Modelling the volatility of Bitcoin returns using Nonparametric GARCH models

Mestiri, Sami

13 December 2021

Modelling the volatility of Bitcoin returns using Nonparametric GARCH models

Sami Mestiri [1]

Research Unit EAS-Mahdia
Faculty of Science Mahdia Management and Economics,
University of Monastir, Tunisia.

Abstract : Bitcoin has received a lot of attention from both investors and analysts, as it forms the highest market capitalization in the cryptocurrency market. The use of parametric GARCH models to characterise the volatility of Bitcoin returns is widely observed in the empirical literature. In this paper, we consider an alternative approach involving non-parametric method to model and forecast Bitcoin return volatility. We show that the out-of-sample volatility forecast of the non-parametric GARCH model yields superior performance relative to an extensive class of parametric GARCH models. The improvement in forecasting accuracy of Bitcoin return volatility based on the non-parametric GARCH model suggests that this method offers an attractive and viable alternative to the commonly used parametric GARCH models.
JEL codes : C01

Keywords : Bitcoin ; volatility ; GARCH ; Nonparametric ; Forecasting.

1 Introduction

The cryptocurrency market continues to be a potential source of financial instability and its impact on the financial market still remains uncertain. Different from other financial assets which are regularized, there is no formal regulation for cryptocurrencies. Cryptocurrencies also differ significantly from other financial assets on the financial market and thus creates great prospects for investors and market players in terms of portfolio analysis, risk management and even consumer sentiment analysis.

In the cryptocurrency market, volatility modelling is important in measuring the riskiness of an investment. Volatility can be define as a measure of the dispersion in a probability density. Market players and investors are therefore interested in accurate estimation of volatility in the cryptocurrency market. This is as a result of the correlation between volatility and returns on investment. It is notable that volatility is not directly observable and as a result there is increasing need for efficient model that can capture the price volatility in the cryptocurrency market. Estimating the volatility of Bitcoin is very

1. Email :mestirisami2007@gmail.com

crucial since Bitcoin has the highest market capitalization in the cryptocurrency market.

A few studies have already been conducted on the financial and statistical characteristics of Bitcoin. One group of economists has been focusing on price discovery in the Bitcoin market, for example, Brandvold et al. (2015) and Bouoiyour et al. (2016) reveal some lead-lag relationship between Bitcoin prices, transactions use, and investors' attractiveness. Other studies also show that Bitcoin price is subject to unique factors which are substantially different from those affecting conventional, financial assets, such as internet search, information on google trends, and word-of-mouth information on social media.In fact, as Bitcoin is mainly used and viewed as an asset rather than a currency, and the Bitcoin market is currently highly speculative, and more volatile and susceptible to speculative bubbles than other currencies. Moreover, the presence of long memory and persistent volatility justifies the application of GARCH-type models.

The purpose of this paper is to utilize time series techniques to predict the future returns and prices of Bitcoin. At the same time, we want to examine the effectiveness of the popular GARCH model in economics and financial world. As Bitcoin gradually has had a place in the financial markets and in portfolio management, time series analysis is a useful tool to study the characteristics of Bitcoin prices and returns, and extract meaningful statistics in order to predict future values of the series.

In this paper, we will leave the functional form of the variance process as unspecified and attempt to estimate it as an additive nonparametric mean. We show that the nonparametric model can capture the leverage effect from the negative news and outperform two of the parametric GARCH family models most commonly considered. In their paper, Tjøstheim and Auestad (1994) worked the possibility of identifying nonlinear time series models using nonparametric methods. Härdle and Chen (1995) present a selective review of the approaches that based on nonparametric model building procedure in time series analysis. They point that nonlinear and nonparametric time series analysis is useful in order to deal with the limitations of the ARMA models with constant mean. Härdle, et al., (1997) review some developments in modern nonparametric techniques for time series analysis.

Engle and Gonzalez- Rivera (1991) addresses semi-parametric ARCH model by introducing a more efficient estimator based on a nonparametric estimated density. They also evaluate the loss of efficiency of the quasi- maximum likelihood estimator, which falsely assumes normality. Buhlmann and McNeil (2002) proposed a nonparametric approach to GARCH modeling. Hou and Suardi (2012) considered Buhlmann and McNeil (2002)'s nonparametric approach to model and forecast crude oil price return volatility. They use 4845 daily observations on crude oil spot prices from West Texas Intermediate, from 6 January 1992 to 30 July 2010, in their application. According to their results on forecasting accuracy, the nonparametric GARCH model has superior performance to parametric GARCH models. They prefer their nonparametric approaches because of the non normality of distribution of oil prices.

Another important reason in the development of nonparametric models is the lag selection procedure. The usual nonparametric models have less than satisfactory performance when dealing with more than one lag especially in the curse of dimensionality case. Alternative lag selection criteria have been studied for nonlinear autoregressive processes. Tjøstheim and Auestad (1994) mention heteroscedasti city in financial returns and propose to use a nonparametric version of the final prediction error (FPE). Tschernig and Yang (2000) derived a nonparametric version of the Final Prediction Error for lag selection in nonlinear autoregressive time series under very general conditions including heteroscedasticity. Wang et al. (2012) proposed a new efficient semi-parametric GARCH modeling of volatility by taking account lag selection procedure.

The rest of the paper is organised as follows. Section 2 describes the parametric and nonparametric GARCH models. Section 3 provides the forecast methodology. Section 4 provides a summary statistics of the data and the empirical results. Section 5 concludes.

2 Econometric models

2.1 GARCH model

The GARCH model of Bollerslev (1986) is the most widely used model for the volatility estimation.The GARCH models have been very successful in the literature because of their simple specification and easy interpretability. As pointed out by Bera and Higgins (1993), most of the applied financial works show that GARCH (1,1) provides a flexible and parsimonious approximation to the conditional variance dynamics and is capable of representing the majority of financial series. The GARCH (1,1) model is written as,

$$R_t = \mu + \varepsilon_t, \qquad with \qquad \varepsilon_t = \sigma_t.z_t; \qquad z_t \sim N(0,1)$$

The equation for the conditional variance of the residuals is defined as :

$$\sigma_t^2 = \alpha_0 + \alpha_1.\varepsilon_{t-1}^2 + \beta_1.\sigma_{t-1}^2 \tag{1}$$

Where $\alpha_0 > 0$,$\alpha_1 \geq 0$ and $\beta_1 \geq 0$ are constants and the ARCH(1) model corresponds to $\beta_1 = 0$. The constraint $\alpha_1 + \beta_1 < 1$ implies that the unconditional variance of the return series ε is finite and the conditional variance σ_t^2 evolves over time. It also provides the necessary and sufficient condition for the stochastic process $\sigma_t; t \in Z$ to be a unique strictly stationary process with $E(\sigma_t^2) < \infty$.
Two key properties can be noted from (1). First, a large ε_{t-1}^2 or σ_{t-1}^2 gives rise to a large

σ_t^2 and this generates the volatility clustering that is commonly known in financial time series. Second, the tail distribution is thicker than that of a normal distribution.

2.2 EGARCH Model

Despite their popularity, ARCH and GARCH models suffer from several weaknesses and drawbacks. Nelson (1991) criticized the GARCH models in three aspects : First, parameters are restricted to be positive at every time point ; Second, it fails to accommodate asymmetry effect (or leverage effect) ; and Third, measuring the persistence of the shocks on volatility is difficult. Nelson (1991) proposed the exponential GARCH (EGARCH) that accommodates the drawbacks of a standard GARCH model. The the first-order EGARCH (or EGARCH(1,1)) process specifies the model as

$$R_t = \mu + \varepsilon_t, \qquad with \qquad \varepsilon_t = \sigma_t.z_t; \qquad z_t \sim N(0,1)$$

The equation for the conditional variance of the residuals is defined as :

$$\log(\sigma_t^2) = \alpha_0 + \alpha_1.g(\varepsilon_{t-1}) + \beta_1.\log(\sigma_{t-1}^2) \tag{2}$$

Where ε_t follows the normal law is a weak white noise and the function $g(.)$ verified.

$$g(\varepsilon_{t-1}) = \alpha.\varepsilon_{t-1} + \gamma(|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)) \tag{3}$$

Here the coefficient γ signifies the leverage effect of shocks on the volatility. The key advantage of the EGARCH model is that the positive restrictions are not needed to be imposed on the variance coefficients. The coefficients γ need to be negative for evidence of asymmetric effects.

2.3 GJR-GARCH Model

In the simple GARCH (1,1) approach good news and bad news, i.e. positive and negative shocks, have the same impact on the conditional variance. Many studies have found evidence of asymmetry in stock price behavior, i.e., negative surprises seem to increase volatility more than positive surprises. To allow asymmetric effects in the volatility, Glosten et al. (1993) add an additional term in the conditional variance and formulate the so called GJR model. The GJR (1,1) is specified as follows,

$$R_t = \mu + \varepsilon_t, \qquad with \qquad \varepsilon_t = \sigma_t.z_t; \qquad z_t \sim N(0,1)$$

The equation for the conditional variance of the residuals is defined as :

$$\sigma_t^2 = \alpha_0 + \alpha_1.\varepsilon_{t-1}^2 + \gamma.(I_{\varepsilon_{t-1}<0}.\varepsilon_{t-1}^2) + \beta_1.\sigma_{t-1}^2 \tag{4}$$

Where z_t denotes a weak white noise of zero mean and constant variance over time, and the coefficients α_1, β_1 and γ are real parameters et $I_{\varepsilon_{t-1}<0}$ denotes the indicator function such that

$$\begin{aligned} I_{\varepsilon_{t-1}<0} &= 1 \qquad si \quad \varepsilon_{t-1} < 0 \\ &= 0 \qquad sinon. \end{aligned}$$

The structure of this model indicates that a positive ε_{t-1} contributes $\alpha_1.\varepsilon_{t-1}^2$ to σ_t, whereas a negative ε_{t-1} has a larger impact of $(\alpha_1 + \gamma).\varepsilon_{t-1}^2$ with $\gamma > 0$. Therefore, if parameters γ is significantly positive, then negative innovations generate more volatility than positive innovations of equal magnitude. The main feature of this model is that a negative shock has a larger impact than a positive shock and hence, it captures the leverage effect. Like the GARCH model, the GJR-GARCH model captures the volatility clustering. Also, it can be shown that the unconditional distribution presents excess kurtosis even under the Gaussian distribution.

2.4 Nonparametric ARCH Models

The starting point of the data generating process of a strictly stationary discrete-time stochastic process R_t defined on some probability space is the general univariate non-linear stochastic regression model given by

$$R_t = m(R_{t-1}, \ldots, R_{t-p}) + \sigma(R_{t-1}, \ldots, R_{t-p})\epsilon_t, \quad t = 1, ..., T \tag{5}$$

Where $m(R_{t-1}, ..., R_{t-p}) = E(R_t/R_{t-1} = r_1, \ldots, R_{t-p} = r_p)$ is the nonlinear autoregressive conditional mean (smooth) function, $\sigma^2(R_{t-1}, \ldots, R_{t-p}) = Var(R_t/R_{t-1} = r_1, \ldots, R_{t-p} = r_p)$ represents the nonlinear autoregressive conditional variance (smooth) function, and ϵ_t is an independent and identically distributed (i,i,d) sequence of random variables with $E(\epsilon_t/R_{t-1}, \ldots, R_{t-p}) = 0$,$Var(\epsilon_t/R_{t-1}, \ldots, R_{t-p}) = 1$ and independent of $R_{t-1}, \ldots, R_{t-p}$.

The model (5) is known as the Nonparametric Autoregressive Conditional Heteroscedastic NARCH-model see Fan and Yao (1998).This model is the most flexible nomparametric time series model because it does not impose any (parametric) particular form on the conditional mean and volatility functions. However,due to the well-known "curse of dimensionality" problem, to assume a certain level of structure on the conditional function $m(.)$ and $\sigma(.)$. In this current study, we employ the first-order conditional heteroscedastic nonlinear autoregressive NARCH (1,1) model

$$R_t = m(R_{t-1}) + \sigma(R_{t-1})\epsilon_t, \quad t = 2, ..., T \tag{6}$$

where R_t are observed and depend on R_{t-1} with lag 1, $m(R_{t-1})$ is the trend function of NARCH-model, $\sigma(R_{t-1})$ is the heteroscedastic function of NARCH-model, and ϵ_t denotes a random variable in the error term, with mean zero and variance one. Following Fan and Yao (1998) if R_t is a stationary process, the conditional variance function can be decomposed as

$$\sigma^2 = E(R_t^2/R_{t-1} = r) - (E(R_t/R_{t-1} = r)^2 = g(r) - m(r)^2$$

such that the conditional variance estimate is based on the nonparametric estimation of $g(r)$ and $m(r)$ given by $\hat{\sigma}^2(r) = \hat{g}(r) - \hat{m}(r)^2$.
A way to obtain estimates of functions $g(r)$ and $m(r)$ is by applying the popular Nadaraya-Watson estimator given by :

$$\hat{m}(R_{t-1}) = \frac{\sum_{t=2}^{T} K(R_{t-1} - r)/h)R_t}{\sum_{t=2}^{T} K(R_{t-1} - r)/h)}$$

$$\hat{g}(R_{t-1}) = \frac{\sum_{t=2}^{T} K(R_{t-1} - r)/h)R_t^2}{\sum_{t=2}^{T} K(R_{t-1} - r)/h)}$$

The function $K(\cdot)$ is usually a symmetric probability density and examples of commonly used kernel functions are the Gaussian kernel $K(t) = (\sqrt{2\pi})^{-1}\exp(-t^2/2)$ and the *Epanechnikov* kernel $K(t) = \max\{\frac{3}{4}(1-t^2), 0\}$ and h is bandwidth parameter (smoothing parameter) .

2.5 Nonparametric GARCH Models

We propose to apply this nonparametric method that does not require the specification of the functional form of the volatility and that does not regard to the distributional form of the innovation distribution. Moreover, nonparametric GARCH models allow the conditional covariance matrix of the dependent variables to follow a flexible dynamic structure. The stationary stochastic process $\{\varepsilon_t; 1 < t < n\}$ has the nonparametric GARCH(1,1) form given in (Bühlmannand McNeill, 2002) :

$$R_t = \mu + \varepsilon_t, \qquad with \qquad \varepsilon_t = \sigma_t.z_t; \qquad z_t \sim N(0,1)$$

The equation for the conditional variance of the residuals is defined as :

$$\sigma_t^2 = f(\varepsilon_{t-1}, \sigma_{t-1}^2) \tag{7}$$

In the nonparametric GARCH approach the exact form of f is unspecified and is estimated using a bivariate nonparametric smoothing technique which is less sensitive to model misspecification such as neglected asymmetric volatility.
Assuming that $\{\varepsilon_t; 1 < t < n\}$ coming from a process satisfying (7), the estimation of a nonparametric GARCH model is applied with the following steps as proposed in (Bühlmann and McNeill, 2002) :
1- Firstly, at the m=0 step, an estimate of volatility $\{\hat{\sigma}_{t,0}^2; 1 < t < n\}$ is obtained by fitting an ordinary parametric GARCH(1,1). Then the predictions from the GARCH(1,1) model are extracted which gives the $\{\hat{\sigma}_{t,0}^2; 1 < t < n\}$ estimates for the m=0 step of the algorithm. Since the first value is not estimated in returns, it is set as equal to the mean.
2- In the m=1 step, ε_t^2 is regressed with a nonparametric smoothing technique against ε_{t-1} and $\hat{\sigma}_{t-1,0}^2$ which are obtained from the parametric GARCH(1,1). The squared values

of the residuals are obtained from the ARIMA model and the lagged values are the first lag of the residuals of the ARIMA model. The estimated variance of the return series is obtained from the previous step of the algorithm.
3- At the m'th step, the algorithm is repeated and the $\hat{\sigma}^2_{t-1,m}$ is estimated by ε_{t-1} and $\hat{\sigma}^2_{t-1,m-1}$.

3 Forecast performance measures

While there are several different measurements for evaluating volatility forecasting performances,the mean square error (MSE) and the mean absolute error (MAE) are used in this study. When the true underlying volatility process is unobservable, we adopt the suggestion to use $(\sigma_t^2 = R_t - \bar{R})^2$ as a proxy for latent volatility in this scenario. The MAE and MSE for n step ahead forecast are defined as follows :

$$MSE = \frac{1}{N}\sum_{t=1}^{N}((R_{t+n} - \bar{R})^2 - \hat{h}_t(n))^2$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|(R_{t+n} - \bar{R})^2 - \hat{h}_t(n)|$$

where R_{t+n} : the return over horizon n steps ahead at current time t ,
$\bar{R}$: the mean of return ,
$\hat{h}_t(n)$: the forecasted conditional variance over horizon n steps ahead at current time t.

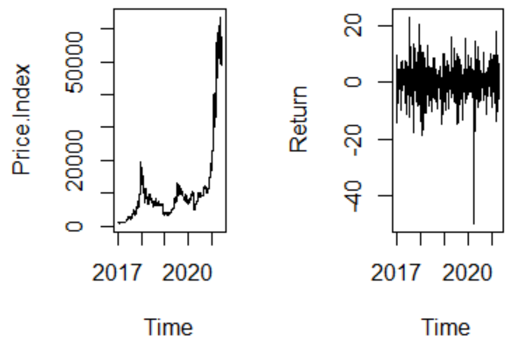
4 Empirical Results

4.1 Data description

In this study, we apply the previously described different parametric and nonparametric GARCH to estimate variance function for Bitcoin returns. The data used are the closing price was selected as the price of Bitcoin because it reflected all the activities of Bitcoin for each trading day. Historical daily closing price of Bitcoin was extracted from 02/01/2017 to 30/04/2021 at https ://finance.yahoo.com/ and consisted of 1580 trading days.

In order to assess and compare the predictive performance of the Nonparametric GARCH model with various parametric models, the data is further divided into an in-sample group (from January 2, 2017 to April 30, 2021) and an out-of-sample group (from May 1, 2021 to June 10, 2021). The whole sample has 1540 observations and the last 40 are used for out-of-sample forecasts.

FIGURE 1 – Closing price and return series of Bitcoin.



Titre	Mean	Min	Max	St.dev	Skew	Kurt	JB
Bitcoin	0.2636	-49.7278	22.7602	4.338	-0.968	16.626	12108.26

TABLE 1 – Descriptive statistics of return series of Bitcoin

Now, assume P_t and P_{t-1} represents the current day and previous day price of Bitcoin, then the return series/log returns (R_t) and multiplied by 100 as follows

$$R_t = ln(\frac{P_t}{P_{t-1}}) * 100.$$

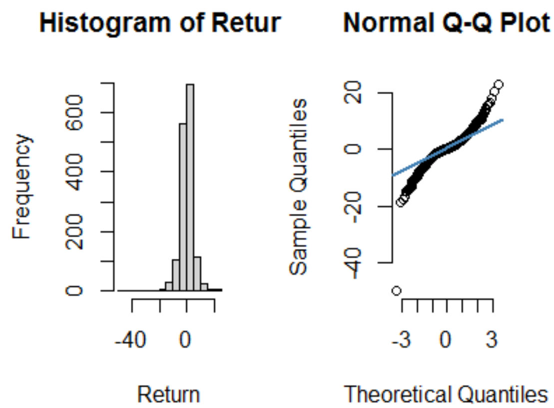
Table 1 presents the descriptive statistics of the return series of Bitcoin with the Jarque-Bera test for normality which is calculated as

$$JB = T.(\frac{skew^2}{6} + \frac{(kurt-3)^2}{24})$$

where T is the sample size, skew and kurt are the sample skewness and kurtosis respectively.
Under the null that the data is normal iid, JB is asymptotically distributed as chi-square with 2 degrees of freedom. The tests rejected the normality at 5% significance level. The return series of Bitcoin is negatively skewed. This indicates that the returns of Bitcoin is non-symmetric. The negative value of the skewness indicates that the distribution of Bitcoin return series is skewed to the left. The positive excess kurtosis (16.626) indicates that the returns are leptokurtic. That is, the returns series has a fatty tail.

Figure 1 shows the time series plot of Bitcoin price (left Figure) and the return series (right Figure) of Bitcoin for the time period. Figure 2 is the histogram and the normal quantile-quantile (q-q) plot of the return series for the same time period.

FIGURE 2 – Histogram and normal q-q plot of return series of Bitcoin.



	Ljung box test	LM test	ADF test
P-value	0.0019	0.00152	0.01

TABLE 2 – Test of Auto Regressive Conditional Heteroscedasticity (ARCH) effect.

The Augmented Dickey Fuller (ADF) test (Dickey and Fuller, 1979) is used to test for stationarity. From Table 2, the null hypothesis of statonarity is accepted at 5% -level of significance. Hence, there is no need to difference the return series. To apply GARCH models to the Bitcoin returns series, the presence of stationarity and ARCH effects in the residual return series are tested. The Ljung-box and Lagrange multiplier (LM) test (Engle, 2001) are used to test for the presence of ARCH effects in the data.The Ljung-box and LM test are presented in Table 2. From the Ljung box test, the null hypothesis of "no autocorrelation" in the squared residuals is rejected at 5% significance level. That is, there is dependency in the squared returns series of Bitcoin. Using the LM test, the null hypothesis of "no ARCH effects" is rejected at 5% significance level. From the Ljung box and LM test, it can be concluded that the volatility ARCH effect is very much present in the return series. Hence, the GARCH models are used to model the returns series data.

4.2 Estimation results

The data descriptive statistics indicate that an appropriate model of Bitcoin returns volatility should account for its time-varying nature and the departure from normality in Bitcoin returns distribution. All estimations and computations are done in R Statistical Environment (R, 2008) using " rugarch" R package developed by and Ghalanos (2013) the "KernSmooth" R package developed by Wand and Ripley (2007). The parametric GARCH models are estimated with the Bollerslev and Wooldridge (1992) quasi-maximum

	GARCH		EGARCH		GJR-GARCH	
	Normal	Student-t	Normal	Student-t	Normal	Student-t
μ	0.327 (3.385)	0.264 (3.870)	0.282 (2.848)	0.237 (3.645)	0.257 (2.633)	0.267 (3.916)
ω	1.243 (5.464)	0.316 (1.921)	4.973 (1.107)	0.046 (1.582)	1.402 (5.214)	0.258 (1.474)
α	0.128 (6.290)	0.113 (5.442)	0.114 (4.369)	0.1357 (5.503)	0.084 (4.612)	0.120 (5.190)
β	0.818 (35.613)	0.885 (39.055)	0.7545 (16.155)	0.910 (58.616)	0.802 (31.269)	0.896 (33.643)
γ			0.1858 (1.313)	-0.1402 (-1.574)	0.101 (3.257)	-0.0369 (-1.384)
$log.likelihood$	-4338.704	-4135.845	-4329.967	-4123.559	-4331.529	-4134.896
AIC	5.661	5.398	5.653	5.385	5.653	5.398
BIC	5.67	5.416	5.674	5.409	5.671	5.419
$Q(20)$	4.889	5.2496	5.8857	5.176	5.278	4.9664

Table 3 – In-sample estimations of the GARCH, EGARCH, and GJR models

likelihood method which gives robust standard errors.

We first fit the series from 02/01/2017 to 30/04/2021 with the standard GARCH(1,1) model. Considering the existence of the asymmetry effects in the cyrpto markets, we also fit the data with the EGARCH and GJR models. For all these models, the innovations are assumed to be both Gaussian,and student-t distributed. The estimated parameters and Ljung-Box Q-statistics tests of the standardized residuals are presented in Table 3.

Table 3 shows the results of the maximum likelihood estimate (MLE) of GARCH(1,1), EGARCH(1,1), and GJR-GARCH(1,1) models for Bitcoin returns using Normal and student t-distribution. From the table, the log-likelihood value (-4123.559) is maximum for GJR-GARCH(1,1) model. The values of the two information criterions (AIC= 5.385 , BIC=5.409) of EGARCH(1,1) are minimum as compared to GARCH(1,1)-t and GJR.GARCH(1,1)-t. These results indicate thatEGARCH(1,1)-t model is the optimal model to describe the volatility of the return series of Bitcoin.

Note that all parameters of the conditional volatility are significant at the 5% significance level. The coefficient of lagged variance β shows very high volatility persistence. The sum of α and β from the GARCH model are close to 1, which supports the evidence of volatility clustering. The P-values of Ljung-Box Q-statistic test at the lag 20 of standardized residual series from all models fail to suggest the autocorrelation at a 5% significance level. Thus all models appear to be adequate in describing the linear dependence in the return and volatility series.

The estimated value of the leverage parameters γ of the EGARCH and GJR models

Model	Law	MSE	MAE
GARCH	Normal	456.52	10.65
	Student-t	525.85	10.69
EGARCH	Normal	433.8	10.63
	Student-t	449.32	10.97
GJR-GARCH	Normal	450.65	10.65
	Student-t	511.81	10.68
N-ARCH	Normal	433.30	10.3
N-GARCH	Normal	430.73	10.29

TABLE 4 – Goodness-of-fit for out-of-sample forecasts

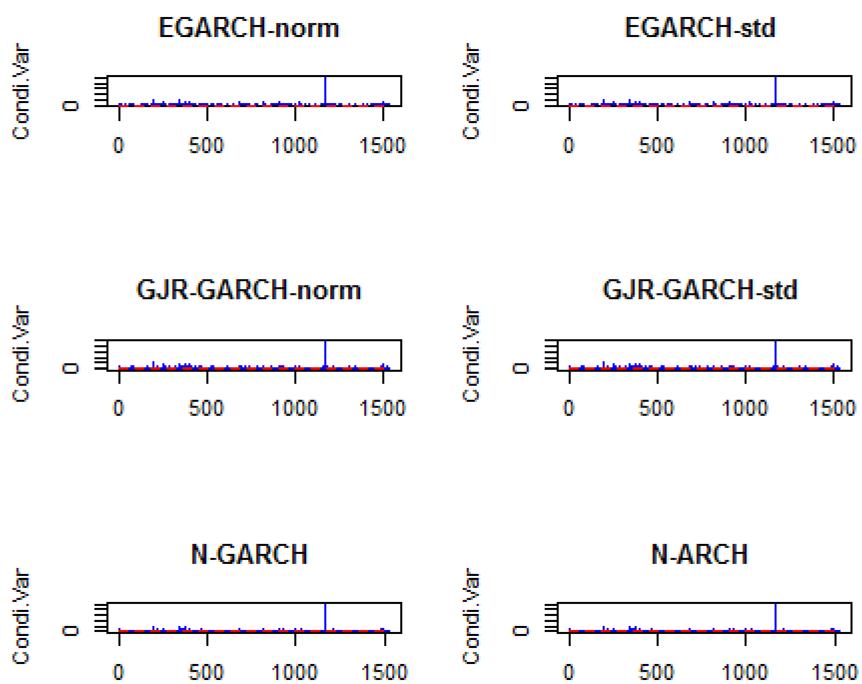
with Gaussian/t distributed innovations is : $0.1858/-0.1402$ and $0.10/-0.0369$, respectively. All these parameters are significant at the 5% level with the exception of the γ from the EGARCH model with Gaussian errors. The significance of the parameters indicates the existence of asymmetry effect i.e., bad news (negative shock) has a larger impact on return volatility than good news (positive shock). It is also worth noting that the leverage effect estimated from models fitted with t distributed innovation is higher than the ones with normal distributed innovations. The existence of the asymmetry effect as in other mature stock markets in the world may be a positive sign for market efficiency and completeness.

4.3 Forecast results

The performance of the out-of-sample volatility forecasts of various models are summarized in Table 4. It is clear from this table that among different models, the GARCH model performs the worst according to all goodness-of-fit measures and the N-GARCH model performs the best in delivering the lowest forecast error. Compared with the GARCH model, the EGARCH model improves the volatility estimation by capturing the leverage effects. For the GJR model, it slightly improves the result from the GARCH estimation. This is perhaps not surprising because the asymmetric effect in Bitcoin market is not as strong . However, this may indicate that the EGARCH model can capture more leverage effect than the GJR model. When looking at the N-ARCH model, we observe a significant improvement of the N-ARCH model compared with the EGARCH model with Gaussian errors. In addition, all loss functions from the N-GARCH model do not differ from the ones with N-ARCH model.

To demonstrate the importance of our results and the application of the N-GARCH model in practice, we calculate the forecasted return intervals which are based on one day ahead out-of-sample forecasts. The out-of-sample period is from May 1, 2021 to June 10, 2021. The realized volatility is calculated as $\hat{\sigma}_t = \sqrt{(R_t - \bar{R})^2}$, where R_t are the log

Figure 3 – The estimated volatility from in-sample volatility estimation



return at time t. Figure 3 plots the volatility for the in-sample period. The red lines are the realized volatility , while the blue lines are the estimated volatility. The three volatility plots are for the N-GARCH model , EGARCH, and GJR models. Again, it can be seen from these plots that the N-GARCH model performs better than the EGARCH and the GJR models in capturing the rise and fall movements of return volatility.

5 Conclusion

GARCH modeling builds on advances in the understanding and modeling of volatility. It takes into account excess kurtosis (i.e. fat tail behavior) and volatility clustering, two important characteristics of financial time series, which are also observable in the Bitcoin case. It's theoretically able to provide accurate forecasts of variances and covariances of returns through modeling time-varying conditional variances. As a consequence, GARCH models have become quite popular in diverse fields as risk management, portfolio management and asset allocation, option pricing, foreign exchange, and the term structure of interest rates.

In this paper, we intend to predict the future prices of Bitcoin, one of the most widely used and traded cryptocurrency, and study the predictive power of GARCH model on the Bitcoin return/price series. From the predicted results, we have realized that although GARCH models are useful across a wide range of financial and economical applications, they are not a quite effective and suitable model candidate in studying the Bitcoin return/price series. One of the main reasons is that GARCH models are parametric specifications that operate best under relatively stable market conditions. Although GARCH is explicitly designed to model timevarying conditional variances, GARCH models often fail to capture highly irregular phenomena, including wild market fluctuations (e.g., crashes and subsequent rebounds), and other highly unanticipated events that can lead to significant structural change, which are exactly what has been going on in the Bitcoin market recently.

In this paper an application of a nonparametric GARCH model to Bitcoin return volatility has been proposed. The nonparametric smoothing technique uses a general additive function of lagged innovations and volatilities to estimate the unobserved diffusion process. Although the volatility model is estimated nonparametrically, its specification resembles the widely used parametric GARCH models of Bitcoin return volatility. The empirical applications to daily Bitcoin prices document significant improvement in the out-of-sample predictive power of the non- parametric model over an extensive class of GARCH models. Careful attention is paid to the use of loss functions that are robust to the squared returns proxy which is used to measure realised volatility. The superiority in the out-of-sample volatility predictive performance of the nonparametric GARCH model is further verified. The results suggest that the nonparametric GARCH model with its improved forecasting accuracy over the parametric counterparts is worthy to be considered as a useful alternative method of modelling Bitcoin return volatility. For future research, we intend to apply this technique to Value-at-Risk computations and illustrate the potential benefits that can be derived from applying this model in the context of hedging against extreme price fluctuation in cryptocurrency market.

References

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327

Bollerslev, T., Wooldridge, J., 1992. Quasi-maximum likelihood estimation and infer- ence in dynamic models with time-varying covariances. *Econ. Rev.* 11, 143-172.

Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. ARCH models. In Engle, R.F., McFadden, D. (Eds.), *Handbook of Econometrics, Amsterdam Elsevier Science,* pp. 2959–3038.

Brandvold, M., Molnár, P., Vagstad, K., and Valstad, O. C. A. 2015. Price discovery on Bitcoin exchanges. *Journal of International Financial Markets, Institutions and Money* 36 : 18– 35.

Bouoiyour, J., Selmi, R., Tiwari, A.K. and Olayeni, O.R. 2016. What drives Bitcoin price ?. *Economics Bulletin* 36(2) : 843-850.

Bossaerts P., Härdle W., Hafner C. (1996) Foreign Exchange Rates Have Surprising Volatility. In : Robinson P.M., Rosenblatt M. (eds) Athens Conference on Applied Probability and Time Series Analysis. Lecture Notes in Statistics, vol 115. Springer, New York, NY.

Bühlmann, P., McNeil, A., 2002. An algorithm for nonparametric GARCH modelling. *Computational Statistics and Data Analysis* 40, 665–683.

Engle R ,2001, Garch 101 : The use of arch/garch models in applied econometrics. *J Econ Perspect*15 : 157–168.

Engle, R.F., Bollerslev, T., 1986. Modelling the persistence of conditional variances. *Econ. Rev.* 5, 1–50.

Engle, R.F., Gonzalez-Rivera, G. 1991. Semiparametric ARCH models. *Journal of Business and Economic Statistics*, 9(4), 345-359.

Fan, J, Yao, Q. 1998, Efficient estimation of conditional variance functions in stochastic regression, *Biometrika*, Volume 85, Issue 3, Pages 645–660,

Ghalanos. A .2011, rgarch : A package for flexible garch modelling in r. Version 1.89.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Finance* 48, 1779–1801.

Härdle, W., Chen, R. 1995. Nonparametric Time Series Analysis, a Selective Review with Examples. *Proceedings of the 50th Session of the ISI.*

Härdle, W., Lütkepohl, H., Chen, R. 1997.A Review of Nonparametric Time Series Analysis. *International Statistical Review/Revue Internationale de Statistique*, 65(1), 49-72.

Hou, A., Suardi, S., 2012. A Nonparametric GARCH Model Of Crude Oil Price Return Volatility. *Energy Economics*, 34(2), 618-626.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns : a new approach. *Econometrica* 59, 347–1313.

Pagan, A.R., Schwert, G.W., 1990. Alternative models for conditional stock volatility. *J. Econ.* 45, 267–290.
Tjøstheim, D., Auestad, B.H. 1994. Nonparametric Identification of Nonlinear Time Series : Projections. *Journal of the American Statistical Association,* 89(428), 1398-1409.
Tschernig, R., Yang, L. 2000. Nonparametric Lag Selection for Time Series. *Journal of Time Series Analysis*, 21(4), 457-487.
Wand, M., Ripley, B , 2007 . KernSmooth : Functions for kernel smoothing for Wand and Jones (1995). R package version 2.22-20.
Wang, L., Feng, C., Song, Q., Yang, L. 2012. Efficient Semiparametric GARCH Modeling Of Financial Volatility. *Statistica Sinica,* 22, 249-270.