# Problem Set #5
# Co-integration and Error Correction

This problem set deals with co-integration. Exercise #5.1 presents the ideas of co-integration for a simple example, while exercise #5.2 takes you through the steps of an Engle-Granger analysis for Danish aggregate consumption.

## #5.1 Unit Roots and Co-integration

Let $r_t$ denote the effective US Federal funds rate (which is an overnight interest rate), and let $b_t$ denote a 1 year bond rate. Assume that the bond rate $b_t$ follows a unit root process,

$$b_t = b_{t-1} + \epsilon_{bt}, \quad t = 1, 2, ..., T, \tag{5.1}$$

and that the Federal Funds rate depends on the lagged level of the bond rate, as given by,

$$r_t = \beta_2 \cdot b_{t-1} + \epsilon_{rt}, \quad t = 1, 2, ..., T, \tag{5.2}$$

where the initial values $b_0$ and $r_0$ are given and the error terms are independent over time and normally distributed, i.e. $\epsilon_{bt} \stackrel{d}{=} N(0, \sigma_b^2)$ and $\epsilon_{rt} \stackrel{d}{=} N(0, \sigma_r^2)$.

(1) Derive the moving average representation for $b_t$ and $r_t$.
[Hint: Derive the moving average representation for $b_t$ first. Then use the expression you find to derive the moving average representation for $r_t$.]

(2) Use the moving average representations of $r_t$ and $b_t$ to show that the linear combination $\beta' x_t$,

$$\beta' x_t = \begin{pmatrix} 1 & -\beta_2 \end{pmatrix} \begin{pmatrix} r_t \\ b_t \end{pmatrix} = r_t - \beta_2 b_t, \tag{5.3}$$

is a stationary process. Explain how this is related to the concept of co-integration.

(3) Assume that both $b_t$ and $r_t$ are found to be unit root processes based on Dickey-Fuller unit root tests. We are interested in testing whether the interest rate spread,

$$s_t = b_t - r_t, \tag{5.4}$$

is stationary. Table 5.1 contains the output of the regression,

$$\Delta s_t = \delta + c\Delta s_{t-1} + \pi s_{t-1} + \epsilon_t, \tag{5.5}$$

for the interest rate spread over the effective sampe $1988:1-2004:10$.
Use the information in the table to test the hypothesis that $s_t$ has a unit root.
Are $b_t$ and $r_t$ co-integrated with the assumed known co-integration vector $\beta = (1,-1)$?

Table 5.1: Modelling $\Delta s_t$ by OLS for $1988:1-2004:10$

| | Coefficient | | Std.Error | $t-$value |
|---|---|---|---|---|
| Constant | 0.0121 | | 0.0142 | 0.8490 |
| $\Delta s_{t-1}$ | 0.2318 | | 0.0690 | 3.3600 |
| $s_{t-1}$ | −0.1060 | | 0.0287 | −3.7000 |
| $\widehat{\sigma}$ | 0.1966 | $RSS$ | | 7.6919 |
| $R^2$ | 0.0940 | $F(2,199)$ | | 10.3200 |
| No. of observations | 202 | | | |

## #5.2 Engle-Granger Analysis for Danish Consumption

In this exercise we want to construct a single equation co-integration model for the Danish private consumption. We define the vector of variables

$$Z_t = \begin{pmatrix} c_t \\ y_t \\ w_t \\ \mathsf{ARBLOS}_t \end{pmatrix} = \begin{pmatrix} \log(\mathsf{FCP}_t) \\ \log(\mathsf{FYD\_H}_t) \\ \log(\mathsf{FWCP}_t) \\ \mathsf{ARBLOS}_t \end{pmatrix},$$

where FCP is real Danish private consumption, FYD_H is real disposable income, FWCP is real wealth including the value of owner occupied housing, and ARBLOS is the income loss from changes in unemployment. The variables are located in `ConsumptionData.In7`. All the variables are seasonally adjusted and are taken from the data base from the model MONA of the Danish Central Bank.

(1) Load the dataset and apply the suggested transformations.
Make a graph of the four time series.
Test for unit roots in each of the four variables in $Z_t$. Make sure that you use a sensible deterministic specification.
What do your conclusions imply for the estimation of a co-integration relation for the variables in $Z_t$?

(2) Use OLS to estimate the static long-run relation

$$c_t = \beta_0 + \beta_1 y_t + \beta_2 w_t + u_t, \tag{5.6}$$

and interpret the signs and the magnitudes of the coefficients. Do you think it is reasonable to use the output to test statistical hypothesis on the coefficients?

(3) Perform an Engle-Granger residual based test for whether (5.6) is a co-integrating relation. Explain the test and the outcome.

(4) Irrespective of the conclusions in question (3), define the error correction term, $\mathsf{ecm}_t = \hat{u}_t$, corresponding to the relation in (5.6).

Construct a graph of $\mathsf{ecm}_t$ and related it to your finding regarding co-integration.

(5) Construct single equation error correction models for $\Delta c_t$, $\Delta y_t$, and $\Delta w_t$, where you include also $\mathsf{ARBLOS}_t$ as an explanatory variable, e.g.

$$\begin{aligned}\Delta c_t &= \alpha_0 + \alpha_1 \Delta c_{t-1} + \alpha_2 \Delta y_t + \alpha_3 \Delta y_{t-1} + \alpha_4 \Delta w_t + \alpha_5 \Delta w_{t-1} \\ &\quad + \alpha_6 \mathsf{ARBLOS}_t + \alpha_7 \mathsf{ARBLOS}_{t-1} + \alpha_8 \mathsf{ecm}_{t-1} + \epsilon_t. \end{aligned} \tag{5.7}$$

Do any of the variables seem to error correct? Is that in line with your expectations based on economic theory?

(6) Concentrate on the model for $\Delta c_t$. Delete insignificant variables and reestimate the model (5.7). Perform the usual misspecification tests. Does the model appear well specified?

(7) If you find outliers in the model, construct dummy variables and augment the model with dummy variables to account for the outliers.

Discuss how the results change.

Look at the time series of estimated residuals, $\hat{\epsilon}_t$, and explain how the dummy variables change the residuals.

Now we introduce the concept of recursive estimation, which is very useful in analyzing the structural stability of an estimated model. Recursive estimation is done by estimating a model like (5.7) for increasing samples $t = 1973:2 - T_0$, where $T_0$ takes the values

$$1973:2+N, \quad 1973:2+N+1, \quad 1973:2+N+2, \quad ..., \quad 2018:3.$$

That is, we first estimate the model with only the first $N$ observations; and then we successively add a new observation and reestimate the model. For each sample we do an OLS estimation and obtain all the usual statistics. Afterwards we can consider the sample paths of the different statistics calculated for each sample. For example we can consider the estimated coefficients for the expanding samples

$$\hat{\alpha}_i(T_0), \quad \text{for} \quad T_0 = 1973:2+N, ..., 2018:3.$$

The model (5.7) is estimated under the assumption of constant coefficients, which implies that a graph for $\hat{\alpha}_i(T_0)$, $i = 0, ..., 8$, should not fluctuate too much.

(8) Do a recursive estimation of your error correction model and look at the recursive parameter estimates. Does the model look stable?

[Hint: To do this, select `Recursive Graphics` in the `Test` menu.]