Table of contents

1. OLS and origin of heteroskedasticity
2. Heteroskedasticity
3. Volatility
4. ARCH/GARCH Models
5. Rolling Forecast
6. Value at Risk
7. Conclusion

**1) OLS and origin of heteroskedasticity**

OLS stands for Ordinary Least Squares, which is a method for estimating the parameters in a Multiple Linear Regression (MLR) model. A multivariate linear regression represents a relationship of proportionality between dependent(y) and multiple independent variables(xi) and is modeled as follows:

where i = 1,2,3,4 …. n,

The above linear regression model states that the expectation of the variable y is β times E(X) plus a constant α. The proportionality relationship between y and x is not exact but also contains an error ε. Standard regression theory assumes that the error ε has a zero mean and a constant standard deviation σ independent of all variables. The standard deviation is the square root of the variance, which is the expectation of the squared error:

It is a positive number that measures the size of the error.When the expectedsize of this error is constant and does not depend on the sizeof the variable x, we call this assumption, homoskedasticity.

**Heteroskedasticity**

Financial return data, in which the variances of the error terms are not equal, error terms may reasonably be expected to be larger for some points or ranges of the data than for others, are said to suffer from heteroskedasticity, which literally means “various dispersion” in ancient Greek. In the context of regression analysis, heteroskedasticity occurs when the variance of the errors (or residuals) is not constant across all levels of the independent variables. This violates one of the assumptions of ordinary least squares (OLS) regression, which assumes that the variance of the errors is constant (homoskedasticity).

Heteroskedasticity can lead to inefficient estimates of the regression coefficients. Standard errors and hypothesis tests based on these standard errors can become unreliable. If heteroskedasticity is present in your regression analysis, it can affect the validity of statistical inferences drawn from the model.

A diagram of a graph

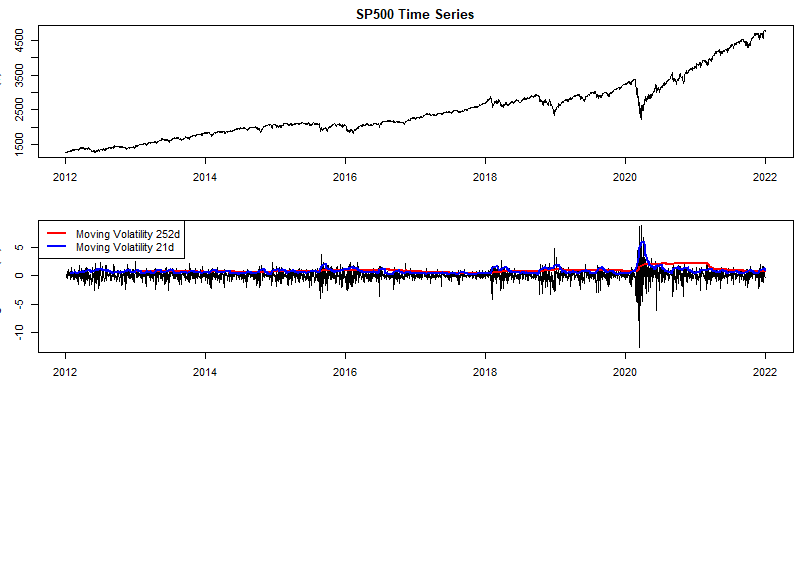
Description automatically generated

**Volatility**

Volatility is a statistical measure of asset return dispersion across time used in finance. The standard deviation or variance of price returns is frequently used to calculate it. Volatility is used to characterize the uncertainty surrounding the price movement of financial assets. It’s a crucial notion that is used in risk management, portfolio optimization, and other areas. It’s also one of the most active fields of empirical finance and time series analysis study. The risk and volatility of assets are directly proportional to each other which means the riskier a financial asset is, the higher the volatility.

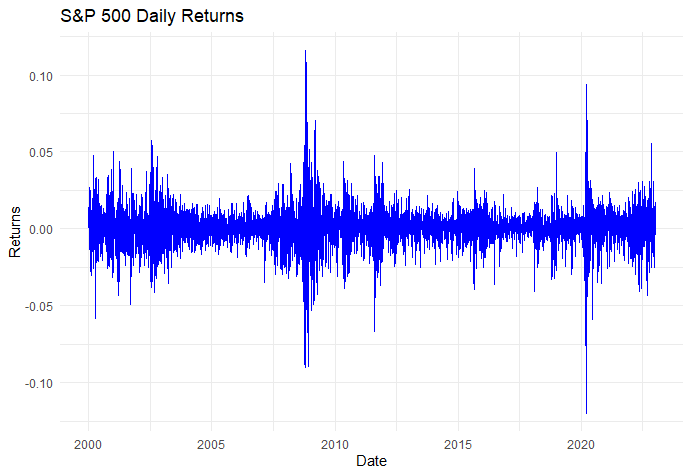
**Leverage-effect**

Volatility is a response to the news which must be a surprise. However, the timing of the news may not be a surprise and gives rise to predictable components of volatility such as economic announcements. Volatility tends to rise or fall in a predictable pattern over time. Black (1976) first noted that changes in stock returns display a tendency to be negatively correlated with changes in returns volatility, i.e., volatility tends to rise in response to “bad news” and to fall in response to “good news”. This phenomenon is termed the “leverage effect”. We can observe the phenomenon of “leverage effect” by plotting the market prices and their volatility as below:

****

**Volatility Clustering**

In time series modeling, it’s a frequent assumption that volatility remains constant throughout time. Looking at financial data suggests that some time periods are riskier than others; that is, the expected value of the magnitude of error terms at sometimes is greater than at others. Moreover, these risky times are not scattered randomly across quarterly or annual data. Instead, there is a degree of autocorrelation in the riskiness of financial returns. Financial analysts, looking at plots of daily returns such as in Figure 1, notice that the amplitude of the returns varies over time and describe this as “volatility clustering.”



Modeling and forecasting the volatility of financial assets, such as the S&P 500 index, is an important task in quantitative finance. Forecasting accuracy is generally enhanced with the addition of more data points. Unfortunately, nobody can create more financial data without a time machine handy. Unlike a clinical trial where you can sample another 1,000 people, you simply must wait for each interval (be it a tick, day, month, etc.) to pass. Hence, a prediction model, no matter how rudimentary or advanced, is necessary to digest the historical information and forecast the next day’s value. Given T trading days of data, the model will digest a certain amount of information, leading to a prediction value for the T+1 day.

**ARCH/GARCH Models**

The ARCH/GARCH models, which stand for autoregressive conditional heteroskedasticity and generalized autoregressive conditional heteroskedasticity, are designed to deal with just this set of issues. Instead of considering this as a problem to be corrected, ARCH and GARCH models treat heteroskedasticity as a variance to be modeled. As a result, not only are the deficiencies of least squares corrected, but a prediction is computed for the variance of each error term. Modeling time-varying volatility and correlation in financial time series is an important element in pricing equity, risk management, and portfolio management. Higher volatilities increase the risk of assets and higher correlations between assets cause an increased risk in portfolios. ARCH/GARCH models have been applied to model volatility with great success in capturing the heteroskedastic of financial time series and behavior such as the leverage effect and volatility clustering.

The Autoregressive Conditional Heteroskedasticity (ARCH) model was first introduced in the seminal paper of Engle (1982). Bollerslev (1986) generalized the ARCH model (GARCH) by modeling the conditional variance to depend on its lagged values as well as squared lagged values of disturbance.

**ARCH model**

In financial and economic models, conditioning is often stated as regressions of the future values of the variables based on the present and past values of the same variable. For example, if we assume that time is discrete, we can express conditioning of returns as an autoregressive (AR(n)) model:

The error term is conditional on the information that, in this example, is represented by the past n values of the returns.

The return can now be modeled as:

Where:

,

And the formulation of the ARCH model is:

where the coefficient must be estimated from empirical data. Thus, ARCH(p) is a forecasting model as it forecasts the error variance at time *t* based on the known information in the form of squared residuals of previous p times.

**GARCH Model:**

A useful generalization of this model is the GARCH parameterization introduced by Bollerslev (1986). This model is also a weighted average of past squared residuals, but it has declining weights that never go completely to zero. In its most general form, it is not a Markovian model, as all past errors contribute to forecast volatility. The most widely used GARCH specification asserts that the best predictor of the variance in the next period is a weighted average of

1) the long-run average variance,

2) the variance predicted for this period, ()

3) the new information in this period that is captured by the most recent squared residual ().

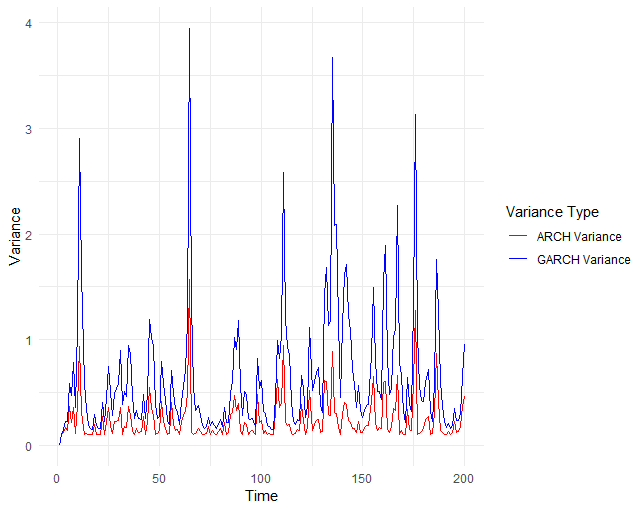
Such an updating rule is a simple description of adaptive or learning behavior from past information. Therefore, a GARCH (p,q) model for variance looks like this:

The econometrician must estimate the coefficients . Then, updating simply requires knowing the previous forecasted or actual q values of and p values of the residual .

The GARCH(p,q) model described above derives its name from the fact that the p,q in parentheses is a standard notation in which the first number refers to the number of autoregressive lags (or ARCH terms) that appear in the equation and the second number refers to the number of moving average lags specified (often called the number of GARCH terms). The process is quite straightforward: For any set of parameters and a starting estimate for the variance of the first observation, which is often taken to be the observed variance of the residuals, it is easy to calculate the variance forecast for the second observation. The GARCH updating formula takes the weighted average of the unconditional variance, the squared residual for the first observation, and the starting variance and estimates the variance of the second observation. This is input into the forecast of the third variance, and so forth. Eventually, an entire time series of variance forecasts is constructed.

**ARCH vs. GARCH**

Below is a simulation of ARCH(1) and GARCH(1,1) to illustrate the effect of including the GARCH term in volatility modeling. It can be observed in the figure below that the variance becomes more sensitive and accurate as we include the past information of variance on the model. At times of high volatility, the difference can be seen to be significant and hence the GARCH model captures the volatility more accurately.

****

**GARCH Model on Financial Data**

SP500 index is extracted for 10 years from *2012-01-01* to *2022-01-01*. Daily returns are calculated using the closing prices and the ADF (Augmented Dickey-Fuller) test is performed on the daily returns which give a p-value less than 0.01 (Appendix, result 1) and thus, we can conclude that the returns are stationary. Further Breusch-Pegan test for heteroskedasticity is performed and a p-value close to 0 was obtained(Appendix, result 2), concluding the presence of heteroskedasticity in the returns time series data. Further, to implement the GARCH model, the following variations are considered:

1. **Distribution Assumption**

The GARCH model makes distribution assumptions about the residuals and the mean return. However, financial time series data often does not follow a normal distribution as extreme positive and negative values are observed at times of shock. To be more representative of real financial data we can specify the model’s distribution assumption to be a Skewed student's t-distribution(sstd).

Originating from the original student-t distribution, Hansen (1994) proposed the skewed-t distribution. By adding the skewness parameter based on the student-t distribution, the skewed-t distribution can better describe the asymmetric and fat tail features of financial asset data as compared to the normal distribution. It is be defined by two parameters, degrees of freedom(shape) and skewness.

1. **Asymmetricity**

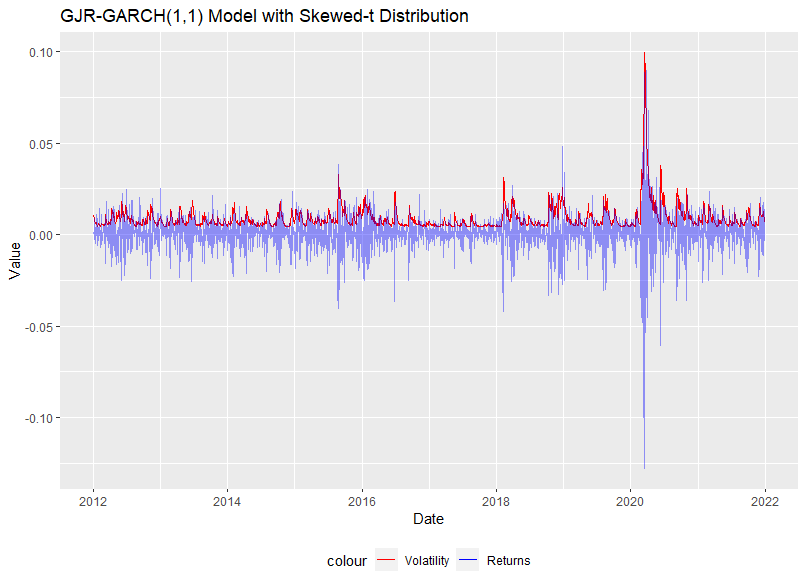
Glosten-Jagannathandan-Runkle (1993) introduced the asymmetric GARCH model, the GJR-GARCH model. The advantage of the GJR-GARCH model is that it can measure volatility due to the different effects of bad news and good news. The GJR-GARCH is defined as follows:

where ,

This is same as GARCH(p,q) model, the only difference is that a term has been added to account for negative returns. If the returns are negative, then the squared residuals are given more weight by an additional parameter and the coefficient associated with negative shock becomes . This is how GJR-GARCH model captures the volatility clustering.

GJR-GARCH model was fitted using skewed student-t distribution and with constant mean assumption, and the estimated parameters are shown in(Appendix, result 3), the estimated parameters can be used to write the following equation for volatility at time T+1, using the information at time T:

Using the above equation, volatility is forecasted and is plotted with respective returns as show below. The model can be observed to be highly accurate in forecasting the volatility as the red line moves in vicinity to the returns. The model accuracy will be tested using VaR and QPS later.

****

**GARCH Rolling Window Forecast**

* **Fixed rolling window forecast:** new data points are added while old ones are dropped from the sample.

**A graph of progress bar

Description automatically generated**

**Value at Risk**

ARCH/GARCH models play a key role in Value at Risk calculations. VaR is a statistical measure used to quantify the level of financial risk within a firm's investments or portfolio. Using ARCH/GARCH models, financial institutions can estimate the potential losses (or gains) at different confidence levels, providing valuable insights into the risks faced by their portfolios.

Three ingredients:

1. Portfolio value
2. Time horizon
3. Probability

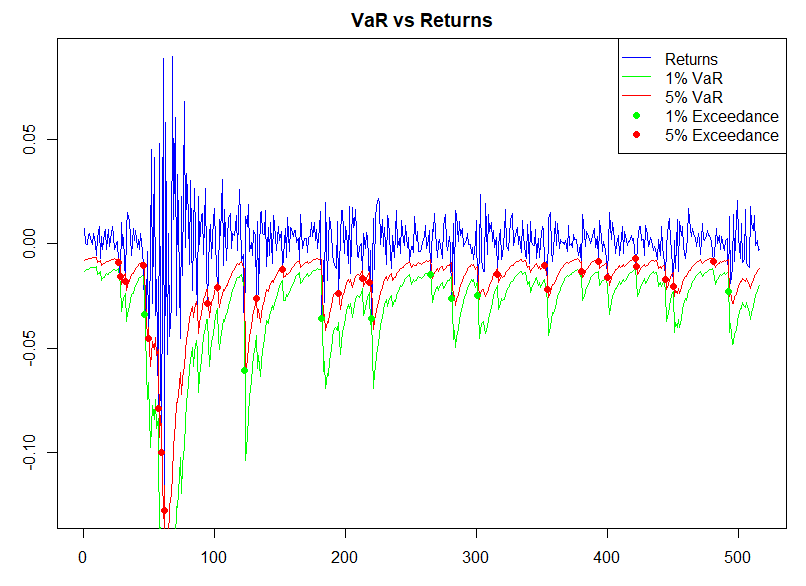
For example,1-day 1% VaR of $1 million means that there is 1% probability the portfolio will fall in value by 1 million dollars or more over a 1-day period. Stated the other way, there is 99% probability that the portfolio will not lose the calculated value or more in one day period. Exceedance points are the time points at which the returns are more negative than the calculated value of VaR.

**VaR in risk management terms**

Suppose a 5% daily VaR and 252 trading days in a year. A valued VaR model should have less than 13 VaR exceedance in a year, that is 5% \* 252 if there are more exceedances the model is underestimating the risk. The VaR back-testing report ((Appendix, result 4) shows that there are exceedance points 1.6% exceedance points at 1% level. The model is underestimating the risk a little bit.

**Dynamic VaR with GARCH**

* GARCH gives more realistic estimation of VaR.
* VaR = mean + (GARCH vol) \* quantile.



**Back-testing**

Quadratic Performance Score is used to measure the performance of VaR that has been estimated. If 𝑟𝑡 is the return at time 𝑡 and 𝑉a𝑅𝑡 is a VaR prediction at time 𝑡. The following size-adjusted frequency approach was introduced by Lopez in 1998:

The QPS statistic is calculated using the following equation:

Where n is the number of datapoints, 𝑝 is a probability value of VaR(0.01). The QPS value is between the [0,2] range with 0 being the minimum value that occurs when 𝑟𝑡 ≤ 𝑉a𝑅𝑡 for all values of t, and 2 is the maximum value that occurs when 𝑟𝑡 > 𝑉a𝑅𝑡 for all values of t. VaR performance is said to be good when small QPS approach 0.

**Results and Conclusion**

**Appendix**

**Result 1:**

A screenshot of a computer

Description automatically generated

**Result 2:**

A white background with black text

Description automatically generated

**Result 3:**

**A screenshot of a computer

Description automatically generated**

**Result 4:**

A white paper with black text and numbers

Description automatically generated

**Limitations**

when using GARCH models for financial analysis, especially in scenarios where asymmetric responses, non-symmetric distributions, and the interpretation of volatility shocks are critical considerations, researchers in the field of finance need to be aware of the following limitations:

if we were able to make conditional forecasts of returns, then the ARCH model describes the behavior of the errors and it is no longer true that the unconditional variance of errors coincides with the unconditional variance of returns. Thus, the statement that ARCH models describe the time evolution of the variance of returns is true only if returns have a constant expectation.

N**eglect of Volatility Asymmetry**: GARCH models assume that only the magnitude of unanticipated excess returns influences volatility, not their positivity or negativity. However, empirical evidence suggests that volatility tends to rise more sharply in response to bad news (negative excess returns) than to good news (positive excess returns). This suggests a limitation of GARCH models in capturing asymmetric responses to positive and negative shocks.

**Symmetric Distribution Assumption**: If the distribution of the error term zt is symmetric, the change in variance is uncorrelated with excess returns. This means that GARCH models assume that volatility changes are not influenced by the direction of excess returns (positive or negative). This assumption might not align with the observed behavior of financial markets.

**Nonnegativity Constraints**: GARCH models include constraints to ensure that the conditional variance remains nonnegative. While this ensures stability in the model, it limits the possibility of capturing certain types of random behavior in the variance process. Additionally, these constraints can complicate the estimation process.

**Difficulty in Estimation**: The nonnegativity constraints can lead to difficulties in estimating GARCH models. Researchers have had to impose specific structures on model coefficients to prevent them from becoming negative, indicating potential challenges in the estimation process.

**Interpretation of Volatility Shocks**: GARCH models are used to analyze the persistence of shocks to conditional variance. The duration for which these shocks persist is crucial in understanding their impact on asset pricing and investment decisions. Long-lasting volatility shocks can significantly influence risk premia and long-term investment strategies.

References:

1. <https://web-static.stern.nyu.edu/rengle/GARCH101.PDF>
2. <https://www.jstor.org/stable/4479615>
3. <https://www.sciencedirect.com/science/article/abs/pii/0304407686900631>
4. <https://www.tandfonline.com/doi/abs/10.1080/16843703.2004.11673078>
5. <https://pages.stern.nyu.edu/~rengle/ARCHGARCH.pdf>
6. <https://www.ripublication.com/gjpam16/gjpamv12n2_03.pdf>
7. <https://goldinlocks.github.io/ARCH_GARCH-Volatility-Forecasting/#GARCH-rolling-window-forecast>
8. <https://rpubs.com/ionaskel/VaR_Garch_market_risk>
9. <https://saltfinancial.com/insights/blog/stack-the-deck-with-rolling-realized-volatility/>
10. <https://ieomsociety.org/proceedings/2021indonesia/701.pdf>
11. <https://www.researchgate.net/publication/334766450_Estimation_of_Value-at-Risk_Adjusted_under_the_Capital_Asset_Pricing_Model_Based_on_ARMAX-GARCH_Approach>
12. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1029673>

Appendix: