WORKING PAPER SERIES: WPS/0058

Analysing the distribution properties of Bitcoin returns

Afees A. Salisu, Aviral Kumar Tiwari and Ibrahim D. Raheem

Cite as:

Salisu A. A., Tiwari A. K. and Raheem I. D. (2018): Analysing the distribution properties of Bitcoin returns - *Centre for Econometric and Allied Research, University of Ibadan Working Papers Series, CWPS 0058*

Analysing the distribution properties of Bitcoin returns

Afees A. Salisu[1], Aviral Kumar Tiwari[2], Ibrahim D. Raheem[3]

[1] Centre for Econometric & Allied Research, University of Ibadan.
Email: aa.salisu@cear.org.ng; Phone: +234(0) 8034711769.

[2] 2300 avenue des moulins, 3400 Montpellier, France Montpellier Business School, Montpellier, France.
Email: a.tiwari@montpellier-bs.com; Phone: +33 75 87 58 648.

[3] School of Economics, University of Kent, Canterbury.
Email: idr6@kent.ac.uk; Phone: +44(0)1227 824308.

Abstract

This study exploits several conditional heteroskedasticity models with various supported distributions in order to find the best distribution as well as the best GARCH-type model that may be used to model volatility of Bitcoin returns. Innovatively, the study is able to establish that pre-testing the residuals of Bitcoin returns for the best distribution can help to identify the appropriate distribution when modelling with GARCH-type models regardless of the data frequency.

Keywords: Bitcoin returns; GARCH-type models; Error distributions

1. Introduction

Bitcoin has recently received significant consideration from investors, speculators, and policymakers and it has the highest market share in terms of market capitalization and volume of trade (about 37%) over 1595 cryptocurrencies as of 17th May 2018[1]. The attractiveness towards Bitcoin in recent years has also awakened the curiosity of researchers with increasing attempts to analyse the various characteristics of its returns. For example, Baek and Elbeck (2015), Dyhrberg (2016) and Katsiampa (2017) analyse the Bitcoin price volatility; Urquhart (2016), Nadarajah and Chu (2017), Bariviera (2017), Jiang et al. (2017) and Tiwari et al. (2018) have tested for the efficient market hypothesis (EMH) and Khuntia and Pattanayak (2018) attempt to validate the adaptive market hypothesis (AMH), etc.

This study advances the literature on Bitcoin returns (particularly, we extend Baek and Elbeck, 2015; Dyhrberg, 2016; Katsiampa, 2017) by examining the distribution properties of the series using several GARCH-type models. Studies on Bitcoin returns have assumed different distributions without any empirical justification for their choice and this may have implications on its predictability. Thus, our contribution to the existing literature is four fold. First, we subject the Bitcoin return series to various alternative distributions about eleven of them and their comparative performance is evaluated. This helps us to determine the distribution that best fits the return series. Second, based on the underlying statistical features of the return series that suggest the presence of ARCH effects, each distribution is subjected to several GARCH-type models including those with asymmetries and nonlinearities in order to determine the best GARCH model for Bitcoin returns. Third, we further examine whether pre-testing for the distribution property of the Bitcoin returns matters in choosing appropriate distribution for the GARCH-type models. This involves two stages: (1) we specify a typical mean equation that ignores GARCH error for Bitcoin returns and thereafter we test for the appropriate distribution of its residuals; (2) We then investigate whether the best distribution for the GARCH-type models picks the best distribution obtained from the pre-test. This enables us to establish the need to pre-test for the distribution of the Bitcoin as a precondition for the choice of distribution when modelling and forecasting its return volatility. Fourth, the robustness of our results has been tested using daily, weekly and monthly datasets.

---

[1] https://coinmarketcap.com

Following the introduction, we organize the rest of the paper as follows. The methodology and data issues are highlighted in Section 2; the results are presented and discussed in Section 3 while the conclusion is rendered in Section 4.

2. Methodology, Data and Preliminary Analyses

The methodology can be partitioned into two: First the pre-testing for the distribution of Bitcoin returns and Second, the main GARCH analyses with different models and distributions. For consistency, we use the same set of distributions for both stages. Thus, we limit our distributions to those that are largely supported by most of the GARCH-type models particularly those considered in this paper.

The underlying specification for the first stage involving pre-testing of the Bitcoin returns residuals is given as $r_t = \alpha + \rho r_{t-1} + \varepsilon_t$ where $r_t = \log\left(p_t / p_{t-1}\right)$ and p_t is the Bitcoin price; we then fit ε_t with the following distributions: Gaussian; Student-t; Generalised Error; Modified Cauchy; Laplace; Logistic; Hansens Skew-t; Gram-Charlier expansion series with constant higher moments; Rayleigh; and Extreme Value Distribution Distribution Type 1.

For the second stage, we consider the following GARCH-type models whose choice is determined largely by the selected distributions and for convenience, we allow for first lag in the relevant variables both in the mean and variance equations. Thus, the mean equation common to all the selected GARCH-type models is ARMA (1,1): $r_t = \alpha + \rho r_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$ where $\rho \neq 0$; $\theta \neq 0$ and $\varepsilon_t = \sigma_t e_t$; but e_t follows different distributions as previously mentioned while σ_t is the conditional standard deviation. The corresponding variance equation of the latter varies for the different GARCH-type models based on the underlying characteristics for ARCH and GARCH terms but with a common lag combination (1,1) respectively. The considered GARCH-type models are: GARCH: $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$ where $\omega > 0,\ \alpha \geq 0,\ \beta \geq 0$; GJR-GARCH: $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I_{t-1} + \beta\sigma_{t-1}^2$ where $I_{t-1} = 1$ if $\varepsilon_{t-1} < 0$ and zero(0) if $\varepsilon_{t-1} \geq 0$; EGARCH: $\ln\sigma_t^2 = \omega + \alpha\left|u_{t-1} \big/ \sqrt{\sigma_{t-1}^2}\right| + \gamma\left(u_{t-1} \big/ \sqrt{\sigma_{t-1}^2}\right) + \beta\ln\sigma_{t-1}^2$; NARCH (Nonlinear ARCH): $\sigma_t^\delta = \omega + \alpha\left|\varepsilon_{t-1}\right|^\delta + \beta\sigma_{t-1}^\delta$; NGARCH (Nonlinear GARCH): $\sigma_t^\delta = \left(\omega + \alpha\left|\varepsilon_{t-1}\right|^\delta + \beta\sigma_{t-1}^\delta\right)^{1/\delta}$; AGARCH (Asymmetric GARCH): $\sigma_t^2 = \omega + \alpha\left(\varepsilon_{t-1} + \gamma\right)^2 + \beta\sigma_{t-1}^2$; APGARCH (Asymmetric

Power GARCH): $\sigma_t^{\delta} = \left(\omega + \alpha\left(\left|\varepsilon_{t-1}\right| - \gamma\varepsilon_{t-1}\right)^{\delta} + \beta\sigma_{t-1}^{\delta}\right)^{1/\delta}$ and NAGARCH (Nonlinear Asymmetric GARCH): $\sigma_t^2 = \omega + \alpha\left(\left(\varepsilon_{t-1}/\sigma_{t-1}\right) + \gamma\right)^2 + \beta\sigma_{t-1}^2$ where $-1 < \gamma < 1$.

The study has used daily, weekly and monthly data frequencies of Bitcoin for the period July 19, 2010 – May 03, 2018 and all the data are sourced from the Bloomberg terminal. Table 1 presents the descriptive statistics of Bitcoin returns. It could be deduced that the variable has an average mean value of 4.6216 and considered volatile, judging by its high value of standard deviation. The variable is negatively skewed and platykurtic.

[Insert Table 1 about here]

3. Results discussion and findings

About 10 distributions and 8 various GARCH Models are estimated in order to achieve the objective of this study as previously highlighted. Table 2 presents the results of the first stage that involves pre-testing the return series for the best distribution. For the want of space, only statistics for the Alkaike Information Criteria (AIC) are presented. The distribution with the least value of AIC is assumed to have the best fit. An overview of the table shows that the Generalized Error Distribution (GED) is the most preferred distribution, which is distantly followed by Cauchy and Hansens Skew-t distribution.

[Insert Table 2 about here]

Moving to the second stage, similar scenario as depicted in stage one also plays out in the second stage. Thus, our results suggest that GED is preferred whether for the residuals or any of the GARCH-type models considered. This confirms our hypothesis that pre-testing the residuals of relevant series for their distribution properties may help determine the appropriate distribution to use for GARCH modelling of such series. Table 3 has the results of the estimates for the various GARCH type models. Instructively, irrespective of the GARCH-type model, the GED is consistently chosen while the results of the EGARCH model obtained with this distribution seem to give the best fit for bitcoin returns.

As a robustness test, we further examine if these results are sensitive to the choice of data frequencies. As such, the analyses are being replicated for weekly and monthly data frequencies. It is instructive to infer that our results stand the test of changes in data frequencies. In other words, regardless of the choice of data frequency, the GED as well as the associated EGARCH model best fit Bitcoin returns while the results of EGARCH model as well as other GARCH-types obtained under different distributions other than GED are found to perform worse.

[Insert Table 3 about here]

4. Conclusions

This study sets out to add to the emerging literature on Bitcoin returns by justifying the need to understand the underlying distribution property of the series for improved GARCH model results. Previous studies have assumed different distributions in their empirical analyses of Bitcoin returns; we however show that the choice of distribution has implications on the predictability of the return series. Specifically, we demonstrate that pre-testing the residuals of the return series for the distribution properties helps to determine the appropriate distribution for GARCH analyses. Our results show that the GED chosen from the pre-testing also gives the best fit for the GARCH analyses irrespective of the GARCH-type. In addition, the EGARCH model associated with the GED is found to offer the best fit for Bitcoin returns. Our results are robust to alternative data frequencies.

References

Baek, C., Elbeck, M., 2015. Bitcoins as an investment or speculative vehicle? A first look. Applied Economics Letters, 22 (1), 30–34.

Bariviera, A.F., 2017. The inefficiency of Bitcoin revisited: A dynamic approach. Economics Letters 161, 1–4.

Bariviera, A.F., Basgall, M.J., Hasperué, W., Naiouf, M., 2017. Some stylized facts of the Bitcoin market. Physica A 484, 82–90.

Dyhrberg, A.H., 2016. Bitcoin, gold and the dollar - A GARCH volatility analysis. Finance Research Letters 16, 85–92.

Jiang, Y., Nie, H., Ruan, W., 2017. Time-varying long-term memory in Bitcoin market. Finance Research Letters, https://doi.org/10.1016/j.frl.2017.12.009.

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

Khuntia, S., Pattanayak, J.K., 2018. Adaptive market hypothesis and evolving predictability of bitcoin. Economics Letters 167, 26–28

Nadarajah, S., Chu, J., 2017. On the inefficiency of Bitcoin. Economics Letters 150, 6–9.

Tiwari, A.K., Jana, R.K., Das, D., Roubaud, D., 2018. Informational efficiency of Bitcoin-An extension. Economics Letters 163, 106–109.

Urquhart, A., 2016. The inefficiency of Bitcoin. Economics Letters 148, 80–82.

Table 1: Descriptive Statistics of Bitcoin

Mean	Std Dev	Min	Max	Skew.	Kurt.	Q Stat	Q^2 Stat	ARCH-LM
4.6216	2.9533	-2.9957	9.834	-0.688	2.835	1886[a]	29.423[a]	29.767[a]

Source: Authors' computation.

Note: [a] implies level of statistical significance at 1%. The reported values for the serial correlation are the Ljung-Box Q-statistics and ARCH-LM test F-statistics for the heteroscedasticity. We consider three different lag lengths (k) of 10, 20 and 36. We presented statistics for 36 lags.

Table 2: Residual Analysis of GARCH Effect

Type of Test	Daily	Weekly	Monthly
Gaussian Normal	-5113	-323	78
Student-T	-6315	-470	60
Generalised Error	-9392	-1180	-121
Cauchy	-6330	-448	74
Laplace	-6222	-466	61
Logistic	-5785	-410	63
Hansens Skew-t	-6323	-470	60
Gram-Charlier	-4414	-251	137
Rayleigh	-6003	-495	28
Extreme Value Distribution	-3887	-114	132

Source: Authors' computation.

Table 3: Distribution property of Bitcoin returns using various types of GARCH models

	GARCH	GJR	EGARCH	NARCH	NGARCH	AGARCH	APGARCH	NAGARCH
Generalized Error Distribution								
Daily	-10570	-10599	-10635[b]	-10599	-10458	-10594	-10475	-10435
Weekly	-1388	-1431	-1446[b]	-1392	-1384	-1401	-1420	-1384
Monthly	-127	-137	-195[b]	-135	-129	-136	-144	-129
Cauchy								
Daily	-6773	-6766	-6785	-6823	-6776	-6777	-6993[b]	-6698
Weekly	-493	-496	-495	-491	-487	-497[b]	-494	-488
Monthly	71	69	66[b]	71	73	69	69	71
Logistic								
Daily	-6695	-6700	-6719[b]	-6704	-6606	-6697	-6579	-6586
Weekly	-542	-558[b]	-556	-541	-532	-554	-551	-532
Monthly	54[b]	50	40[b]	54	54	47	46	48
Extreme Value Distribution								
Daily	-4891	-5316	-5517[b]	-5075	-4971	-5405	-5463	-5217
Weekly	-315	-367	-431[b]	-264	-267	-359	-400	-349
Monthly	133	101	60[b]	124	124	125	117	102
Gaussian Normal								
Daily	-6268	-6292	-6315	-6284	-6518[b]	-6287	-6166	-6138
Weekly	-508	-520	-536[b]	-507	-498	-520	-520	-499
Monthly	74[b]	64	28[b]	70	66	62	57	61
Gram-Gharlie								
Daily	-5882	-5315	-5198	-5238	-6367	-6848[b]	-6539	-4983
Weekly	-386	-343	-330	-185	-312	-431	-726[b]	-527
Monthly	43	100	73	14[b]	79	52	44	52
Hansen's Skew T								
Daily	-6874	-6876	-6965[b]	-6647	-6884	-6878	-6943	-6825
Weekly	-548	-561[b]	-561[b]	-548	-538	-561[b]	-558	-543
Monthly	52	37	33[b]	54	54	47	48	49
Laplace								
Daily	-6921	-6922	-6941[b]	-6937	-6849	-6921	-6849	-6827
Weekly	-564	-573[b]	-567	-563	-549	-571	-564	-556
Monthly	56	51	37	57	56	51	50	25[b]
Rayleigh								
Daily	-6912[b]	-6324	-6729	-5191	-6580	-6629	-6800	-5762
Weekly	-677[b]	-659	-622	-661	-630	-640	-602	-624
Monthly	11[b]	25	59	37	51	20	46	36
Student-T								
Daily	-6875	-6878	-6950	-6960[b]	-6885	-6880	-6844	-6849
Weekly	-549	-562[b]	-562[b]	-549	-538	-562[b]	-556	-544
Monthly	55	52	33[b]	57	56	50	51	52

Source: Authors' computation

Note: Statistics reported in this table are Akaike Information Criteria; [b] denotes the best GARCH model under each distribution.