Problem Set 4: Forecasting the Unemployment Rate
due Friday April 28, 2023

Instructions: Submissions are individual, R code must be readable, commented and attached at the end of your problem set. Plots, tables and other outputs should be given in the answers or at the end of the problem set.

Problem 1: Forecasting the Current Unemployment Rate

The dataset you will construct consists of weekly initial jobless claims[1] aggregated to monthly data and the monthly unemployment rate in the US.[2] Jobless claims are very valuable for policymaking because they are released more often and earlier than the unemployment rate, as a result they are often used to predict the current level of unemployment during recessions.

  i. Load the following packages: *fredr, dynlm, timetk*

 ii. Go to `https://fredaccount.stlouisfed.org/`, create an account. Click on My Account (top right corner), API Keys, and Request API Key.[3] Once you have generated a key, enter it in R using
     `fredr_set_key('******')`

iii. Now download the data using the following commands:
     ```
     Claims <-fredr( series_id = "ICSA",frequency = "m", observation_start=as.Date("1990-01-01"), observation_end=as.Date("2023-04-01"))
     Unempl <-fredr( series_id = "UNRATE",frequency = "m", observation_start=as.Date("1990-01-01"), observation_end=as.Date("2023-04-01"))
     Claims_w <-fredr( series_id = "ICSA",frequency = "w", observation_start=as.Date("1990-01-01"), observation_end=as.Date("2023-04-01"))
     ```
     You can read the latest weekly Claims number using `tail(Claims_w)`. Now convert the data into the zoo format used by dynlm:
     ```
     Claims = tk_zoo_(Claims)
     Unempl = tk_zoo_(Unempl)
     ```

---

[1] `https://fred.stlouisfed.org/series/ICSA`
[2] `https://fred.stlouisfed.org/series/UNRATE`
[3] For "Describe the application or program you intend to write:", you can say this is for a problem set.

```
data = merge(Claims,Unempl)
colnames(data)= c('Claims','Unempl')
```
You can take a peek at the data using `head(data)` and `tail(data)`. What was the unemployment rate in March 2023?

iv. Initial claims number measure how many individuals filed for unemployment in a given week/month while unemployment numbers measures the total numbers who are unemployed at a given point in time. We will thus try to predict *changes in the unemployment rate* using claims rather than the level of unemployment itself. Explain why this is a good strategy using the notion of flow and stock variables from macroeconomics?

Hint: Claims is the number of individuals who file for unemployment benefits in a given week or month (Claims_w, Claims). Unemployment rate is the fraction of individuals in the labour force who are unemployed in a given month.

v. We will consider ADL models of the form:

$$\Delta U_t = \beta_0 + \beta_1 \Delta U_{t-1} + \cdots + \beta_p \Delta U_{t-p} + \delta_1 X_t + \cdots + \delta_q X_{t-q+1} + e_t,$$

where $\Delta U_t = U_t - U_{t-1}$ is the change in the unemployment rate from $t-1$ to t and X_t are the jobless claims. Estimate the ADL(p,q) model with $p = 0$ and $q = 1, 2, 3$ as well as $p = 1$ and $q = 1, 2, 3$. Report the OLS estimates.

Hint: to estimate the ADL model described above, you can use the following command from the dynlm package:
```
dynlm( diff(Unempl)~ L(diff(Unempl),1:p)+ L(Claims,0:(q-1)), data = data )
```

vi. Using the `window` command, create two subsets of data. The first consists of observations from January 1990 to December 2005 and the second will consist of observations from January 2006 to March 2023. To create the first subset, you could use `data1 = window(data,start = "1990-01-01", end = "2005-12-01")`.

Estimate the regression models in ii. on the first subset of data and compute the forecasted values on the second subset. The forecasts are computed manually using:
```
forecast1 = model1$coef[1] + model1$coef[2]*lag(data2$Claims,0)
```
for an ADL(0,1) model for instance.

vii. Compute the RMSFE for the specifications you estimated above and choose the best model for out-of-sample forecasting.

viii. Using the preferred model from the previous question and the full dataset, predict the change in the unemployment rate in April 2023. Then, using the March 2023 numbers, predict the level of unemployment in April 2023. Do you predict an increase or a decline in the unemployment rate?

Hint: The jobless claims number for April 2023 are in Claims_w.

2