VOLATILITY REGIMES IN CRYPTOCURRENCY RETURNS

Master's thesis
for the Master's degree programme Economics in the Faculty of
Business, Economics and Social Sciences at the
Christian-Albrechts-Universität zu Kiel

Submitted by

Nhu, Le Quynh

Student ID: stu218511

First assessor: Prof.Dr. Markus Haas
Second assessor: Prof.Dr. Stefan Reitz
Vietnam, January, 2022

Table of Contents

Contents

List of Figures

List of Tables

1   Abstract

The paper evaluates whether there exists any regime changes in volatility dynamics of Cryptocurrencies, namely Bitcoin by using Markov-switching GARCH (MSGARCH) approach. There is also a comparison between MSGARCH process and traditional single-regime GARCH model and prediction of one-day-ahead Value-at-Risk (VaR). We use both Bayesian approach and Maximum Likelihood model to forecast the parameters, compute the VaR, so that we can examine which approach is better. The result supports the evidence of regime changes in GARCH process of Bitcoin and shows that MSGARCH can outperform single-regime GARCH process in estimating the VaR but not much.

2   Introduction

Since the recent global financial crisis in 2008, financial institutions have been under pressure to employ proper models for risk monitoring, with modelling volatility being critical for risk management. As a result of that crisis, the Basel III regulation has been implemented the risk management's mechanism, requiring a more stringent management system as well as higher capital requirements (Basel Comittee on Banking Supervision, 1996). Since then, the financial system has also constantly faced a fresh threat in the form of the creation of decentralized Cryptocurrencies, the first of which is Bitcoin. It was such electronic cash that could be traded without intermediaries, which was regarded its key benefit over regular payment systems, when it was first launched by Nakamoto (2008) in 2009. High liquidity, minimal transaction costs, and anonymity are further advantages.

Since its introduction up to now, there are more than 8000 types of Cryptocurrencies (CCs) (source: coinmarketcap.com). At the beginning of 2021, market capitalization of CCs has increased from approximately 2 trillion USD to more than 2.5 trillion USD at the end of the year (source: coinmarketcap.com). Despite the recent COVID-19 epidemic, the phenomenal expansion of CCs has attracted a great number of investors, corporations, financial authorities, and legislators. There have been several talks on the benefits of CCs, such as their use as a medium of exchange or as hedging tools against gold and dollars (Dyhrberg, 2016). Bitcoin delivers diversification benefits, according to Corbet et al. (2018), because of its remarkable high return. Furthermore, Kim (2017) demonstrated that Bitcoin's transaction costs differed by 2% from those of other traditional market places.

Nevertheless, because of its decentralized properties and its anonymity, CCs marketplaces have yet to be formally regulated, and risk assessments have yet to be properly estimated. According to Corbet et al. (2018); Baur et al. (2018); Fry (2018), Bitcoin is mostly utilized for speculative motives, resulting in significant volatility and bubbles in market itself. Thus, finding models that effectively cap-

ture volatility dynamics and forecast the risk of CCs accurately is essential. In fact, several previous publications have concentrated on predicting the volatility clustering of CCs.

The paper is organized as follows. Immediately after the introduction, the literature review are presented in Section 3. Then, the methodology of Markov-switching GARCH process and the backtesting methods are given in Section 4. In Section 5, we discuss the empirical results. Conclusion is given in Section 6.

3   Literature review

A number of researches have been studied about the volatility dynamics of Bitcoin returns. In particular, Dyhrberg (2016) has performed the symmetric GARCH model with conclusion that Bitcoin and gold possess several similar characteristics, implicating hedging capabilities and advantages as a medium of exchange. Chu et al. (2017) analysed seven of the most common CCs, compared goodness-of-fit of each model, and concluded that the I-GARCH and GJR-GARCH models presented the best fits in terms of modelling volatility. These findings explain the use of single-regime (SR) Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models to model CCs volatility. However, several researchers discovered that volatility dynamics led to structural breaks and regime changes, which then prompted them to improve volatility forecasts. Bouri et al. (2017) and Katsiampa (2017), for instance, applied asymmetric GARCH models to study the conditional variance's reaction to prior positive and negative shocks and discovered an inverted leverage effect. According to Haas et al. (2004) and Bauwens et al. (2014), they concluded that not accounting for regime transitions or leverage impact could lead to bias estimates on volatility forecasts. As a result, many researches have been conducted to compare regime-switching (RS), particularly, Markov-Switching GARCH (MSGARCH) with SRG models. They have discovered that MSG(K) models tend to outperform SRG models in estimating CCs risk. Thus, Ardia et al. (2019) and Caporale and Zekokh (2019) recommend implementing Markov-Switching GARCH (MSGARCH), a methodology for evaluating conditional variance of Bitcoin as well as other CCs, whose parameters might change over time according to a discrete latent variable.

Research for modelling volatility on the use of time-series models was first introduced in the concept of ARCH process byEngle (1982) and then generalized by Bollerslev (1986) as GARCH process. Many more concepts of GARCH have been proposed, in particular, (Nelson, 1991) developed so-called exponential GARCH (EGARCH), the Student's t-GARCH by Bollerslev (1987), GJR-GARCH model of Glosten et al. (1993) and so on. As mentioned above, Bauwens et al. (2014)concluded that there are structural breaks or regime changes in volatility dynamics of financial asset and ignoring these features can lead to biased volatility estimation or even systemic consequences. Therefore, MSGARCH model, whose parameters

can change over time according to unobservable variable, is known as one of the best ways to address the switch of regimes among return process.

Due to path-dependence problem, which means the conditional variance depends entirely on the history of regimes Haas et al. (2004), first attempts to forecast regime changes conditional on volatility dynamics were Regime-switching ARCH models by Cai (1994) and Hamilton and Susmel (1994). However, Gray (1996) shows that, regimes are latent, in order to evaluate the likelihood function for Markov-switching GARCH model, all K^N possible regime paths have to be integrated (N is number of observations), making this estimation extremely unfeasible. To tackle this issue, several MSGARCH approximations were proposed. Specifically, Dueker (1997) and Klaassen (2002) used an adhoc approximation to reduce path dependence by integrating out the latent regime path from the lagged conditional volatility instead of in the likelihood function. However, using this model is difficult because of computing multi-step volatility forecasts and missing covariance stationary conditions (Haas et al., 2004). Therefore, Ardia et al. (2018) specified a MSGARCH model of K separate GARCH processes for the conditional variance so that the variance solely depends on past data and the current state, which solves the path-dependence problem and makes the model analytically tractable. In addition, in the research, two estimation approaches, namely Markov chain Monte Carlo simulation and Maximum Likelihood approach are compared and find that MSGARCH outperform single regime models in estimating VaR and ES for daily, weekly, and ten-day log returns (Caporale and Zekokh, 2019).

Considering recent breakthroughs in the CCs market, this paper investigate two main points. Firstly, whether there are regime changes in CCs dynamics and secondly, whether Markov-switching GARCH models outperform Single-regime GARCH models in forecasting CCs risk accurately.

4 Methodology

4.1 Model specification

Let s_t be the state variable which is equal to k if the market is in state k at time t. The regimes are not observable and Ardia et al. (2018) proposes that the stochastic mechanism generating the sequence of states s_t is a latent first-order ergodic homogeneous Markov-chain with K states and transition probability matrix $\mathbf{P} \equiv \{p_{ij}\} = \mathbb{P}[s_t = j | s_{t-1} = i]_{i,j=1}^{K}$. Let y_t be the percentage log-return of the CCs at time t. Here, it is assumed that log-returns have μ_t - the conditional mean is zero and are not autocorrelated. Then following equations describe MSGARCH

specifications:

$$y_t = \mu_t + e_t \tag{1}$$

$$e_t = \sigma_{k,t} \eta_{k,t} \tag{2}$$

$$\eta_{k,t} \overset{\text{iid}}{\sim} \mathcal{D}(0, 1, \xi_k), \qquad k = 1, \ldots, K, \tag{3}$$

where e_t is the error term, $\eta_{k,t}$ are the innovations at specific regimes and $\sigma_{k,t}$ is the time varying conditional standard deviation in regime k. $\mathcal{D}(.)$ is considered as continuous distribution of innovation term with zero mean, unit variance and shape parameters ξ_k. The return distribution conditional on state k is:

$$y_t | (s_t = k, \mathcal{I}_{t-1}) \sim \mathcal{D}(0, \sigma_{k,t}^2, \xi_k) \tag{4}$$

where $\mathcal{I}_{t-1} \equiv y_{t-1}, t > 0$ is the return history up to time $t - 1$.

According to Hamilton and Susmel (1994) and Paolella et al. (2012), the stochastic mechanism governs a latent first order ergodic homogeneous Markov-chain with K regimes. Latent means that we cannot reliably monitor or forecast regimes or regime transitions with certainty. A first-order Markov chain indicates that the probability that $s_t = j$ at time t is determined solely by the most recent previous state s_{t-1}; $s_{t-2}, s_{t-3}...$ will not provide any new information. Homogeneous indicates that the transition probabilities are time-invariant, that is, they are independent of the time point t or the history of previous returns, $\mathcal{I} = y_{t-1}, y_{t-2}, \ldots$. The Markov-chain transition probabilities may therefore be stated as follows:

$$\mathbb{P} = [s_t = j | s_{t-1} = i, s_{t-2}, \ldots, \mathcal{I}_{t-1}] = \mathbb{P}[s_t = j | s_{t-1} = i] = p_{ij} \tag{5}$$

then its **Transition probability matrix P** including $K \times K$ transition probabilities is:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{K1} \\ p_{12} & p_{22} & \cdots & p_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1K} & p_{2K} & \cdots & p_{KK} \end{bmatrix} \tag{6}$$

where $\sum_{j=1}^{K} = p_{kj} = 1, k = 1, \ldots, K$. The staying probabilities is $p_{kk} = \mathbb{P}[s_t = k | s_{t-1} = k]$. For a Markov-chain to be ergodic, two characteristics must be met, namely irreducibility and aperiodicity. In the two-state case, the former holds if $p_{kk} < 1$; $k = 1, 2$, i.e. there is a non-zero probability to go from one state to another. The latter holds if $p_{kk} > 0$ for at least one $k = 1, 2$, i.e. no periodic chain is obtained.

The **unconditional regime probabilities** are defined as $\lambda_k = \mathbb{P}[s_t = k], k = 1, \ldots, K$. In the long run, the probability λ_k is expected to be in regime k. Here we focus on a derivation K = 2. From the assumption of stationary unconditional regime probabilities, so-called time-invariant property $\lambda_k = \mathbb{P}[s_t = k] = \mathbb{P}[s_{t-1} =$

k], k = 1,2 for any t, we can write the unconditional regime probabilities in terms of transition probabilities as follows:

$$\lambda_1 = \mathbb{P}[s_t = 1]$$
$$= \mathbb{P}[s_{t-1} = 1]\mathbb{P}[s_t = 1|s_{t-1} = 1] + \mathbb{P}[s_{t-1} = 2]\mathbb{P}[s_t = 1|s_{t-1} = 2]$$
$$= \lambda_1 p_{11} + \lambda_2 p_{21}$$
$$= \lambda_1 p_{11} + (1 - \lambda_1)(1 - p_{22}) \tag{7}$$

Solving this equation, we get the long-run probabilities λ_1, λ_2:

$$\lambda_1 = \frac{1 - p_{22}}{2 - p_{11} - p_{22}} \qquad \lambda_2 = 1 - \lambda_1 = \frac{1 - p_{11}}{2 - p_{11} - p_{22}} \tag{8}$$

Given the parameterization of $\mathcal{D}(.)$ mentioned above in equation (3), we have $\mathbb{E}[y_t^2|s_t = k, \mathcal{I}_{t-1}] = \sigma_{k,t}^2$; that is, $\sigma_{k,t}^2$ is the variance of y_t conditional on the realization of s_t and the information set $\mathcal{I}_{t-1}$. As Haas et al. (2004) indicated that conditional on current regime k, the conditional variance of y_t is a function of past data and the additional regime-dependent vector of parameters θ_k. $\sigma^2(.)$ is a $\mathcal{I}_{t-1}$ - measurable function, defining the GARCH process in state k and under different forms of $\sigma^2(.)$, we obtain different skedastic specifications. In particular, considering the symmetric GARCH$(1,1)$ of Bollerslev (1986):

$$\sigma_{k,t}^2 \equiv \omega_k + \alpha_k y_{t-1}^2 + \beta_k \sigma_{t-1}^2, \qquad k = 1, \ldots, K \tag{9}$$

According to Paolella et al. (2012), ω_k is the long-run unconditional volatility level in state k, α_k is the preceding period's instant response parameter to a shock, β_k the memory parameter of $\sigma_{k,t}^2$ in regime k and $\theta_k \equiv (\omega_k, \alpha_k, \beta_k)'$. $\omega_k > 0$, $\alpha_k \geq 0$ and $0 \leq \beta_k < 1$ ensure positivity of $\sigma_{k,t}^2$ in regime k. $\alpha_k + \beta_k < 1$ ensures the covariance stationary within each regime k.

The asymmetric response of the conditional variance to both prior positive and negative shocks of the same size would be different if we consider for inverted leverage effects, as Bouri et al. (2017) and Katsiampa (2017) highlighted. As a result, we will consider different type of GARCH process, which is GJR$(1,1)$ model of Glosten et al. (1993) as below:

$$\sigma_{k,t}^2 \equiv \omega_k + (\alpha_k + \gamma_k \mathbb{I}\{y_{t-1} < 0\})y_{t-1}^2 + \beta_k \sigma_{k,t-1}^2, \qquad k = 1, \ldots, K \tag{10}$$

where $\mathbb{I}$ is an indicator function, which equals to 1 if the condition holds, and 0 otherwise. γ_k is the asymmetric parameter and $\theta_k \equiv (\omega_k, \alpha_k, \beta_k)'$. $\omega_k > 0$, $\alpha_k \geq 0$ and $0 \leq \beta_k < 1$ ensure positivity of $\sigma_{k,t}^2$ and $\alpha_k + \gamma_k E\left[\eta_{k,t}^2 \mathbb{I}\{\eta_{k,t} < 0\}\right] + \beta_k < 1$ ensures the covariance stationary of $\sigma_{k,t}^2$ in regime k. If $\gamma_k > 0$, negative shocks have a stronger effect than positive ones.

In the empirical study, we will examine 2 conditional variance specifications sGARCH and GJR-GARCH, with 3 innovations distributions, which are standard Normal ($\mathcal{N}$), the standardized Student-t ($\mathcal{S}$) and the standardized skewed-Student's t (sk$\mathcal{S}$), and up to three regimes, $K = 1, 2, 3$. In total, we look at total 18 models.

4.2 Estimation method

According to Haas et al. (2004), parameters of MSGARCH process can be estimated by either frequentist or Bayesian approach via Monte-Carlo simulation. In particular, Bayesian has some advantages such as avoiding the problem of convergence to local maxima or providing more accurate left-tail forecasts Ardia et al. (2018). Whether using frequentist approach or Bayesian method, we also need the likelihood function. The one-step ahead probability density function (PDF) of y_t is a mixture of K state dependent distributions with the mixing weights $z_{k,t|t-1} \equiv \mathbb{P}[s_t = k|\Psi, \mathcal{I}_{t-1}]$, $k = 1, \ldots, K$ being the one-step-ahead predictive probabilities:

$$f(y_t|\Psi, \mathcal{I}_{t-1}) \equiv \sum_{k=1}^{K} z_{k,t|t-1} f_{\mathcal{D}}(y_t|s_t = k, \Psi, \mathcal{I}_{t-1}) \tag{11}$$

In particular, $\Psi \equiv (\theta_1, \ldots, \theta_k, \xi_1, \ldots, \xi_k, \mathbf{P})$ is a mixture of model parameter. $f_{\mathcal{D}}(y_t|s_t = k, \Psi, \mathcal{I}_{t-1})$ is the conditional density of y_t in state $s_t = k$ given Ψ and $\mathcal{I}_{t-1}$. Mixing weights $z_{k,t|t-1} \equiv \sum_{k=1}^{K} p_{i,k} \eta_{i,t-1}$, $\eta_{i,t-1} \equiv \mathbb{P}[s_t = k|\Psi, \mathcal{I}_{t-1}]$ is the filtered probability at time $t-1$. The ML estimator $\hat{\Psi}$ is obtained by numerically maximizing the log-likelihood as follows:

$$ln\ \mathcal{L}(\Psi|\mathcal{I}_T) = \sum_{t=1}^{T} ln\ [f(y_t|\Psi, \mathcal{I}_{t-1})] \tag{12}$$

As Ardia et al. (2018) stated that when using Bayesian approach, we must take into consideration the kernel of the posterior distribution $f(\Psi|\mathcal{I}_T)$, depending on the prior distribution $f(\Psi)$ and the likelihood function. According to Trottier and Ardia (2016), getting the prior density for the transition matrix requires K rows to be independent and follow a Dirichlet prior with all hyper-parameters set to two. Then the following is the prior distribution:

$$f(\Psi) \propto f(\theta_1, \xi_1) \ldots f(\theta_K, \xi_K) f(\mathbf{P}) \mathbb{I}\{\bar{h_1} < \ldots < \bar{h_K}\} \tag{13}$$

In particular, $\mathbf{0}$ and $\mathbf{I}$ are vector of zeros and identity matrix respectively. $f_{\mathcal{N}}(.; \mu, \Sigma)$ is the multivariate normal density with μ is the mean vector, Σ is the covariance matrix. $\xi_{k,1}$ is the asymmetric parameter and $\xi_{k,2}$ is the tail parameter of the skewed Student's t-distribution in regime k. $\bar{h_1} < \ldots < \bar{h_K}$ is the unconditional variance while $\mathcal{CSC}_k$ is covariance stationary condition in regime k (Ardia et al., 2018). We have:

$$f(\boldsymbol{\theta_k}, \boldsymbol{\xi_k}) \propto f(\boldsymbol{\theta_k}) f(\boldsymbol{\xi_k}) \mathbb{I}\{(\boldsymbol{\theta_k}, \boldsymbol{\xi_k}) \in \mathcal{CSC}_{\|}\} \qquad (k = 1, \ldots, K)$$

$$f(\boldsymbol{\theta_k}) \propto f_{\mathcal{N}}(\boldsymbol{\theta_k}; \mathbf{0}, 1000 \times \boldsymbol{I}) \mathbb{I}\{\boldsymbol{\theta_k} > 0\} \qquad (k = 1, \ldots, K)$$

$$f(\boldsymbol{\xi_k}) \propto f_{\mathcal{N}}(\boldsymbol{\xi_k}; \mathbf{0}, 1000 \times \boldsymbol{I}) \mathbb{I}\{\xi_{k,1} > 0, \xi_{k,2} > 2\} \qquad (k = 1, \ldots, K)$$

$$f(\boldsymbol{P}) \propto \prod_{i=1}^{K} \Big(\prod_{j=1}^{K} p_{i,j} \Big) \mathbb{I}\{0 < p_{i,j} < 1\},$$

$$\tag{14}$$

We estimate model parameters using Bayesian method through Markov-chain Monte Carlo simulation since the form of posterior is unknown. We use diffusion priors as Ardia et al. (2018) and generate draws from the posterior by using adaptive random-walk Metropolis sampler of Vihola (2012). Moreover, some parameters are put under some constraints to ensure that volatility under the MS-GARCH specification cannot be generated by a single-regime specification Ardia et al. (2018).

Despite having some advantages over Maximum Likelihood Estimator, in terms of estimating VaR, Ardia et al. (2018) still mentioned that MLE performs as well as Bayesian method. As a result, we estimate model using both approaches and make some comparison between the two in our empirical analyses.

4.3 Downside risk forecasting

In the empirical study, we will evaluate so-called downside risk measure VaR. In order to determine the one-step-ahead VaR, we need the one-step-ahead cumulative distribution (CDF) of y_t; particularly, CDF is derived by integrating the PDF f over a range of value, defined in equation (11):

$$F_{t+1} = F(y_{T+1}|\Psi, \mathcal{I}_{T-1}) = \int_{-\infty}^{y_{t+1}} f(u|\Psi, \mathcal{I}_t)du \tag{15}$$

In the case of frequentist framework, the PDF and CDF is easily obtained by using the ML estimator $\hat{\Psi}$, while under Bayesian approach, we integrate out the parameter uncertainty to get CDF. As in Ardia et al. (2018), the predictive PDF that is based on a posterior sample $\Psi^{[m]}, m = 1, \ldots, M$ is as follows:

$$f(y_{T+1}|\mathcal{I}_T) = \int_{\Psi} f(y_{T+1}|\Psi, \mathcal{I}_T) f(\Psi|\mathcal{I}_T)d\Psi \approx \frac{1}{M} \sum_{m=1}^{M} f(y_{T+1}|\Psi^{[m]}, \mathcal{I}_T), \tag{16}$$

And the CDF is given by:

$$F(y_{T+1}|\mathcal{I}_T) = \int_{-\infty}^{y_{T+1}} f(u|\mathcal{I}_T)du \tag{17}$$

VaR(α) defines as the maximum loss we expect to occur with probability α over a period of time. Since we focus on the left tail where the negative returns are located, $\alpha \in (0,0.5)$ is the shortfall probability. Caporale and Zekokh (2019) estimated the one-step-ahead VaR at level α risk level by inverting the predictive CDF, which is equal to:

$$VaR_\alpha(X) = inf\{xF_X(x) \geq \alpha\} = F_{\bar{X}}(x) \tag{18}$$

where $F_X(x)$ is the cumulative distribution function of X. For the continuous distribution, VaR is the α-quantile of the return distribution.

Nevertheless, VaR has some disadvantages. Lacking of sub-additivity characteristics causes VaR not a coherent risk measure. In particular, Acerbi and Tasche (2002); Christoffersen and Pelletier (2004) concluded that VaR could only estimate a threshold for the losses but ignored the amount of losses above the threshold. Therefore, another adhoc risk measure, Expected Shortfall, which is both coherent and takes the real extent of severe losses into consideration, seems to be the appropriate one for estimating the risk. However, in practice VaR is still widely used because of its simplicity and Nieto and Ruiz (2016) argued that for a large class of distributions, missing sub-additivity characteristics could be erased. Therefore, in this paper, we use VaR at different risk levels to mitigate the second shortcoming mentioned.

4.4 Backtesting methods

Philippe (2001) stated that Backtesting is a statistical methodology for determining if realized losses match expected VaR estimations. According to Christoffersen (2008), backtesting implies testing the sequence of ex-ante VaR forecasts, $\{VaR_t^\alpha\}_{t=1}^{t=T}$, computed against the observed returns, $\{r_t\}_{t=1}^{t=T}$. Specifically, Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements (1996) has organized backtesting process as the crucial aspect of the internal risk management validation. In order to certify that our model estimates BTC risk accurately, we will use two popular VaR Backtesting approaches: the Likelihood Ratio (LR) test of Unconditional Coverage (UC) (LRUC) of Kupiec (1995) and the LR test of Conditional Coverage (LRCC) of Christoffersen (1998). The test of *unconditional coverage* by Kupiec (1995), which does not take into account when the violations occur, is to test whether during the holding period, the densities of exceptions are correlated to the defined confidence level α. Exceptions are defined as the situation when the ex-post loss of returns surpass the ex-ante VaR value. However, in addition to counting the number of violations, an ideal VaR model should potentially be able to measure the distribution of the exceptions, which should be unrelated to each other, namely *independence property*. Christoffersen (1998) indicated that during the recession time, if large losses occurred, devastating events were more likely to happen such as bankruptcy. Therefore, not accounting for market volatility or its correlations among other sectors can make VaR model to be invalid. Thus, VaR back-testing techniques testing for the independence property are classified as tests of *conditional coverage* (Katsenga, 2013).

The LR_{UC} and LR_{CC} tests are built on a hit sequence, a dummy variable that indicates a loss that exceeds the VaR level $\{I_t \equiv \mathbb{I}\{y_t \leq VaR_t^\alpha\}$. Where VaR_t^α indicates the VaR prediction at risk level of α at time t, $\mathbb{I}(.)$ is the hit sequence, which equals to one if r_t is smaller than VaR_t^α on day t, i.e. the actual loss exceeds the expected loss, and zero otherwise. Christoffersen (1998) stated that in order to examine the accuracy of a VaR model, the hit sequence should

meet two properties: UC property and the independence property. Thus, the null hypothesis of an adequate risk model can be formulated as:

$$H_0 : I_t \overset{\text{iid}}{\sim} Bernoulli(\alpha) \tag{19}$$

$$f(I_t, \lambda) = (1 - \lambda)^{1-I_t} \lambda_t^I \tag{20}$$

According to the test proposed by Kupiec (1995), namely the Proportion of Failure (POF), the LR_{UC} test focuses on the UC property, testing whether the promised shortfall probability α is equal to the actual value, λ, particularly, $\mathrm{E}[I_t] = \mathbb{P}(I_t = 1) = \lambda = \alpha$. Based on equation (199, the conditional likelihood function of the hit sequence $\{I_t\}_{t=1}^{T}$ with I_t follows Bernoulli(λ) distribution:

$$\mathcal{L}(\lambda | I_1, \ldots, I_T) = \prod_{t=1}^{T} \lambda_t^I (1 - \lambda)^{1-I_t} = \lambda^x (1 - \lambda)^{T-x} \tag{21}$$

where $x = \sum_{t=1}^{T} I_t$ is the total number of violations. λ is unknown and being estimated by maximizing the logarithm of $\mathcal{L}(\lambda | I_1, \ldots, I_T)$. Then the maximum estimator is equal to $\hat{\lambda} = \frac{x}{T}$. Then we get the optimized likelihood function as follows:

$$\mathcal{L}(\hat{\lambda} | I_1, \ldots, I_T) = \hat{\lambda}^x (1 - \hat{\lambda})^{T-x} \tag{22}$$

Under the null hypothesis that $\lambda = \alpha$, we get the new likelihood function:

$$\mathcal{L}(\alpha | I_1, \ldots, I_T) = \alpha^x (1 - \alpha)^{T-x} \tag{23}$$

Then the likelihood ratio under unconditional coverage LR_{UC} test that has a Chi-square distribution with one degree of freedom is defined as:

$$LR_{UC} = -2[\ln \mathcal{L}(\alpha | I_1, \ldots, I_T) - \ln \mathcal{L}(\hat{\lambda} | I_1, \ldots, I_T)$$
$$= -2 \ln(\alpha^x (1 - \alpha)^{T-x}) + 2 \ln(\hat{\lambda}^x (1 - \hat{\lambda})^{T-x}) \sim \chi^2(1) \tag{24}$$

However, this test only focuses on how often violations occur over a given period but not on how those violations are distributed. Consequently, it has low power in detecting VaR measures that fulfills the UC property but has time-dependent violations Campbell (2005); Christoffersen (1998). Therefore, according to Christoffersen (1998), one of the most commonly known test to tackle this issue, *conditional coverage*, which examines the independence property, or exception clustering, is the independence test, namely Markov test. Markov test examines whether the dependence of VaR exceptions is affected by what happened on the preceding day (Campbell, 2005). In particular, suppose that the hit sequence $\{I\}_{t=1}^{T}$ follows first-order Markov-chain with transition probability matrix:

$$\Pi = \begin{bmatrix} \pi_{00} & \pi_{10} \\ \pi_{01} & \pi_{11} \end{bmatrix} = \begin{bmatrix} 1 - \pi_{01} & 1 - \pi_{11} \\ \pi_{01} & \pi_{11} \end{bmatrix}, \tag{25}$$

where $\pi_{ij} = P(I_t = j|I_{t-1} = i)$, i.e π_{01} is the probability of no exceptions tomorrow given no violations today. Then the corresponding conditional likelihood function of the hit sequence $\{I\}_{t=1}^T$ is as:

$$\mathcal{L}(\Pi|I_1,\ldots,I_T) = \pi_{01}^{n_{01}}(1-\pi_{01})^{n_{00}}\pi_{11}^{n_{11}}(1-\pi_{11})^{n_{10}}, \tag{26}$$

where n_{ij} is the number of observations with a j coming after a value i and $i,j = 0,1$.

Maximizing the logarithm of the above equation $\mathcal{L}(\Pi|I_1,\ldots,I_T)$, we get the following matrix of ML estimates:

$$\hat{\Pi} = \begin{bmatrix} \hat{\pi_{00}} & \hat{\pi_{10}} \\ \hat{\pi_{01}} & \hat{\pi_{11}} \end{bmatrix} = \begin{bmatrix} \frac{n_{00}}{n_{00}+n_{01}} & \frac{n_{10}}{n_{10}+n_{11}} \\ \frac{n_{01}}{n_{01}+n_{01}} & \frac{n_{11}}{n_{10}+n_{11}} \end{bmatrix}, \tag{27}$$

and the optimized likelihood function is:

$$\mathcal{L}(\hat{\Pi}|I_1,\ldots,I_T) = \hat{\pi}_{01}^{n_{01}}(1-\hat{\pi}_{01})^{n_{00}}\hat{\pi}_{11}^{n_{11}}(1-\hat{\pi}_{11})^{n_{10}} \tag{28}$$

Under the null hypothesis, the violations are independent across days, then the probabilities should be the same, $\pi = \pi_0 = \pi_1$. That is, the probability that an exception occur after a day of no exception is the same as occurring after a day of a violation Campbell (2005). We have the likelihood function:

$$\mathcal{L}(\alpha|I_1,\ldots,I_T) = \alpha^{n_{01}} + n_{11}(1-\alpha)^{n_{01}} + n_{11} \tag{29}$$

The relevant test statistic for independence of exceptions is given by:

$$LR_{Ind} = -2ln\left[\frac{\mathcal{L}(\hat{\pi})}{\mathcal{L}(\hat{\Pi})}\right] \overset{iid}{\sim} \chi^2(1) \tag{30}$$

However, Markov test still has some drawbacks, which is detecting clustering behavior of returns. In particular, Markov test can only examine the independence of violations on two consecutive days, while the likelihood of VaR exception might determine not by whether one occurred the day before, but by whether one occurred one week or two weeks before. Therefore, in order to detect VaR with the most accurate result, Christoffersen (1998) proposed a joint test statistic for simultaneously testing the number of violations and the independence of exceptions, namely *conditional coverage test*. Under an adequately specified risk model, the probability of a hit in t should be independent of the outcome of the previous I_{t-1} and equal to the promised shortfall probability α. The LR_{CC} test is defined as:

$$LR_{CC} = -2\left[ln\,\mathcal{L}(\alpha|I_1,\ldots,I_T) - ln\,\mathcal{L}(\hat{\Pi}|I_1,\ldots,I_T) \rightarrow \chi^2(2) \tag{31}$$

Under null hypothesis, the LR_{CC} test follows Chi-square distribution with two degrees of freedom, given that there are two separate LR-statistics in the test, one

Figure 1: Percentage differences in prices and returns between Yahoo and BitFinex datasources
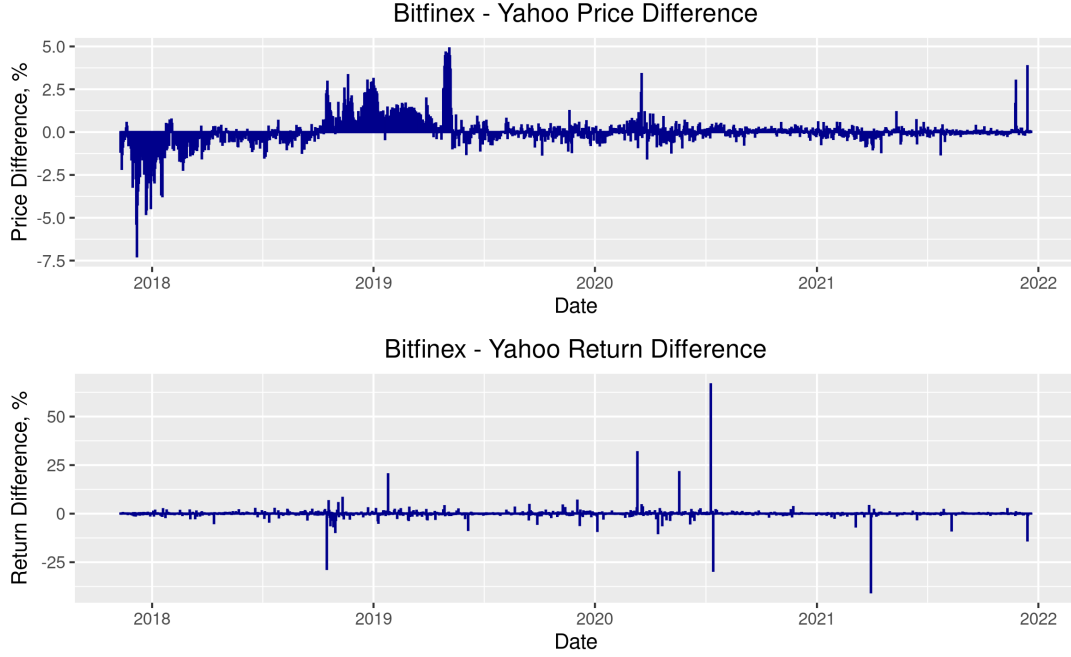
Notes: Yahoo BTC-USD data is only available from November 10, 2018, while BitFinex BTC-USD data is available from January 01, 2015.

for the correct number of exceptions and one for the independence of violations. However, Campbell (2005) mentioned that combing two tests could decrease the ability in detecting VaR measures which violated only one of the two properties (Katsenga, 2013). Therefore, in case of the null hypothesis rejection, we do not know which property is violated, in particular, the model passes the joint test while still failing either the unconditional coverage test or independence test (Campbell, 2005). Another shortcoming of the LR_{CC} test is that it considers only one specific dependence structure, i.e. a first-order Markov dependence, and ignores more complex temporal dependencies of the violations Campbell (2005).

5 Empirical analysis

5.1 Data description

According to Alexander and Dakos (2020), more than 60% papers published in academic journals use data from unreliable sources, so-called price-tracking website. There are a lot of inconsistencies among traded prices on different sources. We sourced BTC-USD data from 2 popular sources: BitFinex and Yahoo Finance. Figure 1 shows the differences in percentage of daily reported closing prices, and

derived log-return between 2 sources. Interestingly, price differences are not random, especially for the period of 2018 to mid-2019, indicating some systematic differences in either data collection, consolidation, or reporting. Differences in returns can sometimes be very substantial: in some cases reach ±25%. In the end, we go with data from BitFinex due to higher availability. Data used is daily closing prices of Bitcoin (in $) from 2015-01-01 to 2021-12-20, for a total of 2,543 observations.

Table 1: Summary statistics of Bitcoin percentage log-returns

| Min	Max	Mean	Median	Std	Skewness	Kurtosis	Jb p-val
-48.09	23.72	0.197	0.176	4.078	-0.966	12.792	0

Notes: The total number of returns is 2,543. Max stands for maximum; Min for minimum; Std for Standard deviation; Skew for Skewness; Kurt for Kurtosis and Jb for Jarque-Bera Normality test.

Table 1 describes some descriptive statistics of Bitcoin returns. In particular, kurtosis is 12.8 and skewness is -0.966, which show that returns are leptokurtic and negative skewed. Together with the value of JB p-value, being equal to 0, it shows that Bitcoin returns are not normally distributed. As we can see from Figure 5, there is no strong evidence that indicates autocorrelation. However, the ACF of the absolute returns shows long diminishing autocorrelation over many lags, indicating possible variance clustering, where large changes happen close to each other. Furthermore, figure percentage log-returns of Figure 2 confirms conditional heteroscedasticity. The annualized empirical volatility is $\sqrt{365} \times 4.08 = 78\%$.

5.2   In-sample Analysis

We will first begin with the in-sample analysis, in which 18 models are fitted to the whole dataset of demeaned and AR(1)-filtered log returns using Maximum Likelihood Estimation (MLE). To assess the model's goodness-of-fit for models estimated with MLE, Akaike Information Criterion (AIC) (Akaike, 1998) and Bayesian Information Criterion (BIC) (Schwarz, 1978) will be used to select the most appropriate model in terms of a trade-off between the goodness of fit and the number of parameters (Guidolin and Pedio, 2018). In addition to MLE, Bayesian Estimation can also be used to estimate parameters of MSGARCH models, and can bring good results (Ardia et al., 2019). In order to verify this, the same routines and tests will also be done with the aforementioned models, estimated using Bayesian Estimation (BE). Since the posterior distribution of parameters do not have an analytical form, we sample estimated parameters using Markov Chain Monte Carlo (MCMC) methods (Ardia et al., 2019). On this front, Deviance Information Criterion (DIC) is used to evaluate goodness-of-fit (Spiegelhalter et al., 2002).

Figure 2: Summary statistics of Bitcoin percentage log-returns
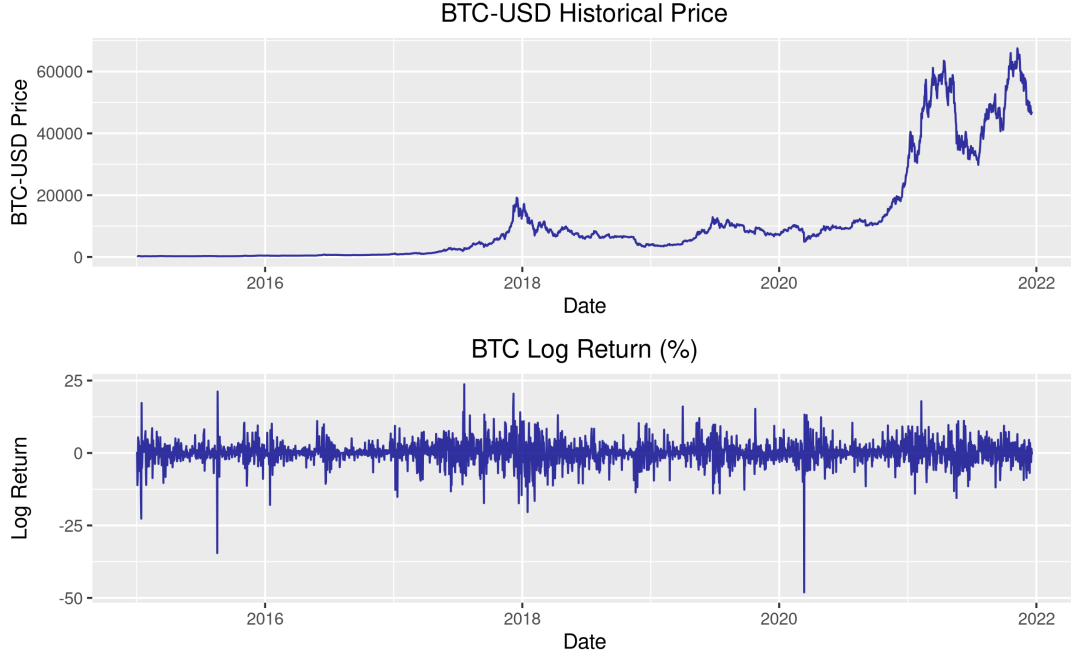
Notes: The time series range from January 01, 2015 to Dec 20, 2021 for the closing prices shown in the upper plot and from January 01, 2015 to Dec 20, 2021 for the percentage log-returns in the lower plot

As we can see from Table 2, for all distribution and variance specifications, except for the case GJR(2)-sk$\mathcal{S}$ model, MSG(K) models in both 2 and 3 regimes almost outperform SRG models, which is the same as the conclusions of Ardia et al. (2019) and Maciel (2021). Particularly, for all states, sGARCH models display lower values than GJR-GARCH. Regarding the volatility specifications, the best in-sample models are sGARCH(2)-sk$\mathcal{S}$ model and sGARCH(3)-sk$\mathcal{S}$ model based on AIC, which shows no evidence of inverted leverage effects. Bayesian Estimation, on the other hand, prefers GJR-GARCH with three regimes.

Table 2: In-sample model selection for Bitcoin

Model	$\mathcal{N}$			$\mathcal{S}$			sk$\mathcal{S}$		
	AIC	BIC	DIC	AIC	BIC	DIC	AIC	BIC	DIC
Regime $k = 1$									
sGARCH	14004.3	14021.9	14004.8	13245.0	13268.4	13246.0	13240.5	13269.7	13241.4
GJR-GARCH	13988.5	14011.9	13988.4	13247.0	13276.3	13250.4	13242.6	13277.6	13276.5
Regime $k = 2$									
sGARCH	13302.1	13348.9	13304.1	13170.7	13220.7	13220.7	**13164.1**	13234.2	13205.2
GJR-GARCH	13455.5	13513.9	13304.1	13199.8	**13209.9**	13209.9	13329.6	13411.3	13199.0
Regime $k = 3$									
sGARCH	13236.8	13324.4	13339.6	13174.8	13279.9	13215.1	13168.6	13291.3	13297.0
GJR-GARCH	13352.6	13457.7	13277.6	13189.5	13312.1	38241.3	13174.8	13315.0	**13183.5**

Notes: The table reports the Bayesian information criterion (BIC) and the Akaike information criterion (AIC) for the 18 specified models. The smallest values are printed in bold and highlight the best in-sample models. sGARCH is the symmetric GARCH of Bollerslev (1986) and GJR is the asymmetric GARCH model of Glosten et al. (1993). $\mathcal{N}$, $\mathcal{S}$ and sk$\mathcal{S}$ are the standard normal, standardized Student's t-distribution and standardized skewed-Student's t-distribution, respectively.

According to the result from Table 2, the parameter estimates of the best in-sample model, namely sGARCH(2)-sk$\mathcal{S}$, and its counterpart are presented in Table 3. ω_k, α_k, β_k are the sGARCH parameters, ν_k is the tail parameters and δ_k is the asymmetry parameter of the standardized skewed Student's t-distribution in regime k. ν_1 and δ_1 indicate excess kurtosis and right skewness, respectively. The sum of α_1 and β_1 is close to 1, indicating high persistence of the volatility dynamics.

Unlike Ardia et al. (2019), we find that a MSG(2) and skewed Student's t-distribution to be the best in-sample model based on AIC. While BIC points to standard Student's t-distribution, two-regime is still the common ground, indicating that volatility process can switch between two regimes. And since standard Student's t-distribution is a special case of skewed Student's t-distribution, we can make use of the latter. The MSG(2) parameters show that the BTC volatility process switch between a low- and high-volatility regime with an annualized uncondition volatility. of 99% in regime 1 and 54% in regime 2. Both regimes are extremely durable with the staying probabilities $p_{kk} = 0.98$, $k = 1, 2$. On average, low- and high-volatility regimes are equally likely as indicated by the stable probabilities $\lambda_1 = 0.55$ and $\lambda_2 = 0.45$. In particular, ω_2 is considerably smaller than ω_1, which shows the same pattern as annualized unconditional volatility between two states. Further, both α_2 and β_2 are higher than α_1 and β_1, indicating a stronger reaction to shocks and more volatility memory in the low-volatility regime. This situation can be presumably explained by finding of Haas et al. (2004), the effect of the pressured market circumstances, COVID-19 to the market. The estimated probability transition matrix of the sGARCH(2)-sk$\mathcal{S}$ model is:

$$P = \begin{bmatrix} 0.9877 & 0.0123 \\ 0.0152 & 0.9848 \end{bmatrix} \tag{32}$$

To understand more about the volatility regimes of BTC-USD returns, we can use the estimated probability of being in a certain regime for certain points in time (Guidolin and Pedio, 2018), here denoted $z_{k,t|\Psi,\mathcal{I}}$. There are 2 types of probabilities indicating current regimes: filtered probability and smoothed probability. Filtered probability $z_{k,t|\Psi,\mathcal{I}_{t-1}}$, which is a real-time inference on the state probabilities, depends only on information up to time t. Smoothed probability $z_{k,t|\Psi,\mathcal{I}_T}$, on the other hand, depends on the entire information set available, and is more suitable for retrospective analysis conducted here. Figure 3 shows the ML-estimated smoothed probability of returns being in high-volatility regime $k = 1$. There are 2 long periods that has continuously high (smoothed) probability of being in the high-volatility regime, that is around the year 2017 to 2019, and from 2021 to 2022. This aligns quite well with the development ob BTC-USD price, which exhibited wild fluctuation during the same periods.

As for Bayesian Estimation, our largest model in the repertoire achieves the lower DIC, hence is used for in-sample analysis. The chosen model has GJR-GARCH as the underlying process for return volatility, which has a leverage effect

Table 3: Parameter estimates

	Single-regime		Two-regime		Three-regime	
Regime k = 1						
ω_1	0.1967***	(0.0627)	0.9424***	(0.3635)	0***	$(< 1e - 10)$
α_1	0.1115*	(0.0731)	0.0721**	(0.0413)	0.0145***	$(< 1e - 10)$
β_1	0.8873***	(0.0008)	0.8924***	(0.0178)	0.9854***	$(< 1e - 10)$
ν_1	3.3585***	(0.1494)	5.2787***	(1.0495)	2.8459***	$(< 1e - 10)$
δ_1	0.9513***	(0.0187)	0.9149***	(0.0324)	0.9344***	$(< 1e - 10)$
$p_1 1$	1		0.9877		0.9762	
aUV$_1$	89.2114		98.1903		0.6639	
Regime k = 2						
ω_2	-		0.2971**	(0.1467)	0.0038***	$(< 1e - 10)$
α_2	-		0.0781	(0.061)	0.0098***	$(< 1e - 10)$
β_2	-		0.9213***	(0.0004)	0.9842***	$(< 1e - 10)$
ν_2	-		2.3258***	(0.0814)	2.2125***	$(< 1e - 10)$
δ_2	-		0.954***	(0.027)	1.0019***	$(< 1e - 10)$
$p_2 2$	-		0.9848		0.9720	
$textaUV_2$	-		59.2302		9.4739	
Regime k = 3						
ω_3	-		-		4.0207***	$(< 1e - 10)$
α_3	-		-		0.0958***	$(< 1e - 10)$
β_3	-		-		0.7831***	$(< 1e - 10)$
ν_3	-		-		5.6673***	$(< 1e - 10)$
δ_3	-		-		0.8618***	$(< 1e - 10)$
$p_3 3$	-		-		0.9803	
aUV$_3$	-		-		110.1367	

Notes: The table presents the maximum-likelihood parameter estimates of the K-regime sGARCH-skS models, K = 1; 2; 3. sGARCH is the symmetric GARCH model of Bollerslev (1986). skS is the standardized skewed Student's t-distribution with tail parameter ν and asymmetry parameter δ. p_{kk} is the staying probability of state k. aUV is the annualized unconditional volatility (in %). In parentheses are the standard errors. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively

parameter γ, and is inline with findings from Ardia et al. (2019). $K = 3$ indicates that the volatility process of out BTC-USD returns follows 3 distinct regimes, which low (26.98%), medium (71.18%), and high (123.7%) unconditional volatility, as shown in Table 6. In-sample predicted smoothed probability shows that this model specification changes regimes very eagerly and disruptively, without staying at one regime for too long. As can be seen in Figure 6, the model stays medium-volatility regime the most, followed by relatively long periods of low-volatility regime. High-volatility regime, on the other hand, appears sparsely and only shortly, interloping between medium-volatility periods.

Figure 3: Smoothed probabilities and annualized volatilities
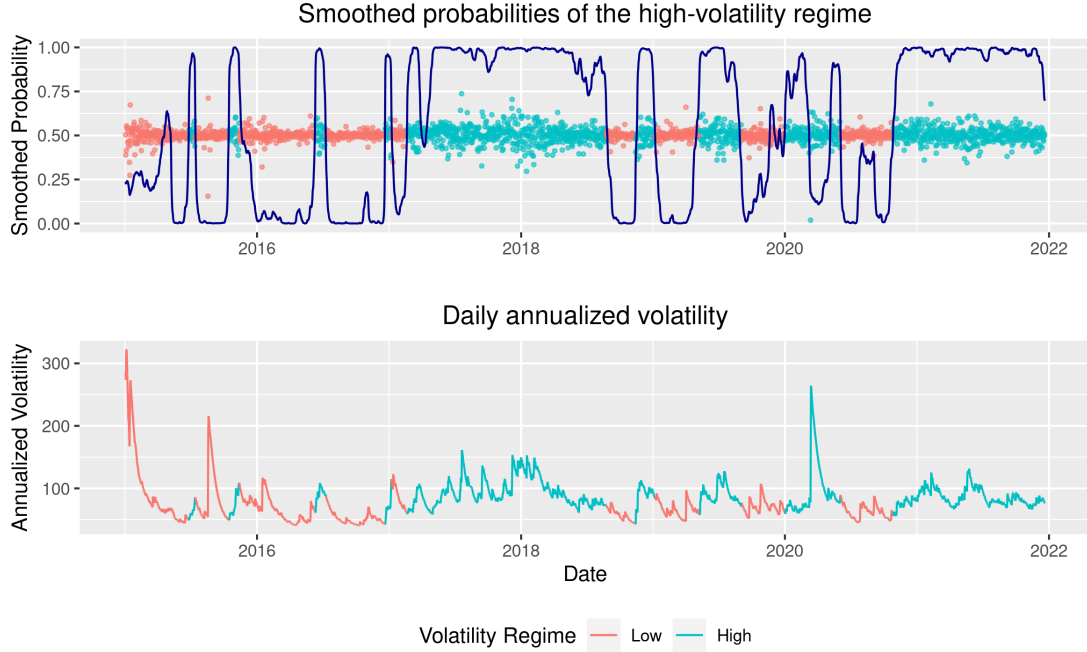
Notes: The time series range from January 01, 2015 to Dec 20, 2021. The graphs are computed from the two-regime sGARCH model with a standardized skewed Student's t-distribution. Parameters in the models are estimated using MLE. The estimated smoothed probabilities of the high-volatility regime (1st regime) are shown in the upper graph and the BTC log-returns (points). The bottom graph shows the annualized volatilities in percentage. The colors of each point and color segments of the annualized volatility series indicate the regime to which the points or segments are predicted to belong.

5.3 Out-of-sample Analysis

For out-of-sample analysis, we use a rolling window of size 2,189 for both MLE and BE. We construct 354 out-of-sample one-step VaR predictions from January 01, 2021 to December 20, 2021. We update the parameters of our models after every five one-step-ahead forecasts. The reason we use a rather large rolling window for estimation is to see if data in the long past can be indicative of recent record-breaking swings in BTC-USD returns. In 2021 alone, BTC-USD prices went from around \$30,000 to more than \$60,000, then crashed back to \$30,000 before spiking up to yet another new record of around \$68,000 and slid down from there.

Table 4 presents the backtesting results for one-step-ahead VaR at the 0.5%, 1%, 5% and 10% risk levels. As mentioned above, the purpose of backtesting here is to consider whether the VaR calculation in practice is significantly dfferent from VaR from model estimation. We can see that among three regimes, single-regime GARCH have more rejections than two- and three-regime GARCH. Both LR_{UC} and LR_{CC} test reject the adequacy of MSGARCH(1) at the 5% and 10% level of significance for the 10% risk level. Moreover, we also reject the null hypothesis of

Figure 4: Value-at-risk forecasts
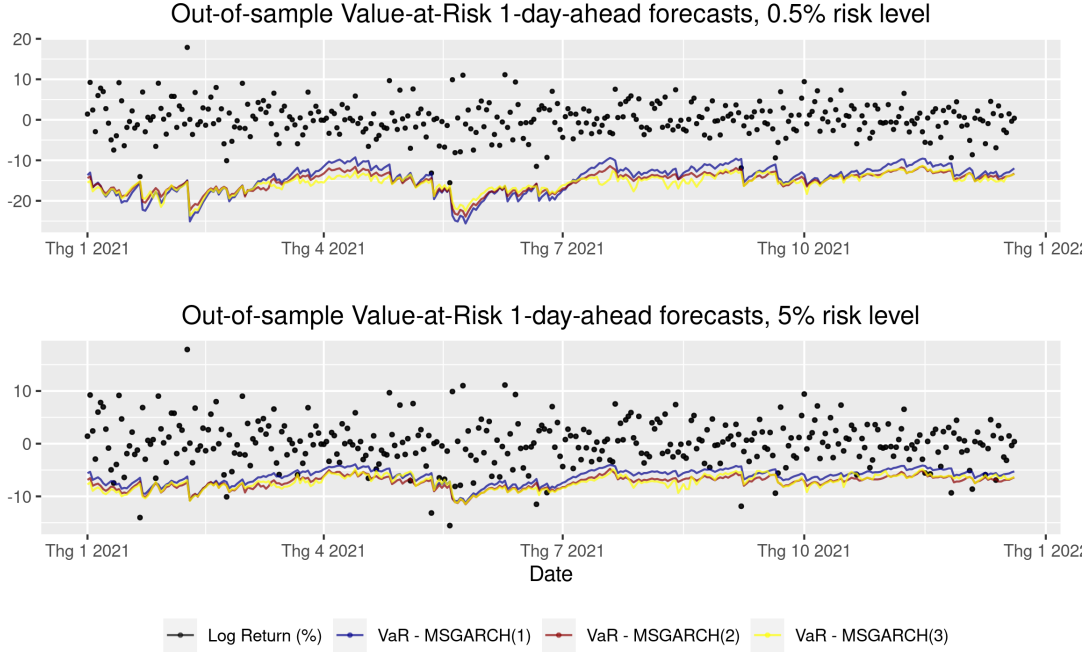
Notes: The time series range from January 01, 2015 to Dec 20, 2021. The out-of-sample period ranges from Jan 01, 2021 to Dec 20, 2021. The top graph shows the one-day-ahead Value-at-risk (VaR) forecasts at the 0.5% risk level computed with the single regime (blue line), two-regime (red line) and three-regime (green line) sGARCH-sk$\mathcal{S}$ model. The percentage log-returns are depicted as the dots. sGARCH is the symmetric GARCH model of Bollerslev (1986) and sk$\mathcal{S}$ is the standardized skewed Student's t-distribution. The bottom graph depicts the one-day-ahead VaR forecasts at the 5% risk level. The single-regime GARCH tends to lead to (absolute) smaller VaR values and thus, the observed number of violations is higher than for the two-regime GARCH model.

correct VaR forecasting for two–regime and three–regime MSGARCH models for the 0.5% VaR at the 10% level of significance. According to findings of Ardia et al. (2019), Maciel (2021), SGR model is not a valid risk model since it underestimates the BTC risk. However, based on the result in Table 5, it seems that all three models have almost the same power in detecting the VaR violations. Particularly, in Table 5, the number of hits of three models are all lower than the expected number of hits for all risk levels. Again, this can be seen graphically in Figure 4. We can conclude that the models tend to overestimate the BTC risk, which contradicts to the findings of Ardia et al. (2019).

The models of which parameters are estimated with Bayesian method tell a similar story. According to Table 7, the two- and three-regime models, however, see rejection of unconditional test, at 10% and 0.5% risk levels while single-regime model rejects both tests at 10% risk level. These test results stem from overestimation of risks for all models, more so for the multiple regime ones, showed in the Table 5. Again, empirical results of Ardia et al. (2019) disagree with the results here.

Table 4: Backtesting results for 1-step-ahead volatility forecasts

	VaR risk level	0.005	0.01	0.05	0.1
MSGARCH(1)	LR_{UC}	0.8652	0.2319	0.5823	**0.0220**
	LR_{CC}	0.9745	0.4412	0.2584	**0.0722**
MSGARCH(2)	LR_{UC}	**0.0596**	0.8098	0.3498	0.6675
	LR_{CC}	0.1696	0.9279	0.3622	0.9105
MSGARCH(3)	LR_{UC}	**0.0596**	0.8098	0.8636	0.6675
	LR_{CC}	0.1696	0.9279	0.4168	0.9105

Notes: The table presents the backtesting results for the 1-step-ahead value-at-risk (VaR) at risk levels 0.5%, 1%, 5% and 10%. Shown are the p-values of the Likelihood Ratio test of Unconditional Coverage (LR_{UC}) and the Likelihood Ratio test of Conditional Coverage (LR_{CC}) for the K-regime symmetric sGARCH (Bollerslev, 1986) models, K = 1,2,3, with a standardized skewed Student's t-distribution.

There are a few possible explanation for the disparity in empirical results. First, BTC-USD historical price data is obtained from different sources. As we have discussed earlier in this paper, the return differences derived from prices between sources can be substantial. Figure 1 is a demonstration. secondly, the data periods used are different. Since 2018, BTC has experienced some significant growth in both popularity and value. In 2018, BTC-USD price peaked around less then $20,000 and eventually hovers around $7,000. BTC-USD reached $60,000 in April 2021 before crashing down to $32,000 in May 2021 and bounced back to more than $65,000 towards the end of 2021 before another dip to $50,000. Such consistent large swings may empirically favor a single high-volatility regime instead of multiple regimes. Lastly, the rolling window size used in our study may not be optimal, since it includes a large portion of data from a relatively distant past, the signals of which may overpower more recent patterns in return volatility.

Table 5: Expected and observed number of violations

	VaR risk level	0.005	0.01	0.05	0.1
Expected number of hits		5.31	10.62	53.1	106.2
Observed number of hits					
	MSGARCH(1)	2	6	20	49
	MSGARCH(2)	0	4	14	33
	MSGARCH(3)	0	4	17	33

6 Conclusion

Taking recent rising popularity of Crytocurrencies, in particular Bitcoin, into account, it is crucial to select the accurate models for estimating risk. The aim of

this paper is to assess whether there are regime changes in the BTC volatility dynamics. In-sample analyses highly favor multiple regimes for models estimated by either Maximum Likelihood Estimation or Bayesian estimation. Only BE points towards model with leverage effect using GJR-GARCH as the underlying process of return volatility. The backtesting results, on the other hand, are less decisive on the matter of single vs multiple regimes. In the out-of-sample analysis, 2- and 3-regime models selected and estimated by both MLE and BE only give somewhat better backtesting results at 10% risk level, while single-regime model is slightly better at 0.5% risk level. In all cases, models tend to overestimate risks for the backtesting period used. Our findings contradict that of relevant papers, such as Ardia et al. (2019), Caporale and Zekokh (2019), Maciel (2021). The differences may be attributed to different amount of data used, different backtesting period, and rolling window size used for parameter estimation.

References

Acerbi, C. and D. Tasche (2002). Expected shortfall: a natural coherent alternative to value at risk. *Economic notes 31*(2), 379–388.

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike*, pp. 199–213. Springer.

Alexander, C. and M. Dakos (2020). A critical investigation of cryptocurrency data and analysis. *Quantitative Finance 20*(2), 173–188.

Ardia, D., K. Bluteau, K. Boudt, and L. Catania (2018). Forecasting risk with markov-switching garch models: A large-scale performance study. *International Journal of Forecasting 34*(4), 733–747.

Ardia, D., K. Bluteau, K. Boudt, L. Catania, and D.-A. Trottier (2019). Markov-switching garch models in r: The msgarch package. *Journal of Statistical Software 91*(4), 1–38.

Ardia, D., K. Bluteau, and M. Rüede (2019). Regime changes in bitcoin garch volatility dynamics. *Finance Research Letters 29*, 266–271.

Basel Comittee on Banking Supervision (1996, Jan). Amendment to the capital accord to incorporate market risks. `https://www.bis.org/publ/bcbs24.htm`.

Baur, D. G., K. Hong, and A. D. Lee (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money 54*, 177–189.

Bauwens, L., A. Dufays, and J. V. Rombouts (2014). Marginal likelihood for markov-switching and change-point garch models. *Journal of Econometrics 178*, 508–522.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, 542–547.

Bouri, E., G. Azzi, and A. H. Dyhrberg (2017). On the return-volatility relationship in the bitcoin market around the price crash of 2013. *Economics 11*(1).

Cai, J. (1994). A markov model of switching-regime arch. *Journal of Business & Economic Statistics 12*(3), 309–316.

Campbell, S. D. (2005). A review of backtesting and backtesting procedures.

Caporale, G. M. and T. Zekokh (2019). Modelling volatility of cryptocurrencies using markov-switching garch models. *Research in International Business and Finance 48*, 143–155.

Christoffersen, P. (2008). Backtesting. *Available at SSRN 2044825*.

Christoffersen, P. and D. Pelletier (2004). Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics 2*(1), 84–108.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, 841–862.

Chu, J., S. Chan, S. Nadarajah, and J. Osterrieder (2017). Garch modelling of cryptocurrencies. *Journal of Risk and Financial Management 10*(4), 17.

Corbet, S., A. Meegan, C. Larkin, B. Lucey, and L. Yarovaya (2018). Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters 165*, 28–34.

Dueker, M. J. (1997). Markov switching in garch processes and mean-reverting stock-market volatility. *Journal of Business & Economic Statistics 15*(1), 26–34.

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

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.

Fry, J. (2018). Booms, busts and heavy-tails: The story of bitcoin and cryptocurrency markets? *Economics Letters 171*, 225–229.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance 48*(5), 1779–1801.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of financial economics 42*(1), 27–62.

Guidolin, M. and M. Pedio (2018). *Essentials of time series for financial applications*. Academic Press.

Haas, M., S. Mittnik, and M. S. Paolella (2004). A new approach to markov-switching garch models. *Journal of financial Econometrics 2*(4), 493–530.

Hamilton, J. D. and R. Susmel (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of econometrics 64*(1-2), 307–333.

Katsenga, G. Z. (2013). *Value at risk (VaR) backtesting'Evidence from a South African market portfolio'*. Ph. D. thesis.

Katsiampa, P. (2017). Volatility estimation for bitcoin: A comparison of garch models. *Economics Letters 158*, 3–6.

Kim, T. (2017). On the transaction cost of bitcoin. *Finance Research Letters 23*, 300–305.

Klaassen, F. (2002). Improving garch volatility forecasts with regime-switching garch. In *Advances in Markov-switching models*, pp. 223–254. Springer.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The J. of Derivatives 3*(2).

Maciel, L. (2021). Cryptocurrencies value-at-risk and expected shortfall: Do regime-switching volatility models improve forecasting? *International Journal of Finance & Economics 26*(3), 4840–4855.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*, 21260.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347–370.

Nieto, M. R. and E. Ruiz (2016). Frontiers in var forecasting and backtesting. *International Journal of Forecasting 32*(2), 475–501.

Paolella, M. S., M. Haas, L. Bauwens, C. M. Hafner, and S. Laurent (2012). Mixture and regime-switching garch models. *Wiley handbooks in financial engineering and econometrics* (3), 71–102.

Philippe, J. (2001). Value at risk: the new benchmark for managing financial risk. *NY: McGraw-Hill Professional*.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461–464.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the royal statistical society: Series b (statistical methodology) 64*(4), 583–639.

Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements (1996, Jan). Amendment to the capital accord to incorporate market risks. `https://www.bis.org/publ/bcbs22.htm`.

Trottier, D.-A. and D. Ardia (2016). Moments of standardized fernandez–steel skewed distributions: Applications to the estimation of garch-type models. *Finance Research Letters 18*, 311–316.

Vihola, M. (2012). Robust adaptive metropolis algorithm with coerced acceptance rate. *Statistics and Computing 22*(5), 997–1008.

7 Appendix

Figure 5: Autocorrelation Functions
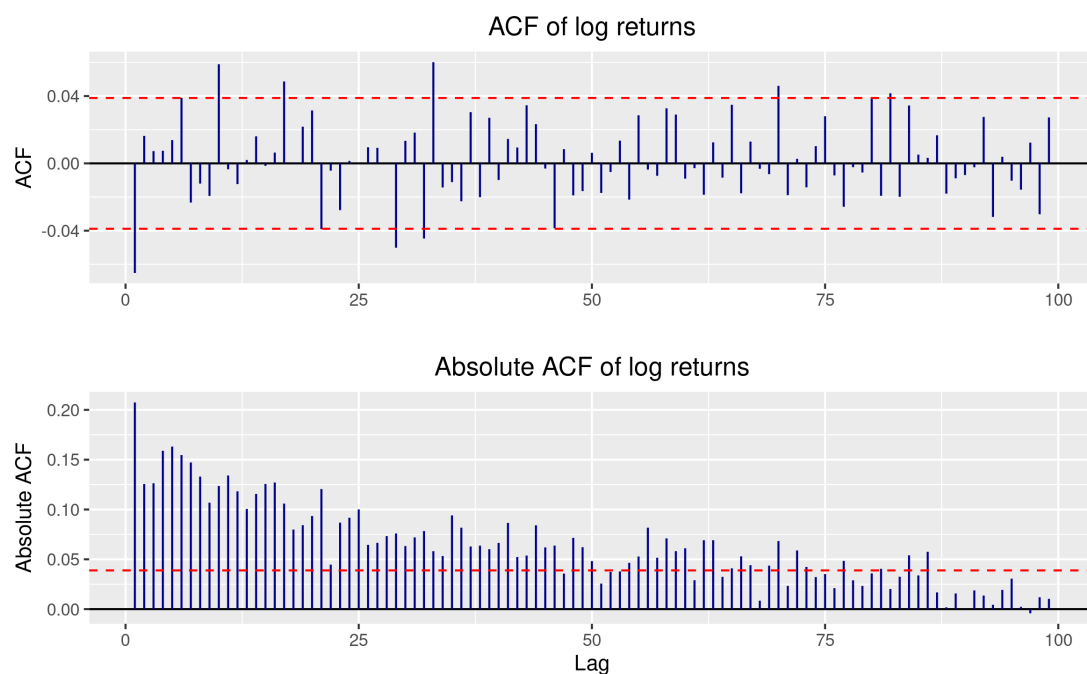
Notes: The upper plot shows the autocorrelation function (ACF) of the raw percentage log-returns up to lag 100. The lower plot presents the ACF of the absolute percentage log-returns up to lag 100. The horizontal dashed blue lines represent the 95% (individual) significance bounds given by $\pm\frac{1.96}{\sqrt{T}}$. T is the total number of returns. Note that we do not account for multiple testing or conditional heteroscedasticity. In general, both would lead to larger significance bounds

Table 6: Parameter estimates, Bayesian Estimation

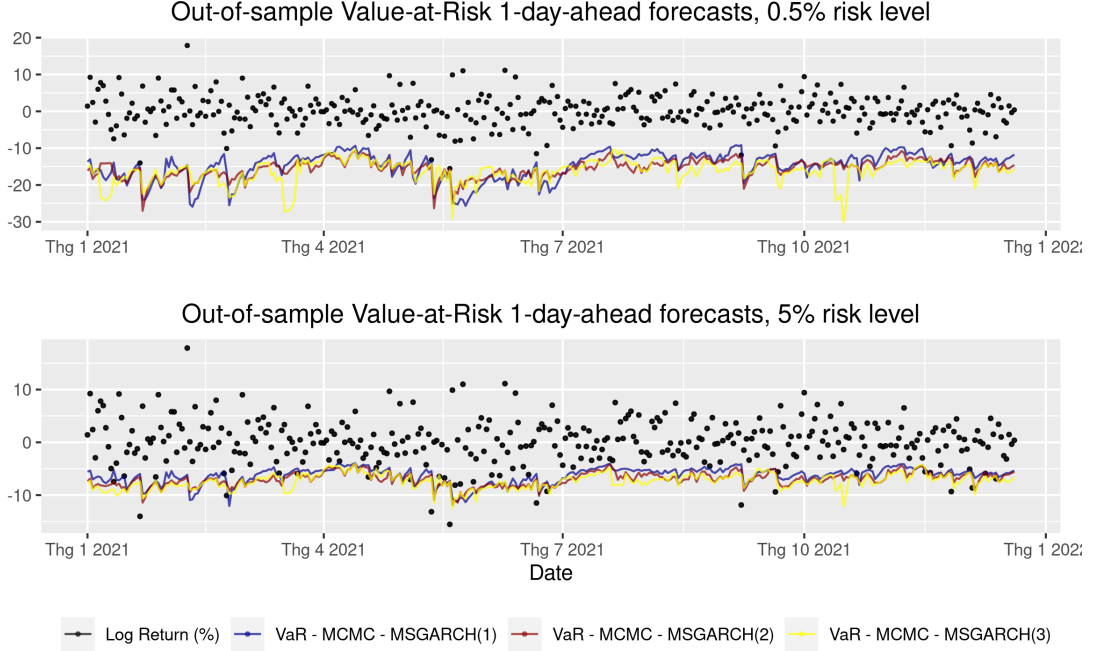|	Single-regime			Two-regime			Three-regime		
	Mean	*Q1*	*Q3*	*Mean*	*Q1*	*Q3*	*Mean*	*Q1*	*Q3*
				Regime $k = 1$					
ω_1	0.01	0.01	0.02	5.92	0.15	12.18	0.07	0.05	0.08
α_1	0.07	0.06	0.07	0.08	0.07	0.10	0.01	0.01	0.01
γ_1	0.01	0.01	0.01	0.10	0.00	0.19	0.00	0.00	0.00
β_1	0.93	0.92	0.93	0.67	0.40	0.91	0.96	0.95	0.96
ν_1	3.84	3.72	3.95	4.03	2.78	5.08	2.87	2.53	3.04
δ_1	0.95	0.94	0.97	0.91	0.89	0.96	0.99	0.95	1.02
$p_1 1$	-	-	-	0.95	0.96	0.98	0.95	0.95	0.96
aUV$_1$	43.38	-	-	89.64	-	-	26.98	-	-
				Regime $k = 2$					
ω_1	-	-	-	6.01	0.26	11.06	1.92	0.08	0.13
α_1	-	-	-	0.09	0.07	0.11	0.05	0.03	0.04
γ_1	-	-	-	0.12	0.00	0.22	0.01	0.00	0.00
β_1	-	-	-	0.68	0.47	0.90	0.89	0.95	0.96
ν_1	-	-	-	4.17	2.64	5.28	4.27	3.51	4.19
δ_1	-	-	-	0.90	0.87	0.95	0.93	0.91	0.96
$p_2 2$	-	-	-	0.92	0.96	0.98	0.97	0.97	0.98
$textaUV_2$	-	-	-	113.80	-	-	71.18	-	-
				Regime $k = 3$					
ω_1	-	-	-	-	-	-	27.16	23.44	32.73
α_1	-	-	-	-	-	-	0.17	0.10	0.22
γ_1	-	-	-	-	-	-	0.16	0.11	0.20
β_1	-	-	-	-	-	-	0.11	0.02	0.08
ν_1	-	-	-	-	-	-	8.39	5.74	10.66
δ_1	-	-	-	-	-	-	0.75	0.68	0.81
$p_3 3$	-	-	-	-	-	-	0.91	0.90	0.93
aUV$_3$	-	-	-	-	-	-	123.70	-	-

Notes: The table presents the Bayesian Estimation parameter estimates of the K-regime GJR-GARCH-sk$\mathcal{S}$ models, $K = 1, 2, 3$. GJR-GARCH is an extension of GARCH model proposed in Glosten et al. (1993). sk$\mathcal{S}$ is the standardized skewed Student's t-distribution with tail parameter ν and asymmetry parameter δ. p_{kk} is the staying probability of state k. aUV is the annualized unconditional volatility (in %). Here, 1st and 3rd Quartiles are showed as uncertainty estimates instead of traditional standard deviation and t-statistics.

Figure 6: Smoothed probabilities and annualized volatilities, Bayesian Estimation
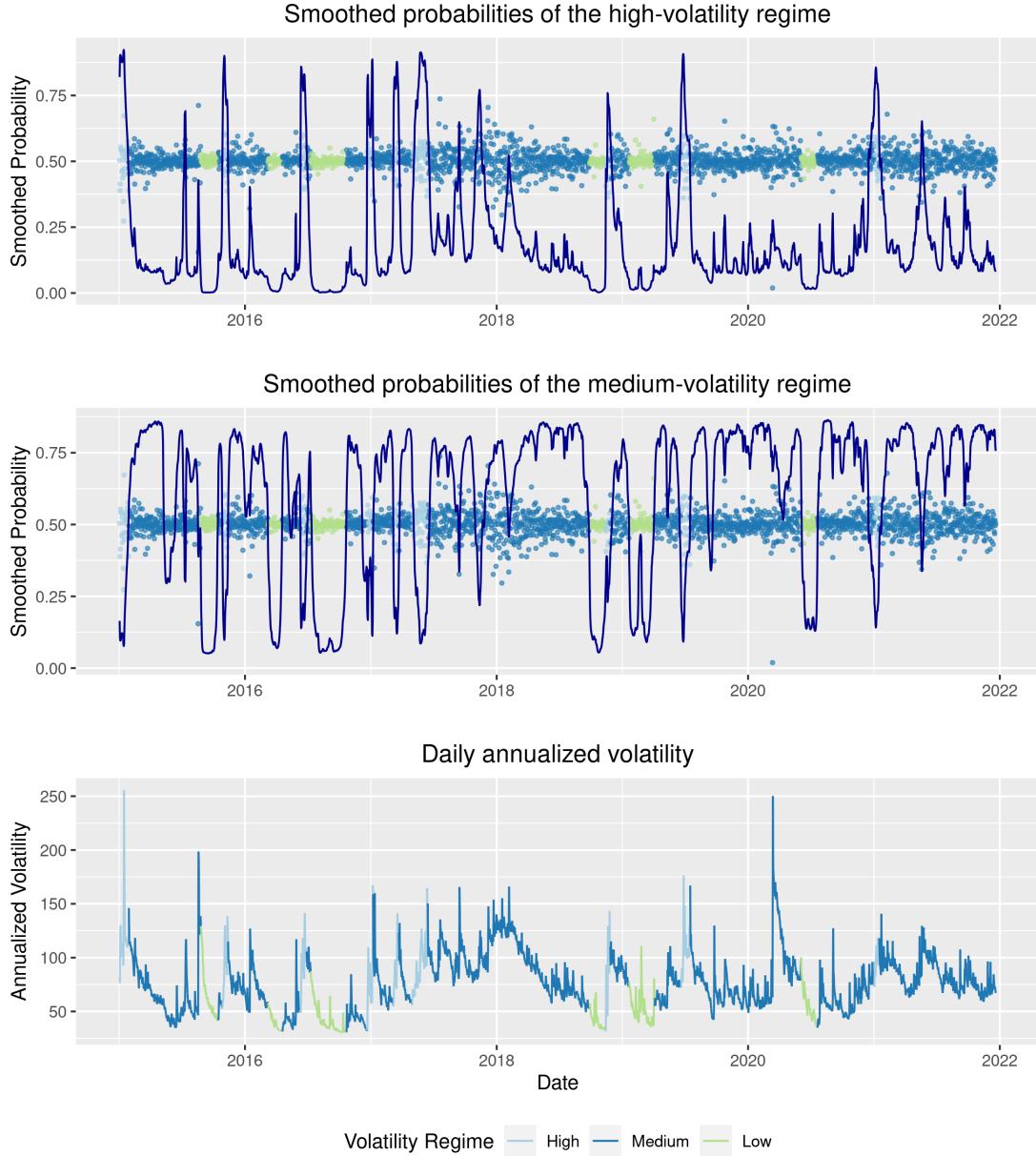
Notes: The time series range from January 01, 2015 to Dec 20, 2021. The graphs are computed from the three-regime SJRGARCH model with a standardized skewed Student's t-distribution. Parameters in the models are estimated using Bayesian Estimation. The estimated smoothed probabilities of the high- and medium-volatility regimes (3rd and 2nd regimes) are shown in the upper and middle graphs, respectively, together with the BTC-USD log-returns (points). The bottom graph shows the annualized volatilities in percentage. The colors of each point and color segments of the annualized volatility series indicate the regime to which the points or segments are predicted to belong.

Figure 7: Value-at-risk forecasts, Bayesian Estimation

Out-of-sample Value-at-Risk 1-day-ahead forecasts, 0.5% risk level

Out-of-sample Value-at-Risk 1-day-ahead forecasts, 5% risk level

— Log Return (%) — VaR - MCMC - MSGARCH(1) — VaR - MCMC - MSGARCH(2) — VaR - MCMC - MSGARCH(3)

Notes: The time series range from January 01, 2015 to Dec 20, 2021. The out-of-sample period ranges from Jan 01, 2021 to Dec 20, 2021. The top graph shows the one-day-ahead Value-at-risk (VaR) forecasts at the 0.5% risk level computed with the single regime (blue line), two-regime (red line) and three-regime (green line) GJR-GARCH-sk$\mathcal{S}$ models. The percentage log-returns are depicted as the dots. GJR-GARCH is an extenstion of GARCH model of Glosten et al. (1993) and sk$\mathcal{S}$ is the standardized skewed Student's t-distribution. The bottom graph depicts the one-day-ahead VaR forecasts at the 5% risk level.

Table 7: Backtesting results for 1-step-ahead volatility forecasts, Bayesian Estimation

	VaR risk level	0.005	0.01	0.05	0.1
MCMC-MSGARCH(1)	LR_{UC}	0.5272	0.2319	0.2157	**0.0142**
	LR_{CC}	0.8165	0.4412	0.4180	**0.0495**
MCMC-MSGARCH(2)	LR_{UC}	0.8652	0.8098	0.8636	**0.0491**
	LR_{CC}	0.9745	0.9279	0.4168	0.1209
MCMC-MSGARCH(3)	LR_{UC}	**0.0596**	0.7671	0.4994	0.6486
	LR_{CC}	0.1696	0.9328	0.4090	0.7991

Notes: The table presents the backtesting results for the 1-step-ahead value-at-risk (VaR) at risk levels 0.5%, 1%, 5% and 10%. Shown are the p-values of the Likelihood Ratio test of Unconditional Coverage (LR_{UC}) and the Likelihood Ratio test of Conditional Coverage (LR_{CC}) for the K-regime GJR-GARCH (Glosten et al., 1993) models, $K = 1, 2, 3$, with a standardized skewed Student's t-distribution.

Table 8: Expected and observed number of violations, Bayesian Estimation

VaR risk level	0.005	0.01	0.05	0.1
Expected number of hits	5.31	10.62	53.1	106.2
Observed number of hits				
MCMC-MSGARCH(1)	1	6	23	50
MCMC-MSGARCH(2)	2	4	17	47
MCMC-MSGARCH(3)	0	3	15	38

Affirmation

I hereby declare that I have composed my Master's thesis "Volatility regime in cryptocurrency returns" independently using only those resources mentioned, and that I have as such identified all passages which I have taken from publications verbatim or in substance. I agree that the work will be reviewed using plagiarism testing software.

Neither this thesis, nor any extract of it, has been previously submitted to an examining authority, in this or a similar form

I have ensured that the written version of this thesis is identical to the version saved on the enclosed storage medium.

2022-01-23