Research on properties of Bitcoin Based on GARCH model

Jiacong Yuan*

University College London, University of London, London, United Kingdom

*Corresponding author. Email: jiacong.yuan.21@ucl.ac.uk

Abstract. Cryptocurrencies have become a world-class phenomenon, with governments, companies and investors facing huge challenges and opportunities. Bitcoin is the one of most widely known cryptocurrencies. People prefer to use Bitcoin as an asset to invest for profit and hedge risk than for its payment function. This is the reason for the study of bitcoin market is so important. Many scholars had used daily data on bitcoin to construct different GARCH models to analyse the volatility of its returns. This research builds a GARCH model for the logarithmic return series of the bitcoin price to understand the volatility of the bitcoin price over the experimental time horizon. The experimental results show that the price of bitcoin is more volatile and vulnerable to external shocks. The analysis suggests that Bitcoin is more clearly a speculative financial instrument and that the price of Bitcoin is highly frothy.

Keywords: Bitcoin, Cryptocurrencies, GARCH, Price fluctuations, Logarithmic rate of return.

1. Introduction

Today many government, auditing firms and internet companies are conducting research to cryptocurrencies and blockchain.The first cryptocurrency, Bitcoin, was launched as a by-product of the blockchain. Unlike other traditional financial assets, cryptocurrencies have no physical form or government backing, and their creation relies solely on technology. The particular form in which cryptocurrencies are created and traded, makes them slightly different in nature from ordinary financial products. The nature of cryptocurrencies can be divided into two aspects: the transactional and monetary aspects. Due to its reliance on a decentralised system, cryptocurrency transactions, once stale, are irreversible and remain in the blockchain forever. Transactions are also anonymous and all transactions and records are not linked to real identities, which make them a popular form of transaction on the dark web. Cryptocurrencies rely on the internet and are not governed by governments or institutions, so their transactions are not restricted by location and are confirmed quickly, typically taking only a few minutes. Cryptocurrencies are secured by keys. As a new type of financial asset, cryptocurrencies have had some impact on the traditional financial markets and have been regulated by governments in various countries. In 2013, China released a document on the regulation of trading in the Bitcoin market, and other cryptocurrency organisations such as the Federal Reserve and the European Central Bank had issued numerous policies on the subject, but this has not stopped the frenzy over cryptocurrencies. In this paper, Bitcoin is chosen as a representative of cryptocurrencies. In the financial markets, Bitcoin can be used as a means of payment for cross-border transactions and as an asset to diversify risk in investments. Bitcoin is not strongly correlated with traditional assets and alternative investments such as real estate and hedge funds. Besides, fluctuations in the stock market do not affect the returns of Bitcoin.

Autoregressive conditional heteroskedasticity (ARCH) and generalised autoregressive conditional heteroskedasticity (GARCH) models [1] have been widely used in the estimation of volatility in time series in finance. Both models were proposed by Engle and Bollerslev in 1982 and 1986 respectively, and were based on least squares models. The differences are the least squares model assumes that the variance of the residual terms is the same, whereas the ARCH and GARCH models take into account the heteroskedasticity of the residual terms of asset returns [2]. The ARCH and GARCH models take into account the heteroskedasticity of the residuals of asset returns. The combination of heteroskedasticity and autoregression is a good representation of the volatility of asset returning. The

combination of heteroskedasticity and autoregression can well represent the volatility of asset returns as a feature of clustering [3].

Several determinants of virtual currency prices have been identified as important factors in the previous literature, including supply and demand market forces [4], speculations [5], and macro-financial developments [6]. Fang used the GARCH-MIDAS approach [7] to show that global economic policy uncertainty improves the hedging efficacy of Bitcoin and the prediction of Bitcoin volatility. Previous research had used GARCH-based models to simulate Bitcoin volatility, with the majority of predictions relying solely on previous returns as conditional information [8]. The GARCH model for the logarithmic rate is established in this study.

2. DATA

The closing price is the price at which no further trading takes place and is the price that is commonly accepted by market participants price. Therefore, the research will use the daily closing price of Bitcoin as the data. The paper will therefore use the daily closing price of Bitcoin as the data object. Bitcoin has been existed around since 2009, but because it was initially little known and the market demand was so low that it was volatile. The paper will use the daily closing price of Bitcoin as the data object. The paper is based on the period 2710 daily closing prices from September 17, 2016 to February 16, 2022.

As the bitcoin price is non-stationary, the independent variables and variables need to satisfy the assumption of smoothness to perform regression analysis. If we want to make the data trend smooth and linear, without changing the original relationship, we will log the time. (Bitcoin's price fluctuations inevitably lead to fluctuations in returns)) Therefore, the pre-processing is done by taking the logarithmic return on the closing price of bitcoin and constructing the series with a sample of data. Sequence M are constructed with 2710 sample data. Suppose that the close price of bitcoin at t'clock is M_t. We will use the logarithmic form of M to construct a logarithmic yield series N. Suppose the logarithmic rate of return is N_t, then the logarithmic rate of return takes the form: $N_t = \ln(M_t) - \ln(M_{t-1})$. The full sample will be consisted of 2710 log-return data through pre-processing (Figure 1).

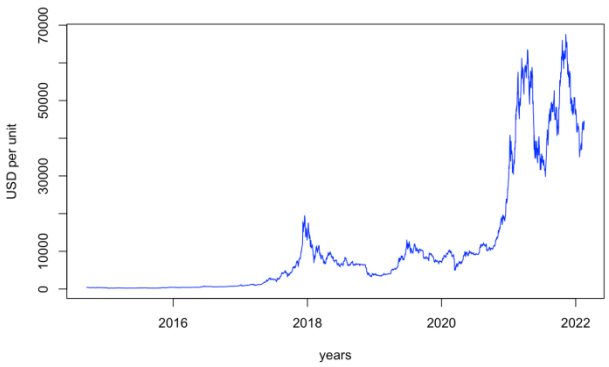


Figure 1. The time series plot of yields

3. Method

The general case for modelling a smooth time series is to build an autoregressive moving average model (ARMA (p, q)). For some non-smooth series, although the individual series values have uncertainty, the overall variation may still be represented by a smooth stochastic process. For example, after differential transformation, the model will have smoothness, then we can use ARIMA (p, d, q) model to fit this time series. The return series of the index is r_t. For the return series r_t, the ARMA (p, q) model is normally used to analyze the autocorrelation of the series. The ARMA(p, q) model takes the form:

$$r_t = \varphi_0 + \varphi_1 r_t + \cdots + \varphi_p r_{t-p} + u_t - \theta_1 u_{t-1} - \cdots - \theta_q u_{t-q} \quad (1)$$

$$\varphi_p \neq 0, \theta_q \neq 0 \quad (2)$$

$$E(u_t) = 0, Var(u_t) = \sigma_u^2, E(u_t u_s) = 0, s \neq t \quad (3)$$

$$E(r_t u_t) = 0, \forall s < t \quad (4)$$

Where$\varphi_0, \varphi_0, \varphi_1, \cdots, \varphi_{p,} \theta_1, \cdots, \theta_q$are polynomial coefficients, with the parameter to be estimated later. $\varphi_0, \varphi_0, \varphi_1, \cdots, \varphi_p$ are autoregressive partial coefficients and $\theta_1, \cdots, \theta_q$ are moving average partial coefficients. u_t is a white noise series with mean 0 and variance σ_u^2.

Since the ARMA(p, q) model assumes a white noise series which makes the model ineffective in explaining the volatility clustering that often occurs in the index return series, the GARCH(m, n) model needs to be introduced to analyze the volatility clustering that cannot be explained by the ARMA(p, q) model.

Bollerslov proposed the generalised autoregressive conditional heteroskedasticity (GARCH) model in 1985. The GARCH (p, q) model, also known as the generalised ARCH model, is a generalised form of the ARCH model. When q=0, the GARCH model is equivalent to the ARCH model. In fact, the ARCH model is only applicable to the short-term autocorrelation process of the heteroskedasticity function. Therefore, the GARCH model is formed by adding the p-order autoregression of the heteroskedasticity function to the ARCH model.

Let us take the log-return series as r_t and denote the perturbation at moment t as α_t. The μ_t model is known as the mean equation of r_t, and the σ_t^2 model is known as the volatility equation of r_t.

$$a_t = r_t - \mu_t \quad (5)$$

$$a_t = \sigma_t \varepsilon_t \quad (6)$$

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

Where ε_t is a sequence of independent identically distributed random variables with mean 0 and variance 1. Then we have $\alpha_0 > 0, \alpha_i > 0$ $\sum_{i=1}^{max(m,s)} \alpha_i + \beta_j < 1$. Suppose ε_t is normally distributed.

Stability check: Unit root tests are commonly used to test the smoothness of time series, and Augmented Dickey-Fuller Test (Table.1) is used for log-return series. The ADF test assumes independent identical distribution for random perturbations, with the original hypothesis: there is a unit root (non-stationary) in the original time series.

Normality test: When describing the morphological and distributional characteristics of the data, in addition to the JB statistic, the data can also be analysed by and histograms (Figure 2) and QQ plots (Figure 3).

Conducting Autocorrelation test: Based on the smooth series, it can be seen that the autocorrelation (ACF) and partial autocorrelation (PACF) plots (Figure 4) of the yield series both have obvious trailing phenomena. Therefore, an attempt is made to develop an ARMA model to fit the mean equation of the series and portray the linear trend of the return series.

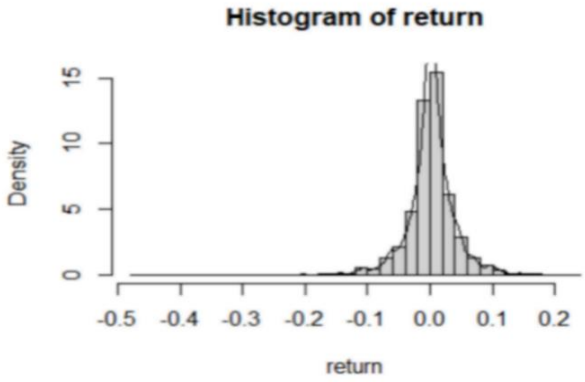


Figure 2. Normality test

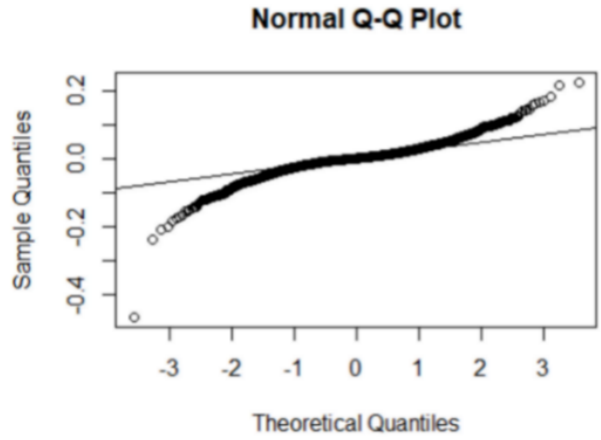


Figure 3. Normal *Q-Q plot*

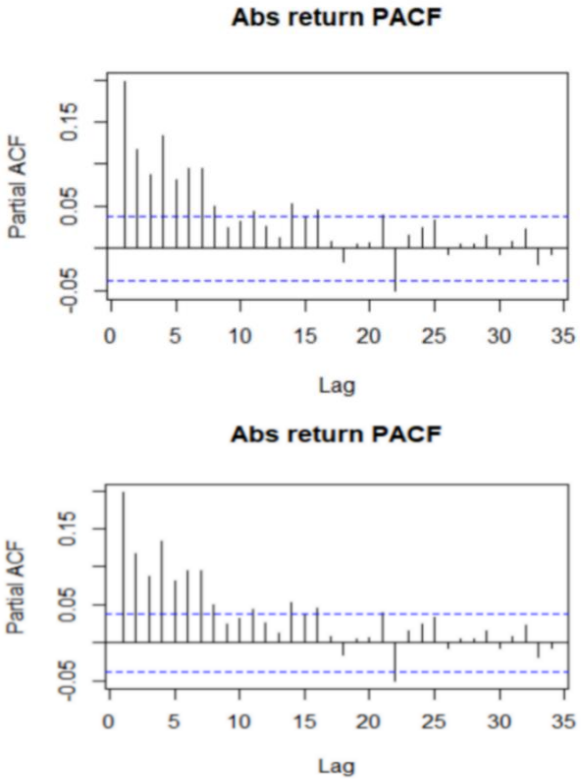


Figure 4. Autocorrelation test

Table 1. ADF test

Variable	ADF value	Prob	result
Return	-53.08531	0	stable

After that, fitting the GARCH model, the mean equation μ_t is developed for the series of returns:

$$r_t = -0.0196r_{t-1} + 0.0025r_{t-2} + \varepsilon_t \tag{8}$$

Then, let's test the ARCH effect. The ARCH effect test hypothesis is:

$$H_0: \rho_1 = \rho_2 = \cdots \rho_m = 0, \forall m \geq 1 \tag{9}$$

$$H_1: \ \prod_{k=1}^{m} \rho_k \neq 0 \tag{10}$$

The test statistic is the LB (Ljung-Box) statistic:

$$LB = \mathrm{n}(n+2) \sum_{k=1}^{m} \left(\frac{\hat{\rho}_k^2}{n-k}\right) \sim \chi^2(m), \forall_m > 0 \tag{11}$$

Establishing GARCH (1, 1) parameter estimation, variance equation(condition) is followed $\sigma_t^2 = \omega + \alpha {\mu_{t-1}}^2 + \beta\sigma_{t-1}^2$ In the conditional variance equation, ω is the constant term. μ_{t-12} is the ARCH term, which is the squared perturbation term of the previous period; α is the information parameter, representing the size of the impact of the emergence of new information in the stock spot market on the conditional variance of the current period. The larger α indicates the stronger effect by the new information; σ_{t-12} is the forecast variance of the previous period. β is the σ_{t-12} parameter, which indicating the impact of historical market news on the conditional variance of the current period. The larger β indicates the stronger effect of historical news.

4. Results and discussion

The time series plot of yields (Figure 1) shows that yields are less volatile in some periods and more volatile in others. This shows the typical "volatility aggregation phenomenon" of financial time series, so it can be considered for GARCH modelling.

Now we descriptive statistical analysis of the return (Figure 5). The mean value of the return is 0.001683; the skewness is -0.792858 <0 return left skewed distribution; the kurtosis is 14.23398 greater than 3; the return distribution is concentrated with a spike trend; the J-B statistic is corresponding to a p-value of 0. Overall the return series has spike thick tail non-normal characteristics.

To verify the validity of the model, a smoothness test is performed before modelling the time series. The test reveals that the hypothesis of the existence of a unit root is rejected at the 1% significance level. The p-value corresponding to the value of the test statistic is 0, indicating that the index return series is a smooth series and there is no unit root.

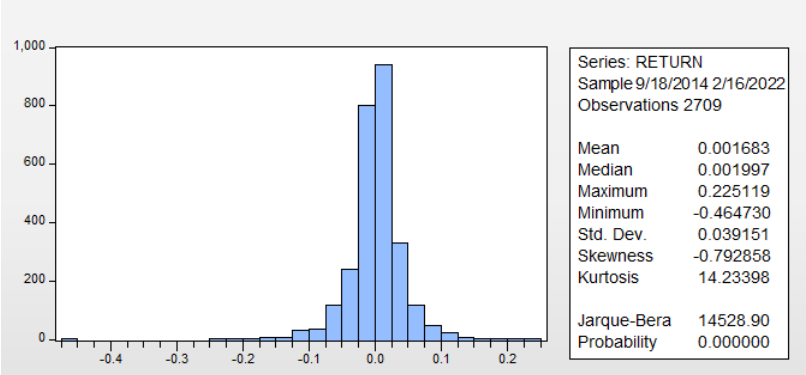


Figure 5. Descriptive statistical analysis of the return

In Figure 4, it is observed that the QQ plot shows a clear tendency for the distribution, indicating that the index return distribution is non-normal. This indicates that the data fluctuates dramatically in price over a certain time interval, and the distribution has a spiky and thick-tailed pattern. Relative to the standard normal distribution, the actual distribution of returns is more kurtic and the tails of the two segments are thicker, implying that volatility is likely to be higher.

The mean value equation ($r_t = -0.0196r_{t-1} + 0.0025r_{t-2} + \varepsilon_t$) is a model to fit a linear trend to a series of returns. The number of orders p, q in an ARMA model is generally determined according to the autocorrelation, partial autocorrelation graph coefficient trailing and truncation tails. Fixing the order to compare the orders in the model based on the criterion that the smaller the AIC and BIC are, the better the model is. In R Studio, let's set the order to p=2 and q=0 and build the $ARMA(2,0)$ model, then run auto arima in R to output the following results (Figure 6).

```
ARIMA(2,0,0) with non-zero mean

Coefficients:
          ar1     ar2    mean
      -0.0196  0.0025  0.0017
s.e.   0.0192  0.0192  0.0007

sigma^2 estimated as 0.001533:  log likelihood=4935.18
AIC=-9862.36   AICc=-9862.34   BIC=-9838.74
```

Figure 6. AMIRMA (2,0,0)

Table 2. Different parameters

Parameters	Estimated value	Standard error
AR(1)	-0.0196	0.0192
AR(2)	0.0025	0.0192

Table 3. Result of ARCH effect test

series	Lagging order	Chi-square	P-value
return	5	53.716,	0
	10	77.262	0
	15	81.756	0
	20	83.625	0

If the original ARCH effect test hypothesis is rejected, it indicates that there is still relevant information in the residual series, and further GARCH models are built to portray data heteroskedasticity. Now we will study the volatility of spot historical data, and portray volatility characteristics with fitted marginal distributions. After determining the mean equation, it is tested for ARCH effect and the residual test of the return is studied to determine whether there is an ARCH effect in the return series. The method used is the commonly called Lagrange multiplier test for the residuals of the mean equation. The results are showed in the Result of ARCH effect test (Table 3)

The p-value of the chi-square statistic is significantly less than 0.01. The lagged residuals jointly reject the original hypothesis that the series does not have an ARCH effect at lags 5, 10, 15 and 20, indicating that the model has an ARCH effect. There is conditional heteroskedasticity and the variance changes over time. This result reflects the riskiness of yield movements.

Considering that the distribution has non-normality, a GARCH (1,1) model based on normal cloth is considered in building the model to compare the model effect.

Table 4. GARCH (1, 1) based on normal distribution

Variable	Estimate	Std. Error	t value	P-value
ω	0.000069	0.000026	2.682590	0.007305
alpha	0.126766	0.035503	3.570551	0.000356
Beta	0.840862	0.028336	29.674545	0.000000

As can be seen from the GARCH (1, 1) model based on normal distribution (Table 4), the ARCH term coefficient α is 0.126766 and the GARCH term coefficient β is 0.840862. The P-values corresponding to the parameters indicate the significance of the parameters, and the model constraints satisfied by the sum of the ARCH term coefficient α, and the GARCH term coefficient β ensures the stability of the model application. The sum of the two coefficients is close to 1, indicating the long memory of the price index. The value of β indicates the impact of past return fluctuations on current return fluctuations, with larger values indicating a stronger effect. The value of α indicates the sensitivity of the market to new information, with larger values indicating a more sensitive reflection. Therefore, the volatility is mainly affected by perturbations in past information, with little impact from new information.

In order to further verify whether the ARCH effect still exists after the model is constructed, another ARCH-LM test is conducted on the basis of this model.

The test results (Table 5) show that the original hypothesis of "the series does not have ARCH effect" is accepted, indicating that there is no ARCH effect in the model. The ARCH effect does not exist after the ARCH model is fitted, and the conditional heteroskedasticity problem is solved. The useful information of the sample data is extracted, and the ARMA-GARCH (1,1) based on the Normal distribution is established. GARCH (1,1) is reliable. It is reasonable to construct a GARCH model to characterize the return movements of the variables.

5. Conclusion

External shocks have a large impact on the price of Bitcoin and an equally large impact on the sustainability of its rise and fall. External shocks in this case are mainly policies enacted regarding bitcoin, with favourable policies (state support) increasing the price of bitcoin and negative policies (state bans) decreasing the price of bitcoin.

Table 5. ARCH-LM test

Series	Lagging order	Chi-square	P-value
Res	5	2.5738	0.7653
	10	3.3723	0.9713
	15	4.2046	0.997
	20	5.165	0.9996

Both a ten year view of Bitcoin data and a graph of its log-return series, or the coefficients in a GARCH model, show that Bitcoin's price is very volatile and changes very rapidly. In particular, the price of bitcoin reacts strongly to an external shock.

The very high price volatility of bitcoin, the existence of a cap on its size and its fungibility have led most investors to engage in bitcoin trading due to a follow-the-leader or speculative mindset [9]. Moreover, the lack of regulatory instruments that Bitcoin has been subjected since its existence has led to Bitcoin being traded in a form that is actually a speculative financial instrument [10].

The bubble in the price of Bitcoin is a very serious problem, meaning that the price has climbed sharply and the price has deviated from the value. In fact, the deviation is the bubble. Bitcoin is created as a result of public demand for decentralisation, but Bitcoin is not backed by a real economy, which leads to the conclusion that the soaring price of Bitcoin is in fact inflated. Because the price of Bitcoin has fallen so much in the face of negative news, it is even more clear that bitcoin is seriously bubbly.

References

[1] R. F. Engle, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica, vol. 50, 1982, pp. 987–1007. DOI:10.2307/1912773.

[2] D. G. Baura, K. Hong, A. D. Lee, Medium of exchange or speculative assets? Journal of International Financial Markets, Institutions and Money. vol. 54, 2018, pp.177–189. DOI:10.1016/j.intfin.2017.12.004.

[3] T. Bollerslev, E. Ghysels, Periodic Autoregressive Conditional Heteroscedasticity, Journal of Business & Economic Statistics, vol. 14 , 2016, pp.139–150.

[4] A. H. Aarhus, M. Peter, D. Erik, What can explain the price, volatility and trading volume of Bitcoin? Finance Research Letters, vol. 29, 2019, pp. 255–265. DOI:10.1016/j.frl.2018.08.010.

[5] A. Urquhart, The inefficiency of Bitcoin, Economics Letters, vol. 148, 2016, pp. 80–82. DOI:10.1016/j.econlet.2016.09.019.

[6] E. Bouri, A. H. Dyhrberg, On the return-volatility relationship in the Bitcoin market around the price crash of 2013. Economics: The OpenAccess, Open-Assessment E-Journal, vol. 11, 2017, pp. 1–17. DOI:10.5018/economics-ejournal.ja.2017-2.

[7] C. Conrad, A. Custovic, E. Ghysels, Long-and Short-Term Cryptocurrency Volatility Components: A GARCH-MIDAS Analysis, Journal of Risk and Financial Management, vol.11, 2018, pp.1–12. DOI:10.3390/jrfm11020023.

[8] J. Chu, S. Chan, S. Nadarajah, GARCH Modelling of Cryptocurrencies, Journal of Risk and Financial Management, vol.10, 2017, pp.17-25. DOI :10.3390/jrfm10040017.

[9] S. Corbet, B. M. Lucey, L. Yarovaya, Cryptocurrencies as a financial asset: A systematic analysis, International Review of Financial Analysis, 2018. DOI:10.1016/j.irfa.2018.09.003.

[10] L. Catania, S. Grassi, Modelling Crypto-Currencies Financial Time-Series. 2017, DOI:10.2139/ssrn.3028486.