Assignment 4

Filip Mellgren, Ismael Moreno Martinez, Athina Swahn

2019-03-10

Part I

1 a

Rewriting the equation in the following way:

$$-\alpha_s(r_{Lt-1} - \beta r_{St-1} - \mu) = -\Delta r_{St} + \sum_{i=1}^{2} a_{i,11}\Delta r_{St-1} + \Delta a_{i,12}\Delta r_{Lt-i} + \epsilon_{St}$$

It is clear from the equation above that all RHS variables are I(0). Therefore, the linear combination $r_{Lt-1} - \beta r_{St-1}$ must be a stationary process. That would mean that r_{St} and r_{Lt} must have a stochastic trend in common. Hence, the cointergration vector is

$$B = \begin{bmatrix} 1 \\ -\beta \end{bmatrix}$$

The long run equilibrium is defined by $r_{Lt} = \beta r_{St}$

1 b

The notion that r_{Lt} does not Granger cause r_{St} means that past realisations of $r_{Ls}, s < t$ don't predict r_{St}. Formally:

$H_0 : \alpha_S = \alpha_{1,12} = \alpha_{2,12} = 0$

1 c

Because of the cointegrating relationship, any deviations from the long run equilibrium will revert back, as the residual series follow a stationary process. The return to the equilibrium is described by the bivariate error correction system.

Starting in equilibrium at time t, the error correction is as follows for case i) and ii):

- $\Delta r_{St} = \alpha_S(r_{Lt-1} - \beta r_{St-1} - \mu)$
- $\Delta r_{Lt} = -\alpha_L(r_{Lt-1} - \beta r_{St-1} - \mu)$

At time $t+1$ we obtain the expected change:

- $\Delta r_{St+1} = \alpha_S(r_{Lt} - \beta r_{St} - \mu) + a_{1,11}\Delta r_{St} + a_{1,12}\Delta r_{Lt}$
- $\Delta r_{Lt+1} = -\alpha_L(r_{Lt} - \beta r_{St} - \mu) + a_{1,21}\Delta r_{St} + a_{1,22}\Delta r_{Lt}$

And so on until we have that the change in 0 arising from $r_{Lt} = \beta r_{St} + \mu$ and that the two prior lags are also zero, which will happen in a converging way, with changes diminishing over time.

In i), the system long run rate is too high relative to the short run rate, and in ii) the opposite holds. The two rates will adjust until they reach equilibrium.

Assuming $\alpha_L = 0$ simplifies the model. In period t:

- $\Delta r_{St} = \alpha_S(r - L_{t-1} - \beta r_{St-1} - \mu)$
- $\Delta r_{Lt} = 0$

In $t + 1$:

- $\Delta r_{St+1} = \alpha_S(r_{Lt} - \beta r_{St} - \mu) + a_{1,11}\Delta r_{St} + a_{1,12}\Delta r_{Lt}$
- $\Delta r_{Lt+1} = a_{1,21}\Delta r_{St}$

Assuming for simplicity, that the disequilibrium is of magnitude 1 in all of the cases. In that case, controlling for previous adjustments, the average impact of this disequilibrium on the adjustments at time t will be of α_S for r_{St} and of $-\alpha_L$ for r_{Lt}. Notice that in case (i), given the signs of these effects, this reduces the LHS and augments the RHS. This makes the series to come closer to the equilibrium where $r_{Lt} = \beta r_{St} + \mu$. At the same time, this also happens in case (ii), but in the opposite direction. If the time series remain apart from the long-run relation, this process of adjustments will be re-iterated, reducing the disequilibrium until it comes to meet the co-integration vector.

Cases (iii) and (iv) are pretty similar, nevertheless, here only r_{St} makes a direct response towards the equilibrium. See Figure 1 for a graphic representation of the adjustments departing from disequilibriums (i), (ii), (iii) and (iv).

1d

We can't run the regression because it's a spurious correlation. Given that $r_{Lt} - \beta r_{St} - \mu = \epsilon_t \sim I(1)$, you cannot assume stationarity of the error term. Hence, none of the regular assumptions for a valid OLS error term apply in this case (e.g. $E(\epsilon_t) \neq 0$). On the context of an VECM, this would lead to an imbalance as you would be estimating an stationationary process (Δr_{St}) through an I(1) regressor.

Part II

Do exercise 4 (but not 4f) in the textbook (pp.402-403). Remark: It is possible that the values you obtain differ from those reported in the text to the exercise since the sample is extended. However, the main conclusions should be the same.

```
## -- Attaching packages --------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.1.0       v purrr   0.2.5
## v tibble  2.0.1       v dplyr   0.8.0.1
## v tidyr   0.8.2       v stringr 1.3.1
## v readr   1.3.1       v forcats 0.3.0

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Begin by plotting the series:

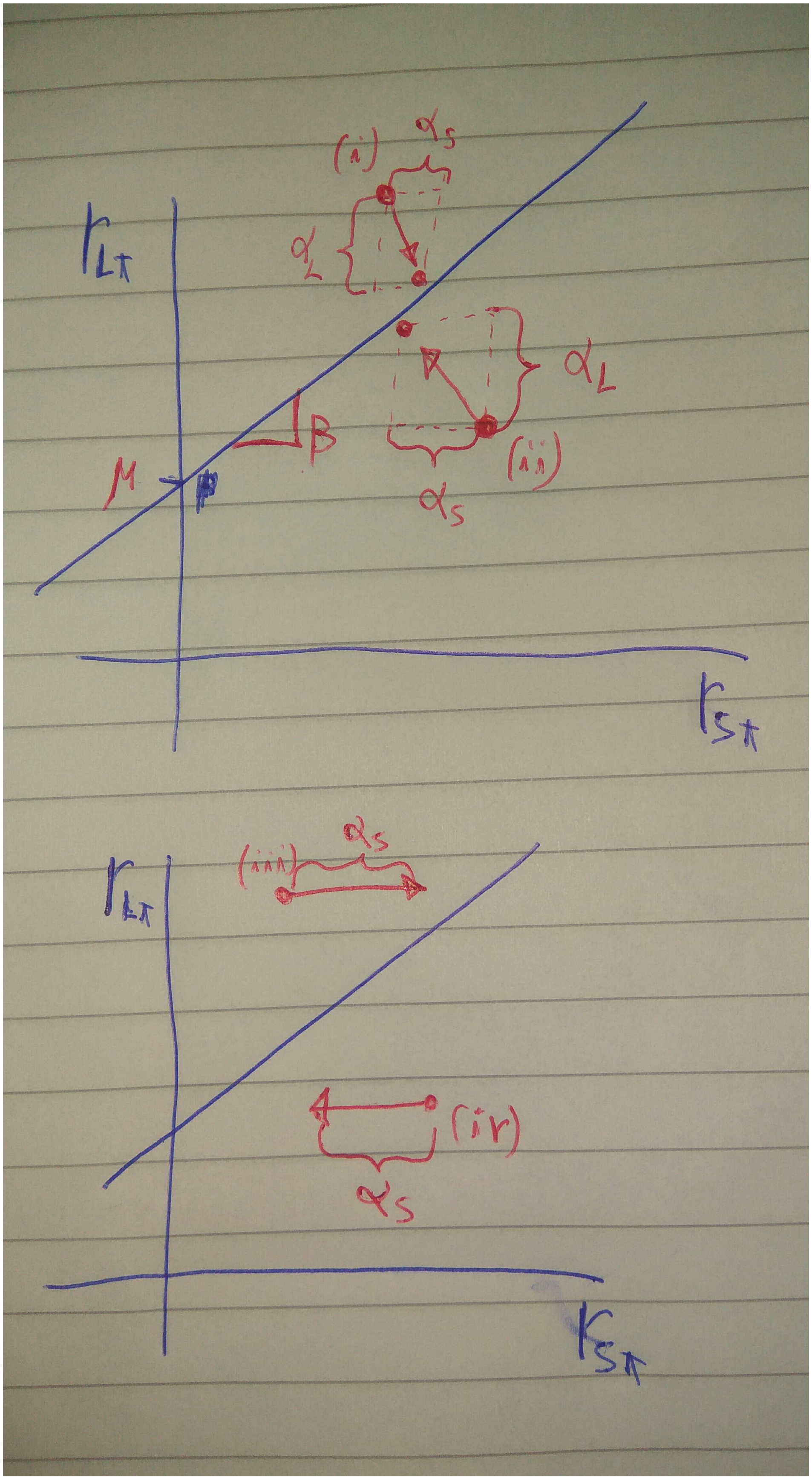
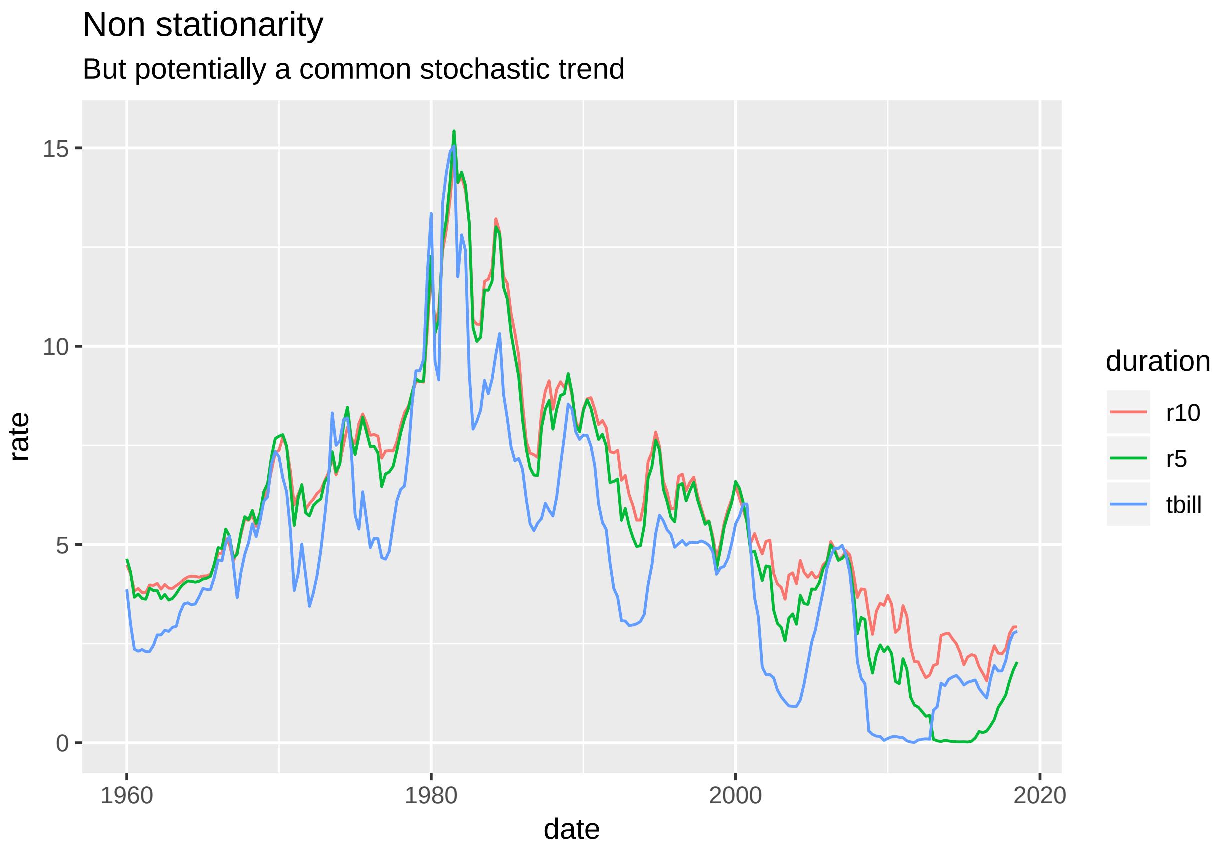


Figure 1: Error Correction Adjustments



4a

Pretest to show that all variables act as unit root processes using ADF with lag length equal to the longest lag length with significicant at the 5% level, including an intercept but not a time trend.

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3429 -0.2288  0.0241  0.3020  3.2625
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.13340    0.08227   1.621  0.10636
## z.lag.1     -0.02805    0.01484  -1.890  0.06011 .
## z.diff.lag1  0.38160    0.06547   5.829 1.99e-08 ***
## z.diff.lag2 -0.34302    0.07012  -4.892 1.94e-06 ***
```

```
## z.diff.lag3  0.39176    0.07312   5.357 2.14e-07 ***
## z.diff.lag4 -0.11781    0.07713  -1.527  0.12810
## z.diff.lag5  0.19271    0.07357   2.619  0.00943 **
## z.diff.lag6 -0.05730    0.07031  -0.815  0.41600
## z.diff.lag7 -0.21134    0.06613  -3.196  0.00160 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6244 on 218 degrees of freedom
## Multiple R-squared:  0.2843, Adjusted R-squared:  0.258
## F-statistic: 10.83 on 8 and 218 DF,  p-value: 8.21e-13
##
##
## Value of test-statistic is: -1.8898 1.7859
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81

##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.21077 -0.28442  0.01613  0.26136  1.58504
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.052676   0.073040   0.721  0.47156
## z.lag.1     -0.010683   0.011091  -0.963  0.33651
## z.diff.lag1  0.297352   0.066791   4.452 1.36e-05 ***
## z.diff.lag2 -0.216159   0.069725  -3.100  0.00219 **
## z.diff.lag3  0.200865   0.070757   2.839  0.00496 **
## z.diff.lag4 -0.001707   0.072040  -0.024  0.98111
## z.diff.lag5 -0.112661   0.070956  -1.588  0.11379
## z.diff.lag6  0.046019   0.069413   0.663  0.50805
## z.diff.lag7 -0.165468   0.066977  -2.471  0.01426 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5189 on 218 degrees of freedom
## Multiple R-squared:  0.1517, Adjusted R-squared:  0.1206
## F-statistic: 4.874 on 8 and 218 DF,  p-value: 1.512e-05
##
##
```

```
## Value of test-statistic is: -0.9632 0.5005
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.03942 -0.27816  0.01328  0.24547  1.43204
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.07605    0.07287   1.044   0.2978
## z.lag.1     -0.01292    0.01072  -1.205   0.2294
## z.diff.lag1  0.28211    0.06624   4.259 3.03e-05 ***
## z.diff.lag2 -0.11886    0.06885  -1.726   0.0857 .
## z.diff.lag3  0.14655    0.06882   2.129   0.0343 *
## z.diff.lag4 -0.01911    0.06883  -0.278   0.7815
## z.diff.lag5 -0.14290    0.06655  -2.147   0.0329 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4485 on 222 degrees of freedom
## Multiple R-squared:  0.1174, Adjusted R-squared:  0.09357
## F-statistic: 4.923 on 6 and 222 DF,  p-value: 9.491e-05
##
##
## Value of test-statistic is: -1.2053 0.7363
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

The appropriate lag lengths for the extended data set are 7 for the tbill and r5, and 5 for r10, which differs somewhat from the lag lengths in Enders p.402. Comment, can we reject the null of a unit root?

4b

Use the Engle-Granger procedure to estimate cointergrating relationships.

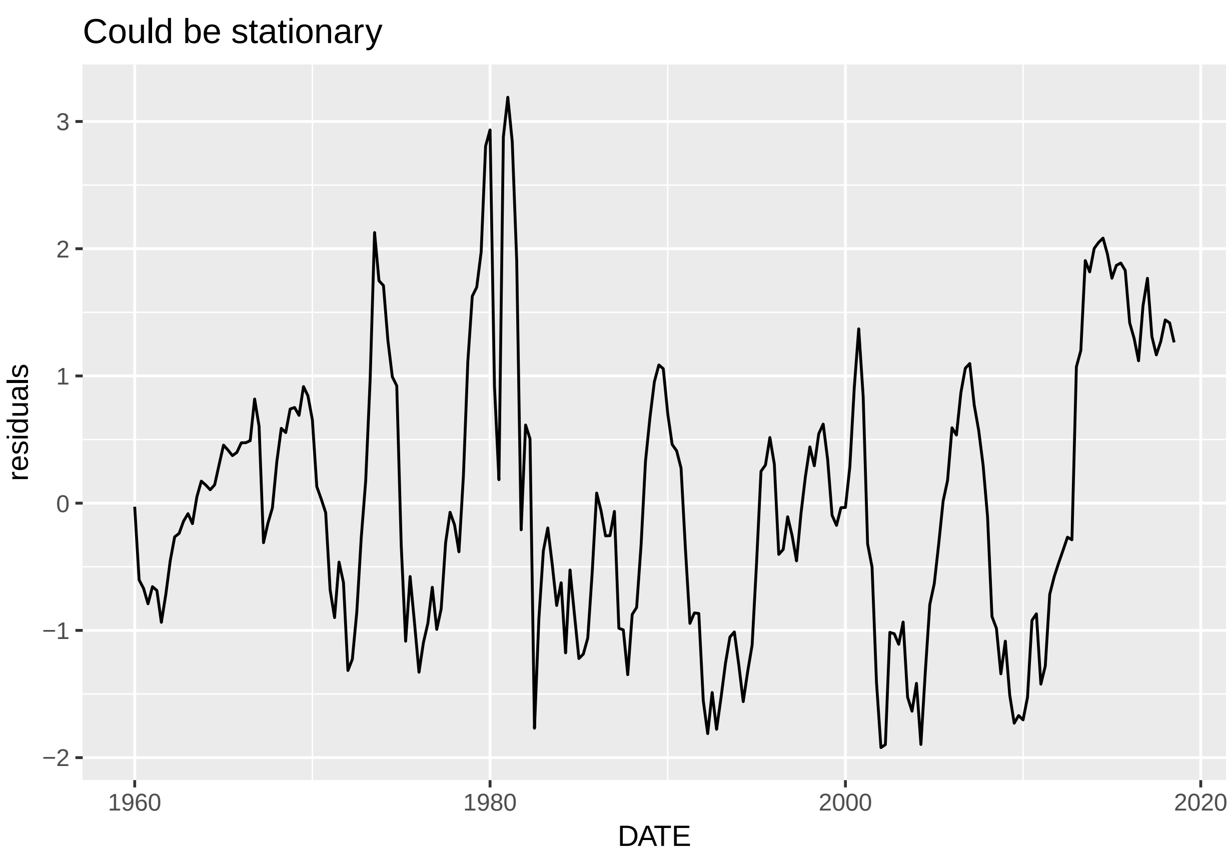
We begin by running the regression:

$Tbill_t = a_0 + a_1 R5_t + a_2 R10_t$

```
##
## Call:
## lm(formula = tbill ~ r5 + r10, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9214 -0.8683 -0.1057  0.6633  3.1906
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.07301    0.23512  -0.311    0.756
## r5           0.98553    0.13691   7.198  8.4e-12 ***
## r10         -0.13409    0.15578  -0.861    0.390
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.082 on 232 degrees of freedom
## Multiple R-squared:  0.8725, Adjusted R-squared:  0.8714
## F-statistic: 793.5 on 2 and 232 DF,  p-value: < 2.2e-16
```

The estimates differ from those in the book. However, when we estimated the same regression using the data used in the book, we obtianed similar values. We proceed by testing if the residuals are serially conintegrated using ADF-tests.

We plot the residual series:



Next, we perform ADF test on residuals (i.e. Engle Granger). We select lag length using the AIC criteria, with a maximal lag length of 10.

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.72942 -0.19718  0.02017  0.23527  2.48827
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## z.lag.1     -0.1743042  0.0395451  -4.408 1.65e-05 ***
## z.diff.lag1  0.2940184  0.0688267   4.272 2.92e-05 ***
## z.diff.lag2 -0.0973810  0.0716596  -1.359 0.175597
## z.diff.lag3  0.2434532  0.0703014   3.463 0.000645 ***
## z.diff.lag4  0.0005642  0.0705065   0.008 0.993623
## z.diff.lag5  0.1734476  0.0705443   2.459 0.014738 *
## z.diff.lag6 -0.0444056  0.0706463  -0.629 0.530306
## z.diff.lag7 -0.0374750  0.0705873  -0.531 0.596037
## z.diff.lag8  0.1411249  0.0680282   2.075 0.039228 *
## z.diff.lag9  0.1243237  0.0683708   1.818 0.070405 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4585 on 214 degrees of freedom
## Multiple R-squared:  0.2005, Adjusted R-squared:  0.1632
## F-statistic: 5.368 on 10 and 214 DF,  p-value: 4.48e-07
##
##
## Value of test-statistic is: -4.4077
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

We need Engle - Granger critical values to evaluate the test statistic,-4.408 against -3.785. The test statistic is small enough for us to reject the null of the series having a unit root, hence the series is stationary and there is cointegration between the variables.

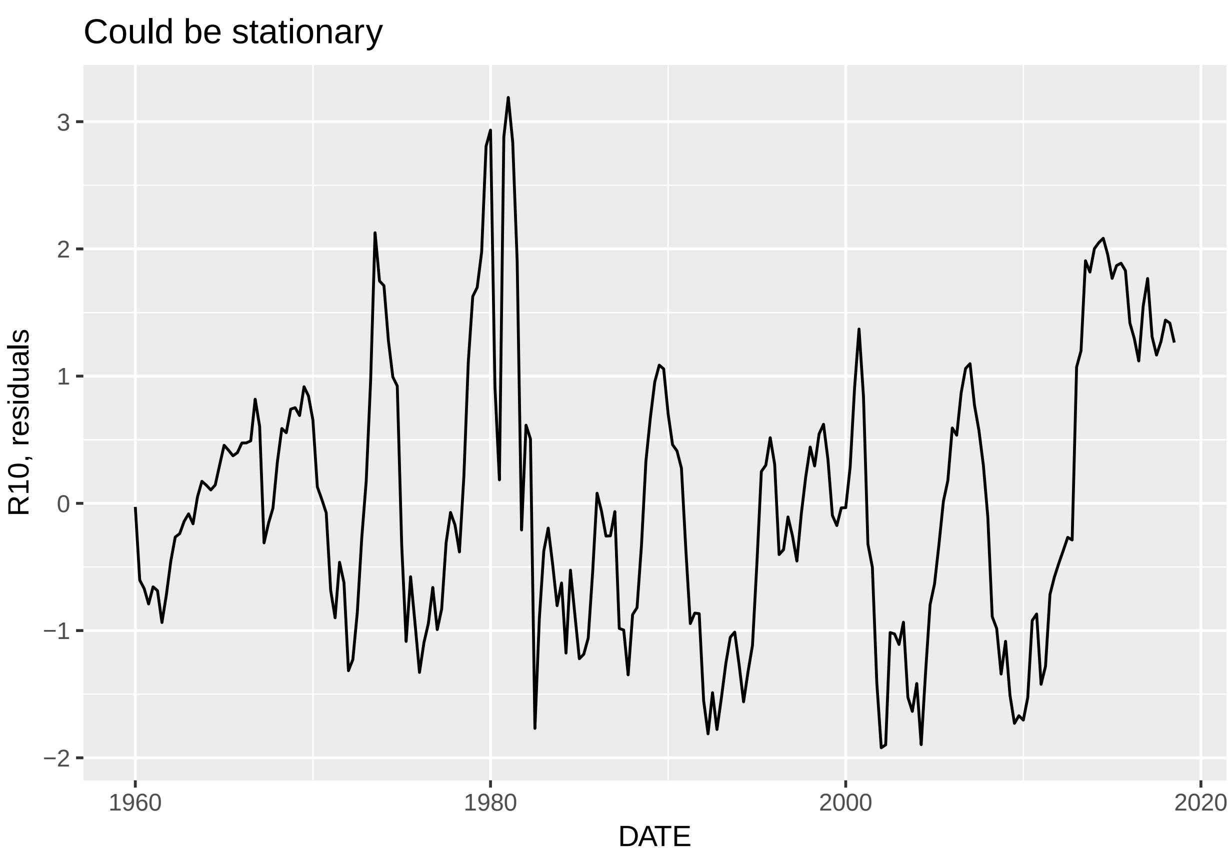
Next, we test whether the residuals follow a white noise process. We use the Ljung box test with a lag length of 9 (the optimal number by the AIC in the ADF test).

```
##
##  Box-Ljung test
##
## data:  residuals
## X-squared = 582.14, df = 9, p-value < 2.2e-16
```

Given the low p-value, we reject the null hypothesis of independence, meaning the residuals are correlated. Anyway, we move on.

4c

```
##
## Call:
## lm(formula = r10 ~ tbill + r5, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.73394 -0.39245 -0.01576  0.26473  1.56788
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.19415    0.06038  19.778   <2e-16 ***
## tbill       -0.02374    0.02758  -0.861     0.39
## r5           0.88829    0.02567  34.601   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4554 on 232 degrees of freedom
## Multiple R-squared:  0.9747, Adjusted R-squared:  0.9745
## F-statistic:  4464 on 2 and 232 DF,  p-value: < 2.2e-16
```




```
##
## ################################################
```

```
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.45318 -0.07318 -0.00511  0.07413  0.78798
##
## Coefficients:
##            Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -0.06210    0.02166  -2.867  0.00454 **
## z.diff.lag  0.12906    0.06660   1.938  0.05391 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1428 on 222 degrees of freedom
## Multiple R-squared:  0.04517,    Adjusted R-squared:  0.03657
## F-statistic: 5.251 on 2 and 222 DF,  p-value: 0.005912
##
##
## Value of test-statistic is: -2.8674
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

A test statistic value of -2.8673815 is obtained and evaluated against -3.785, and we are not able to reject the null of a unit root. This means we can't draw the conclusion that the three variables are cointegrated form this test. This contradicts what we found earlier and indicates that the Engle Granger test might not be the most suitable test for cointegration.

4d

Use the Johansen procedure:

```
##
## ######################
## # Johansen-Procedure #
## ######################
##
## Test type: trace statistic , without linear trend and constant in cointegration
##
## Eigenvalues (lambda):
## [1] 1.356102e-01 3.228729e-02 4.987719e-03 2.922355e-17
##
## Values of teststatistic and critical values of test:
##
##           test 10pct  5pct  1pct
## r <= 2 |  1.14  7.52  9.24 12.97
```

```
## r <= 1 |  8.62 17.85 19.96 24.60
## r = 0  | 41.85 32.00 34.91 41.07
##
## Eigenvectors, normalised to first column:
## (These are the cointegration relations)
##
##            tbill.l7     r5.l7     r10.l7  constant
## tbill.l7  1.0000000  1.000000  1.0000000  1.000000
## r5.l7    -0.1046992 -6.964447 -1.5725779 -2.126845
## r10.l7   -0.7718248  7.146788 -0.2333131  1.023022
## constant  0.5526140 -9.421788  4.7112363  4.789007
##
## Weights W:
## (This is the loading matrix)
##
##            tbill.l7         r5.l7      r10.l7      constant
## tbill.d -0.13398724 -0.0102461035 0.011556133 -1.988484e-16
## r5.d     0.04611983 -0.0007305243 0.010908509 -4.830457e-17
## r10.d    0.06616040 -0.0074301156 0.008508246 -1.137027e-16
```

i) By the trace test, we can reject that there are 0 distinct cointegrating vectors, meaning there is some cointegration. However, we can't reject that there are less or equal to 1 cointegrating relationships, so we conclude that there is only one cointegrating relationship present in the data.

```
##
## ######################
## # Johansen-Procedure #
## ######################
##
## Test type: maximal eigenvalue statistic (lambda max) , without linear trend and constant in cointegra
##
## Eigenvalues (lambda):
## [1] 1.356102e-01 3.228729e-02 4.987719e-03 2.922355e-17
##
## Values of teststatistic and critical values of test:
##
##           test 10pct  5pct  1pct
## r <= 2 |  1.14  7.52  9.24 12.97
## r <= 1 |  7.48 13.75 15.67 20.20
## r = 0  | 33.23 19.77 22.00 26.81
##
## Eigenvectors, normalised to first column:
## (These are the cointegration relations)
##
##            tbill.l7     r5.l7     r10.l7  constant
## tbill.l7  1.0000000  1.000000  1.0000000  1.000000
## r5.l7    -0.1046992 -6.964447 -1.5725779 -2.126845
## r10.l7   -0.7718248  7.146788 -0.2333131  1.023022
## constant  0.5526140 -9.421788  4.7112363  4.789007
##
## Weights W:
## (This is the loading matrix)
##
##            tbill.l7         r5.l7      r10.l7      constant
## tbill.d -0.13398724 -0.0102461035 0.011556133 -1.988484e-16
```

```
## r5.d     0.04611983 -0.0007305243 0.010908509 -4.830457e-17
## r10.d    0.06616040 -0.0074301156 0.008508246 -1.137027e-16
```

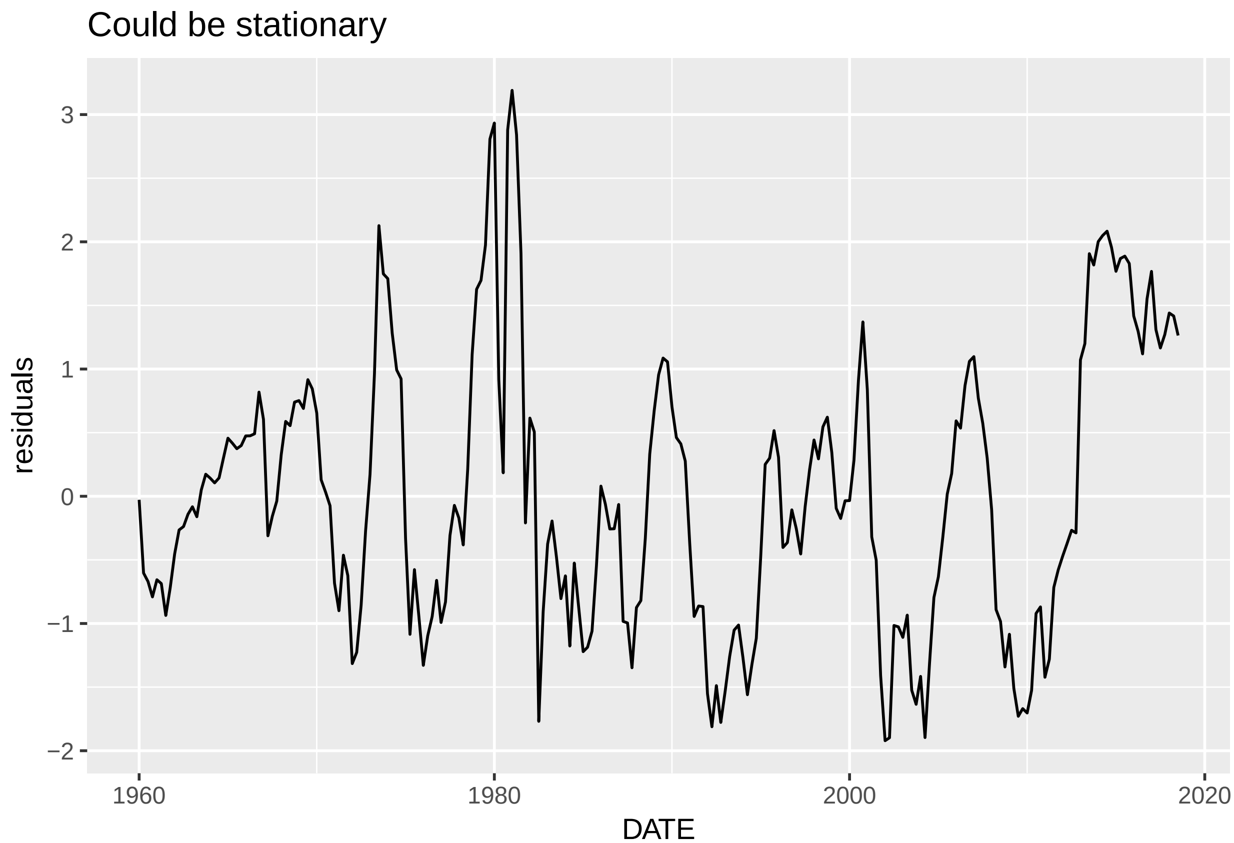
ii) The observed test statistics support the notion that there is one cointegrating relationship present in the data.

We don't verify that the cointrgrated vector is as was stated in the book since the data has changed. In theory, we'd obtain a zero vector after applying the constants to the data, but we don't have the proper constants.

4e

Check to determine whether the individual interest pairs are cointegrated. In particular, is R5 with cointegrated with R10.

```
##
## Call:
## lm(formula = r5 ~ r10, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.85635 -0.30065  0.02157  0.43399  0.81289
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.20322    0.08027  -14.99   <2e-16 ***
## r10          1.12325    0.01188   94.54   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5179 on 233 degrees of freedom
## Multiple R-squared:  0.9746, Adjusted R-squared:  0.9745
## F-statistic:  8937 on 1 and 233 DF,  p-value: < 2.2e-16
```




```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.87751 -0.08317  0.00650  0.08715  0.53146
##
## Coefficients:
##            Estimate Std. Error t value Pr(>|t|)
## z.lag.1    -0.06355    0.02193  -2.898  0.00413 **
## z.diff.lag  0.13623    0.06657   2.046  0.04191 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1659 on 222 degrees of freedom
## Multiple R-squared:  0.04709,    Adjusted R-squared:  0.0385
## F-statistic: 5.485 on 2 and 222 DF,  p-value: 0.004731
##
##
```

```
## Value of test-statistic is: -2.8985
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

The observed value of the test statistic is -2.8984804 which is not small enough to reject the null of no cointegrating relationship at the 5% level. Remember, we do not use the critical values above, but the Engle Granger cirtical values. For our sample with two variables, the 5% critical value is about -3.37.