THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF MATHEMATICS

PREDICTING VOLATILITY OF STOCKS USING TIME SERIES MODELS

ZITAO SONG
SPRING 2022

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Mathematics
with honors in Mathematics

Reviewed and approved* by the following:

Mauricio Nascimento
Assistant Teaching Professor of Statistics
Thesis Supervisor

Sergei Tabachnikov
Professor of Mathematics
Honors Adviser

* Electronic approvals are on file.

ABSTRACT

The volatility of an individual stock plays an important role in the pricing and risk management of various financial instruments. Volatility does not behave like stock price, which is close to a random walk, on the contrary, volatility has its own mathematical and statistical properties. In this paper, I modeled and forecasted the volatility of selected US stocks by using Generalized Autoregressive Heteroscedasticity (GRACH) and the variations of the GARCH model. I showed the fitting result, prediction, and loss function of each model and concluded their strengths and weaknesses.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. Nascimento for his inspiring and patient supervision as well as guidance throughout the research. I would also like to thank Dr. Tabachnikov for his thoughtful advising. In addition, I am grateful for everything that my family has done for me. Lastly, I will not able to complete this thesis without the help and support from everyone that I mentioned, so thank you all!

Chapter 1

Introduction

Volatility in the financial market is a statistical measurement of the dispersion of returns for a given asset in a given time interval. It indicates the expected range of return movement. Volatility is a part of risk which, by definition, is the degree of uncertainty inherent in the investment decision. In most cases, high volatility will increase risk in finance. In a volatile market, an asset's price tends to rise or fall significantly in a short period. Dramatic fluctuation brings risk to investors. Forecasting volatility can reduce the impact of this uncertainty in price change by analyzing it in investment decisions, that is, risk management. If the model gives a good prediction of volatility, the investors can adjust or rebalance their portfolios to reduce risk using these metrics.

Volatility is more important in the short term due to its clustering effect which is a tendency of large changes in returns to follow large changes, and small changes in returns to follow small changes. In statistics, volatility clustering can refer to heteroskedasticity. This statistical property enables the forecasting of volatility. Knowing expected volatility beforehand can also contribute to trading opportunities. For example, if the model gives a continuous volatility trend in the next few days, then investors can construct option strategies to gain profits.

Volatility is also a key factor to option pricing. In the Geometric Brownian motion, the volatility of the stock price is the variance rate. The Brownian motion follows Ito's Lemma; thus, the value of an option can be expressed as a differential equation. The Black-Sholes model is a solution to this equation. Theoretically, the volatility parameter in the Black-Sholes formulas refers to historical volatility which is the standard deviation of the percentage change in the stock price in a time interval.

Common types of volatilities are market volatility, stock volatility, historical volatility, and implied volatility. Market volatility measures the overall dispersion of value changes including commodities, forex, and the stock market. Stock volatility often referred to the systematic risk or beta, is a measure of the correlation between a stock's volatility and that of the whole market. Implied volatility is the future volatility of a stock assessed by the market. Historical volatility or realized volatility describes how much volatility a stock has had in a period. It can be obtained by calculating the standard deviation of returns for a stock in a given time period. Forecasting future realized volatility is the main objective of this thesis.

Time series is a collection of observations of data items obtained through repeated measurement over time. Similar to other analyses, time series can also be modeled. Normally, a time series model contains two parts: deterministic trend and stochastic trend. The deterministic part, also called signal, can be modeled using standard modeling techniques like linear regression. The main purpose is to remove the deterministic trend and seasonal trend contained in the time series data. After removing these trends, the stochastic part or noise is left. This part can be modeled by fitting the times series model, such as ARMA or ARIMA, with the assumption of stationarity. Moving average and autoregression are two main processes of the stochastic part. However, if the stochastic trend is still not stationary which implies that the variance over time is not constant, then the data has heteroskedasticity and it can be explained by conditional heteroskedasticity using the ARCH or GARCH model. ARCH model is a class of models used to model changing variance of a time series.

Chapter 2

Methodology

Autoregressive (AR) and Moving average (MA) are two basic models used to analyze time series data. The MA(q) model, moving average of order q, specifies that the response variable is written as the linear combination of the current and previous error terms and can be written as:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

where Y_t is the observed time series point, ε_t is the error term assumed to be independent and $N(0, \sigma_\varepsilon^2)$, also called the white noise, and θ_q is the q-th model parameter.

AR(p) models, autoregressive of order p, specify that the response variable is written as the linear combination of the previous p response and the current error. This model can be written as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

where Y_{t-i} is the observed response i lags before the current time t, ε_t is the error term, which is assumed to be white noise, and ϕ_p is the p-th parameter.

Autoregressive Moving Average of order p and q, or ARMA (p, q) model is a combination of an AR(p) and MA(q) processes. This means the observed time series is written as a linear combination of the p previous observations, the current error, and the q previous errors. This model can be written as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

The order of the models can be determined by the autocovariance function (ACF), the partial autocovariance function (PACF), and the EEEEE autocovariance function (EACF). The

ACF measures the correlation of the time series with observation r steps before. ACF of Y_t at a lagged value k is denoted as:

$$r_k = \frac{\sum_{t=k+1}^{n}\left(Y_t - \overline{Y}\right)(Y_{t-k} - \overline{Y})}{\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)^2}$$

The PACF measures the correlation of time series with its lagged values with the control of shorter lag length. PACF of Y_t at lag k is $Cor(Y_t, Y_{t-k}|Y_{t-1}, \dots, Y_{t-k+1})$. In other words, PACF shows the more exact correlation between the current data point and its lagged point. Thus, PACF is used for suggesting the order of the autoregressive part and it decreases exponentially to 0 for the MA model.

Normally, if the PACF graph shows cuts off after a lagged value k, then the order of the autoregressive part or p would be k. Similarly, if the ACF graph shows cuts off after a lagged value k, then the order of the moving average part would be k. If there are no obvious cuts off in ACF and PACF graph, then we consider the time series to follow an ARMA(p, q), and EACF can be used for suggesting orders for the model.

The time series model can be written into two parts: deterministic and stochastic. Let Y_t be a time series, and it can be written as:

$$Y_t = D_t + W_t$$

where D_t is the deterministic trend and W_t is the stochastic trend. The deterministic part removes any deterministic trend or seasonality contained in the data. Techniques like regression are normally used to fit this part. For example, if the deterministic part shows a constant increase or decrease, D_t can be modeled using standard linear regression, and it can be written as:

$$D_t = \beta_0 + \beta_1 t$$

After fitting the deterministic part, the stochastic part or residuals can be obtained. In most cases, these residuals are expected to meet a few assumptions: normality, constant variance, and independence. When the assumption of independence is not met, the stochastic part can be modeled using time series models. For example, the stochastic part can be modeled using the ARMA model, and it can be written as:

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

Residual analysis is also required after the stochastic fitting; the assumptions are normal distributed, mean of 0, constant variance, and independence needed to be satisfied. If the residual analysis is failed, it is either because the selection of the model went wrong, or the stochastic part is not stationary, and it is caused by non-constant variance and dependence. GARCH model now can be introduced to explain heteroskedasticity.

Generalized Autoregressive Conditional Heteroskedasticity, or GARCH, is an extended version of the ARCH model. GARCH has two components: moving average and autoregressive. Each component has an order. The order of the moving average part is denoted by q, and the order of the autoregressive part is denoted by p. GARCH (p, q) would be a p order and q order, model, and it can be denoted as:

$$r_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

where r_t is the stochastic part, or normally residuals from the mean function, σ_t^2 is the conditional variance, $\varepsilon_t \sim i.i.d. N(0, 1)$ is the error term, and ω is a constant. Furthermore, the order of the moving average part, or q, represents the number of lags in squared residuals in the

GARCH model, and the order of the autoregressive part, or p, is the number of lags in squared conditional variance in the GARCH model. Similar approaches like ACF and PACF are used for determining orders in the GARCH model. After the fitting, residuals are supposed to be normally distributed with a mean of 0 and a variance of 1, and independent of each other.

There are many known variations of the GARCH model. For example, EGARCH, TGARCH, and GJR-GARCH. However, the methodology I applied mainly focuses on GARCH and its simple variations. Four models are been used: ARMA (p, q) + GARCH (p, q), Weighted GARCH (p, q), GARCH (p, q) with an exogenous variabel, Seasonal GARCH (p, q).

ARMA (1, 1) + GARCH (1, 1) can be written as:

$$Y_t = \phi_1 Y_{t-1} + r_t - \theta_1 r_{t-1}$$

$$r_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

where Y_t is the return for a stock at time t, r_t is the residual from ARMA(1,1) at time t, and σ_t^2 is the conditional variance, that is, the volatility of a stock at time t.

Weighted GARCH (1, 1) can be written as:

$$r_t = \frac{1}{w_t} \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

where w_t is a weighted factor at time t. In this case, I chose the trading volume of a stock as the weighted factor. For a good amount of stocks, the trading volume is correlated with volatility. Weighing returns by a correlated factor, which is trading volume, can emphasize the variances of return so that the volatility would have a better capture of volatility when it is significantly changing. Normally, a weighted factor is in percentage, so I applied normalization by dividing the mean of the trading volume in a selected period.

GARCH (1, 1) with a exogenous variable can be denoted as:

$$r_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \beta_1 x_{t-1}$$

where x_{t-1} is an external variable. In this model, I also applied the trading volume, which means x_{t-1} is the daily trading volume of a stock at time t-1. Since the exogenous variable is independent of the conditional variance, I used the ARIMA model to forecast. Instead of weighing returns, the trading volume is directly modeled into the volatility as an external variable in the GRACH model.

Seasonality is another potential cofactor of the volatility in a relatively short time window. Seasonal GARCH (1, 1) with the cosine trends can be written as:

$$r_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \beta_1 \cos(2\pi f t) + \beta_2 \sin(2\pi f t)$$

where f is the frequency, and β_1, β_2 are amplitudes.

In general, following steps are incorporated into my methodology: getting the latest 500 day's closing prices of the stock; data preprocessing to make sure that there is no any missing values or outliers, also the volatility is a measure of dispersion of returns, so I take the logarithm difference of the data; splitting the data into train and test sets which I use the last 5 days as the test sets since my prediction would be in the range of next few days; calculating the historical volatility to be the comparison or actual volatility; first removing the deterministic part if there is any trend or seasonality contained in the time series; residual analysis to see if the data is satisfied to fit the ARMA model; checking autocorrelation using ACF, PACF, and EACF to determine orders of the model; fitting the ARMA model; residual analysis which is likely to be failed due to the clustering effect of volatility; checking GARCH effect with residuals and squared residuals from ARMA model using ACF, PACF, and Ljung-Box test; using squared residuals is because residuals now have a mean of 0 which implies that the variance is just simply squared residuals, so by looking at the ACF and PACF of squared residuals, I can have a better judgement of the autocorrelation of variance, and Ljung-Box test is for testing independence between squared residuals; fitting GARCH model; checking the summary of model, and make sure that p-values of parameters are significant and residuals pass the test; show the fitting result, such as the conditional variance graph, the entire model with parameters, and AIC/BIC or loss function; making predictions and compare with the test set of historical volatility.

Chapter 3

Data Analysis

Tesla (TSLA), Apple (AAPL), and Disney (DIS) are stocks of interest. The stock price is close to a random walk which is unlikely to be modeled, so I focused on the dispersion of returns. In the following graph, figure 1 is the daily returns of Tesla stock in 500 days.

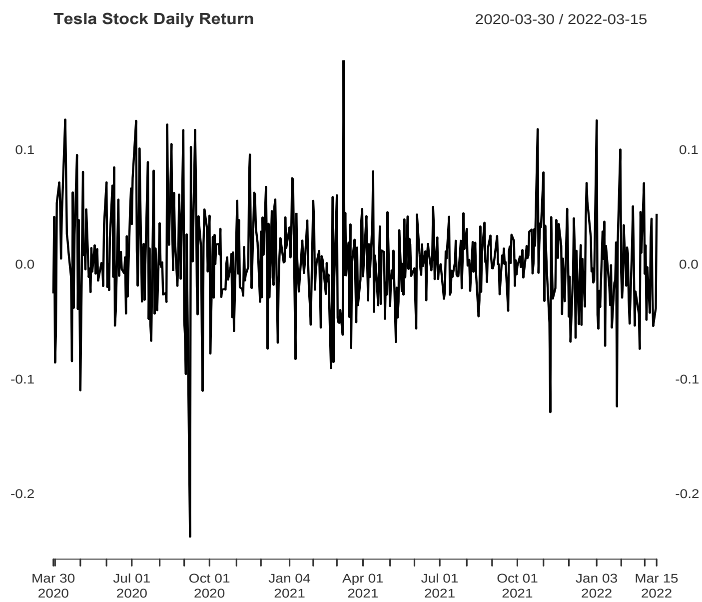

Figure 1: Tesla Stock Daily Returns

There is no obvious deterministic trend, and I checked ACF and PACF where both graphs tail off. This is an indication of the ARMA model, and EACF suggested using ARMA (1, 1). After fitting the model, the residual analysis failed as expected. The Ljung-Box test showed squared residuals are autocorrelated, that is, the GARCH effect. By looking at the ACF and PACF of squared residuals, I chose ARMA (2, 2) + GARCH (1, 1) to be my model. In the model summary, each estimated parameter is significant, and the residual analysis is passed. Figure 2 is the conditional SD, or fitted volatility, and the historical volatility, or actual volatility.

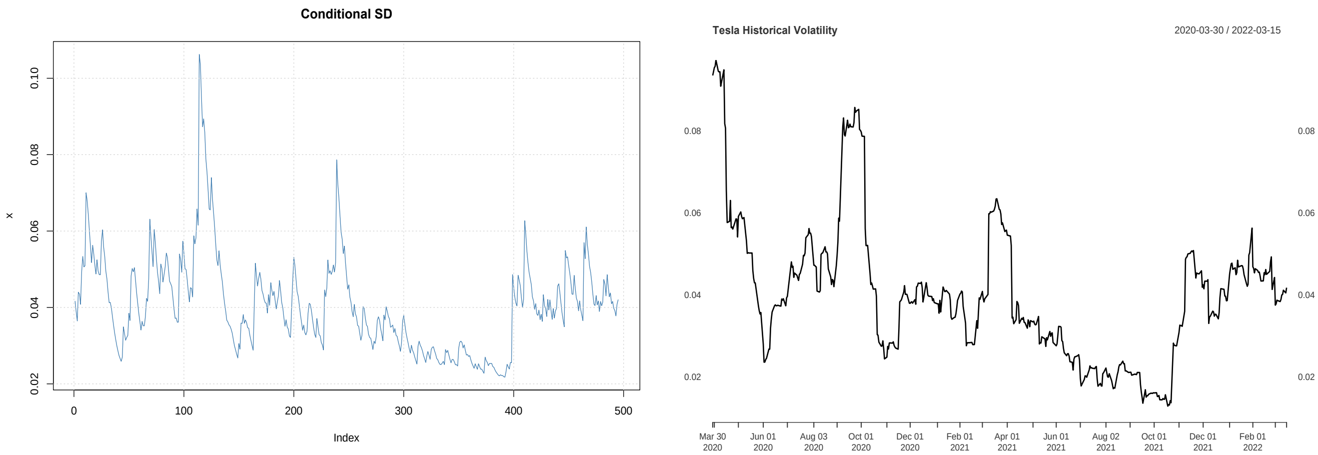

Figure 2: Tesla Conditional SD (GARCH) vs. Tesla Historical Volatility

Besides the beginning, ARMA (2, 2) + GARCH (1, 1) model showed a good fitting result. Next, I predicted the volatility for the next consecutive 5 days and compared them to the test actual volatility. MSE is 6.1×10^{-7}. For Apple and Disney stock, I used ARMA (2, 2) + GARCH (1, 1) and ARMA (5, 5) + GARCH (1, 1), respectively. MSE for Apple stock is 6.35×10^{-8}, and that for Disney stock is 6.1×10^{-6}. A Summary of every model can be found in the appendix.

Furthermore, I added an exogenous variable, the daily trading volume, into ARMA (p, q) + GARCH (p, q) model. I checked the correlation between historical volatility at time t and the trading volume at time t-1 for Apple stock first. The correlation coefficient is 0.53. Similarly, by looking at the ACF, PACF, and the optimal parameters table. I chose the mean model of ARMA (2, 2), and the variance model of GARCH (2, 2) with an external variable. The following graph is the conditional standard deviation and the actual volatility. In addition, the Ljung-Box test supported that the model does not show a lack of fit. In order to forecast future volatility, the trading volume needed to be modeled as well. I chose ARIMA (2, 1, 2) to be the model. The result passed the residual test. Then I forecast the volatility, and the MSE for the future 5 days is 1.2×10^{-5}. I repeated the same procedures for Disney and Tesla. I chose ARMA (2, 2) + GARCH (1, 1) for Tesla stock, and MSE is 6.9×10^{-4}. Disney stock did not show any significance on the external variable since its trading volume and its historical volatility are weakly correlated.

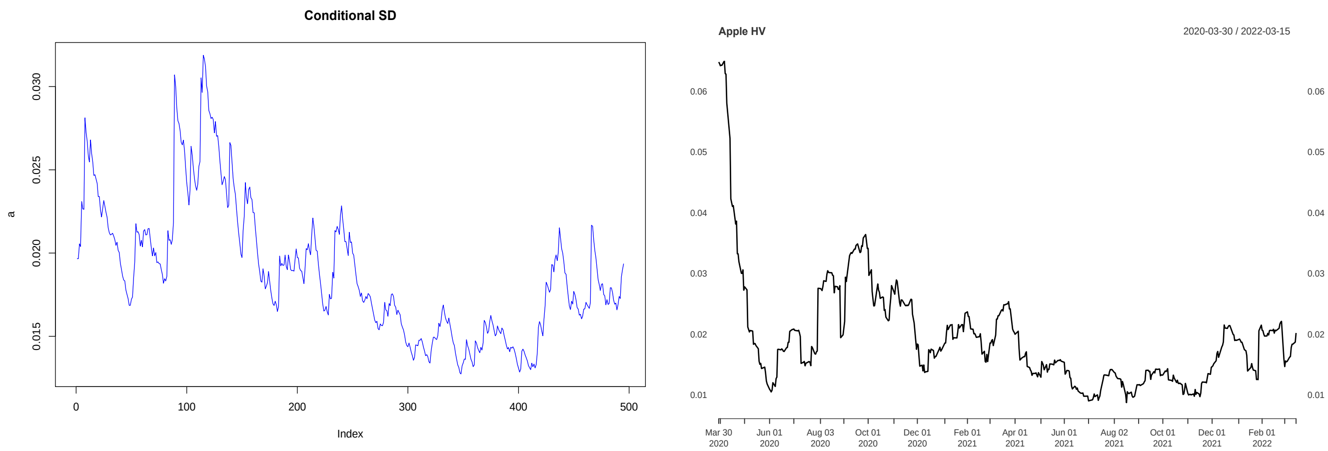

Figure 3: Apple Conditional SD (GARCH with covariant) vs. Apple Historical Volatility

Now I narrowed my data size down to the latest 200 days. GARCH+ARMA model still worked fine with Tesla and Apple stock. However, the fitting result for Disney stock is not ideal. I chose ARMA (1,1) + GARCH (1,1), and the following graph is the conditional standard deviation and the actual volatility

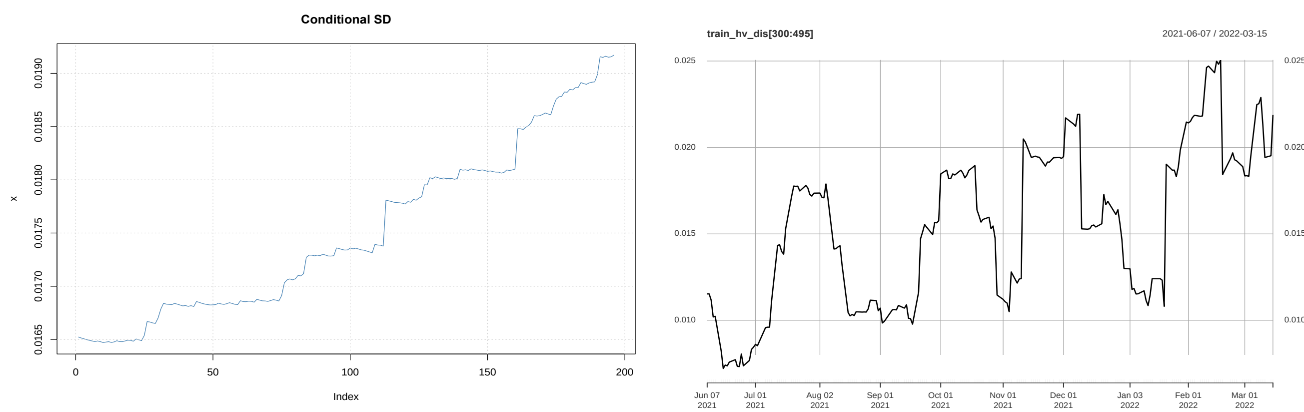

Figure 4: Disney Conditional SD (GARCH) vs. Disney Historical Volatility

There is an obvious seasonality trend contained in the historical volatility, and it caused the poor performance of the GARCH model. In this case, I decided to add harmonic series into the GARCH model to fit the data. The harmonic series is more straightforward and accessible compared to the seasonal GARCH. However, I found some shortages of this model. For example, the frequency and the period of data will greatly affect the fitting result. By observing and analyzing the data, I selected a frequency of 42 days and a period of 125 days. The following graph is the conditional standard deviation from ARMA (1, 1) + GARCH (1, 1) with the harmonic series and the historical volatility of Disney stock.

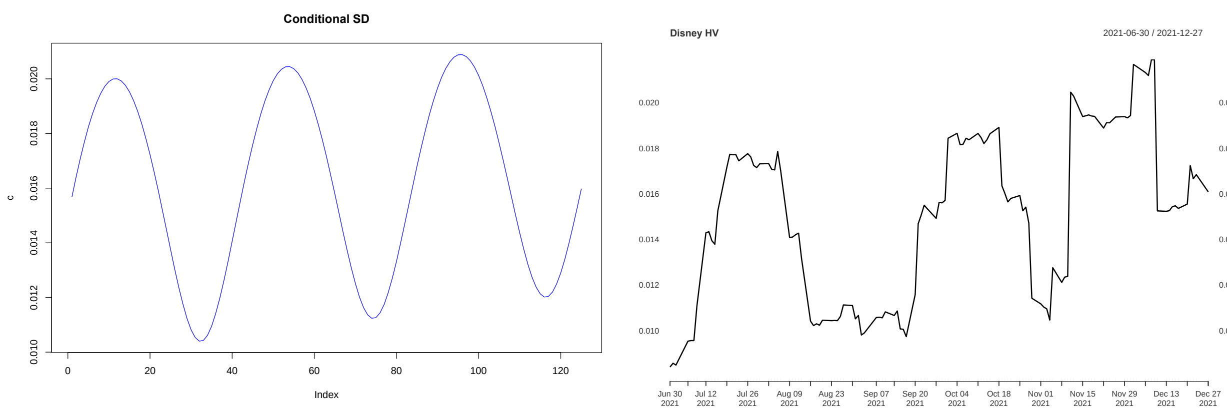

Figure 5: Disney Conditional SD (GARCH with harmonic series) vs. Disney Historical Volatility

Furthermore, I forecasted the conditional standard deviation for the next 5 days, and the MSE is 4.132×10^{-6} which is slightly smaller than the MSE of the ARMA (1,1) + GARCH (1,1) model.

The last model is the weighted GARCH. I weighed each day's return by the trading volume of the day before. Similarly, I used ARIMA (2,1,2) to predict the weighted factor. The variance model that I selected is GARCH (1, 1). In addition, I added the plot of conditional SD from an unweighted GARCH model in the same period. Figure 6 showed the historical volatility, conditional SD from GARCH, and conditional SD from weighted GARCH.

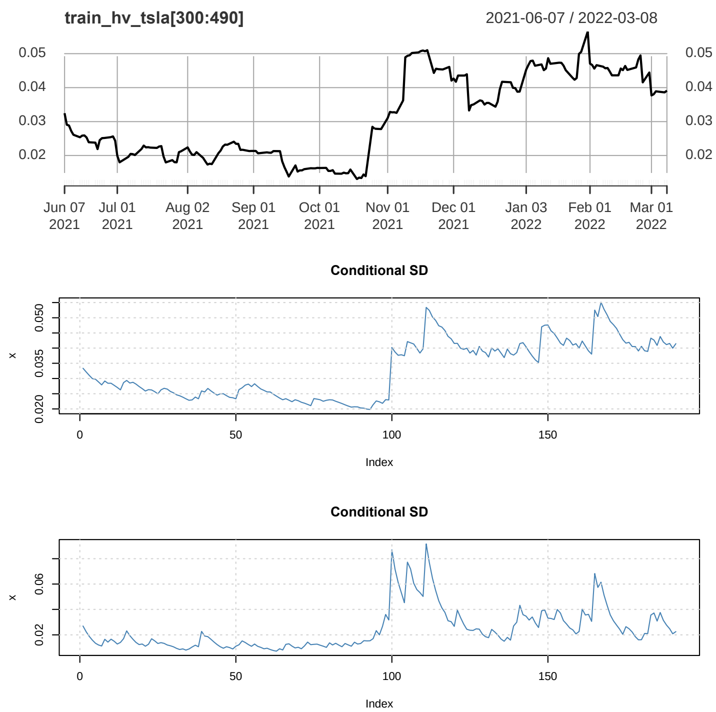

Figure 6: Tesla Historical Volatility vs. Conditional SD(GARCH) vs. Conditional SD (weighted GARCH)

Chapter 4

Conclusion

ARMA + GARCH model generally works great in most cases, it has the smallest loss function among all the models based on my previous data analysis. Also, its fitted conditional standard has the closest pattern to the actual volatility or historical volatility. On the other hand, GARCH with an external variable model resulted in a slightly greater MSE than the GARCH model, but AIC/BIC is less. An external variable will help to explain the conditional variance so that it is likely to reduce the number of parameters needed in the model. These two models both work when the data size is sufficiently large and the interested stock's price does not have a special pattern. If, for example, a stock's movement is seasonal in a short period like Disney, GARCH would not be the ideal model to fit that seasonality. Instead, GARCH with harmonic series would be a better choice of model. The seasonality of the volatility can be modeled by this model. However, it is restricted by the frequency and the period of data. To predict future volatility, the ARMA + GARCH is still recommended. Lastly, weighted GARCH focuses on capturing the degrees of volatility. Its reaction to the change of volatility is the most sensitive one.

Furthermore, since ARMA + GRACH and GACH with an exogenous model showed good predictions, we can use the future volatility for asset pricing, asset allocation, or risk management. GARCH with the harmonic series worked in certain periods and frequency, but it showed a good prediction in that period to the seasonality of volatility. It can be used for strategy testing or back testing. Weighted GARCH measured degrees of volatility, we can use it to make guidelines or thresholds in the risk profile.

Appendix A

Summary of Models

Table 1: ARMA (2,2) + GARCH (1,1) for Apple

```
Coefficient(s):
mu          ar1          ar2          ma1          ma2          omega        alpha1       beta1
1.9608e-03  7.8638e-01  -8.7659e-01  -7.8863e-01  8.4947e-01   4.6139e-06   4.6782e-02   9.4167e-01

Error Analysis:
        Estimate   Std. Error  t value Pr(>|t|)
mu      1.961e-03  8.497e-04    2.308  0.0210 *
ar1     7.864e-01  7.555e-02   10.409  <2e-16 ***
ar2    -8.766e-01  6.311e-02  -13.889  <2e-16 ***
ma1    -7.886e-01  9.183e-02   -8.588  <2e-16 ***
ma2     8.495e-01  6.849e-02   12.403  <2e-16 ***
omega   4.614e-06  3.523e-06    1.310  0.1903
alpha1  4.678e-02  1.781e-02    2.627  0.0086 **
beta1   9.417e-01  2.229e-02   42.252  <2e-16 ***
---

Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  81.91877  0
 Shapiro-Wilk Test  R    W      0.9825707 1.179216e-05
 Ljung-Box Test     R    Q(10)  4.653007  0.9131207
 Ljung-Box Test     R    Q(15)  7.395952  0.9457171
 Ljung-Box Test     R    Q(20)  18.25339  0.5707206
 Ljung-Box Test     R^2  Q(10)  7.347119  0.6923185
 Ljung-Box Test     R^2  Q(15)  9.492555  0.850389
 Ljung-Box Test     R^2  Q(20)  11.45684  0.9335044
 LM Arch Test       R    TR^2   8.602928  0.7364191

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
-5.088332 -5.020379 -5.088843 -5.061656
```

Table 2: ARMA (2, 2) + GARCH (1,1) for Tesla

```
Coefficient(s):
mu          ar1          ar2          ma1          ma2          omega        alpha1       beta1
1.0382e-02  -1.0000e+00  -9.4082e-01  9.7852e-01  8.9920e-01   6.5998e-05   1.3545e-01   8.3483e-01

Error Analysis:
        Estimate   Std. Error  t value Pr(>|t|)
mu      1.038e-02  4.445e-03    2.336  0.01951 *
```

ar1	-1.000e+00	3.694e-02	-27.072	< 2e-16	***
ar2	-9.408e-01	3.612e-02	-26.044	< 2e-16	***
ma1	9.785e-01	5.738e-02	17.052	< 2e-16	***
ma2	8.992e-01	4.227e-02	21.274	< 2e-16	***
omega	6.600e-05	3.693e-05	1.787	0.07392	.
alpha1	1.355e-01	4.329e-02	3.129	0.00175	**
beta1	8.348e-01	5.167e-02	16.157	< 2e-16	***

Standardised Residuals Tests:

			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	68.64156	1.221245e-15
Shapiro-Wilk Test	R	W	0.9798784	2.403142e-06
Ljung-Box Test	R	Q(10)	7.094817	0.7164643
Ljung-Box Test	R	Q(15)	10.00274	0.8195672
Ljung-Box Test	R	Q(20)	11.19066	0.9411273
Ljung-Box Test	R^2	Q(10)	4.803197	0.9039308
Ljung-Box Test	R^2	Q(15)	16.7801	0.3321789
Ljung-Box Test	R^2	Q(20)	22.0032	0.340337
LM Arch Test	R	TR^2	16.41942	0.172772

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-3.614844	-3.546892	-3.615356	-3.588168

Table 3: ARMA (5,5) + GARCH (1,1) for Disney

Coefficient(s):

mu	ar1	ar2	ar3	ar4	ar5	ma1	ma2
-3.1171e-04	-2.2194e-01	-1.4788e-01	-2.6414e-01	-9.4077e-01	-4.5330e-02	1.2656e-01	2.1288e-01

2.8564e-01

ma3	ma4	ma5	omega	alpha1	beta1
2.8564e-01	9.9999e-01	1.9244e-02	6.8192e-06	2.4682e-02	9.5649e-01

Error Analysis:

	Estimate	Std. Error	t value	Pr(>|t|)	
mu	-3.117e-04	6.274e-07	-496.851	<2e-16	***
ar1	-2.219e-01	2.629e-05	-8441.228	<2e-16	***
ar2	-1.479e-01	2.648e-05	-5584.442	<2e-16	***
ar3	-2.641e-01	2.580e-05	-10239.202	<2e-16	***
ar4	-9.408e-01	2.575e-05	-36537.282	<2e-16	***
ar5	-4.533e-02	2.644e-05	-1714.357	<2e-16	***
ma1	1.266e-01	2.888e-05	4382.684	<2e-16	***
ma2	2.129e-01	2.836e-05	7505.663	<2e-16	***
ma3	2.856e-01	2.785e-05	10254.653	<2e-16	***
ma4	1.000e+00	2.857e-05	34997.434	<2e-16	***
ma5	1.924e-02	2.906e-05	662.166	<2e-16	***
omega	6.819e-06	3.460e-06	1.971	0.0487	*
alpha1	2.468e-02	1.088e-02	2.269	0.0233	*
beta1	9.565e-01	1.459e-02	65.558	<2e-16	***

Standardised Residuals Tests:

			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	917.1423	0
Shapiro-Wilk Test	R	W	0.9272821	9.288336e-15
Ljung-Box Test	R	Q(10)	11.14886	0.3460383
Ljung-Box Test	R	Q(15)	22.13637	0.1042778
Ljung-Box Test	R	Q(20)	29.81281	0.07294411
Ljung-Box Test	R^2	Q(10)	1.833093	0.9974667
Ljung-Box Test	R^2	Q(15)	2.199364	0.9999445
Ljung-Box Test	R^2	Q(20)	3.069201	0.999995
LM Arch Test	R	TR^2	2.0453	0.9993332

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-5.094989	-4.976072	-5.096531	-5.048306

Table 4: ARFIMA (2,0, 2) + GARCH (2,2) with an exogenous variable for Apple

Optimal Parameters

	Estimate	Std. Error	t value	Pr(>|t|)
ar1	-1.760579	0.010495	-167.754761	0.000000
ar2	-0.989982	0.005348	-185.101111	0.000000
ma1	1.752076	0.005872	298.403312	0.000000
ma2	0.994427	0.000464	2141.573972	0.000000
omega	0.000000	0.000002	0.000000	1.000000
alpha1	0.000369	0.031692	0.011634	0.990718
alpha2	0.043109	0.031321	1.376344	0.168715
beta1	0.881731	0.492770	1.789334	0.073561
beta2	0.002406	0.487443	0.004936	0.996061
vxreg1	0.000026	0.000012	2.258409	0.023920

Information Criteria

Akaike	-5.0827
Bayes	-4.9978
Shibata	-5.0835
Hannan-Quinn	-5.0494

Weighted Ljung-Box Test on Standardized Residuals

----

	statistic	p-value
Lag[1]	0.02101	0.8848
Lag[2*(p+q)+(p+q)-1][11]	2.19548	1.0000
Lag[4*(p+q)+(p+q)-1][19]	5.06052	0.9935

d.o.f=4

Weighted Ljung-Box Test on Standardized Squared Residuals

----

	statistic	p-value
Lag[1]	0.1822	0.6695
Lag[2*(p+q)+(p+q)-1][11]	1.9483	0.9636
Lag[4*(p+q)+(p+q)-1][19]	3.7309	0.9846

Table 5: ARIMA (2,0,2) + GARCH (1,1) with an exogenous variable for Tesla

Optimal Parameters

	Estimate	Std. Error	t value	Pr(>|t|)
ar1	-1.141271	0.013215	-8.6361e+01	0.000000
ar2	-0.980115	0.010543	-9.2966e+01	0.000000
ma1	1.161558	0.012274	9.4635e+01	0.000000
ma2	0.988182	0.000001	1.7543e+06	0.000000
omega	0.000180	0.000153	1.1789e+00	0.238446
alpha1	0.000000	0.001627	0.0000e+00	1.000000
beta1	0.181974	0.222720	8.1705e-01	0.413899
vxreg1	0.001238	0.000363	3.4095e+00	0.000651

Information Criteria

Akaike	-3.6203
Bayes	-3.5523
Shibata	-3.6208
Hannan-Quinn	-3.5936

Weighted Ljung-Box Test on Standardized Residuals

----

	statistic	p-value
Lag[1]	0.0836	0.7725
Lag[2*(p+q)+(p+q)-1][11]	2.2416	1.0000
Lag[4*(p+q)+(p+q)-1][19]	3.6720	0.9998

d.o.f=4

H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals

----

	statistic	p-value
Lag[1]	0.07116	0.7897
Lag[2*(p+q)+(p+q)-1][5]	0.58663	0.9435
Lag[4*(p+q)+(p+q)-1][9]	2.58305	0.8253

Table 6: ARIMA (1,0,1) + GARCH (1,1) with the harmonic series for Disney

Optimal Parameters

	Estimate	Std. Error	t value	Pr(>|t|)
ar1	0.799472	0.112181	7.1267e+00	0.00000
ma1	-0.860209	0.076942	-1.1180e+01	0.00000
omega	0.000000	0.000002	2.5358e-01	0.79982
alpha1	0.000003	0.001265	2.0680e-03	0.99835
beta1	0.999992	0.000011	9.4045e+04	0.00000
vxreg1	0.000022	0.000001	1.7938e+01	0.00000
vxreg2	0.000002	0.000000	2.0777e+01	0.00000

```
Robust Standard Errors:
        Estimate  Std. Error     t value Pr(>|t|)
ar1     0.799472    0.095478  8.3733e+00 0.000000
ma1    -0.860209    0.060055 -1.4324e+01 0.000000
omega   0.000000    0.000007  6.4063e-02 0.948920
alpha1  0.000003    0.000631  4.1470e-03 0.996691
beta1   0.999992    0.000009  1.1039e+05 0.000000
vxreg1  0.000022    0.000005  4.7507e+00 0.000002
vxreg2  0.000002    0.000000  2.0554e+01 0.000000

LogLikelihood : 347.0563

Information Criteria
------------------------------------

Akaike       -5.4409
Bayes        -5.2825
Shibata      -5.4467
Hannan-Quinn -5.3766

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1853  0.6669
Lag[2*(p+q)+(p+q)-1][5]    2.7815  0.6116
Lag[4*(p+q)+(p+q)-1][9]    4.5342  0.5630
d.o.f=2
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                  4.603e-05  0.9946
Lag[2*(p+q)+(p+q)-1][5] 1.976e+00  0.6241
Lag[4*(p+q)+(p+q)-1][9] 3.368e+00  0.6971
```

Table 7: ARMA (2,2) + weighted GARCH (1,1) for Tesla

```
Coefficient(s):
        mu          ar1          ar2          ma1          ma2        omega       alpha1        beta1
1.7413e-03   4.9265e-01  -9.6356e-01  -4.5584e-01   1.0000e+00   4.2345e-05   1.0000e+00   4.4978e-01


Error Analysis:
         Estimate  Std. Error  t value Pr(>|t|)
mu      1.741e-03   1.434e-03    1.214 0.224665
ar1     4.926e-01   1.693e-02   29.098  < 2e-16 ***
ar2    -9.636e-01   1.611e-02  -59.823  < 2e-16 ***
ma1    -4.558e-01   6.069e-03  -75.114  < 2e-16 ***
ma2     1.000e+00   9.264e-03  107.940  < 2e-16 ***
omega   4.234e-05   2.071e-05    2.045 0.040876 *
```

alpha1 1.000e+00 3.318e-01 3.014 0.002577 **
beta1 4.498e-01 1.173e-01 3.835 0.000126 ***

Standardised Residuals Tests:

			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	835.635	0
Shapiro-Wilk Test	R	W	0.8811303	3.790614e-11
Ljung-Box Test	R	Q(10)	8.463701	0.5836395
Ljung-Box Test	R	Q(15)	17.01858	0.3177551
Ljung-Box Test	R	Q(20)	26.93699	0.137046
Ljung-Box Test	R^2	Q(10)	2.420636	0.9919825
Ljung-Box Test	R^2	Q(15)	7.961265	0.9253262
Ljung-Box Test	R^2	Q(20)	9.055669	0.9822535
LM Arch Test	R	TR^2	6.684935	0.8777129

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
-4.386613	-4.250392	-4.389937	-4.331437

BIBLIOGRAPHY

Bollerslev, Tim, 1986, Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.

Alford, A.W. and Boatsman, J.R. (1995). Predicting long-term stock return volatility: Implications for accounting and valuation of equity derivates. Accounting Review, 70, 4, 599-618.

Black, F. (1976). Studies of stock market volatility changes. Proceedings of the American Statistical Association, Business and Economic Statistics Section, 177–81.

Figlewski, S. (1997). Forecasting volatility, Financial Markets, Institutions and Instruments. New York University Salmon Center, 6, 1, 1-88.

Lopez, J.A. (2001). Evaluation of predictive accuracy of volatility models. Journal of Forecasting, 20(1).

Andersen, T. G., & Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review,39*(4), 885. doi:10.2307/2527343

Day, T. E., & Lewis, C. M. (1993). Forecasting Futures Market Volatility. *The Journal of Derivatives, 1*(2), 33-50. doi:10.3905/jod.1993.407876

Figlewski, S. (1997). Forecasting Volatility. *Financial Markets, Institutions and Instruments, 6*(1), 1-88. doi:10.1111/1468-0416.00009

Bollerslev T. 2009. ”Glossary to ARCH (GARCH),” In Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle, Boller- slev T, Russell JR, Watson M (eds). Oxford University Press: Oxford, UK; pp. 137164.

Engle, Robert F. 2001. ”GARCH 101: The Use of ARCH/GARCH Mod- els in Applied Econometrics,” Journal of Economic Perspectives, Vol. 15, No. 4, pp. 157-168.

Abraham, B. and Ledolter, J. (1983). *Statistical Methods for Forecasting*. New York: John Wiley & Sons.

Taylor, S. J. (1986). *Modeling Financial Time Series*. Chichester: John Wiley & Sons.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time Series Analysis, Forecast- ing and Control*, 2nd ed. New York: Prentice-Hall.

Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*, 2nd ed. New York: Springer.

Harvey, A. C. (1993). *Time Series Models*, 2nd ed. New York: Harvester Wheatsheaf.

Amadeo, Kimberly. “5 Types of Volatility.” *The Balance*, The Balance, 10 May 2021, https://www.thebalance.com/volatility-definition-and-types-3305968.

Hayes, Adam. “Volatility.” *Investopedia*, Investopedia, 8 Feb. 2022, https://www.investopedia.com/terms/v/volatility.asp.

ACADEMIC VITA

Zitao Song

EDUCATION

Pennsylvania State University—University Park

B.S. in Mathematics | Minor in Statistics and Economics *August 2018-May 2022*

Honors: Schreyer Honors College

Relevant Course: Financial Mathematics, Financial Modeling, Corporation Finance, Probability Theory, Numerical Analysis, Computational Statistics

Research department leader, International Students Investment Association | Treasure Department leader, Chinese Undergraduate Student Association | Member, Penn State Actuarial Science Club | Analyst, Penn State Investment Association

---

PROFESSIONAL EXPERIENCE

Moyi-Tech **New York, NY**

Data Analyst Intern *March 2020-June 2020*

- Conducted research on the mechanism of risk parity allocation and improved Python code of using risk parity for calculating targeted assets' weights based on historical market data.
- Wrote a couple of research on topics like "Negative crude oil price", "Quantitative Trading", and "Machine learning using technical indicators".
- Managed research and created correlation analyses and correlation heat maps based on S&P500 stocks' historical data. The company demonstrated my result in the internal seminar.
- Practiced a machine learning pipeline covering algorithms including Decision Trees and Logistic Regression on S&P 500's historical market data.

Wisdom Exchange Investment **Shenzhen, GD**

Buy-Side Quantitative Research Intern *November 2021-Present*

- My research presentations, quantitative analysis, and trading reports of listed American companies helped the company to make decisions.
- During institutional and buy-side meetings, I contributed ideas and answers that are based on circumspect researches.
- Participated in collating surveys, collecting data, and assisting in the management of investment community operations.

PROJECTS & ACTIVITIES

- Participated in a few course projects: "The convergence rate of the binomial option pricing model", "Statistical analysis of World Happiness Index", "Black-Scholes-Merton equation and the heat equation", and "Implementation of the quadratic sieve algorithm for factorization".
- Published around 30 articles on company's analysis and held lectures on quantitative strategies as well as trading tactics. Have over one million views and 6000 followers on a brokerage platform.

SKILLS & INTERESTS

Programming: Python, SQL, R, VBA, C++, MATLAB Languages: Chinese and English