Cryptocurrency Volatility Modeling with Leverage Effect and Heavy-Tailed Distributions

Min Seo Kim, Yigang Hou, Abdullah Naser

Spring 2025

Abstract

This paper explores volatility modeling in cryptocurrency markets, focusing on the leverage effect and the heavy-tailed nature of return distributions. Using daily Bitcoin data, we evaluate the performance of both classical and extended GARCH-type models—including EGARCH and GJR-GARCH—that account for asymmetric volatility responses. To better capture extreme return events, we incorporate heavy-tailed distributions such as the Student's t and Generalized Error Distribution (GED). We further investigate hybrid extensions, including CVX-GARCH, RT-GARCH, and RNN-GARCH, which integrate exogenous variables and nonlinear dynamics. Our results show that models capturing both leverage effects and fat tails significantly improve volatility forecasts, with real-time and neural network variants offering additional gains under certain conditions.

Keywords: Cryptocurrency, volatility modeling, GARCH models, leverage effect, heavy tails, Cryptocurrency Volatility Index (CVX), RNN

Contents

1 Introduction

Over the past decade, cryptocurrencies have emerged as a significant asset class, attracting widespread interest from both retail and institutional investors. Among them, Bitcoin remains the most prominent, often regarded as a benchmark for the broader crypto market. Despite its growing adoption, however, Bitcoin and other digital assets exhibit extreme price volatility, non-stationary behavior, and non-normal return distributions, all of which are characteristics that pose considerable challenges for risk forecasting, portfolio management, and derivative pricing.

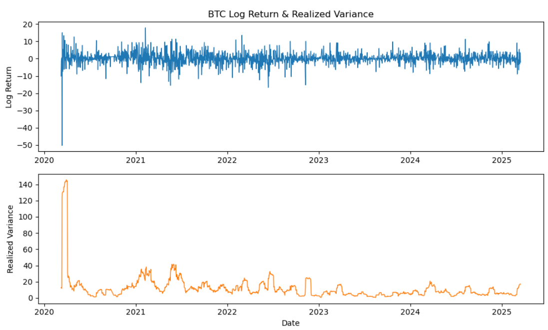


Figure 1: BTC Log Return and Realized Variance

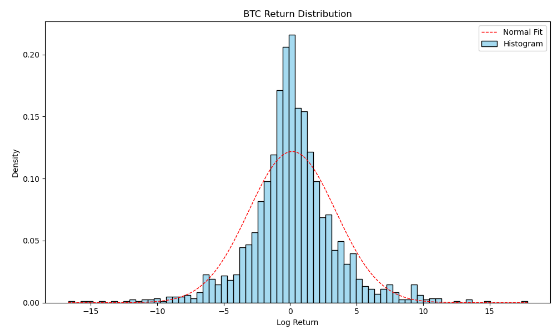


Figure 2: BTC Return Distribution

Volatility modeling plays a crucial role in financial risk management, particularly in highly speculative markets such as cryptocurrencies. As illustrated in Figure 1 and Figure 2, Bitcoin returns exhibit both sharp jumps in realized variance and a return distribution that deviates substantially from normality. The left plot highlights volatility clustering and extreme spikes, while the right plot reveals strong leptokurtosis and asymmetry in the return distribution. These empirical features underscore the need for more flexible volatility models that can account for nonlinearity, asymmetry, and heavy tails.

Traditional models like the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) framework are widely used due to their ability to capture volatility clustering and persistence. However, standard GARCH models assume symmetric response to return shocks and rely on the normal distribution, limiting their effectiveness in environments characterized by leverage effects and fat tails.

In this study, we explore enhanced volatility models that address these limitations. Specifically, we pursue three complementary approaches: (1) extending GARCH-type models such as EGARCH and GJR-GARCH to incorporate asymmetric responses to negative shocks; (2) replacing the normal distribution with heavy-tailed alternatives such as the Student's t and Generalized Error Distribution (GED) to better capture extreme events; and (3) developing hybrid extensions that integrate exogenous variables—including the Cryptocurrency Volatility Index (CVX) and real-time shock measures—and nonlinear mappings using neural networks to improve predictive accuracy.

By evaluating a range of volatility models on empirical data and comparing their in-sample fit and out-of-sample forecast performance, this paper aims to identify which combinations of model structure and distributional assumptions best reflect the unique characteristics of cryptocurrency markets.

2 Data Collection and Preprocessing

This study uses daily BTC/USD price data from Binance, spanning from March 2020 to March 2025. To prepare the data for volatility forecasting, we follow three preprocessing steps: (1) compute daily log returns and examine key summary statistics to transform price data into a stationary series suitable for modeling; (2) conduct the Augmented Dickey-Fuller (ADF) test to formally assess stationarity; and (3) fit an ARMA model to remove serial correlation, isolating the residuals for input into GARCH-type models.

2.1 Log Returns and Summary Statistics

Daily log returns are computed using the standard transformation:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{1}$$

This stabilizes the variance and captures essential properties for modeling financial time series. Summary statistics for the resulting return series are presented in Table 1.

Ticker	**Date Range**	**Max**	**Min**	**Mean**	**Std. Dev.**	**Skewness**	**Kurtosis**
BTC	3/8/20 – 3/16/25	17.84%	-50.26%	0.13%	3.48%	-1.66	26.39

Table 1: Summary Statistics for BTC/USD Daily Log Returns (3/8/2020 – 3/16/2025)

The return distribution is notably left-skewed, suggesting that large negative returns occur more frequently than positive ones. It also exhibits extreme kurtosis, indicating a high probability of observing large deviations from the mean—far more than would be expected under a normal distribution.

2.2 ADF Stationarity Test

To validate the suitability of GARCH-type models, we test the return series for stationarity using the Augmented Dickey-Fuller (ADF) test. The test yields a statistic of -14.31 with a p-value less than 0.01, allowing us to reject the null hypothesis of a unit root. This confirms that the return series is stationary.

2.3 ARMA Demeaning Procedure

Before fitting GARCH-type models, we first remove any linear dependence in the return series using an Autoregressive Moving Average (ARMA) model, which expresses the return series as a linear function of past values and past shocks:

$$r_t = \mu + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{2}$$

where μ is the mean, ϕ_i and θ_j are the respective AR and MA coefficients, and ϵ_t is white noise. This step ensures that the volatility models capture time-varying conditional variance rather than predictable mean dynamics.

Figure 3 presents the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the BTC log return series. These plots provide initial insight into the presence of serial correlation in the data.

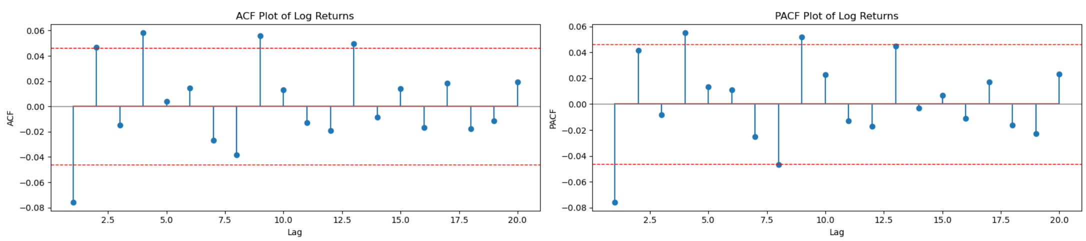


Figure 3: ACF and PACF Plots of BTC Log Returns

The final ARMA model order was determined through a grid search over a range of (p, q) combinations using the Akaike Information Criterion (AIC). The optimal model selected was ARMA(0,4), which effectively removes serial correlation and produces residuals suitable for use in GARCH-variant and hybrid volatility models.

3 Baseline and Asymmetric Volatility Models

We begin our analysis by applying three GARCH model variants to capture the volatility dynamics of Bitcoin returns, focusing on identifying the presence of leverage effects. These models include the standard GARCH(1,1) and its asymmetric extensions—EGARCH(1,1) and GJR-GARCH(1,1)—which allow for differential responses to positive and negative shocks. Each model progressively incorporates mechanisms to account for the empirical observation that negative returns tend to amplify volatility more than positive ones. To ensure robust evaluation, all models are estimated under the assumption of normally distributed residuals and assessed using an 80:20 train-test split.

3.1 GARCH(1,1)

The GARCH model, introduced by Bollerslev [1], captures the volatility clustering commonly observed in financial return series. In the GARCH(1,1) specification, the conditional variance at time t depends on two components: the squared return shock from the previous period (ϵ_{t-1}^2) and the past conditional variance (σ_{t-1}^2). The parameter ω represents the long-run average level of volatility, so a higher ω implies a higher baseline variance when recent shocks and past volatility have little influence. For the model to be stationary, it is required that $\alpha + \beta < 1$.

The model is formally defined as follows:

$$r_t = \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_t^2) \tag{3}$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{4}$$

Fitting the model under the assumption of normally distributed residuals yields parameter estimates of $\omega = 0.4097$, $\alpha = 0.0927$, and $\beta = 0.8812$, with an in-sample log-likelihood of -3813.59 and an AIC of 7633.17. The corresponding out-of-sample volatility forecast is shown in Figure 4.

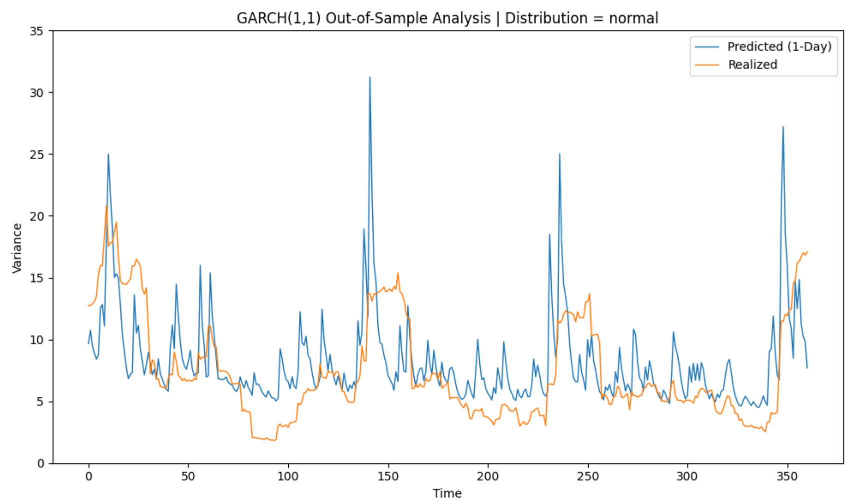


Figure 4: GARCH(1,1) Volatility Forecast (Normal Distribution)

3.2 EGARCH(1,1)

To account for asymmetric volatility responses, particularly the tendency of negative returns to induce larger increases in volatility than positive ones, Nelson [2] proposed the Exponential GARCH (EGARCH) model. Unlike standard GARCH, EGARCH models the logarithm of the conditional variance, which naturally enforces non-negativity and allows for more flexible dynamics. The specification introduces a leverage term γ, which directly captures the asymmetry in how shocks affect volatility.

The model is defined as:

$$r_t = \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_t^2) \tag{5}$$

$$z_t = \frac{\epsilon_t}{\sigma_t} \tag{6}$$

$$\log(\sigma_t^2) = \omega + \alpha(|z_{t-1}| - \mathbb{E}[|z_{t-1}|]) + \gamma z_{t-1} + \beta \log(\sigma_{t-1}^2) \tag{7}$$

Here, ω, α, and β play similar roles as in GARCH(1,1), while $\gamma < 0$ implies that negative shocks increase volatility more than positive shocks of the same magnitude, thereby capturing the so-called leverage effect. When fit to Bitcoin residuals under a normal distribution, the model yields $\omega = 0.1271$, $\alpha = 0.1673$, $\beta = 0.9548$, and $\gamma = -0.0496$, with an in-sample log-likelihood of -3805.39 and AIC of 7618.79. The resulting volatility forecast is shown in Figure 5.

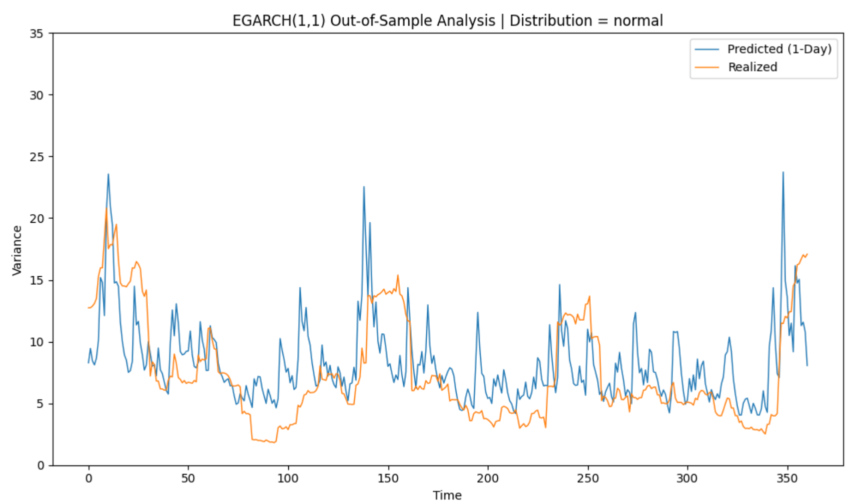


Figure 5: EGARCH(1,1) Volatility Forecast (Normal Distribution)

3.3 GJR-GARCH(1,1)

To better capture asymmetries in volatility reactions to positive and negative shocks, Glosten, Jagannathan, and Runkle [3] proposed the GJR-GARCH model. It builds on the standard GARCH framework by incorporating an indicator function that magnifies the effect of negative return shocks, enabling the model to reflect the empirical regularity that “bad news” increases volatility more than “good news”.

Formally, the model is given by:

$$r_t = \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_t^2) \tag{8}$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \gamma\epsilon_{t-1}^2 \cdot \mathbf{1}_{\{\epsilon_{t-1}<0\}} + \beta\sigma_{t-1}^2 \tag{9}$$

Rather than transforming the variance like EGARCH, GJR-GARCH retains the original GARCH structure and adds a threshold effect via an asymmetric term. Specifically, the conditional variance equation includes a component that is only activated when the previous shock is negative, allowing the model to adapt the size of volatility updates based on the direction of returns.

Again, ω, α, and β follow the same interpretation as in GARCH(1,1). The key parameter γ governs the size of this asymmetric adjustment: when $\gamma > 0$, the model responds more strongly to negative shocks, consistent with the leverage effect. In our estimation under the normal distribution, we obtain $\omega = 0.5485$, $\alpha = 0.0657$, $\beta = 0.8604$, and $\gamma = 0.0813$, with an in-sample log-likelihood of -3807.51 and an AIC of 7623.01. The out-of-sample volatility forecast is shown in Figure 6.

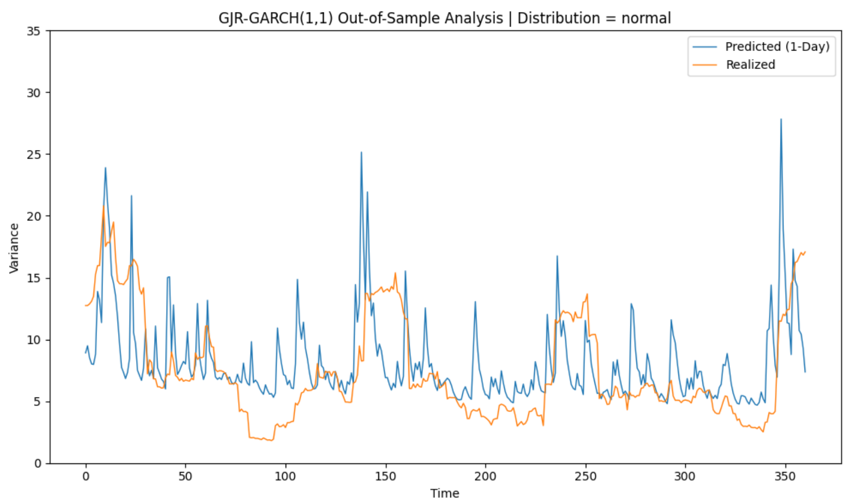


Figure 6: GJR-GARCH(1,1) Volatility Forecast (Normal Distribution)

3.4 Model Comparison under Normal Distribution

Table 2 summarizes the performance of the GARCH, EGARCH, and GJR-GARCH models under the assumption of normally distributed residuals. We evaluate each model using log-likelihood (LLK), Akaike Information Criterion (AIC), mean squared error (MSE), mean absolute error (MAE), and heteroskedasticity-adjusted MSE (HMSE).

Model	Distribution	LLK	AIC	MSE	MAE	HMSE
GARCH	Normal	−3813.59	7633.17	13.7797	2.7479	0.4163
EGARCH	Normal	−3805.39	7618.79	12.4011	2.7190	0.4777
GJR-GARCH	Normal	−3807.51	7623.01	14.6338	2.8247	0.5259

Table 2: Model Comparison under Normal Distribution

As shown in the table, models that account for the leverage effect—EGARCH and GJR-GARCH—offer improved in-sample fit compared to the symmetric GARCH model. EGARCH achieves the highest log-likelihood and lowest AIC, indicating the best overall in-sample fit. It also exhibits the lowest MSE and MAE, suggesting better predictive performance. Meanwhile, GJR-GARCH's relatively higher forecast errors suggest that the EGARCH model provides a more flexible and effective framework for modeling Bitcoin volatility under normal distribution assumptions.

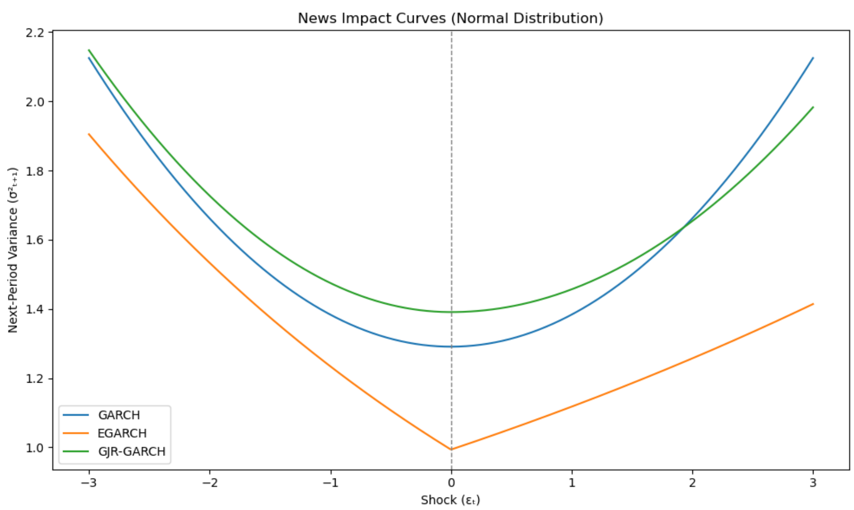


Figure 7: Comparison of News Impact Curves (Normal Distribution)

Furthermore, Figure 7 displays the news impact curves for the GARCH, EGARCH, and GJR-GARCH models under the assumption of normally distributed residuals. These curves show how each model adjusts next-period variance in response to shocks of different magnitudes and signs. The GARCH model produces a symmetric U-shaped curve, indicating equal sensitivity to positive and negative shocks. In contrast, EGARCH and GJR-GARCH capture asymmetry. The EGARCH curve is notably steeper for negative shocks, reflecting the leverage effect—greater volatility following bad news. The GJR-GARCH curve also shows asymmetry, though more moderately, due to the threshold effect from the indicator function.

4 Modeling Heavy Tails

As previously discussed, a key limitation of using the normal distribution in financial modeling, especially in cryptocurrency markets, is its inability to account for fat tails and extreme return values. To address this, we incorporate two alternative distributions: the Student's t distribution and the Generalized Error Distribution (GED). Our objective is to assess whether these distributions improve model robustness and forecasting accuracy by better capturing heavy-tailed behavior.

4.1 Student's t and GED Distributions

Table 3 summarizes key properties of the Student's t and GED distributions, both of which are widely used to account for heavy tails in financial return data. These heavy-tailed alternatives serve as substitutes for the normal distribution within GARCH-type volatility models. In the following subsections, we evaluate their effect on GARCH, EGARCH, and GJR-GARCH models.

Property	**Student's t Distribution**	**Generalized Error Distribution (GED)**
PDF	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	$\frac{\beta}{2\alpha\Gamma(1/\beta)}\exp\left(-\left\lvert\frac{x-\mu}{\alpha}\right\rvert^{\beta}\right)$
Tail Behavior	Polynomial decay	Exponential decay
Control Parameter	Degrees of freedom (ν)	Shape parameter (β)
Symmetry	Cannot model skewness	Supports skewed variants
Special Case	Normal as $\nu \to \infty$	Normal if $\beta = 2$

Table 3: Comparison of Heavy-Tailed Distributions

In our empirical analysis, we implement the GED with two different shape parameters: $\beta = 1$, which corresponds to the Laplace distribution, and $\beta = 1.5$, which offers an intermediate tail thickness between Laplace and Gaussian. These settings allow us to test the sensitivity of GARCH-type models to varying degrees of tail behavior.

4.1.1 Log-Likelihood Functions

To estimate model parameters under each distributional assumption, we derive the corresponding log-likelihood functions. These are computed over a sample of T observations, omitting constant terms where appropriate.

Normal Distribution. Assuming normally distributed residuals with variance σ_t^2, the log-likelihood is derived from the standard Gaussian density:

$$L_T = \sum_{t=1}^{T} l_t(\sigma_t, \epsilon_t) \tag{10}$$

$$l_t(\sigma_t, \epsilon_t) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{\epsilon_t^2}{2\sigma_t^2} \tag{11}$$

Student's t Distribution. Introducing degrees of freedom ν, the Student's t distribution generalizes the normal distribution to allow for heavier tails, and the corresponding log-likelihood function is given by:

$$L_T = \sum_{t=1}^{T} l_t(\sigma_t, \epsilon_t, \nu) \tag{12}$$

$$l_t(\sigma_t, \epsilon_t, \nu) = \log\left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi}}\right) - \frac{1}{2}\log(\sigma_t^2) - \frac{\nu+1}{2}\log\left(1 + \frac{\epsilon_t^2}{(\nu-2)\sigma_t^2}\right) \tag{13}$$

Generalized Error Distribution (GED). Assuming residuals follow a GED with shape parameter β, the log-likelihood incorporates a scale parameter α_t that depends on the conditional variance:

$$L_T = \sum_{t=1}^{T} l_t(\sigma_t, \epsilon_t, \beta) \tag{14}$$

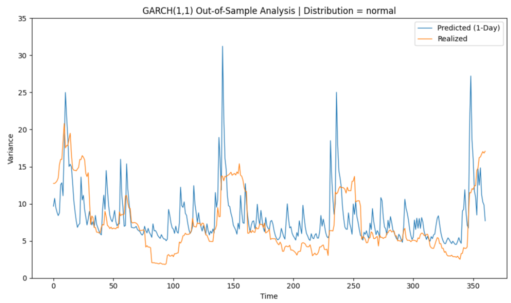
$$\alpha_t = \sqrt{\frac{\Gamma(1/\beta)}{\Gamma(3/\beta)}} \cdot \sigma_t \tag{15}$$

$$l_t(\sigma_t, \epsilon_t, \beta) = \log\left(\frac{\beta}{2\alpha_t\Gamma(1/\beta)}\right) - \left|\frac{\epsilon_t}{\alpha_t}\right|^{\beta} \tag{16}$$

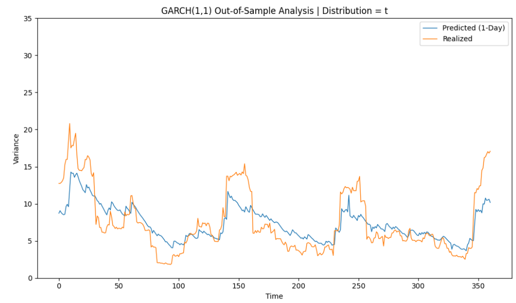
4.2 Distributional Impact on Volatility Forecasts

We first evaluate the GARCH(1,1) model under four distributional assumptions: Normal, Student's t, Laplace, and GED. Each distribution shapes the model's response to return shocks in distinct ways:

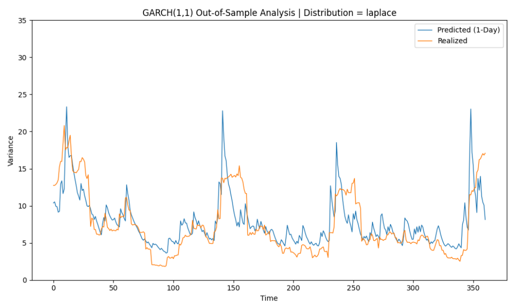
- The **normal** distribution often exaggerates volatility during stable periods.
- The **Student's t** distribution smooths forecasts by accommodating large shocks.
- The **Laplace** and **GED** distributions balance tail sensitivity and stability.



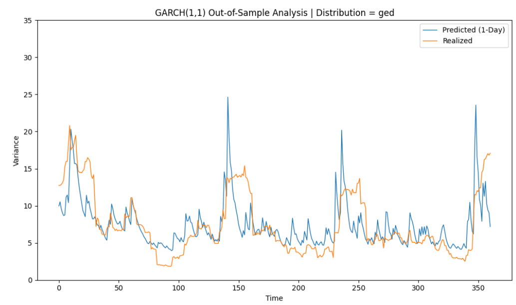
(a) Normal Distribution



(b) Student's t Distribution



(c) Laplace Distribution



(d) Generalized Error Distribution

Figure 8: Impact of Distributional Assumptions on Volatility Forecasts

Figure 8 illustrates how each distribution affects the shape and magnitude of the forecasted volatility path in the GARCH(1,1) model. Under the normal distribution, the model tends to overestimate volatility during relatively calm periods, resulting in sharp, exaggerated spikes that often diverge from realized values. In contrast, the Student's t distribution, with its heaviest tails, smooths the forecast and better accommodates large return shocks, leading to more stable and realistic predictions. The Laplace distribution, with its sharper peak and moderately heavy tails, produces forecasts that strike a balance between responsiveness and robustness, often closely aligning with realized variance. The GED distribution, when parameterized with $\beta = 1.5$, behaves similarly to the Laplace distribution but produces slightly sharper volatility spikes, offering a moderately heavier tail while preserving responsiveness to large shocks.

Overall, the use of heavy-tailed distributions clearly improves the GARCH model's ability to track realized variance, underscoring the importance of distributional choice in modeling highly volatile assets like cryptocurrencies.

4.3 Model Comparison under Heavy-Tailed Distributions

Table 4 highlights the impact of distributional assumptions on model performance across all three GARCH-type specifications. Across all metrics—log-likelihood, AIC, and forecast accuracy (MSE, MAE, HMSE)—models incorporating heavy-tailed distributions consistently outperform their normally distributed counterparts.

Model	**Distribution**	**LLK**	**AIC**	**MSE**	**MAE**	**HMSE**
GARCH	Normal	-3813.59	7633.17	13.7797	2.7479	0.4163
GARCH	Student's t	-3616.91	7241.82	6.3158	1.9791	0.1836
GARCH	Laplace	-3616.22	7238.45	6.8451	1.9401	0.1755
GARCH	GED	-3686.03	7378.07	9.1096	2.1628	0.2076
EGARCH	Normal	-3805.39	7618.79	12.4011	2.7190	0.4777
EGARCH	Student's t	-3605.48	7220.95	5.5644	2.1102	0.2387
EGARCH	Laplace	-3609.33	7226.65	9.1527	2.3334	0.3000
EGARCH	GED	-3678.16	7364.32	10.5427	2.4032	0.3292
GJR-GARCH	Normal	-3807.51	7623.01	14.6338	2.8247	0.5259
GJR-GARCH	Student's t	-3614.95	7239.89	5.2241	1.7337	0.1472
GJR-GARCH	Laplace	-3615.87	7239.74	8.3310	2.0970	0.2353
GJR-GARCH	GED	-3685.91	7379.83	10.3448	2.3174	0.2740

Table 4: Model Comparison under Heavy-Tailed Distributions

In particular, the EGARCH model with Student's t distribution achieves the best in-sample fit, indicated by the highest LLK and lowest AIC values. However, in terms of out-of-sample forecasting, the GJR-GARCH model with Student's t distribution performs best, yielding the lowest MSE, MAE, and HMSE among all configurations. This result underscores the advantage of combining asymmetric volatility structures with fat-tailed error distributions for capturing the extreme and skewed nature of cryptocurrency return dynamics. Notably, the Laplace and GED distributions also offer improvements over the normal case, though to a slightly lesser extent than Student's t.

Lastly, note that the GARCH models with Student's t and Laplace distributions produce nearly identical log-likelihood values, yet their predicted volatility paths diverge meaningfully in Figure 8. This discrepancy stems from structural differences in how each distribution handles extreme observations. The Student's t distribution features polynomially decaying tails, which allow it to accommodate large return shocks more gradually. In contrast, the Laplace with exponential tails assigns less probability mass to extreme events, leading to sharper, more localized adjustments in the estimated variance. Although both models fit the data similarly in likelihood terms, their implied variance dynamics differ, highlighting the importance of evaluating the temporal behavior of volatility forecasts alongside statistical fit when selecting among heavy-tailed innovations.

5 Hybrid Model Enhancements

This last section explores enhancements to traditional GARCH models by incorporating additional information and nonlinear dynamics. First, we extend the baseline model by introducing the CVX-GARCH, which includes an implied volatility factor derived from the Cryptocurrency Volatility Index (CVX) introduced by Woebbeking [4]. Second, we consider the Real-Time GARCH (RT-GARCH) model proposed by Smetanina [5], which updates variance estimates using both past and contemporaneous information. Finally, we implement a recurrent neural network (RNN) architecture to capture nonlinear relationships and sequential dependencies in the volatility structure.

5.1 CVX-GARCH(1, 1)

The motivation for incorporating implied volatility stems from the forward-looking nature of the Cryptocurrency Volatility Index (CVX). Similar in construction to the VIX, which is derived from S&P 500 options, the CVX captures the market's expectation of Bitcoin's future volatility using at-the-money option prices with 30 days to maturity. Given its strong correlation with realized volatility, we hypothesize that CVX contains predictive information useful for improving volatility forecasts.

We formalize the model as follows:

$$r_t = \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_t^2) \tag{17}$$

$$\sigma^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \psi \mathrm{CVX}_{t-1}^2 \tag{18}$$

Fitting the model under normal distribution yields parameter estimates of $\omega = 0.1952$, $\alpha = 0.0479$, $\beta = 0.8915$, and $\psi = 0.0441$, with an in-sample log-likelihood of -3810.16 and an AIC of 7628.32. The out-of-sample forecast, shown in Figure 9, exhibits frequent volatility spikes, likely driven by fluctuations in the CVX term. However, it remains unclear whether CVX contributes additional predictive power beyond the information already captured by lagged squared returns and conditional variance.

We experimented with alternative CVX-derived signals to mitigate overreaction and improve model performance, including:

- First Differences: $CVX_{t-1}^2 - CVX_{t-2}^2$
- Simple Moving Average Momentum: $SMA_{t-1} - SMA_{t-2}$, for $N = 1, \ldots, 10$
- Normalized Changes: $\frac{SMA_{t-1} - SMA_{t-j}}{SMA_{t-j}}$, for $j = 2, \ldots, 10$

Despite these efforts, empirical results remain inconclusive regarding the incremental value of CVX in improving volatility prediction beyond traditional GARCH components.

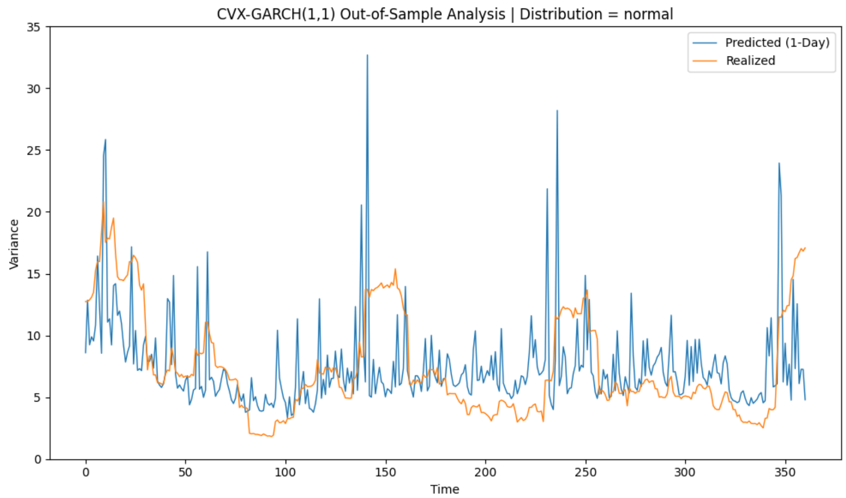


Figure 9: CVX-GARCH(1,1) Volatility Forecast (Normal Distribution)

5.2 RT-GARCH(1,1)

Unlike previous models that rely solely on lagged information, the RT-GARCH model incorporates current-period shocks to update volatility. By introducing the term z_t, which standardizes the return innovation using contemporaneous variance, the model captures instantaneous deviations in return behavior. This feature reflects the reality that financial markets react immediately to new information, allowing the conditional variance to spike on the same day rather than waiting until the next period. Such responsiveness is particularly useful during crises or jump events, when abrupt increases in risk must be quickly reflected in volatility forecasts.

The mode is therefore given by:

$$r_t = \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_t^2) \tag{19}$$

$$z_t = \frac{\epsilon_t}{\sigma_t} \tag{20}$$

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \psi z_t^2 \tag{21}$$

Fitted under a normal distribution, the model produces parameter estimates of $\omega = 0$, $\alpha = 0.0938$, $\beta = 0.8228$, and $\psi = 0.4481$, along with an in-sample log-likelihood of -3596.29 and an AIC of 7200.58. As shown in Figure 10, the model demonstrates a notable reduction in volatility overestimation despite assuming a normal distribution. The additional term z_t^2 dynamically adjusts variance based on current-period shocks, improving responsiveness to abrupt changes.

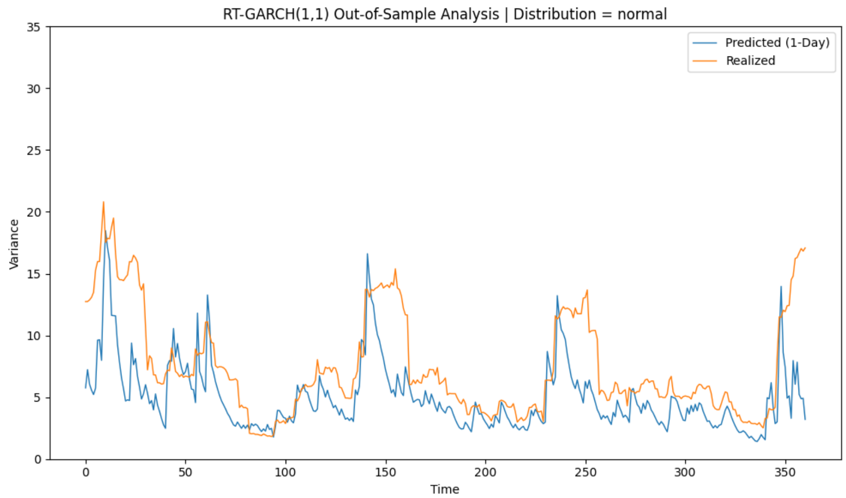


Figure 10: RT-GARCH(1,1) Volatility Forecast (Normal Distribution)

Note that since z_t depends on the very variance it helps update, forecasting σ_t^2 requires solving a quadratic equation, which yields the following formula:

$$\sigma_t^2 = \frac{b_{t-1} + \sqrt{b_{t-1}^2 + 4\phi\epsilon_t^2}}{2}, \quad \text{where } b_{t-1} = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{22}$$

5.3 RNN-GARCH(1, 1)

Finally, the RNN-GARCH model replaces the fixed parametric structure of GARCH with a data-driven, nonlinear mapping learned by a recurrent neural network. Instead of relying on predefined recursion formulas, the model learns to forecast conditional variance from sequences of past squared returns and past variances. This flexibility allows it to capture more complex temporal dependencies and nonlinear volatility dynamics that are common in financial time series.

To ensure positivity of the conditional variance, the RNN outputs are passed through a softplus activation function, defined as $f(x) = \log(1 + e^x)$. This guarantees strictly positive variance estimates during training and inference. Additionally, to maintain consistency with traditional GARCH training objectives, the model is trained by minimizing the negative log-likelihood rather than mean squared error (MSE), enabling fair comparisons across all volatility models. Figure 11 is a visualization of the resulting out-of-sample volatility forecast using the RNN-GARCH model.

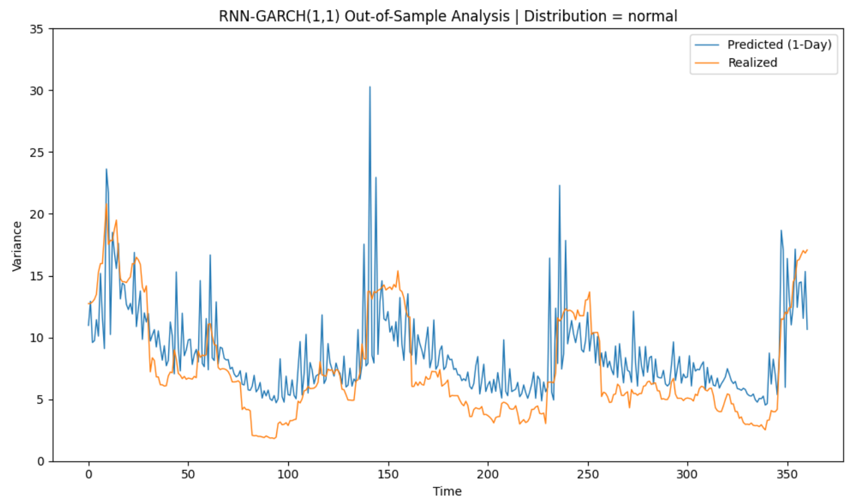


Figure 11: RNN-GARCH Volatility Forecast (Normal Likelihood)

5.4 Hybrid Model Comparison

Table 5 summarizes the performance of hybrid volatility models—CVX-GARCH, RT-GARCH, and RNN-GARCH—compared against standard GARCH under three distributional assumptions: Normal, Laplace, and GED.

Model	**Distribution**	**LLK**	**AIC**	**MSE**	**MAE**	**HMSE**
GARCH	Normal	-3813.59	7633.17	13.7797	2.7479	0.4163
CVX-GARCH	Normal	-3810.16	7628.32	17.7839	3.0801	0.3885
RT-GARCH	Normal	-3596.29	7200.58	14.4856	2.7182	0.1347
RNN-GARCH	Normal	-3781.76	N/A	9.5409	2.4760	0.3517
GARCH	Laplace	-3616.22	7238.45	6.8451	1.9401	0.1755
CVX-GARCH	Laplace	-3605.86	7219.72	15.6248	2.9471	0.3223
RT-GARCH	Laplace	-3598.58	7205.17	11.2055	2.2933	0.1102
RNN-GARCH	Laplace	-3613.48	N/A	16.5850	2.8486	0.1660
GARCH	GED	-3686.03	7378.07	9.1096	2.1628	0.2076
CVX-GARCH	GED	-3675.68	7359.35	16.2042	2.9239	0.2806
RT-GARCH	GED	-3654.89	7317.79	11.9115	2.4874	0.1395
RNN-GARCH	GED	-3680.65	N/A	20.0456	3.1657	0.1486

Table 5: Hybrid Model Comparison

Across all metrics, the RT-GARCH model consistently achieves strong results. It records the best in-sample log-likelihood and AIC under all three distributions, indicating superior model fit. Additionally, RT-GARCH under Laplace produces the lowest HMSE (0.1102), capturing tail behavior effectively with minimal overreaction. Although it does not always attain the lowest MSE or MAE, its performance remains competitive in striking the best balance between in-sample fit and out-of-sample stability. Its ability to incorporate current-period shocks through the z_t term proves especially valuable in volatile cryptocurrency markets, where timely adjustments to variance are critical.

Next, the RNN-GARCH model performs well in certain settings—most notably achieving the lowest MSE (9.5409) and MAE (2.4760) under the normal distribution. However, its performance deteriorates under Laplace and GED, where it shows significantly higher error and variance, likely due to overfitting or sensitivity to data distribution shifts.

Finally, the CVX-GARCH model shows mixed results. While it improves over standard GARCH in terms of LLK and AIC, it fails to outperform RT-GARCH or RNN-GARCH in MSE or HMSE. These findings suggest that although CVX (a forward-looking implied volatility measure) may contain useful information, its contribution is not substantial when lagged returns and variances are already present in the model.

6 Discussion

This study provides a detailed investigation into how various volatility models perform when applied to the distinctive characteristics of cryptocurrency returns. By evaluating traditional GARCH models, asymmetric extensions, and hybrid approaches across different distributional assumptions, several important insights emerged.

First, models that incorporate asymmetric responses to shocks—including EGARCH and GJR-GARCH—consistently outperformed the baseline GARCH model. This aligns with the well-documented leverage effect, where markets tend to react more strongly to negative returns than to positive ones. Among these, EGARCH stood out by delivering a strong trade-off between in-sample fit and forecast accuracy, highlighting the value of explicitly modeling asymmetry.

Second, replacing the normal distribution with heavy-tailed alternatives like Student's t and GED significantly improved the models' ability to capture extreme return events. Although both distributions yielded comparable log-likelihood scores, their impact on volatility dynamics differed. Student's t generated smoother, more stable forecasts due to its heavier tails, while GED responded more abruptly to outliers.

Turning to hybrid approaches, the results were more mixed. Incorporating the Cryptocurrency Volatility Index (CVX) as an exogenous regressor in CVX-GARCH yielded modest improvements in some settings but often led to increased forecast variance, suggesting redundancy with existing GARCH terms. In contrast, the RT-GARCH model, which updates volatility using current-period shocks, showed robust performance under both normal and heavy-tailed distributions, thanks to its ability to capture abrupt volatility spikes in real time. Finally, the RNN-GARCH model leveraged deep learning to capture complex nonlinear patterns in the data. While it achieved strong in-sample fit, its out-of-sample performance deteriorated, indicating potential overfitting. This suggests that while machine learning models offer expressive power, they require careful regularization and validation to generalize effectively.

Overall, our findings reinforce that capturing asymmetry and heavy tails is critical when modeling cryptocurrency volatility. The best-performing models combined these features—particularly EGARCH and RT-GARCH under Student's t innovations. More broadly, the results emphasize the importance of aligning model flexibility with sound distributional assumptions and careful attention to generalization.

7 Future Work

While this study provides a comprehensive comparison of volatility models for cryptocurrency markets, several promising avenues remain for future exploration. One direction is to incorporate asymmetric behavior into hybrid models like CVX-GARCH, RT-GARCH, and RNN-GARCH. Introducing leverage effects in these frameworks could enhance their ability to distinguish between the market impact of positive and negative shocks, especially during stress periods when volatility rises disproportionately in response to downside movements.

Another extension involves exploring skewed heavy-tailed distributions, such as the skewed Student's t or skewed GED. These could improve tail modeling by accounting for the observed asymmetry in return distributions, where extreme negative returns are more frequent and severe than positive ones—a characteristic particularly pronounced in cryptocurrency markets.

Lastly, on the machine learning front, extending the RNN architecture to asymmetric GARCH variants like RNN-EGARCH or RNN-GJR-GARCH could combine the flexibility of neural networks with more realistic volatility dynamics. Additionally, incorporating auxiliary inputs like macroeconomic indicators, sentiment measures, or blockchain-based metrics (e.g., on-chain volume, wallet activity) could provide hybrid models with a more holistic view of the market, potentially enhancing their predictive accuracy.

Collectively, these directions represent a path toward more expressive and resilient volatility models tailored to the complexities of digital asset markets.

References

[1] Bollerslev, Tim (1986). *Generalized Autoregressive Conditional Heteroskedasticity.* Journal of Econometrics, Journal of Econometrics, vol. 31, no. 3, Apr. 1986, pp. 307–327.

[2] Nelson, Daniel (1991). *Conditional Heteroskedasticity in Asset Returns: A New Approach.* Econometrica, vol. 59, no. 2, Mar. 1991, pp. 347–370.

[3] Glosten, Lawrence, et al (1993). *On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks.* The Journal of Finance, vol. 48, no. 5.

[4] Woebbeking, Fabian (2020). *Cryptocurrency Volatility Markets.* Digital Finance.

[5] Smetanina, Ekaterina (2022). *Real-Time GARCH.* Journal of Financial Econometrics, vol. 15, no. 4, 2017, pp. 561–601.