# PREDICTIVE ANALYSIS OF RETURN VOLATILITY FOR THE S&P 500 INDEX

**Yuchen Gu**
57126260
yuchengu4-c@my.cityu.edu.hk

**Yulin Cai**
57126990
yulincai2-c@my.cityu.edu.hk

**Yiyao Jin**
57124923
yiyaojin2-c@my.cityu.edu.hk

**Yang Li**
56843298
yli669-c@my.cityu.edu.hk

April, 2024

**ABSTRACT**

This study utilizes a range of advanced econometric models to analyze the volatility dynamics of the S&P 500 Index, including ARCH(1), GARCH(1,1), AR-X-GARCH, and Threshold GARCH. Applied to historical data from Yahoo Finance covering April 8, 2019, to April 5, 2024, the models were assessed using Akaike's Information Criterion (AIC) and log-likelihood values, with further validation through the analysis of Volatility Forecasting Plots. The AR-X-GARCH model demonstrated the best performance, with the highest log-likelihood and lowest AIC scores. However, the Threshold GARCH model showed a propensity for overreaction to market extremes, indicating a need for further model calibration. The study confirms the critical role of advanced models in capturing market volatility and underscores the importance of continuous enhancement and the inclusion of more volatility factors for improved performance, especially in extreme market conditions. The findings aim to aid financial analysts and investors in making informed decisions and provide a foundation for future research in financial time-series analysis.

## 1 Motivation and Background

The S&P 500 Index, created in 1957 by Standard & Poor's Dow Jones Indices LLC, stands as a principal benchmark for the U.S. stock market. It tracks the performance of 500 leading companies listed on major U.S. exchanges such as the New York Stock Exchange and NASDAQ. A specialized committee selects these constituents and updates them quarterly to mirror market shifts. The index encompasses well-known industrial stocks like Boeing and General Motors, consumer stocks such as Coca-Cola and Procter & Gamble, and technology stocks including Apple and Google. Known for its broad market representation, market-cap weighting, and dynamic updates, the S&P 500 serves as a vital tool for quickly grasping the overall trends in the U.S. stock market.

## 2 Objectives

The primary objective of this report is to utilize various advanced econometric models to analyze and forecast the volatility dynamics of the S&P 500 (SPX). By comparing different models, we aim to offer a comprehensive understanding of its volatility patterns. This analysis will provide valuable insights to financial analysts and investors, aiding them in making informed decisions.

## 3 Methods

Preliminary tests were conducted to confirm the appropriateness of the volatility models chosen for analysis. By executing the White test, the p-value of 0.0016, which is less than the threshold of 0.05, indicates the presence of heteroscedasticity in the residual errors.

In light of this, four econometric models were employed to analyze and forecast the volatility dynamics of the S&P 500 Index: Autoregressive Conditional Heteroskedasticity (ARCH(1)), Generalized Autoregressive Conditional Heteroskedasticity (GARCH(1,1)), Component Generalized Autoregressive Conditional Heteroskedasticity (Component GARCH), and Threshold Generalized Autoregressive Conditional Heteroskedasticity (Threshold GARCH). The fit of these models is assessed using the Akaike Information Criterion (AIC) and Log-likelihood values. Additionally, the predictive performance of the models is validated through the analysis of Volatility Forecasting Plots.

# 4 Data Collection and Data Prepossessing

## 4.1 Data Collection

The historical data utilized in this report was sourced from Yahoo Finance, spanning from April 8, 2019, to April 5, 2024. The dataset encompasses opening and closing prices, highest and lowest prices, adjusted closing prices, and trading volumes, totaling 1259 effective trading days.

## 4.2 Data Prepossessing

To calculate daily returns, we applied the formula

$$R_n = \frac{C_{n+1} - C_n}{C_n}$$

where $R_n$ represents the return on day $n$, and $C_n$ denotes the closing price on day $n$. We partitioned the entire dataset into two subsets: a training set and a test set. The test set encompasses the last 252 trading days of the time series, roughly equivalent to one full trading year. The training set comprises all the data preceding the test set. Models were trained and their fit assessed on the training set, followed by predictions and accuracy evaluations conducted on the test set. This partitioning ensures a rigorous assessment of model performance over an unseen period of market activity.

# 5 Model Fitting and Forecasting

## 5.1 ARCH(1)

The autoregressive conditional heteroskedasticity (ARCH) model is a statistical model of conditional variance. It is usually used for predicting asset return volatility. The ARCH(1) model is defined as:

$$h_t = \omega + \alpha r_{t-1}^2$$

In ARCH(1), the length of ARCH lags is 1, meaning that the conditional variance of a given day is not constant and is a function of the previous day's news .

Constant Mean - ARCH Model Results

| | | | |
|---|---|---|---|
| Dep. Variable: | returns(%) | R-squared: | 0.000 |
| Mean Model: | Constant Mean | Adj. R-squared: | 0.000 |
| Vol Model: | ARCH | Log-Likelihood: | -1655.21 |
| Distribution: | Normal | AIC: | 3316.42 |
| Method: | Maximum Likelihood | BIC: | 3331.16 |
| | | No. Observations: | 1005 |
| Date: | Sat, Apr 13 2024 | Df Residuals: | 1004 |
| Time: | 21:10:59 | Df Model: | 1 |

Mean Model

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| mu | 0.1356 | 4.006e-02 | 3.385 | 7.121e-04 | [5.708e-02, 0.214] |

Volatility Model

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| omega | 1.0120 | 0.110 | 9.170 | 4.744e-20 | [ 0.796, 1.228] |
| alpha[1] | 0.5595 | 0.143 | 3.915 | 9.038e-05 | [ 0.279, 0.840] |

Figure 1: Constant Mean - ARCH Model Results

The table presented in Figure 1 summarizes the results of an ARCH(1) model. This model was trained on data collected from the 9th of April, 2019 to the 4th of April, 2023. It is utilized to predict the volatility over the next year. The

Log-likelihood value is recorded at -1655.21, while the Akaike Information Criterion (AIC) is 3316.42. Both the p-values for omega and alpha are less than 0.05, which indicates that these parameters are statistically significant and the effects are non-random.

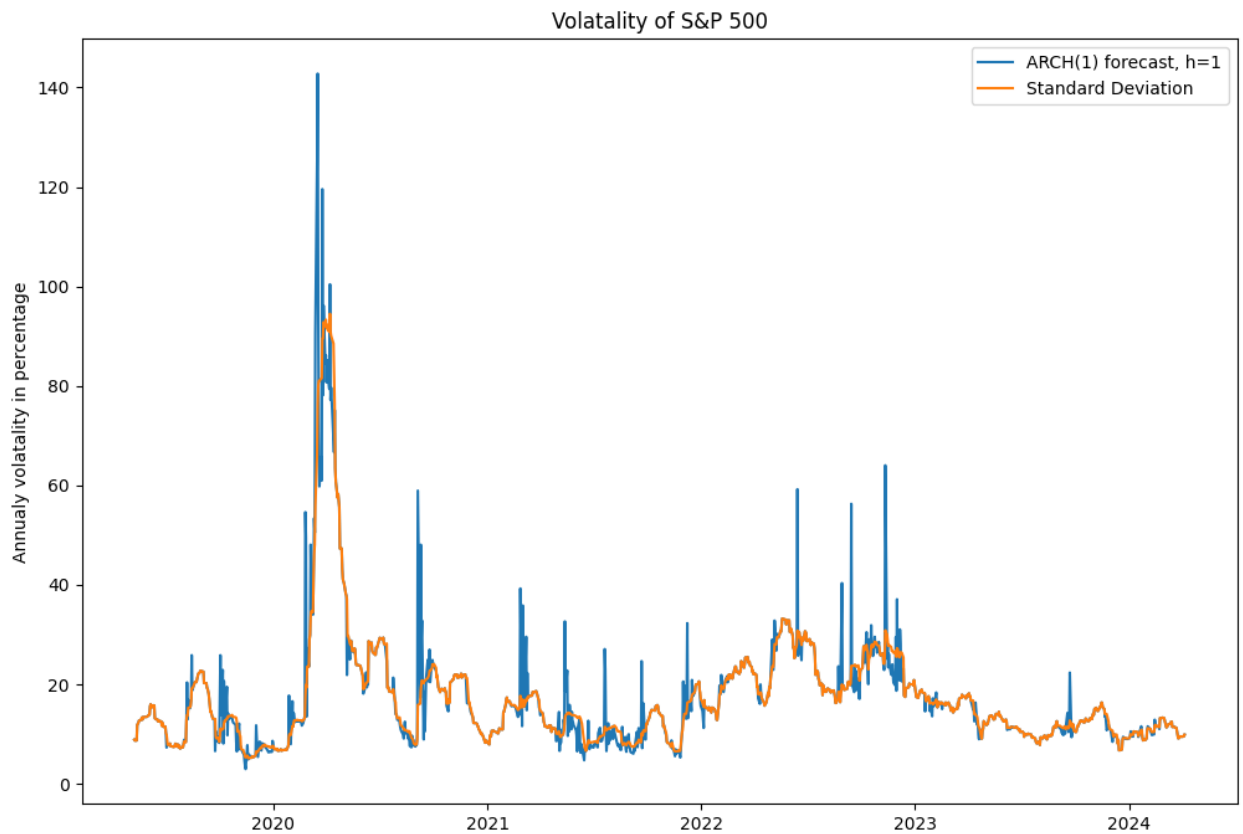


Figure 2: S&P 500 Volatility: ARCH(1) Predictions vs. Actual

Figure 2 depicts the comparison between the predicted and actual volatility of the S&P 500. The blue line represents the daily volatility forecast generated by an ARCH(1) model trained on the preceding 21 days' data. The orange line indicates the actual observed volatility. While both trends appear to align closely, the predicted volatility occasionally overshoots the actual values, especially noticeable during periods of sharp volatility increases.

### 5.2 GARCH(1,1)

GARCH, which stands for Generalized AutoRegressive Conditional Heteroskedasticity, is a statistical model that is useful for analyzing time-series data in the presence of heteroskedastic effects, which means the variance of the error term is not constant. GARCH models assume that the variance of the error term is serially autocorrelated following an autoregressive moving average process. In general, the model is denoted as GARCH (p, q), where $p$ represents the number of past squared terms of conditional variance considered in the model, and $q$ represents the number of past conditional variance terms considered in the model. The simplest version of the model is GARCH (1, 1), which captures the impact of the squared residuals from the past one time point and the impact of the conditional variances from the past one time point. The formula is as follows:

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

Where $r_t$ is the t-th log return and $\sigma_t$ is the t-th volatility estimate in the past.

Below, the fit summary of the GARCH(1,1) model is shown.

Constant Mean - GARCH Model Results

| | | | |
|---|---|---|---|
| Dep. Variable: | returns(%) | R-squared: | 0.000 |
| Mean Model: | Constant Mean | Adj. R-squared: | 0.000 |
| Vol Model: | GARCH | Log-Likelihood: | -1517.91 |
| Distribution: | Normal | AIC: | 3043.83 |
| Method: | Maximum Likelihood | BIC: | 3063.48 |
| | | No. Observations: | 1005 |
| Date: | Sat, Apr 13 2024 | Df Residuals: | 1004 |
| Time: | 21:10:59 | Df Model: | 1 |

Mean Model

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| mu | 0.0911 | 2.865e-02 | 3.180 | 1.472e-03 | [3.496e-02, 0.147] |

Volatility Model

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| omega | 0.0464 | 1.698e-02 | 2.731 | 6.316e-03 | [1.309e-02,7.964e-02] |
| alpha[1] | 0.2043 | 4.567e-02 | 4.473 | 7.722e-06 | [ 0.115, 0.294] |
| beta[1] | 0.7830 | 3.758e-02 | 20.835 | 2.072e-96 | [ 0.709, 0.857] |

Figure 3: GARCH(1, 1) Model Summary

The table in Figure 3 presents a summary of the GARCH(1, 1) model that has been trained on data from the 9th of April, 2019 to the 4th of April, 2023, and is used for predicting volatility over the next year. The model exhibits a Log-likelihood value of -1517.91 and an AIC of 3043.83, suggesting improved performance over the ARCH(1) model. P-values for the parameters omega, alpha, and beta are all below the 0.05 threshold, denoting their statistical significance. The long-run average variance omega is about 0.0464.

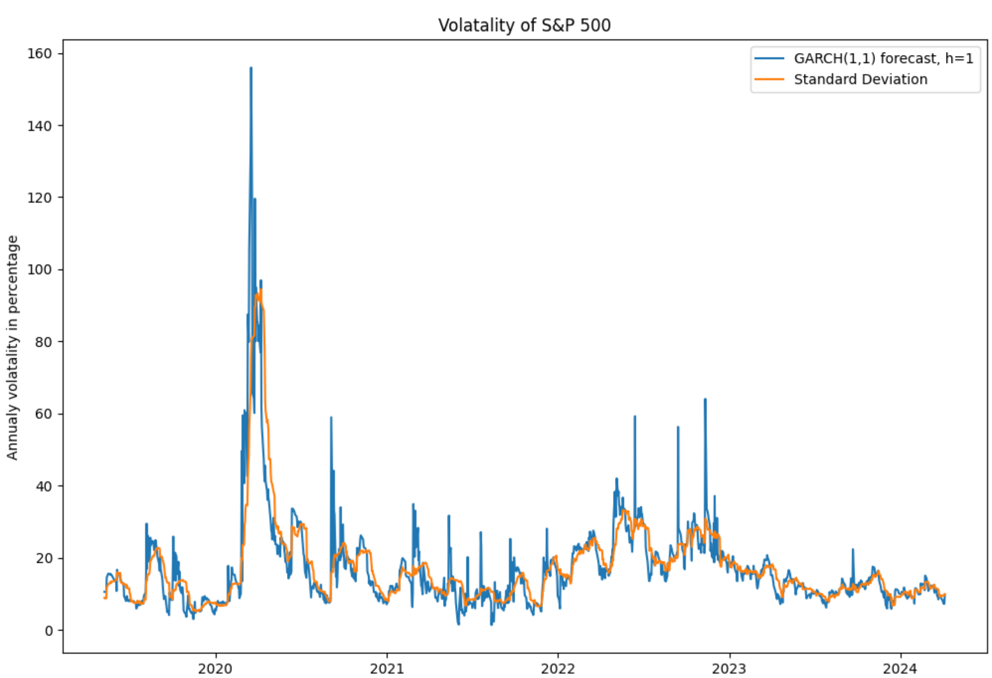


Figure 4: GARCH(1,1) Volatility Predictions

The illustration in Figure 4 compares the GARCH(1,1) model's forecasted volatility against the observed standard deviation for the S&P 500. Both lines display similar overall trends. Notably, at several points, the predicted volatilities are significantly higher than the actual volatilities. These peaks in the predicted data occur at positions similar to those in the ARCH model.

### 5.3 Component GARCH

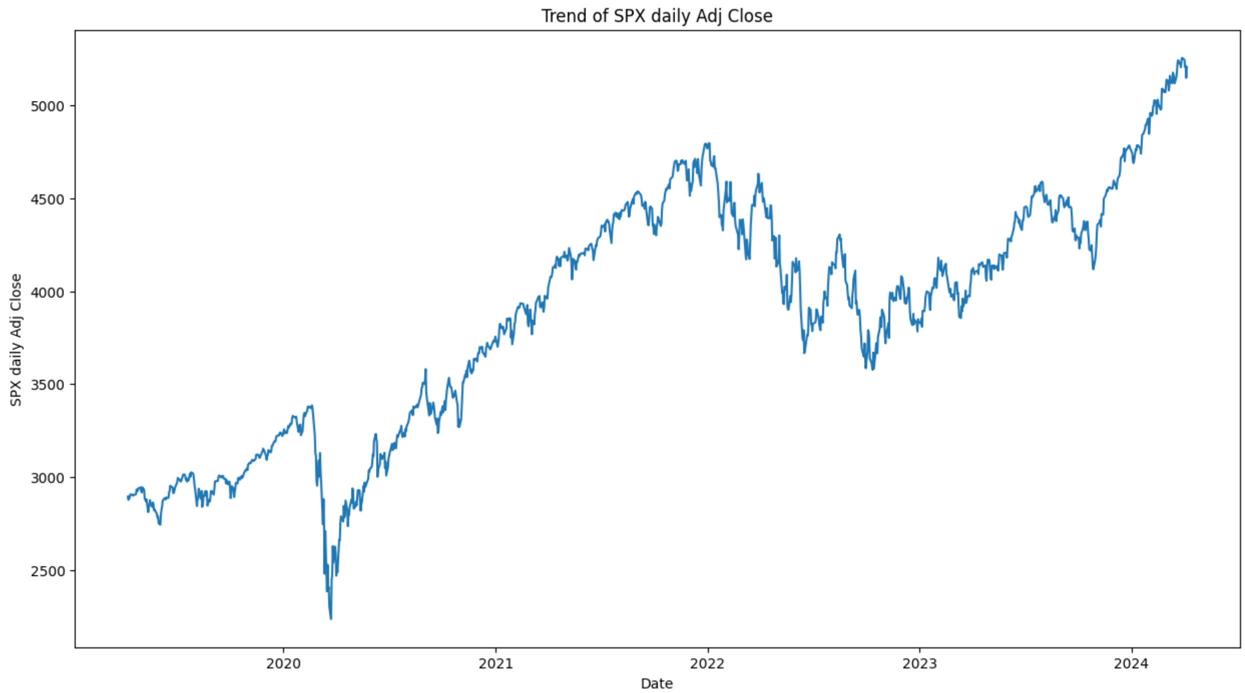


Figure 5: S&P 500 Adjusted Close Trend

Observably, the overall trend of the S&P 500 is upward Engle [1982], Bollerslev [1986], although downward trends were observed in early 2020, and again in the latter halves of 2022 and 2023. Volatility modeling is crucial in economics, serving as a pivotal metric in the study of investment portfolios, asset pricing, and hedging strategies Lai and Xing [2013]. For a nuanced analysis of volatility over multiple time scales, we applied the Component GARCH model. This advanced model extends GARCH by incorporating multiple volatility components, enabling the examination of different frequencies of volatility.

Our mean model of choice for the Component GARCH analysis was the Autoregressive Exogenous Variable (ARX) model. ARX combines the Autoregressive (AR) approach with the Exogenous Variable (X) model to capture both the dynamic nature of time series data and the influence of external factors.

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_{t-1} + \varepsilon_t \tag{2}$$

where:

- $Y_t$ represents the dependent variable (or the target variable) at time $t$.
- $X_{t-1}$ represents the exogenous variable at time $t-1$.
- $\beta_0$ is the intercept term of the model.
- $\beta_1$ is the coefficient of the autoregressive term, used to measure the time-dependence or autocorrelation of the dependent variable.
- $\beta_2$ is the coefficient of the exogenous variable, used to measure the impact of the exogenous variable on the dependent variable.
- $\varepsilon_t$ represents the error term at time $t$, which accounts for the random errors not explained by the model.

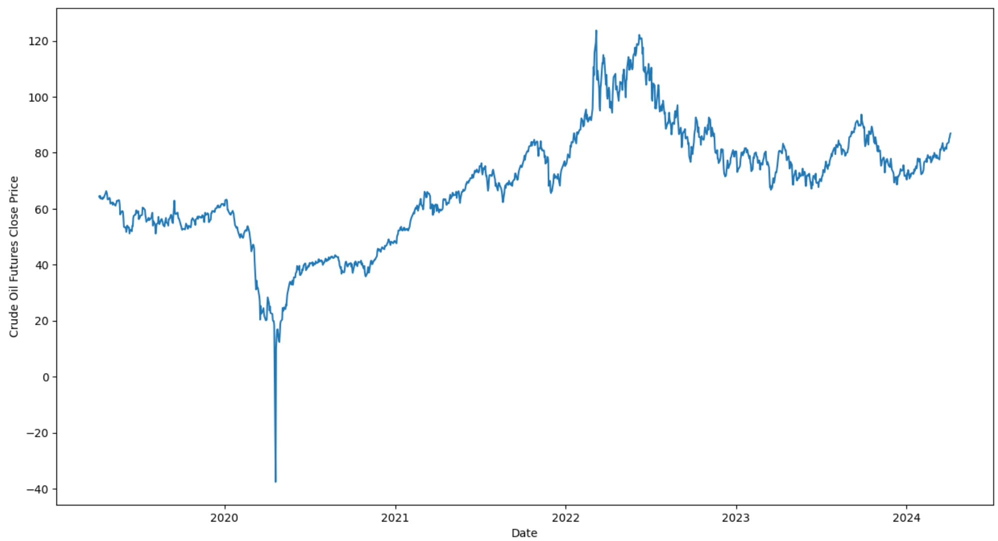


Figure 6: Crude Oil Futures Close Price

We chose Crude Oil Futures as the exogenous variable, considering petroleum's significant role as a primary energy source and industrial raw material. With 67.2% of US petroleum consumption utilized for transportation and 26.9% for industrial purposes in 2021, as reported by the EIA (2022), oil price fluctuations critically affect global economic conditions. The Crude Oil Futures benchmark plays a central role in oil price discussions and market strategies.

Below, the fit summary of the Component GARCH model is shown.

AR-X - GARCH Model Results

| | | | |
|---|---|---|---|
| Dep. Variable: | returns(%) | R-squared: | 0.034 |
| Mean Model: | AR-X | Adj. R-squared: | 0.029 |
| Vol Model: | GARCH | Log-Likelihood: | -1497.29 |
| Distribution: | Normal | AIC: | 3012.58 |
| Method: | Maximum Likelihood | BIC: | 3056.76 |
| | | No. Observations: | 1001 |
| Date: | Sat, Apr 13 2024 | Df Residuals: | 995 |
| Time: | 21:10:59 | Df Model: | 6 |

Mean Model

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| Const | 0.1078 | 3.045e-02 | 3.539 | 4.013e-04 | [4.809e-02, 0.167] |
| returns(%)[1] | -0.0663 | 3.455e-02 | -1.920 | 5.486e-02 | [ -0.134,1.382e-03] |
| returns(%)[2] | 0.0134 | 3.498e-02 | 0.382 | 0.703 | [-5.520e-02,8.191e-02] |
| returns(%)[3] | -0.0623 | 3.391e-02 | -1.838 | 6.602e-02 | [ -0.129,4.125e-03] |
| returns(%)[4] | -0.0244 | 3.473e-02 | -0.702 | 0.483 | [-9.243e-02,4.370e-02] |
| Change | 0.0316 | 1.618e-02 | 1.955 | 5.064e-02 | [-8.778e-05,6.333e-02] |

Volatility Model

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| omega | 0.0403 | 1.467e-02 | 2.747 | 6.022e-03 | [1.154e-02,6.907e-02] |
| alpha[1] | 0.2063 | 5.067e-02 | 4.072 | 4.660e-05 | [ 0.107, 0.306] |
| beta[1] | 0.7856 | 4.167e-02 | 18.854 | 2.728e-79 | [ 0.704, 0.867] |

Figure 7: AR-X-GARCH Model Results

The summary table in Figure 7 outlines the AR-X-GARCH model that has been trained on data from April 9, 2019, to April 4, 2023. The model is applied to forecast the volatility for the forthcoming year. It exhibits a Log-likelihood value of -1497.29 and an AIC of 3012.58. The p-values for omega, alpha, and beta are all below the 0.05 level, signifying the statistical significance of these parameters.

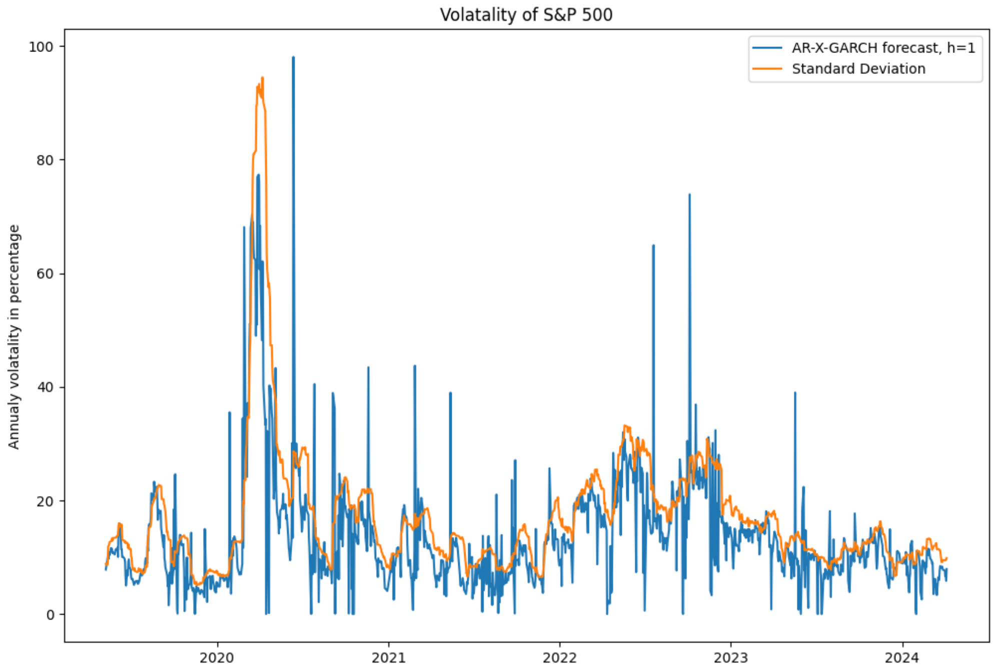


Figure 8: AR-X-GARCH Volatility Forecast

As depicted in Figure 8, the trends of the forecasted (blue line) and actual (orange line) volatility of the S&P 500 are similar, yet the forecasted volatility exhibits a higher fluctuation amplitude. Because in the ARX component, we set lag=4 to capture the time series short-run dependency.

In comparison to the earlier models, this AR-X-GARCH model results in a greater number of forecasts that deviate significantly below the actual observed values.

The actual volatility of the S&P 500 experienced a significant increase in 2020. While all models captured this change, only the AR-X-GARCH model's predicted volatility closely matched the actual volatility, while other models provided overly high forecasts. Similarly, the Crude Oil Futures Close Price underwent a drastic fluctuation around the same time period but quickly reverted to relatively normal levels. Combining these observations, we can speculate that the incorporation of exogenous variables constrained the model, leading to forecasts that better aligned with real-world trends.

### 5.4 Threshold GARCH

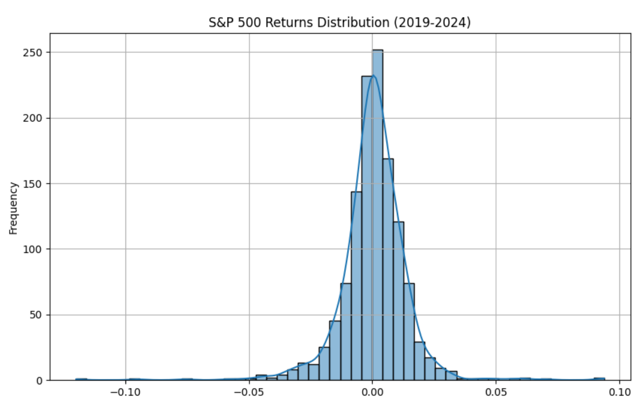


Figure 9: S&P 500 Returns Distribution (2019-2024)

In our preliminary analysis of the S&P 500 index returns from 2019 to 2024, the distribution graph and data results indicate a skewness of -0.523 in the return distribution, signifying a pronounced left tail. This negative skewness suggests that negative returns occur more frequently than positive returns of the same magnitude. Furthermore, the kurtosis of the return distribution stands at 13.308, pointing to a leptokurtic or "fat-tailed" distribution, which implies a higher probability of extreme values compared to a normal distribution. The TGARCH (Threshold Generalized Autoregressive Conditional Heteroskedasticity) model, owing to its adeptness at capturing the asymmetry in volatility and the "leverage effect," along with its flexibility in handling fat-tailed data, proves to be more suitable than the standard GARCH model. This constitutes the rationale behind our selection of the TGARCH model for the analysis. The Threshold GARCH (T-GARCH) model equation is expressed as follows:

$$\sigma_t^2 = \omega + (\alpha + \gamma I_{t-1})\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

where $I_{t-1}$ is an indicator function that takes the value of 1 if $\epsilon_{t-1} < 0$ and 0 otherwise, which allows the model to capture the different impacts of positive and negative shocks on volatility.

Below, the fit summary of the TGARCH model is shown.

Constant Mean - GJR-GARCH Model Results

| | | | |
|---|---|---|---|
| Dep. Variable: | returns(%) | R-squared: | 0.000 |
| Mean Model: | Constant Mean | Adj. R-squared: | 0.000 |
| Vol Model: | GJR-GARCH | Log-Likelihood: | -1485.35 |
| Distribution: | Standardized Student's t | AIC: | 2982.71 |
| Method: | Maximum Likelihood | BIC: | 3012.19 |
| | | No. Observations: | 1005 |
| Date: | Sun, Apr 14 2024 | Df Residuals: | 1004 |
| Time: | 11:36:30 | Df Model: | 1 |

Mean Model

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| mu | 0.0872 | 2.505e-02 | 3.479 | 5.035e-04 | [3.806e-02, 0.136] |

Volatility Model

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| omega | 0.0321 | 1.057e-02 | 3.040 | 2.362e-03 | [1.142e-02,5.285e-02] |
| alpha[1] | 0.0422 | 2.964e-02 | 1.423 | 0.155 | [-1.590e-02, 0.100] |
| gamma[1] | 0.2588 | 6.766e-02 | 3.825 | 1.307e-04 | [ 0.126, 0.391] |
| beta[1] | 0.8198 | 3.101e-02 | 26.437 | 5.116e-154 | [ 0.759, 0.881] |

Distribution

| | coef | std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| nu | 6.8634 | 1.437 | 4.776 | 1.790e-06 | [ 4.047, 9.680] |

Covariance estimator: robust

Figure 10: Threshold GARCH Model Summary

Figure 10 presents the fit summary of the Threshold GARCH (TGARCH) model, trained on data spanning from April 9, 2019, to April 4, 2023, and is designed for forecasting the forthcoming year's volatility. The model reports a Log-likelihood of -1510.67 and an AIC of 3031.3. The $\omega$, $\gamma$, and $\beta$ parameters have statistically significant $p$-values below 0.05, whereas the $\alpha$ parameter has a $p$-value exceeding the 0.05 threshold, indicating it is not statistically significant.

Figure 11 illustrates the volatility of the S&P 500, with the TGARCH model forecast represented by the blue line, and the actual standard deviation by the orange line. The trends of both lines are comparable; however, the fluctuations within the TGARCH forecast are more pronounced, suggesting a heightened sensitivity to market extremes which

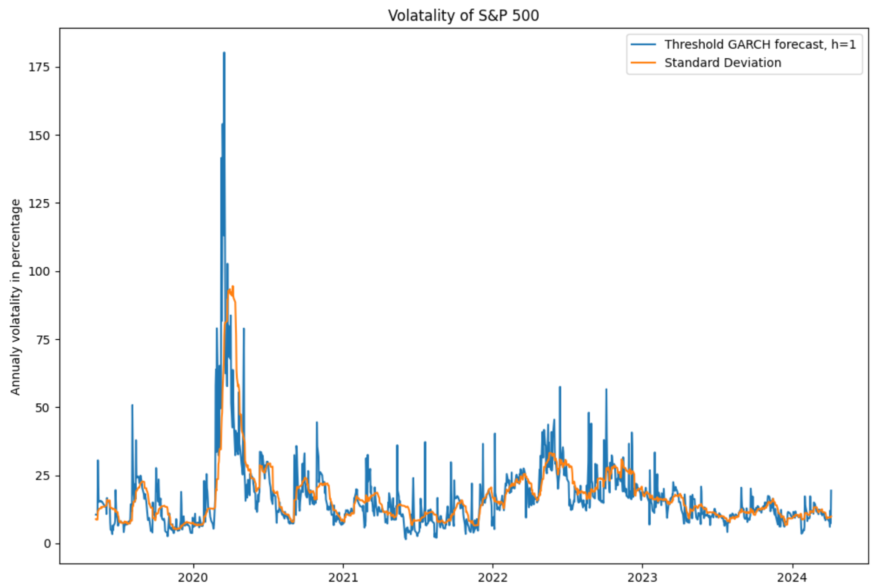


Figure 11: S&P 500 Volatility: Threshold GARCH vs. Standard Deviation

may be linked to its leverage effect—a common phenomenon in financial markets where negative shocks have a disproportionate impact.

The advantage of the TGARCH model lies in its ability to differentiate between the impacts of positive and negative shocks on volatility, thanks to the incorporation of an indicator function $I_{t-1}$. This feature makes the model adept at capturing the asymmetric volatility often observed in the actual markets. Nonetheless, the non-significance of the $\alpha$ parameter indicates a potential shortfall in the model's capacity to capture all the dynamics of volatility through its autoregressive structure. Future work may benefit from exploring models that include a broader range of volatility determinants or combining different types of GARCH models to enhance predictive accuracy. Additionally, further calibration and refinement of the model, especially in predicting extreme market conditions, may reduce the propensity for overreaction and yield smoother volatility forecasts.

### 5.5 Result Analysis and Model Comparison

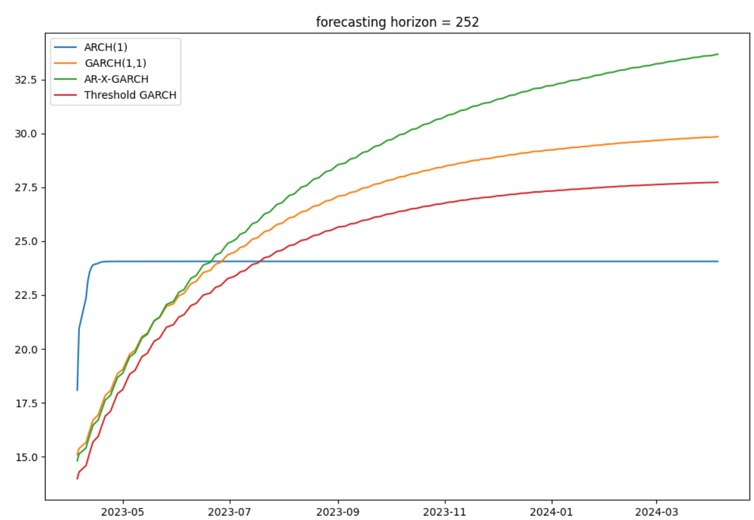


Figure 12: Volatility Prediction Comparison

The graph above presents a comparative forecast of the S&P 500 Index's volatility over the forthcoming year. The blue line, representing the Autoregressive Conditional Heteroskedasticity model (ARCH(1)), initially indicates a steep rise in volatility, which then levels out, maintaining a rate below 18%. The Generalized Autoregressive Conditional Heteroskedasticity model (GARCH(1,1)) and the Component GARCH model (AR-X-GARCH), represented by the orange and green lines respectively, show a consistent and gradual increase in predicted volatility. Notably, the red line, depicting the Threshold GARCH model, presents a marked volatility forecast that begins more conservatively but eventually exhibits similar trends to the GARCH(1,1) and AR-X-GARCH models. This convergence suggests a potential overestimation of market reactions to recent events in the Threshold GARCH model. Future refinements could focus on calibrating this model to dampen initial overreactions and align more closely with established volatility patterns, or the development of hybrid models that could temper the fluctuations for more balanced volatility forecasting.

According to the fit summary tables provided, the Component GARCH model (AR-X-GARCH) outshines the other models. Specifically, the AR-X-GARCH model boasts the highest log-likelihood and the lowest Akaike Information Criterion (AIC) values, indicating a superior fit to the data. The results highlight its efficiency and accuracy in capturing the volatility structure of the S&P 500 Index.

These quantitative metrics provide strong evidence of the AR-X-GARCH model's applicability and predictive capacity in complex financial time-series analysis. Compared to other models, AR-X-GARCH better reflects the intricacies of market variations, possibly due to its advanced approach in capturing volatility across multiple time scales. Future research may delve further into the performance of such models under diverse financial conditions and consider how model parameter adjustments could enhance robustness in forecasting extreme market events. This analysis should be included in the "Result Analysis and Model Comparison" section of the academic paper to discuss the comparative performance and implications for future research.

## 6 Conclusion

This study has employed a variety of advanced econometric models to analyze and forecast the volatility dynamics of the S&P 500 Index, an essential benchmark for the U.S. stock market. The ARCH(1), GARCH(1,1), Component GARCH (AR-X-GARCH), and Threshold GARCH models have been rigorously tested and evaluated, providing insights into the volatility patterns of the index. The historical data from Yahoo Finance served as the foundation for our analysis, contributing to a robust examination of market trends over 1259 effective trading days.

The AR-X-GARCH model has been identified as the superior model, exhibiting the highest log-likelihood and the lowest Akaike Information Criterion (AIC) values, indicating a more precise fit compared to the others. This model adeptly captures the multi-scaled volatility structure of the S&P 500 Index, which is crucial for understanding market variations.

Despite the strengths of the AR-X-GARCH model, the volatility forecasts, particularly from the Threshold GARCH model, have shown areas of overreaction to market extremes. This highlights the need for further calibration and potentially integrating a broader range of volatility determinants to enhance the models' robustness, especially in the context of extreme market conditions. Future research should also consider the application of these models across different financial indices and environments to verify their efficacy and to fine-tune the model parameters for even more reliable predictions.

## 7 Appendix

Data source: https://finance.yahoo.com/quote/%5ESPX?.tsrc=fin-srch
Our `Python` source code can be found here:
`https://github.com/Fiona1224/GARCH-1-1--SPX`
`https://github.com/YOLANDALI/Stock3022/blob/main/C_GARCH.ipynb`
Additional dataset: `https://cn.investing.com/commodities/crude-oil-historical-data`

## References

Robert F Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1008, 1982.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

Tze Leung Lai and Hai Xing. Stochastic change-point arx-garch models and their applications to econometric time series. *Statistica Sinica*, 23(4):1573–1594, 2013.