Annals of Data Science
https://doi.org/10.1007/s40745-024-00517-4

Assessing the Risk of Bitcoin Futures Market: New Evidence

Anupam Dutta[1]

Received: 15 September 2023 / Revised: 12 January 2024 / Accepted: 17 January 2024

Abstract

The main objective of this paper is to forecast the realized volatility (RV) of Bitcoin futures (BTCF) market. To serve our purpose, we propose an augmented heterogenous autoregressive (HAR) model to consider the information on time-varying jumps observed in BTCF returns. Specifically, we estimate the jump-induced volatility using the GARCH-jump process and then consider this information in the HAR model. Both the in-sample and out-of-sample analyses show that jumps offer added information which is not provided by the existing HAR models. In addition, a novel finding is that the jump-induced volatility offers incremental information relative to the Bitcoin implied volatility index. In sum, our results indicate that the HAR-RV process comprising the leverage effects and jump volatility would predict the RV more precisely compared to the standard HAR-type models. These findings have important implications to cryptocurrency investors.

Keywords Bitcoin futures market · Realized volatility · Jump-induced volatility · Bitcoin implied volatility index · Leverage effects · HAR-RV models

JEL Classification C01 · G13 · G17

1 Introduction

Cryptocurrencies represent decentralized payment systems involving technological innovation. Notably, such payment process provides cryptocurrencies an incomparable transparency. Transactions of digital currencies are all recorded on the open public ledger known as 'blockchain'. Note that in blockchain technology, data science guarantees that all the transactions are safe and protected. It ensures the integrity and security of blockchain transactions. In particular, data science effectively detects any

✉ Anupam Dutta
adutta@uwasa.fi

1 School of Accounting and Finance, University of Vaasa, Vaasa, Finland

sort of suspicious activities on the blockchain network. Besides, it also assures that transactions are executed promptly. Due to this transparency and 24-h accessibility, digital currencies have gained immense popularity among the investors [1].

Over the past few years, digital currencies have grown rapidly in price, attractiveness, and mainstream adoption [2]. As of July, 2021, there are over 6000 cryptocurrencies being traded and currently, the global market capitalization amounts to $1.6 T. Among these digital currencies, Bitcoin still leads the market with a total market capitalization of 700bn, corresponding to 44% of the overall market capitalization.

Given the popularity and market capitalization of this leading crypto, it is of paramount importance to precisely predict its volatility so that investors who take large positions in it could be able to make proper investment decisions [3]. Note that the price formation for Bitcoin does not match with other asset classes because of its own uniqueness. A number of studies document that Bitcoin appears to be more volatile than equities and maintains a low correlation with other financial markets [4, 5]. This promotes Bitcoin as an efficient hedging instrument which can diversify the risk linked to investor portfolio [6, 7]. However, Bitcoin has been highly volatile in recent years and therefore, understanding the volatility dynamics of this digital currency plays a pivotal role in portfolio optimization and developing appropriate hedging strategies.[1]

It is noteworthy that in order to reduce the volatility of Bitcoin prices and increase its efficiency, Bitcoin futures (henceforth, BTCF) market was introduced in December 2017. Since its inception numerous studies have investigated the volatility dynamics of this new futures market. Note that the BTCF being a new asset class could be highly volatile [8]. Therefore, the patterns of volatility for this market have received ample attention among the academics, investors and policymakers [9].

The earlier works assessing the volatility dynamics of BTCF can be divided into three different categories. The first group of literature is mainly focused on how the volatility of BTCF impacts the price and volatility of other cryptocurrencies [10, 11]. The second category examines the quantitative risk management of BTCF market [12, 13], while the last category has assessed its hedging effectiveness [14, 15].

The present study joins the existing literature to propose an extended heterogeneous autoregressive (HAR) process which could be beneficial for predicting the realized volatility (henceforth, RV) of the BTCF market more precisely. It is worth mentioning that while a growing body of literature has explored the volatility dynamics and risk management of the BTCF market, prediction of realized volatility for this asset remains understudied. Given that volatility plays a key role in making appropriate asset allocation decisions and managing the risk of investor portfolio, finding accurate estimates of future RV for the BTCF market is crucial. To this end, this paper aims to

[1] Data science plays a pivotal role in estimating the risk involved in Bitcoin investments. In fact, one of the key applications of data science in cryptocurrency trading is predictive analytics. In doing so, it employs cutting-edge algorithms and statistical methods for predicting future price movements based on prior data. Such forecasts are important for cryptocurrency investors when estimating volatility and formulating effective trading strategies. In sum, data science has emerged as a key tool for interpreting complex market behaviors, predicting trends and managing portfolio risk.

fill this void in the existing literature by investigating the role of time-varying jumps occurring in BTCF returns when predicting the RV using the HAR-type models.

In doing so, our study makes several noteworthy contributions to the exiting literature on the cryptocurrency futures market. First, to the best of our knowledge, this is the initial attempt to asses if jump volatility has predictive contents for forecasting the RV of BTCF market. In particular, we examine whether and to what extent the volatility of time-varying jumps occurring in BTCF returns can predict the RV of BTCF market.[2] Such investigation is crucial given that large jumps in financial asset prices indicate an upsurge in market uncertainty [16]. Besides, jump-induced volatility, which refers to sudden, unexpected and abrupt price changes, temporarily disrupts market mechanisms, strains the capital markets, and brings huge losses to investors. Since Bitcoin is highly volatile in nature [17], the volatility of jumps may have predictive information for the RV of BTCF market. Hence, using the information content of jump-induced volatility would improve the forecast accuracy for the BTCF return volatility. Moreover, as the jump-induced volatility might carry important information for understanding the potential risk, time-dependent jumps should be detected properly. Hence, our analysis could be useful for deriving appropriate asset pricing models which could minimize the risk linked to cryptocurrency investments.

Second, we check the predictive ability of Bitcoin implied volatility index (henceforth, BVIN) when modeling the RV of BTCF market. It is now well-documented in the finance literature that the implied volatility offers incremental information for equity market volatility forecasts [18]. However, such literature is scarce in the context of cryptocurrency markets. We thus extend the prior literature by investigating whether the information content of BVIN is useful for the volatility forecasts of the BTCF market. In doing so, we also verify if the HAR process including the volatility of time-varying jumps outperforms the HAR model containing the information on the implied volatility index.

In our empirical analysis, we obtain the estimates of (time-varying) jump volatility by applying the GARCH-jump process to BTCF returns. We then use the information content of jump volatility in the HAR model to forecast the RV of BTCF returns. Our findings indicate the importance of jump-induced volatility. In fact, the HAR approach considering the information on jump-induced volatility outperforms all other models used in this study. Both in-sample and out-of-sample analyses support this novel finding. Our analysis thus reveals that financial time series alone would include predictive contents for forecasting the future RV and that using the jump process could be useful in this regard. Given that precise estimation of time-varying volatility plays a pivotal role in portfolio optimization and hedging strategies, our findings offer important implications to participants in cryptocurrency markets.

[2] Jumps are often observed in Bitcoin market. Previous studies have also documented the presence of time-varying jumps in cryptocurrency markets [17, 19, 20]. In this paper, unlike the prior works, we study the occurrence of such jumps in Bitcoin futures market and their role in forecasting the RV.

2 Data

Our data include daily observations from January 2019 to March 2022. The beginning of our sample period is detected by the availability of the BVIN index. The data on CME Bitcoin futures prices and BVIN are retrieved from the Bloomberg terminal, while the information on realized volatility (RV), based on the 5-min intra-day squared returns for the BTCF, is collected from Professor Dacheng Xiu's risk lab (https://dachxiu.chicagobooth.edu/#risklab).

Table 1 shows the descriptive statistics for the Bitcoin futures index. The results suggest that returns are negatively skewed with a leptokurtic distribution. The Jarque–Bera (J-B) test also shows the violation of normality assumption. Finally, the augmented Dickey Fuller (ADF) and Philips Perron (PP) tests confirm that Bitcoin futures returns appear to be stationarity.

Figure 1 exhibits the bitcoin futures prices for the sample period considered in this paper. It demonstrates that although the prices seem to be low initially, the market experiences substantial increments in 2021. However, we also witness a notable drop after the inception of Russo-Ukrainian war.

Next, Fig. 2 plots the realized volatility of Bitcoin futures prices. This graph reveals that while the futures index is, in general, volatile, the risk increases substantially during the COVID-19 pandemic periods.

Table 1 Summary statistics

Mean	Standard deviation	Skewness	Kurtosis	J-B test	ADF test	PP test
0.0031	0.0476	– 0.0417	7.67	707.05***	– 28.38***	– 28.37***

***, ** and * indicate statistical significance at 1%, 5% and 10% levels, respectively

Fig. 1 Time series plot of Bitcoin futures prices

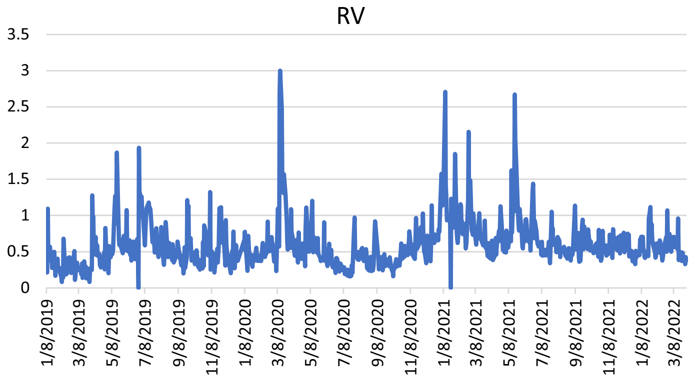

Fig. 2 The realized volatility of Bitcoin futures prices

3 Methodology

In our empirical analysis, we follow a two-step procedure. The first step estimates the conditional jump-induced volatility (JV) using the GARCH-jump approach developed by Chan and Maheu [16]. In the second step, we extend the HAR model considering the information on JV.

3.1 GARCH-Jump Approach

Following Chan and Maheu [16], we employ the GARCH-jump model as follwos[3]:

$$r_t = \pi + \mu r_{t-1} + \epsilon_t \tag{1}$$

where r_t indicates the logarithmic difference for the BTCF index at time t, and ϵ_t refers to the innovation term which is specified as:

$$\epsilon_t = \epsilon_{1t} + \epsilon_{2t} \tag{2}$$

where ϵ_{1t} will follow the GARCH (1,1) specification:

$$\int_{1t} = \sqrt{h_t} z_t,\ z_t \sim NID(0,\ 1) h_t = \omega + \alpha \epsilon_{1t-1}^2 + \beta h_{t-1} \tag{3}$$

In addition,ϵ_{2t} denotes a jump innovation defined as:

$$\epsilon_{2t} = \sum_{l=1}^{n_t} J_{tl} - \theta \lambda_t \tag{4}$$

[3] The choice of the AR(1) specification is based on the AIC and BIC values.

where J_{tl} is the jump size with a mean value θ and a variance ϑ^2, $\sum_{l=1}^{n_t} J_{tl}$ refers to the jump factor, and n_t represents the number of jumps at time$t$, following a Poisson distribution given by:

$$P(n_t = j|I_{t-1}) = \frac{e^{-\lambda_t}\lambda_t^j}{j!}, \; j = 0, 1, 2, \ldots \tag{5}$$

With an autoregressive conditional jump intensity (ARJI) given as[4]:

$$\lambda_t = \lambda_0 + \rho\lambda_{t-1} + \gamma\xi_{t-1} \tag{6}$$

In Eq. 6, λ_t indicates the time-varying conditional jump intensity parameter, λ_0 is the constant jump intensity, and ξ_{t-1} denotes the intensity residual. Chan and Maheu [16] assume that $\lambda_t > 0, \lambda_0 > 0, \rho > 0,$ and $\gamma > 0$.

We define the log-likelihood as:

$$L(\Theta) = \sum_{t=1}^{T} \log f(X_t|I_{t-1}; \Theta)$$

where $\Theta = (\pi, \mu, \omega, , \beta, \theta, \vartheta, \lambda_0, \rho, \gamma)$ and I_{t-1} is the information set.

Note that the jump-induced volatility (henceforth, JV) is given as:

$$JV_t = (\theta^2 + \vartheta^2)\lambda_t \tag{7}$$

3.2 HAR Models

The HAR process has received considerable attention in earlier research works as a suitable method for predicting the realized volatility of financial markets. The advantage of employing this approach is that it considers separating the realized volatility into short-, medium-, and long-term volatility components, thereby producing more accurate forecasts [23]. Previous studies [24–27] also demonstrate its dominance over the GARCH-type, SV-type, VAR-RV, MIDAS-RV, and ARFIMA-RV models.

Following Corsi [28] and Busch et al. [29], we define the baseline HAR-RV model as follows:

3.2.1 HAR-RV

$$RV_{t,t+h} = \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} + \varepsilon_t \tag{8}$$

[4] The ARJI process is an extension of the constant jump intensity (CJI) process, which is proposed by Jorion [21]. In the CJI process, it is assumed that $\lambda_t = \lambda_0$, implying that jump intensity is independent of time. However, a number of recent studies [22, 23] document that such intensity is time-dependent and hence prefer ARJI model to CJI process.

with h being equal to 1, 5 and 22 depending on the daily, weekly and monthly volatility components, respectively and

$$RV_{t_1,t_2} = \frac{1}{t_2 - t_1} \sum_{t=t_1+1}^{t_2} RV_t \tag{9}$$

In our paper, a number of extensions to the baseline HAR-RV model have been considered.[5] We first define the LHAR process, proposed by Corsi and Renò [30], where leverage effects are used to extend the HAR model:

3.2.2 LHAR-RV

$$\begin{aligned} RV_{t,t+h} &= \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} \\ &\quad + \psi_d r_t^- + \psi_w r_{t-5,t}^- + \psi_m r_{t-22,t}^- + \varepsilon_t \end{aligned} \tag{10}$$

where, $r_t^- = \min(r_t, 0)$, $r_{t-5,t}^- = \min((r_{t-4} + r_{t-3} + \cdots + r_t)/5,\ 0)$ and $r_{t-22,t}^- = \min((r_{t-21} + r_{t-20} + \cdots + r_t)/22,\ 0)$.

As mentioned earlier, the HAR model is also extended using the information content of BVIN index. This process, called LHAR-RV-IV, is given as:

3.2.3 LHAR-RV-IV

$$\begin{aligned} RV_{t,t+h} &= \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} \\ &\quad + \psi_d r_t^- + \psi_w r_{t-5,t}^- + \psi_m r_{t-22,t}^- + \delta BVIN_t + \varepsilon_t \end{aligned} \tag{11}$$

Next, we introduce the jump-induced volatility (JV) term, specified in Eq. (7), to the LHAR-RV process as follows:

3.2.4 LHAR-RV-JV

$$\begin{aligned} RV_{t,t+h} &= \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} + \psi_d r_t^- + \psi_w r_{t-5,t}^- \\ &\quad + \psi_m r_{t-22,t}^- + \phi_d JV_t + \phi_w JV_{t-5,t} + \phi_m JV_{t-22,t} + \varepsilon_t \end{aligned} \tag{12}$$

Finally, we propose the LHAR-RV-IV-JV model to simultaneously consider the information on leverage effects, BVIN index and jump-induced volatility:

[5] Prior literature [23, 24, 29, 31] shows that separating RV into a continuous and a jump component improves the predictive ability of the HAR-RV models. Yu [32], however, documents that considering these components is not useful for the Bitcoin market. Hence, the jump and continuous components are ignored in this paper.

3.2.5 LHAR-RV-IV-JV

$$RV_{t,t+h} = \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} + \psi_d r_t^- + \psi_w r_{t-5,t}^- \\ + \psi_m r_{t-22,t}^- + \phi_d JV_t + \phi_w JV_{t-5,t} + \phi_m JV_{t-22,t} + \delta BVIN_t + \varepsilon_t \quad (13)$$

Notably, our objective is to forecast one-month ahead volatility only given that options expire at a monthly frequency. Hence, h equals 22 in our analysis. Busch et al. [29] also advocate this when forecasting the RV of bond, currency and equity markets using the VIX index.

3.3 Out-of-Sample Forecast s

3.3.1 Forecast Evaluation

The forecasting performance of different models used in this study is evaluated using the heteroskedasticity adjusted root mean square error (HRMSE) proposed by Bollerslev and Ghysels [33]. We define this statistic as:

$$HRMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{RV_t - \widehat{RV_t}}{RV_t}\right)^2} \quad (14)$$

With T indicating the number of data points to be forecasted, while RV_t and $\widehat{RV_t}$ specify the true and estimated volatility for day t, respectively.

For robustness check, the mean absolute error (MAE) is also estimated. We define MAE as:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|RV_t - \widehat{RV_t}\right| \quad (15)$$

3.3.2 Diebold and Mariano Test

With a view to testing the null hypothesis that two forecasts have the same accuracy, we apply the Diebold and Mariano (hereafter, DM) test [34]. This test assumes that $\boldsymbol{e_{it}} = \boldsymbol{RV_t} - \widehat{\boldsymbol{RV_t}}\,(\boldsymbol{i} = 1, 2)$ refers to the forecast errors and that $\boldsymbol{d_t} = \boldsymbol{f}(\boldsymbol{e}_{1t}) - \boldsymbol{f}(\boldsymbol{e}_{2t})$, with $\boldsymbol{f}(.)$ denoting a function of forecast errors. We then wish to test:

$$H_0 : E(d_t) = 0$$

DM [34] shows that the approximate asymptotic variance of $\overline{d}$ is given as:

$$Var\left(\overline{d}\right) \approx k^{-1}\left[\eta_0 + 2\sum_{l=1}^{p-1}\eta_l\right] \tag{16}$$

where η_l indicates the l-th autocovariance of d_t, which is estimated as:

$$\widehat{\eta} = k^{-1}\sum_{t=l+1}^{k}(d_t - \overline{d})(d_{t-l} - \overline{d}) \tag{17}$$

The DM test statistic is then defined as:

$$\mathbf{DM} = \left(\widehat{\boldsymbol{Var}\left(\overline{\boldsymbol{d}}\right)}\right)^{-\frac{1}{2}}\overline{\boldsymbol{d}} \tag{18}$$

Given that $\boldsymbol{H}_0$ is true, the probability distribution of DM statistic is asymptotically normal.

3.3.3 Mincer–Zarnowitz Regression

The out-of-sample predictions are also compared applying the Mincer–Zarnowitz (MZ) [35] regression approach. This technique is beneficial as it examines if the proposed models provide incremental information relative to the baseline HAR-RV model. We define the MZ regression approach as follows:

$$RV_t = \varphi_0 + \varphi_1\widehat{RV_t} + \epsilon_t \tag{19}$$

where, RV_t and $\widehat{RV_t}$ are the true and estimated volatility for day t, respectively. We then compare the HAR models on the basis of R^2 (coefficient of determination) statistics. Notably, our in-sample estimation period spans from January 2019 to March 2021 and the out-of-sample period from April 2021 to March 2022.

4 Empirical Findings

4.1 In-Sample Estimates

The findings of the HAR models are shown in Table 3. Before discussing these results, we briefly explain the estimates of out GARCH-jump approach presented in Table 2. It is evident from these estimates that jumps exist in BTCF returns as the parameters of the ARJI process are mostly significant. Notably, the positive coefficient of the jump variance (see the estimate of.$\boldsymbol{\vartheta}^2$) infers that volatility driven by abnormal information has a positive effect on the volatility of Bitcoin returns. In addition, the high value of the intensity parameter ($\boldsymbol{\rho} = 0.98$) indicates that the time-varying jump intensity

Table 2 Estimates of the GARCH-ARJI model

Variables	Estimates	Standard errors	t-statistics	p-values
π	0.0027**	0.0012	2.16	0.03
μ	− 0.0181	0.0358	− 0.50	0.61
ω	0.0001	0.0001	1.01	0.29
α	0.0484***	0.0107	4.54	0.00
β	0.8904***	0.0238	37.36	0.00
θ	0.0031	0.0079	0.39	0.69
ϑ^2	0.0783***	0.0107	7.28	0.00
λ_0	0.0036***	0.0012	2.95	0.00
ρ	0.9813***	0.0071	138.84	0.00
γ	0.0466	0.0316	1.47	0.14
Log-likelihood	1357.82			

This Table presents the findings of the GARCH-ARJI model for the BTCF index. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels, respectively

is persistent [16]. The nonzero value of the parameter $\boldsymbol{\rho}$ also confirms that jumps occurring in Bitcoin futures index do vary over time. This finding is consistent with Chaim and Laurini [19] and Gronwald [17]. However, the analysis of these earlier studies is focused on the Bitcoin spot index, whereas our paper examines the jump dynamics of Bitcoin futures index. This finding is crucial given that time-varying jump risk could play a key role in portfolio allocation decisions and risk management inferences in the cryptocurrency markets [36, 37]. In particular, the results of our analysis could guide investors to form better portfolios which could diversify the jump-induced risk. Hence, this strand of empirical research is useful for developing optimal investment strategies [38].

Moreover, Fig. 3 depicting the jump-induced volatility also confirms presence of high jumps during the COVID-19 crisis periods. Hence, time-varying jumps may include predictive content for forecasting the RV of BTCF market.

Now, the numbers shown in Table 3 indicate several interesting findings. First, including leverage effects in the HAR-RV model improves the accuracy of the baseline HAR process as the $\mathbf{R}^2$ (%) statistic increases from 23.80 to 25.60. Second, we find a negative association between leverage effects and RV, which is not an exception in finance literature [30]. Third, the BVIN index, although exerts a significant impact on the RV, does not increase the $\mathbf{R}^2$ statistic substantially. This is not consistent with earlier studies given that implied volatilities, in general, contain predictive information for the financial markets. Forth, inserting the JV factor to the LHAR-RV model increases the $\mathbf{R}^2$ statistic markedly, revealing the importance of using the information on jumps when modeling the RV of BTCF market. Fifth, the RV is sensitive only to the monthly component of the jump-induced risk factor revealing the long memory behavior of volatility. Summarizing, our analysis suggests that including leverage effects and time-dependent jumps in the HAR-RV model would increase its accuracy.

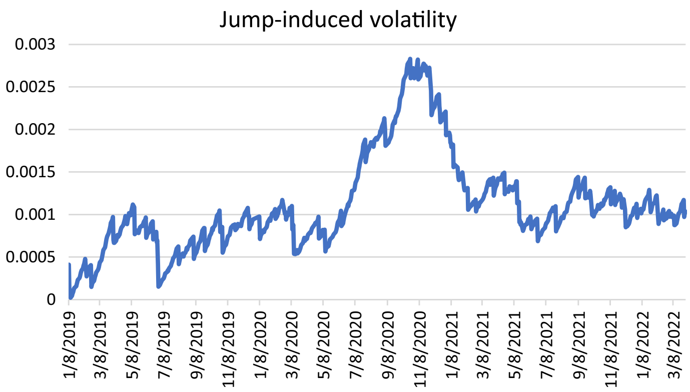

Fig. 3 The conditional jump-induced volatility of Bitcoin futures prices

4.2 Out-of-Sample Analysis

The out-of-sample forecast results based on the HMSE and MAE statistics along with MZ regression model are exhibited in Table 4. Our in-sample estimation period spans from January 2019 to March 2021 and the out-of-sample period from April 2021 to March 2022. The results reveal that the LHAR-RV-IV-JV process produces the lowest HRMSE and MAE statistics. The Diebold and Mariano (DM) [34] test further supports these findings by rejecting the null hypothesis of equal forecast accuracy. However, the DM test cannot tell us whether LHAR-RV-IV-JV outperforms LHAR-RV-IV in case when we use HRMSE statistic. But for the MAE statistic, the results confirm the superiority of LHAR-RV-IV-JV model over the LHAR-RV-IV process. Hence, the extended models including the information content of jump-induced volatility predict the volatility of BTCF market more precisely than the standard HAR-RV models.

Next, the R^2 (%) statistics, produced by the MZ regression process, further confirm that the HAR-type models considering the jumps (i.e., LHAR-RV-JV and LHAR-RV-IV-JV) surpass other approaches by yielding higher R^2 values. Our overall results thus suggest that the information on time-dependent jumps in BTCF returns is essential for increasing the precision of the HAR-RV model. Therefore, participants in Bitcoin futures market should analyze such jumps while forecasting the realized volatility of this new asset class.

4.3 Forecasting Value-at-Risk (VaR)

We now conduct ta simple VaR analysis for finding the best forecast model. In doing so, we test for all the HAR-RV models with a VaR estimated for the quantile level q. To this end, the likelihood ratio (LR) test, developed by Kupiec [39], is applied.

Table 3 Estimates of different HAR-RV models

Models →	HAR-RV		LHAR-RV		LHAR-RV-IV		LHAR-RV-JV		LHAR-RV-IV-JV	
	Estimate	S.E	Estimate	S.E	Estimate	S.E	Estimate	S.E	Estimate	S.E
c	0.3233***	0.2060	0.3071***	0.0207	0.3514***	0.0294	0.1387***	0.0369	0.1727***	0.0445
τ_d	0.1343***	0.0297	0.1239***	0.0349	0.1358***	0.0353	0.1140***	0.0333	0.1222***	0.0338
τ_w	0.0958**	0.0449	0.1175**	0.0474	0.1650***	0.0523	0.0097	0.04513	0.0179	0.0551
τ_m	0.2244***	0.0451	0.2621***	0.0462	0.2621***	0.0462	0.2969***	0.0446	0.3336***	0.0520
ψ_d			− 0.0930	0.2813	− 0.0740	0.2807	− 0.0588	0.2663	− 0.0443	0.2664
ψ_w			− 0.8466	0.6865	− 0.6801	0.6894	− 1.6462**	0.6567	− 1.5363**	0.6611
ψ_m			− 6.8661***	1.4574	− 7.1275***	1.4592	− 5.6752***	1.3831	− 5.8501***	1.3882
δ					0.0013**	0.0006			0.0008	0.0006
ϕ_d							2.3784	7.2349	2.4113	7.2306
ϕ_w							10.4959	8.4094	10.8669	8.4088
ϕ_m							19.8081***	3.7034	20.0575***	3.7057
R^2(%)	23.80		25.60		25.94		33.65		33.73	

This table presents the estimates of different HAR-type specifications. Our sample runs from January 2019 to March 2022. We estimate five HAR models and the R^2 statistics are reported accordingly. $\tau\prime$ s, $\psi\prime$ s, $\phi\prime$ s and δ represent the coefficients of different volatility components, leverage effects, jump-induced volatility and Bitcoin volatility index, respectively. S.E. indicates the standard error. ***, ** and * indicate statistical significance at 1%, 5% and 10% levels, respectively

Table 4 Out-of-sample forecast results

Models →	HAR-RV	LHAR-RV	LHAR-RV-IV	LHAR-RV-JV	LHAR-RV-IV-JV
HRMSE	0.2367**	0.2277**	0.2280**	0.2229	0.2206
MAE	0.0332**	0.0301**	0.0289**	0.0251**	0.0242
R^2 (%)	12.81	16.25	17.12	18.94	21.18

This table shows the HRMSE and MAE statistics along with the R^2 (%) statistics provided by the MZ regression model. Our in-sample estimation period spans from January 2019 to March 2021 and the out-of-sample period from April 2021 to March 2022. ** indicates that the Diebold-Mariano (DM) test is statistically significant at 5% level

We begin with the following hit sequence:

$$Hit_t = \begin{cases} 1, \ if r_t < VaR_t \\ 0, \ if r_t \geq VaR_t \end{cases} \tag{20}$$

Where r_t refers to the return on day t and VaR_t is specified as:

$$VaR_t = Z_q \sqrt{g_t} \tag{21}$$

where Z_q indicates the quantile at $100 \times q\%$ of the standardized probability distribution and g_t denotes the risk predicted by the HAR-RV approaches under study [40–42].

We then assume that N computes the frequency of VaR violations and T refers to the data points. In order to investigate $H_0 : f = q$, with f measuring the failure rate, we use the following LR test statistic proposed by Kupiec [39]:

$$LR = -2\ln\{(1-q)^N q^{T-N} / (1 - N/T)^{T-N} (N/T)^N\} \sim \chi^2(1) \tag{22}$$

The p-values of this test are provided in Table 5. Our findings show failure rates for both left and right quantiles. Given that the accuracy of the HAR models increases with the increment in p-values, these numbers further reveal the significance of jumps in forecasting the RV of BTCF market. Overall, our analysis suggests that both leverage effects and time-varying jumps are crucial for predicting the VaR with precision.

5 Conclusions

This paper aims to forecast the realized volatility of Bitcoin futures market. In doing so, we propose an augmented HAR-RV model to consider the information on time-dependent jumps observed in BTCF returns. In particular, we estimate the jump-induced volatility using the GARCH-jump process and then consider this information in the HAR model. Both the in-sample and out-of-sample analyses show that jumps offer extra information which is not provided by the existing HAR models. Besides, a novel finding is that the jump-induced volatility offers incremental information relative to the Bitcoin implied volatility index. In sum, our results indicate that the HAR-RV

Table 5 Forecasting value-at-risk

Models ↓	LQ = 10%	LQ = 5%	LQ = 1%	RQ = 10%	RQ = 5%	RQ = 1%
Panel A: CARBONEX						
HAR-RV	0.223	0.296	0.282	0.351	0.388	0.431
LHAR-RV	0.331	0.368	0.377	0.391	0.428	0.488
LHAR-RV-IV	0.356	0.390	0.401	0.411	0.447	0.509
LHAR-RV-JV	0.429	0.465	0.449	0.476	0.501	0.561
LHAR-RV-IV-JV	0.445	0.481	0.463	0.493	0.527	0.579

In this Table, we present the *p*-values of the likelihood ratio test specified in Eq. 18. The accuracy of the HAR models increases with the increment in *p*-values. *LQ* left quantile *RQ* right quantile. Numbers in bold indicate the highest *p*-values

process comprising the leverage effects and jump volatility would forecast the RV more precisely compared to the standard HAR-type models.

Our analysis could be of interest to cryptocurrency investors given that the information on time-varying jumps in BTCF returns could be useful for measuring the risk of investor portfolio. In addition, as jumps often provide early signals of market crashes, they might play a crucial role in risk management, portfolio optimization and hedging strategies. However, it is also worth noting that since the use of jump-induced volatility in forecasting the realized volatility of Bitcoin market is a novel area of empirical finance, it could be difficult to fully understand the role of such jumps for predicting the risk of cryptocurrency investments. As investors and policymakers are very keen to have further knowledge about this domain, future research could employ more sophisticated statistical tools to shed more light on how the information on jump-induced volatility can be used to obtain improved volatility forecasts during the crisis periods. To this end, applications of big data methods and artificial intelligence such as neural networks could provide superior forecasts for Bitcoin volatility. Artificial neural networks are particularly useful for handling high frequency trading data and predicting financial risk. Given that real-time decision making also benefits from the use of artificial intelligence deep learning methods are highly recommended for forecasting the price of Bitcoin in real-time. Prior studies such as Shi [43], Olson and Shi [44], Shi et al. [45] and Tien [46] discuss a number of important big data methods which are quite useful for this sort of analysis. Note that future studies could also investigate whether jumps in Bitcoin returns can predict the volatility of other cryptos. In addition, attempts could be made to identify the key determinants of jumps occurring in Bitcoin price index. Given that jump-induced volatility represents a large fraction of market volatility, identifying such determinants might be essential for hedging cryptocurrency market risk amid the crisis periods.

Author Contributions Statement I am the solo author of this manuscript

Funding Open Access funding provided by University of Vaasa. This research has received funding from the European Union's Horizon 2020 research and innovation programme, within the OpenInnoTrain project under the Marie Sklodowska-Curie grant agreement n°823971. The content of this publication does not reflect the official opinion of the European Union. Responsibility for the information and views expressed in the publication lies entirely with the author.

Data Availablity The data will be available from the author upon request.

Code Availablity The RATS code are used to produce the results.

Declarations

Ethical Statement This work was not copied from any source, and it does not harm humans or society in any way.

Conflict of interest The author declares no conflict of interest.

References

1. Grobys K, Sapkota N (2020) Predicting cryptocurrency defaults. Appl Econ 52:5060–5076. https://doi.org/10.1080/00036846.2020.1752903
2. Foley S, Karlsen JR, Putninš TJ (2019) Sex, drugs, and Bitcoin: How much illegal activity is financed through cryptocurrencies? Rev Financ Stud 32:1798–1853. https://doi.org/10.1093/rfs/hhz015
3. Aras S (2021) On improving GARCH volatility forecasts for Bitcoin via a meta-learning approach. Knowl-Based Syst 230:107393. https://doi.org/10.1016/j.knosys.2021.107393
4. Yermack D (2015) Is Bitcoin a real currency? An Economic appraisal. In: Lee DKC (ed) Handbook of digital currency. Academic Press, pp 31–43
5. Briere M, Oosterlinck K, Szafarz A (2015) Virtual Currency, tangible return: portfolio diversification with Bitcoin. J Asset Manag 16:365–373
6. Lee DKC, Guo L, Wang Yu (2018) Cryptocurrency: a new investment opportunity? J Altern Invest 20:16–40
7. Dutta A, Das D, Jana RK, Vo XV (2020) COVID-19 and oil market crash: revisiting the safe haven property of gold and Bitcoin. Resour Policy 69:101816. https://doi.org/10.1016/j.resourpol.2020.101816
8. Park BJ (2022) The COVID-19 pandemic, volatility, and trading behavior in the Bitcoin futures market. Res Int Bus Financ 59:101519. https://doi.org/10.1016/j.ribaf.2021.101519
9. Bouteska A, Mefteh-Wali S, Trung Dang T (2022) Predictive power of investor sentiment for Bitcoin returns: evidence from COVID-19 pandemic. Technol Forecast Soc Chang 184:121999. https://doi.org/10.1016/j.techfore.2022.121999
10. Ruozhou L, Shanfeng W, Zili Z, Zhao X (2019) Is the introduction of futures responsible for the crash of Bitcoin? Financ Res Lett 34:101259. https://doi.org/10.1016/j.frl.2019.08.007
11. Kim W, Lee J, Kang K (2020) The effects of the introduction of Bitcoin futures on the volatility of Bitcoin returns. Financ Res Lett 33:101204. https://doi.org/10.1016/j.frl.2019.06.002
12. Guo ZY (2021) Price volatilities of Bitcoin futures. Financ Res Lett 43:102022. https://doi.org/10.1016/j.frl.2021.102022

13. Guo ZY (2022) Risk management of Bitcoin futures with GARCH models. Financ Res Lett 45:102197. https://doi.org/10.1016/j.frl.2021.102197
14. Alexander C, Deng J, Zou B (2022) Hedging with automatic liquidation and leverage selection on Bitcoin futures. Eur J Oper Res 306:478–493. https://doi.org/10.1016/j.ejor.2022.07.037
15. Sebastião H, Godinho P (2020) Bitcoin futures: an effective tool for hedging cryptocurrencies. Financ Res Lett 33:101230. https://doi.org/10.1016/j.frl.2019.07.003
16. Chan WH, Maheu JM (2002) Conditional jump dynamics in stock market returns. J Bus Econ Stat 20:377–389. https://doi.org/10.1198/073500102288618513
17. Gronwald M (2019) Is Bitcoin a commodity? On price jumps, demand shocks and certainty of supply. J Int Money Financ 97:86–92. https://doi.org/10.1016/j.jimonfin.2019.06.006
18. Kambouroudis DS, McMillan DG, Tsakou K (2021) Forecasting realized volatility: the role of implied volatility, leverage effect, overnight returns, and volatility of realized volatility. J Futures Mark 41:1618–1639. https://doi.org/10.1002/fut.22241
19. Chaim P, Laurini MP (2018) Volatility and return jumps in Bitcoin. Econ Lett 173:158–163. https://doi.org/10.1016/j.econlet.2018.10.011
20. Bouri E, Roubaud D, Shahzad SJH (2020) Do Bitcoin and other cryptocurrencies jump together? Q Rev Econ Finance 76:396–409. https://doi.org/10.1016/j.qref.2019.09.003
21. Jorion P (1988) On jump processes in the foreign exchange and stock markets. Rev Financ Stud 1:427–445. https://doi.org/10.1093/rfs/1.4.427
22. Dutta A, Bouri E, Roubaud D (2021) Modelling the volatility of crude oil returns: jumps and volatility forecasts. Int J Financ Econ 26:889–897. https://doi.org/10.1002/ijfe.1826
23. Dutta A, Das D (2022) Forecasting realized volatility: new evidence from time-varying jumps in VIX. J Futures Mark 42:2165–2189. https://doi.org/10.1002/fut.22372
24. Andersen TG, Bollerslev T, Diebold FX (2007) Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility. Rev Econ Stat 89:701–720. https://doi.org/10.1162/rest.89.4.701
25. Andersen TG, Bollerslev T, Huang X (2011) A reduced form framework for modeling volatility of speculative prices based on realized variation measures. J Econom 160:176–189. https://doi.org/10.1016/j.jeconom.2010.03.029
26. Giot P, Laurent S (2007) The information content of implied volatility in light of the jump/continuous decomposition of realized volatility. J Futures Mark 27:337–359. https://doi.org/10.1002/fut.20251
27. Forsberg L, Ghysels E (2007) Why do absolute returns predict volatility so well? J Financ Economet 5:31–67. https://doi.org/10.1093/jjfinec/nbl010
28. Corsi F (2009) A simple approximate long memory model of realized volatility. J Financ Econom 7:174–196. https://doi.org/10.1093/jjfinec/nbp001
29. Busch T, Christensen BJ, Nielsen MO (2011) The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. J Econom 160:48–57. https://doi.org/10.1016/j.jeconom.2010.03.014
30. Corsi F, Renò R (2012) Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. J Bus Econ Stat 30:368–380. https://doi.org/10.1080/07350015.2012.663261
31. Corsi F, Pirino D, Renò R (2010) Threshold bipower variation and the impact of jumps on volatility forecasting. J Econom 159:276–288. https://doi.org/10.1016/j.jeconom.2010.07.008
32. Yu M (2019) Forecasting Bitcoin volatility: the role of leverage effect and uncertainty. Phys A 533:120707. https://doi.org/10.1016/j.physa.2019.03.072
33. Bollerslev T, Ghysels E (1996) Periodic autoregressive conditional heteroscedasticity. J Bus Econ Stat 14:139–151. https://doi.org/10.1080/07350015.1996.10524640
34. Diebold FX, Mariano RS (1995) Comparing predictive accuracy. Journal of Business & Economic Statistics 13:253–263. https://doi.org/10.1198/073500102753410444
35. Mincer J, Zarnowitz V (1969) The evaluation of economic forecasts. In: Mincer J (ed) Economic forecasts and expectations. NBER, New York, pp 3–46
36. Xu F, Bouri E, Cepni O (2022) Blockchain and crypto-exposed US companies and major cryptocurrencies: the role of jumps and co-jumps. Financ Res Lett 50:103201. https://doi.org/10.1016/j.frl.2022.103201
37. Zhou C, Wu C, Wang Y (2019) Dynamic portfolio allocation with time-varying jump risk. J Empir Financ 50:113–124. https://doi.org/10.1016/j.jempfin.2019.01.003

38. Zhang L, Bouri E, Chen Y (2023) Co-jump dynamicity in the cryptocurrency market: a network modelling perspective. Financ Res Lett 58:104372. https://doi.org/10.1016/j.frl.2023.104372
39. Kupiec P (1995) Techniques for verifying the accuracy of risk measurement models. J Deriv 2:173–184
40. Giot P (2005) Implied volatility indexes and daily value-at-risk models. J Deriv 12:54–64
41. Giot P, Laurent S (2003) Value-at-risk for long and short trading positions. J Appl Econome 18:641–664. https://doi.org/10.1002/jae.710
42. Giot P, Laurent S (2004) Modelling daily value-at-risk using realized volatility and ARCH type models. J Empir Financ 11:379–398. https://doi.org/10.1016/j.jempfin.2003.04.003
43. Shi Y (2022) Advances in Big Data Analytics: Theory, Algorithm and Practice, Springer, Singapore.
44. Olson DL, Shi Y (2007) Introduction to business data mining. McGraw-Hill/Irwin, New York
45. Shi Y, Tian YJ, Kou G, Peng Y, Li JP (2011) Optimization based data mining: theory and applications. Springer, Berlin
46. Tien JM (2017) Internet of things, real-time decision making, and artificial intelligence. Ann Data Sci 4(2):149–178