THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF STATISTICS

TIME SERIES METHODOLOGIES TOWARD FINANCIAL DATA ANALYSIS

YUNJIE JIANG
SPRING 2013

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Statistics and Economics
with honors in Statistics

Reviewed and approved* by the following:

Zhibiao Zhao
Assistant Professor of Statistics
Thesis Supervisor

David Hunter
Professor of Statistics Department Head
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

The question of predicting future value based on historical data always motivates econometricians and investors in both academia and real industry. The volatility of financial data challenges thousands of investors from making the right decision while investing. Finding the right techniques to solve the problems is critical to investors and scholars.

Time series analysis, therefore, has been developed to diagnose the correlated relationships among variables. In this thesis, I will introduce some techniques of time series from a statistical perspective including MA model, AR model, ARIMA model, and ARCH/GARCH models that are widely used in econometrics and financial world. These statistical methods help us to understand the relationships between previous and future data.

To better understand the methodologies, I applied these time series models on weekly Dow Jones index and analyzed the characteristics and behavioral trends of it. The statistical coding package, R, was used as a tool during my analysis. At the end, discussion and conclusion will be driven to summarize the application of the techniques as well as some important findings of the Dow Jones Index.

TABLE OF CONTENTS

Abstract i

List of Figures iii

List of Tables iv

Acknowledgements v

Chapter 1: Method 1

- 1.1 Moving Average Model 2
- 1.2 Autoregressive Model 3
- 1.3 Autoregressive Integrated Moving Average Model 4
- 1.4 Autocorrealtion Function and Partial Autocorrealtion Function 5

Chapter 2: Application on Dow Jones Index 7

- 2.1 Data Overlook and Moving Average Smoothing 7
- 2.2 Cutting Points 10
- 2.3 Period 1 from 1950 to 1966 11
- 2.4 Period 2 from 1967 to 1982 13
- 2.5 Period 3 from 1983 to 1993 14
- 2.6 Period 4 from 1994 to 2000 17

Chapter 3: ARCH and GARCH Models 21

- 3.1 ARCH and GARCH Models 22
- 3.2 Period 1 from 1950 to 1996 23
- 3.3 Period 3 from 1983 to 1993 25
- 3.4 Period 4 from 1994 to 2000 27

Chapter 4: Conlusion and Discussion 29

Appendix A R Outputs for ARMIA Model Check for Period 1950-1966 30

Appendix B R Outputs for ARMIA Model Check for Period 1967-1993 31

Appendix C R Outputs for ARMIA Model Check for Period 1994-2000 32

Appendix D R Outputs for GARCH Models 33

REFERENCES 36

LIST OF FIGURES

Figure 2.1-1. General Time Series Plot on Average Weekly Price 8

Figure 2.1-2. Residual Plot with Filter 12.. 9

Figure 2.1-3. Residual Plot after Transformation. 9

Figure 2.2-1. Cutting Points of Logged Price. 10

Figure 2.3-1. Period 1 Time Series Plot and Residual Plot... 11

Figure 2.3-2. Model Diagnose for Period 1 12

Figure 2.4-1. Period 2 Time Series Plot and Residual Plot... 14

Figure 2.5-1. Period 3 Time Series Plot and Residual Plot... 15

Figure 2.5-2. Model Diagnose for Period 3 16

Figure 2.6-1. Period 2 Time Series Plot and Residual Plot... 18

Figure 2.6-2. Model Diagnose for Period 4 19

Figure 3-1. Volatile Finanical Data over Time 22

Figure 3.2-1. ACF amd PACF Plots 23

Figure 3.3-1: ACF and PACF Plots 25

Figure 3.4-1: ACF and PACF Plots 27

LIST OF TABLES

Table 2.3-1. Model Selection for Period 1 13

Table 2.4-1. Summary for Period 2 14

Table 2.5-1. Model Selection for Period 3. 17

Table 2.6-1: Model Selection for Period 4. 20

Table 3.2-1. Box-Pierce, Shapiro-Wilk, and KPSS 24

Table 3.2-2: Summary of GARCH (1,1) 24

Table 3.3-1: Box-Pierce, Shapiro-Wilk, and KPSS. 26

Table 3.2-2: Summary of GARCH (1,1) 26

Table 3.3-1: Box-Pierce, Shapiro-Wilk, and KPSS 28

Table 3.4-2: Summary of GARCH (1,1) 28

ACKNOWLEDGEMENTS

I would like to thank Dr. Zhibiao Zhao for all of his helps towards my study of time series during the writing period of this thesis and his guidance as my thesis advisor. Dr. David Hunter, being my honor's advisor, also generously provides me his supports of my study. I would also like to thank my dear friend, Xi Tian, who has always been supportive to me with R coding strategies. Without their helps, there would not be the completion of this thesis.

Additionally, I would like to thank my dearest parents who have given me so much spiritual comforts and loves all the time.

Last but not the least, I would like to thank the Schreyer Honor's College for the efforts of providing students great opportunities to stimulate their thoughts and release their potentials. It is such an honor of being one of the members in the Schreyer Honor's College.

Chapter 1: Method

In many studies, regression analysis has been wildly used to figure out the relationships between multiple variables. However, when we have variables that are dependent on time, it is misleading to simply use regression analysis without including time. Therefore, time series analysis, as a method of tracking data over time has been invented and applied in econometrics and financial studies. In statistical time series analysis, multiple methods have been developed including autoregressive modeling, moving average modeling, kernal smoothing, box jenkins, bootstrap and so on. In the financial market, tracking the amount of changes for different time periods is essential. Many methods such as autoregressive and moving average models have been developed to catch the features of financial data. Other methodologies such as markov chain and stochastic volatility models are also available for time series analysis. In this Chapter, I will mainly focus on the discussion of MA (moving average) model, AR (autoregressive) model, and ARIMA (autoregressive-integrated-moving-average) model. Moreover, I will introduce the usage of ACF (autocorrelation function) and PACF (partial autocorrelation function) to propose possible models.

1.1 Moving Average Model

For a simple time series data $\{X_t\}_{t=1}^{n}$ with a smooth trend of f_t and stationary noise w_t, we can denote a simple equation as:

$$X_t = f_t + \omega_t$$

An example of smooth trend $f(\cdot)$ could be$\beta_0 + \beta_1 i + \beta_2 i^2$. That way, the time series trend equation can be converted into linear regression problem by estimating the coefficients of $\beta_0, \beta_1, \beta_2$. [5]

However, real life cases are usually more complicated than simple linear or quadratic. One of the most popular smoothing methods to determine an unknown form is MA (moving average) method. The main idea of MA model is capturing a series of subgroups of the full data points and averaging the points in each subgroup. By taking the average of each subset bandwidth, we are able to smooth out short-term oscillations and focus on long-term trend so that it will be easier for us to observe the relationship of variables. For data with natural window size such as monthly data, moving average method may have a strong performance.

Assume white noise sequence w_t that are identically and independently distributed, each with 0 mean and same variance $w_t \sim i.i.d N(0, \sigma_w^2)$. The qth order of moving average model could be explained by MA (q) as:

$$X_t = \mu + w_z + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_1 W_{t-q}$$

Aware that MA models are always weekly stationary because they are finite linear combinations of a white noise sequence for which the first two moments are time-invariant. [4]

Even though the sliding windows of moving average can provide us a better idea of long-term trend across time, certain substantial amount of variability in our estimation may still exist. For data without natural window size, deciding an appropriate priori window size may also be a hard work.

1.2 Autoregressive Model

The autoregressive model is another common model used in time series. It is a group of linear function that predicts the output variable of a system based on its own previous values. Consider a time series of X_t, a general autoregressive model of order p could be explained by AR (p) as:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + W_t$$

Assuming white noise follows $w_t \sim i.i.d N(0, \sigma_w^2)$.

AR model is also known as an infinite impulse response filter estimating the current term of the series by a linear weighted sum of previous terms in the series. AR analysis is a method to obtain the most appropriate weights (autoregression coefficients"ϕ_p"). The higher the order (p) of the system, the more accurate the estimations will be. There are many techniques observing the AR coefficients. Two of the main methods are least squares (Yule-Walker) and Burg method [3].

AR model is wildly used in applications such as speech processing, image compression, redundancy removal, and so on. With the straightforward AR model, we can easily determine current output by observing the relationship with its previous values. However, because we do not consider the past disturbance and process model, AR model can sometimes be misleading. [4]

1.3 Autoregressive Integrated Moving Average Model

I have discussed moving average model (section 1.1) for smoothing time series trend by averaging each window bandwidth and autoregressive model (section 1.2) for describing linear relationship of current values based on previous values. In some practices, AR model or MA model alone may become complicated to analyze high-order model with many parameters. It is hard to observe the dynamic structure of the data. Therefore, finding a method to overcome the difficulty is necessary. In modern time series, a method that combines both the ideas of AR model and MA model has been discovered, named ARMA model (autoregressive moving average). [5] With the difference term, a more general model is called ARIMA (autoregressive integrated moving average).

ARIMA model, also named Box-Jenkin model. It keeps the number of parameters small by compacting autoregressive terms and moving average terms.

Assuming a time series of X_t and white noise sequence $w_t \sim i.i.d N(0, \sigma_w^2)$, a general ARMA model with order of (p, q) can be denoted as:

$$X_t = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} + W_t + \sum_{i=1}^{q} \theta_i W_{t-i}$$

ARMA model defines the model with p autoregressive terms and q moving average terms. The minimized orders of p and q simplify the ARMA model and make it easier for representation.

With the difference terms and the lag operator L^i, we can write the general ARIMA (p, d, q) model as:

$$\left(1 - \sum_{i=1}^{p} \alpha_i L^i\right)(1-L)^d X_t = \left(1 + \sum_{i=1}^{q} \beta_i L^i\right) w_t$$

To find an appropriate ARIMA (p, d, q) model, autocorrelation functions (ACF) and partial autocorrelation functions (PACF) can be deployed (section 1.4).

1.4 Autocorrelation Function and Partial Autocorrelation Function

In time series analysis, we use ACF and PACF plots to determine whether we should use AR, MA or ARIMA model.

ACF measures the correlation between values of a time series process at different time points. [6] The correlation of coefficient between two time point x_t and $x_{t-k}$is also called the lag k autocorrelation of x_t. At lag k, we can describe the ACF as:

$$ACF(k) = \frac{E\left[\left(x_t - E(x_t)\right)\left(x_{t-k} - E(x_{t-k})\right)\right]}{\sqrt{var(x_t) var(x_{t-k})}}$$

PACF is the partial autocorrelation of a time series at lag k. It is the autocorrelation between two different time pints x_t and x_{t-k} after removing the effects of other time lags 1, 2, …, k-1. We can define the partial autocorrelation of k^{th}order as:

$$\text{PACF (k)} = Cor\left(x_{t+k} - P_{t,k}(x_{t+k}), x_t - P_{t,k}(x_t)\right)$$

The basic rules of choosing appropriate model based on ACF and PACF plots are described as following:

- ACF tails off gradually, PACF cuts off after p lags → AR(p) model
- ACF cuts off after q lags, PACF tails off gradually → MA(q) model
- ACF tails off gradually, PACF tails off gradually → ARIMA (p, d, q) model

When observing time series models, it is essential to check the patterns of ACF and PACF plots. We also need to aware that ACF and PACF are significant under the condition of stationary data. [6]

Chapter 2: Application on Dow Jones Index

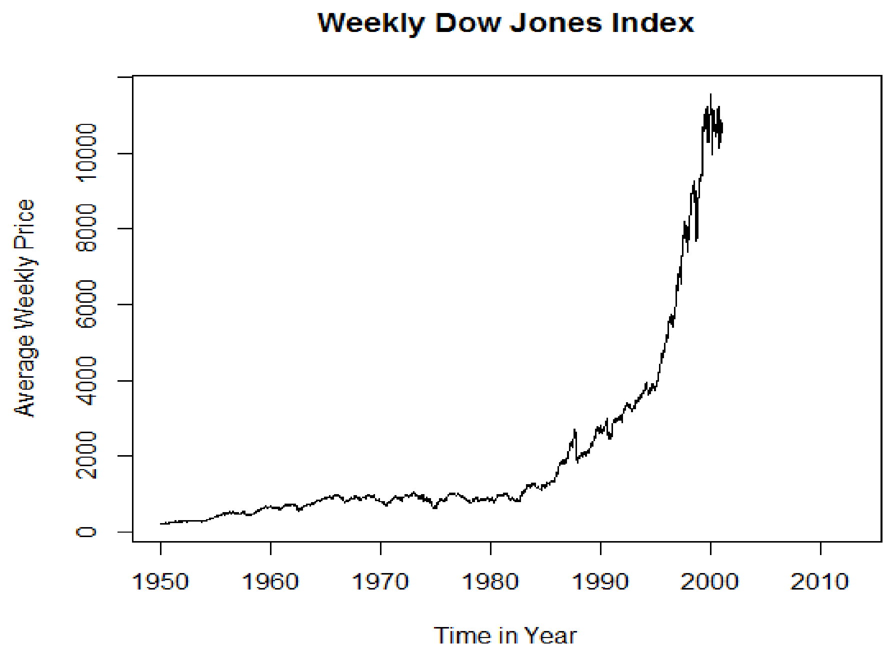
From Chapter 1, I have briefly introduced three useful methods (AR, MA, ARIMA models) for analyzing time series problem. In this chapter, I will apply the three methods on Dow Jones index and observe the features of the data. The average weekly price of Dow Jones index starting from 1/6/1950 to12/22/2000 will be used in this thesis as an example.

Dow Jones index, as the earliest financial index, witnesses the development of the financial market and represents the change of it. Analyzing its price change could be necessary to catch important information from the market, which allows analysts and investors to better react toward market shocks. Many people argue that predictions on stock prices are not meaningful due to the high volatility of the market. However, if we cut the historical Dow Jones index into different periods, we are able to obtain certain meaningful trends for each of those periods. Therefore, it is significant to study its characteristics in order to have a better understanding on its movement.

2.1 Data Overlook and Moving Average Smoothing

Before selecting the best model for Dow Jones Index, it is essential to check the general feature of the data. To do that, we can simply fit a time series plot on the original data. As we can see from *Figure 2.1-1*, the average weekly price of the Dow has a general upward trend along the time axis.

Figure2.1- 1: General Time Series Plot on Average Weekly Price



To give an example of moving average smoothing technique and display a typical non-equal variance situation of a lot of data, I filtered the index in a monthly base by dividing it by 12 (1 year = 12 months). *Figure 2.1-2* is the residual plot from the original data with the filter or sliding window size of 12 from moving average. It is obvious that the variance is not stationary and has an expanding trend over time.

Due to the constant variance assumption condition in time series, it is important for us to smooth the data and get more consistent variance trend before applying time series techniques in our analysis. One of the most common methods of reducing variance difference is taking logarithm of the data. *Figure 2.1-3* shows that the residual plot after transferring the data by taking logarithm is much more constant than the original one.

Figure 2.1-2: Residual Plot with Filter 12

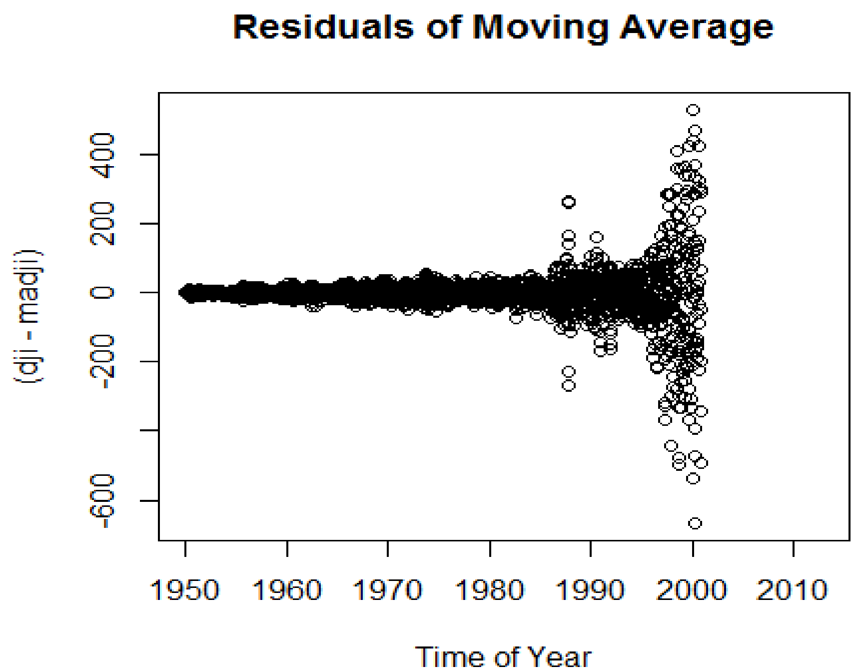
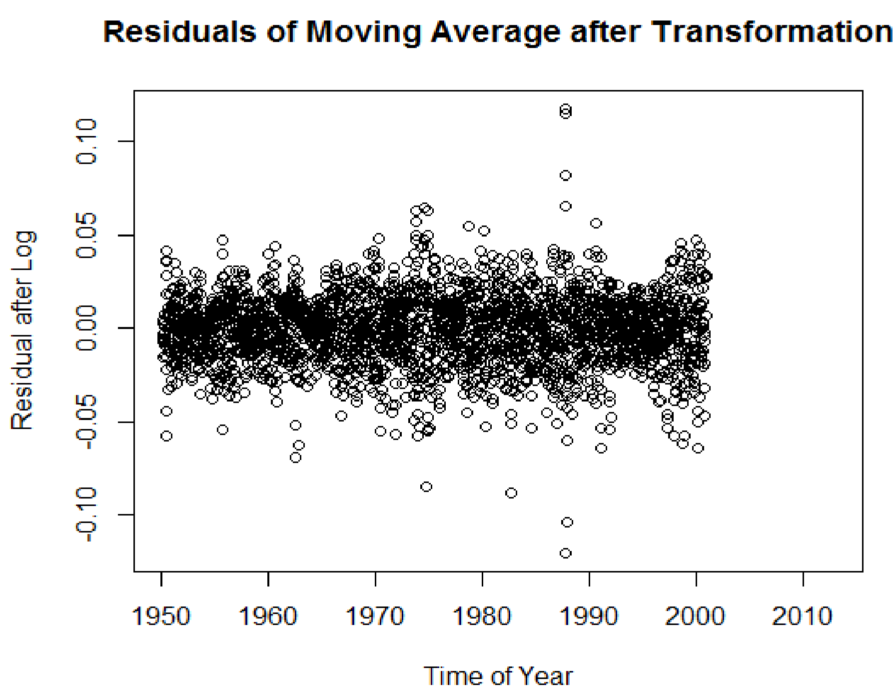


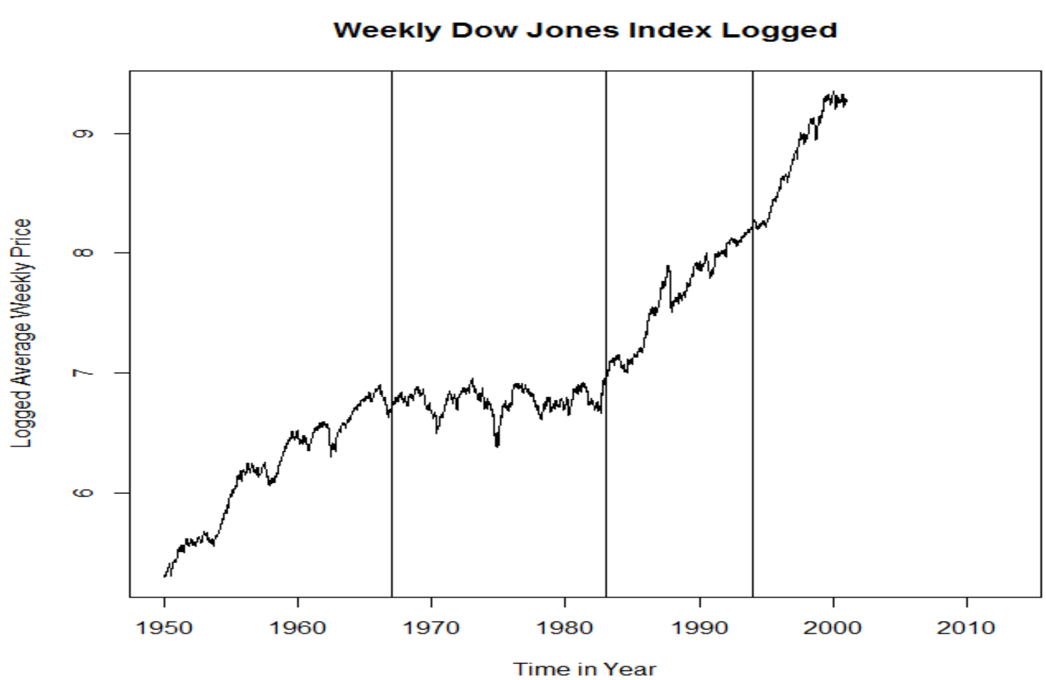
Figure2.1- 3: Residual Plot after Transformation



2.2 Cutting Points

I have showed how to decrease the fluctuation of variance by taking logarithm above. Now, I will start to use the Logged average weekly price of the Dow Jones Index so that the variance trend of our data will be decreased. *Figure 2.2-1*exhibits the shape of the average price after logarithm transformation. Noticing that the price has different behavior trends, I have divided it into four periods. From 1950 to 1966, the stock market was growing steadily due to the recovery of the economy after World War II. From 1967 to 1982, the stock market went into an adjusting period with roughly flat trend of growth. From 1983 to 1993, the financial market was acting robustly after the dull period. Due to the Dot-com Bubble effect, the stock market was growing extraordinarily from 1994 to 2000.

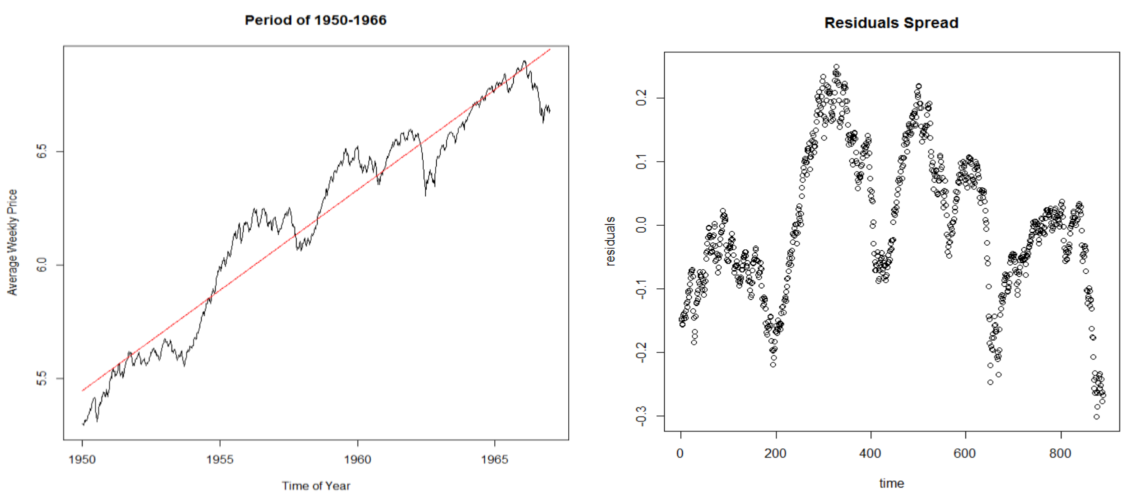
Figure 2.2-1: Cutting Points of Logged Price



2.3 Period 1 from 1950 to 1966

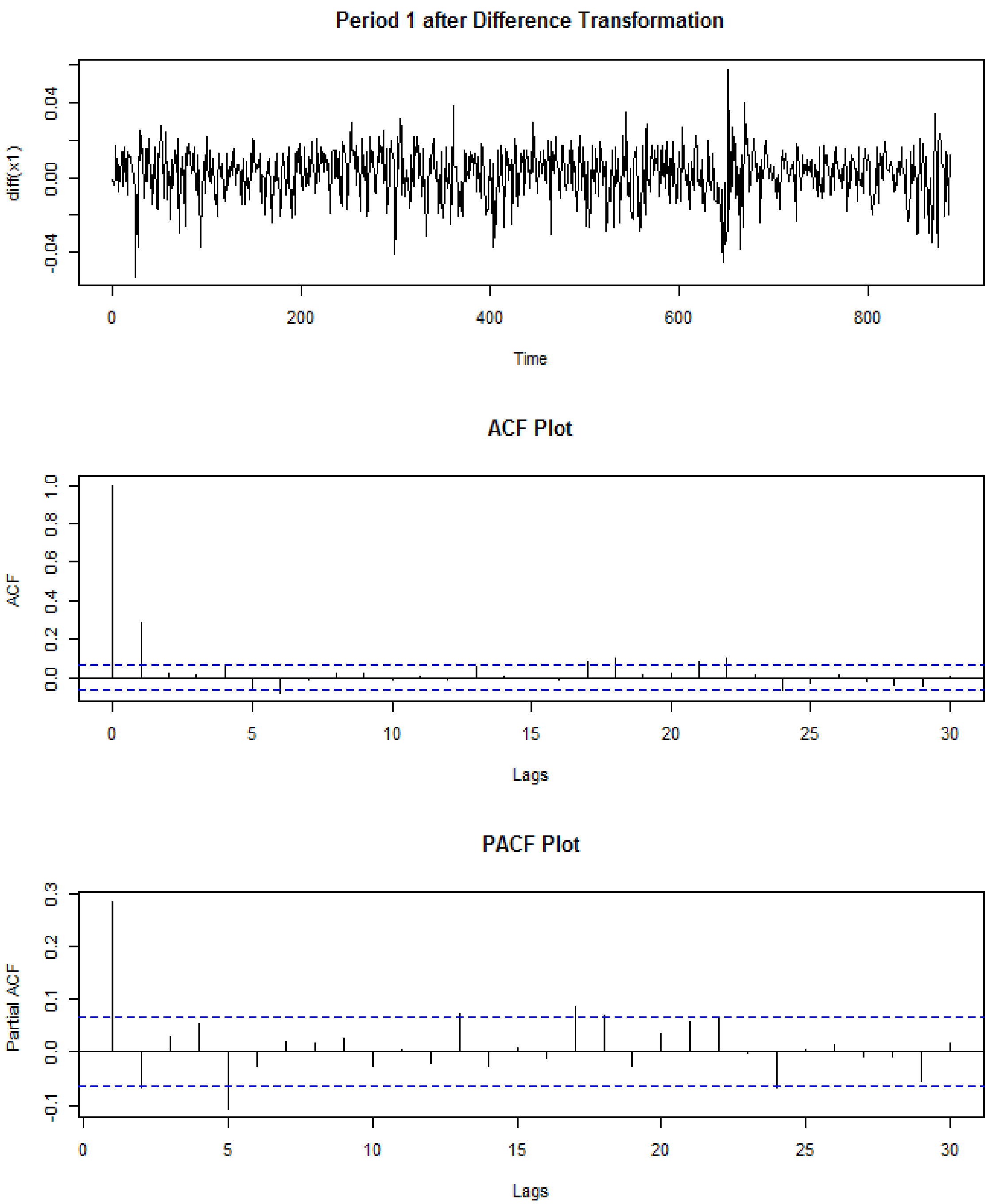
Now that we have divided the data into four different periods, let us diagose the four periods one by one. From *Figure 2.3-1*, we can see that the first period from 1950 to 1966 has a pretty steady growing trend and spreads nicely long the fitted line (upward red line). However, the variance from period 1 seem to fluctuate with some obvious spikes. One of the most common way of reducing the trend of variance is to take difference of the data. The top plot in *Figure 2.3-2* displays the time series plot after taking the first difference and it is clear that the fluctations have been largely reduced.

Figure 2.3-1: Period 1 Time Series Plot and Residual Plot



Follwing by the model selecting rule, lag 1 from the ACF plot and lag 5 from the PACF plot are significant (*Fiugre 2.3-2*). Therefore, we diagnose the possible models could be ARIMA (0,1,1), ARIMA (5,1,0) and ARIMA (5,1,1). Later, I will further check the fit of our suspected models by comparing AIC values, coefficients and their standard errors. Keep in mind that when the coefficient is less than two times its standard error, the model need to be reduced.

Figure 2.3-2: Model Diagnose for Period 1



After checking all the models (detail in Appendix A), a summary of the comparision of the potential models is given in *Table 2.3-1*. From theory, the best model always occurs with the smallest AIC value. Even though ARIMA (5,1,0) gives us the smallest AIC value, its coefficient for AR(2) is less than two times of its standard error. Thus, we reduce the model to ARIMA (1,1,0). As we can see, ARIMA (0,1,1) has the smallest AIC value among all. Therefore, we can conclude that the ARIMA (0,1,1) could be the best model explaining the period from 1950 to 1966 with the MA(1) coefficient of 0.2999.

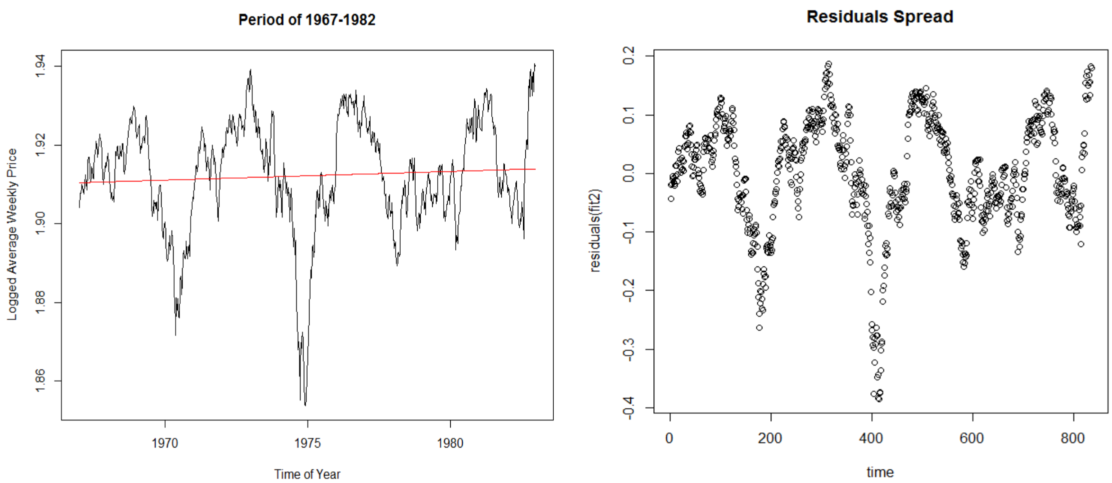
Model	**AIC**	**Coefficients vs. SE**
ARIMA (0,1,1) **	-5235.56	Good
ARIMA (5,1,0)	-5242.07	AR(2) coefficient < 2SE
ARIMA (1,1,0)	-5233.54	Good
ARIMA (1,1,1)	-5234.91	AR(1) coefficient < 2SE MA(1) coefficient < 2SE

Table 2.3-1: Model Selection for Period 1

2.4 Period 2 from 1967 to 1982

Different than the upward trend of the first period, the second period does not have an obvious trend. Fluctating around the fitted line, we notice that the average prices move almost constantly. Therefore, we can roughly conclude that period two has approximately constant behavior with the mean weekly price of $874. *Table 2.4-1* summarizes the basic features of the data.

Figure 2.4-1: Period 2 Time Series Plot and Residual Plot



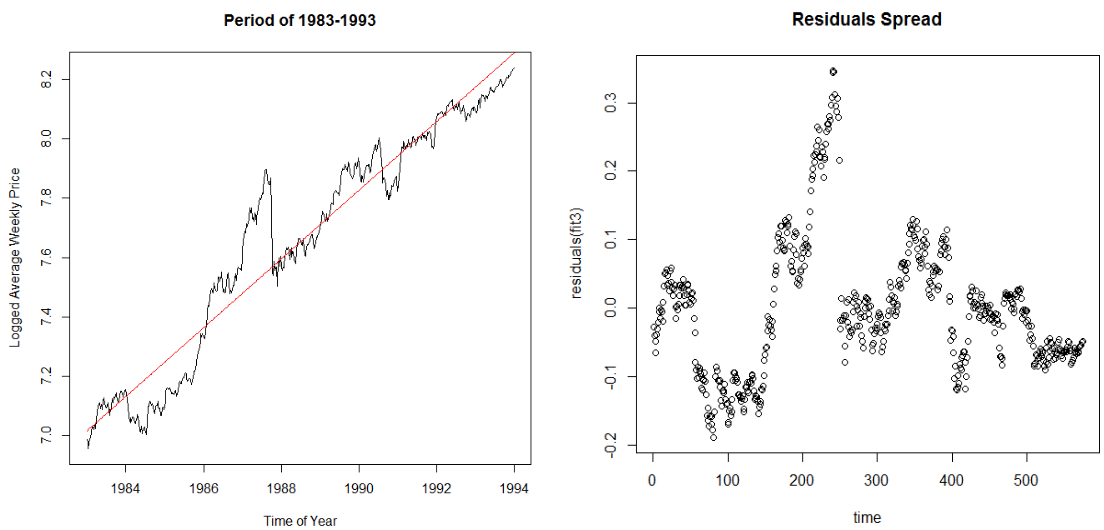
Min	1st Quantile	Median	Mean	3rd Quantile	Max
591.7	828.2	874.5	874.0*	936.5	1057.0

Table 2.4-1: Summary for Period 2

2.5 Period 3 from 1983 to 1993

From 1983 to 1993, the financial market went into a new developing period after the non-growing period of 1967 to 1982. Similarly to what we did for period one, it is important for us smooth the data due to the instability of its residuals. *Figure 2.5-1* shows the general trend of the logged price and the residual spread of the data. Apparently, the variance is not constant with some oscillations in the plot. Therefore, I took and the first difference of the logged price to remove the fluctuations.

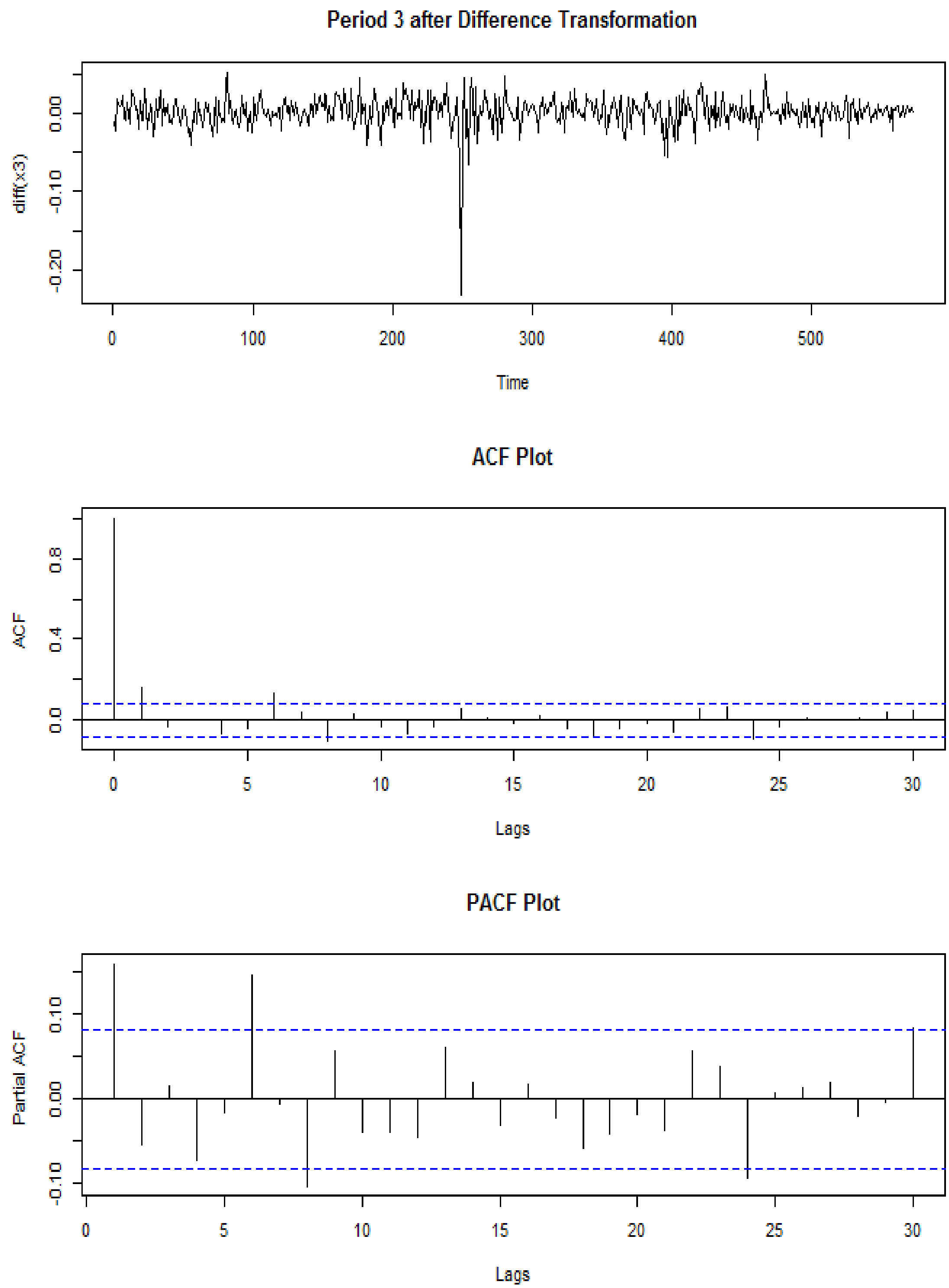
Figure 2.5-1: Period 3 Time Series Plot and Residual Plot



Once we have the data smoothed out, we are able to start our model diagnosis and selection. Repeat the steps as we did in period 1, it is necessary for us to propose potential models with based on the autocorrelation and partial autocorrelation functions. As we can see from ACF plot in *Figure 2.5-2*, the first lag obviously exceeds the dotted blue line. Since both lag 6 and lag 8 exceed the blue line, it is essential for us to check both AR (6) and AR (8). Therefore, our initiated models could be ARIMA (0,1,1), ARIMA(6,1,0), ARIMA(8,1,0).

After fitting the models, we notice that both ARIMA (6,1,0) and ARMIA (8,1,0) have AR2 coefficients smaller than two times of their standard errors. Therefore, we need to reduce the model into lower order and propose a new potential model ARIMA (1,1,0) and ARIMA (1,1,1).

Figure 2.5-2: Model Diagnose for Period 3



As we can see from the summarized output from *Table 2.5-1*, ARIMA(1,1,1) has both AR1 and MA1 coefficients smaller than two times their standard errors. Therefore, we elimiate it and compare the AIC values between ARMIA (0,1,1) and ARIMA(1,1,0). Since ARIMA(0,1,1) has the smallest AIC value, we can conclude that model ARIMA (0,1,1) best describes the weekly price of period 3 with the MA1 coefficient of 0.1837.

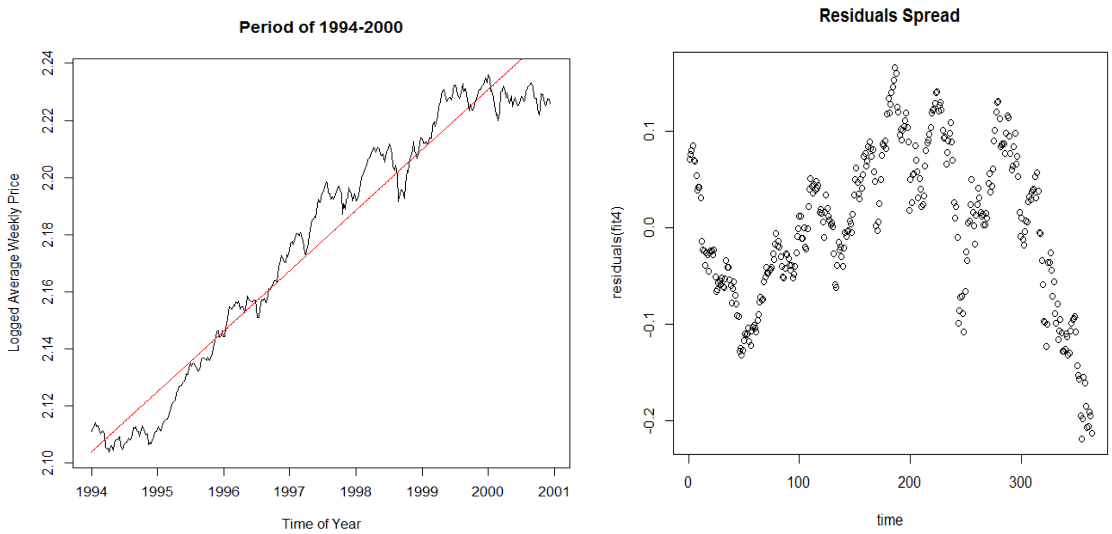
Model	**AIC**	**Coefficients vs. SE**
ARIMA (0,1,1) **	-2908.95	Good
ARIMA (6,1,0)	-2915.36	AR2 coefficient < 2SE
ARIMA (8,1,0)	-2916.45	AR2 coefficient < 2SE
ARIMA (1,1,0)	-2907.74	Good
ARIMA (1,1,1)	-2907.36	AR1 coefficient < 2SE MA1 coefficient < 2SE

Table 2.5-1: Model Selection for Period 3

2.6 Period 4 from 1994 to 2000

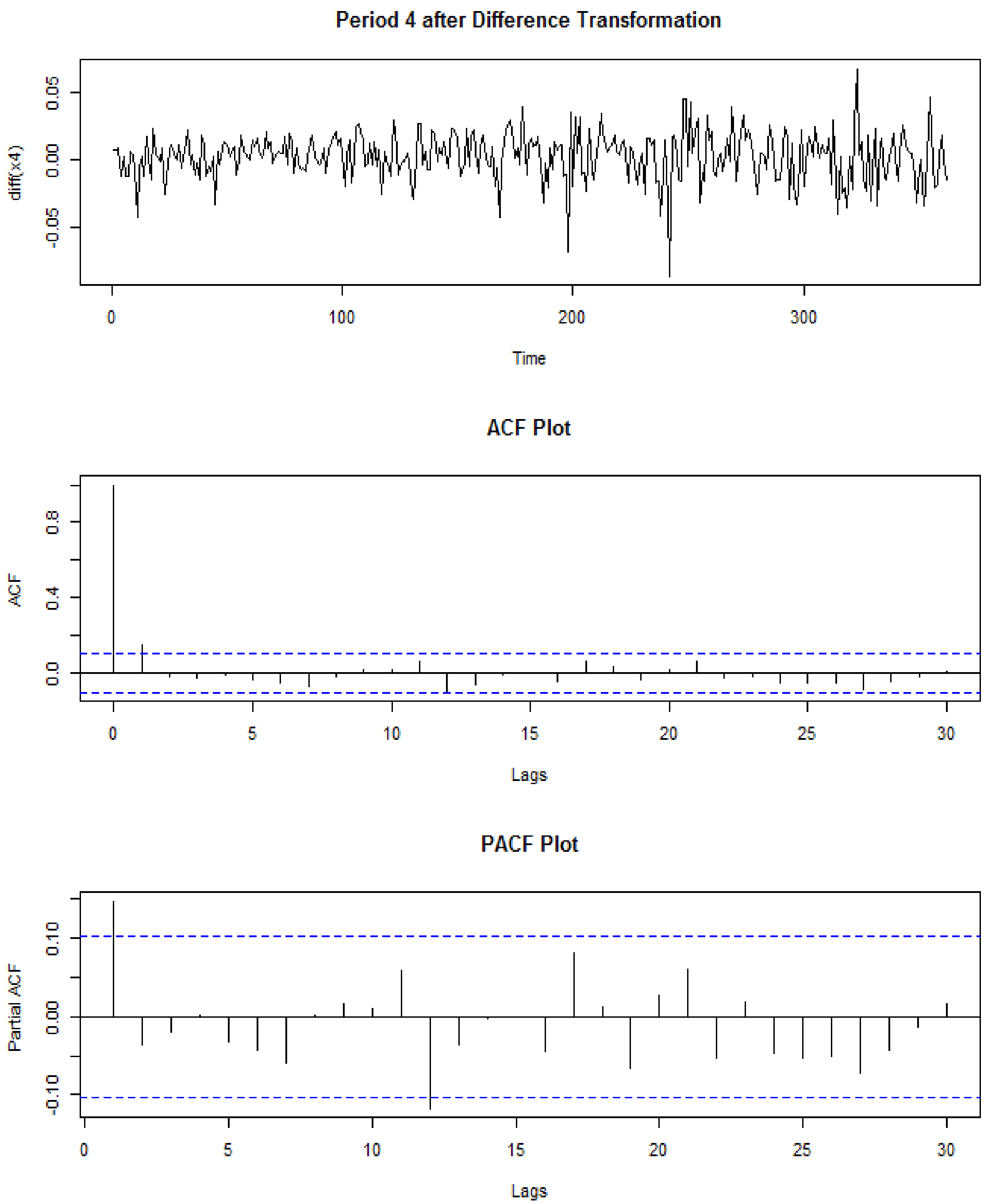
Due to the Dot- Com Bubble Effect, the financial market rocketed up during the fourth period from 1994 to 2000. Following the same technique that we used for the previous periods, I also took the first difference of the data to remove the fluctuations. *Figure 2.6-1* contains the general trend plot and the residual spread plot.

Figure 2.6-1: Period 2 Time Series Plot and Residual Plot



From the ACF plot in *Figure 2.6-2*, we notice that lag 1 has higher ACF than the fitted blue line, which suggests the model of MA 1. In PACF plot, lag 12 exceeds the fitted blue line, which suggests the model of AR 12. Therefore, our initial potential models are ARIMA (0,1,1), ARIMA (12,1,0). However, after checking the fits of the models, we relaize that the AR 2 coefficent in ARIMA (12,1,0) is smaller than two time of its standard error. Thus, reduction of the model was processed until we have all of the coefficients bigger than two times of each of their standard error. After decreasing the order, we get the models of ARIMA (1,1,0) and ARIMA (1,1,1). Since the coefficients of AR1 and MA1in ARIMA (1,1,1) are smaller than two times of their coefficients, we automatically elimiate the model. Because ARIMA (0,1,1) has the smallest AIC value, we conclude that ARIMA (0,1,1) is our best model that describes period 4 with MA1 coefficient of 0.169.

Figure 2.6-2: Model Diagnose for Period 4



Model	AIC	Coefficients vs. SE
ARIMA (0,1,1) **	-1902.24	Good
ARIMA (12,1,0)	-1888.18	AR2 coefficient < 2SE
ARIMA (1,1,0)	-1902.2	Good
ARIMA (1,1,1)	-1900.3	AR1 coefficient < 2SE MA1 coefficient < 2SE

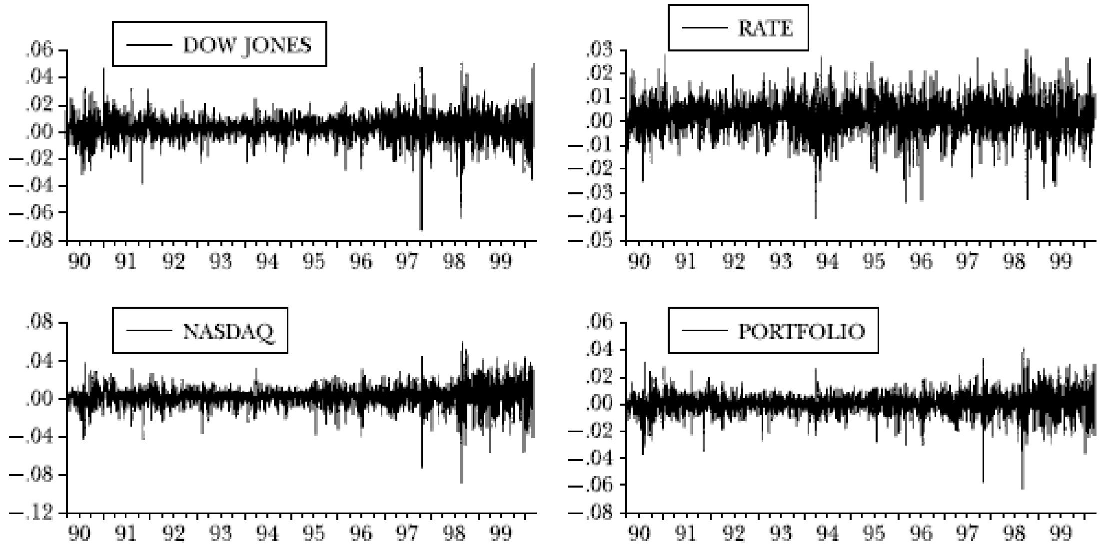
Table 2.6-1: Model Selection for Period 4

Chapter 3: ARCH and GARCH Models

In modern financial econometrics practices, it is common that we have inconsistent variances. Therefore, specific types of nonlinear dependence models called ARCH (autoregressive conditionally heteroscedastic) and GARCH (Generalized autoregressive conditional heteroskedastic) models have been developed to predict the changes in volatility of time series data. ARCH/GARCH models are applied to observe the changes of volatile variance, especially for time series with generally increasing variance over time (heteroscedastic). ARCH/GARCH models treat heterskedasticity as a variance and include it in the model. [10] Once a suitable financial model has been built, econometricians and investors are usually required to analyze and predict the error terms or volatility of the model in order to provide a more accurate forecast on the range of future price or return.

In financial industry, different errors tend to happen in different cluster/serially correlated. For example, small returns are followed by smaller returns, while large returns are followed by larger returns. [11] Similarly, when we look at the financial data, some periods of time have constant high price while some other periods are very volatile and risky. *Figure 3-1* shows different financial data from Nasdaq, Dow Jones and bond return that volatile over time. [10] It is obvious to notice that in certain time periods, the prices and returns fluctuate largely which implies that these periods are risker than others. In Chapter 2, I have generally discovered the best model for each period. Based on the models we observed, we are able to find appropriate ARCH or GARCH models for each period.

Figure 3-1: Volatile Financial Data over Time



Engle, R. (2001)

3.1 ARCH and GARCH Models

Knowing that the variance of different time t is under the condition of the observations at previous times, we are able to drive the general variance for ARCH (p) model as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2$$

p is the past times that can be explained in the process; α_p is the coefficients and ε_t is the errors for the observations at previous times. Therefore, the ARCH process can be converted into AR(p) model.

If we combine both the AR(p) terms and MA(q) terms, we are able to generate the general expression for GARCH (p,q) model as:

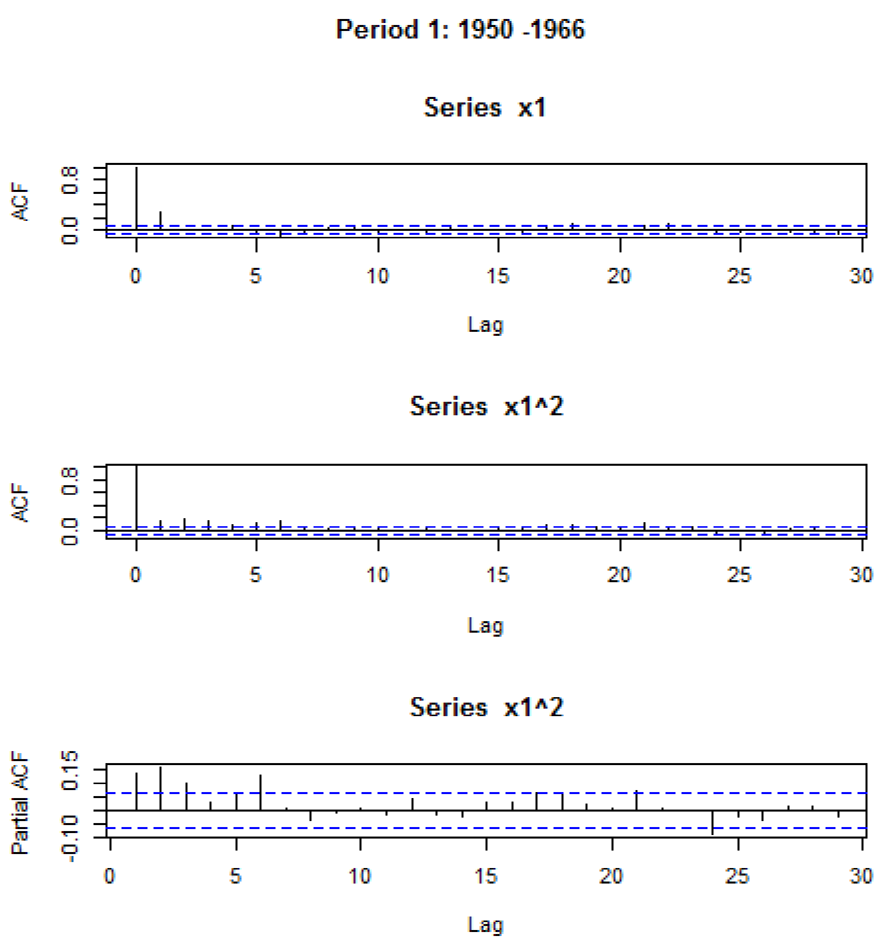
$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_q \sigma_{t-q}^2$$

To define the appropriate ARCH/GARCH models, we can also autocorrelation function (ACF) and partial autocorrelation function (PACF) plots.

3.2 Period 1 from 1950 to 1966

To estimate the appropriate variance model, we need to diagnose the ACF and PACF plots in order to observe the cutting lags that could possibly explain the model.

Figure 3.2-1: ACF and PACF Plots



Before checking the suitable models, let us check the Box-Pierce test for non-correlation, the Shapiro-Wilk Normality test for normal distribution, and the KPSS test for stationarity. The three tests give us the p-values as shown in *Table 3.2-1*, which imply that the differenced logged Dow Jones data in period 1 is approximately non-correlated with normal distribution and stationary trend. The results from these tests are consistent with our theoretical properties.

Test	**P-value (compared with 0.05 significance value)**
Box-Pierce (x1)	< 2.2e-16
Box-Pierce (x1^2)	3.473e-05
Shapiro-Wilk (x1)	2.629e-09
KPSS (x1)	0.1

Table 3.2-1: Box-Pierce, Shapiro-Wilk, and KPSS

Based on the ACF and PACF plots, we are able to propose multiple potential models. After fitting and comparing the possible models, we find a good model could be GARCH (1,1). *Table 3.2-2* below summarizes the coefficients of the model.

Coefficients	**Estimate**	**Std. Error**	**Pr(>|t|)**
a0	9.591e-06	4.724e-06	0.042322
a1	7.671e-02	2.020e-02	0.000146
b1	8.688e-01	4.160e-02	< 2e-16

Table 3.2-2: Summary of GARCH (1,1)

3.3 Period 3 from 1983 to 1993

In Period 3, we observe the ACF and PACF plots as shown in *Figure 3.3-1*. From the ACF (x3) plot, lag 1 has an obvious line higher than the blue line. Based on the three plots, we propose possible GARCH models including GARCH (0, 1), GARCH (1, 0), and GARCH (1, 1).

Figure 3.3-1: ACF and PACF Plots

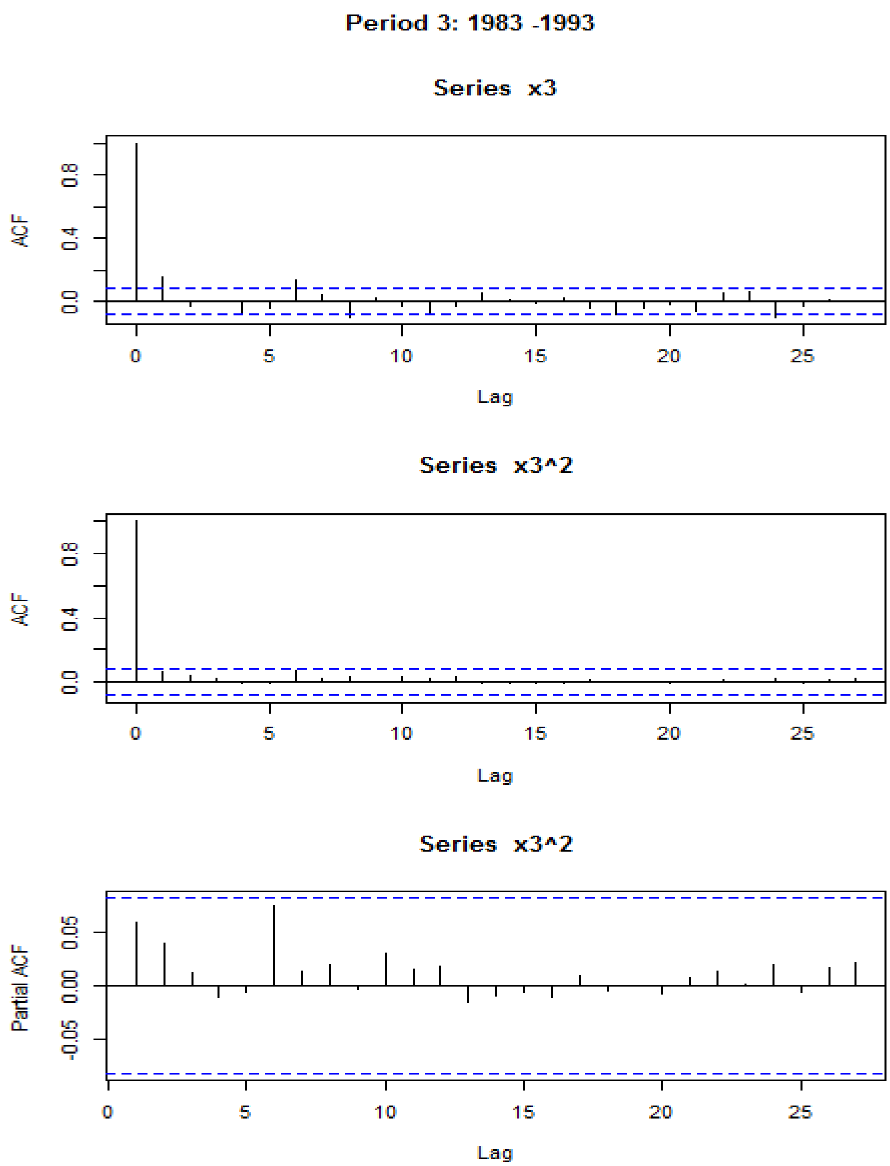


Table 3.3-1 checks the correlation, normality and stationarity for the data. Noticing that for the second power of x3, the Box-Pierce test has the p-value greater than 0.05, which shows that the data for the second power of x3 is correlated. Further research on the correlated variables could be studied in the future.

Test	P-value (compared with 0.05 significance value)
Box-Pierce (x3)	0.0001265
Box-Pierce (x3^2)	0.1566
Shapiro-Wilk (x3)	< 2.2e-16
KPSS (x3)	0.1

Table 3.3-1: Box-Pierce, Shapiro-Wilk, and KPSS

After trying all the possible models, we notice that GARCH (1, 1) provide us the best fit of the data as well as the smallest AIC value. Therefore, GARCH (1, 1) could be a good model that explains the variance change of the Dow Jones price in the third period. *Table 3.2-2* is the summary of the coefficients for the GARCH (1, 1) model.

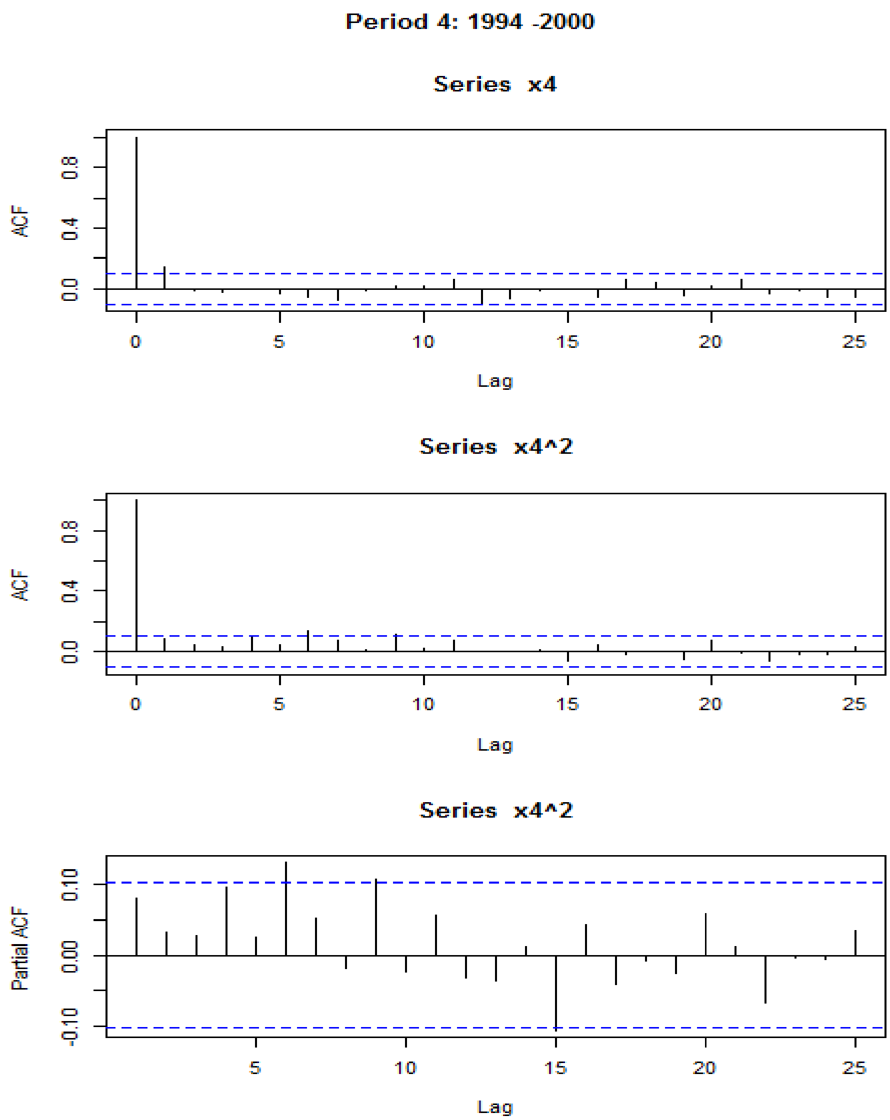
Coefficients	Estimate	Std. Error	Pr(>|t|)
a0	7.854e-05	1.552e-05	4.16e-07
a1	3.055e-01	1.964e-02	< 2e-16
b1	4.857e-01	5.442e-02	< 2e-16

Table 3.2-2: Summary of GARCH (1,1)

3.4 Period 4 from 1994 to 2000

In Period 4, we notice that the ACF for x4 has a significant lag 1 in the series. Considering the lag features for ACF and PACF plots, we are able to propose possible GARCH models to explain the variance change for the fourth period.

Figure 3.4-1: ACF and PACF Plots



The Box-Pierce, Shapiro-Wilk, and KPSS tests results are summarized in *Table 3.3-1*. Similar with the third period, the second power of x4 has p-value larger than the significance level of 0.05, which suggests that there are possible correlation regards to it. Further study on the correlation could be done in the future research.

Test	**P-value (compared with 0.05 significance value)**
Box-Pierce (x4)	0.005036
Box-Pierce (x4^2)	0.1275
Shapiro-Wilk (x4)	4.643e-06
KPSS (x4)	0.1

Table 3.3-1: Box-Pierce, Shapiro-Wilk, and KPSS

After testing all the potential models, we notice that GARCH (1, 1) has approximately the best fit with the coefficients shown in the table below.

Coefficients	**Estimate**	**Std. Error**	**Pr(>|t|)**
a0	9.720e-06	8.763e-06	0.26732
a1	5.647e-02	1.874e-02	0.00258
b1	9.138e-01	3.957e-02	< 2e-16

Table 3.4-2: Summary of GARCH (1,1)

Chapter 4: Conclusion and Discussion

In Chapter 2, we notice that the four different periods of the Dow Jones Index can be represented by different models. After taking the first different of the overall data and fitting the potential models, we can see that period 1 from 1950 to 1966, period 3 from 1983 to 1993 and period 4 from 1994 to 2000 have the best model of ARIMA (0,1,1), which includes the first different term and MA (1) term. The MA (1) term for period 1, 3 and 4 are 0.2999, 0.1837 and 0.169, respectivley. The second period from 1967 to 1982 can be treated as a constant trend with the average weekly price of $874.

In Chapter 3, the change of variance for each period was diagosed by GARCH model. After fitting all potential models, we conclude that GARCH (1,1) model could be an appropriate model for period 1, period 3, and period 4 with different coefficients summarized in the tables in Chapter 3.

Among all the time series techniques, AR, MA and ARIMA are the most commonly used ones. ARCH/GARCH models are wildely used to further anlayze the volatility changes. In the complicated and extremly volatile market, more advanced time series models have been developed for more accurate anlysis. Finding the general trend of the data is just the first step in my research. In my future study, I will focus on narrowing down the window sizes until each small window has the highest explaination on the trend of its data within its bandwidth. This could be done by applying locally stationary methodlogy. The success of finding the appropriate window size will provide us stronger power and accuracy of future value prediction.

Appendix A

R Outputs for ARMIA Model Check for Period 1950-1966

```
> arima(x=x1,order=c(0,1,1))

Call:
arima(x = x1, order = c(0, 1, 1))

Coefficients:
         ma1
      0.2999
s.e.  0.0303

sigma^2 estimated as 0.0001592:  log likelihood = 2619.78,  aic = -5235.56
> arima(x=x1,order=c(5,1,0))

Call:
arima(x = x1, order = c(5, 1, 0))

Coefficients:
         ar1      ar2     ar3     ar4      ar5
      0.3185  -0.0645  0.0109  0.0936  -0.0996
s.e.  0.0334   0.0349  0.0350  0.0350   0.0335

sigma^2 estimated as 0.0001567:  log likelihood = 2627.03,  aic = -5242.07
> arima(x=x1,order=c(1,1,0))

Call:
arima(x = x1, order = c(1, 1, 0))

Coefficients:
         ar1
      0.2950
s.e.  0.0321

sigma^2 estimated as 0.0001596:  log likelihood = 2618.77,  aic = -5233.54
> arima(x=x1,order=c(1,1,1))

Call:
arima(x = x1, order = c(1, 1, 1))

Coefficients:
         ar1     ma1
      0.1163  0.1968
s.e.  0.1001  0.0980

sigma^2 estimated as 0.000159:  log likelihood = 2620.45,  aic = -5234.91
>
```

Appendix B

R Outputs for ARMIA Model Check for Period 1983-1993

```
> arima(x=x3,order=c(0,1,1))

Call:
arima(x = x3, order = c(0, 1, 1))

Coefficients:
         ma1
      0.1837
s.e.  0.0423

sigma^2 estimated as 0.0003628:  log likelihood = 1456.48,  aic = -2908.95
> arima(x=x3,order=c(6,1,0))

Call:
arima(x = x3, order = c(6, 1, 0))

Coefficients:
         ar1      ar2     ar3      ar4      ar5     ar6
      0.1813  -0.0415  0.0300  -0.0521  -0.0333  0.1552
s.e.  0.0412   0.0419  0.0419   0.0419   0.0419  0.0412

sigma^2 estimated as 0.0003525:  log likelihood = 1464.68,  aic = -2915.36
> arima(x=x3,order=c(8,1,0))

Call:
arima(x = x3, order = c(8, 1, 0))

Coefficients:
         ar1      ar2     ar3      ar4      ar5     ar6     ar7      ar8
      0.1811  -0.0267  0.0272  -0.0572  -0.0304  0.1508  0.0194  -0.0938
s.e.  0.0416   0.0423  0.0418   0.0418   0.0417  0.0417  0.0422   0.0415

sigma^2 estimated as 0.0003493:  log likelihood = 1467.22,  aic = -2916.45
> arima(x=x3,order=c(1,1,0))

Call:
arima(x = x3, order = c(1, 1, 0))

Coefficients:
         ar1
      0.1708
s.e.  0.0411

sigma^2 estimated as 0.0003636:  log likelihood = 1455.87,  aic = -2907.74
> arima(x=x3,order=c(1,1,1))

Call:
arima(x = x3, order = c(1, 1, 1))

Coefficients:
          ar1     ma1
      -0.1557  0.3352
s.e.   0.2259  0.2157

sigma^2 estimated as 0.0003626:  log likelihood = 1456.68,  aic = -2907.36
>
```

Appendix C

R Outputs for ARMIA Model Check for Period 1994-2000

```
> arima(x=x4,order=c(0,1,1))

Call:
arima(x = x4, order = c(0, 1, 1))

Coefficients:
         ma1
       0.169
s.e.   0.051

sigma^2 estimated as 0.0003024:  log likelihood = 953.12,  aic = -1902.24
> arima(x=x4,order=c(12,1,0))

Call:
arima(x = x4, order = c(12, 1, 0))

Coefficients:
         ar1      ar2     ar3     ar4      ar5      ar6      ar7     ar8     ar9    ar10    ar11     ar12
      0.1745  -0.0164  0.0001  0.0292  -0.0115  -0.0154  -0.0442  0.0179  0.0352  0.0230  0.0954  -0.1014
s.e.  0.0522   0.0527  0.0528  0.0527   0.0527   0.0528   0.0527  0.0531  0.0531  0.0532  0.0538   0.0533

sigma^2 estimated as 0.0002956:  log likelihood = 957.09,  aic = -1888.18
> arima(x=x4,order=c(1,1,0))

Call:
arima(x = x4, order = c(1, 1, 0))

Coefficients:
         ar1
      0.1687
s.e.  0.0518

sigma^2 estimated as 0.0003024:  log likelihood = 953.1,  aic = -1902.2
> arima(x=x4,order=c(1,1,1))

Call:
arima(x = x4, order = c(1, 1, 1))

Coefficients:
         ar1     ma1
      0.0754  0.0961
s.e.  0.2938  0.2928

sigma^2 estimated as 0.0003023:  log likelihood = 953.15,  aic = -1900.3
```

Appendix D

R Outputs for GARCH Models

```
> f1=garch(x1,order=c(1,1))

 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****

     I     INITIAL X(I)        D(I)

     1     1.553548e-04     1.000e+00
     2     5.000000e-02     1.000e+00
     3     5.000000e-02     1.000e+00

    IT   NF      F         RELDF    PRELDF    RELDX   STPPAR   D*STEP   NPRELDF
     0    1 -3.395e+03
     1    8 -3.395e+03  3.99e-06  7.53e-06  1.2e-05  1.8e+10  1.2e-06  6.63e+04
     2    9 -3.395e+03  4.62e-08  5.04e-08  1.2e-05  2.0e+00  1.2e-06  1.15e+00
     3   18 -3.399e+03  1.16e-03  2.12e-03  4.5e-01  2.0e+00  8.2e-02  1.14e+00
     4   20 -3.405e+03  1.70e-03  2.31e-03  7.3e-01  1.9e+00  3.3e-01  8.96e-02
     5   29 -3.406e+03  1.64e-04  3.63e-04  7.6e-06  4.8e+00  5.8e-06  3.03e-03
     6   30 -3.406e+03  5.90e-06  5.67e-06  7.5e-06  2.0e+00  5.8e-06  1.00e-03
     7   31 -3.406e+03  4.35e-08  8.99e-08  7.6e-06  2.0e+00  5.8e-06  1.06e-03
     8   40 -3.407e+03  3.49e-04  7.22e-04  2.4e-01  7.0e-01  2.4e-01  1.06e-03
     9   41 -3.410e+03  8.41e-04  4.73e-04  4.6e-02  0.0e+00  7.5e-02  4.73e-04
    10   43 -3.411e+03  2.36e-04  2.44e-04  2.0e-02  1.9e+00  3.0e-02  6.99e-03
    11   45 -3.412e+03  4.12e-04  4.23e-04  3.8e-02  1.6e+00  6.0e-02  9.84e-03
    12   46 -3.413e+03  3.36e-04  7.24e-04  6.9e-02  1.5e+00  1.2e-01  3.84e-03
    13   54 -3.413e+03  2.97e-05  5.90e-05  2.3e-07  8.6e+00  4.1e-07  9.87e-04
    14   55 -3.413e+03  8.62e-08  1.10e-07  2.3e-07  2.0e+00  4.1e-07  5.63e-04
    15   64 -3.414e+03  1.59e-04  3.04e-04  9.8e-03  8.5e-01  1.7e-02  5.62e-04
    16   65 -3.414e+03  1.75e-05  3.57e-05  8.8e-03  7.8e-01  1.7e-02  5.62e-05
    17   66 -3.414e+03  1.24e-06  1.67e-06  2.3e-03  0.0e+00  4.2e-03  1.67e-06
    18   69 -3.414e+03  3.06e-10  2.76e-09  1.6e-05  1.8e+00  2.9e-05  2.24e-08
    19   72 -3.414e+03  4.88e-10  4.70e-10  3.0e-06  1.9e+00  5.5e-06  1.76e-08
    20   76 -3.414e+03  2.59e-11  2.25e-11  1.6e-07  2.0e+00  3.0e-07  1.70e-08
    21   80 -3.414e+03  4.37e-13  7.90e-13  5.7e-09  2.0e+00  1.1e-08  1.70e-08
    22   87 -3.414e+03 -5.41e-14  8.82e-18  1.9e-14  2.6e+01  3.5e-14  1.70e-08

 ***** FALSE CONVERGENCE *****

 FUNCTION    -3.413874e+03   RELDX        1.925e-14
 FUNC. EVALS      87         GRAD. EVALS      22
 PRELDF       8.820e-18      NPRELDF      1.701e-08

     I      FINAL X(I)        D(I)          G(I)

     1    9.591438e-06     1.000e+00     8.232e-01
     2    7.671099e-02     1.000e+00    -2.048e-02
     3    8.688334e-01     1.000e+00    -2.791e-01

> summary(f1)

Call:
garch(x = x1, order = c(1, 1))

Model:
GARCH(1,1)

Residuals:
    Min       1Q  Median      3Q     Max
-4.8930 -0.4800  0.2122  0.8228  2.8637

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0 9.591e-06   4.724e-06    2.030 0.042322 *
a1 7.671e-02   2.020e-02    3.797 0.000146 ***
b1 8.688e-01   4.160e-02   20.885  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals
X-squared = 63.1565, df = 2, p-value = 1.932e-14


        Box-Ljung test

data:  Squared.Residuals
X-squared = 0.0604, df = 1, p-value = 0.8059
```

```
> f3=garch(x3,order=c(1,1))

 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****


     I     INITIAL X(I)        D(I)

     1     3.333576e-04     1.000e+00
     2     5.000000e-02     1.000e+00
     3     5.000000e-02     1.000e+00

    IT   NF      F         RELDF    PRELDF    RELDX   STPPAR   D*STEP   NPRELDF
     0    1 -1.995e+03
     1    7 -1.998e+03  1.33e-03  2.12e-03  4.1e-04  2.5e+09  4.1e-05  2.68e+06
     2    8 -1.998e+03  3.86e-05  5.36e-05  4.0e-04  2.0e+00  4.1e-05  4.66e+01
     3    9 -1.998e+03  1.29e-05  1.32e-05  4.1e-04  2.0e+00  4.1e-05  4.60e+01
     4   16 -2.019e+03  1.07e-02  2.14e-02  5.0e-01  2.0e+00  1.0e-01  4.55e+01
     5   17 -2.025e+03  2.94e-03  4.05e-03  3.4e-01  2.0e+00  1.0e-01  1.73e-02
     6   20 -2.029e+03  2.06e-03  1.56e-03  4.9e-01  1.4e+00  3.0e-01  2.34e-02
     7   22 -2.030e+03  3.97e-04  4.72e-04  6.2e-02  2.0e+00  5.9e-02  8.47e+00
     8   23 -2.031e+03  4.06e-04  9.22e-04  5.5e-02  2.0e+00  5.9e-02  1.73e-01
     9   24 -2.032e+03  5.84e-04  6.57e-04  4.9e-02  2.0e+00  5.9e-02  4.02e-02
    10   26 -2.033e+03  5.59e-04  8.00e-04  9.8e-02  1.9e+00  1.4e-01  1.70e-02
    11   28 -2.033e+03  4.77e-05  2.60e-04  4.4e-03  1.6e+00  7.4e-03  1.15e-03
    12   29 -2.034e+03  2.13e-04  2.28e-04  4.7e-03  1.9e+00  7.4e-03  1.63e-03
    13   33 -2.036e+03  1.22e-03  1.42e-03  2.7e-01  0.0e+00  3.4e-01  1.42e-03
    14   34 -2.039e+03  1.38e-03  1.23e-03  1.6e-01  0.0e+00  1.6e-01  1.23e-03
    15   35 -2.040e+03  3.47e-04  2.39e-04  4.9e-02  0.0e+00  6.5e-02  2.39e-04
    16   36 -2.040e+03  1.45e-04  1.21e-04  4.8e-02  0.0e+00  5.7e-02  1.21e-04
    17   37 -2.040e+03  4.41e-05  3.25e-05  1.1e-02  0.0e+00  1.1e-02  3.25e-05
    18   38 -2.040e+03  2.19e-05  1.67e-05  7.7e-03  0.0e+00  9.5e-03  1.67e-05
    19   39 -2.040e+03  6.84e-06  5.71e-06  1.3e-03  0.0e+00  1.4e-03  5.71e-06
    20   40 -2.040e+03  8.15e-07  8.30e-07  1.4e-03  4.4e-01  1.4e-03  8.75e-07
    21   41 -2.040e+03  1.40e-07  1.52e-07  1.2e-03  0.0e+00  1.3e-03  1.52e-07
    22   42 -2.040e+03  4.81e-09  1.74e-09  1.0e-04  0.0e+00  1.1e-04  1.74e-09
    23   43 -2.040e+03  2.31e-09  8.85e-11  4.2e-05  0.0e+00  4.5e-05  8.85e-11
    24   44 -2.040e+03 -1.14e-10  4.14e-13  3.7e-06  0.0e+00  3.9e-06  4.14e-13

 ***** RELATIVE FUNCTION CONVERGENCE *****

 FUNCTION    -2.040382e+03   RELDX        3.670e-06
 FUNC. EVALS      44         GRAD. EVALS      24
 PRELDF       4.141e-13      NPRELDF      4.141e-13

     I      FINAL X(I)        D(I)          G(I)

     1    7.854092e-05     1.000e+00     2.637e+00
     2    3.054526e-01     1.000e+00     1.616e-04
     3    4.856587e-01     1.000e+00     9.901e-04
```

```
> summary(f3)

Call:
garch(x = x3, order = c(1, 1))

Model:
GARCH(1,1)

Residuals:
    Min      1Q  Median      3Q     Max
-6.2061 -0.4001  0.1577  0.7245  3.3447

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0 7.854e-05   1.552e-05    5.061 4.16e-07 ***
a1 3.055e-01   1.964e-02   15.551  < 2e-16 ***
b1 4.857e-01   5.442e-02    8.925  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals
X-squared = 258.2785, df = 2, p-value < 2.2e-16


        Box-Ljung test

data:  Squared.Residuals
X-squared = 1.8577, df = 1, p-value = 0.1729

> AIC(f3)
[1] -3023.497
```

```
> f3=garch(x4,order=c(1,1))

 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****


     I     INITIAL X(I)        D(I)

     1     2.739855e-04     1.000e+00
     2     5.000000e-02     1.000e+00
     3     5.000000e-02     1.000e+00

    IT   NF      F         RELDF    PRELDF    RELDX   STPPAR   D*STEP   NPRELDF
     0    1 -1.279e+03
     1    9 -1.279e+03  3.32e-06  7.01e-06  2.1e-05  2.0e+09  2.1e-06  6.97e+03
     2   10 -1.279e+03  6.45e-08  7.29e-08  2.1e-05  2.0e+00  2.1e-06  3.86e-01
     3   11 -1.279e+03  4.67e-08  5.96e-08  4.2e-05  2.2e+00  4.2e-06  3.86e-01
     4   19 -1.281e+03  1.72e-03  2.96e-03  5.5e-01  2.0e+00  1.2e-01  3.85e-01
     5   21 -1.282e+03  6.78e-04  6.79e-04  3.3e-01  1.6e+00  1.2e-01  1.71e-02
     6   22 -1.283e+03  7.31e-04  1.14e-03  4.0e-01  1.8e+00  2.4e-01  1.22e-02
     7   24 -1.283e+03  2.61e-04  6.28e-04  9.3e-02  1.6e+00  8.6e-02  9.64e-04
     8   26 -1.283e+03  2.81e-04  2.67e-04  7.8e-02  0.0e+00  8.6e-02  3.58e-04
     9   28 -1.284e+03  1.76e-04  3.72e-04  4.5e-02  1.7e+00  5.9e-02  2.55e-03
    10   30 -1.285e+03  9.19e-04  5.29e-04  9.3e-02  0.0e+00  1.5e-01  5.29e-04
    11   32 -1.285e+03  5.01e-04  7.81e-04  2.9e-02  1.8e+00  5.6e-02  3.96e-03
    12   33 -1.286e+03  1.09e-04  4.22e-04  2.6e-02  1.4e+00  5.6e-02  1.27e-03
    13   35 -1.286e+03  3.19e-04  1.99e-04  3.1e-02  3.0e-01  5.6e-02  5.26e-04
    14   45 -1.286e+03  2.64e-06  1.16e-04  6.3e-07  2.8e+00  1.1e-06  2.22e-01
    15   46 -1.286e+03  3.81e-05  3.78e-05  3.1e-07  2.0e+00  5.7e-07  2.52e-01
    16   51 -1.286e+03  1.47e-09  1.15e-07  1.4e-08  2.2e+00  2.6e-08  3.01e-01
    17   59 -1.286e+03 -2.25e-14  7.75e-15  3.9e-15  1.9e+00  7.1e-15 -1.12e-02

 ***** FALSE CONVERGENCE *****

 FUNCTION    -1.286041e+03   RELDX        3.908e-15
 FUNC. EVALS      59         GRAD. EVALS      17
 PRELDF       7.752e-15      NPRELDF     -1.116e-02

     I      FINAL X(I)        D(I)          G(I)

     1    9.720096e-06     1.000e+00    1.395e+03
     2    5.646959e-02     1.000e+00   -2.392e+01
     3    9.137573e-01     1.000e+00   -2.102e+01


> summary(f3)

Call:
garch(x = x4, order = c(1, 1))

Model:
GARCH(1,1)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.5603 -0.4042  0.2369  0.7975  3.3291 

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)    
a0 9.720e-06   8.763e-06    1.109  0.26732    
a1 5.647e-02   1.874e-02    3.013  0.00258 ** 
b1 9.138e-01   3.957e-02   23.091  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals 
X-squared = 104.9094, df = 2, p-value < 2.2e-16


        Box-Ljung test

data:  Squared.Residuals 
X-squared = 0.1621, df = 1, p-value = 0.6872

> AIC(f3)
[1] -1902.609
```

REFERENCES

[1] Ganeshi, U. , Krishna, M., & Hariprasad S. (2011). Review on Models for Generalized Predictive Controller. *CCSEA.* 418-424.

[2] Davidson, J. E. H., Hendry, D.F., Srba, F. , & Yeo, S. (1978). Econometric Modelling of the Aggregate Tim. *Economic Journal*. 661–692.

[3] Horvath, L., & Leipus, R. (2005). Effect of Aggregation On Estimators in AR(1) Sequence. *Preprint*.

[4] Brockwell, P. J., & Richard A. D. (1987).Time Series: Theory and Methods. *New York: Springer–Verlag.*

[5] Williams, C. (2010). Probabilistic Modelling and Reasoning Time Series Modelling: AR, MA, ARMA and All That. *School of Informatics University of Edinburgh.*

[6] Chatfield, C. (1989). The Analysis of Time Series: An Introduction. *Chapman and Hall, London.*

[7] Choi, H., Hernando, O., & Bonnie, R. (2008). Sequential Change-Point Detection Methods for Nonstationary Time Series. *Technometrics*. 40-53.

[8] Robert, E. (2001). The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives.* 157-168.

[9] Yamanishi, K., & Takeuchi, J.I. (2002). A Unifying Framework for Detecting Outliers and Change Points from Non-Stationary Time Series Data. *Proc. of KDD 2002* . 676-681.

[10] Engle, R. (2001). The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives*, Vol. 15, No. 4, 157-168.

[11] Bera, A. K., & Higgins, M.L. (1993). ARCH Models: Properties, Estimation and Testing. *Journal of Economic Surveys.* Vol. 7. No. 4. 307-366.

ACADEMIC VITA

JOYCE YUNJIE JIANG

Address: 921 Oakwood Ave., State College, PA 16803   Email: yoj5029@psu.edu

---

Education

Master of Applied Statistics, Penn State University, Spring 2013

Bachelor of Science in Statistics, Penn State University, Spring 2013

Bachelor of Science in Economics, Penn State University, Spring 2013

Honors and Awards

The Schreyer Honor's College, Penn State University

Graduate Assistantship, Penn State Research Hub

The National Honor Society

Related Experience

Summer Intern, Franklin Park Asset Management, Philadelphia, PA

Financial Analyst Intern, Citic Bank, Shanghai, China

Honor Thesis Research, Penn State Statistic Department, University Park, PA

Association Memberships/Activities

Hands-on Project Chair, The Schreyer Consulting Group

Co-founder and President, Penn State Speech Association

Participant, Global Engagement and Leadership Conference

Participant, PNC Leadership Assessment Program

Member, Penn State International Business Association

Member, Economics Association

Member, Penn State Martial Art Group