Model Risk

Group 3 Model Risk of Group 2 – GJR-GARCH
Quantitative Risk Management – ECON 6295_10
By: Melda Akan | Oluwatobi Ige | Jeff Matthew Sam | Rakesh Jayswal | Salmah Animashaun

1. Business Case

Group 2 evaluated a new client's 10-day Value-at-Risk (VaR) for a portfolio consisting of \$10,000 cash, 100 shares of Apple Inc. (AAPL) at \$186.79, and 100 shares of SPDR S&P (SPY) at \$281.12. They utilized GARCH volatility modeling, focusing on the GJR-GARCH model as recommended. Group 2 sourced one year of AAPL and SPY daily price data from Yahoo Finance to calculate returns. The GJR-GARCH model forecasted 10-day volatility, with data used for VaR estimates through parametric methods and Monte Carlo simulations, ensuring robust risk evaluation as per the risk managers' requirements.

As Group 3, we evaluated the GJR-GARCH model's suitability for business use, focusing on the model's conceptual soundness, data quality, and documentation clarity. We identified key issues, assigned severity ratings, and assessed Group 2's responses to these concerns. We also evaluated back-testing performance for predictive accuracy and conducted independent verification of assumptions, parameter choices, and results.

2. Description of Data

Group 2 data analysis examines the historical performance of AAPL (Apple Inc.) and SPY (S&P 500). The dataset comprises daily closing prices over 252 trading days, covering a year of market activity. AAPL's prices range from \$164.41 to \$236.22, with an average of \$198.67 and a standard deviation of \$21.50, indicating moderate volatility. SPY shows a broader price range, with a minimum of \$417.02, a maximum of \$584.59, and an average of \$515.55, while its standard deviation of \$41.18 reflects greater variability than AAPL. Both datasets exhibit upward price trends, although AAPL shows more significant fluctuations.

Regarding returns, AAPL demonstrates higher variation, evident in more significant price spikes than SPY. This volatility highlights AAPL's sensitivity to market changes, making it a riskier asset. Despite this, a correlation exists between AAPL and SPY returns, indicating some synchronized market behavior. In Group 3, we delve into data for risk analysis and modeling. The dataset includes daily closing prices and percentage changes for AAPL and SPY, forming a basis for portfolio risk evaluation. AAPL's higher volatility presents unique challenges for assessment, while SPY serves as a stable benchmark with lower variability but still influences overall portfolio risk due to its correlation with AAPL.

3. Model Description

In compliance with the risk managers' request not to use a simple GARCH method, Group 2 is advised by GJR-GARCH to forecast the volatility of the new client's portfolio of equity holdings. The GJR (Glosten-Jagannathan-Runkle Generalized Autoregressive Conditional Heteroskedasticity) model is a nonlinear extension of the standard GARCH model that captures the asymmetric effects of disturbances on conditional variance, meaning it captures the "leverage effect" commonly observed in financial markets. The leverage effect is a phenomenon where negative shocks tend to increase future volatility more than positive shocks of the same magnitude. The risk modeling team emphasized capturing asymmetrical volatility responses. The following GARCH variants were tested:

Model Name	Description
Responsive GJR-GARCH	A variant of the GJR-GARCH model that responds quickly to market changes.
High Memory GARCH	A GARCH model with a high memory parameter, capturing long-term dependencies.
Lower Order GARCH	A GARCH model with a lower order, typically simpler and faster to compute.
Lower Order normal GARCH	A normal GARCH model with a lower order, assuming normal distribution of errors.
Normal GARCH	A standard GARCH model assuming normal distribution of errors.
Conservative GJR-GARCH	A GJR-GARCH model that is more conservative in its responses to market changes.
Short Memory GARCH	A GARCH model with a short memory parameter, focusing on short-term dependencies.

Table I : Garch Models

Back-testing performance assessment:

The estimation results reveal distinct volatility dynamics between the individual stock (AAPL) and the market index (SPY). SPY shows more substantial leverage effects ($\gamma = 0.2327846$) and higher volatility persistence ($\beta = 0.799885$), both with strong statistical significance. In contrast, AAPL exhibits weaker leverage effects and minimal volatility persistence, suggesting more independent volatility patterns. We found the Group 2 model's performance validated through back-testing using two different rolling windows on their presentation.

Window Size	MAE	RMSE
50-day	1.5972	2.0701
80-day	0.5125	0.6165

TableII:Model Validation and Back-testing:

The 50 day window size has the loss function of 2.0701 but the 80-day rolling window demonstrated superior accuracy, with substantially lower error metrics of 0.6165. This finding informed the final model implementation, favoring the longer estimation window for more stable and accurate volatility forecasts.

Forecast Generation:

The validated model generated 10-day volatility forecasts with 95% confidence intervals for both securities. These forecasts show higher predicted volatility levels for AAPL than SPY and gradually widening confidence intervals reflecting increasing forecast uncertainty with slight downward trends in volatility for both securities.

Risk Measurement Implementation

Group 2 final risk measurement implementation utilized 10,000 Monte Carlo simulations Brownian geometric motion for price path modeling - 95% confidence level for VaR calculation.

This comprehensive estimation process resulted in a 10-day VaR estimate of $9,454.08, representing the maximum expected portfolio loss at a 95% confidence level. This figure provides a robust risk metric for portfolio management and regulatory compliance purposes.

4. Model Evaluation

Issues found during evaluation and their ratings (Level 1, Level 2, Level 3):

We found that Group 2's presentation lacked clarity on their decision-making process and could have been supported with more research-based evidence. Group 2 selected the Conservative GJR-GARCH model as the best for both AAPL and SPY, not solely based on the Akaike Information Criterion (AIC), but rather due to the model's overall performance, despite its tendency to overfit as indicated in Tables IV and V. We found this issue to be a 'Level 1' problem and needs to be addressed urgently.

In the context of GARCH models, strict normality of returns is not a necessary assumption for the model to function effectively, but it is an important consideration when interpreting results, particularly for financial time series. Group 2 did not address this issue in their presentation nor did we find any relevant information in their presentation. Analysing the GARCH model results we found the residuals are not normal, violating a basic assumption of the GARCH model. We found this to be a 'Level 2' problem.

During the Model Risk presentation, Group 2 evaluated AIC, BIC, and MAE for the Conservative GJR-GARCH model without finding significant differences among the metrics. They assessed all tests collectively and concluded that Conservative GJR-GARCH performed better, primarily due to its complexity. While this approach aligns with machine learning modeling practices, in risk economics and time-series forecasting, it is standard to evaluate one metric at a time. Group 2 did not address this issue in their presentation nor did we find any relevant information in their presentation. We found this issue to be a 'Level 3' problem.

To analyze these issues, we conducted independent verification using the provided ipynb file and the shared presentation.

AAPL:

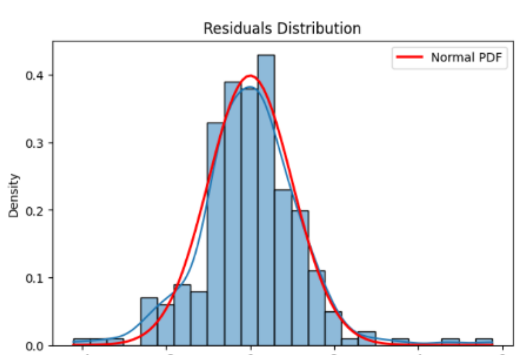

Fig i.

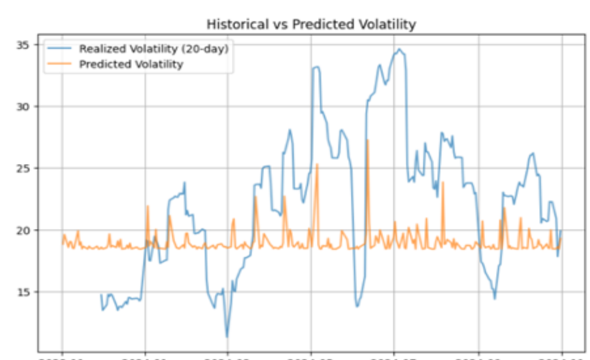

Fig ii

	AIC	BIC	Log-Likelihood	Max Forecast Error	Mean Absolute Error
Responsive GJR-GARCH	870.058326	894.736496	-428.029163	6.000937	0.748713
High Memory GARCH	875.970622	911.225151	-427.985311	6.045548	0.744603
Lower Order GARCH	868.130131	889.282848	-428.065065	6.023531	0.748736
Lower Order normal GARCH	888.937482	903.039294	-440.468741	5.960028	0.743159
Normal GARCH	890.452516	908.079780	-440.226258	5.921023	0.743688
Conservative GJR-GARCH	870.058326	894.736496	-428.029163	6.000937	0.748713
Short Memory GARCH	888.937482	903.039294	-440.468741	5.960028	0.743159

Table IV: Model Metrics for SPY Model

SPY:

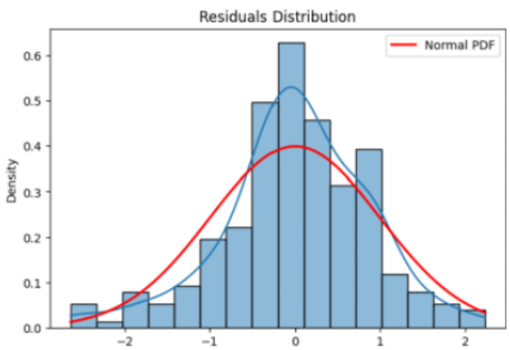

Fig iii

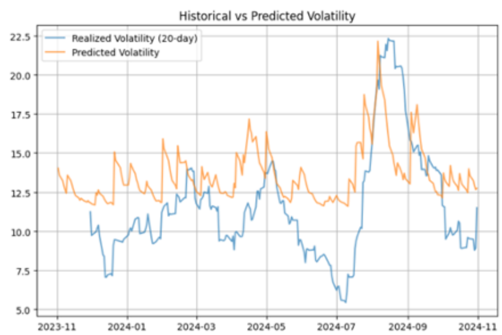

Fig iv

	AIC	BIC	Log-Likelihood	Max Forecast Error	Mean Absolute Error
Responsive GJR-GARCH	556.785007	581.463177	-271.392503	1.719973	0.487941
High Memory GARCH	561.883689	597.138218	-270.941844	1.705366	0.487619
Lower Order GARCH	567.475290	588.628007	-277.737645	1.869269	0.501133
Lower Order normal GARCH	572.109304	586.211116	-282.054652	1.881067	0.501599
Normal GARCH	562.433616	580.060881	-276.216808	1.779193	0.487935
Conservative GJR-GARCH	556.785007	581.463177	-271.392503	1.719973	0.487941
Short Memory GARCH	572.109304	586.211116	-282.054652	1.881067	0.501599

Table V: Model Metrics for SPY Model

Recommendation to use or not use the model in business:

Model diagnostics for AAPL (Figures i and ii) indicated that the Conservative GJR-GARCH model captured volatility effectively and provided mean-reverting forecasts, making it suitable for forecasting, albeit with non-normal residuals and potential underestimation of extreme events. In contrast, SPY diagnostics (Figures iii and iv) revealed model instability, including non-mean-reverting forecasts and insignificant parameters (alpha[1], lambda).

Tables IV and V showed that, based on AIC, the best models for AAPL and SPY were Lower Order GARCH and Responsive GJR-GARCH, respectively. However, we agree with the decision to implement Conservative GJR-GARCH for AAPL due to the model's ability to capture leverage effects, which may result in higher VaR estimates during market downturns. For SPY, we recommend considering Responsive GJR-GARCH, as it better captures asymmetric volatility responses and adapts more quickly to changing market conditions than Conservative GJR-GARCH. Using Conservative GJR-GARCH for an asset class with lower relative volatility, like SPY, could lead to underestimating volatility during rapid market shifts.

Despite the noted shortcomings, we recommend employing the model in this business context. Although the model is conceptually robust, Group 2 should have explicitly outlined their assumptions and the limitations of their approach, and the shortcomings we highlighted in their presentation.

Assumptions

The GARCH model uses several key assumptions to ensure accurate and meaningful results. First, stationarity assumes that the time series' statistical properties, such as mean, variance, and autocorrelation, remain constant over time. This stability is crucial for generating reliable forecasts. Second, the normality of innovations presumes that error terms are normally distributed, facilitating straightforward estimation and inference. Finally, the model relies on constant parameters, assuming that the coefficients in the variance equation do not change over the sample period, ensuring the consistency of volatility forecasts.

Limitations

Despite its usefulness, the GARCH model has notable limitations. Model misspecification can occur if the correct order of the model (p, q) is not selected, leading to biased results. Parameter instability may arise during periods of market turbulence or structural changes, reducing the model's accuracy. Additionally, the non-normality of innovations can undermine parameter estimates if returns exhibit skewness or kurtosis. Another challenge is model order selection, as choosing the appropriate lag structure can be complex. Lastly, conditional heteroskedasticity, while modeled explicitly, may not capture all real-world volatility patterns, limiting the model's applicability in specific financial environments.

6. Conclusion

Group 2's evaluation of the GJR-GARCH model demonstrates its effectiveness in capturing asymmetric volatility dynamics, particularly for AAPL and SPY. The Conservative GJR-GARCH model accurately incorporates leverage effects, essential for robust risk estimation during market downturns, and performs well for AAPL with an 80-day rolling window, validating its use for 10-day volatility forecasting and Value-at-Risk (VaR) estimation. However, its limitations with SPY, such as insignificant parameters and underestimation of rapid volatility shifts, suggest a Responsive GJR-GARCH model may better suit assets requiring quicker adaptability to market changes.

The 10-day VaR estimate of $9,454.08, obtained via Monte Carlo simulations, provides a reliable measure of portfolio risk. Still, the model's assumptions of stationarity and normality highlight the need for cautious interpretation and potential alternative models. To enhance the framework's reliability, aligning model selection with asset-specific characteristics, incorporating scenario analysis for extreme events, and regularly validating assumptions are recommended to ensure accurate and actionable risk metrics for decision-making.