I choose the United Kingdom’s real GDP data from FRED as my project dataset. This dataset has been seasonally adjusted when posted on FRED. I use the ts function to convert it to Time Series (Exhibit 1). By drawing the TS plot, I find the dataset shows an obvious ascending trend (Exhibit 2). Except for the years around 1980 and 2008 financial crisis, the growth rate (Exhibit 3) did not have a big variation which can explain why the ts plot has a comparatively smoothing rising trend. Then, I checked ACF (Exhibit 4) and PACF (Exhibit 5) for the data. The ACF decays slowly so it suggests that the AR model would be a good choice than the MA model. Since the GDP number is too big, not easy to read and plot, I take the log of the numbers. To process the forecasting models, I divide the data into two parts. 80% for the training data (1975-2010), 20% for the valid data (2011-2019).

Before doing the forecasting, I use the ADF test on unit root data to test the stationarity. Since my data has a trend, I choose to use Test 3 (with trend), using the auto lag selection feature with the BIC. According to the test result, I find my data non-stationary, failing to reject the null. For non-stationary model, the average is dependent on time. The ARMA(p,q) does not trend up or down over time (stationarity). Non-stationary data cannot be implemented in a typical ARMA format, but I can fit trend into an ARMA model. Therefore, I am going to use the ARIMA model, taking the first difference of the data to reduce the possible bias. Then I check the ACF (Exhibit 6) and PACF (Exhibit 7) plots for the first-differenced data and based on the plots, I believe my data fit an ARIMA (1,1,0) model. I choose to use the ARIMA model on my training data but take the first difference to avoid some biases due to its non-stationarity. Then, I conduct the forecast with a 95% confidence interval on my validation data and plotted the graphs (Exhibit 8) as well as using the accuracy function to identify the RMSE. The RMSE for this model is 0.016.

Then, I run another three baseline models for comparison. The first one is the auto.arima model (Exhibit 9). I choose to use BIC information to find out the best model since BIC is a larger penalty at longer time horizons. My data has a longtime horizon. The second baseline model is the Naive model (Exhibit 10). The naive model is suitable for being a baseline model. Since my data has a trend, I use the Naive forecast with drift (rwf). I believe this would be better than the pure Naive model. The third baseline model (Exhibit 11) is the Holts linear model (ETS, model = ‘AAN’ (level+trend)). The steps for three baseline models are pretty much the same. I train the model using the training data and do the forecast using the test data. I find out their RMSEs through the accuracy function. I also conduct the DM Test three times to compare the sqrt error of my model and three baseline models.

The plots for my model and three baseline models are presented in the exhibits page and comparison results of RMSEs and DM Tests are shown in the tables below. According to the results, one of the baseline models, auto.arima model (2,1,0), has the smallest RMSE, meaning a better model performance than other models. However, based on the DM Tests, my model, ARIMA (1,1,0), outperforms than others, that is having a smaller sqrt error than others. As you can see in the plots, the forecasting plots of auto.arima model (2,1,0) and my model ARIMA(1,1,0) are pretty similar. Therefore, I think auto.arima model (2,1,0) and my model ARIMA(1,1,0) both perform well and are doing much better than other models, like naive and ETS models.

Exhibit 1.

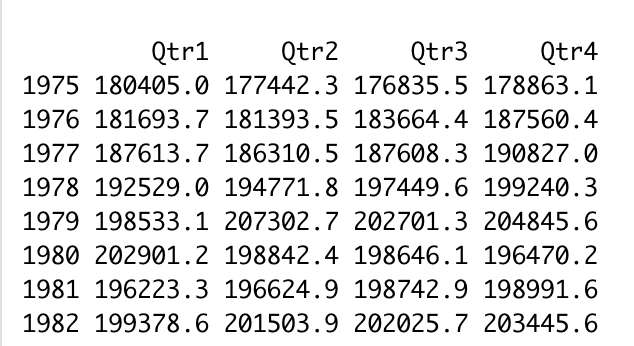


Exhibit 2.

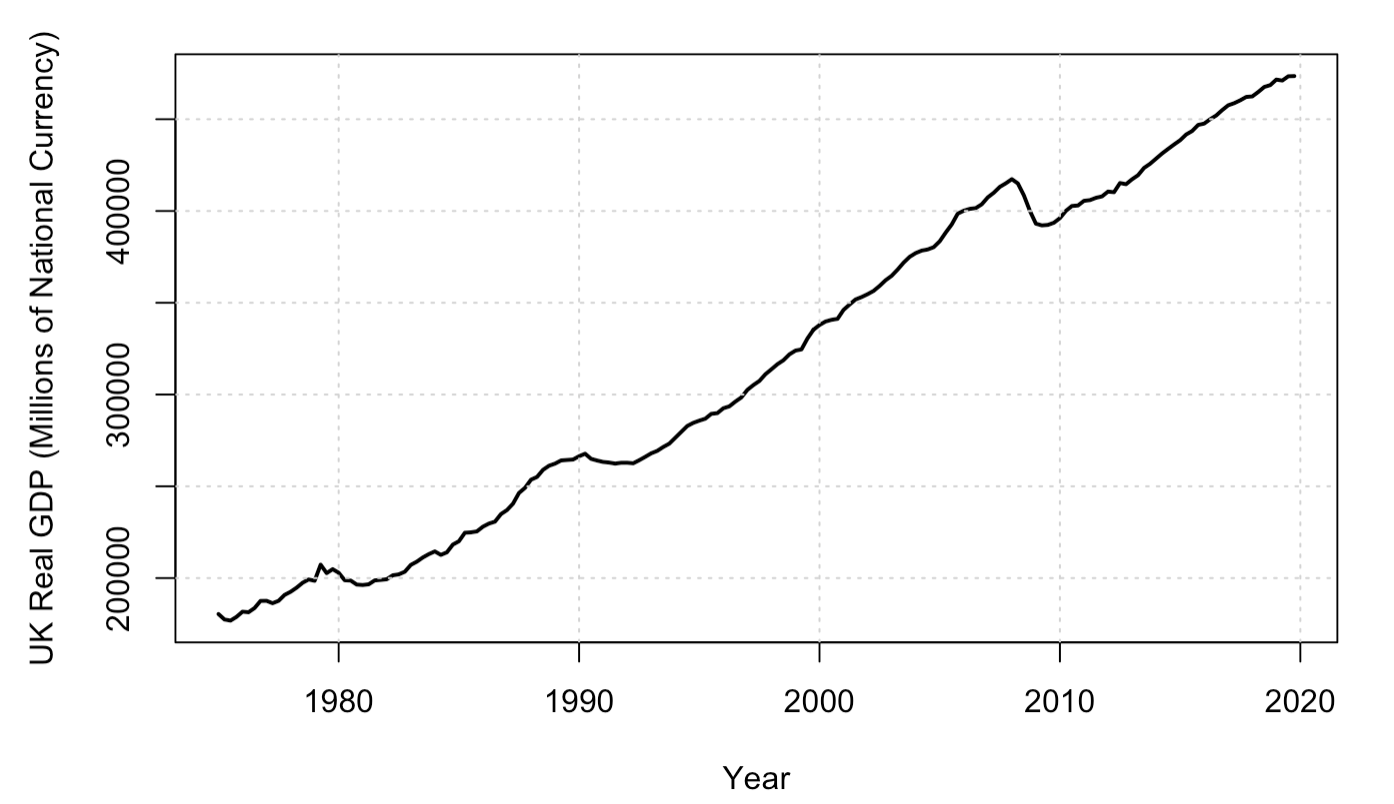


Exhibit 3.

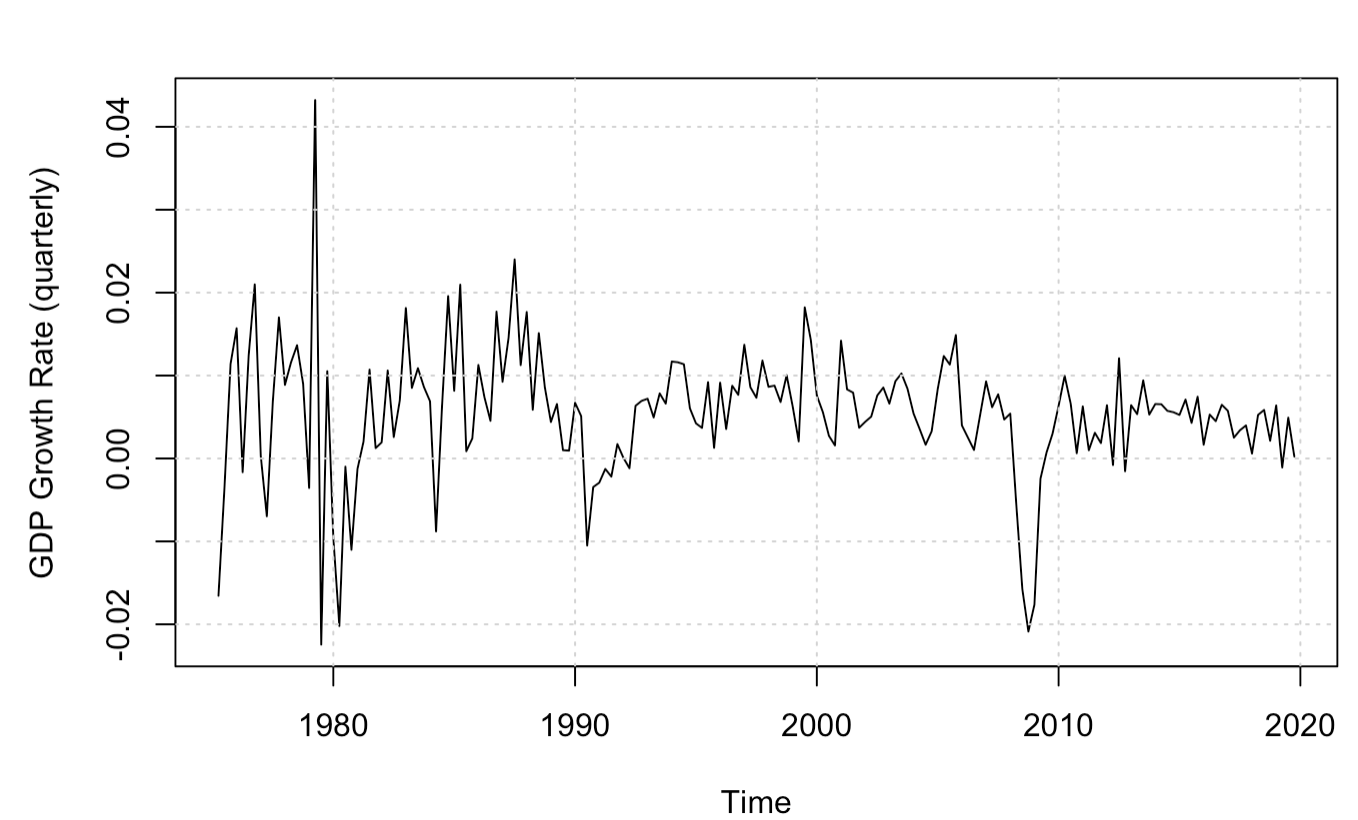


Exhibit 4.

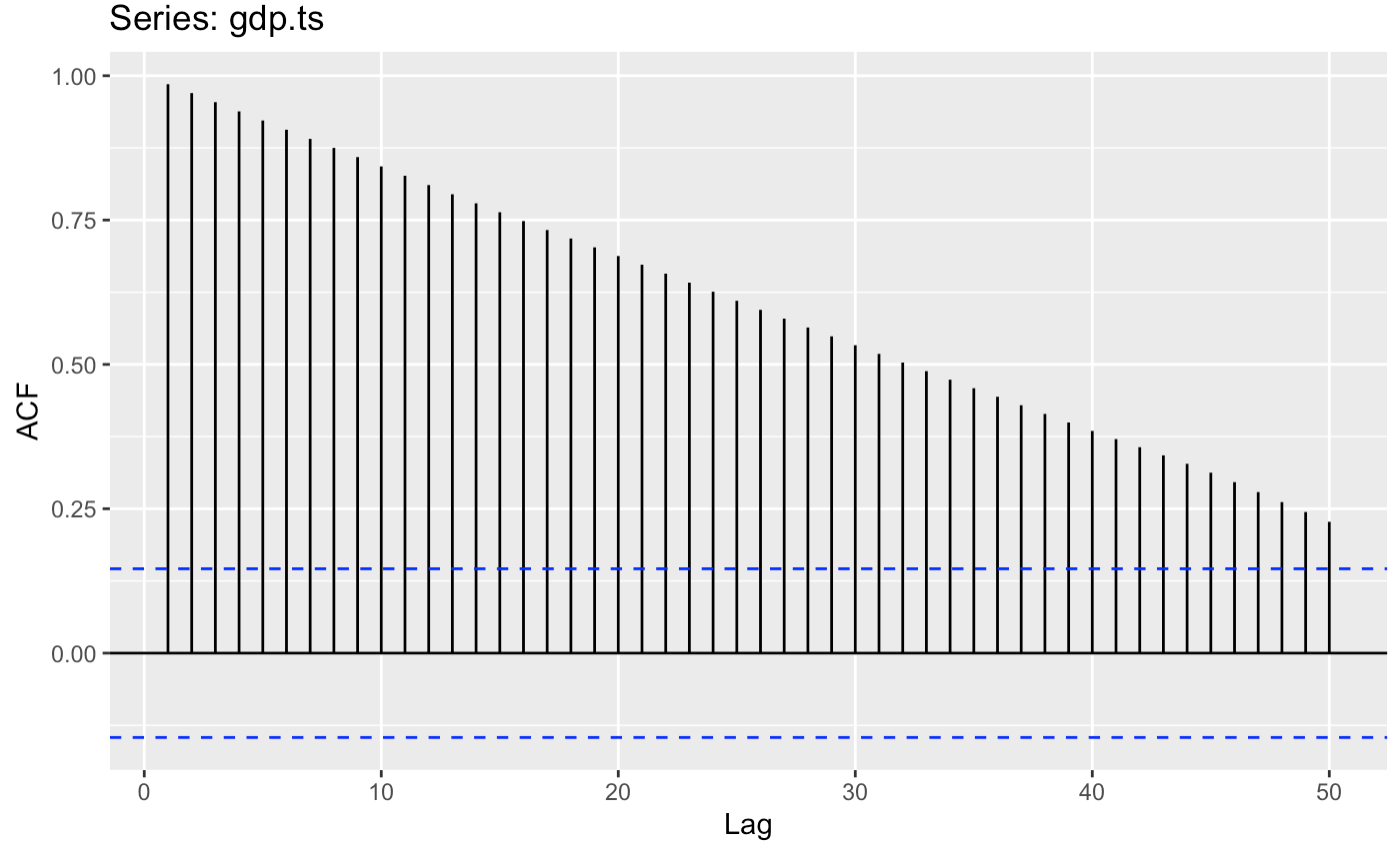


Exhibit 5.

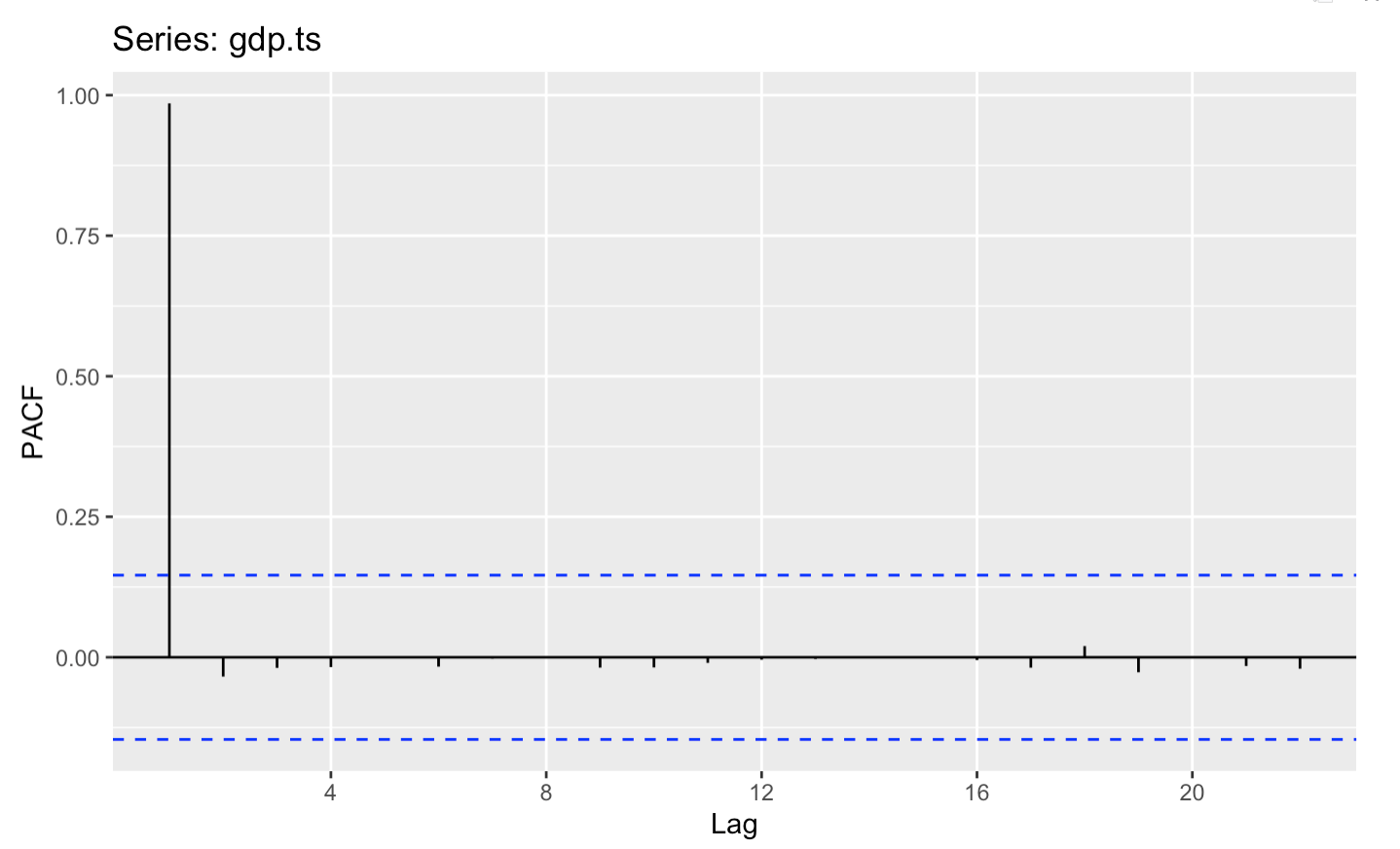


Exhibit 6.

ACF plot for first difference

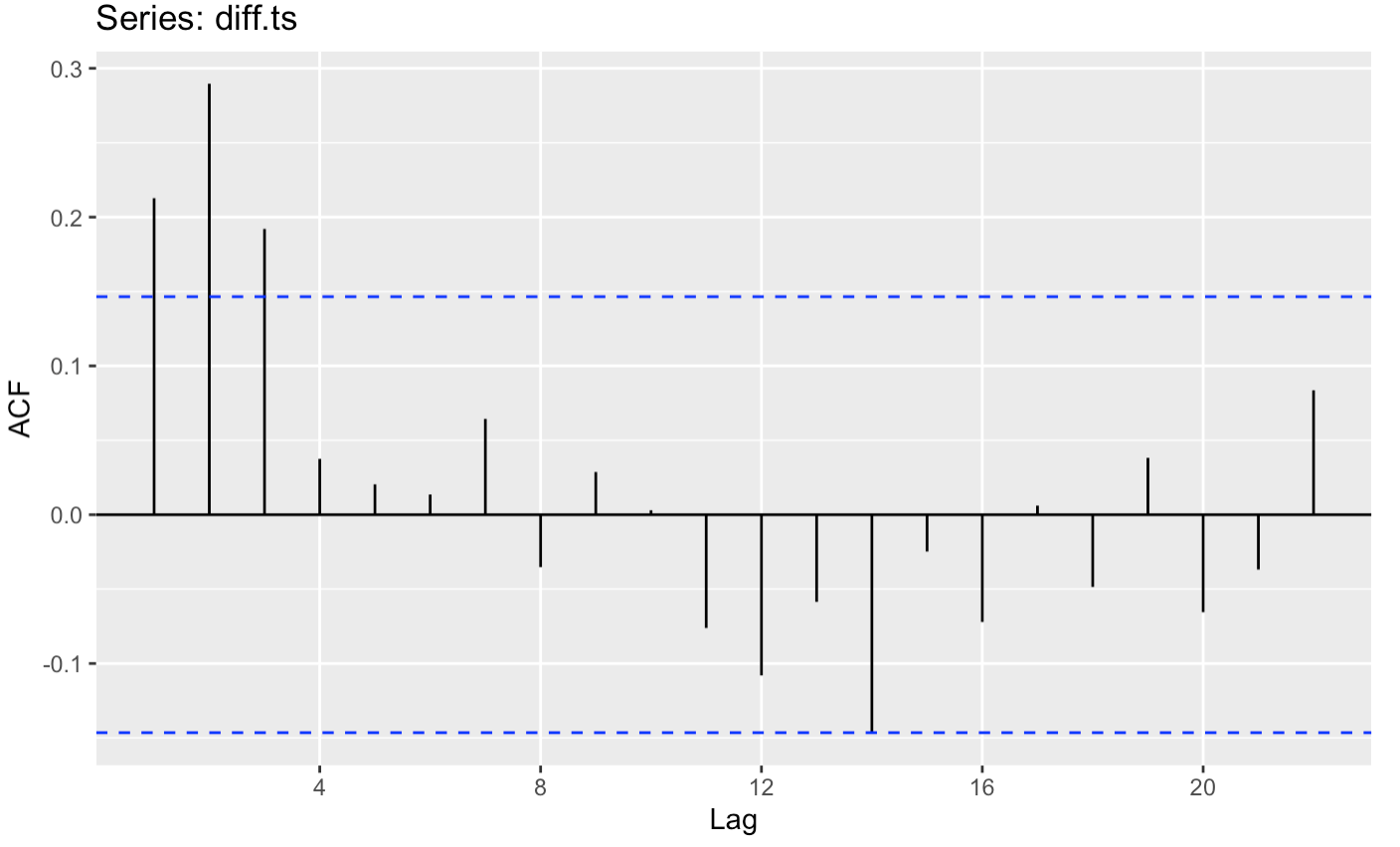


Exhibit 7.

PACF plot for first difference

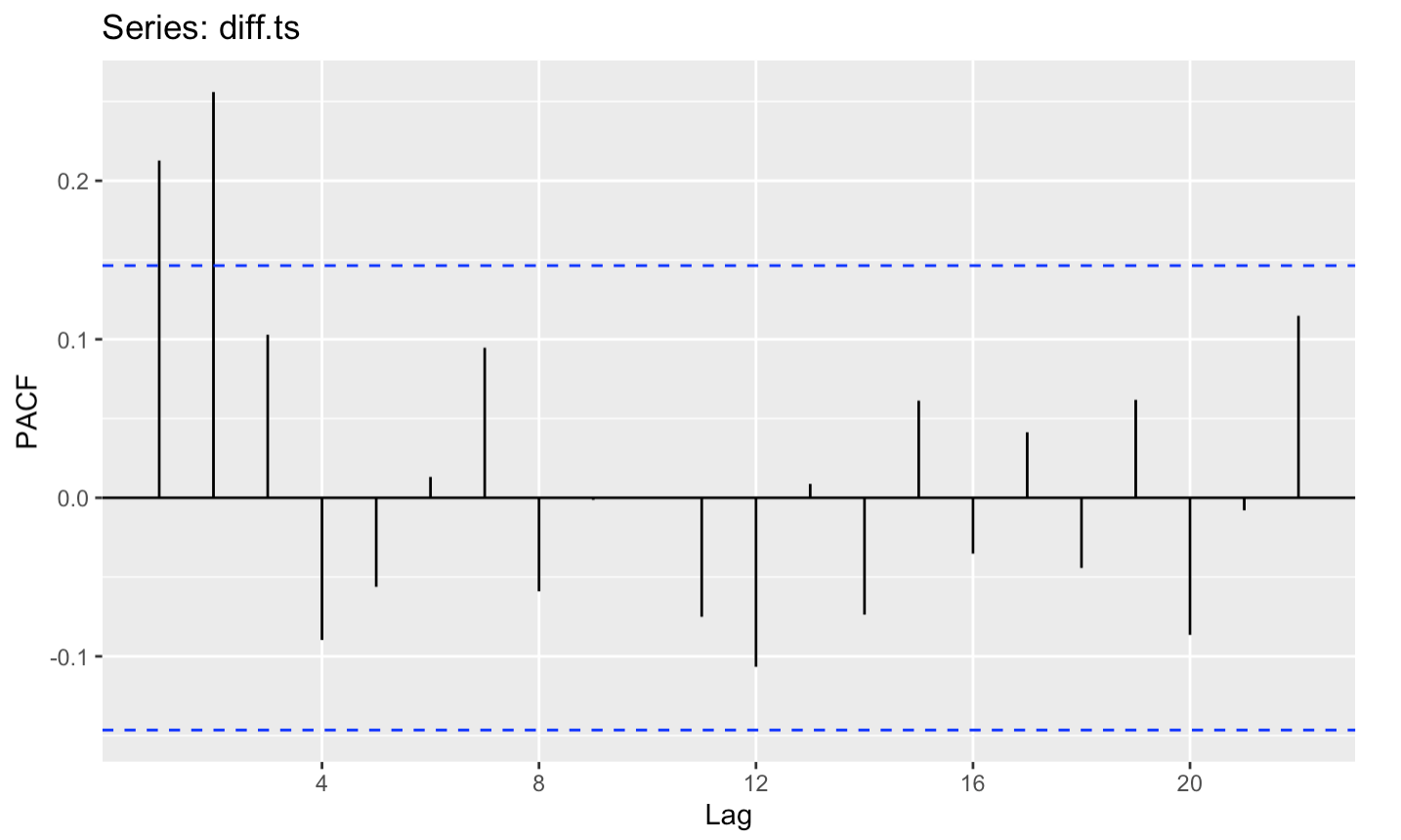


Exhibit 8.

My model: ARIMA (1,1,0)

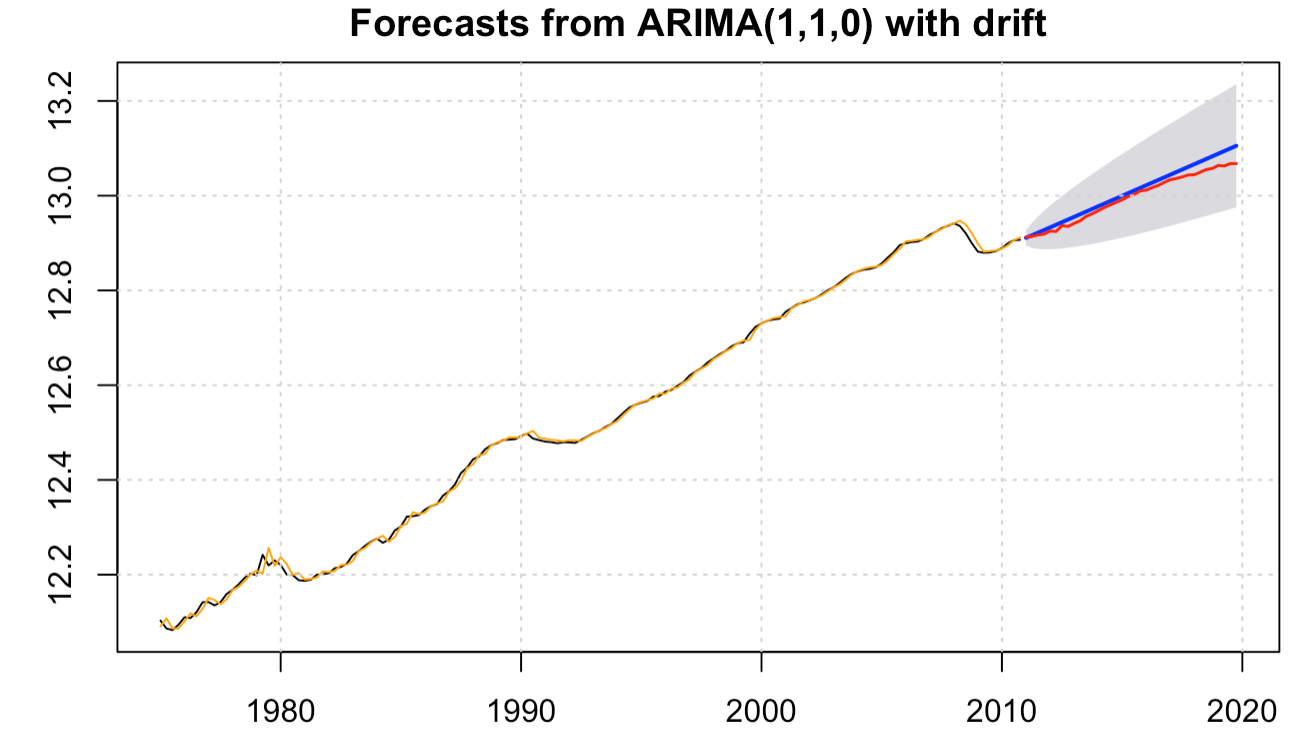


Exhibit 9.

AUTO.ARIMA

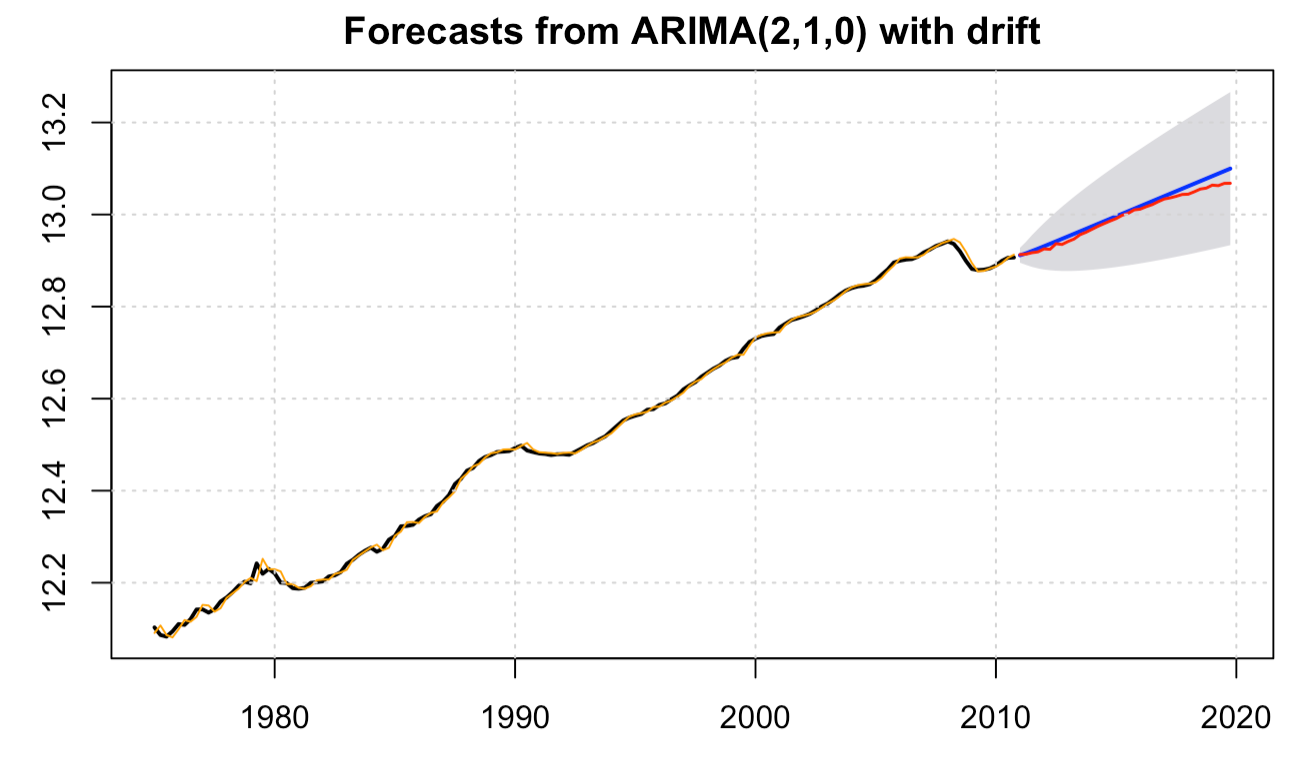


Exhibit 10.

Naive with drift

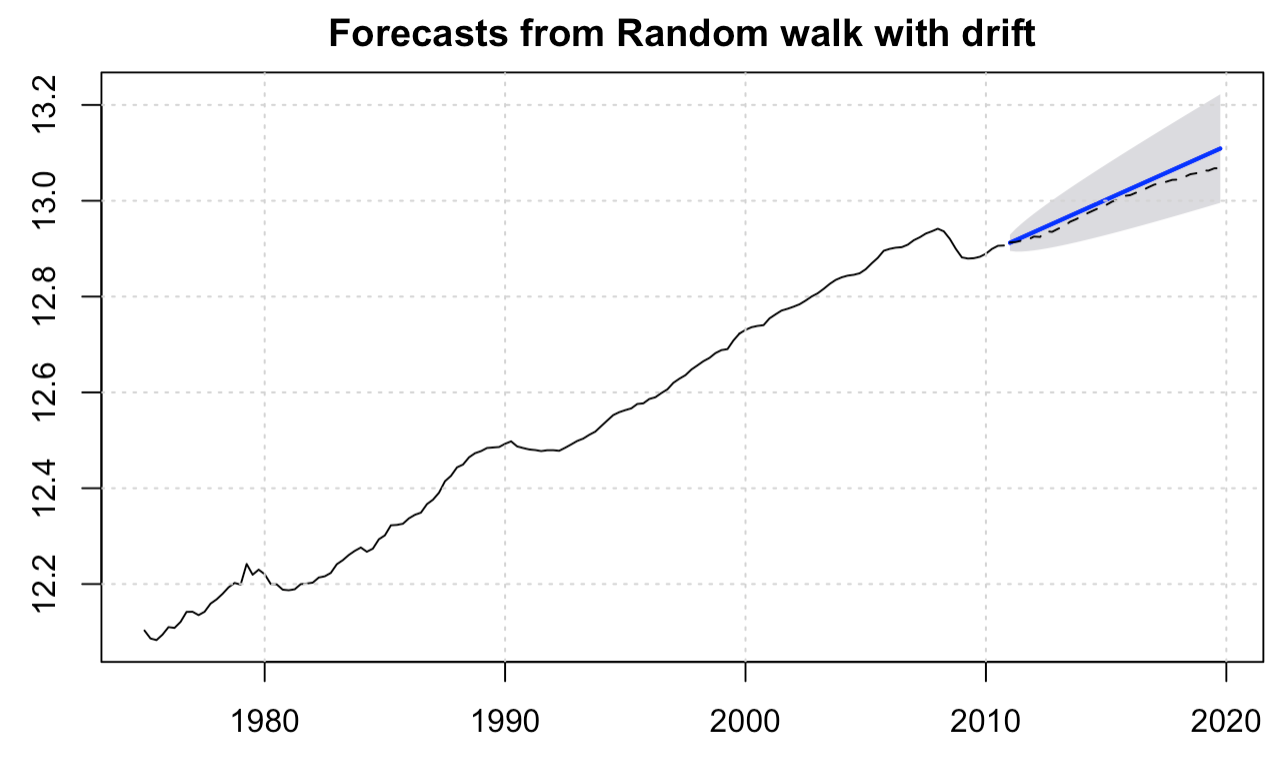
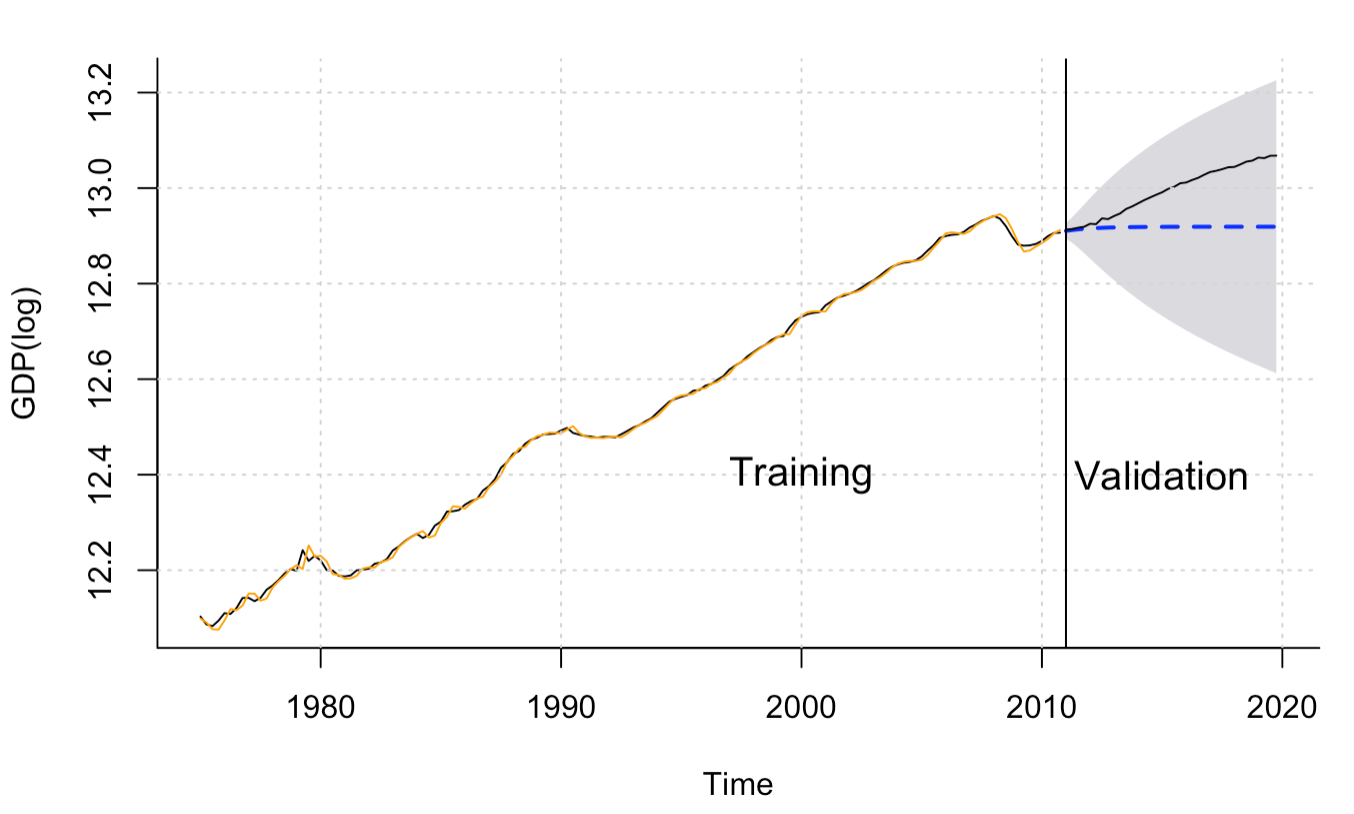
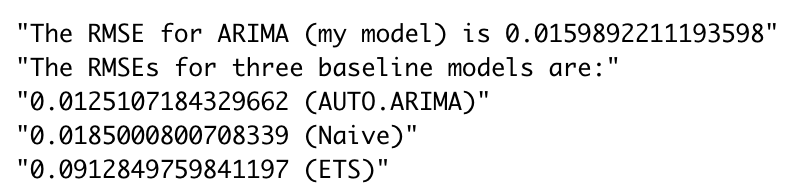


Exhibit 11.

Holts linear model, AAN



RMSE Comparison:



DM Test:

**Arima (1,1,0) V.S. auto.arima (2,1,0)**

P-value of auto.arima = 0.1263 in DM\_test test, fail to reject the null at 5% level.

So, the sqrt residual of the auto.ARIMA model is larger than the sqrt residual of the ARIMA model. The ARIMA model is better than auto.ARIMA model based on the DM Test.

**Arima (1,1,0) V.S. Naive with drift**

P-value of the naive model = 0.7726 in DM\_test test, fail to reject the null at 5% level.

So, the sqrt residual of the naive model is larger than the sqrt residual of the ARIMA model.

The ARIMA model is better than the naive model based on the DM Test.

**Arima (1,1,0) V.S. ETS**

P-value of ETS = 0.3452 in DM\_test test, fail to reject the null at 5% level.

So, the sqrt residual of the Holts linear model is larger than the sqrt residual of the ARIMA model. The ARIMA model is better than the naive model based on the DM Test.