香 港 大 學

THE UNIVERSITY OF HONG KONG

STAT4601 Time-series Analysis
Group Project

Statistical modeling of customer price index in the US

Group 6 members:

Tsang Edmund Chiu Yat 3035667261
So Ka Ho (Kyle) 3035690048
Law Hoi Ting (Alexandra) 3035708649
Chan Kui Wun 3035687352

Abstract

Consumer Price Index (CPI), which is a price index of a weighted average market basket of consumer goods and services purchased by households, plays a crucial role in indicating a country's inflation level and economy. We modeled monthly CPI in the U.S. using the data from January, 2000 to May, 2019 in our report. We first carried out difference transformations to the raw data to make it stationary. Several seasonal ARIMA models were specified and the parameters are estimated, to forecast the CPI of the next 5 months. All forecasts were within 95% confidence intervals of our specified models. We later extended our project to include all data from 1980 to 2019 to suggest seasonal ARIMA models, but the prediction was not as satisfactory due to economic fluctuations in the 1980s. We later applied the Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model to predict the CPI in the coming 5 months. Evaluation and limitation of our predicting will be explored at the last part of the report.

Keywords: Consumer Price Index, Seasonal ARIMA, GARCH model

Contents

1 Introduction

A consumer price index (CPI) is a price index of a weighted mean market basket of consumer goods and services purchased by households, including but not limited to foods, transportation, housing and medical care. A CPI is not just a price index, it indicates Inflation, which is defined as the yearly percentage change in the CPI. A CPI is also applied to index the real price value of, for example, salaries and pensions to regulate prices. With its wide application, CPI is one of the most important indicators of the national economy. Apart from the overall index, CPI can also be computed in the sub-categories, to understand the price change of that particular category of consumer goods and services.

Seasonally Adjusted U.S. CPI from January, 1980 to October, 2022 is collected from FRED Economic. The data is already seasonally-adjusted based on weather, school year, production cycles, and holidays, and thus no additional seasonal adjustment is needed. Due to the significant scale of economic fluctuations observed during the 40-year observation period, including recession in the early 1980s, economic boom in the 1990s and the COVID-19 pandemic in the 2020s, we first include data from January, 2000 to May, 2019, with a total of 233 months, for training only. Corresponding models and forecasts are also explored with the dataset from 2000 to 2019.

Exploration is then extended to data from 1980 to 2019, for the sake of investigating how aforementioned economics fluctuations in the 1980s and 1990s affect the time-series analysis. Apart from the ARIMA models, we also employ the GARCH model, a statistical model in analyzing time-series data where the variance error is serially autocorrelated, in the analysis. Forecasting is made with ARIMA and the GARCH models.

This report will be arranged into 4 sections: Section 2 is the model for 2000-2019 data. Stationarity and seasonality of the collected data is explored in section 2.1. We will then cover model specification, diagnostic and parameters estimations in section 2.2, with respective suggested models, in which we will cover checking of models using Ljung-Box test and overparameterized method to check the adequacy of the fitted model. Forecasting will be conducted for each model in section 2.3. We will then discuss the criteria for model selection in section 2.4. Section 3 will be the model for 1980-2019 data. Structure of this section is similar to that of section 2, but we will talk about the GARCH model in section 3.4, before conducting forecasting in section 3.5. At last, section 4 will conclude the report.

2 Model for 2000-2019 dataset

2.1 Stationary and Seasonality

The first step of model fitting involves the test for stationarity. First, the time series plot of the original data is investigated carefully, and there is a clear sign of upward trend which is confirmed by the slow decaying sample ACF plot. On the contrary, the variance does not change over time, and thus log transformation is not necessary.

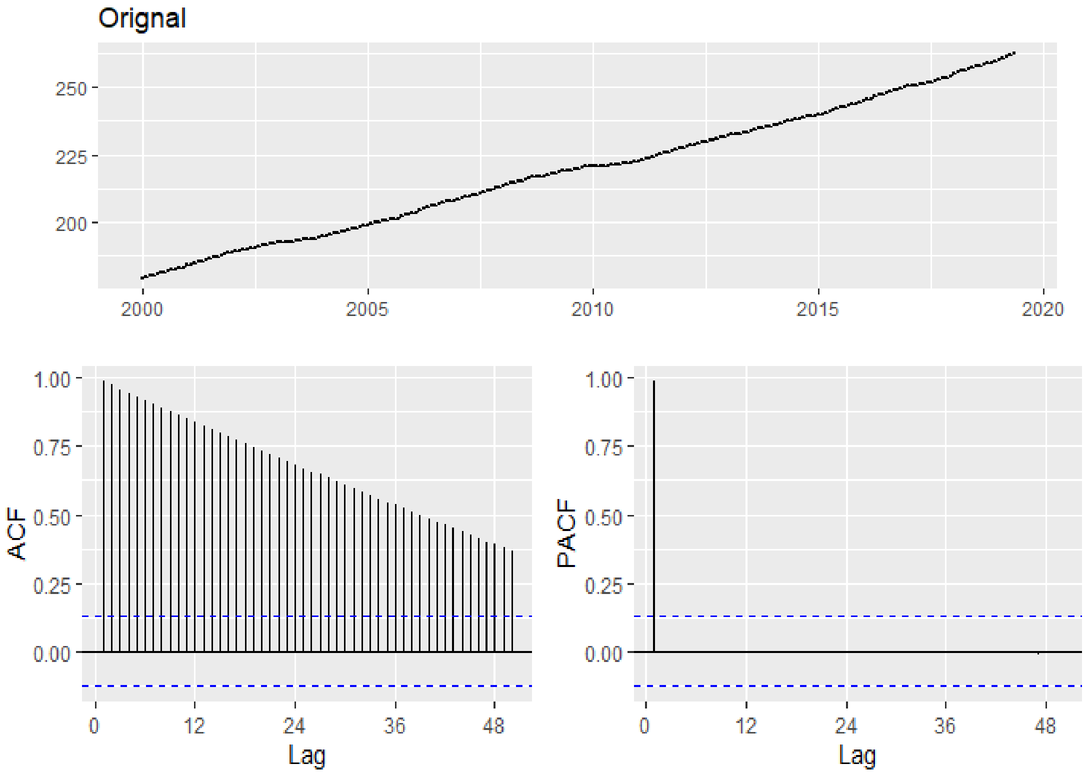


Fig. (1). The plot of time series (above); the ACF and PACF plots of the data (below)

To further confirm the non-stationarity of the original dataset, an Augmented Dickey Fuller Test (ADF Test) is conducted in Python, with the null hypothesis that there is a unit root present in a time series sample. The result gives a p-value of 0.9979, indicating that the data is non-stationary. As a result, we decided to take the first difference on the original dataset.

After first differencing, we observed that the original upward trend is removed in the plot of the data with the first difference, and no strong seasonal effect is observed in the ACF plot. The p-value of the ADF test is smaller than 0.01, suggesting that the null hypothesis is rejected. Therefore, we conclude that the time series has been transformed into a stationary series which can be used for model-building.

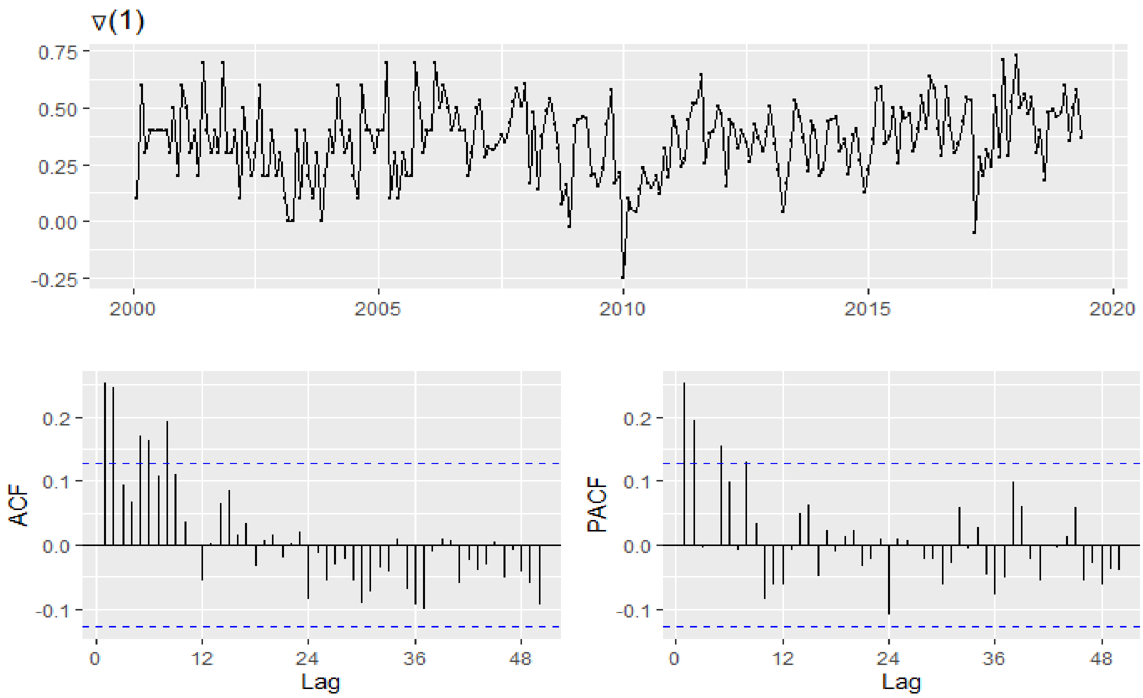


Fig. (2). The plot of time series of the first differencing (above); the ACF and PACF plots of the data (below)

2.2 Model specification, diagnostic and parameter estimation

2.2.1 Model 1 - ARIMA (1,1,1)

$$Z_t = \phi_1 Z_{t-1} + a_t - \theta_1 a_{t-1}$$

Based on the pattern of the ACF and PACF plot, we have noticed that there are several significant autocorrelations on the ACF plot and PACF plot, mainly at lag 1,2,5,6,8 for the ACF plot and lag 1,2,4 for the PACF plot. We decided to start with a simple ARIMA (1,1,1) and use it as our baseline model.

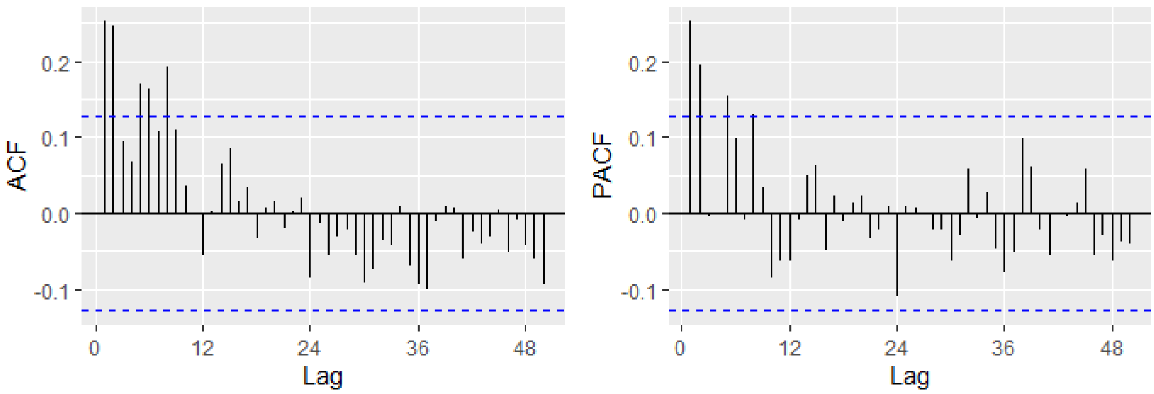


Fig. (3). The ACF and PACF plots of the data

Using Maximum Likelihood Estimation (MLE) to estimate Φ_p and θ_q. Both estimates, AR1 and MA1, are derived and proved to be statistically significant by the p-value in the Z test. Giving us the model :

$$Z_t = 0.9979Z_{t-1} + a_t - 0.8385a_{t-1} \quad \{a_t\} \sim WN(0,0.02572)$$

```
z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)
ar1  0.9978911  0.0028504 350.092 < 2.2e-16 ***
ma1 -0.8384612  0.0480553 -17.448 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
sigma^2 = 0.02572:  log likelihood = 95.53
AIC=-185.07   AICc=-184.96   BIC=-174.73
```

Fig. (4). Summary of the parameter estimation result and AIC and BIC

For the first step of the model diagnostics of the ARIMA(1,1,1) model, we plot the standard residuals sequence and look for possible trends and patterns. Below is the plot of the standard residuals sequence, where it seems to center around zero with constant variance. We cannot identify any special patterns or trends from the standard residuals sequence, and thus, we conclude that there is nothing special found from the residuals.

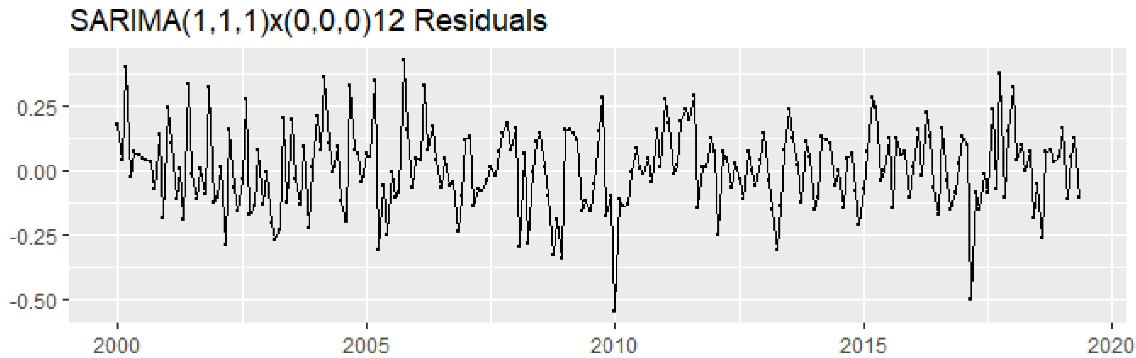


Fig. (5). The plot of residuals of ARIMA (1,1,1)

In the second step, we analyze the normality of residuals to confirm our belief that there is no trend or patterns in the standard residuals sequence. The unimodal and bell-shaped histogram on the left suggests that the residuals are approximately normally distributed. Also, the QQ plot of residuals on the right implies that the straight line is a good fit for the residuals. The Shapiro-Wilk normality test with the null hypothesis of: the residuals is normally distributed, yields a p-value of 0.2939, indicating that the residuals is normally distributed.

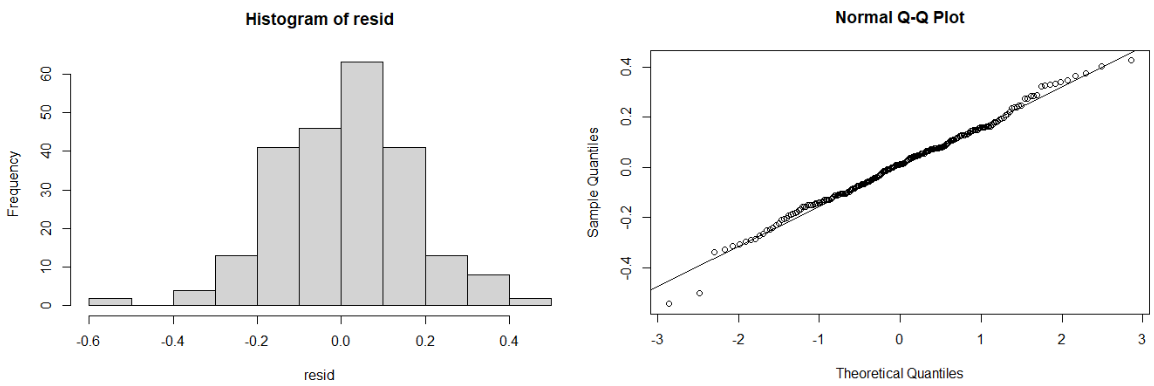


Fig. (6). The histogram of residuals (left); the normal Q-Q plot of residuals (right)

In the third step, we then extend our model diagnostic to the autocorrelations of residuals after fitting the baseline model. We choose K from 10 to 20 for the Ljung-Box test, and they all yield p-values larger than 0.05, suggesting that no autocorrelation within residuals. Also from the sample ACF plot, they are all inside the bound except for lag 12. We decided to ignore it as it is marginal and insignificant, hence we recognize this model as adequate.

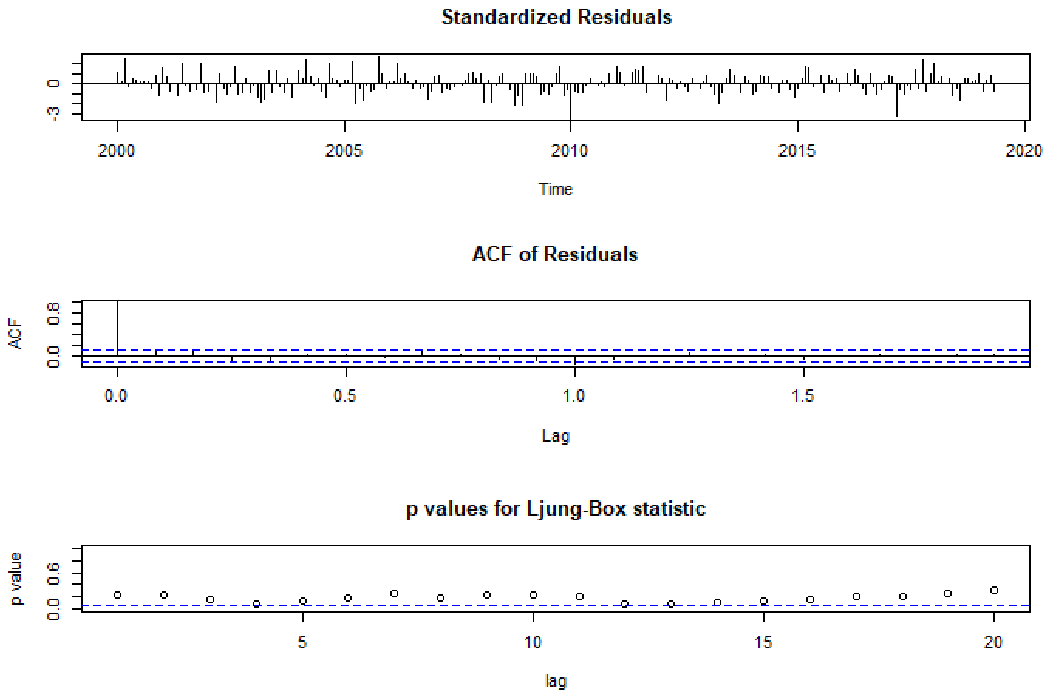


Fig. (7). The plots of standardized residuals (top), ACF of residuals (middle), p values for Ljung-Box statistics (bottom) of the ARIMA (1,1,1) model

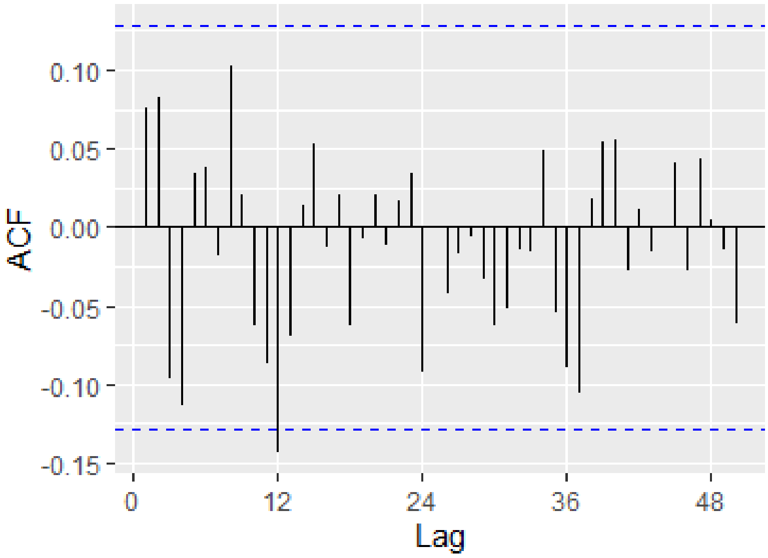


Fig. (8). The ACF of the sample standardized residuals of the ARIMA (1,1,1) model

The last step to confirm our baseline model is by using the over-parameterized method. When 1 is added to the MA part, giving us the overfitting model of ARIMA (1,1,2), the original estimates do not change to a large extent. And more importantly, the p-value of the newly added MA, namely MA2 is not significant and both AIC and BIC increase compared to original value.

```
z test of coefficients:

      Estimate Std. Error  z value Pr(>|z|)    
ar1  0.9984402  0.0022513 443.4860   <2e-16 ***
ma1 -0.7912790  0.0613238 -12.9033   <2e-16 ***
ma2 -0.0695084  0.0629975  -1.1034   0.2699    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

sigma^2 = 0.02569:  log likelihood = 96.16
AIC=-184.31   AICc=-184.13   BIC=-170.52
```

Fig. (9). Summary of the parameter estimation result and AIC and BIC of ARIMA (1,1,2)

On the other hand, when 1 is added to the AR part, suggesting another overfitting model of ARIMA (2,1,1), we observed that the newly added parameter is small in value and does not alter the original estimates to a large extent. Therefore, we conclude that the newly added AR, namely AR2 is not significant either.

```
z test of coefficients:

     Estimate Std. Error z value Pr(>|z|)    
ar1  1.111269   0.087619  12.683   <2e-16 ***
ar2 -0.112413   0.087007  -1.292   0.1964    
ma1 -0.886604   0.053516 -16.567   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

sigma^2 = 0.02563:  log likelihood = 96.38
AIC=-184.76   AICc=-184.58   BIC=-170.97
```

Fig. (10). Summary of the parameter estimation result and AIC and BIC of ARIMA (2,1,1)

2.2.2 Model 2 - ARIMA (2,1,8)

Apart from the ARIMA (1,1,1) model we have discussed in session 2.2.1, ARIMA (2,1,8) is another possible model, as it is observed that there is a significant correlation at lag 8 of the ACF plot. Since we do not take any seasonal differencing, we decided to consider the PACF plot in the hope of covering more signals than the ARIMA (1,1,1) model by introducing a ARIMA (2,1,8) model.

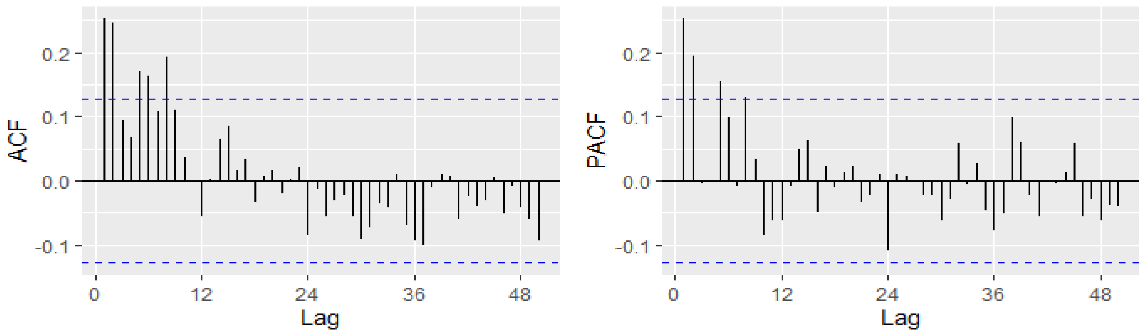


Fig. (11). The ACF and PACF plots of the data

By using the Maximum Likelihood Estimation (MLE) to estimate Φ_p and θ_q. We obtain

$$Z_t = 0.352105Z_{t-1} + 0.647468Z_{t-2} + a_t - 0.137594a_{t-1} - 0.47017a_{t-2} - 0.193427a_{t-3} - 0.156648a_{t-4} + 0.132859a_{t-5} + 0.064408a_{t-6} - 0.124872a_{t-7} - 0.021564a_{t-8} \quad \{a_t\} \sim WN(0,0.02517)$$

```
z test of coefficients:

     Estimate Std. Error z value Pr(>|z|)
ar1  0.352105   0.479632  0.7341  0.46288
ar2  0.647468   0.479624  1.3499  0.17703
ma1 -0.137594   0.476482 -0.2888  0.77276
ma2 -0.470170   0.386692 -1.2159  0.22403
ma3 -0.193427   0.075700 -2.5552  0.01061 *
ma4 -0.156648   0.124647 -1.2567  0.20885
ma5  0.132859   0.089567  1.4834  0.13798
ma6  0.064408   0.112190  0.5741  0.56590
ma7 -0.124872   0.068653 -1.8189  0.06893 .
ma8 -0.021564   0.104508 -0.2063  0.83653
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
sigma^2 = 0.02517:  log likelihood = 101.73
AIC=-181.47   AICc=-180.27   BIC=-143.55
```

Fig. (12). Summary of the parameter estimation result and AIC and BIC of ARIMA (2,1,8)

In the first step of the model diagnostics for the ARIMA(2,1,8) model, we plot the residuals and we observed that the trend and variance are nearly constant. No particular pattern is observed in the residual plot.

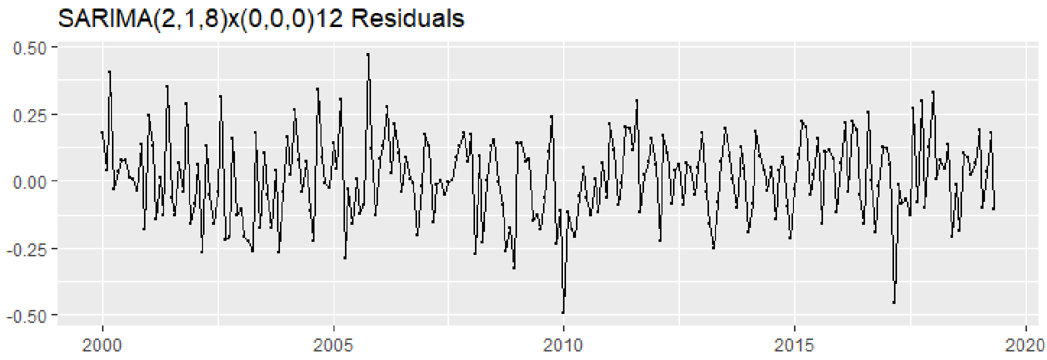


Fig. (13). The plot of residuals of ARIMA (2,1,8)

In the second step, we then move on to analyze the normality of residuals. The unimodal and bell-shaped histogram implies that the residuals are approaching normally distributed. The QQ plot of residuals suggests that the straight line is a good fit for the residuals. The Shapiro-Wilk normality test yields a p-value of 0.8072, which is larger than 0.05. Thus, we assume that the residuals follow normal distribution.

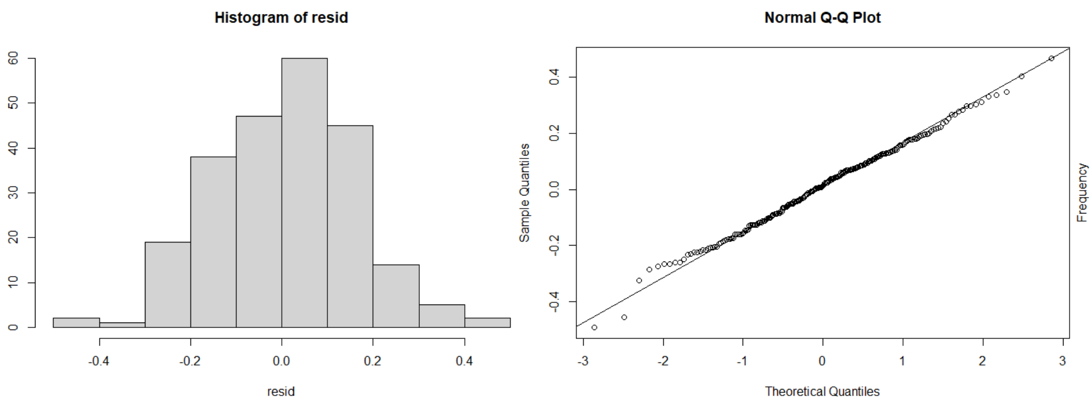


Fig. (14). The histogram of residuals (left); the normal Q-Q plot of residuals (right)

In the third step, we then extend our model diagnostic to the autocorrelations of residuals after fitting the baseline model by choosing K from 10 to 20 for the Ljung-Box test, and they all yield p-values larger than 0.05, suggesting that no autocorrelation within residuals. From the sample ACF plot, all lags are inside the bound, and hence we recognize this model as adequate.

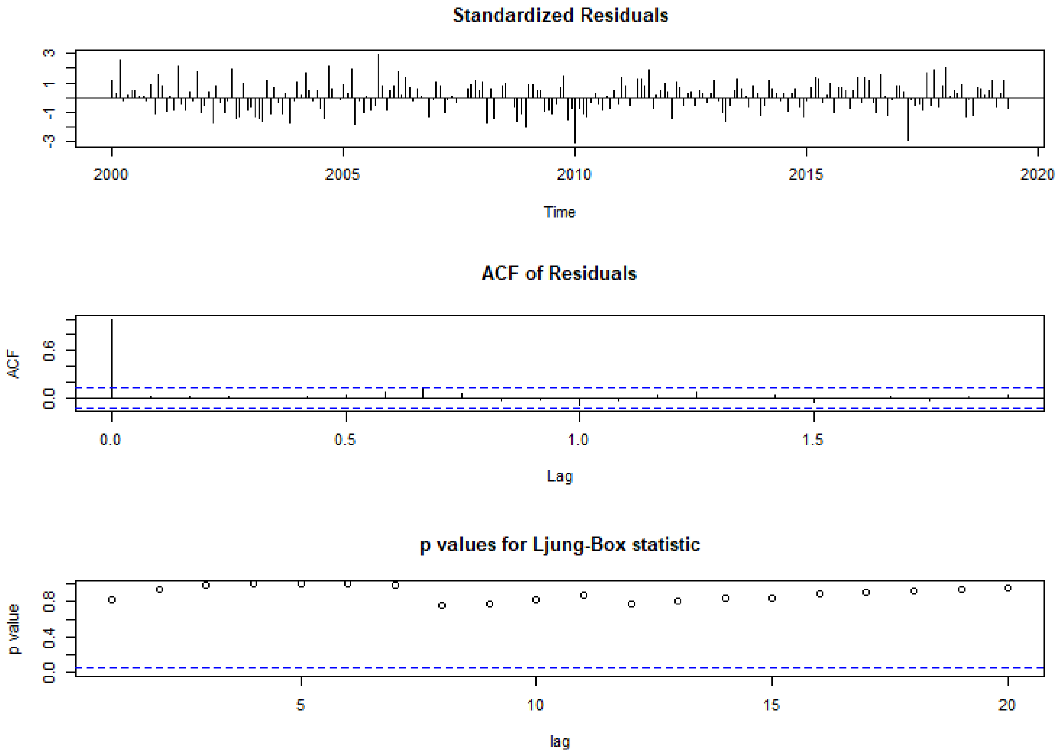


Fig. (15). The plots of standardized residuals (top), ACF of residuals (middle), p values for Ljung-Box statistics (bottom) for the ARIMA (2,1,8) model

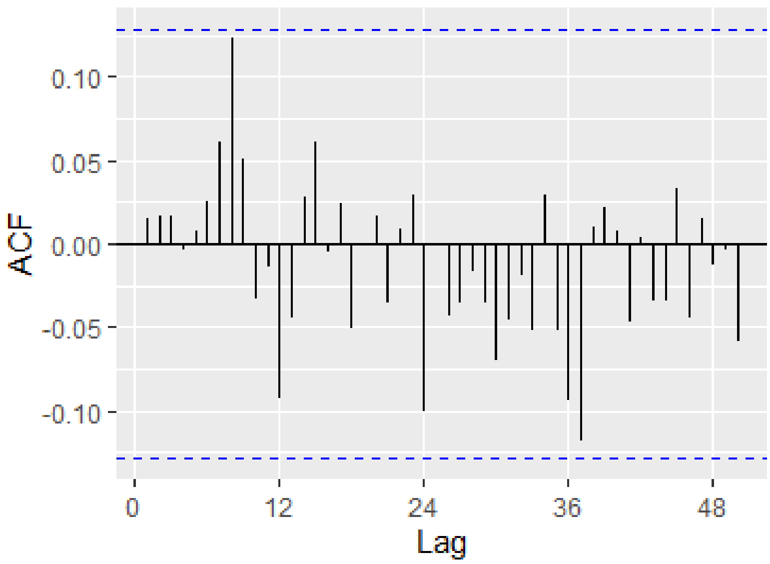


Fig. (16). The ACF of the sample standardized residuals of the ARIMA (2,1,8) model

The final step to confirm the model is by the over-parameterized method. We added MA1 and AR1 to the ARIMA(2,1,8) model and when 1 is added to the AR, i.e. the overfitting model of ARIMA (3,1,8), the p-value of the new added parameter is significant and both AIC and BIC decrease compared to those of ARIMA (2,1,8). All evidence suggests that ARIMA (3,1,8) may be a better model compared to ARIMA (2,1,8). Therefore, we will further investigate ARIMA (3,1,8) in the next session.

$$
\begin{aligned}
Z_t = & -0.143408Z_{t-1} + 0.187209Z_{t-2} + 0.954320Z_{t-3} + a_t + 0.372874a_{t-1} + 0.129545a_{t-2} \\
& - 0.894321a_{t-3} - 0.229691a_{t-4} - 0.123273a_{t-5} + 0.042096a_{t-6} \\
& + 0.015973a_{t-7} - 0.021365a_{t-8} \quad \{a_t\} \sim WN(0,0.025)
\end{aligned}
$$

```
z test of coefficients:

      Estimate Std. Error  z value  Pr(>|z|)    
ar1 -0.143408   0.066712  -2.1497  0.031583 *  
ar2  0.187209   0.059651   3.1384  0.001699 ** 
ar3  0.954320   0.057390  16.6287 < 2.2e-16 ***
ma1  0.372874   0.093368   3.9936 6.508e-05 ***
ma2  0.129545   0.090435   1.4325  0.152010    
ma3 -0.894321   0.085785 -10.4251 < 2.2e-16 ***
ma4 -0.229691   0.098910  -2.3222  0.020220 *  
ma5 -0.123273   0.105538  -1.1680  0.242789    
ma6  0.042096   0.100519   0.4188  0.675374    
ma7  0.015973   0.087969   0.1816  0.855920    
ma8 -0.021365   0.075334  -0.2836  0.776709    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
sigma^2 = 0.025:  log likelihood = 102.94
AIC=-181.88   AICc=-180.45   BIC=-140.52
```

Fig. (17). Summary of the parameter estimation result and AIC and BIC of ARIMA (3,1,8)

2.2.3 Model 3 - ARIMA (3,1,8)

In the first step of the model diagnostics for the ARIMA(3,1,8) model, we again plot the residuals and we observed constant trend and variance. We conclude that the residual plot does not contain any special patterns.

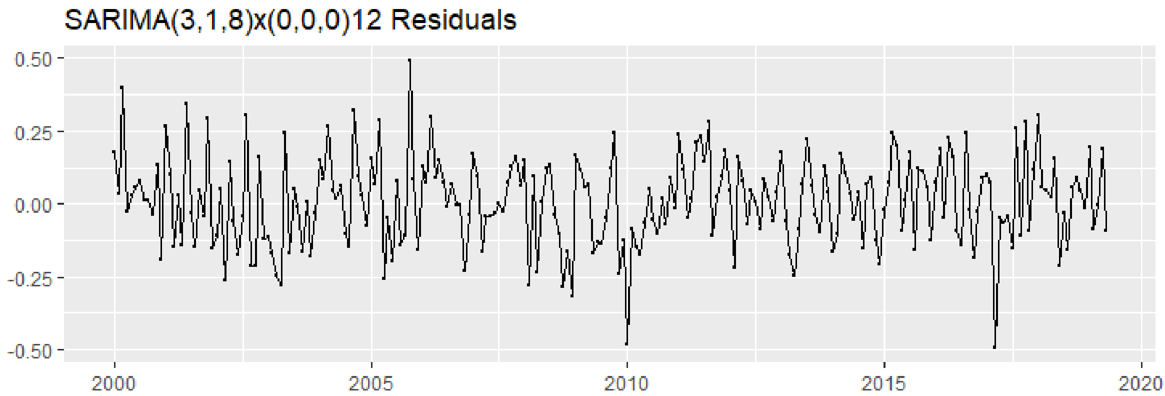


Fig. (18). The plot of residuals of ARIMA (3,1,8)

In the second step, the normality of residuals are analyzed. The unimodal and bell-shaped histogram on the left suggests that the residuals are approximately normally distributed. Also, the QQ plot of residuals on the right implies that the straight line is a good fit for the residuals. The Shapiro-Wilk normality test yields a p-value of 0.5552, which is larger than 0.05. Therefore, residuals are assumed to follow normal distribution.

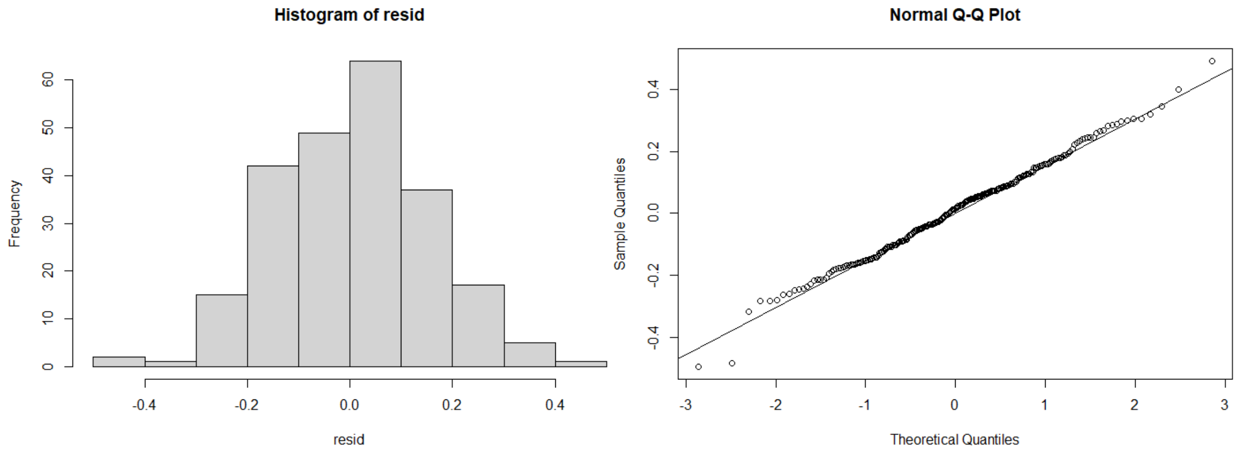


Fig. (19). The histogram of residuals (left); the normal Q-Q plot of residuals (right)

We then extend our model diagnostic to the autocorrelations of residuals after fitting the baseline model. Ljung-Box test is performed for the first 20 lag, which all yield p-values larger than 0.05, suggesting that no autocorrelation within residuals. From the sample ACF plot, all lags are inside the bound, and hence we recognize this model as adequate.

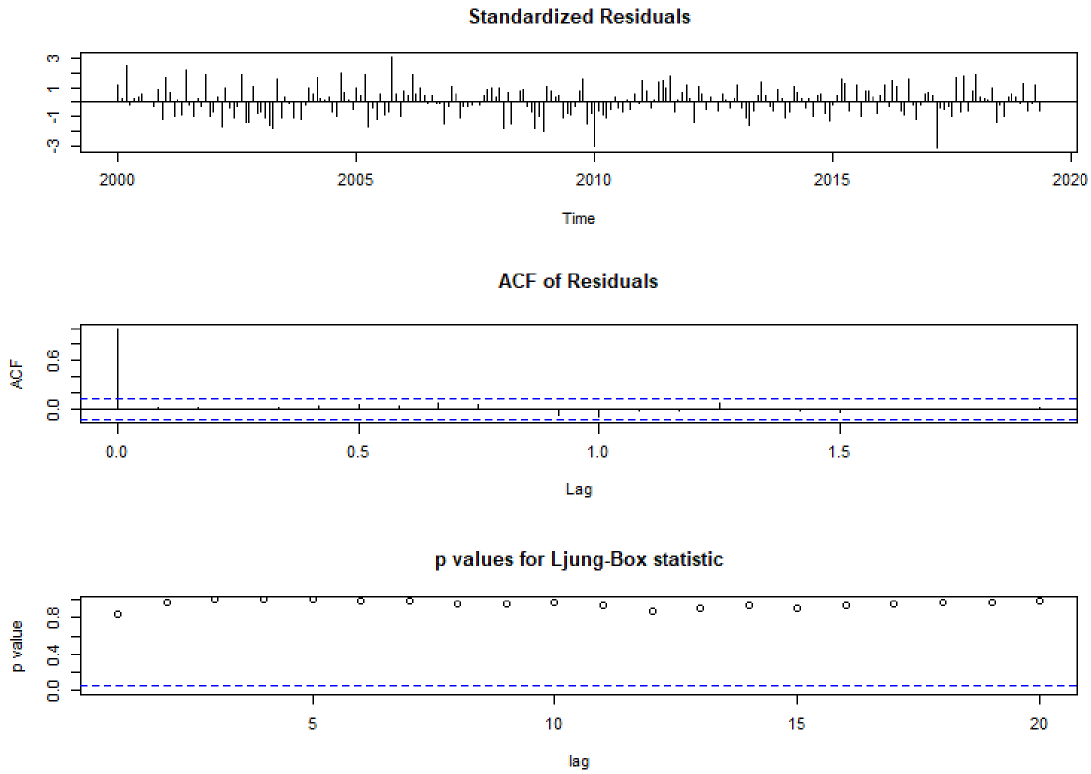


Fig. (20). The plots of standardized residuals (top), ACF of residuals (middle), p values for Ljung-Box statistics (bottom) for the ARIMA (3,1,8) models

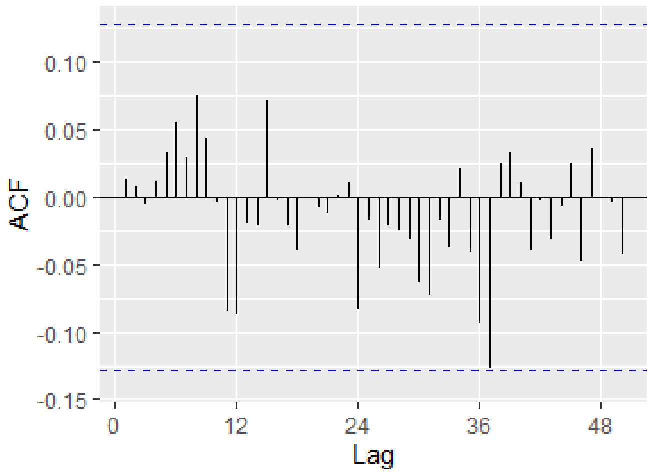


Fig. (21). The ACF of the sample standardized residuals of the ARIMA (3,1,8) model

2.3 Forcasting

2.3.1 Forecasting with ARIMA (1,1,1) model

The ARIMA (1,1,1) model gives a satisfactory prediction. All five actual data are within the 95% confidence interval of the predicted value. The mean square error between the predicted value and the actual value is 0.14751679.

Actual value

```
         Jun     Jul     Aug     Sep     Oct
2019 262.945 263.491 264.135 264.588 265.019
```

Fig. (22). The actual value for June, 2019 to October, 2019

	Forecasting value	95% forecasting limit	
		Lo 95	Hi 95
Jun 2019	262.7723	262.4580	263.0867
Jul 2019	263.2227	262.7415	263.7040
Aug 2019	263.6722	263.0371	264.3073
Sep 2019	264.1207	263.3339	264.9074
Oct 2019	264.5682	263.6284	265.5081

Fig. (23). The forecasting value and 95% forecasting limit with ARIMA (1,1,1) model for June, 2019 to October, 2019

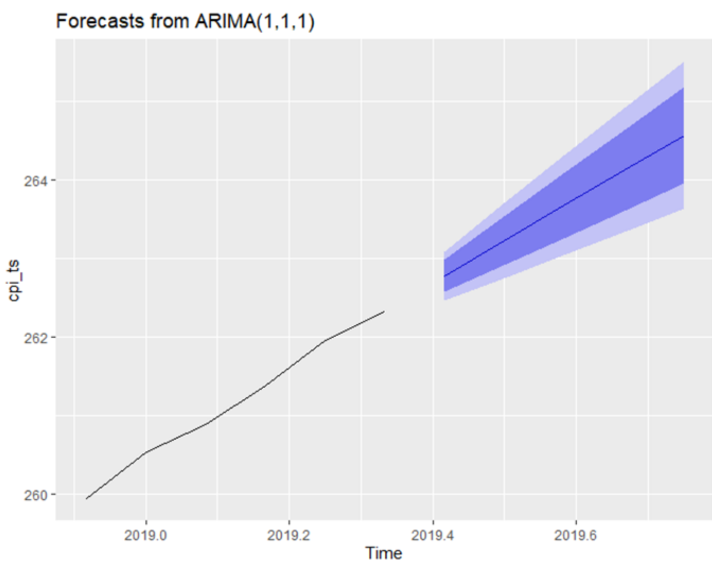


Fig. (24). The plot of forecasting value and 95% forecasting limit (blue) with ARIMA (1,1,1) model for June, 2019 to October, 2019

2.3.2 Forecasting with ARIMA (2,1,8) model

The ARIMA (2,1,8) model also generates a satisfactory prediction. All five actual data are within the 95% confidence interval of the predicted value. The mean square error between the predicted value and the actual value is 0.20291811, which is slightly larger than that of ARIMA(1,1,1).

	Point Forecast	95% forecasting limit	
		Lo 95	Hi 95
Jun 2019	262.7864	262.4754	263.0973
Jul 2019	263.1853	262.6961	263.6745
Aug 2019	263.5927	262.9237	264.2617
Sep 2019	264.0458	263.2300	264.8616
Oct 2019	264.4641	263.5211	265.4071

Fig. (25). The forecasting value and 95% forecasting limit with ARIMA (2,1,8) model for June, 2019 to October, 2019

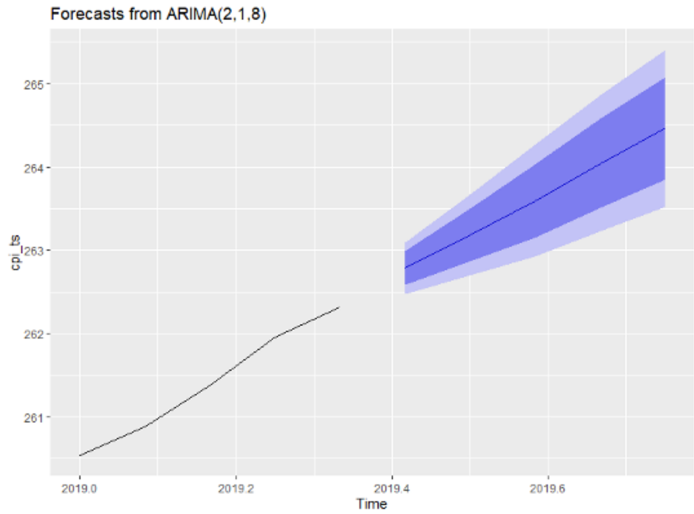


Fig. (26). The plot of forecasting value and 95% forecasting limit (blue) with ARIMA (2,1,8) model for June, 2019 to October, 2019

2.3.3 Forecasting with ARIMA (3,1,8) model

The ARIMA (3,1,8) model also generates a satisfactory prediction. All five actual data are within the 95% confidence interval of the predicted value. The mean square error between the predicted value and the actual value is 0.14840698, which is slightly larger than that of ARIMA(1,1,1), but it is smaller than that of ARIMA(2,1,8).

	Point Forecast	95% forecasting limit	
		Lo 95	Hi 95
Jun 2019	262.8411	262.5312	263.1510
Jul 2019	263.2224	262.7313	263.7136
Aug 2019	263.6491	262.9700	264.3283
Sep 2019	264.1474	263.3110	264.9839
Oct 2019	264.5406	263.5668	265.5144

Fig. (27). The forecasting value and 95% forecasting limit with ARIMA (3,1,8) model for June, 2019 to October, 2019

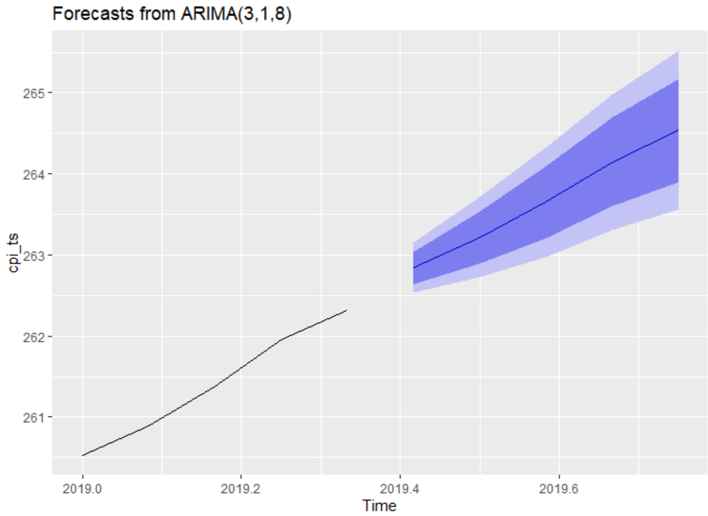


Fig. (28). The plot of forecasting value and 95% forecasting limit (blue) with ARIMA (3,1,8) model for June, 2019 to October, 2019

2.4 Model selection

The table below compares all 3 proposed models for CPI data extracted from 2000-2019.

Models	ARIMA (1,1,1)	ARIMA (3,1,8)	ARIMA (2,1,8)
AIC	-185.07	-181.88	-180.27
BIC	-174.73	-180.45	-143.55
Forecasting MSE	0.1475	0.1484	0.2029

From the table above, we observe that the ARIMA (1,1,1) model has the lowest AIC and BIC among the three models that we have proposed. Moreover, when comparing the forecasting accuracy, we observed that ARIMA (1,1,1) has the lowest mean square error. Following the principle of Parsimony, and the result in the above table, we decided to select ARIMA (1,1,1) as our parsimonious, best-fitted model for the 2000-2019 CPI dataset.

3 Model for 1980-2019 datasets

3.1 Stationary

With a view to increase the robustness and generalizability of the model, in this part, more data, which demonstrates strong time-varying volatility clustering effect, will be included in the training and parameter estimation. Based on previous data, the time series data is extended from 1980 to 2019, in which (event). The GARCH model will be applied to describe the variance error that is believed to be serially auto-correlated (literature).

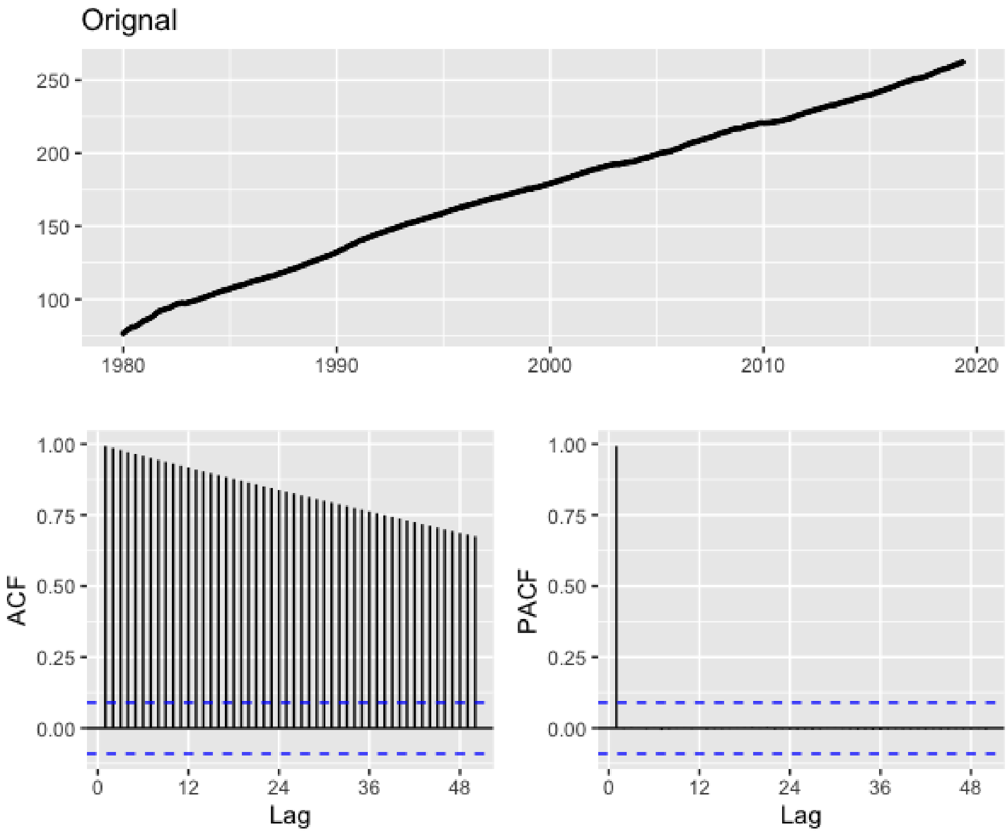


Fig. (29). The plot of time series of data from 1980 to 2019 (above); the ACF and PACF plots of the data (below)

Above shows the original time series plot from 1980/01/01 to 2019/05/01, as well as the ACF and PACF plot of the original time series plot. Given that the original time series plot demonstrates a upward increasing trend and the ACF of that shows a slowly decaying pattern, the time series is indicated as non-stationary, which is supported by the ACF test with a p-value equal to 0.998, with no evidence to reject the null hypothesis that a unit root is present in a time series sample. Therefore, differencing is considered to be necessary to attain stationary.

```
        Augmented Dickey-Fuller Test

data:  cpi_ts.rdiff
Dickey-Fuller = -4.3795, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

Fig. (30). The ADF Test results

According to the ADF test, the p-value is less than 0.01, meaning that the null hypothesis of non-stationarity is rejected. Hence, we will adopt first order differences to achieve stationarity conditions. Based on the time series plot, it is obvious to see neither the increasing dispersity or auto-variance that are associated with the increased level of the series nor the seasonal influence that appears in the ACF plot, thus logarithm transformation and 12-months seasonal difference is not considered.

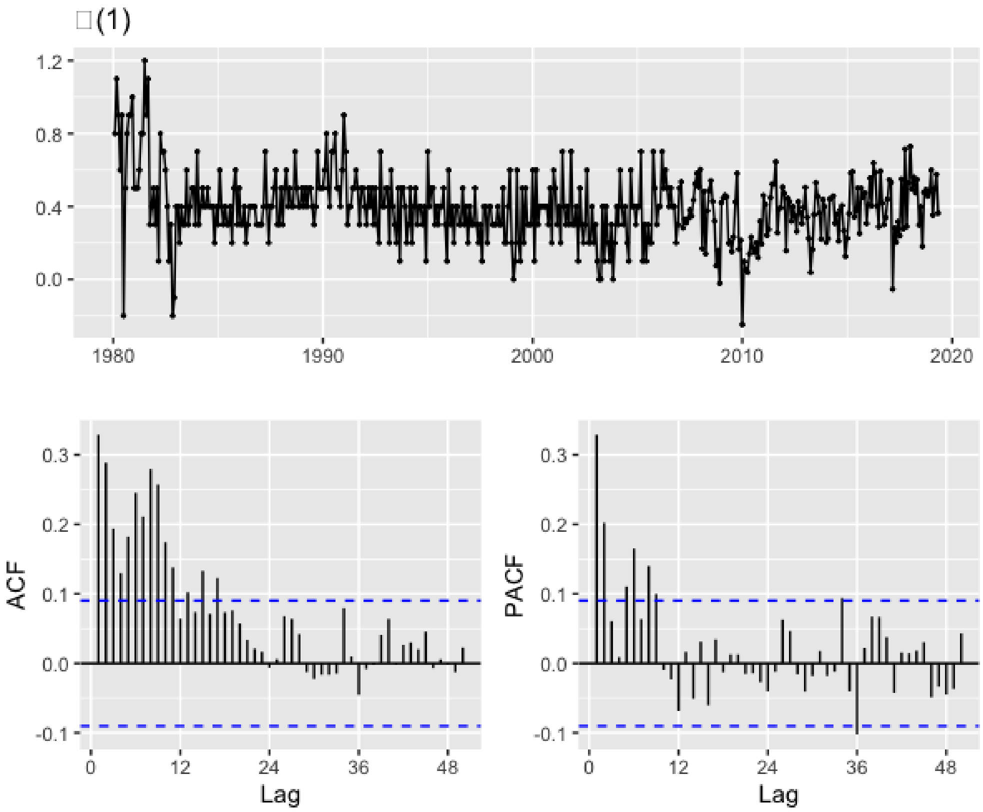


Fig. (31). The plot of time series of the first differencing (above); the ACF and PACF plots of the data (below)

Observed from the ACF plot, we can observe the cut-off effect is approximately at lag 11. Hence, proposing an ARIMA (0,1,11) model is appropriate. Given that the signal at lag 9 is also beyond the 95% confidence interval, we can propose an ARIMA (9,1,0) as the second model.

3.2 Model specification, diagnostic and parameter estimation

3.2.1 Model 1: ARIMA (0,1,11)

```
> coeftest(model_fit)

z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)
ma1   0.204774   0.046164  4.4358 9.173e-06 ***
ma2   0.208717   0.047314  4.4113 1.027e-05 ***
ma3   0.112245   0.049333  2.2753  0.022890 *
ma4   0.014406   0.048859  0.2948  0.768112
ma5   0.081498   0.049003  1.6631  0.096289 .
ma6   0.138625   0.050958  2.7204  0.006520 **
ma7   0.067752   0.050052  1.3536  0.175850
ma8   0.206041   0.047786  4.3117 1.620e-05 ***
ma9   0.200457   0.051143  3.9195 8.872e-05 ***
ma10  0.132643   0.051380  2.5816  0.009834 **
ma11  0.101039   0.050923  1.9842  0.047238 *
drift 0.397315   0.018720 21.2241 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> model_fit
Series: cpi_ts
ARIMA(0,1,11) with drift

Coefficients:
         ma1     ma2     ma3     ma4     ma5     ma6     ma7     ma8     ma9    ma10    ma11   drift
      0.2048  0.2087  0.1122  0.0144  0.0815  0.1386  0.0678  0.2060  0.2005  0.1326  0.1010  0.3973
s.e.  0.0462  0.0473  0.0493  0.0489  0.0490  0.0510  0.0501  0.0478  0.0511  0.0514  0.0509  0.0187

sigma^2 = 0.02824:  log likelihood = 177.84
AIC=-329.69   AICc=-328.9   BIC=-275.65
```

Fig. (32). Summary of the parameter estimation result and AIC and BIC of ARIMA (0,1,11)

As the MLE method shown above, most except for 3 coefficients of ARIMA(0,1,11) are statistically significant, which indicates the estimates are all expected to have fair impacts on the model and the standard errors are all reasonably small.

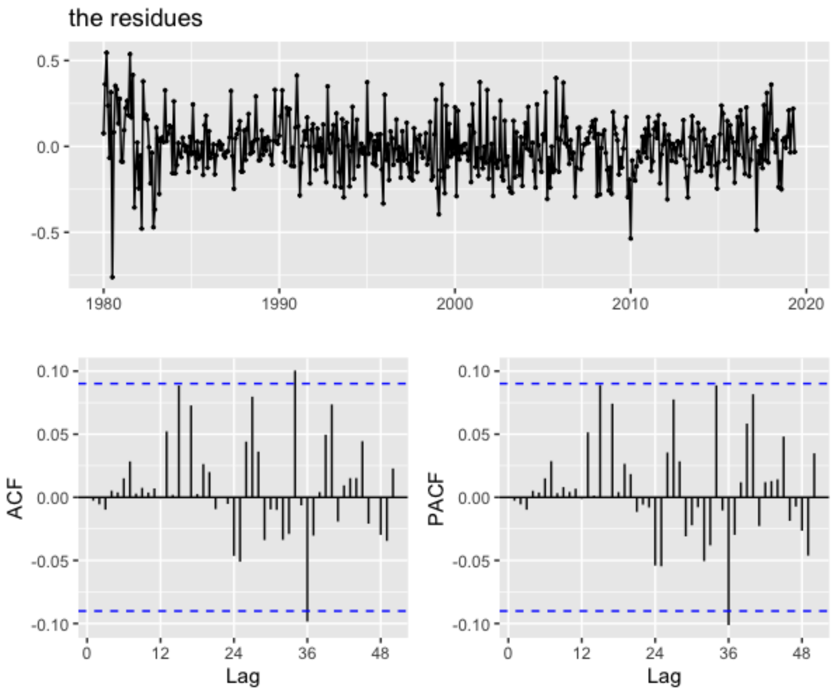


Fig. (33). The plot of residuals (top), ACF (bottom left) and PACF (bottom right)

In the first step of model diagnostics, we firstly check the autocorrelation of the residuals. As we can see from the above ACF and PACF plots, the ACF and PACF of the residuals are mostly within the 95% of confidence interval boundaries which implies that the residuals do not possess significant autocorrelation with any lags. Since the signals that exceed the 95% confidence interval in the ACF plots are marginal, it is believed that the autocorrelations in these lags can be seens as coincidence and hence be ignored. This is supported by the Ljung-Box test illustrated below, in which the test statistics are all higher than 0.05, which means we have no evidence to reject the null hypothesis of independent distribution of the residuals.

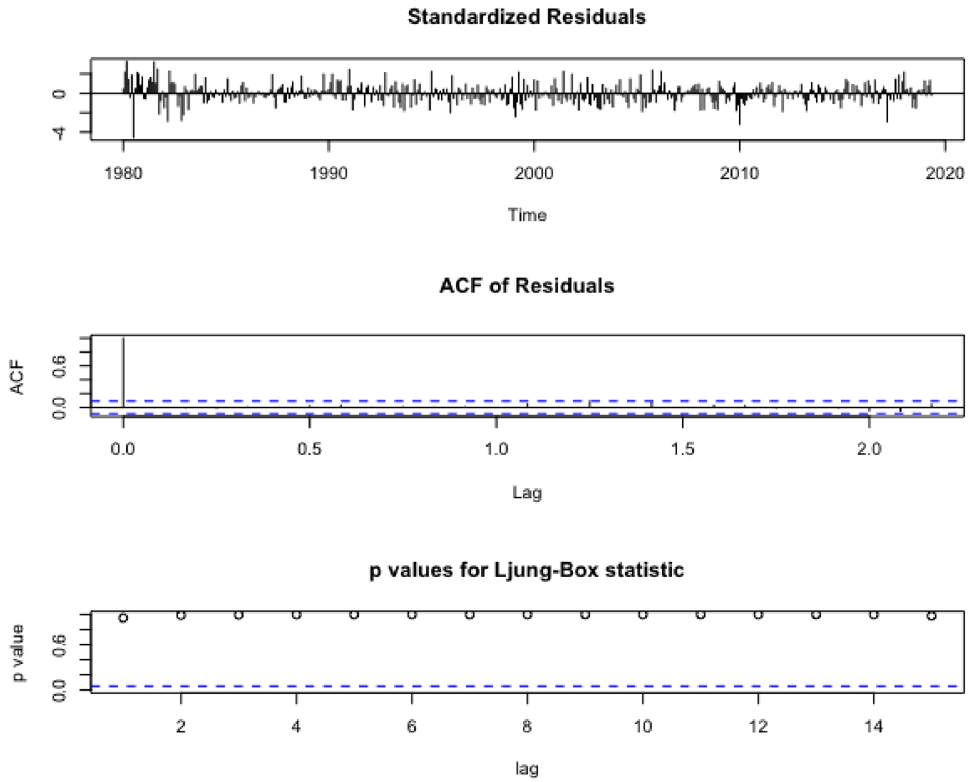


Fig. (34). The plots of standardized residuals (top), ACF of residuals (middle), p values for Ljung-Box statistics (bottom)

Next, the normality of the residuals will be checked. According to the histogram, it is shown that it is not very skewed (-0.05730657) but with a higher kurtosis (4.214159) compared to the normal distribution. This is supported by both the QQ plot and shapiro test. The QQ plot shows that the tails deviates from the theoretical quantiles quite a lot, meaning that it has more data peaked in the middle and fatter tails. With this line of thinking, we can see that the residuals do not pass the shapiro test with a p-value less than 0.01, which rejects the null hypothesis of following normal distribution.

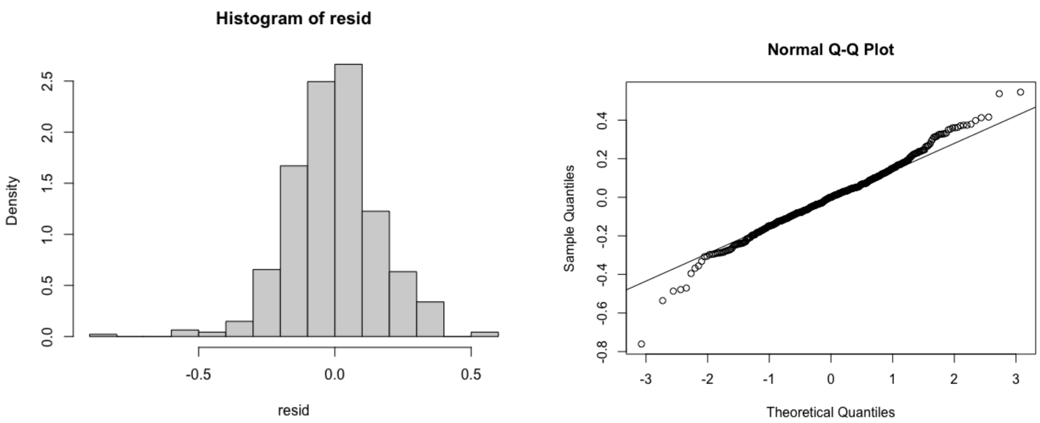


Fig. (35). The histogram of residuals (left); the normal Q-Q plot of residuals (right)

```
> shapiro.test(resid) #H0: the variable is normally distributed

	Shapiro-Wilk normality test

data:  resid
W = 0.98798, p-value = 0.0006166
```

Fig. (36). The result of Shapiro-Wilk normality test

The final step to confirm the model is the over-parameterized method. When 1 is added to the AR, i.e. the overfitting model of ARIMA (1,1,11), the p-value of the new added parameter is significant and the AIC decreases a little. However, it is noteworthy that the addition of the AR part makes a bunch of coefficients of the MA part insignificant, so adding one AR parameter is not considered. In addition, adding one more MA parameter gives a small coefficient and insignificant test result, so it also does not take into consideration. Eventually, we conclude that ARIMA(0,1,11) is a fairly adequate model for our data. The details of the analysis is shown below.

```
> coeftest(model_fit)

z test of coefficients:

       Estimate Std. Error  z value Pr(>|z|)    
ar1    0.9659395  0.0358348  26.9553  < 2e-16 ***
ma1   -0.7797721  0.0591980 -13.1723  < 2e-16 ***
ma2    0.0143591  0.0595448   0.2411  0.80944    
ma3   -0.0931295  0.0602913  -1.5447  0.12243    
ma4   -0.0659146  0.0589236  -1.1186  0.26329    
ma5    0.0650300  0.0606192   1.0728  0.28338    
ma6    0.0846869  0.0657255   1.2885  0.19757    
ma7   -0.0778440  0.0673327  -1.1561  0.24764    
ma8    0.1373742  0.0694180   1.9789  0.04782 *  
ma9   -0.0091208  0.0699305  -0.1304  0.89623    
ma10  -0.0573280  0.0661045  -0.8672  0.38581    
ma11  -0.0560100  0.0515344  -1.0868  0.27710    
drift  0.4094374  0.0375289  10.9099  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> autoplot(model_fit)
> model_fit
Series: cpi_ts 
ARIMA(1,1,11) with drift 

Coefficients:
         ar1      ma1     ma2      ma3      ma4     ma5     ma6      ma7     ma8      ma9     ma10
      0.9659  -0.7798  0.0144  -0.0931  -0.0659  0.0650  0.0847  -0.0778  0.1374  -0.0091  -0.0573
s.e.  0.0358   0.0592  0.0595   0.0603   0.0589  0.0606  0.0657   0.0673  0.0694   0.0699   0.0661
         ma11   drift
      -0.0560  0.4094
s.e.   0.0515  0.0375

sigma^2 = 0.02807:  log likelihood = 179.73
AIC=-331.45   AICc=-330.53   BIC=-273.26
```

```
> model_fit = Arima(cpi_ts, order = c(0,1,12), seasonal = c(0,0,0), lambda = NULL, include.constant = TRUE)
> coeftest(model_fit)

z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)    
ma1    0.204082   0.046001  4.4364 9.146e-06 ***
ma2    0.205739   0.047500  4.3313 1.482e-05 ***
ma3    0.113157   0.049278  2.2963 0.0216583 *  
ma4    0.011748   0.049398  0.2378 0.8120224    
ma5    0.085119   0.049663  1.7139 0.0865414 .  
ma6    0.135436   0.050458  2.6842 0.0072713 ** 
ma7    0.067600   0.049914  1.3543 0.1756342    
ma8    0.205875   0.047284  4.3540 1.337e-05 ***
ma9    0.197552   0.050991  3.8743 0.0001069 ***
ma10   0.127485   0.052510  2.4278 0.0151891 *  
ma11   0.100245   0.050616  1.9805 0.0476473 *  
ma12  -0.022813   0.048837 -0.4671 0.6404164    
drift  0.397278   0.018439 21.5450 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> model_fit
Series: cpi_ts 
ARIMA(0,1,12) with drift 

Coefficients:
         ma1     ma2     ma3     ma4     ma5     ma6     ma7     ma8     ma9    ma10    ma11     ma12
      0.2041  0.2057  0.1132  0.0117  0.0851  0.1354  0.0676  0.2059  0.1976  0.1275  0.1002  -0.0228
s.e.  0.0460  0.0475  0.0493  0.0494  0.0497  0.0505  0.0499  0.0473  0.0510  0.0525  0.0506   0.0488
       drift
      0.3973
s.e.  0.0184

sigma^2 = 0.02829:  log likelihood = 177.95
AIC=-327.91   AICc=-326.99   BIC=-269.71
```

Fig. (37). The result of overparameterization by fitting ARIMA(1,1,11) (left) or ARIMA (0,1,11) (right)

3.2.2 Model 2: ARIMA(9,1,0)

```
> coeftest(model_fit)

z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)    
ar1    0.187795   0.045668  4.1122  3.92e-05 ***
ar2    0.158481   0.046380  3.4170 0.0006331 ***
ar3    0.017848   0.046915  0.3804 0.7036157    
ar4   -0.050512   0.046667 -1.0824 0.2790823    
ar5    0.070572   0.047077  1.4991 0.1338498    
ar6    0.120121   0.048427  2.4804 0.0131218 *  
ar7    0.013156   0.048709  0.2701 0.7870830    
ar8    0.138827   0.048178  2.8815 0.0039574 ** 
ar9    0.122737   0.047985  2.5578 0.0105333 *  
drift  0.406995   0.033675 12.0859 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> model_fit
Series: cpi_ts 
ARIMA(9,1,0) with drift 

Coefficients:
         ar1     ar2     ar3      ar4     ar5     ar6     ar7     ar8     ar9   drift
      0.1878  0.1585  0.0178  -0.0505  0.0706  0.1201  0.0132  0.1388  0.1227  0.4070
s.e.  0.0457  0.0464  0.0469   0.0467  0.0471  0.0484  0.0487  0.0482  0.0480  0.0337

sigma^2 = 0.02783:  log likelihood = 180.26
AIC=-338.51   AICc=-337.94   BIC=-292.78
> ggtsdisplay(model_fit$residuals, lag.max = 50, main='the residues')
```

Fig. (38). Summary of the parameter estimation result and AIC and BIC of ARIMA (9,1,0)

The above figure shows the coefficients of the ARIMA(9,1,0) model by maximum likelihood estimation method. We can see that most of the coefficients are significant except for 4 of them. Refitting with fewer parameters does not make more coefficients become significant but lead to more signals in the residual ACF plot, as shown below.

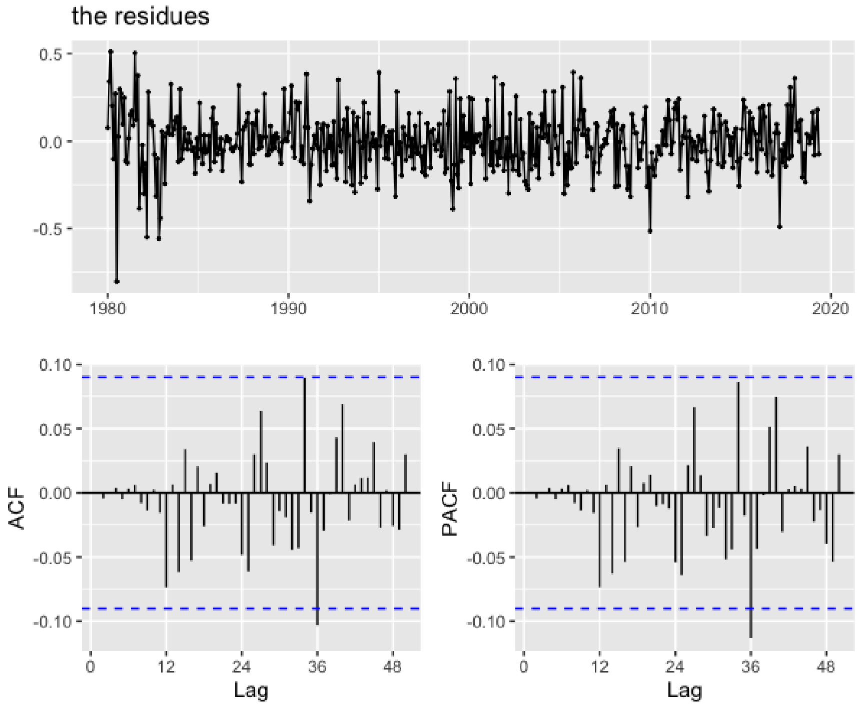


Fig. (39). The plot of time series residuals after fitting into ARIMA(9,1,0) (above) The ACF and PACF plot of the the residuals (below)

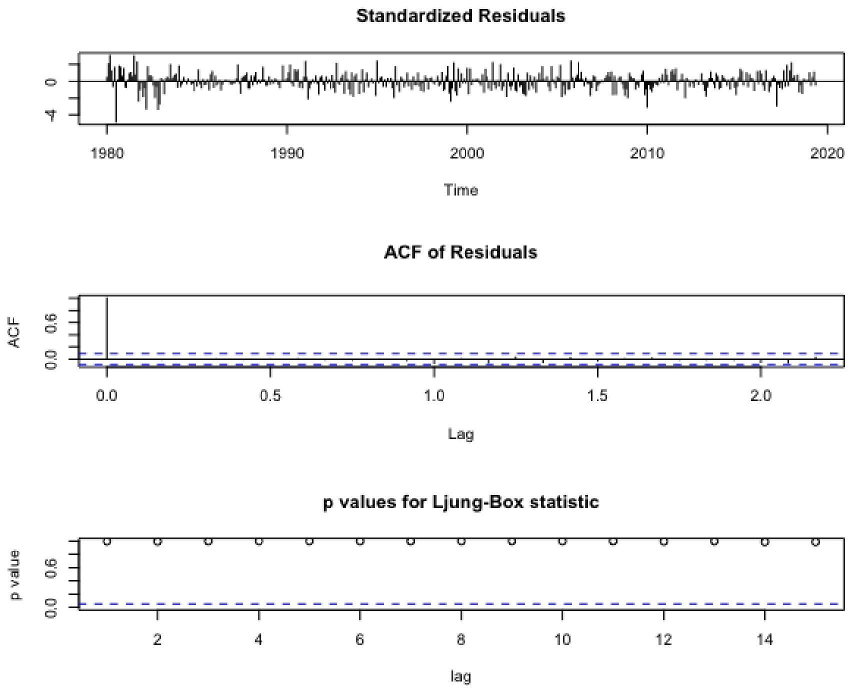


Fig. (40). The plots of standardized residuals (top), ACF of residuals (middle), p values for Ljung-Box statistics (bottom)

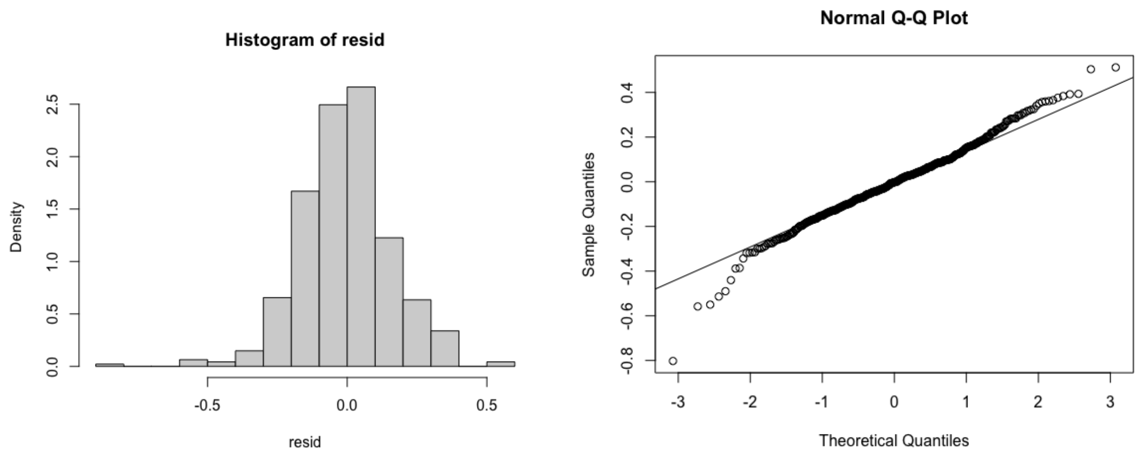


Fig. (41). The histogram of residuals (left); the normal Q-Q plot of residuals (right)

The residual analysis based on the above figures reveals that ACF are mostly within the confidence interval and Ljung-Box tests at different lags are all larger than 0.05, which implies that the residuals are independently distributed. As for the normality test, the skewness and kurtosis of the residuals are -0.2274237 and 4.512347 respectively, which can be observed in the heavy-tailed QQ plot. This observation of non-normality is also supported by the shapiro test with a p-value of 8.006e-05, which strongly rejects the normality assumption.

```
> coeftest(model_fit)

z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)
ar1    0.152666   0.363903  0.4195  0.674834
ar2    0.165693   0.087487  1.8939  0.058236 .
ar3    0.023482   0.074521  0.3151  0.752684
ar4   -0.049469   0.047662 -1.0379  0.299308
ar5    0.069055   0.049348  1.3993  0.161709
ar6    0.122521   0.054090  2.2651  0.023504 *
ar7    0.017528   0.066224  0.2647  0.791260
ar8    0.139960   0.049343  2.8365  0.004561 **
ar9    0.128460   0.074610  1.7217  0.085115 .
ma1    0.035616   0.366340  0.0972  0.922552
drift  0.406847   0.033528 12.1346 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> model_fit
Series: cpi_ts
ARIMA(9,1,1) with drift

Coefficients:
         ar1     ar2     ar3      ar4     ar5     ar6     ar7     ar8     ar9     ma1   drift
      0.1527  0.1657  0.0235  -0.0495  0.0691  0.1225  0.0175  0.1400  0.1285  0.0356  0.4068
s.e.  0.3639  0.0875  0.0745   0.0477  0.0493  0.0541  0.0662  0.0493  0.0746  0.3663  0.0335

sigma^2 = 0.02789:  log likelihood = 180.26
AIC=-336.52   AICc=-335.84   BIC=-286.64
```

```
> coeftest(model_fit)

z test of coefficients:

        Estimate Std. Error z value  Pr(>|z|)
ar1    0.1883925  0.0459800  4.0973  4.18e-05 ***
ar2    0.1591399  0.0467624  3.4032 0.0006661 ***
ar3    0.0178297  0.0469143  0.3800 0.7039096
ar4   -0.0500122  0.0468827 -1.0668 0.2860843
ar5    0.0709241  0.0471877  1.5030 0.1328338
ar6    0.1199773  0.0484424  2.4767 0.0132604 *
ar7    0.0132676  0.0487181  0.2723 0.7853660
ar8    0.1396476  0.0487474  2.8647 0.0041738 **
ar9    0.1236720  0.0487281  2.5380 0.0111487 *
ar10  -0.0053452  0.0484857 -0.1102 0.9122173
drift  0.4068119  0.0334941 12.1458 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> model_fit
Series: cpi_ts
ARIMA(10,1,0) with drift

Coefficients:
         ar1     ar2     ar3      ar4     ar5     ar6     ar7     ar8     ar9     ar10   drift
      0.1884  0.1591  0.0178  -0.0500  0.0709  0.1200  0.0133  0.1396  0.1237  -0.0053  0.4068
s.e.  0.0460  0.0468  0.0469   0.0469  0.0472  0.0484  0.0487  0.0487  0.0487   0.0485  0.0335

sigma^2 = 0.02789:  log likelihood = 180.26
AIC=-336.52   AICc=-335.84   BIC=-286.64
```

Fig. (42). The result of overparameterization by fitting ARIMA(9,1,1) (left) or ARIMA (10,1,0) (right)

Finally, the result of over-parameterizing the model in Fig. (42) by adding one more AR or MA part does not give a better fitting on the data. ARIMA(9,1,1) model greatly lowers the significance of the interpretability of other parameters, meaning that it cannot be a better model under consideration. Adding additional AR parameters gives a very small coefficient of -0.00523 with a p-value of 0.9122, so ARIMA(10,1,0) is rejected as well. In summary, the model 2 ARIMA (9,1,0) is considered as a fairly fitted model at the moment.

3.3 Model selection

Based on the AIC and BIC from the ARIMA(0,1,11) and ARIMA(9,1,0) from the above discussion, we can see that for both AIC and BIC, the ARIMA(9,1,0) has a lower value overall, meaning that it is relatively better model to consider.

```
Series: cpi_ts
ARIMA(0,1,11) with drift

Coefficients:
         ma1     ma2     ma3     ma4     ma5     ma6     ma7     ma8     ma9    ma10    ma11   drift
      0.2048  0.2087  0.1122  0.0144  0.0815  0.1386  0.0678  0.2060  0.2005  0.1326  0.1010  0.3973
s.e.  0.0462  0.0473  0.0493  0.0489  0.0490  0.0510  0.0501  0.0478  0.0511  0.0514  0.0509  0.0187

sigma^2 = 0.02824:  log likelihood = 177.84
AIC=-329.69   AICc=-328.9   BIC=-275.65
```

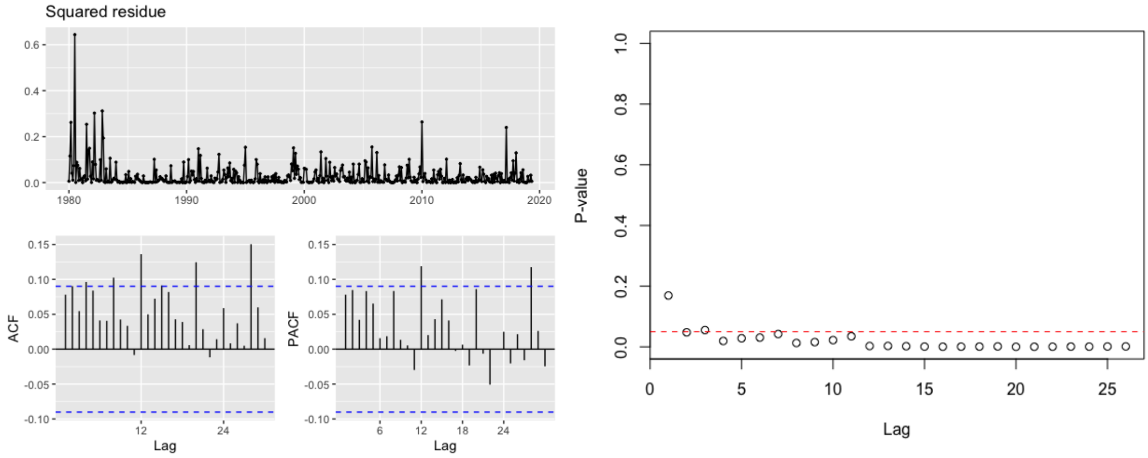
```
Series: cpi_ts
ARIMA(9,1,0) with drift

Coefficients:
         ar1     ar2     ar3      ar4     ar5     ar6     ar7     ar8     ar9   drift
      0.1878  0.1585  0.0178  -0.0505  0.0706  0.1201  0.0132  0.1388  0.1227  0.4070
s.e.  0.0457  0.0464  0.0469   0.0467  0.0471  0.0484  0.0487  0.0482  0.0480  0.0337

sigma^2 = 0.02783:  log likelihood = 180.26
AIC=-338.51   AICc=-337.94   BIC=-292.78
> ggtsdisplay(model_fit$residuals, lag.max = 50, main='the residues')
```

Fig. (43). Summary of the parameter estimation result and AIC and BIC

3.4 GARCH model




```
> ArchTest(resid)

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  resid
Chi-squared = 31.044, df = 12, p-value = 0.00194
```

Fig. (44). The plot of squared residuals, the ACF and PACF (left), the result of McLeod-Li test(right) and the Arch-LM test (below)

Given the failure in the normality test, it is thus suggested that a modal can be fitted in the residuals. Notice that, in figure 44, the residuals demonstrate a high volatility in around 1980 and 2019, here we propose to apply an generalized autoregressive conditional heteroscedasticity (GARCH) model for modeling the changing variance of a time series. In order to determine the presence of the ARCH effect in the residuals, squared-residual ACF, McLeod-Li test and Arch-LM test are carried out. As shown in squared-residual ACF, there are some signals that are out of the 95% confidence interval, with which we can conclude that there are ARCH effect in the residuals and this conclusion is in agreement with the McLeod-Li test result, in which all the p-values except at lag 1 are less than 95% meaning that the null hypothesis of no heteroskedasticity is rejected. The Arch-LM test also gives a p-value 0.00194, which is less than 0.05, so the null hypothesis of no arch effect in the residuals is rejected.

```
 data ~ 1 + garch(1, 1)
<environment: 0x7fb69399d6d0>
 [data = as.numeric(residuals(model_fit))]

Conditional Distribution:
 std

Coefficient(s):
        mu       omega      alpha1       beta1       shape
-0.0042789   0.0017211   0.0624848   0.8707348  10.0000000

Std. Errors:
 based on Hessian

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.004279    0.006927   -0.618   0.5367
omega   0.001721    0.001090    1.579   0.1143
alpha1  0.062485    0.032075    1.948   0.0514 .
beta1   0.870735    0.061227   14.221   <2e-16 ***
shape  10.000000    4.350870    2.298   0.0215 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log Likelihood:
 197.4275    normalized:  0.4173943

Description:
 Tue Dec  6 03:12:50 2022 by user:


Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  9.090644  0.01061675
 Shapiro-Wilk Test  R    W      0.9927909 0.02251892
 Ljung-Box Test     R    Q(10)  2.565553  0.9898856
 Ljung-Box Test     R    Q(15)  7.62858   0.9378091
 Ljung-Box Test     R    Q(20)  9.655786  0.9739927
 Ljung-Box Test     R^2  Q(10)  3.08055   0.9794643
 Ljung-Box Test     R^2  Q(15)  7.70143   0.9351941
 Ljung-Box Test     R^2  Q(20)  9.7551    0.9723983
 LM Arch Test       R    TR^2   9.849168  0.6291912

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
-0.8136470 -0.7696819 -0.8138674 -0.7963546
```

$$r_t = -0.0043 + a_t \quad a_t = \sigma_t \varepsilon_t \quad \sigma_t^2 = 0.0017 + 0.0625\sigma_{t-1}^2 + 0.8787\sigma_{t-2}^2 \quad \varepsilon_t \overset{iid}{\sim} \mathrm{t}(10)$$

Fig.(45). The GARCH(1,1) model and its statistical tests results

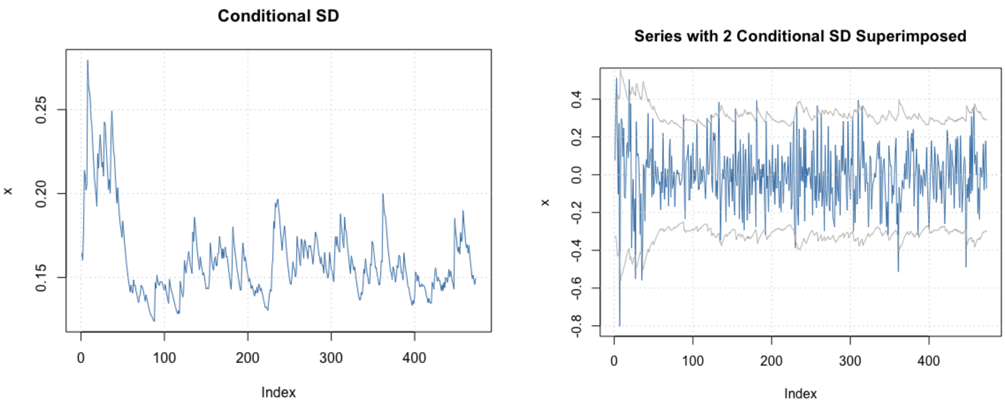


Fig. (46). The fitted volatility of the residuals and the pointwise 95% prediction interval plot

After fitting the ARIMA(9,1,0)-Garch(1,1) model as shown in Fig. 44 and 46, we can see that the presence of heteroscedasticity in residuals from ARIMA(9,1,0) especially in 1980 and the fluctuation of volatility throughout the years. In addition, from the one-step-ahead forecasts are generally within the pointwise 95% prediction interval, indicating our model can capture the heteroscedasticity and give satisfactory predictions.

Next, we will do the model diagnostics. First of all, as shown in Fig. 45, the standardized residuals are stationary and proved by Ljung-Box test with a p-value 0.974 with 20 lags and ACF plot, there are no autocorrelations across different lags with 95% confidence. This means that they are independently distributed.

From Fig. 44, we can see that the ACF of the standardized residuals do not have any signals. This is in agreement with the Ljung-Box test of R in Fig. 45, which gives a p-value of 0.974 with 20 lags. This means that the residuals are independently distributed.

With respect to the normality of residuals, since we observe a heavy-tailed distribution of the residuals after fitted into ARIMA(9,1,0), the residuals are modeled with t distribution. Therefore, the Jarque-Bera test and Shapiro-Wilk test have little reference value. With that being said, the QQ plot shows a much less deviation from normality compared to that of residuals from ARIMA(9,1,0).

Finally, we need to check if there are any ARCH effects remaining in the residuals of the GARCH(1,1) model. By the Arch-LM test as shown in Fig. 44, the p-value of 0.629 indicates there is no ARCH effect remaining in the residuals. The ACF of squared standardized residuals in Fig.44 and Ljung-Box test of R^2 in Fig. 45 also support that there is no presence of heteroscedasticity.

In conclusion, the ARIMA(9,1,0)-GARCH(1,1) model is adequate to fit in the data.

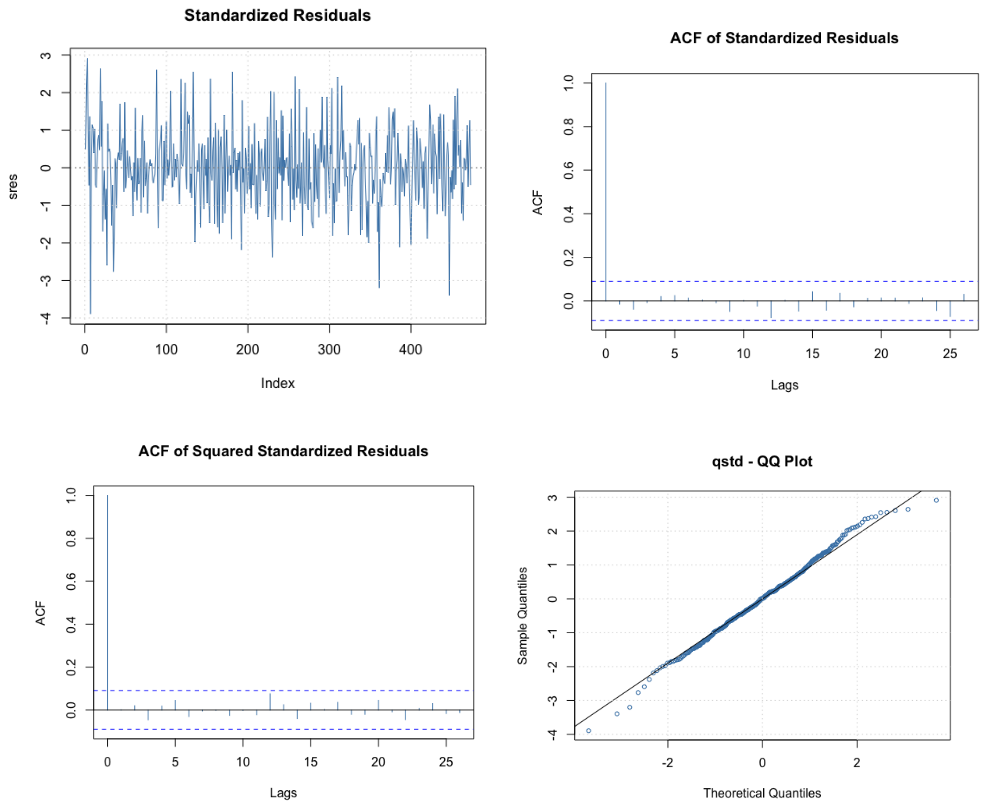


Fig. (47). The plot of standard residuals, ACF, ACF of squared residuals, and QQ plot

3.5 Forecasting

The following figures show the forecasting results of residuals based on the GARCH model. We observe that the mean is the same and the variance-confidence interval is different, we move on to employ the ARIMA model to make forecasting.

	Mean Forecast	Mean Error	Standard Deviation	Lower Interval	Upper Interval
1	-0.004278918	0.1465272	0.1465272	-0.2962941	0.2877362
2	-0.004278918	0.1475044	0.1475044	-0.2982416	0.2896838
3	-0.004278918	0.1484106	0.1484106	-0.3000475	.2914897
4	-0.004278918	0.1492513	0.1492513	-0.3017229	0.2931651
5	-0.004278918	0.1500316	0.1500316	-0.3032780	0.2947202

Fig. (48). Prediction of the residuals

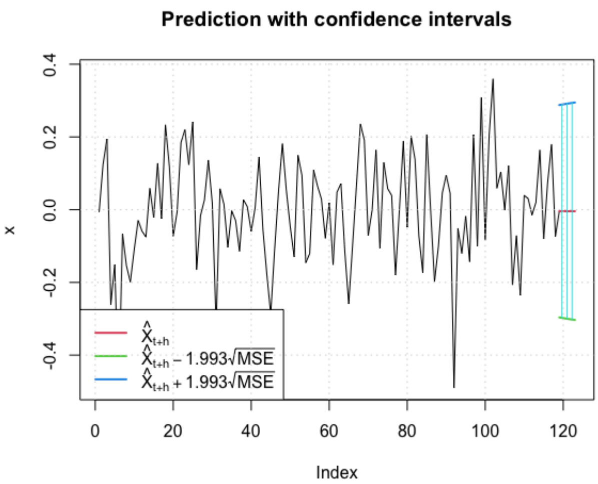


Fig. (49). Prediction plot of the residuals from GARCH(1,1)

The following figures show the forecasting results of residuals based on the ARIMA model. We observe that all the actual values are within the 95% confidence interval of the estimation.

	Point Forecast	Low 80	High 80	Low 95	High 95
Jun 2019	262.7943	262.5805	263.0081	262.4673	263.1212
Jul 2019	263.2429	262.9110	263.5749	262.7353	263.7506
Aug 2019	263.6773	263.2330	264.1217	262.9978	264.3569
Sep 2019	264.1571	263.6135	264.7008	263.3257	264.9886
Oct 2019	264.6138	263.9864	265.2412	263.6543	265.5734

Fig. (50). Prediction of the residuals forecast of ARIMA(9,1,0)

Overall	Forecast result
Jun 2019	262.7900
Jul 2019	263.2386
Aug 2019	263.6730
Sep 2019	264.1528
Oct 2019	264.6095

Fig. (51). The forecast result of ARIMA(9,1,0)

We then minus the value from the forecast of the GARCH model, and the result is shown below in fig. 52.

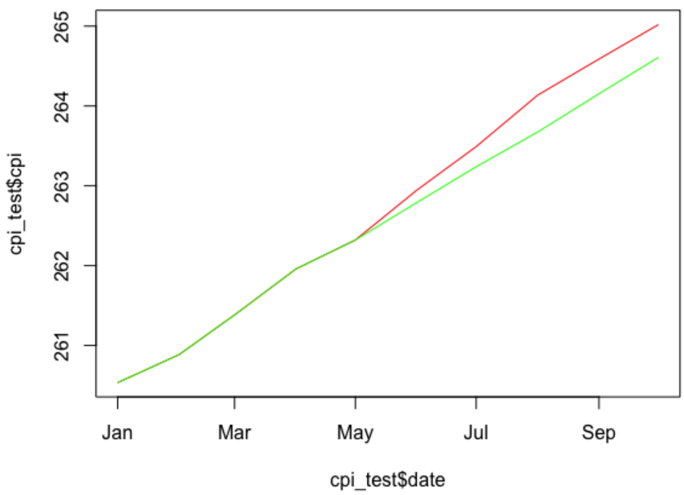


Fig. (52). Plot of the forecast of ARIMA(9,1,0)-GARCH(1,1). Green: forecast, Red: actual data

4 Conclusion

For the monthly Consumer Price Index in the U.S from 2000-01-01 to 2019-10-01:

1. We specified an ARIMA (1,1,1) model and estimated its parameters as:

$$Z_t = 0.9979Z_{t-1} + a_t - 0.8385a_{t-1} \quad \{a_t\} \sim WN(0,0.02572)$$

2. The ARIMA (1,1,1) model gives the forecast of:

	Jun 2019	July 2019	Aug 2019	Sep 2019	Oct 2019
Forecasted Value	262.7723	263.2227	263.6722	264.1207	264.5682
Actual Value	262.945	263.491	264.135	264.588	265.019

Together with the actual value, we can calculate a mean square error of 0.14751679.

For the monthly Consumer Price Index in the U.S from 1980-01-01 to 2019-10-01:

1. We specified an ARIMA(9,1,0)-Garch(1,1) to model the extremely volatility during the 1980 period, the model parameters are:

$$\begin{aligned} Z_t = \ & 0.406995 + 0.187795Z_{t-1} + 0.158481Z_{t-2} + 0.017848Z_{t-3} - 0.050512Z_{t-4} \\ & + 0.070572Z_{t-5} + 0.120121Z_{t-6} + 0.013156Z_{t-7} + 0.138827Z_{t-8} \\ & + 0.122737Z_{t-9} + a_t \quad \{a_t\} \sim WN(0,0.02783) \end{aligned}$$

$$r_t = -0.0043 + a_t \quad a_t = \sigma_t \varepsilon_t \quad \sigma_t^2 = 0.0017 + 0.0625\sigma_{t-1}^2 + 0.8787\sigma_{t-2}^2 \quad \varepsilon_t \overset{iid}{\sim} \mathrm{t}(10)$$

2. The ARIMA(9,1,0)-Garch(1,1) model gives the forecast of:

	Jun 2019	July 2019	Aug 2019	Sep 2019	Oct 2019
Forecasted Value	262.7900	263.2386	263.6730	264.1528	264.6095
Actual Value	262.945	263.491	264.135	264.588	265.019

Together with the actual value, we can calculate a mean square error of 0.13165281.

From the ARIMA (1,1,1) and ARIMA(9,1,0)-Garch(1,1) model, we observe consistent results with the ARIMA(9,1,0)-Garch(1,1) model having a slightly lower mean square error. In real life the great inflation happened during the period of 1970 to 1980 mainly caused by a sequence of events like President Lyndon B. Johnson's “guns and butter” spending on the Vietnam War and the Great Society in the mid-1960’s together with loose monetary policies of the Federal Reserve, the depeg of gold and the U.S. dollar in 1968 and the suspension of the convertibility of the dollar to gold for foreign governments during 1971.

References

Data Source:

[1] FRED. (2022).*Consumer price index for all urban consumers: All items in U.S. city average*. from https://fred.stlouisfed.org/series/CPIAUCSL

Time Series Models:

[2] B. Abroham & J. Ledolter (2005) *Statistical Methods for Forecasting*, Wiley.

[3] Engle, R. F. (1982). *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*. Econometrica 50: 987–1007. https://doi.org/10.2307/1912773

[4] J.D. Cryer & K.S. Chan (2008) *Time Series Analysis with Applications in R*, 2nd edition, Springer.

Information of U.S. great Inflation

[5] Bryan, M. (2013). *The great inflation*. Federal Reserve History from https://www.federalreservehistory.org/essays/great-inflation

[6] Sommer, J. (2022). *Lessons From the ’80s, When Volcker Reigned and RatesWere High*. The New York Times, from https://www.nytimes.com/2022/08/05/business/inflation-investing-paul-volcker.html

[7] U.S. Bureau of Labor Statistics. (2022). U.S. Bureau of Labor Statistics.

Appendix

All R and Python codes for model building can be found on:

https://github.com/edmund844328/STAT4601_TimeSeriesForecasting.git