Lecture 5. Cointegration, Error-Correction Models

Narek Ohanyan

American University of Armenia

February 25, 2024

Outline

Cointegration

Two $\mathrm{I}(1)$ time series y_t and x_t are said to be **cointegrated** if some linear combination of them is $\mathrm{I}(0)$.

That is

$$y_t \sim \mathrm{I}(1) \qquad x_t \sim \mathrm{I}(1) \qquad \text{and} \qquad z_t = c_0 + c_1 y_t + c_2 x_t \qquad z_t \sim \mathrm{I}(0).$$

for some constants c_0, c_1, c_2.

Cointegration implies that y_t and x_t share **similar** stochastic trends and *never diverge* too far from each other.

In other words, cointegration implies that the two series are *not independent*, but rather **move together** in the long run.

Testing for cointegration

The most common test for cointegration is the **Engle-Granger** two-step procedure.

1. Estimate the cointegrating regression

$$y_t = \alpha + \beta x_t + e_t.$$

   and obtain the residuals $\widehat{e}_t$.

2. Test the residuals $\widehat{e}_t$ for stationarity using the ADF test.

If the residuals are **stationary**, then the two series are **cointegrated**.

The test is based on the following regression

$$\Delta\widehat{e}_t = \gamma\widehat{e}_{t-1} + e_t.$$

Critical values for cointegration test

Regression model	1%	5%	10%
(1) $y_t = \beta x_t + e_t$	-3.39	-2.76	-2.45
(2) $y_t = \beta_1 + \beta_2 x_t + e_t$	-3.96	-3.37	-3.07
(3) $y_t = \beta_1 + \delta t + \beta_2 x_t + e_t$	-3.98	-3.42	-3.13

Note: These critical values are taken from J. Hamilton (1994), *Time Series Analysis*, Princeton University Press, p. 766.

Outline

Cointegration

Error-Correction Models

Regression with No Cointegration

ARDL model with cointegrated variables

Consider an ARDL(1,1) model of the form

$$y_t = \delta + \phi_1 y_{t-1} + \delta_0 x_t + \delta_1 x_{t-1} + e_t.$$

where y_t and x_t are **cointegrated**.

If y_t and x_t are cointegrated, then there is a **long-run** relationship between them.

The long-run relationship can be obtained by setting $y_t = y_{t-1} = y$, $x_t = x_{t-1} = x$, and $e_t = 0$.

Then the implied long-run relationship is

$$y = \beta_0 + \beta_1 x$$

where $\beta_0 = \delta/(1-\phi_1)$ and $\beta_1 = (\delta_0 + \delta_1)/(1-\phi_1)$.

Error-correction model (ECM)

Recall that y_t and x_t are related by an ARDL(1,1) model

$$y_t = \delta + \phi_1 y_{t-1} + \delta_0 x_t + \delta_1 x_{t-1} + e_t.$$

Subtracting y_{t-1} from both sides of the equation and adding $\pm\delta_0 x_{t-1}$

$$y_t - y_{t-1} = \delta + (\phi_1 - 1) y_{t-1} + \delta_0 \left(x_t - x_{t-1}\right) + \left(\delta_0 + \delta_1\right) x_{t-1} + e_t.$$

Then we can write the equation as

$$\Delta y_t = (\phi_1 - 1)\left(\frac{\delta}{1-\phi_1} + y_{t-1} + \frac{\delta_0 + \delta_1}{1-\phi_1} x_{t-1}\right) + \delta_0 \Delta x_t + e_t.$$

or

$$\Delta y_t = \alpha \left(y_{t-1} - \beta_0 - \beta_1 x_{t-1}\right) + \delta_0 \Delta x_t + e_t.$$

Error-correction model (ECM)

The **Error-correction model (ECM)** is given by

$$\begin{aligned}\Delta y_t &= \alpha\left(y_{t-1} - \beta_0 - \beta_1 x_{t-1}\right) + \delta_0 \Delta x_t + e_t \\ &= \alpha e_{t-1} + \delta_0 \Delta x_t + e_t.\end{aligned}$$

where $\alpha < 0$ is the *speed of adjustment* and β_0 and β_1 are the *long-run coefficients*.

The ECM has three important features:

- It allows for an underlying or *fundamental (long-run)* link between variables
- It allows for *short-run adjustments* between variables towards the long-run equilibrium

In practice, the ECM is estimated in two steps:

1. Estimate the long-run relationship
2. Estimate the short-run dynamics

Outline

Regression with No-Cointegration

If y_t and x_t are **not cointegrated**, then regression of y_t on x_t is susceptible to *spurious regression*.

In this case, the relationship between y_t and x_t may be estimated by transforming the variables to stationary form.

Variables can be transformed to stationary form by

- taking *first differences* if the variables are *difference stationary* or $\mathrm{I}(1)$

$$\Delta y_t = y_t - y_{t-1} \qquad \Delta x_t = x_t - x_{t-1}.$$

- *detrending* the variables if the variables are *trend stationary*

$$\widetilde{y}_t = y_t - \alpha_1 - \delta_1 t \qquad \widetilde{x}_t = x_t - \alpha_2 - \delta_2 t.$$

Regression with first differences or detrended variables

In a stationary form, the relationship between y_t and x_t can be estimated as an ARDL model

- with *first differences* if the variables are *difference stationary*

$$\Delta y_t = \alpha + \phi_1 \Delta y_{t-1} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + e_t.$$

- with *detrended* variables if the variables are *trend stationary*

$$\widetilde{y}_t = \alpha + \phi_1 \widetilde{y}_{t-1} + \beta_0 \widetilde{x}_t + \beta_1 \widetilde{x}_{t-1} + e_t.$$

Alternatively, if both y_t and x_t are *trend stationary*, a *time trend* may be included in the regression

$$y_t = \alpha + \delta t + \phi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + e_t.$$

Regression with non-stationary variables