The University of Queensland
AUSTRALIA
CREATE CHANGE

ECON3350 - Applied Econometrics for Macroeconomics and Finance

Tutorial 8: Cointegration - I

Tutor: Francisco Tavares Garcia

Report 2 due 9 May

Research Report 2 - Data, Instructions, Rubric and Template

Attached Files: cay.xlsx (20.786 KB)
3350_Research_Report_2_2025.pdf (155.077 KB)
Rubric for Research Report (106.496 KB)

ECON3350 Research Report 2

ECON3350: Applied Econometrics for Macroeconomics and Finance

Research Report 2

Due date: 9^{th} May 2025, 12:59 (12:59pm)

Please upload your report via the "Turnitin" submission link (in the "Assessment / Research Report 2" folder). Please note that hard copies *will not* be accepted. At the moment, the due date is 12:59 PM on 9^{th} May 2025, but please check BlackBoard regularly for announcements regarding any changes to this. Your report should be a write-up of your answers (in PDF format, single-spaced, and in 12 font size).

As with the first research report, please *do not* include or attach any software specific material such as R source code or output. In particular, you should summarize the output in the report, but do not copy-paste the "dump" produced by the software. This "dump" is usually of poor quality in terms of presentation and contains much irrelevant or 'not relevant enough' material and so can cost you marks.

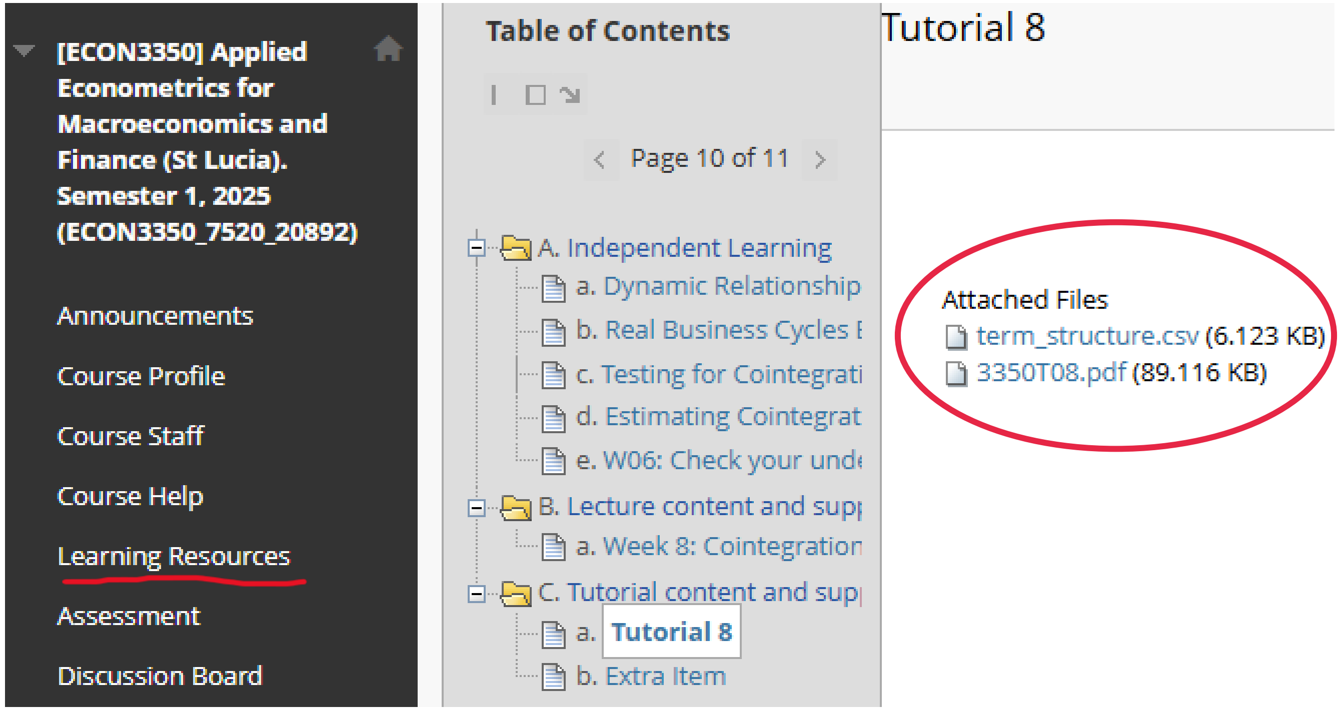
You are allowed to work on this assignment with others, that is, you can discuss how to answer the questions with your classmate(s) and even use AI as a research tool. However, this is not a group assignment, which means that the report must be written individually and by you: you must answer all the questions in your own words and submit your report separately. The marking system will check for similarities and AI content and UQ's student integrity and misconduct policies on plagiarism *strictly apply*.

Tutorial 8: Cointegration - I

At the end of this tutorial you should be able to:

- Automate the task of unit root testing in multiple time-series samples in R;
- Implement the Engle-Granger cointegration test in R;
- Interpret the outcome of an Engle-Granger.
- Use the outcome of the Engle-Granger test to infer possible cointegrating relations.

Let’s download the tutorial and the dataset.



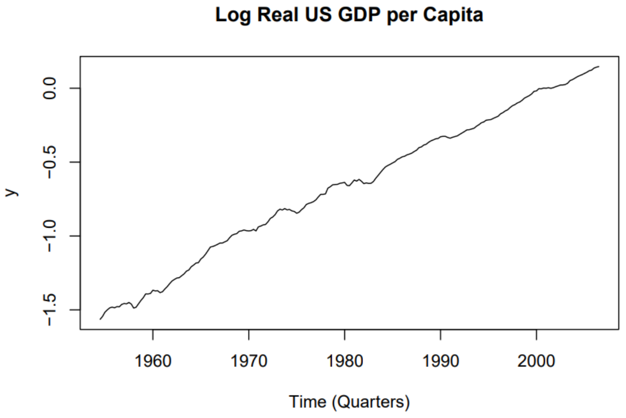
The University of Queensland
Australia

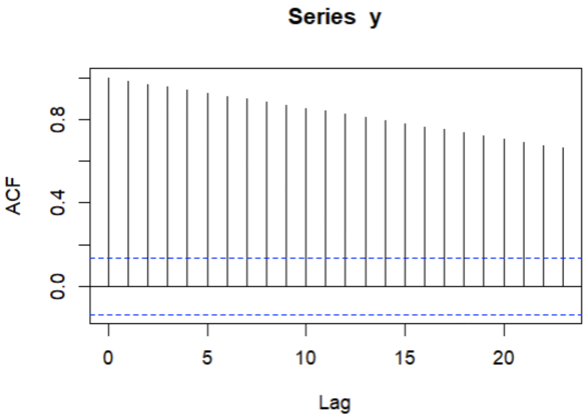
Now, let's download the script for the tutorial.

- Copy the code from Github,
  - https://github.com/tavaresgarcia/teaching
- Save the scripts in the same folder as the data.

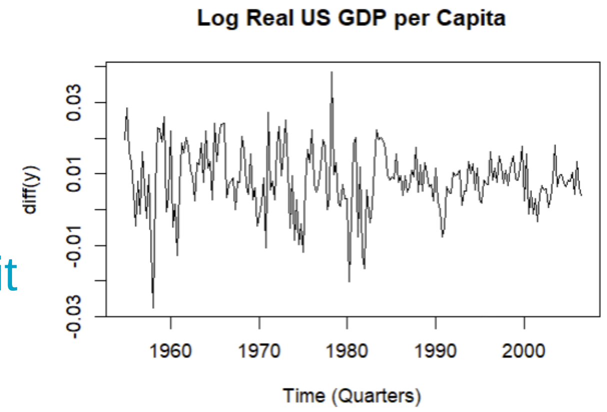
In week 4 we studied the test for unit root ADF

GDP was not empirically distinguishable from an integrated process I(1)



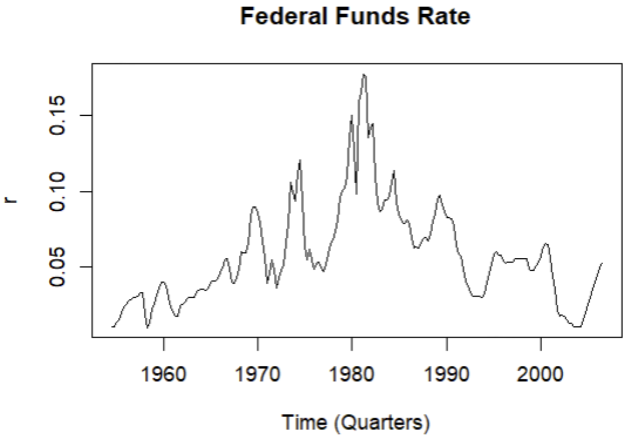


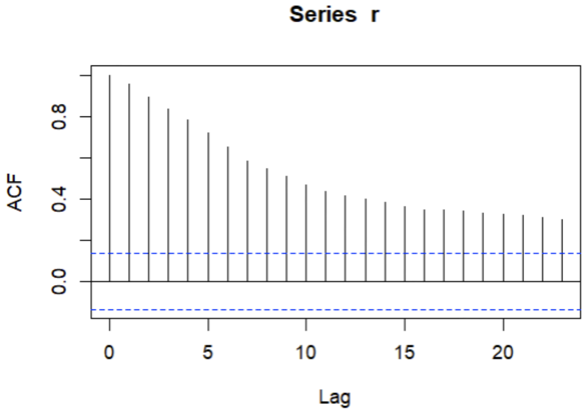
ADF test
Did not reject H_o of the presence of unit root in levels.



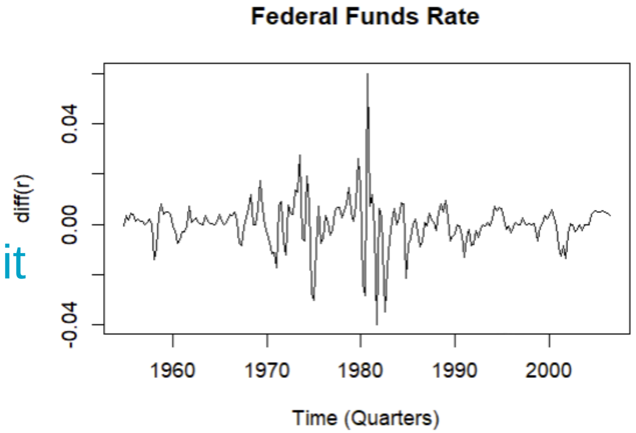
ADF test
Rejected H_o of the presence of unit root in differences.

Interest rates was not empirically distinguishable from an integrated process I(1)





ADF test
Did not reject H_o of the presence of unit root in levels.



ADF test
Rejected H_o of the presence of unit root in differences.

Cointegration

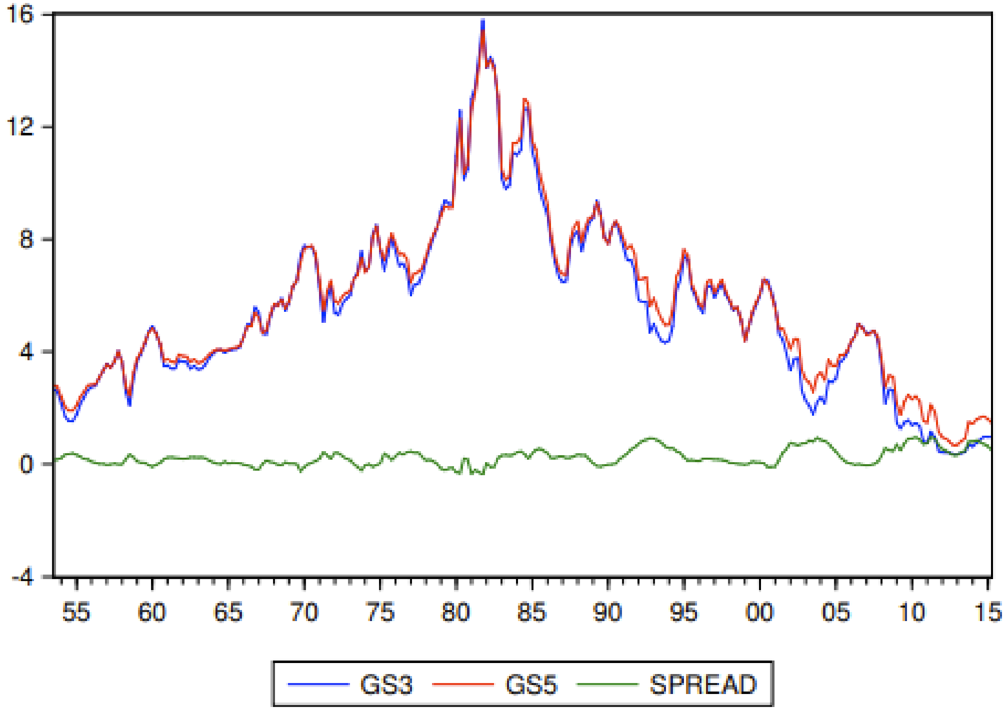
Dynamic Relationships Between I(1) Processes

What are the implications of working with I(1) processes?

In general, linear combinations of I(1) processes yield another I(1) process; however, in some cases a linear combination can result in I(0).

Cointegration: I(1) processes are related such that there exists a linear combination that yields a I(0) process.

Spurious regression: Infer a significant relationship between I(1) processes when in theory one does not exist.



Cointegration and the Engle-Granger test

Testing for Cointegration

When I(1) variables cointegrate, a linear combination will be I(0); otherwise all linear combinations will be I(1).

Basic idea to test for cointegration: given a vector $\mathbf{w}_t$, estimate the regression

$$w_{n,t} = \beta_0 + \beta_1 w_{1,t} + \cdots + \beta_{n-1} w_{n-1,t} + \varepsilon_t,$$

and test the residual

$$\widehat{\varepsilon}_t = w_{n,t} - \widehat{\beta}_0 - \widehat{\beta}_1 w_{1,t} - \cdots - \widehat{\beta}_{n-1} w_{n-1,t}$$

for a unit root.

Can use the Cointegrating ADF or Cointegrating Regression DW test for this: rejecting H_0 implies $\widehat{\varepsilon}_t \sim \mathrm{I}(0)$, and therefore, evidence of cointegration.

Eric Eisenstat Applied Econometrics for Macro and Finance

If we knew β_1 we could use an ADF to test for unit root (like the spread).

But since β_1 is also an estimate, we need to account for the error of this estimation too.

So, today we will use the Engle-Granger test from the function coint.test {aTSA}.

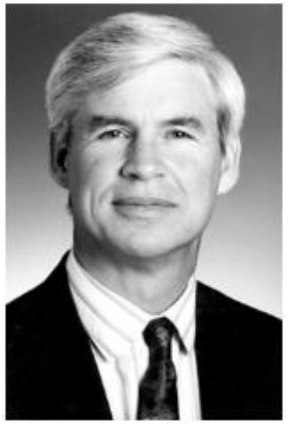


Photo from the Nobel Foundation archive.
Robert F. Engle III
Prize share: 1/2

Photo from the Nobel Foundation archive.
Clive W.J. Granger
Prize share: 1/2

The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2003

Problems

In this tutorial you will test for cointegration using the Engle-Granger method. The data consists of four Australian interest rates: the 5 year (i3y) and 3 year (i3y) Treasury Bond (i.e., Capital Market) rates, along with the 180 day (i180d) and 90 (i90d) day Bank Accepted Bill (i.e., Money Market) rates. The data are annualized monthly rates for the period June 1992—August 2010 ($T = 219$), and are saved in `term_structure.csv`.

	A	B	C	D	E
1	obs	I3Y	I5Y	I90D	I180D
2	1992M06	7.04	7.83	5.64	5.5
3	1992M07	6.22	7.05	5.73	5.66
4	1992M08	7.86	8.41	5.94	5.99
5	1992M09	7.91	8.32	5.86	5.88
6	1992M10	7.89	8.29	5.88	5.92
7	1992M11	8.01	8.48	5.9	5.93
8	1992M12	7.74	8.24	5.88	5.92
9	1993M01	7.49	8	5.82	5.83

1. Analyse the integration properties of each individual process: $\{i3y_t\}$, $\{i5y_t\}$, $\{i90d_t\}$ and $\{i180d_t\}$. Based on the data, what inference can we draw about each of these processes resembling a unit root process?

Solution For this tutorial, we load the following useful packages.

```
library(forecast)
library(dplyr)
library(zoo)
library(aTSA)
```

It is also useful to create some functions to help automate the task of constructing adequate sets for ADF specifications. The following two functions estimate the coefficients and record AIC/BIC values for a range of ADF regressions specified by lags combined with the inclusion and/or excludion of a constant and/or trend.

One function performs the estimation in levels, while the other does the same in differences.

Next, load the data and and extract the four variables.

```
mydata <- read.delim("term_structure.csv", header = TRUE,
                                            sep = ",")

dates <- as.yearqtr(mydata$obs)
i3y <- mydata$I3Y
i5y <- mydata$I5Y
i90d <- mydata$I90D
i180d <- mydata$I180D
```

Now, consider the proximity of $\{i3y_t\}$ to a unit root process. We begin by constructing an adequate set of ADF regressions in the level of $\{i3y_t\}$.

```
i3y_ADF_lev <- ADF_estimate_lev(i3y, p_max = 15)
print(i3y_ADF_lev$ic_aic)
```

```
##       const trend  p      aic      bic
##  [1,]     1     1 10 100.3036 147.6865
##  [2,]     1     1 12 101.0781 155.2301
##  [3,]     1     1 11 101.7485 152.5159
##  [4,]     1     0 12 102.8333 153.6008
##  [5,]     1     1 13 102.9466 160.4830
##  [6,]     1     0 10 103.1882 147.1866
##  [7,]     0     0 12 103.5612 150.9442
##  [8,]     1     1 14 103.5876 164.5085
##  [9,]     1     1  7 103.6625 140.8919
## [10,]     1     0 11 104.2917 151.6746
```

```
print(i3y_ADF_lev$ic_bic)
```

```
##       const trend p      aic      bic
##  [1,]     0     0 1 118.3726 128.5261
##  [2,]     1     0 1 116.0791 129.6171
##  [3,]     1     0 2 113.7241 130.6466
##  [4,]     0     0 2 117.3288 130.8668
##  [5,]     1     0 3 110.8348 131.1417
##  [6,]     1     1 3 107.5159 131.2074
##  [7,]     1     1 2 111.4567 131.7637
##  [8,]     1     1 1 115.0552 131.9776
##  [9,]     0     0 0 125.2605 132.0295
## [10,]     0     0 3 116.1076 133.0301
```

1. Analyse the integration properties of each individual process: $\{i3y_t\}$, $\{i5y_t\}$, $\{i90d_t\}$ and $\{i180d_t\}$. Based on the data, what inference can we draw about each of these processes resembling a unit root process?

	const	trend	p	aic	bic
1	0	0	1	118.3726	128.5261
2	1	0	2	113.7241	130.6466
3	1	0	3	110.8348	131.1417
4	1	0	10	103.1882	147.1866
5	1	1	2	111.4567	131.7637
6	1	1	3	107.5159	131.2074
7	1	1	7	103.6625	140.8919
8	1	1	10	100.3036	147.6865

The AIC and BIC ranking do not have any specifications in common, so we select from both top 10 rankings in a way that reflects some agreement. This is obviously very subjective! The justification we use as follows. From the AIC list, take the most preferred specification along with a few others that have the lowest BIC values. Then, do the same using the BIC list.

As result, we obtain the following set of specifications on which we run our residuals analysis.

```
i3y_adq_set <- as.matrix(arrange(as.data.frame(
                 rbind(i3y_ADF_lev$ic_aic[c(1, 6, 9),],
                       i3y_ADF_lev$ic_bic[c(1, 3, 5:7),])),
                       const, trend, p))
i3y_adq_idx <- match(data.frame(t(i3y_adq_set[, 1:3])),
                 data.frame(t(i3y_ADF_lev$ic[, 1:3])))

for (i in 1:length(i3y_adq_idx))
{
  checkresiduals(i3y_ADF_lev$ADF_est[[i3y_adq_idx[i]]])
}
```

```
        Ljung-Box test

data:  Residuals from Regression with ARIMA(1,0,0) errors
Q* = 20.038, df = 9, p-value = 0.01768

Model df: 1.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(2,0,0) errors
Q* = 18.425, df = 8, p-value = 0.01826

Model df: 2.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(3,0,0) errors
Q* = 16.161, df = 7, p-value = 0.02369

Model df: 3.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(10,0,0) errors
Q* = 5.2834, df = 3, p-value = 0.1522

Model df: 10.   Total lags used: 13
```

```
        Ljung-Box test

data:  Residuals from Regression with ARIMA(2,0,0) errors
Q* = 19.404, df = 8, p-value = 0.01284

Model df: 2.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(3,0,0) errors
Q* = 15.443, df = 7, p-value = 0.03072

Model df: 3.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(7,0,0) errors
Q* = 6.1322, df = 3, p-value = 0.1054

Model df: 7.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(10,0,0) errors
Q* = 4.848, df = 3, p-value = 0.1833

Model df: 10.   Total lags used: 13
```

1. Analyse the integration properties of each individual process: $\{\text{i3y}_t\}$, $\{\text{i5y}_t\}$, $\{\text{i90d}_t\}$ and $\{\text{i180d}_t\}$. Based on the data, what inference can we draw about each of these processes resembling a unit root process?

We reject white noise residuals at the 5% significance level for all models with $p < 7$. Hence, we remove all models except the three with $p = 7, 10$, all containing a constant and two also containing a trend.

Given our adequate set of ADF regressions, we should run the ADF test with `nlag = 11`, but we will use `nlag = 15` just to check how sensitive the results are to including more lags (which the AIC prefers, but the BIC rejects).

```
adf.test(i3y, nlag = 15)
```

For specifications with a constant, no trend, all specifications except $p = 3$, the null cannot be rejected at the 5% significance level. For specifications with a constant and with a trend, the same conclusion holds for all specifications except $p = 3, 4, 7, 9$. Our concern is the one with $p = 7$ since it is in our adequate set.

However, might note that the p-value for $p = 7$ is 0.475, indicating that if we choose 4.75% as the significance level, then we should conclude that the null cannot be rejected for any specification in the adequate set. Is there a great reason to commit to 5% versus 4.75%? That is a question we would need to consider more profoundly in this particular case.

Overall we might lean towards concluding $\{\text{i3y}_t\}$ is not empirically distinguishable from a unit root process, with some ambiguity arising from the specification uncertainty that results from the constant with trend and $p = 7$ specification rejecting a unit root at the 5% significance level (but not the 4.75% level).

For the differenced time series and other variables, please check the solutions pdf.

	const	trend	p	aic	bic
1	0	0	1	118.3726	128.5261
2	1	0	2	113.7241	130.6466
3	1	0	3	110.8348	131.1417
4	1	0	10	103.1882	147.1866
5	1	1	2	111.4567	131.7637
6	1	1	3	107.5159	131.2074
7	1	1	7	103.6625	140.8919
8	1	1	10	100.3036	147.6865

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
      lag    ADF p.value
 [1,]   0 -0.896   0.358
 [2,]   1 -0.697   0.429
 [3,]   2 -1.220   0.242
 [4,]   3 -1.161   0.264
 [5,]   4 -1.103   0.284
 [6,]   5 -1.022   0.313
 [7,]   6 -0.915   0.352
 [8,]   7 -0.950   0.339
 [9,]   8 -0.761   0.406
[10,]   9 -0.773   0.402
[11,]  10 -0.738   0.415
[12,]  11 -0.807   0.390
[13,]  12 -0.709   0.425
[14,]  13 -0.603   0.463
[15,]  14 -0.645   0.448
Type 2: with drift no trend
      lag   ADF p.value
 [1,]   0 -1.74  0.4298
 [2,]   1 -2.05  0.3087
 [3,]   2 -2.80  0.0634
 [4,]   3 -3.10  0.0295
 [5,]   4 -2.85  0.0553
 [6,]   5 -2.38  0.1779
 [7,]   6 -2.22  0.2424
 [8,]   7 -2.65  0.0893
 [9,]   8 -2.32  0.2036
[10,]   9 -2.54  0.1154
[11,]  10 -2.09  0.2924
[12,]  11 -1.97  0.3400
[13,]  12 -1.60  0.4828
[14,]  13 -1.55  0.5018
[15,]  14 -1.37  0.5683
Type 3: with drift and trend
      lag   ADF p.value
 [1,]   0 -2.27  0.4618
 [2,]   1 -2.90  0.1999
 [3,]   2 -3.29  0.0728
 [4,]   3 -3.79  0.0203
 [5,]   4 -3.51  0.0422
 [6,]   5 -2.97  0.1689
 [7,]   6 -2.90  0.1985
 [8,]   7 -3.45  0.0475
 [9,]   8 -3.28  0.0758
[10,]   9 -3.63  0.0310
[11,]  10 -3.05  0.1333
[12,]  11 -2.83  0.2293
[13,]  12 -2.52  0.3550
[14,]  13 -2.65  0.3046
[15,]  14 -2.39  0.4093
```

2. Use the Engle-Granger method to test for a cointegrating relation involving all four processes. Assume the 5 year TB rate is the dependent variable in the initial regression. Hint: Use the `test.coint` function provided by the `aTSA` package.

Solution We need to construct an adequate set of ADF specifications for the estimated residuals from the regression of $i5y_t$ on a constant, $i3y_t$, $i90d_t$, and $i180d_t$. A regression in R is implemented using the `lm` function.

```
eg_reg <- lm( i5y ~ i3y + i90d + i180d, mydata)
eg_res <- eg_reg$residuals
```

Now, use the same approach as in Question 1 but with `eg_res` instead of an observed sample.

```
egr_ADF_lev <- ADF_estimate_lev(eg_res, p_max = 15)
print(egr_ADF_lev$ic_aic)
```

```
##       const trend  p       aic       bic
##  [1,]     0     0  2 -420.1848 -406.6652
##  [2,]     0     0  1 -420.1357 -409.9961
##  [3,]     0     0  0 -419.9250 -413.1652
##  [4,]     0     0  3 -418.5450 -401.6455
##  [5,]     1     0  2 -418.2587 -401.3592
##  [6,]     0     0 15 -418.2065 -360.7482
##  [7,]     1     0  1 -418.1856 -404.6660
##  [8,]     1     0  0 -417.9520 -407.8123
##  [9,]     1     0  3 -416.6074 -396.3280
## [10,]     1     1  0 -416.5675 -403.0479
```

```
print(egr_ADF_lev$ic_bic)
```

```
##       const trend p       aic       bic
##  [1,]     0     0 0 -419.9250 -413.1652
##  [2,]     0     0 1 -420.1357 -409.9961
##  [3,]     1     0 0 -417.9520 -407.8123
##  [4,]     0     0 2 -420.1848 -406.6652
##  [5,]     1     0 1 -418.1856 -404.6660
##  [6,]     1     1 0 -416.5675 -403.0479
##  [7,]     0     0 3 -418.5450 -401.6455
##  [8,]     1     0 2 -418.2587 -401.3592
##  [9,]     1     1 1 -416.5125 -399.6130
## [10,]     1     0 3 -416.6074 -396.3280
```

We will only consider specifications without a constant or trend since we are focusing on residuals.

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

2. Use the Engle-Granger method to test for a cointegrating relation involving all four processes. Assume the 5 year TB rate is the dependent variable in the initial regression. Hint: Use the `test.coint` function provided by the `aTSA` package.

```
egr_adq_set <- as.matrix(arrange(as.data.frame(
                          egr_ADF_lev$ic_bic[c(1, 2, 4, 7), ]),
                          const, trend, p))
egr_adq_idx <- match(data.frame(t(egr_adq_set[, 1:3])),
                     data.frame(t(egr_ADF_lev$ic[, 1:3])))

for (i in 1:length(egr_adq_idx))
{
  checkresiduals(egr_ADF_lev$ADF_est[[egr_adq_idx[i]]])
}
```

```
        Ljung-Box test

data:  Residuals from Regression with ARIMA(0,0,0) errors
Q* = 9.1103, df = 10, p-value = 0.5217

Model df: 0.   Total lags used: 10


        Ljung-Box test

data:  Residuals from Regression with ARIMA(1,0,0) errors
Q* = 7.2714, df = 9, p-value = 0.6089

Model df: 1.   Total lags used: 10


        Ljung-Box test

data:  Residuals from Regression with ARIMA(2,0,0) errors
Q* = 5.1884, df = 8, p-value = 0.7373

Model df: 2.   Total lags used: 10


        Ljung-Box test

data:  Residuals from Regression with ARIMA(3,0,0) errors
Q* = 5.7923, df = 7, p-value = 0.5642

Model df: 3.   Total lags used: 10
```

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

2. Use the Engle-Granger method to test for a cointegrating relation involving all four processes. Assume the 5 year TB rate is the dependent variable in the initial regression. Hint: Use the `test.coint` function provided by the `aTSA` package.

All residuals look OK. Hence, we can continue and use the function `coint.test` from the `aTSA` package to implement the Engle-Granger test for each of the specifications in the adequate set. The inputs to `coint.test` are the dependent variable in the regression, a matrix containing independent variables and the number of lags to use in the unit root test on the residuals series.

We use a `for` loop to implement the test on each of the specifications in the adequate set. Instead of generating output at each iteration, we use the option `output = F` to suppress it and store the p-values in an easy-to-read table.

```
eg_test <- matrix(nrow = 1, ncol = 4)
colnames(eg_test) <- rep("", 4)
rownames(eg_test) <- c("No const, no trend")
for (l in 1:4)
{
  eg_l <- coint.test(i5y, cbind(i3y, i90d, i180d),
                     nlag = l, output = F)
  eg_test[, l] <- eg_l[1, 3]
  colnames(eg_test)[l] <- paste("Lag", l)
}
print(eg_test)
```

```
##                    Lag 1 Lag 2 Lag 3 Lag 4
## No const, no trend  0.01  0.01  0.01  0.01
```

For specifications with no constant and no trend, the unit root in the residuals is rejected at low significance levels. The best inference we can draw is that if the residual in the regression i5y$_t$ on a constant, i3y$_t$, i90d$_t$, and i180d$_t$ is *mean-independent*, then it also does not have a unit root. The reasoning behind this is that if a residual series is mean-independent of the regressors, then its unconditional mean is zero. This scenario matches ADF specifications above that restrict the constant to be zero.

3. Interpret the inference obtained Questions 1 and 2 in terms of empirical evidence of cointegration in the four interest rates.

Solution In Question 1, we concluded that $\mathrm{i3y}_t$ is not empirically distinguishable from I(1), but for the remaining three processes our inference on their proximity to I(1) processes is rather ambiguous.

When we regress $\mathrm{i5y}_t$ on $\mathrm{i3y}_t$, $\mathrm{i90d}_t$ and $\mathrm{i180d}_t$, we find that the residuals process does not have a unit root if we enforce the restriction that residuals are mean-independent. Assuming this restriction is valid, we have the following possibilities:

1. $\mathrm{i3y}_t$, $\mathrm{i5y}_t$, $\mathrm{i90d}_t$ and $\mathrm{i180d}_t$ are all I(0);
2. any three processes are I(1) and cointegrated while a fourth is I(0); for example, we could have that $\mathrm{i3y}_t$, $\mathrm{i5y}_t$ and $\mathrm{i90d}_t$ are cointegrated and $\mathrm{i180d}_t$ is is I(0). The same could hold for any other combination.
3. Any two processes are I(1) and cointegrated while the other two are I(0); for example, we could have that $\mathrm{i3y}_t$ and $\mathrm{i5y}_t$ are cointegrated while $\mathrm{i90d}_t$ and $\mathrm{i180d}_t$ are both I(0).
4. any two processes are I(1) and cointegrated, and the other two processes are also I(1) and cointegrated, but the four processes are not all cointegrated with each other in a single cointegrating relation;
5. all four processes are I(1) and cointegrated in a single cointegrating relation.

Which of these five scenarios prevails? It depends on what we assume about the integration properties of the processes involved. Our unit root tests in Question 1 did not clearly reject a unit root in any of the processes, except $\{\mathrm{i90d}_t\}$. If $\{\mathrm{i90d}_t\}$ is I(0), then we can rule out scenarios 4 and 5.

In terms of scenarios 1-3, we can in principle make the unit root assumption about any combination of $\mathrm{i3y}_t$ on $\mathrm{i5y}_t$ and $\mathrm{i180d}_t$, which will determine the appropriate interpretation. The important thing to remember is that it is *always* an assumption that a unit root exists! Whether or not it is a useful one depends on the application.

4. Repeat Question 2 three more times but each time change the dependent variable. Is the inference regarding cointegration affected?

For solutions, please check the solutions pdf.

5. Next, use the data to test the *expectations theory* of the term structure of interest rates (ETT). Specifically, investigate whether the spreads in the Capital Market (i5y – i3y) and Money Market (i180d – i90d) are stable (and therefore stationary assuming constant variances and auto-covariances).

Solution This is straightforward using the ADF testing approach explained in Question 1. We apply it once to the spread i5y – i3y representing the Capital Market and again to the spread i180d – i90d representing the Money Market. In both cases the spreads are observed samples so there is no special consideration needed to the distribution of the ADF test statistic.

Also note that we do not require any assumptions about unit roots in the DGPs of individual interest rates to draw conclusions about the stationarity of the spreads (such assumptions are only needed if we want to conclude "a stationary spread implies cointegrated interest rates' ').

```
cm_ADF_lev <- ADF_estimate_lev(i5y - i3y, p_max = 15)
print(cm_ADF_lev$ic_aic)
```

```
##       const trend  p       aic       bic
##  [1,]     1     0  2 -557.2660 -540.3435
##  [2,]     1     1  2 -556.7428 -536.4358
##  [3,]     1     0  4 -556.7363 -533.0449
##  [4,]     1     0  3 -556.6236 -536.3166
##  [5,]     1     1  4 -556.4973 -529.4214
##  [6,]     1     1  6 -556.1063 -522.2614
##  [7,]     1     1 15 -556.0468 -491.7414
##  [8,]     1     0 15 -555.8224 -494.9015
##  [9,]     1     1  1 -555.7064 -538.7839
## [10,]     1     1  3 -555.6926 -532.0012
```

```
print(cm_ADF_lev$ic_bic)
```

```
##       const trend p       aic       bic
##  [1,]     0     0 0 -550.2068 -543.4378
##  [2,]     1     0 0 -552.6518 -542.4983
##  [3,]     1     0 1 -555.3603 -541.8224
##  [4,]     0     0 1 -551.9226 -541.7691
##  [5,]     0     0 2 -555.0171 -541.4791
##  [6,]     1     0 2 -557.2660 -540.3435
##  [7,]     1     1 1 -555.7064 -538.7839
##  [8,]     1     1 0 -552.0080 -538.4700
##  [9,]     0     0 3 -555.0132 -538.0907
## [10,]     1     1 2 -556.7428 -536.4358
```

```
cm_adq_set <- as.matrix(arrange(as.data.frame(
                    cm_ADF_lev$ic_bic[c(2:3, 6:8, 10),]),
                    const, trend, p))
cm_adq_idx <- match(data.frame(t(cm_adq_set[, 1:3])),
                   data.frame(t(cm_ADF_lev$ic[, 1:3])))

for (i in 1:length(cm_adq_idx))
{
  checkresiduals(cm_ADF_lev$ADF_est[[cm_adq_idx[i]]])
}
```

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

5. Next, use the data to test the *expectations theory* of the term structure of interest rates (ETT). Specifically, investigate whether the spreads in the Capital Market (i5y − i3y) and Money Market (i180d − i90d) are stable (and therefore stationary assuming constant variances and auto-covariances).

```
        Ljung-Box test

data:  Residuals from Regression with ARIMA(0,0,0) errors
Q* = 19.261, df = 10, p-value = 0.03707

Model df: 0.   Total lags used: 10


        Ljung-Box test

data:  Residuals from Regression with ARIMA(1,0,0) errors
Q* = 12.915, df = 9, p-value = 0.1665

Model df: 1.   Total lags used: 10


        Ljung-Box test

data:  Residuals from Regression with ARIMA(2,0,0) errors
Q* = 10.092, df = 8, p-value = 0.2587

Model df: 2.   Total lags used: 10
```

```
        Ljung-Box test

data:  Residuals from Regression with ARIMA(0,0,0) errors
Q* = 19.346, df = 10, p-value = 0.03608

Model df: 0.   Total lags used: 10


        Ljung-Box test

data:  Residuals from Regression with ARIMA(1,0,0) errors
Q* = 12.354, df = 9, p-value = 0.1941

Model df: 1.   Total lags used: 10


        Ljung-Box test

data:  Residuals from Regression with ARIMA(2,0,0) errors
Q* = 10.031, df = 8, p-value = 0.2629

Model df: 2.   Total lags used: 10
```

```
> adf.test(i5y - i3y, nlag = 3)
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 -2.72  0.0100
[2,]   1 -2.99  0.0100
[3,]   2 -2.13  0.0345
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -3.40  0.0128
[2,]   1 -3.88  0.0100
[3,]   2 -2.82  0.0611
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -3.48   0.045
[2,]   1 -4.09   0.010
[3,]   2 -3.04   0.141
```

	const	trend	p	aic	bic
1	1	0	0	-552.6518	-542.4983
2	1	0	1	-555.3603	-541.8224
3	1	0	2	-557.2660	-540.3435
4	1	1	0	-552.0080	-538.4700
5	1	1	1	-555.7064	-538.7839
6	1	1	2	-556.7428	-536.4358

For i5y − i3y, the test rejects a unit root for models with $p = 1$, but fails to reject it for models with $p = 2$. We are unable to conclusively confirm the ETT in the Capital Market.

5. Next, use the data to test the *expectations theory* of the term structure of interest rates (ETT). Specifically, investigate whether the spreads in the Capital Market (i5y – i3y) and Money Market (i180d – i90d) are stable (and therefore stationary assuming constant variances and auto-covariances).

```
        Ljung-Box test

data:  Residuals from Regression with ARIMA(1,0,0) errors
Q* = 7.5858, df = 9, p-value = 0.5764

Model df: 1.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(2,0,0) errors
Q* = 7.7336, df = 8, p-value = 0.4599

Model df: 2.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(3,0,0) errors
Q* = 7.1189, df = 7, p-value = 0.4166

Model df: 3.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(1,0,0) errors
Q* = 7.8281, df = 9, p-value = 0.5516

Model df: 1.   Total lags used: 10
```

```
        Ljung-Box test

data:  Residuals from Regression with ARIMA(2,0,0) errors
Q* = 7.6639, df = 8, p-value = 0.467

Model df: 2.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(3,0,0) errors
Q* = 6.6809, df = 7, p-value = 0.4628

Model df: 3.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(1,0,0) errors
Q* = 7.8597, df = 9, p-value = 0.5483

Model df: 1.   Total lags used: 10

        Ljung-Box test

data:  Residuals from Regression with ARIMA(2,0,0) errors
Q* = 7.6345, df = 8, p-value = 0.47

Model df: 2.   Total lags used: 10
```

	const	trend	p	aic	bic
1	0	0	1	-539.2831	-529.1434
2	0	0	2	-538.5208	-525.0012
3	0	0	3	-538.1661	-521.2667
4	1	0	1	-540.8512	-527.3316
5	1	0	2	-540.5695	-523.6700
6	1	0	3	-539.8240	-519.5446
7	1	1	1	-539.0038	-522.1043
8	1	1	2	-538.6916	-518.4123

```
> adf.test(i180d - i90d)
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 -3.02    0.01
[2,]   1 -3.84    0.01
[3,]   2 -3.92    0.01
[4,]   3 -4.14    0.01
[5,]   4 -4.00    0.01
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -3.39  0.0135
[2,]   1 -4.26  0.0100
[3,]   2 -4.34  0.0100
[4,]   3 -4.64  0.0100
[5,]   4 -4.52  0.0100
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -3.40  0.0551
[2,]   1 -4.26  0.0100
[3,]   2 -4.33  0.0100
[4,]   3 -4.64  0.0100
[5,]   4 -4.52  0.0100
```

The test rejects a unit root for all models in the adequate set. Note that some specifications lead to a failure to reject, but they are not in our adequate set, so we can ignore them! We can confidently conclude that the money market spread is I(0) and the ETT holds for the Money Market.

Tutorial 8: Cointegration - I

At the end of this tutorial you should be able to:

- Automate the task of unit root testing in multiple time-series samples in R;
- Implement the Engle-Granger cointegration test in R;
- Interpret the outcome of an Engle-Granger.
- Use the outcome of the Engle-Granger test to infer possible cointegrating relations.

THE UNIVERSITY OF QUEENSLAND AUSTRALIA | CREATE CHANGE

Thank you

Francisco Tavares Garcia

Academic Tutor | School of Economics

tavaresgarcia.github.io

Reference

Tsay, R. (2010). Analysis of Financial Time Series, 3rd Edition, John Wiley & Sons.