Time Series assignment 1

Filip Mellgren & Ismael Moreno Martinez & Athina Swahn

2019-01-26

Excercise 1

1a

A necessary condition for a series to be stationary is that the expected value is the same for all values of t. We show that this condition does not hold for all t with an example: $E(Y_1) = E(a_0 + a_1 y_0 + a_2 y_{-1} + \epsilon_t) = E(a_0) + E(a_1 y_0) + E(a_2 y_1) + E(\epsilon_t) = a_0 + a_1 y_0 + a_2 y_{-1}$. Similarly, we have that: $E(Y_2) = a_0 + a_1 E(Y_1) + a_2 y_0$ which simplifies to: $a_0(1 + a_1) + (a_1^2 + a_2)y_0 + a_1 a_2 y_1$. $E(Y_1)$ and $E(Y_2)$ are thus equal if and only if the following holds:
* $a_1 = -1$
* $a_1^2 + a_2 = a_1$
* $a_1 a_2 = a_2 \implies a_1 = 1$ But these can clearly not be true at the same time.

1b

An ARMA(p,q) process, which AR(2) is a special case of, is stationary if all roots of the characteristic polynimial lie outside the unit circle. In this case, we have $A(L) = 1 - a_1 L - a_2 L^2$ and characteristic polynomial $ c(z) = 1 - a_1z-a_2z^2$, which has roots: $z = -\frac{a_1}{2a_2} \pm \frac{\sqrt{a_1^2 + 4a_2}}{2a_2}$ (1).

If at least one solution falls inside the unit circle, the series can't be weakly stationary.

Assume the polynomial has two roots and the first one is outside of the $|z_1| > 1$, therefore (i) $-\frac{a_1}{2a_2} - \frac{\sqrt{a_1^2+4a_2}}{2a_2} > 1$ and (ii) $-\frac{a_1}{2a_2} - \frac{\sqrt{a_1^2+4a_2}}{2a_2} < 1$.

From (i) We get $-\sqrt{a_1^2 + 4a_2} > 2a_2 + a_1$. Elevating both sides to the power of two we get $a_1^2 + a_2 > 4a_2^2 + 4a_1 a_2 + a_2^2$, which gives us the inequality $a_2 < 1 - a_1$

Following similar steps from (ii) we find other inequality $a_2 < 1 + a_1$.

Now assume the polynomial has only one solution $|z^*| > 1$. By one, the polynomial having only one solution implies $a_1^2 = -4a_2$, therefore, $a_2 = -a_1^2/4$ (2). Then the value of the root would be $-\frac{a_1}{2(-a_1^2/4)} = 2/a_1$. Given that this root's absolute value should be greater than 1, gives us the restriction $-2 < a_1 < 2$. This, together with (2) implies the last restriction $-1 < a_2$.

In summary, an AR(2) process is stationary if the pair of coefficients (a_1, a_2) lay on the "stationary triangle" defined by:

a) $a_2 < 1 - a_1$
b) $a_2 < 1 + a_1$
c) $a_2 > -1$

Note that, whenever the pair of coefficients lay below $a_2 = -a_1^2/4$, we need complex numbers to find the roots so these still don't lie in the unit circle and therefore the process is still stationary.

- $\iff \sqrt{(\frac{a_1/a_2}{2})^2 + 1/a_2} > 1 + \frac{a_1/a_2}{2}$

- $\iff (\frac{a_1/a_2}{2})^2 + 1/a_2 > (1 + \frac{a_1/a_2}{2})^2 = 1 + a_1/a_2 + (\frac{a_1/a_2}{2})^2$
- $\iff 1/a_2 > 1 + a_1/a_2$
- $\iff 1/a_2 > 1 + a_1/a_2 \implies 1 > a_2 + a_1$

1c

First, we transform the AR(2) into a mean-zero process.

$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \epsilon_t \implies E(Y_t) = a_0(1 - a_1 - a_2) = \mu$

The auto covariance function is as follows:

$\gamma_s(Y_t Y_{t-1}) = E[(Y_t - \mu)(T_{t-s} - \mu)]$

To get to this expression we add and substract the mean:

$Y_t + \mu - \mu = a_0 + a_1(Y_{t-1} + \mu - \mu) + a_2(Y_{t-2} + \mu - \mu),$

$\implies Y_t - \mu = -\mu + a_0 + a_1\mu + a_1(Y_{t-1} - \mu) + a_2\mu + a_2(Y_{t-2} - \mu) + \epsilon_t,$

$\implies X_t = -\mu + a_0 + \mu(a_1 + a_2) + a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t,$

$\implies X_t = a_1 X_{t-1} + a_2 X_{t_2} + \epsilon_t.$

Where $X_t = Y_t - \mu$.

Next, we use the tecnique made on page 60 to obtain the Yule-Walker equations:

$\gamma_0 = \sigma_y^2 = E(X_t * X_t) = E(X_t(a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t)) = a_1\gamma_1 + a_2\gamma_2 + \sigma_\epsilon^2$ (I)

$\gamma_1 = E(X_t * X_{t-1})E((a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t) * X_{t-1}) = a_1\gamma_0 + a_2\gamma_1$ (II)

$\gamma_s = a_1\gamma_{s-1} + a_2\gamma_{s-2}$ (III)

Therefore, dividing (II) by γ_0 we get $\rho_1 = a_1\rho_0 + a_2\rho_1$. By definition $\rho_0 = 1$, then $\rho_1 = a_1/(1 - a_2)$.

Plug this value in (III) we can solve the equation for $\rho_2 = \frac{a_1^2}{(1-a_2)} + a_2$.

Iterating this process you can find the following values of the autocorrelation function ρ_s for $s > 2$.

Now, we can express (I) as $\gamma_0 = (1 - a_1\rho_1 - a_2\rho_2)\sigma_\epsilon^2$. Hence, substituting the values we already have for the autocorrelations:

$\gamma_0 = \frac{\sigma_\epsilon^2}{[1 - \frac{a_1^2}{(1-a_2)} - \frac{a_2 a_1^2}{(1.a2)} + a_2^2]} = \frac{\sigma_\epsilon^2(1-a_2)}{(1+a_2)(a_2-a_1-1)(a_1+a_2-1)}$

Finally, note that having the autocorrelation function we can find any autocovariance $\gamma_s = \rho_s\gamma_0$.

1d

Stationarity is important in time series because once we have stationarity, we no longer have any systematic bias in the error term; it collapses to random noise. This means we have identified the key underlying data generating process so that we can make good predictions. Non stationarity means biased predictions and larger standard errors of the predictions, which means we'll on average be wrong with much uncertainty.

1e Long run equilibrium of the time series above

$\mu = E(y_t) = E(a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \epsilon_t) = a_0 + a_1 E(y_{t-1}) + a_2 E(y_{t-2}) = a_0 + a_1\mu + a_2\mu \implies \mu = \frac{a_0}{1-a_1-a_2}.$

1f

The provided values give us $\mu = \frac{2}{1-0.5-0.25} = 8$. When we begin with $y_{-1} = y_0 = 8$, the series will remain at 8 forever as it is a stationary point: $y_1 = 2 + 8/2 + 8/4 = 8$ which means we plug in for the next step whatever we had in the previous step. For the other starting values $y_0 = 7, y_{-1} = 6$, the series is going to update every period by: $\Delta y_t = a_0 + (a_1 - 1)y_{t-1} + a_2 y_{t-2} = a_0 - y_{t-1}/2 + y_{t-2}/4$.

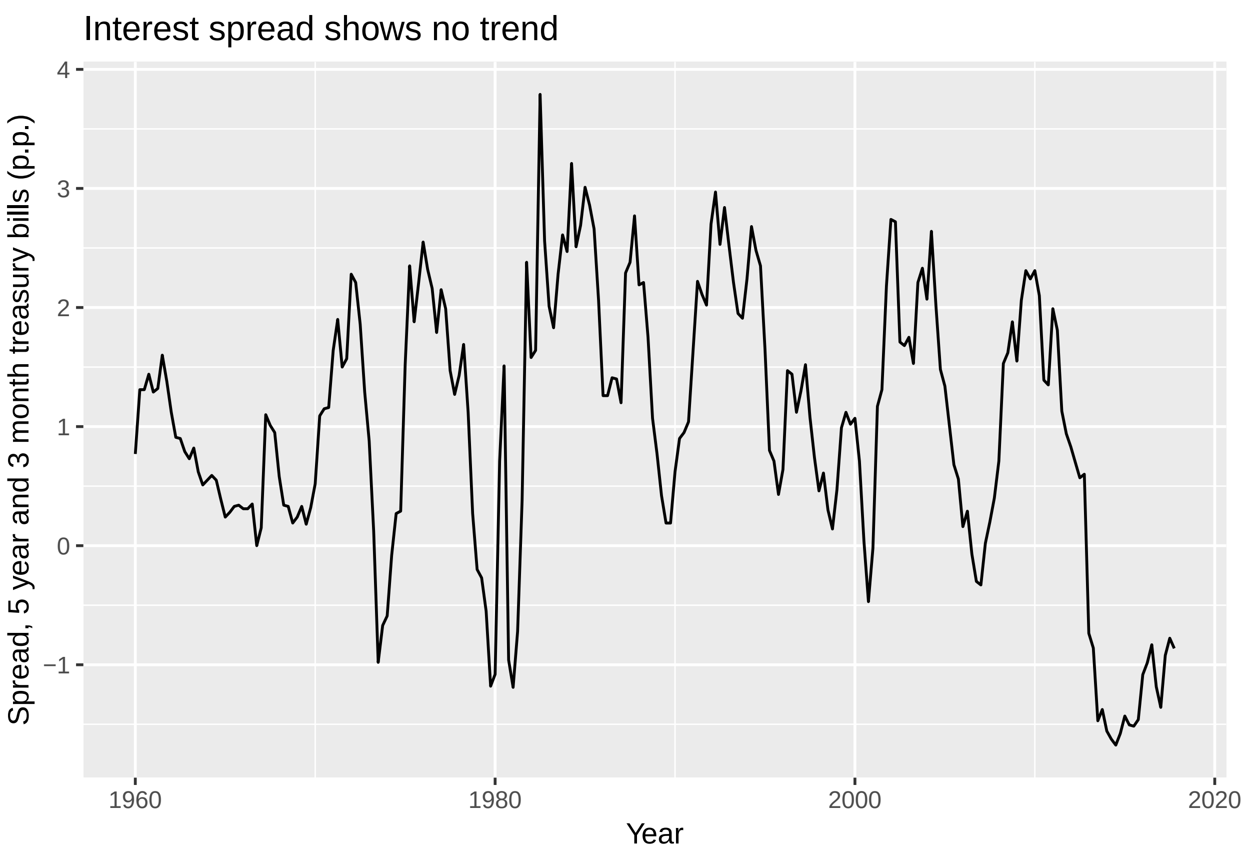
Excercise 2

Question 2a: Go through the example of interest rate spreads Section 2.10 (pp. 88-96) using 23 quarters of additional data (which you download from the course web), i.e. from 1960Q1-2018Q3. Comment on the new results and compare them to the ones in the textbook.

Load and visualise the data

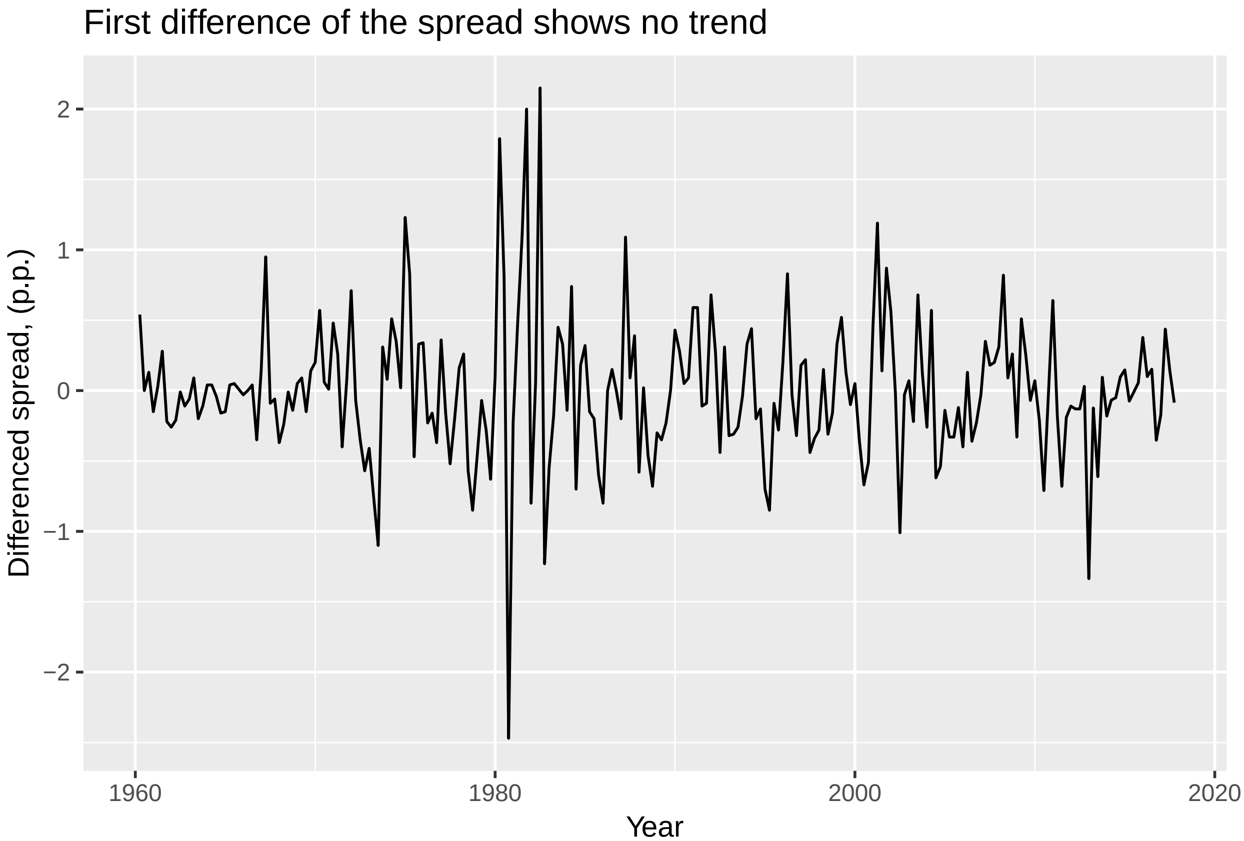
```
##   Tbill   r5
## 1  3.87 4.64
## 2  2.99 4.30
## 3  2.36 3.67
## 4  2.31 3.75
## 5  2.35 3.64
## 6  2.30 3.62
```

We begin our analyses by plotting the observed spread.



We notice that there seems not to be a clear trend appearent in the data. There seems not to be any major structural breaks and we can suspect that the time-seires is covariance stationary. It seems as if there is no notable difference between the textbook spread and our spread that has 23 quarters of additional data. However, one could suspect a downwards going trend from after 1980 instead of an upward going trend mentioned on page 94.

Next, we plot the differenced spread, the quarterly change in spread.



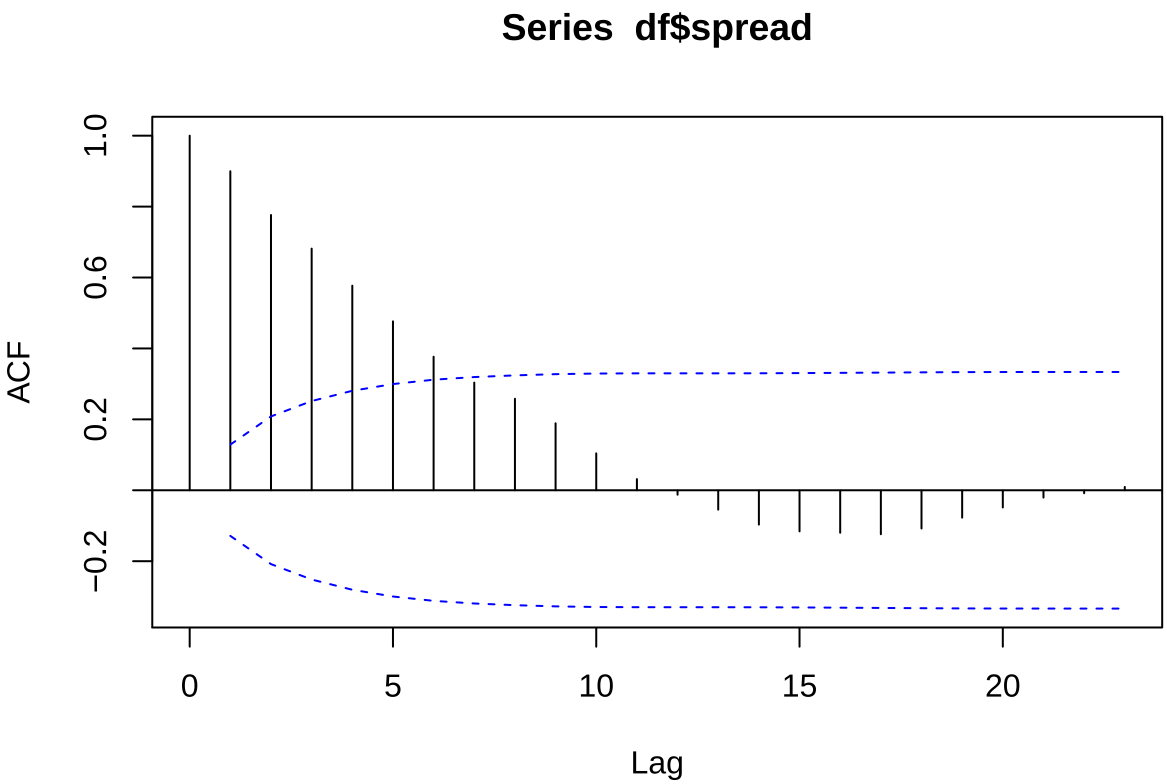
The first difference of the spread seems to be very inconsistant and hence, shows no trend.

Box Jenkins analysis

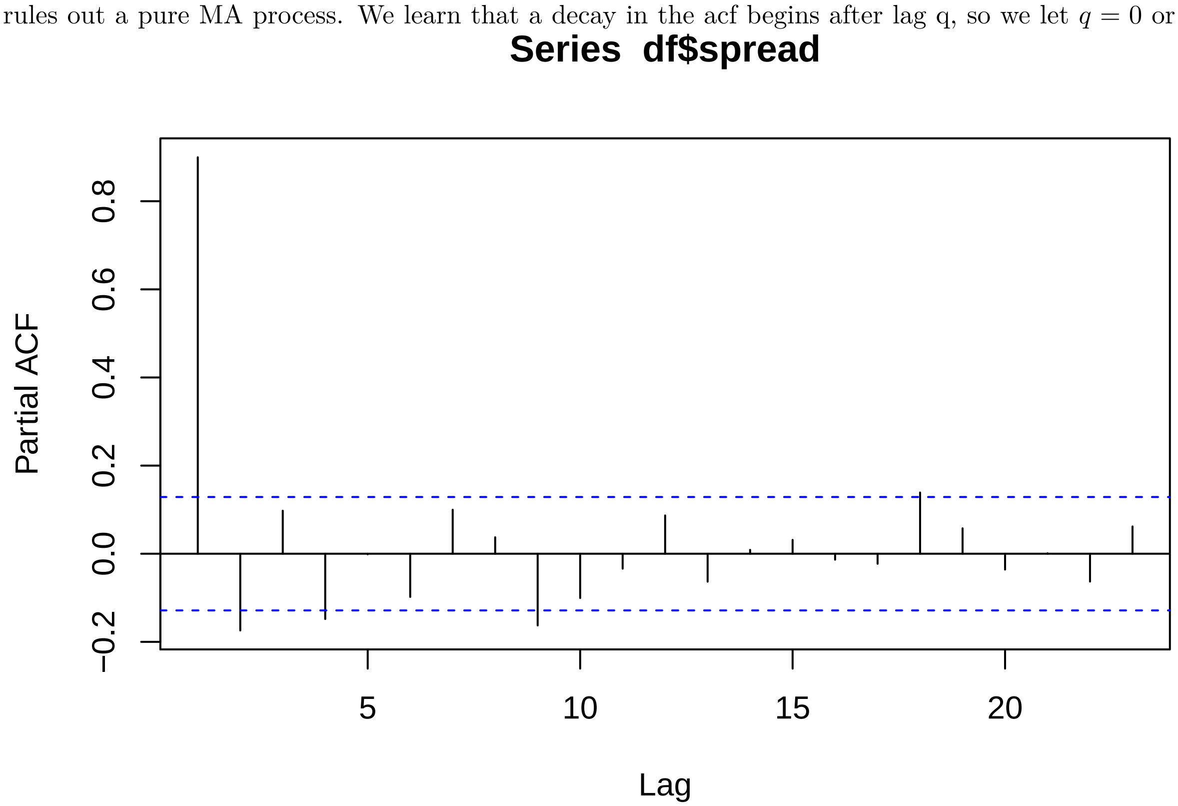
We structure our model selection using the Box Jenkins approach: identification -> estimation -> diagnostics.

Identification: ACF and PACF model selection

We start the identification with ACF and PACF plots:



The ACF shows a decay that begins directly, but it doesn't look geometric. The decay is not sharp, which rules out a pure MA process. We learn that a decay in the acf begins after lag q, so we let $q = 0$ or $q = 1$.



The PACF looks single spiked, which is an indication of an AR(1) process, $p = 1$. On the other side, there are several spikes outside the blue band, suggesting p could be as high as 9. The second spike is also significant, meaning we may want to add at least a second AR term. The oscilation suggests there is a positive MA coefficient, hence we rule out that $q = 0$ and let $q = 1$.

In the end, we test for three different models: ARMA(1,1), ARMA(2,1) and ARMA(9,1).

Estimation of the tentative models

The table shows the setimated models:

Table 1:

Call:	arima(x = df$spread, order = c(1, 0, 1))	Coefficients:	ar1 ma1 intercept	0.8548 0.2881 0.9450	s

Diagnostic checking

Do we capture all serial correlation using the models?

We test the null hypothesis that the residuals are distributed as white noise using lags 1 and 4. We use these values as it seems reasonable that there might be some correlation between the residuals from last quarter and last year, but not for more than one year ago.

For the first model (ARMA(1,1)) we observe the following p-values from the Box-Ljung test:

- Lag 1: 0.6545622
- Lag 4: 0.2253007.

The second model (ARMA(2,1)) gave us:

- Lag 1: 0.9490076
- Lag 4: 0.9045016.

And the third model (ARMA(9,1)) gave us:

- Lag 1: 0.9922003
- Lag 4: 0.9975698.

None of these test statistics are significant, meaning that they all do a good job making the series stationary. The most complicated model is the least significant, but this occurs because it might be overfitting. The ARMA(2,1) looks like a good trade off between parsimony and explanatory power. Accidentaly, we don't select ARMA(2,1) as the "manually" selected model to hold against the automatically selected model using the AIC, because then the comparison wouldn't be interesting as will soon become evident. We therefore go with the simple ARMA(1,1) as our "manual" model.

Model selection with AIC

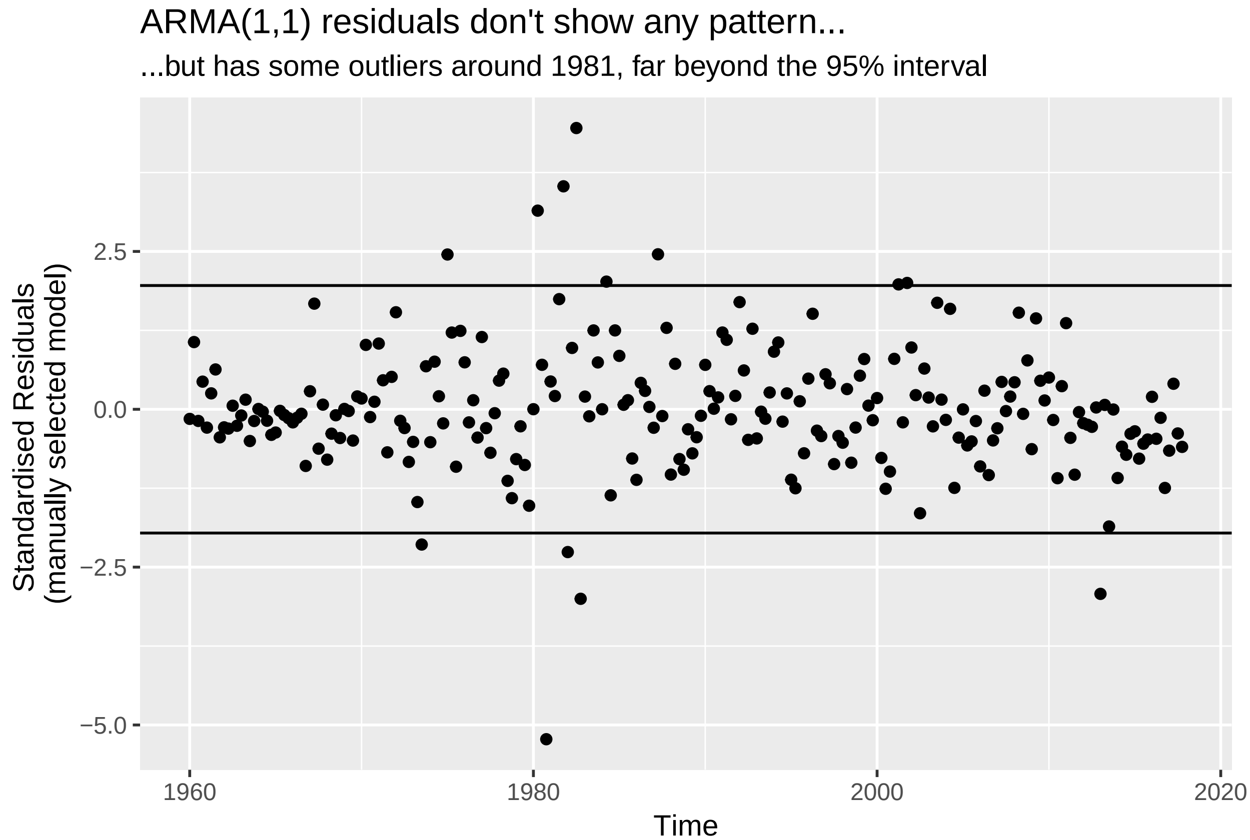
Instead of choosing models manually, we can use the AIC to select a model for us.

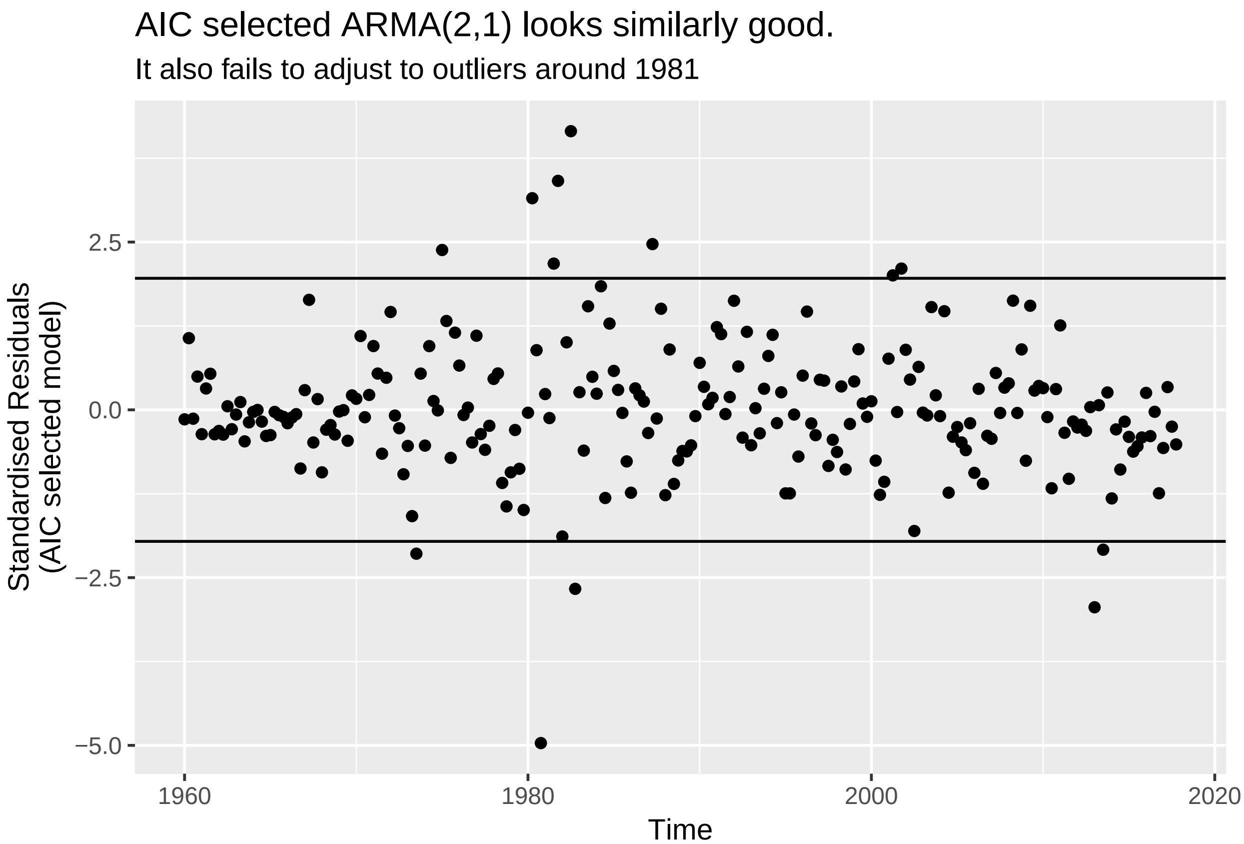
```
## Series: df$spread
## ARIMA(2,0,1) with non-zero mean
##
## Coefficients:
##          ar1     ar2     ma1    mean
##       0.4492  0.3789  0.6807  0.9346
## s.e.  0.1332  0.1272  0.1050  0.2932
##
## sigma^2 estimated as 0.2249:  log likelihood=-155.05
## AIC=320.11   AICc=320.37   BIC=337.34
```

The automatically selected model is ARMA(2,1), with associated $AIC = 320.1070576$. This is the model we thought were best based on the ljung box test statistics above, but chose not to proceed with as we noticed the AIC would select it for us anyway. We now have two models to compare, the ARMA(1,1) and the ARMA(2,1). Note, we did not let R consider differenced models, so it only considered models in the ARMA(p,q) space.

Residual plot

We provide a residuals plot as a quick check the selected models haven't missed anything structurally important. Ideally, the following plot shows no apparent pattern and has most standardised residuals within a 1.96 standard deviation band.



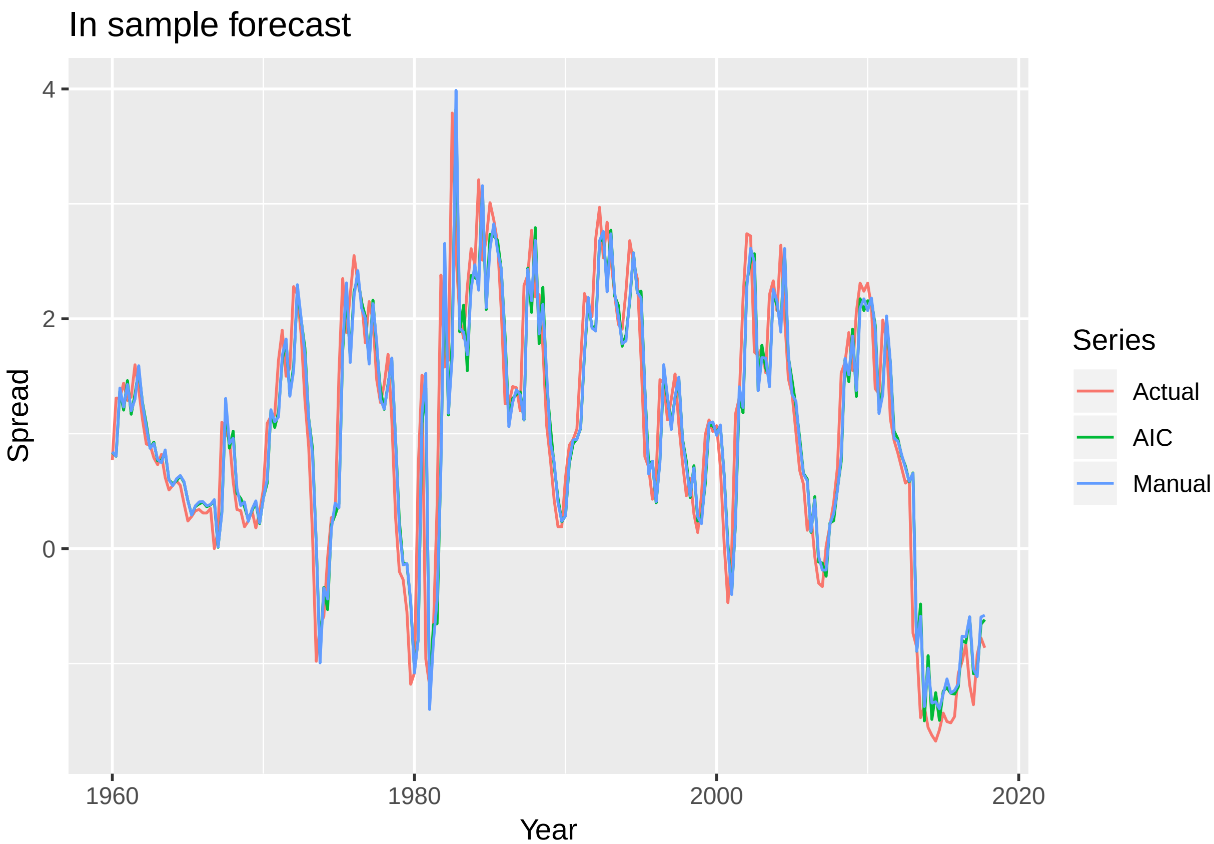


There is some indication in the graphs that they fail to account for variance around 1981. However, the models seem to perform similarly well and overall show no fatal tendencies of bias.

Forecasts

We provide both in sample and out of sample forecasts with corresponding MSPE. ### In sample forecast

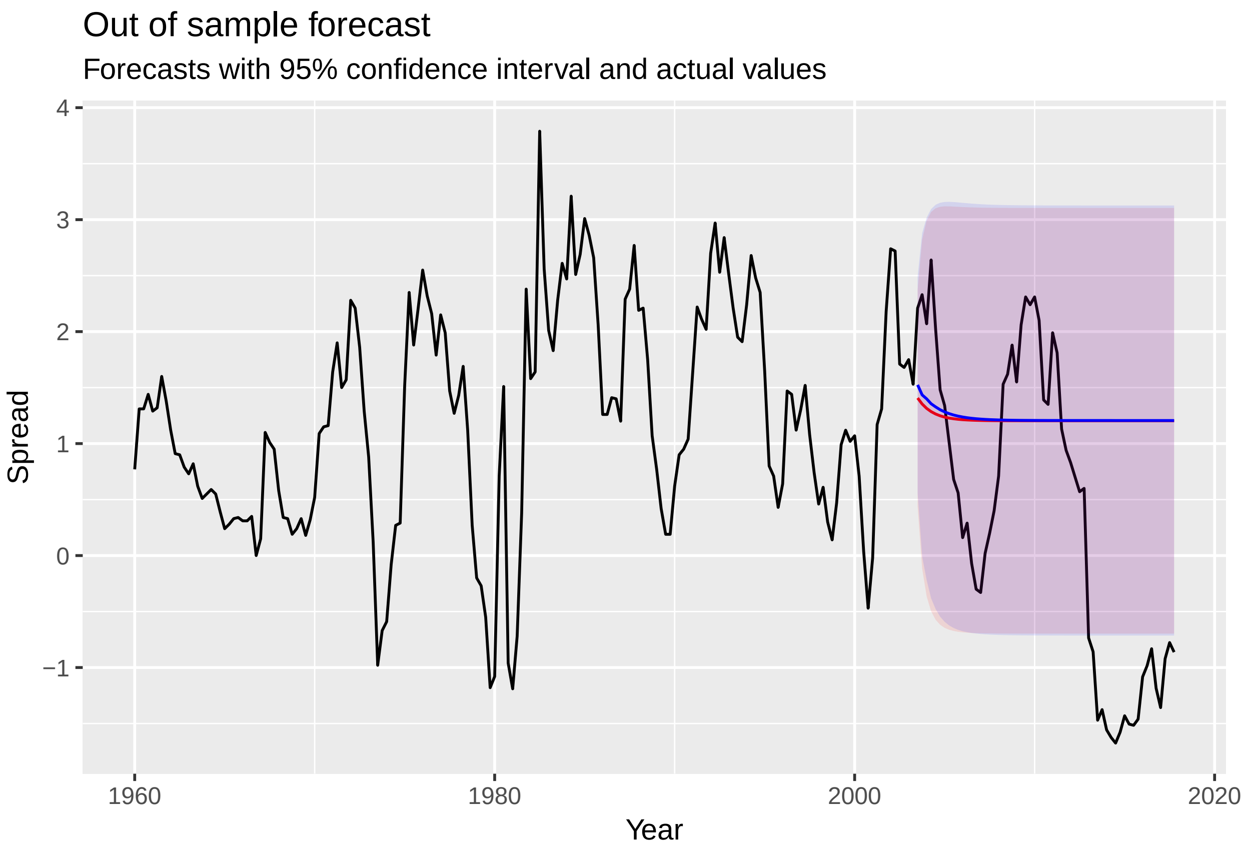
We compare the three models using the MSPE after first showing their performance on in sample forecasting one step ahead.



The in sample $MSPE_{arma(2,1)} = 0.2210262$ and $MSPE_{arma(1,1)} = 0.2260442$. That is, the automatically selected ARMA(2,1) (AIC in the graph) is better than the ARMA(1,1) we choose using the ACF and PACF. However, the forecasts look very similar to one another and the MSPE are very close to one another.

Out of sample forecast

We also want to consider out of sample forecast. This can be seen as another useful tool to determine which of the two models is actually best at predicting future observations.



The red line shows the prediction of ARMA(1,1) and the blue line shows the prediction of the ARMA(2,1). Their forecast intervals overlap to a great extent. The ARMA(2,1) is slower to converge to the mean than the ARMA(1,1), which is not surprising.

The MSPE for the out of sample forecast is $MSPE_{out} = 3.4882006$ while the AIC selected model yields $MSPE_{out, aic)=$4.2854674.

Question 2b

The data indicates there might be a structural change around 1981 Q4. If that is the case, the underlying assumptions of the analysis above might be wrong. We test for this using the Chow test, a version of the regular F test. Here with code provided since the test had to be done manually:

```
# Chow test at breakpoint, t = 1981Q4.
# 1981Q4 corresponds to the following row number in the data:
sc_point <- match(1981.75,df$time) # gives the index position of 1981.75
p <- 2
q <- 2

model <- arima(df$spread, order = c(p,0, q))
model_1 <- arima(df$spread[1:sc_point], order = c(p, 0, q)) # until and incl. SC
model_2 <- arima(df$spread[(sc_point+1):nrow(df)], order = c(p, 0,q)) # after SC
SSR <- sum(model$residuals * model$residuals)
SSR_1 <- sum(model_1$residuals * model_1$residuals)
SSR_2 <- sum(model_2$residuals * model_2$residuals)
n <- p + q + 1
t <- nrow(df)
```

```
Chow <- ((SSR-SSR_1-SSR_2)/n) / ((SSR_1+SSR_2)/(t-2*n))
Chow
```

```
## [1] 3.605993
```

What do you conclude?

If the restriction is not binding, meaning that the coefficients are equal, the F-test should equal zero. If the F-test indicates that there is a sufficient difference between the models used on the pre-break and post-break-data then there has been a structural break. In our case the F-test generates a value of 3.61. The F critical value is 1.84727. We can conclude that since the Chow test of 3.61 is significantly larger than the F critical value of .1.84 the null hypothesis can be rejected and we therfore conclude that there was a structural break.

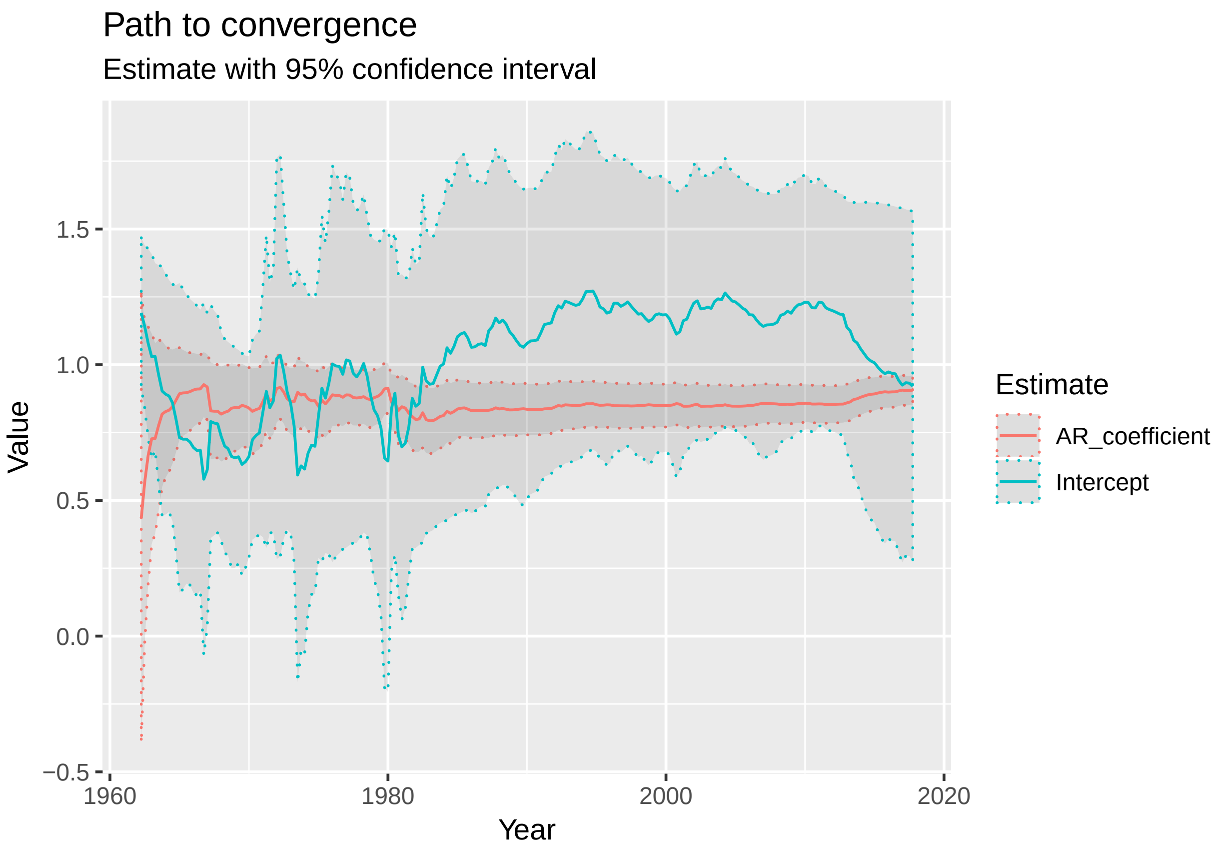
Will your conclusions change if you change the breakpoint?

If we were to change the breakpoint we could end up getting a chow-test showing that there was no structural break. If we would have divided the series into two samples that seperately manages to explain the totals SSR relatively good, The Chow-test would generate a low F-stat and hence make it less the assumption, that the coefficient are equal, less restrictive.

Fortunately, when looking at the data we feel quite confident about picking the correct break point. There can be other major structural breaks in the data that we are not aware of. Thereto, it is unlikely that the structural change is not an effect of a process that has evolved over time. Hopefully, the further analysis below will give us more information about this matter.

Quesiton 2c

Question: Estimate an AR(1) process with intercept recursively for the sample sizes n, n + 1, . . . , T − 1, T where n = 10. Plot the estimation results for the intercept and the AR(1) coefficient with +-2 standard deviation bands (see p. 107 panel (b)). What do you conclude?

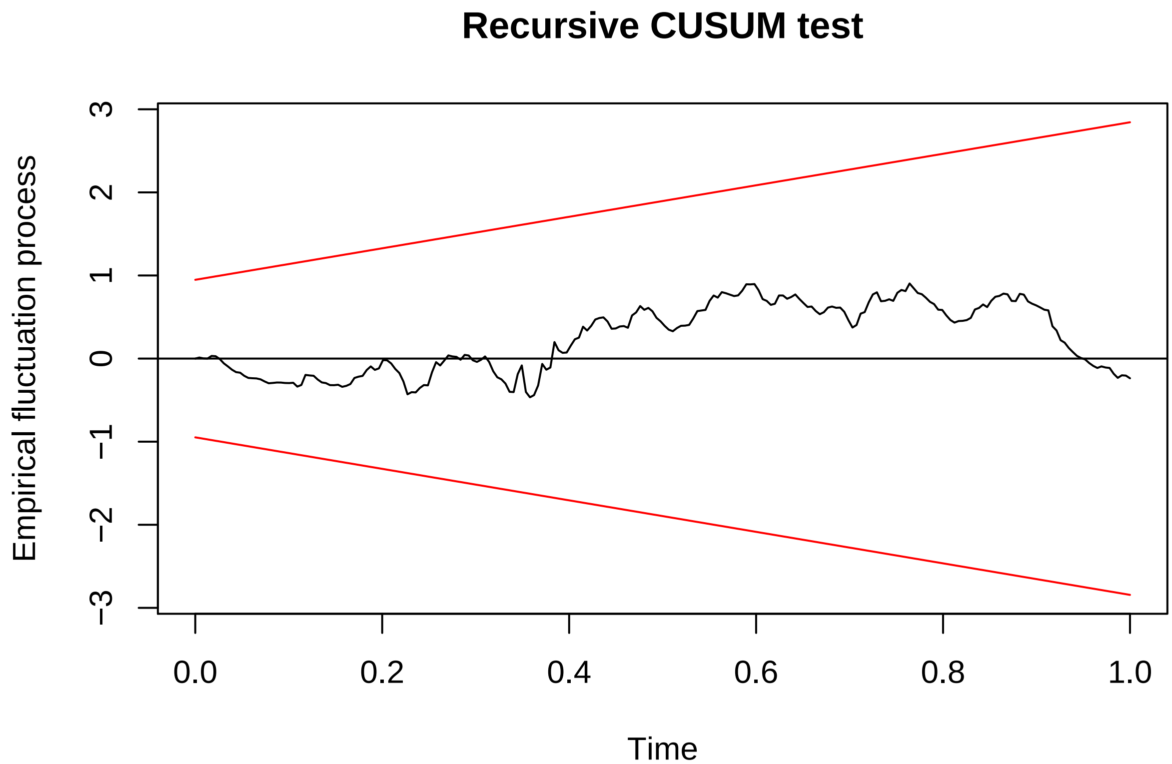


From the graph above it is clear that the estimates seem rather reasonable up until the structural breaking point 1981 Q4. After this point, the estimates drop, however to a very small extent. This could be an indication of a structural change. However, one could have expected a larger indication.

Note that the intercept is much less certain than the ar1 estimate.

Question 2d

Question 2D: For the same sample sizes as in 2c, calculate the CUSUM test accompanied with +-2 standard deviation bands (see p. 107 panel (c)). What do you conclude?



What do you conclude from above?

The CUSUM test shows the CUSUM´s are clearly within the 95 % prediction interval for the entire span of observations. After 1981 Q4, that would be around the point of 0.4 in the graph, the CUSUM's start to rise. As it does not rise above the confidence interval we cannot reject the coefficient stability hypothesis. Another interesting observation is that the CUSUM's start to decrease rapidly right before the end of the series. Maybe, this is an indication of a structural break at a later point in the data.