Volatility Forecasting and Regime Detection in Financial Time Series: An Applied Statistical Approach

Suyash Sohani Sayantan Banerjee

August 8, 2025

Abstract

This project proposes a comprehensive analysis of volatility forecasting and regime detection in stock market indices using advanced statistical models. By applying Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family models and Markov Switching models to long-term financial time series, the study aims to capture key features of financial data, such as volatility clustering and structural shifts. Using daily stock index data from Yahoo Finance, the research contributes to the literature on financial econometrics and has direct applications in risk management and portfolio optimization.

1 Introduction

Financial time series exhibit complex dynamics, including volatility clustering, nonlinearity, and abrupt regime changes, especially during periods of economic stress. Traditional linear models are inadequate for capturing these features, necessitating the use of models that explicitly account for time-varying volatility and nonlinear transitions. The present project aims to study the volatility behavior and regime dynamics of major stock mar- ket indices using statistical models suited for such tasks. Specifically, it investigates the efficacy of GARCH-type models for volatility forecasting and Markov Switching models for detecting structural changes or latent regimes in returns. Forecasting financial volatility is a fundamental problem in finance, with applications in derivative pricing, risk assessment, and capital allocation. Likewise, identifying market regimes can provide valuable insights into asset pricing under uncertainty and inform trading and hedging strategies. The proposed project seeks to examine both aspects through rigorous modeling and empirical analysis of historical index data. All plots in the report use the S&P 500 index, which is representative of the modeling and analysis performed. Other plots are included in the supplementary notebook.

2 Objectives

The primary objectives of this study are threefold. First, it aims to evaluate the ability of GARCH and its extensions to forecast the volatility of equity indices. Second, it seeks to detect and interpret hidden market regimes using Markov Switching autoregressive models. Third, it intends to assess model performance during historical periods of financial turbulence, such as the Global Financial Crisis of 2008 and the COVID-19 pandemic.

3 Data Description

To begin the analysis, we collected historical stock price data for three major equity indices: the S&P 500 (USA), NASDAQ Composite (USA), and Nifty 50 (India). These indices represent a combination of developed and emerging markets, allowing for a diverse analysis of financial time series. We used the `yfinance` library in Python to download daily adjusted closing prices from Yahoo Finance, covering the period from January 1, 2000 to July 20, 2025. We plotted each index to visualize its historical trend and price evolution. This initial step provides a macro level view of market behavior and lays the foundation for further statistical modeling.

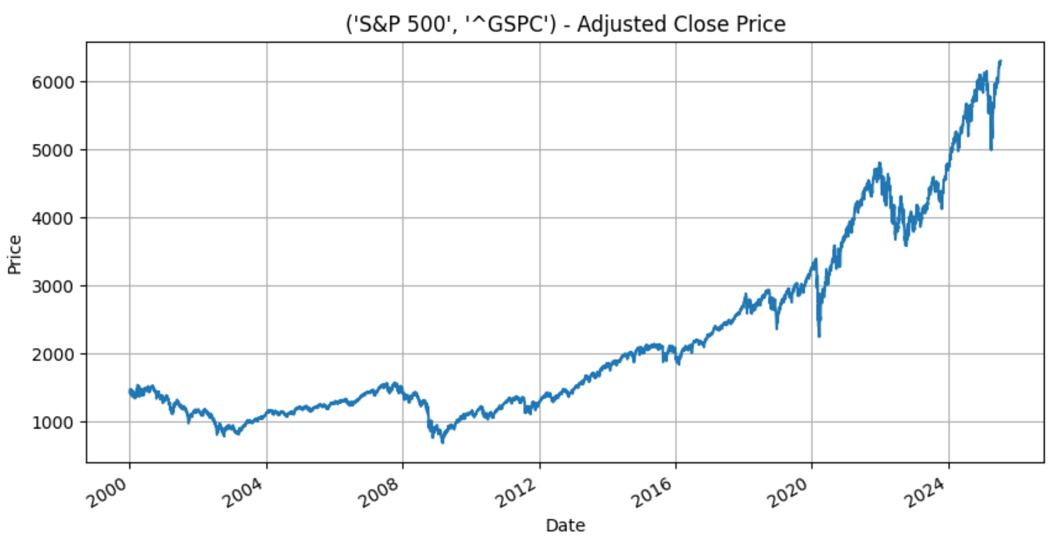

Figure 1: S&P 500 daily adjusted close price from January 2000 to July 2025.

4 Methodology

4.1 Preprocessing

To prepare the data for time series modeling, we converted raw price data into daily log returns. This step helps stabilize the variance and makes the time series more suitable for models like ARIMA and GARCH. We then visualized the log returns of each index to inspect general behavior and check for potential anomalies or patterns that may influence volatility modeling.

After computing the log returns, we conducted autocorrelation analysis using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The ACF plot shows how current returns are correlated with their past values over different lags, while the PACF helps identify the appropriate lag order for autoregressive models by isolating direct effects. This step is particularly useful for diagnosing whether ARIMA-type models are appropriate and determining initial model parameters.

To verify that the return series are suitable for ARIMA and GARCH style modeling, we tested for stationarity using two tests:

- **Augmented Dickey–Fuller (ADF) test**: Null hypothesis H_0: the series has a unit root (non-stationary).
- **Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test**: Null hypothesis H_0: the series is stationary.

Running these tests on the log-returns of all three indices (S&P 500, NASDAQ, Nifty 50) produced the following key results:

- **S&P 500**: ADF statistic ≈ -12.51, p-value $\approx 2.7 \times 10^{-23}$; KPSS statistic ≈ 0.401, p-value ≈ 0.076.
- **NASDAQ**: ADF statistic ≈ -12.42, p-value $\approx 4.2 \times 10^{-23}$; KPSS statistic ≈ 0.329, p-value ≈ 0.10.
- **Nifty 50**: ADF statistic ≈ -11.15, p-value $\approx 2.9 \times 10^{-20}$; KPSS statistic ≈ 0.102, p-value ≈ 0.10.

The ADF tests strongly reject the presence of a unit root for all three indices, while the KPSS tests fail to reject the null of stationarity. Together, these results confirm that the log-return series are stationary, validating the use of ARIMA and GARCH-family models.

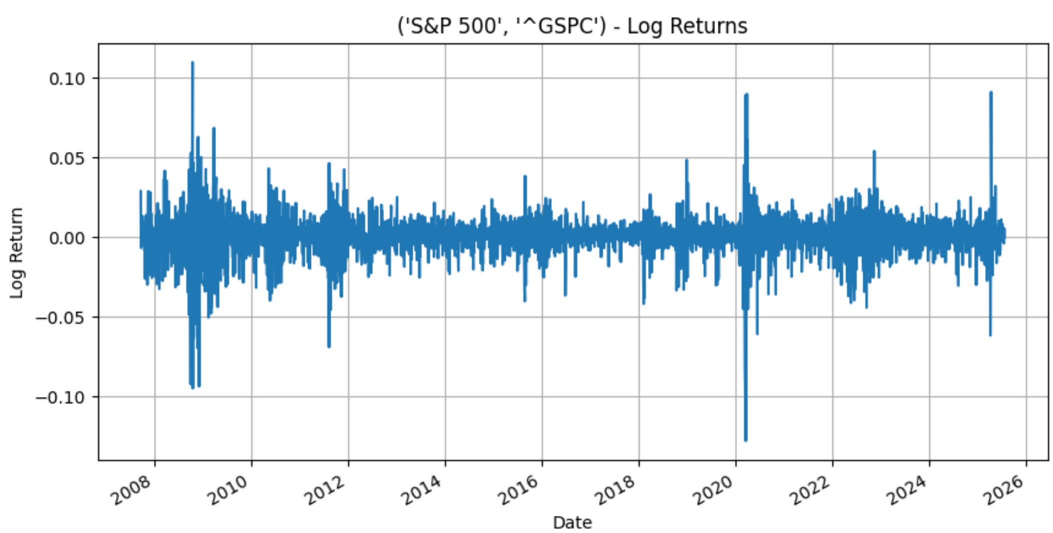

Figure 2: Daily log returns of S&P 500.

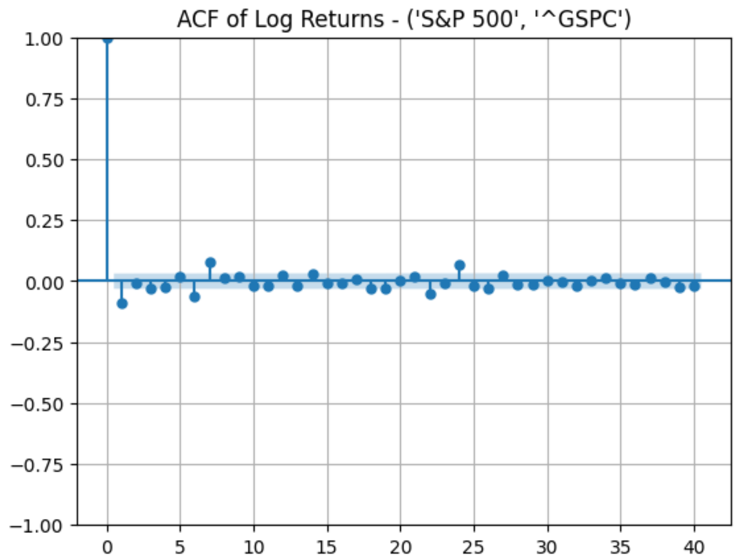

Figure 3: Autocorrelation function (ACF) of S&P 500 log returns.

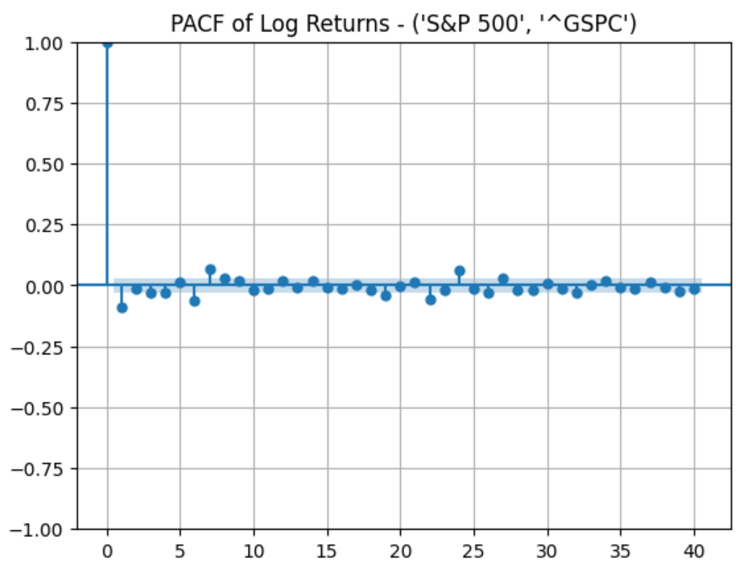

Figure 4: Partial autocorrelation function (PACF) of S&P 500 log returns.

4.2 Modeling Approach

To identify and quantify the underlying patterns and relationships in each return series, we fitted ARIMA models.

We wrote a function to search over a grid of candidate models with parameters (p, d, q) ranging from (0–2, 0–1, 0–2). The best model was selected for each index based on the Akaike Information Criterion (AIC). Once the best-fitting ARIMA model was identified, we plotted the actual versus fitted returns. This step helped isolate the predictable component of the series and allowed us to extract residuals which are then passed to volatility models like GARCH.

The ARIMA fits showed that most of the mean structure is close to white noise i.e. no significant autocorrelation. This further highlights the need to focus on volatility modeling.

To model the volatility of financial returns, we used the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) framework. We applied a GARCH(p,q) model to the residuals from the ARIMA model, where:

- p controls the autoregressive lag of volatility,
- q controls the moving average lag of past squared residuals.

To identify the best configuration, we wrote a function that fits multiple candidate GARCH models and selects the one with the lowest Akaike Information Criterion (AIC). The residuals were scaled to percentage units (multiplied by 100) to improve model stability. The resulting conditional volatility series for each index was plotted to observe how volatility evolved over time.

We then extended the baseline GARCH specification to the Exponential GARCH (EGARCH) model. We first filtered each index's log returns with its AIC-selected ARIMA mean model and used the standardized residuals (scaled by 100 for numerical stability) as inputs to EGARCH. We grid-searched over $(p, q) \in \{1, 2\} \times \{1, 2\}$ (with $o = 0$ since asymmetry is built into EGARCH) and selected the specification with the lowest AIC. For all three indices, the selected models were EGARCH(2,2), with AICs approximately: S&P 500: 46984.7; NASDAQ: 48519.0; Nifty 50: 47595.4. We then plotted the conditional volatility paths produced by the fitted EGARCH models. These series clearly display volatility clustering and asymmetric shock responses, reinforcing the suitability of EGARCH for equity index returns.

To explicitly test for leverage effects i.e. whether negative shocks raise volatility more than positive shocks of the same magnitude, we estimated GJR-GARCH models on the ARIMA filtered residuals (scaled by 100). Operationally, this is implemented by calling the same grid search function as for GARCH, but setting the asymmetry parameter $o = 1$, which activates the threshold (indicator) term. Model selection was again based on the Akaike Information Criterion (AIC). The best-fitting specifications and AICs were approximately:

- S&P 500: GJR-GARCH(1,1,1), AIC $\approx$ 46859.4
- NASDAQ: GJR-GARCH(2,1,1), AIC $\approx$ 48394.2
- Nifty 50: GJR-GARCH(1,1,1), AIC $\approx$ 47518.2

Across indices, the γ (asymmetry) coefficients were positive and statistically significant, confirming the presence of leverage effects, i.e. negative returns tend to amplify subsequent volatility more strongly than positive returns. We also plotted the conditional volatility series produced by each fitted GJR-GARCH model to visualize how these asymmetries manifest over time and to contrast them with the GARCH outputs.

To capture abrupt and nonlinear shifts in market behaviour that conventional linear or conditionally heteroskedastic models cannot fully explain, we estimated a Markov Switching Autoregression (MSAR) on the ARIMA filtered residuals (scaled by 100). Using the filtered series ensures that the detected regimes are primarily linked to volatility and shock dynamics rather than to predictable mean components. For each index, we first selected and fitted the best ARIMA model (by AIC) to remove the mean structure. We then estimated an MS-AR model with $k\_regimes = 2$, order $= 1$, constant mean in each regime, and regime-switching variances (`switching_variance=True`). We then plotted the smoothed marginal probabilities of being in the high-volatility regime over time. The smoothed probability plots identify extended clusters of high-volatility episodes that align with known market stress periods.

To evaluate whether the regime switching behavior of financial returns is detectable without first removing the conditional mean structure, we applied the Markov Switching Autoregression model directly to the raw log return series of each index.

This version of the model uses the same configuration as before. After estimating the model for each index, we extracted and plotted the smoothed marginal probabilities of being in the high volatility regime. These probabilities again aligned with known periods of financial stress, such as the 2008 global crisis and the COVID 19 shock.

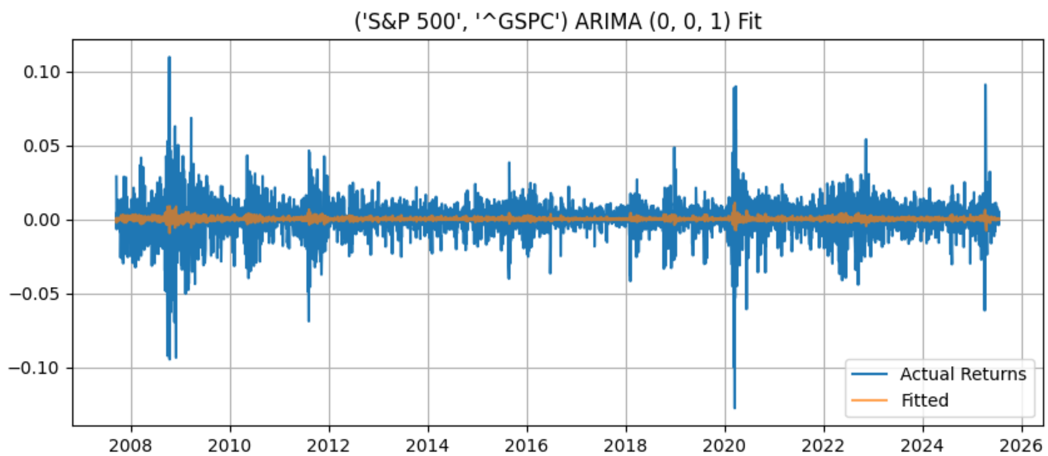

Figure 5: Fitted ARIMA(0,0,1) model for S&P 500 log returns.

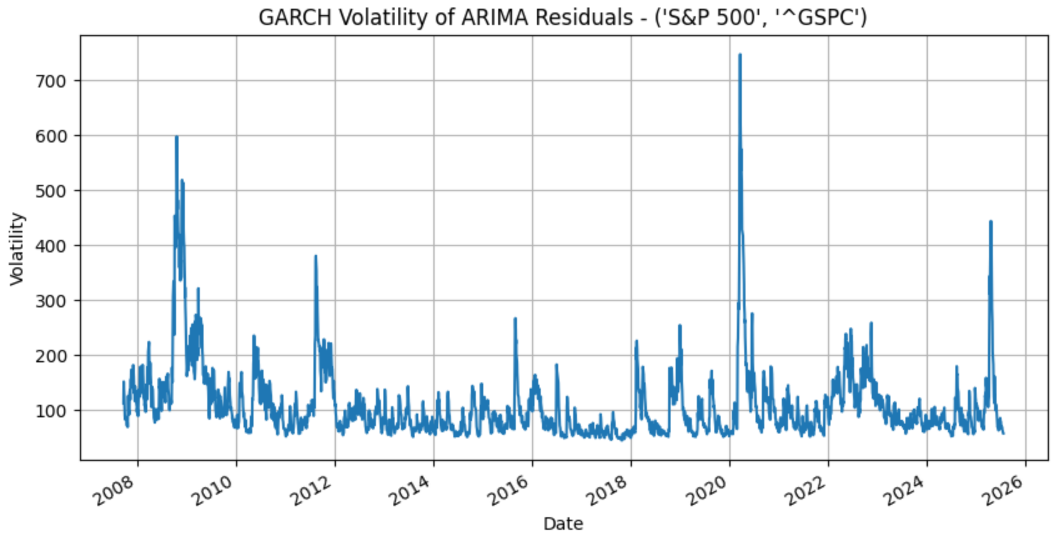

Figure 6: Conditional volatility from GARCH(1,1) model fitted to ARIMA residuals of S&P 500.

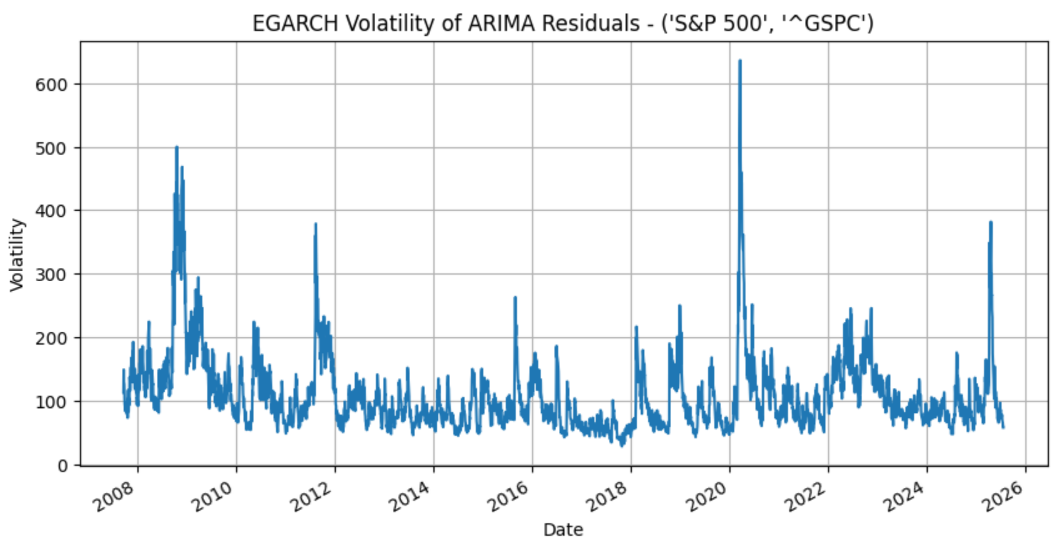

Figure 7: Volatility from EGARCH(2,2) model for S&P 500.

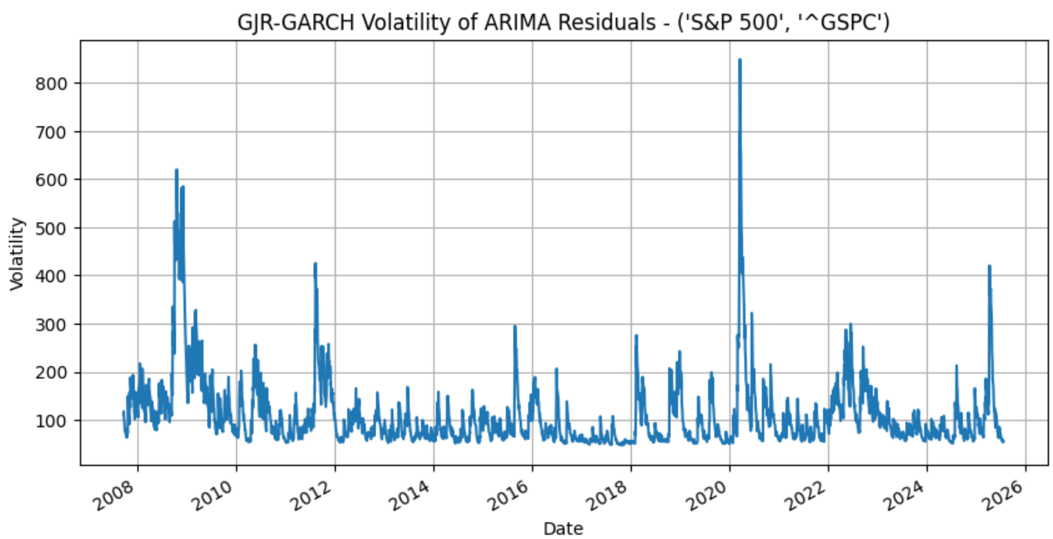

Figure 8: Conditional volatility from GJR-GARCH(1,1,1) model for S&P 500.

4.3 Evaluation

To assess the usefulness of the GARCH type models in predicting future volatility, we conducted a rolling forecast evaluation on each index using the GARCH, EGARCH, GJR-GARCH models. We implemented a rolling window approach as follows:

1. The time series of ARIMA filtered residuals was split into training (80%) and test (20%) sets.
2. For each model, we performed recursive re-estimation: at each time step in the test set, the model was re-fitted using all available data up to that point.
3. A one-step-ahead forecast of conditional volatility was produced at each iteration.

We evaluated forecast performance using three key metrics:

- Mean Squared Error (MSE)
- QLIKE loss
- Value-at-Risk (VaR) exceedance rate

Across all three indices, the rolling forecasts were visualized by comparing absolute residuals with model-predicted volatilities.

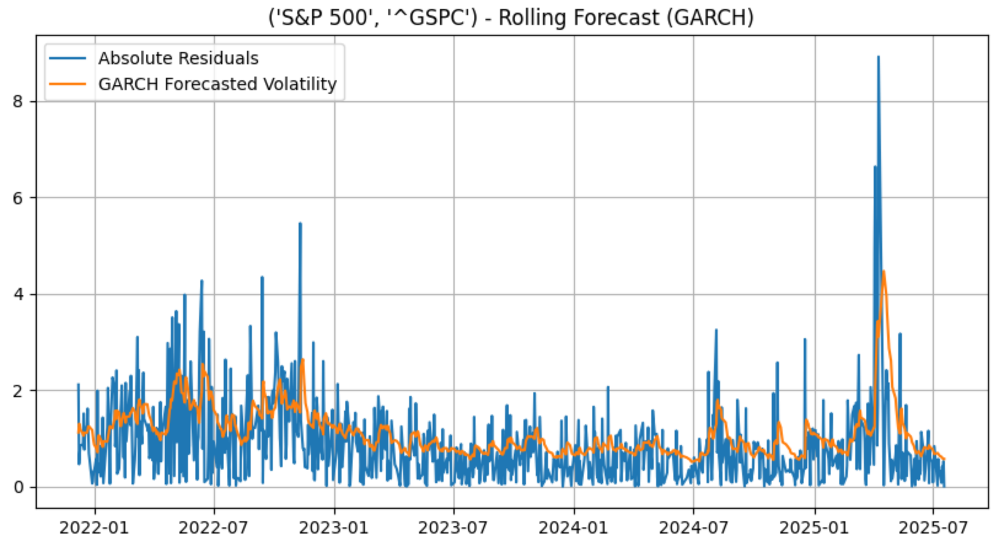

Figure 9: Rolling forecast using GARCH(1,1) for the S&P 500.

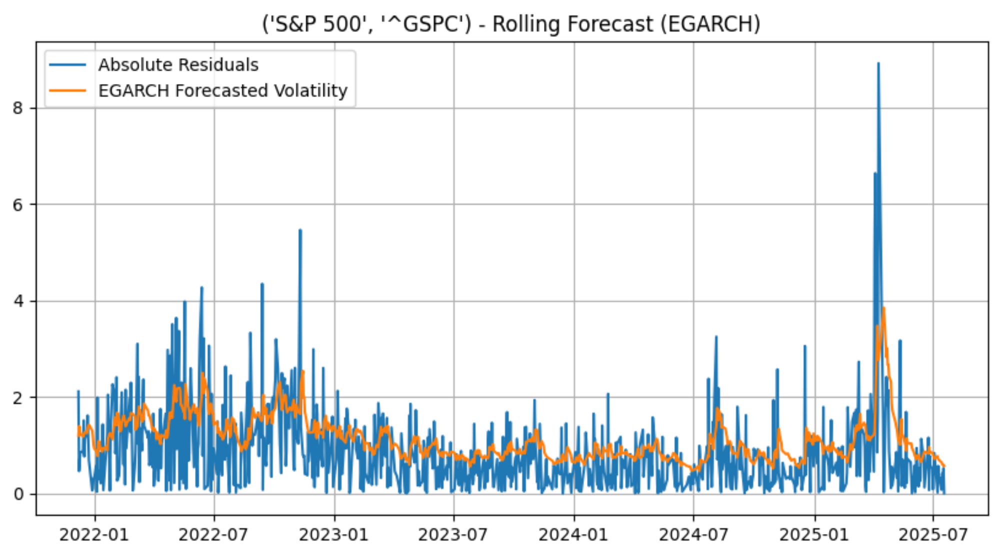

Figure 10: Rolling forecast using EGARCH(1,1) model for S&P 500.

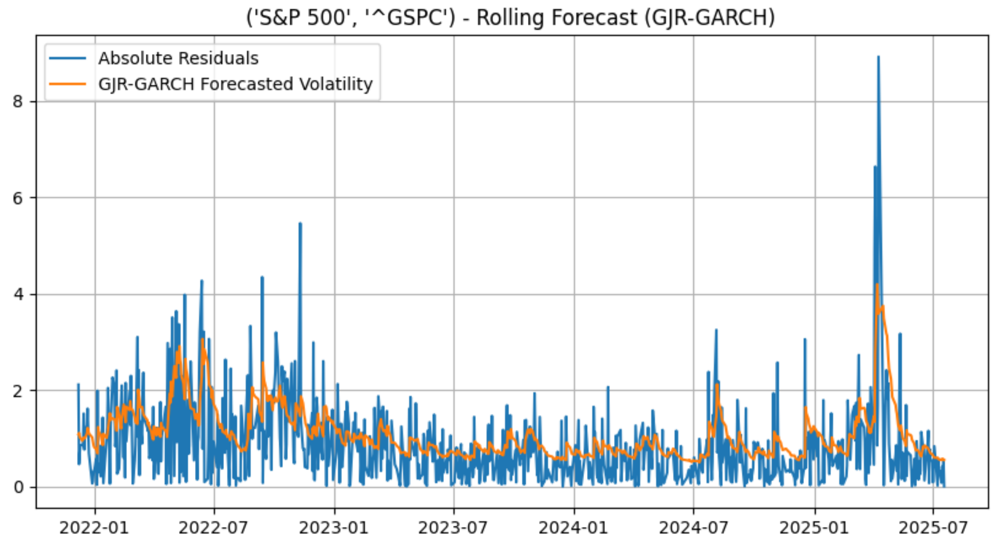

Figure 11: Rolling forecast using GARCH(1,1,1) model for S&P 500.

The table below summarizes the forecast accuracy results:

Index	Model	MSE	QLIKE	VaR Exceedance Rate
S&P 500	GARCH	0.723733	1.08927	0.0598726
S&P 500	EGARCH	0.706920	1.08994	0.0585987
S&P 500	GJR-GARCH	0.685837	1.05384	0.0649682
NASDAQ	GARCH	1.171280	1.70331	0.0636943
NASDAQ	EGARCH	1.135410	1.69820	0.0624204
NASDAQ	GJR-GARCH	1.100130	1.68137	0.0700637
Nifty 50	GARCH	0.450788	0.65578	0.0560510
Nifty 50	EGARCH	0.439608	0.66403	0.0535032
Nifty 50	GJR-GARCH	0.464039	0.69056	0.0522293

5 Results

To further validate the usefulness of GARCH models in financial stress scenarios, we conducted a focused analysis of volatility behavior during two major crisis periods:

- **Global Financial Crisis (2008–2009)**: January 1, 2008 to December 31, 2009
- **COVID-19 Crash (2020)**: January 1, 2020 to December 31, 2020

For each index (S&P 500, NASDAQ, Nifty 50), we refitted the best ARIMA model and used its residuals as input for GARCH, EGARCH, GJR-GARCH volatility models.

We then plotted the predicted conditional volatility paths produced by each model, restricted to each crisis window. This comparison highlights how responsive each model is to periods of market stress.

- The Global Financial Crisis plots showed sharp, sustained volatility spikes, particularly in the S&P 500 and NASDAQ indices. EGARCH and GJR-GARCH models responded faster and more sharply to large shocks compared to the symmetric GARCH model.
- During the COVID 19 period, all three models captured the sudden volatility burst in early 2020. GJR-GARCH showed the highest peak.

These visual comparisons confirm that GARCH family models are capable of adapting to major financial disruptions, with asymmetric variants (EGARCH and GJR-GARCH) generally providing faster and more reactive forecasts in periods of extreme risk.

To further investigate the alignment between statistical volatility modeling and regime detection, we created combined plots that overlay:

- Absolute ARIMA residuals
- Conditional volatilities from GARCH, EGARCH, and GJR-GARCH models
- Smoothed regime probabilities from a 2 state Markov Switching Autoregression (MS AR)

These plots were generated for the two periods of financial stress:

- Global Financial Crisis (2008–2009)
- COVID 19 Crash (2020)

The visualization uses a dual-axis layout:

- The left axis shows volatility estimates and absolute residuals.
- The right axis shows the probability of being in the high-volatility regime.

Findings:

- Regime probabilities surged during market downturns and aligned closely with GARCH-family volatility estimates and absolute residual spikes.
- EGARCH and GJR-GARCH models often captured the onset of turbulence earlier than symmetric GARCH.
- Markov Switching probabilities remained elevated during extended crisis phases, confirming regime persistence.

These results illustrate that volatility and regime-switching models are not only individually informative but also complementary in identifying and characterizing periods of financial stress.

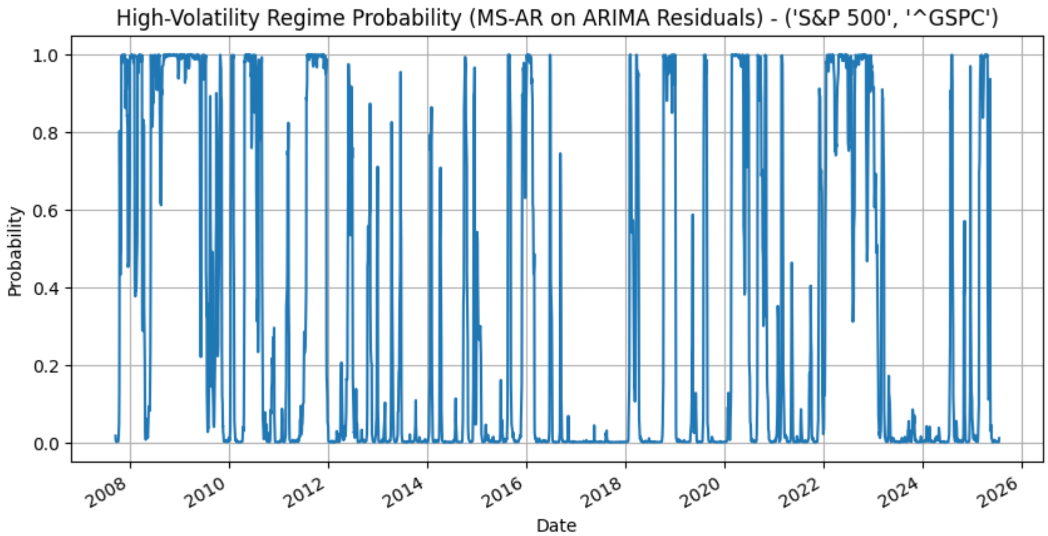

Figure 12: Smoothed probabilities of being in a high-volatility regime using a Markov Switching AR model fitted to ARIMA residuals of the S&P 500.

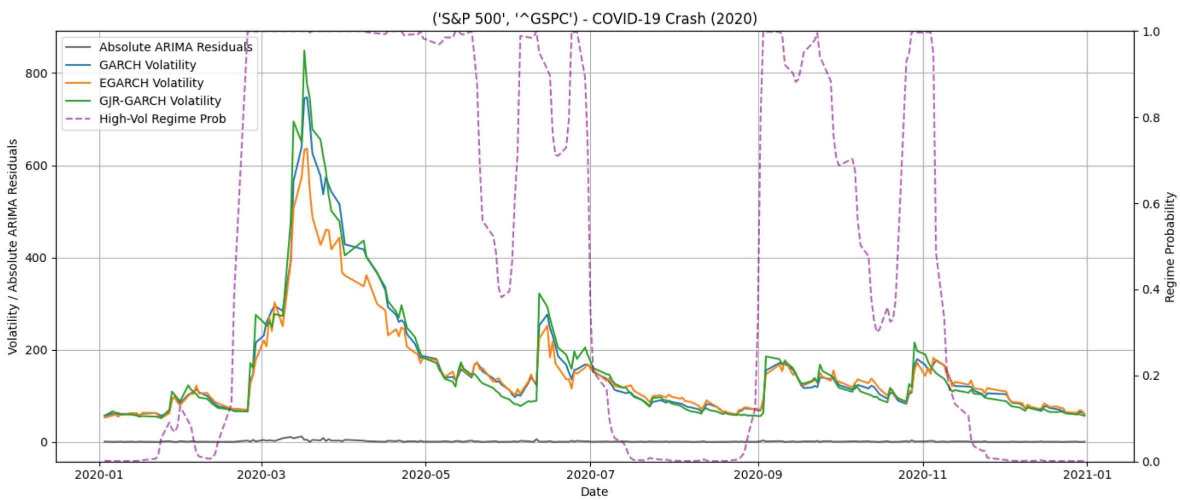

Figure 13: Volatility estimates and regime-switching probabilities during the COVID-19 crash for S&P 500.

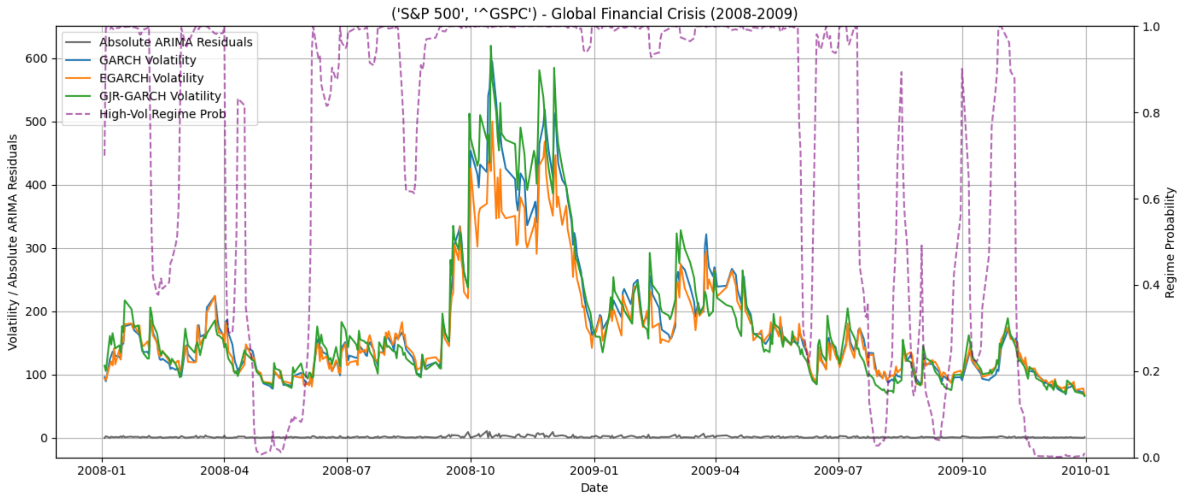

Figure 14: Volatility estimates and regime-switching probabilities during the Global Financial Crisis (2008-09) for S&P 500.

6 Discussion

The results of this project reveal several insights into the volatility dynamics of financial markets. The GARCH models applied to ARIMA residuals successfully captured volatility clustering and asymmetry. Across all three indices (S&P 500, NASDAQ, Nifty 50), GJR-GARCH and EGARCH models outperformed the symmetric GARCH model in terms of forecast accuracy (MSE and QLIKE) and Value-at-Risk (VaR) exceedance coverage. These improvements underscore the importance of modeling asymmetric shock responses in financial time series.

The statistical significance of the γ parameters in GJR-GARCH models confirmed that negative shocks lead to higher volatility increases than positive ones. This effect was particularly pronounced during crisis periods, as seen in the model outputs for 2008–2009 and 2020.

The Markov Switching Autoregressive (MS-AR) models added further value by capturing regime shifts. Smoothed marginal probabilities of being in a high-volatility regime closely aligned with known financial crisis periods, highlighting the model's ability to detect states of market stress. Notably, applying the MS-AR model both to filtered residuals and raw return series yielded qualitatively similar regime dynamics, affirming the robustness of regime detection.

During both the Global Financial Crisis and the COVID-19 crash, regime-switching probabilities surged alongside spikes in GARCH conditional volatility estimates and ARIMA residuals. This reinforces that volatility and regime dynamics are complementary dimensions of financial risk modeling.

7 Conclusion

This study demonstrated the effectiveness of advanced time series models in forecasting financial volatility and identifying structural regimes in equity index returns. The analysis captured both persistent and abrupt changes in market behavior through the application of ARIMA , GARCH models, and Markov Switching Autoregressive model.

The empirical results revealed that GJR-GARCH and EGARCH models outperform symmetric GARCH, particularly in turbulent periods. Leverage effects were confirmed across markets, and rolling forecast evaluations showed superior predictive accuracy for asymmetric models. Additionally, Markov Switching models consistently detected high-volatility regimes that aligned with major financial crises.

By integrating volatility modeling with regime detection, the project provides a framework that is both descriptive and predictive. The findings offer valuable implications for practitioners in risk management, asset allocation, and financial regulation.