arbitrary starting value. The black dot marks the point of an option with the price of the 01.09.2022 as strike price (20,127 USD) and marks its cost at 1,506 USD with starting day chosen as for the blue line.

One-year options

The longest time horizon considered is one year. For an asset like Bitcoin, this is already a huge time span, which is reflected in the prices of these options. The window for the strike price for Figure 28 now ranges from 5,000 USD up to 50,000 USD; however, even this increased window does not create a graph as flat as the ones for the options with shorter maturity were. Now the blue and red line are very distinct, and their minimum distance is 1,788 USD, while the maximum is as high as 2,311 USD. The call option price is at 13,718 USD, which is more than two thirds of a Bitcoin. This indicates, that looking further into the future is a pointless endeavour, as buying a Bitcoin now has almost the same price as buying an option with the current price as strike, so buying the coin outright prevents from spending the whole cost of the coin twice. In the simulation, the lowest occurring price after one year is 1,134 USD, so a loss of almost 95%, while the highest occurring price is 2,005,759 USD, almost the hundredfold of the original price.

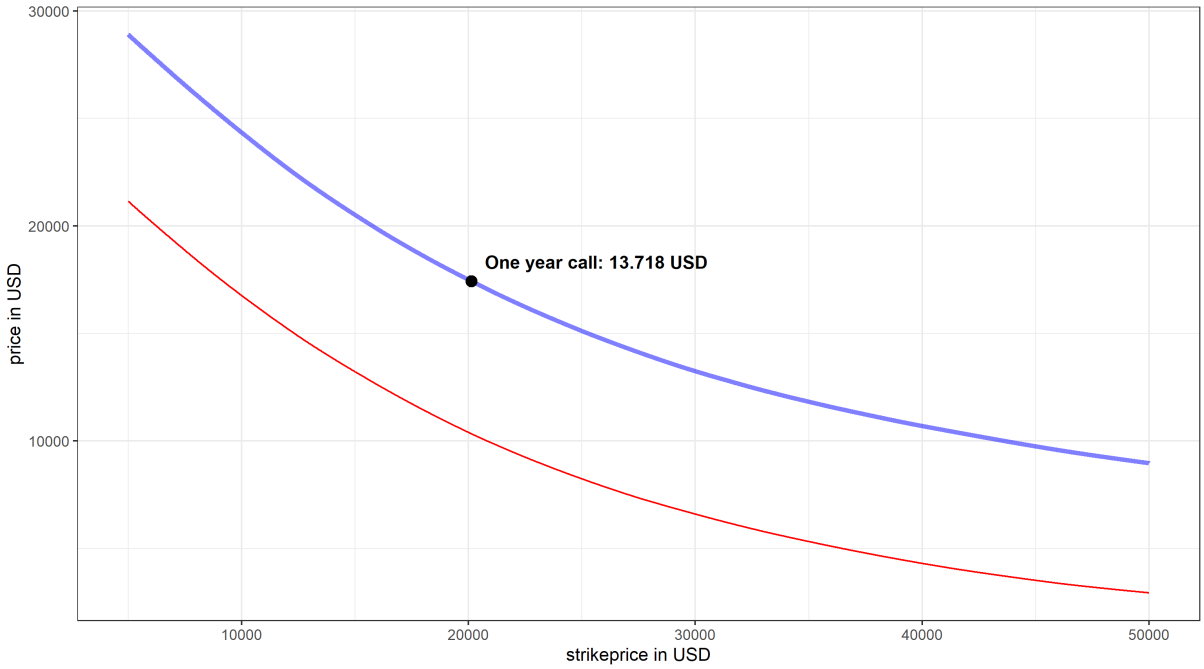


Figure 28: Shown above are the prices of one-year Bitcoin options by strike price. The blue line is the chosen starting value, the red line the arbitrary starting value. The black dot marks the point of an option with the price of the 01.09.2022 as strike price (20,127 USD) and marks its cost at 13,718 USD with starting day chosen as for the blue line.

Empirical distributions

Lastly, the empirical distributions of the simulated Bitcoin price at the maturity of the different options, i.e. 02.09.2022, 08.09.2022, 01.10.2022 and 01.09.2023, respectively, are displayed in Figure 29. The simulated prices at the 01.09.2023 already had many huge outliers, for the purpose of the graph, they were truncated at the price of 50,000 USD, which is why the bar at the 50,000 USD mark is so large, but as mentioned above, the actual prices continue into the million range.

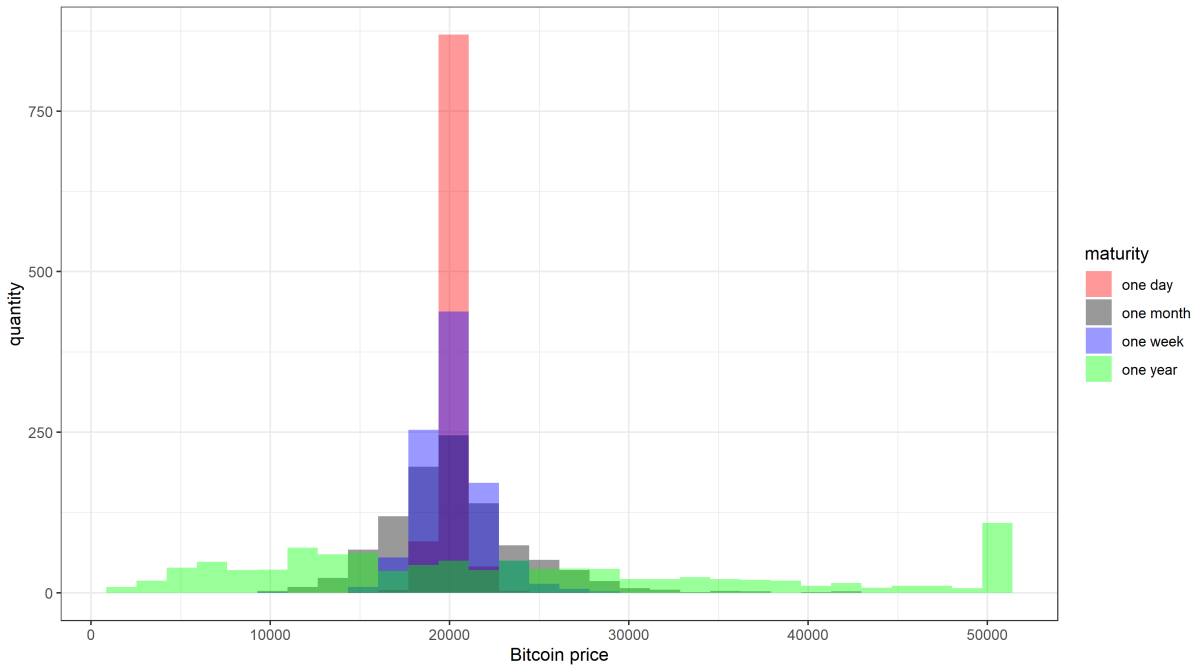


Figure 29: The graph displays the four empirical distributions of the simulated Bitcoin prices, that were used for the option pricing, with the tail of the distribution from the 01.09.2023 cut off.

Implications for Risk management

Looking at the Figures 19, 24 and 29, the hope for a small reserve fades quickly. Looking at the 5% VaR for the one-year horizon, the simulation yields that 73% of the invested amount needs to be held in reserves. The expected shortfall at 5% leads to an astounding 81% of the exposure that needs to be kept as a reserve.

In its most recent consultation [Committee, 2022], the Basel Committee refers to Bitcoin as a type two asset. This more or less means, that banks need to have suffiecient reserves so they can absorb a full write-off of all their capital invested in Bitcoin and other cryptoassets in this classification group. While this is an as careful as possible approach, the modelling results do not give any confidence in lowering this. While diversification could be considered, several studies on this have been conducted, with mixed results (see chapter 2.2). The difficulties in

hedging them further support the prudent approach by the Basel Committee.

6.2 Multivariate GARCH modelling

In this Section, only six out of the seven cryptoassets will be considered, as the data for Solana does not cover the full time period. For these six assets, multiple DCC-GARCH models will be fitted to the data. The models will be fitted using the R-package [Galanos, 2022]. Then, one of these models will be chosen based on its AIC and will be viewed in-depth.

6.2.1 Choice of model

The DCC-GARCH models are implemented via univariate GARCH models in order to reduce the computation amount. This means, that the matrices A_i and B_i in Definition 4.7 are diagonal. There will be two choices of univariate models that will be compared, the standard GARCH model and the TGARCH model. When the standard GARCH model is used, the model corresponds exactly to the one in Definition 4.7, while for the TGARCH model, equation (66) in Definition 4.7 is replaced with equation (40) from the TGARCH model. Also, only two choices of distributions for these univariate models are considered, namely the normal distribution and the student-t distribution; these distributions are only used to fit the matrices A_i and B_i and are not used any further in the model. Regarding the mulivariate distribution, again the multivariate normal distribution and the multivariate student-t distribution are considered. This yields a total of eight models that will be compared, but first, the necessary multivariate distributions will be defined.

Definition 6.3.
The **multivariate normal distribution** of dimension p is defined by

$$f(x) := \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right), \qquad \mu \in \mathbb{R}^p, \Sigma \in \mathbb{R}^{p\times p} \text{ positive definite} \tag{81}$$

The **multivariate student-t distribution** of dimension p is defined by

$$f(x) := \frac{1}{n^{\frac{p}{2}}\pi^{\frac{p}{2}}\Gamma(\frac{n}{2})\det(\Sigma)^{-\frac{1}{2}}}\Gamma(\frac{n+p}{2})\left(1+\frac{1}{n}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)^{-\frac{n+p}{2}}, \tag{82}$$
$$n \in \mathbb{N}^+, \mu \in \mathbb{R}^p, \Sigma \in \mathbb{R}p \times p \text{ positive definite}$$

Below, the eight DCC-GARCH models and their AIC values are displayed, again given divided by the sample size. One can see, that the AIC is most influenced by the choice of the multivariate distribution, where the student-t distribution significantly outperforms the

normal distribution. Furthermore, the TGARCH model is outperformed, while the univariate distributions vary in performance. Since in theory, they should converge to the same GARCH model (used as quasi-likelihood) it is not surprising that their impact is low. The optimal AIC value is marked in red. Therefore, the choice is the model from Definition 4.7, with student-t univariate and student-t multivariate distributions.

Standard GARCH	**normal distribution**	**student-t distribution**
multivariate normal	-23.78788	-23.56667
multivariate student-t	-25.80597	-25.95420
TGARCH	**normal distribution**	**student-t distribution**
multivariate normal	-23.63243	-23.55327
multivariate student-t	-25.60590	-25.55661

The univariate coefficients estimated for the model are displayed in the table below. The parameter values of Bitcoin are very similar to those estimated with the TGARCH model. Also, Bitcoin shows the highest persistence among the assets.

Cryptoasset	ω	α	β
Bitcoin	0.00003807994	0.06858828	0.9284978
Ethereum	0.0001673347	0.08912286	0.8542801
BNB	0.0002248226	0.1676200	0.7627335
XRP	0.0001261342	0.1520014	0.8469985
Cardano	0.0003522731	0.1842818	0.7413640
Dogecoin	0.0003585561	0.3930778	0.6059222

The DCC part of the model consists of 18 parameters, 15 of which are given in the matrix S below. Although it does not give the exact correlation, it is very close, which can be seen by comparing it with the average correlation matrix given in Section 6.2.2.

S	**Bitcoin**	**Ethereum**	**BNB**	**XRP**	**Cardano**	**Dogecoin**
Bitcoin	1	0.8235833	0.7481507	0.6645459	0.7048817	0.5631000
Ethereum	0.8235833	1	0.7898964	0.7135488	0.7609968	0.5697338
BNB	0.7481507	0.7898964	1	0.6683996	0.7243145	0.5314576
XRP	0.6645459	0.7135488	0.6683996	1	0.6839440	0.5412609
Cardano	0.7048817	0.7609968	0.7243145	0.6839440	1	0.5345644
Dogecoin	0.5631000	0.5697338	0.5314576	0.5412609	0.5345644	1

The two parameters that describe equation (68) are

$$\alpha = 0.04973892 \quad \text{and} \quad \beta = 0.9194281$$

The high value of β shows high persistence of shocks to the correlation, i.e. high-correlated periods tend to be followed by high-correlated periods.
The last parameter is the degrees of freedom for the multivariate student-t distribution which is 4.

6.2.2 Model graphs

Here are values describing the model output of the DCC-GARCH model chosen above. As the actual correlation between the assets changes each day, the outputs consist of 1,097 6×6 correlation matrices. The following matrix is the mean over all 1,097 days of the correlations between two assets. As mentioned above, it is quite similar to S. In fact, all differences are below 0.05. The strongest correlation is found between the largest two cryptoassets, Bitcoin and Ethereum, while the weakest correlation is found between Dogecoin and BNB. It is worth noting that the correlation in general is very strong, suggesting that the crypto market is strongly connected, which makes diversification within the market futile.

mean cor.	**Bitcoin**	**Ethereum**	**BNB**	**XRP**	**Cardano**	**Dogecoin**
Bitcoin	1	0.8189741	0.7444730	0.6695164	0.6996167	0.5913520
Ethereum	0.8189741	1	0.7906767	0.7255052	0.7579945	0.5863877
BNB	0.7444730	0.7906767	1	0.6802852	0.7253880	0.5573929
XRP	0.6695164	0.7255052	0.6802852	1	0.6963695	0.5873497
Cardano	0.6996167	0.7579945	0.7253880	0.6963695	1	0.5684471
Dogecoin	0.5913520	0.5863877	0.5573929	0.5873497	0.5684471	1

Figure 30 shows the correlation between Bitcoin and the other five modelled cryptoassets. It can be observed, that most correlations towards Bitcoin behave similarly, indicating that during certain periods, Bitcoin behaved unusually different than the rest of the analyzed market. The minimum correlation is observed between Bitcoin and Dogecoin at the 26.07.2020, where it fell to 0.167, the maximum correlation was observed on 13.03.2020 between Bitcoin and Ethereum, where it reached 0.970.

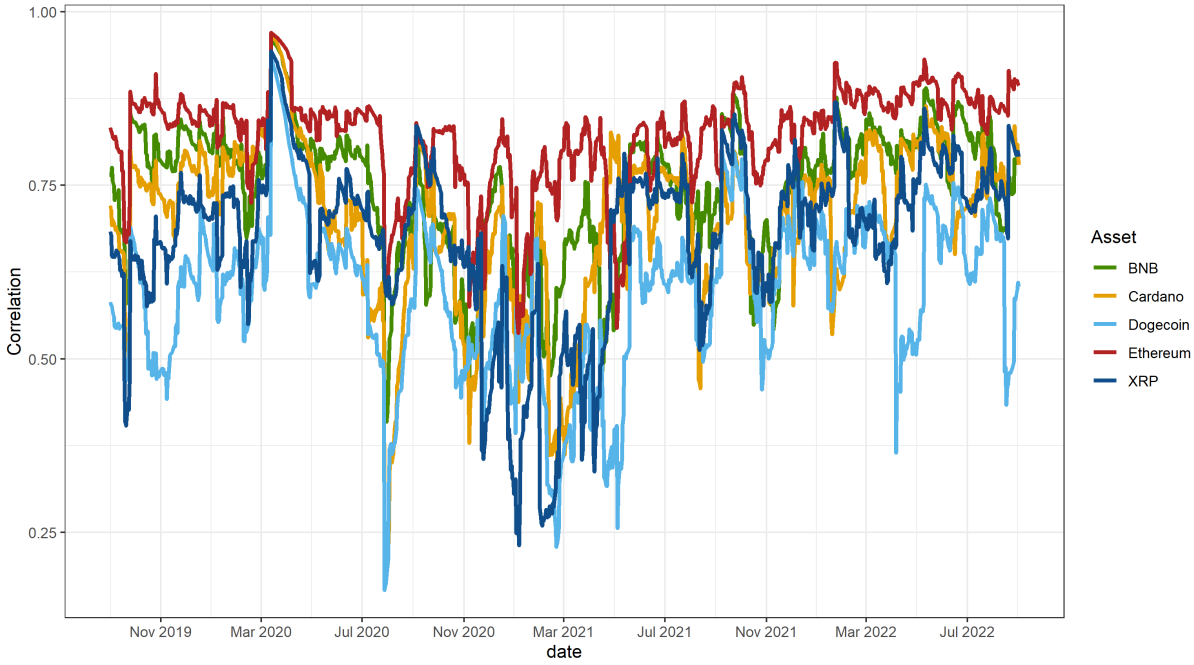


Figure 30: Shows the estimated correlation between Bitcoin and the other 5 modelled cryptoassets between 01.09.2019 and 01.09.2022

7 Conclusion

The results given by the TGARCH model with the student-t distribution that was implemented here seem to be very reasonable. The parameters given by the TGARCH model do reflect the properties of Bitcoin found in previous studies, i.e. high shock persistence and negative returns having a greater impact than positive ones. Furthermore, extreme paths are found in the simulation, but such relative price explosions can be found in Bitcoins early history, and although they are not as likely today, it occurs only in one of the 1,000 paths. Also, the European call option prices estimated by the simulation correspond to the extreme volatility of the underlying. It is found that long term options on Bitcoin are unnecessary, since the call option price is almost as high as the price of a Bitcoin. Moreover, plugging the model data into some risk measures instantly reveals, why the approach of the Basel Committee in [Committee, 2022] seems to be the correct way to go for now. Perhaps the market matures to a point where a smaller reserve would be possible, so far however there is no indication of that happening anytime soon.

The DCC-GARCH model shows strong correlations between the six assets studied. A noteworthy trend can be seen that assets with higher market capitalization are generally more correlated to other assets than those with low market capitalization. Also, the correlation between the assets is generally very high, but can suddenly drop due to events affecting only a single asset directly, e.g. Dogecoin suddenly getting pushed.

As for further research, various other fitting tests, such as in-sample testing or a general goodness of fit test, could be used to continue the investigation of the model fit. Theoretically, there are also more GARCH variations, and other commonly used models such as stochastic volatility or NARDL (Nonlinear Autoregressive Distributed Lags) models could be studied and compared to the GARCH results. While this could certainly lead to an even more improved model, the results concerning risk management and option prices would very likely stand, as well as the heavy correlation within the market.

References

[Aalborg et al., 2019] Aalborg, H. A., Molnár, P., and de Vries, J. E. (2019). What can explain the price, volatility and trading volume of Bitcoin? *Finance Research Letters*, 29:255–265.

[Aielli, 2013] Aielli, G. P. (2013). Dynamic Conditional Correlation: On Properties and Estimation on JSTOR. https://www-1jstor-1org-1000f97f90226.ftubhan.tugraz.at/stable/43702726.

[Aitchison and Silvey, 1958] Aitchison, J. and Silvey, S. D. (1958). Maximum-Likelihood Estimation of Parameters Subject to Restraints. *The Annals of Mathematical Statistics*, 29(3):813–828.

[Andrews, 1991] Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, 59(3):817–858.

[Ardia et al., 2019] Ardia, D., Bluteau, K., and Rüede, M. (2019). Regime changes in Bitcoin GARCH volatility dynamics. *Finance Research Letters*, 29:266–271.

[Billingsley, 1995] Billingsley, P. (1995). *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York, 3rd ed edition.

[Bollerslev, 1986] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

[Bollerslev, 1990] Bollerslev, T. (1990). Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model. *The Review of Economics and Statistics*, 72(3):498–505.

[Bollerslev et al., 1988] Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988). A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*, 96(1):116–131.

[Bouoiyour and Selmi, 2015a] Bouoiyour, J. and Selmi, R. (2015a). Bitcoin Price: Is it really that New Round of Volatility can be on way?

[Bouoiyour and Selmi, 2015b] Bouoiyour, J. and Selmi, R. (2015b). What Does Bitcoin Look Like? *Annals of Economics and Finance*, 16:449–492.

[Bouoiyour and Selmi, 2016] Bouoiyour, J. and Selmi, R. (2016). Bitcoin: A beginning of a new phase? *Economics Bulletin*, 36:1430–1440.

[Bouri et al., 2017a] Bouri, E., Azzi, G., and Dyhrberg, A. H. (2017a). On the return-volatility relationship in the Bitcoin market around the price crash of 2013. *Economics*, 11(1):2.

[Bouri et al., 2018a] Bouri, E., Das, M., Gupta, R., and Roubaud, D. (2018a). Spillovers between Bitcoin and other assets during bear and bull markets. *Applied Economics*, 50(55):5935–5949.

[Bouri et al., 2018b] Bouri, E., Gupta, R., Lahiani, A., and Shahbaz, M. (2018b). Testing for asymmetric nonlinear short- and long-run relationships between bitcoin, aggregate commodity and gold prices. *Resources Policy*, 57:224–235.

[Bouri et al., 2017b] Bouri, E., Molnár, P., Azzi, G., Roubaud, D., and Hagfors, L. I. (2017b). On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters*, 20:192–198.

[BRAIINS, 2021] BRAIINS (2021). How Much Would it Cost to 51% Attack Bitcoin? — Braiins. https://braiins.com/blog/how-much-would-it-cost-to-51-attack-bitcoin.

[Canh et al., 2019] Canh, N. P., Wongchoti, U., Thanh, S. D., and Thong, N. T. (2019). Systematic risk in cryptocurrency market: Evidence from DCC-MGARCH model. *Finance Research Letters*, 29:90–100.

[Catania and Grassi, 2017] Catania, L. and Grassi, S. (2017). Modelling Crypto-Currencies Financial Time-Series.

[Catania et al., 2019] Catania, L., Grassi, S., and Ravazzolo, F. (2019). Forecasting cryptocurrencies under model and parameter instability. *International Journal of Forecasting*, 35(2):485–501.

[Chaim and Laurini, 2019] Chaim, P. and Laurini, M. P. (2019). Nonlinear dependence in cryptocurrency markets. *The North American Journal of Economics and Finance*, 48:32–47.

[Chan, 1999] Chan, W.-S. (1999). A comparison of some of pattern identification methods for order determination of mixed ARMA models. *Statistics & Probability Letters*, 42(1):69–79.

[Charfeddine et al., 2020] Charfeddine, L., Benlagha, N., and Maouchi, Y. (2020). Investigating the dynamic relationship between cryptocurrencies and conventional assets: Implications for financial investors. *Economic Modelling*, 85:198–217.

[Cheikh et al., 2020] Cheikh, N. B., Zaied, Y. B., and Chevallier, J. (2020). Asymmetric volatility in cryptocurrency markets: New evidence from smooth transition GARCH models. *Finance Research Letters*, 35:101293.

[Chen, 2011] Chen, A. (2011). The Underground Website Where You Can Buy Any Drug Imaginable. https://www.gawker.com/the-underground-website-where-you-can-buy-any-drug-imag-30818160.

[Chu et al., 2017] Chu, J., Chan, S., Nadarajah, S., and Osterrieder, J. (2017). GARCH Modelling of Cryptocurrencies. *Journal of Risk and Financial Management*, 10(4):17.

[Chu et al., 2015] Chu, J., Nadarajah, S., and Chan, S. (2015). Statistical Analysis of the Exchange Rate of Bitcoin. *PLOS ONE*, 10(7):e0133678.

[Ciaian et al., 2018] Ciaian, P., Rajcaniova, M., and Kancs, d. (2018). Virtual relationships: Short- and long-run evidence from Bitcoin and Altcoin markets. *Journal of International Financial Markets, Institutions and Money*, 52:173–195.

[coinmarketcap.com, 2022] coinmarketcap.com (2022). Coinmarketcap.com. https://coinmarketcap.com/.

[Committee, 2022] Committee, B. (2022). Prudential treatment of cryptoasset exposures - second consultation.

[Corbet et al., 2020] Corbet, S., Hou, Y. G., Hu, Y., Larkin, C., and Oxley, L. (2020). Any port in a storm: Cryptocurrency safe-havens during the COVID-19 pandemic. *Economics Letters*, 194:109377.

[De Gooijer and Heuts, 1981] De Gooijer, J. G. and Heuts, R. (1981). The Corner Method: An Investigation of an Order Discrimination Procedure for General ARMA Processes. *Journal of the Operational Research Society*, 32:1039/1042.

[Demir et al., 2021] Demir, E., Simonyan, S., García-Gómez, C.-D., and Lau, C. K. M. (2021). The asymmetric effect of bitcoin on altcoins: Evidence from the nonlinear autoregressive distributed lag (NARDL) model. *Finance Research Letters*, 40:101754.

[DIGICONOMIST, 2022] DIGICONOMIST (2022). Bitcoin Energy Consumption Index. https://digiconomist.net/bitcoin-energy-consumption/.

[Ding et al., 1993] Ding, Z., Granger, C. W. J., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106.

[Dwita Mariana et al., 2021] Dwita Mariana, C., Ekaputra, I. A., and Husodo, Z. A. (2021). Are Bitcoin and Ethereum safe-havens for stocks during the COVID-19 pandemic? *Finance Research Letters*, 38:101798.

[Dyhrberg, 2016a] Dyhrberg, A. H. (2016a). Bitcoin, gold and the dollar – A GARCH volatility analysis. *Finance Research Letters*, 16:85–92.

[Dyhrberg, 2016b] Dyhrberg, A. H. (2016b). Hedging capabilities of bitcoin. Is it the virtual gold? *Finance Research Letters*, 16:139–144.

[Engle, 1982] Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. http://www.econ.uiuc.edu/~econ536/Papers/engle82.pdf.

[Engle, 2000] Engle, R. F. (2000). Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. *SSRN Electronic Journal*.

[Engle and Bollerslev, 1986] Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5(1):1–50.

[Engle et al., 1984] Engle, R. F., Granger, C. W. J., and Kraft, D. (1984). Combining competing forecasts of inflation using a bivariate arch model. *Journal of Economic Dynamics and Control*, 8(2):151–165.

[Engle and Kroner, 1995] Engle, R. F. and Kroner, K. F. (1995). Multivariate Simultaneous Generalized Arch. *Econometric Theory*, 11(1):122–150.

[Fisher, 2012] Fisher, T. J. (2012). WeightedPortTest package.

[Francq et al., 2014] Francq, C., Horvath, L., and Zakoian, J.-M. (2014). Variance targeting estimation of multivariate GARCH models. https://mpra.ub.uni-muenchen.de/57794/.

[Francq and Zakoïan, 2004] Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605–637.

[Francq and Zakoian, 2019] Francq, C. and Zakoian, J.-M. (2019). GARCH Models: Structure, Statistical Inference and Financial Applications, 2nd Edition — Wiley. https://www.wiley.com/en-us/GARCH+Models%3A+Structure%2C+Statistical+Inference+and+Financial+Applications%2C+2nd+Edition-p-9781119313489.

[Galanos, 2022] Galanos, A. (2022). Rmgarch: Multivariate GARCH Models.

[Galanos and Kley, 2022] Galanos, A. and Kley, T. (2022). Rugarch: Univariate GARCH Models.

[Georgoula et al., 2015] Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D., and Giaglis, G. M. (2015). Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices. SSRN Scholarly Paper 2607167, Social Science Research Network, Rochester, NY.

[Glosten et al., 1993] Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). 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.

[González et al., 2020] González, M. d. l. O., Jareño, F., and Skinner, F. S. (2020). Nonlinear Autoregressive Distributed Lag Approach: An Application on the Connectedness between Bitcoin Returns and the Other Ten Most Relevant Cryptocurrency Returns. *Mathematics*, 8(5):810.

[González et al., 2021] González, M. d. l. O., Jareño, F., and Skinner, F. S. (2021). Asymmetric interdependencies between large capital cryptocurrency and gold returns during the COVID-19 pandemic crisis. *International Review of Financial Analysis*, 76:101773.

[Guesmi et al., 2019] Guesmi, K., Saadi, S., Abid, I., and Ftiti, Z. (2019). Portfolio diversification with virtual currency: Evidence from bitcoin. *International Review of Financial Analysis*, 63:431–437.

[Guo, 2022] Guo, Z.-Y. (2022). Risk management of Bitcoin futures with GARCH models. *Finance Research Letters*, 45:102197.

[Hou et al., 2019] Hou, A. J., Wang, W., Chen, C. Y.-H., and Härdle, W. K. (2019). Pricing Cryptocurrency Options: The Case of Bitcoin and CRIX.

[Hou et al., 2020] Hou, A. J., Wang, W., Chen, C. Y. H., and Härdle, W. K. (2020). Pricing Cryptocurrency Options. *Journal Of Financial Econometrics*.

[Hu et al., 2021] Hu, Y., Rachev, S. T., and Fabozzi, F. J. (2021). Modelling Crypto Asset Price Dynamics, Optimal Crypto Portfolio, and Crypto Option Valuation. *The Journal of Alternative Investments*, 24(1):75–93. Comment: 25 pages, 6 figures.

[Hubbard, 2022] Hubbard, A. (2022). Autostsm: Automatic Structural Time Series Models.

[Jareño et al., 2020] Jareño, F., González, M. d. l. O., Tolentino, M., and Sierra, K. (2020). Bitcoin and gold price returns: A quantile regression and NARDL analysis. *Resources Policy*, 67:101666.

[Katsiampa, 2017] Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters*, 158:3–6.

[Katsiampa et al., 2019] Katsiampa, P., Corbet, S., and Lucey, B. (2019). Volatility spillover effects in leading cryptocurrencies: A BEKK-MGARCH analysis. *Finance Research Letters*, 29:68–74.

[King and Nadal, 2012] King, S. and Nadal, S. (2012). PPCoin: Peer-to-Peer Crypto-Currency with Proof-of-Stake. page 6.

[Kjærland et al., 2018] Kjærland, F., Khazal, A., Krogstad, E., Nordstrøm, F., and Oust, A. (2018). An Analysis of Bitcoin's Price Dynamics. *Journal of Risk and Financial Management*, 11(4):63.

[Klein et al., 2018] Klein, T., Pham Thu, H., and Walther, T. (2018). Bitcoin is not the New Gold – A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis*, 59:105–116.

[Kristoufek, 2015] Kristoufek, L. (2015). What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis. *PLOS ONE*, 10(4):e0123923.

[Kumar and Suvvari, 2019] Kumar, A. and Suvvari, A. (2019). Volatility spillover in cryptocurrency markets: Some evidence from GARCH and wavelet analysis. *Physica A: Statistical Mechanics and its Applications*, 524.

[Madan et al., 2019] Madan, D. B., Reyners, S., and Schoutens, W. (2019). Advanced model calibration on bitcoin options. *Digital Finance*, 1(1-4):117–137.

[Matic et al., 2021] Matic, J. L., Packham, N., and Härdle, W. K. (2021). Hedging Cryptocurrency Options. SSRN Scholarly Paper 3968594, Social Science Research Network, Rochester, NY.

[Mensi et al., 2019] Mensi, W., Al-Yahyaee, K. H., and Kang, S. H. (2019). Structural breaks and double long memory of cryptocurrency prices: A comparative analysis from Bitcoin and Ethereum. *Finance Research Letters*, 29:222–230.

[Naimy and Hayek, 2018] Naimy, V. and Hayek, M. (2018). Modelling and predicting the Bitcoin volatility using GARCH models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 8:197.

[Nakamoto, 2008] Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. page 9.

[Nelson, 1990] Nelson, D. B. (1990). Stationarity and Persistence in the GARCH(1,1) Model. *Econometric Theory*, 6(3):318–334.

[Nelson, 1991] Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2):347–370.

[Newey and West, 1987] Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3):703–708.

[Nguyen et al., 2019] Nguyen, C., Quang Binh, N., and Su Dinh, T. (2019). Cryptocurrencies and Investment Diversification: Empirical Evidence from Seven Largest Cryptocurrencies. *Theoretical Economics Letters*, 09:431–452.

[Noori and Mohammad, 2021] Noori, N. A. and Mohammad, A. A. (2021). Dynamical Approach in studying GJR-GARCH (Q,P) Models with Application. *Tikrit Journal of Pure Science*, 26(2):145–156.

[Omane-Adjepong and Alagidede, 2019] Omane-Adjepong, M. and Alagidede, I. P. (2019). Multiresolution analysis and spillovers of major cryptocurrency markets. *Research in International Business and Finance*, 49:191–206.

[Paolella, 2018] Paolella, M. S. (2018). *Linear Models and Time-Series Analysis*.

[Phillip et al., 2018] Phillip, A., Chan, J. S. K., and Peiris, S. (2018). A new look at Cryptocurrencies. *Economics Letters*, 163:6–9.

[Poyser, 2017] Poyser, O. (2017). *Exploring the Determinants of Bitcoin's Price: An Application of Bayesian Structural Time Series*. PhD thesis.

[Reiff, 2022] Reiff, N. (2022). What Was the First Cryptocurrency? https://www.investopedia.com/tech/were-there-cryptocurrencies-bitcoin/.

[Scaillet et al., 2017] Scaillet, O., Treccani, A., and Trevisan, C. (2017). High-Frequency Jump Analysis of the Bitcoin Market. *SSRN Electronic Journal*.

[Shi et al., 2020] Shi, Y., Tiwari, A. K., Gozgor, G., and Lu, Z. (2020). Correlations among cryptocurrencies: Evidence from multivariate factor stochastic volatility model. *Research in International Business and Finance*, 53:101231.

[Sovbetov, 2018] Sovbetov, Y. (2018). Factors Influencing Cryptocurrency Prices: Evidence from Bitcoin, Ethereum, Dash, Litcoin, and Monero. SSRN Scholarly Paper 3125347, Social Science Research Network, Rochester, NY.

[Stensås et al., 2019] Stensås, A., Nygaard, M. F., Kyaw, K., and Treepongkaruna, S. (2019). Can Bitcoin be a diversifier, hedge or safe haven tool? *Cogent Economics & Finance*, 7(1):1593072.

[Symitsi and Chalvatzis, 2019] Symitsi, E. and Chalvatzis, K. J. (2019). The economic value of Bitcoin: A portfolio analysis of currencies, gold, oil and stocks. *Research in International Business and Finance*, 48:97–110.

[Trapletti et al., 2022] Trapletti, A., Hornik, K., and test code), B. L. B. (2022). Tseries: Time Series Analysis and Computational Finance.

[Troster et al., 2019] Troster, V., Tiwari, A. K., Shahbaz, M., and Macedo, D. N. (2019). Bitcoin returns and risk: A general GARCH and GAS analysis. *Finance Research Letters*, 30:187–193.

[Tsay, 2010] Tsay, R. S. (2010). *Analysis of Financial Time Series, Third Edition.*

[Tu and Xue, 2019] Tu, Z. and Xue, C. (2019). Effect of bifurcation on the interaction between Bitcoin and Litecoin. *Finance Research Letters*, 31.

[Venter et al., 2020] Venter, P. J., Mare, E., and Pindza, E. (2020). Price discovery in the cryptocurrency option market: A univariate GARCH approach. *Cogent Economics & Finance*, 8(1):1803524.

[WORLDDATAINFO, 2022] WORLDDATAINFO (2022). Energy consumption in Austria. https://www.worlddata.info/europe/austria/energy-consumption.php.

[Yousaf and Ali, 2020] Yousaf, I. and Ali, S. (2020). Discovering interlinkages between major cryptocurrencies using high-frequency data: New evidence from COVID-19 pandemic. *Financial Innovation*, 6(1):45.

[Zakoian, 1994] Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955.