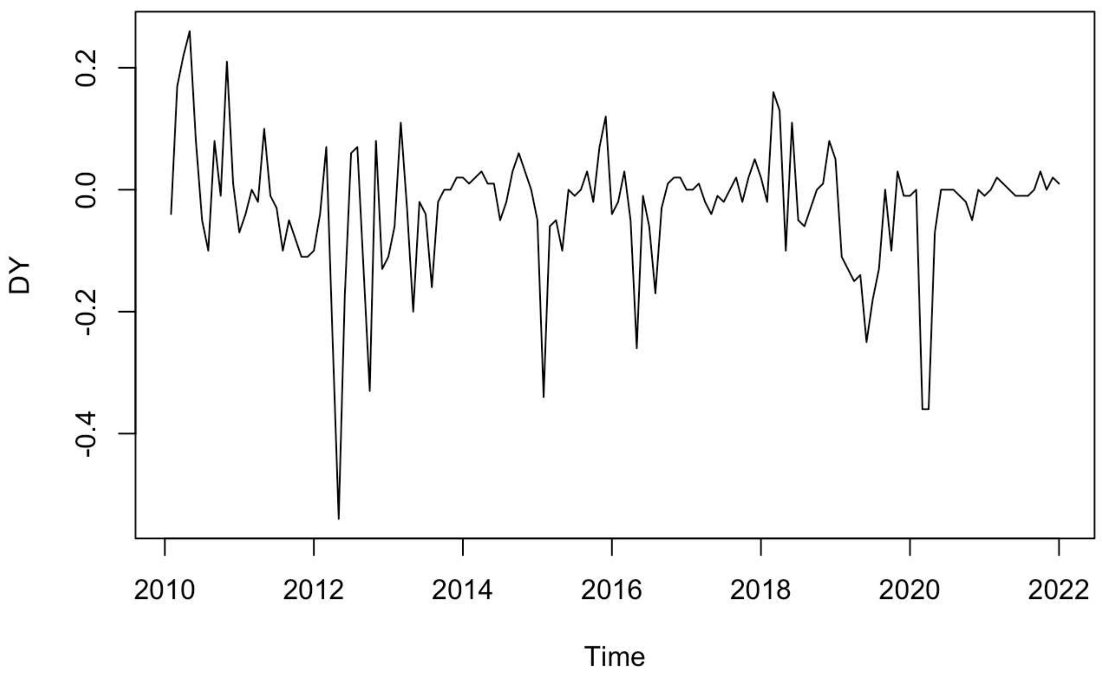
Tutorial 4 Answers

1. Reading in the data and creating the first difference:

```
dt <- read.csv("BAB3mth.csv")
Y <- ts(dt$BAB3, start=c(2010,1), end=c(2025,6), frequency=12)
DY <- diff(Y)
DY <- window(DY, start=c(2010,2), end=c(2022,1))
TimeDY <- round(time(DY),3)
plot(DY)
```



2. Estimation from 2010m2 with $n = 100$:

```
library(forecast)
n <- 100
pmax <- 12
ARorders <- 0:pmax
Statistics <- matrix(nrow=pmax+1, ncol=4)
colnames(Statistics) <- c("LBp", "AICc", "Forecast", "ForecastError")
rownames(Statistics) <- paste0("AR",ARorders)
t_est <- 1:n
for (p in 0:pmax){
  eq <- Arima(DY[t_est], order=c(p,0,0))
  sink(nullfile())
  Statistics[p+1,1] <- checkresiduals(eq, plot=FALSE)$p.value
  sink()
  Statistics[p+1,2] <- eq$aicc
  Statistics[p+1,3] <- forecast(eq, h=1)$mean
  Statistics[p+1,4] <- DY[n+1]-Statistics[p+1,3]
}
print(Statistics)
```

```
           LBp      AICc    Forecast ForecastError
AR0  0.02119014 -153.6739 -0.02240000     0.1324000
AR1  0.45507417 -163.9889 -0.04919461     0.1591946
AR2  0.67033459 -164.8794 -0.08188987     0.1918899
AR3  0.79620749 -164.2343 -0.06612274     0.1761227
AR4  0.76900908 -162.9391 -0.06710616     0.1771062
AR5  0.65085267 -161.1636 -0.06029908     0.1702991
AR6  0.59170527 -159.4486 -0.05636721     0.1663672
AR7  0.62153895 -157.9796 -0.05497310     0.1649731
AR8  0.55215829 -155.7783 -0.05506103     0.1650610
AR9  0.50763811 -154.6281 -0.05884199     0.1688420
AR10 0.58171613 -152.1829 -0.05787059     0.1678706
AR11 0.54100238 -149.6816 -0.05797668     0.1679767
AR12 0.22076405 -147.4861 -0.05556510     0.1655651
```

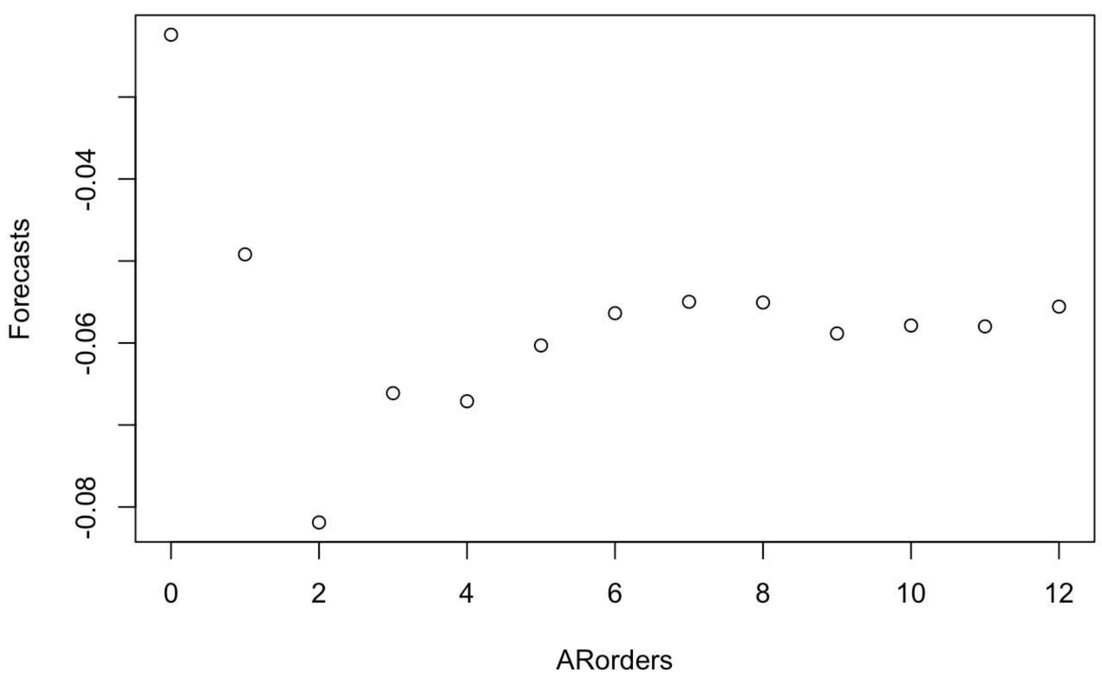
```
# Valid lag orders without autocorrelation:
print(AROrders[which(Statistics[,"LBp"]>0.05)])
```

```
 [1]  1  2  3  4  5  6  7  8  9 10 11 12
```

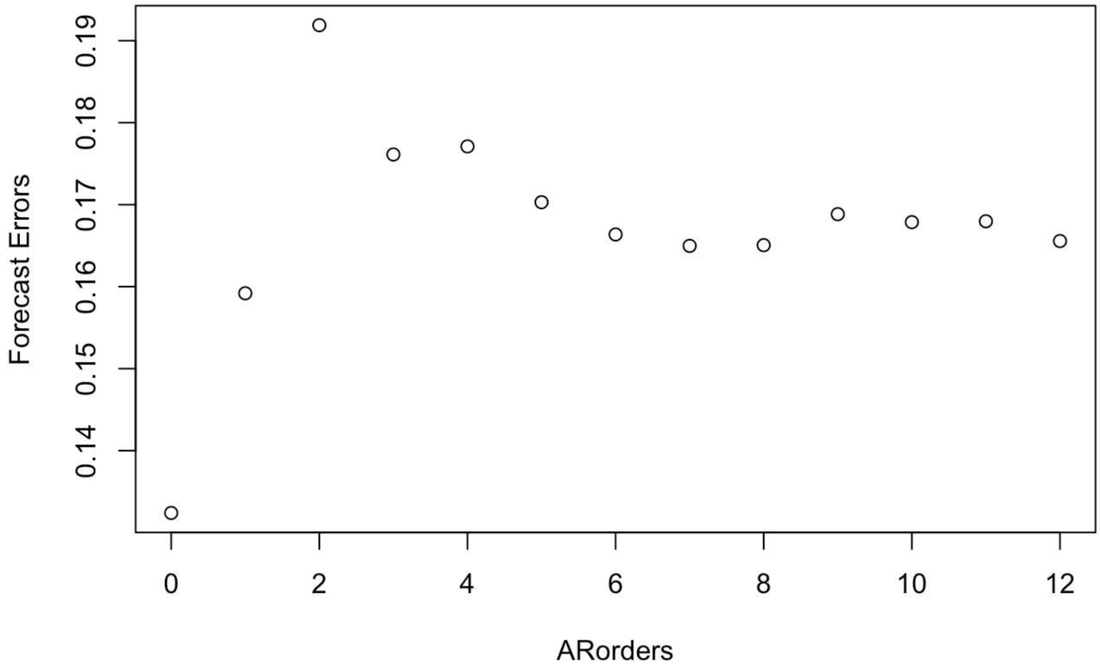
```
# Lag order with minimum AICc:
print(AROrders[which.min(Statistics[,"AICc"])])
```

```
[1] 2
```

```
# Plots
plot(AROrders, Statistics[,3], ylab="Forecasts")
```




```
plot(AROrders, Statistics[,4], ylab="Forecast Errors")
```



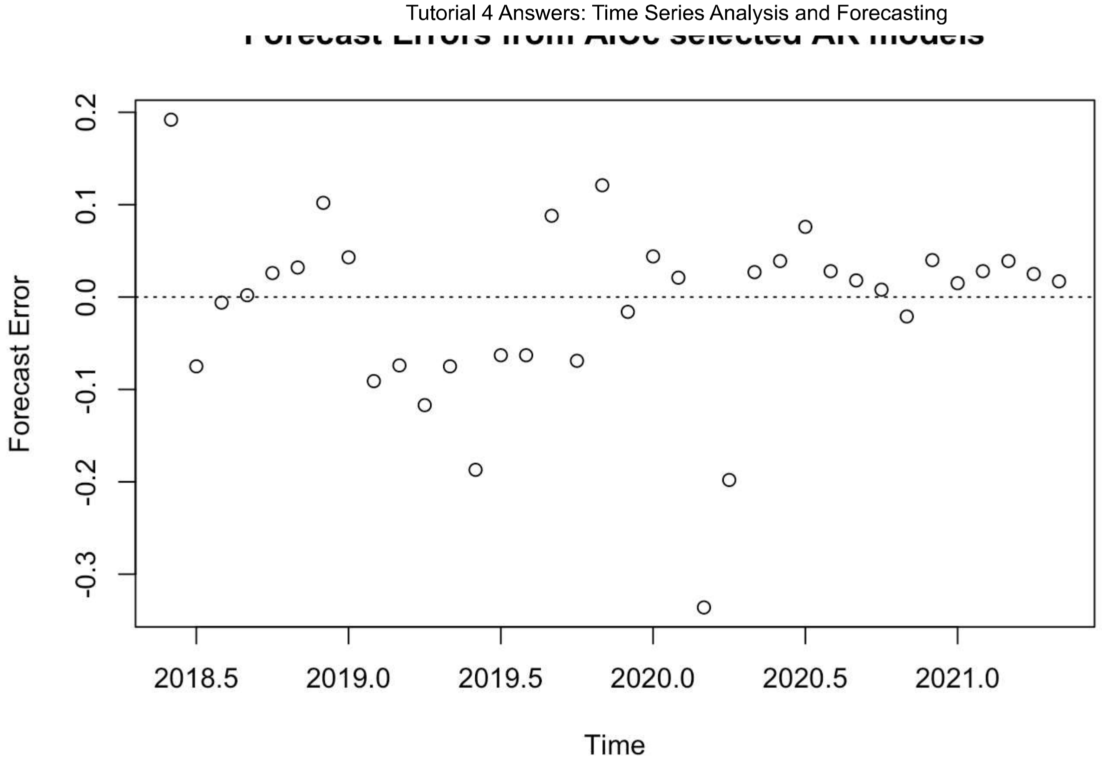
3. Rolling regression

```
RollingStatistics <- matrix(nrow=36, ncol=6)
colnames(RollingStatistics) <- c("EstStart", "EstEnd", "p_AICc",
                                 "NoAuto", "Forecast", "Error")
rownames(RollingStatistics) <- rep("",nrow(RollingStatistics))
for (m in 0:35){
  Statistics <- matrix(nrow=pmax+1, ncol=2)
  colnames(Statistics) <- c("LBp", "AICc")
  t_est <- (1:n)+m
  for (p in 0:pmax){
    eq <- Arima(DY[t_est], order=c(p,0,0))
    sink(nullfile())
    Statistics[p+1,1] <- checkresiduals(eq, plot=FALSE)$p.value
    sink()
    Statistics[p+1,2] <- eq$aicc
  }

  # Start date of estimation
  RollingStatistics[m+1,1] <- TimeDY[1+m]
  # End data of estimation
  RollingStatistics[m+1,2] <- TimeDY[1+m+n]

  # Select p using AICc
  p_AICc <- ARorders[which.min(Statistics[,"AICc"])]
  RollingStatistics[m+1,3] <- p_AICc

  # Check autocorrelation of selected model
  RollingStatistics[m+1,4] <- 1*(Statistics[which.min(Statistics[,"AICc"]),"LBp"]
```

```
  # Forecast from selected model
  eq <- Arima(DY[t_est], order=c(p_AICc,0,0))
  RollingStatistics[m+1,5] <- round(forecast(eq, h=1)$mean,3)
  RollingStatistics[m+1,6] <- round(DY[m+n+1]-forecast(eq, h=1)$mean,3)
}

print(RollingStatistics)
```

```
EstStart   EstEnd p_AICc NoAuto Forecast  Error
2010.083 2018.417      2      1   -0.082  0.192
2010.167 2018.500      1      1    0.025 -0.075
2010.250 2018.583      2      1   -0.054 -0.006
2010.333 2018.667      2      1   -0.032  0.002
2010.417 2018.750      3      1   -0.026  0.026
2010.500 2018.833      1      1   -0.022  0.032
2010.583 2018.917      2      1   -0.022  0.102
2010.667 2019.000      3      1    0.007  0.043
2010.750 2019.083      3      1   -0.019 -0.091
2010.833 2019.167      3      1   -0.056 -0.074
2010.917 2019.250      3      1   -0.033 -0.117
2011.000 2019.333      3      1   -0.065 -0.075
2011.083 2019.417      3      1   -0.063 -0.187
2011.167 2019.500      3      1   -0.117 -0.063
2011.250 2019.583      3      1   -0.067 -0.063
2011.333 2019.667      3      1   -0.088  0.088
2011.417 2019.750      3      1   -0.031 -0.069
2011.500 2019.833      3      1   -0.091  0.121
2011.583 2019.917      3      1    0.006 -0.016
2011.667 2020.000      3      1   -0.054  0.044
2011.750 2020.083      3      1   -0.021  0.021
2011.833 2020.167      3      1   -0.024 -0.336
2011.917 2020.250      3      1   -0.162 -0.198
2012.000 2020.333      3      1   -0.097  0.027
2012.083 2020.417      3      1   -0.039  0.039
2012.167 2020.500      3      1   -0.076  0.076
2012.250 2020.583      1      1   -0.028  0.028
2012.333 2020.667      1      1   -0.028  0.018
2012.417 2020.750      1      1   -0.028  0.008
2012.500 2020.833      1      1   -0.029 -0.021
2012.583 2020.917      1      1   -0.040  0.040
2012.667 2021.000      1      1   -0.025  0.015
2012.750 2021.083      1      1   -0.028  0.028
2012.833 2021.167      1      1   -0.019  0.039
2012.917 2021.250      4      1   -0.015  0.025
2013.000 2021.333      4      1   -0.017  0.017
```

```
plot(RollingStatistics[,2], RollingStatistics[,6],
     xlab="Time", ylab="Forecast Error",
     main="Forecast Errors from AICc selected AR models")
abline(h=0, lty=3)
```



- It is straightforward to select a parsimonious ($p \leq 4$) AR model with $n = 100$ for one-step-ahead forecasting of interest rate changes over the period 2018-2021.

- In the rolling regression exercise, the order of the selected AR model is quite variable as the estimation period steps forward through time. It changes amongst 1,2,3,4 at various times. This could imply some instability in the autocorrelation structure of the time series over this period.

- The forecast accuracy is also variable. Perhaps not surprisingly, forecast accuracy was poor in early 2020, but has was quite stable thereafter with relatively small and generally positive forecast errors.