Lecture: Introduction to Cointegration

Applied Econometrics

Jozef Barunik

IES, FSV, UK

Summer Semester 2010/2011

Readings

1. The Royal Swedish Academy of Sciences (2003): Time Series Econometrics: Cointegration and Autoregressive Conditional Heteroscedasticity, downloadable from:
   http://www-stat.wharton.upenn.edu/~steele/HoldingPen/NobelPrizeInfo.pdf
2. Granger,C.W.J. (2003): Time Series, Cointegration and Applications, Nobel lecture, December 8, 2003
3. Harris Using Cointegration Analysis in Econometric Modelling, 1995
   - (Useful applied econometrics textbook focused solely on cointegration)
4. Almost all textbooks cover the introduction to cointegration
   - Engle-Granger procedure (single equation procedure),
   - Johansen multivariate framework (covered in the following lecture)

Outline of the today's talk

- What is cointegration?
- Deriving Error-Correction Model (ECM)
- Engle-Granger procedure

Outline of the today's talk

- What is cointegration?
- Deriving Error-Correction Model (ECM)
- Engle-Granger procedure

Outline of the today's talk

- What is cointegration?
- Deriving Error-Correction Model (ECM)
- Engle-Granger procedure

Robert F. Engle and Clive W.J. Granger

Robert F. Engle shared the Nobel prize (2003) *"for methods of analyzing economic time series with time-varying volatility (ARCH)* with Clive W. J. Granger who recieved the prize *"for methods of analyzing economic time series with common trends (cointegration).*

Figure: (a) Robert F. Engle (b) Clive W.J. Granger

Introduction

- We learnt that regressing two non-stationary variables (say Y_t on X_t) results in **spurious regression**
- However, if Y_t and X_t are cointegrated, spurious regression no longer arise
- Success of large structural macro models in the 1960s due to trend vs. its failure in 1970s

Introduction

- We learnt that regressing two non-stationary variables (say Y_t on X_t) results in **spurious regression**
- However, if Y_t and X_t are cointegrated, spurious regression no longer arise
- Success of large structural macro models in the 1960s due to trend vs. its failure in 1970s

Introduction

- We learnt that regressing two non-stationary variables (say Y_t on X_t) results in **spurious regression**
- However, if Y_t and X_t are cointegrated, spurious regression no longer arise
- Success of large structural macro models in the 1960s due to trend vs. its failure in 1970s

Introduction cont.

- Assume two time series Y_t , and X_t , are integrated of order $d(Y_t, X_t \sim I(d))$
- If there exists β such that $Y_t - \beta * X_t = u_t$, where u_t is integrated of order less than d (say $d - b$), we say that Y_t and X_t are cointegrated of order $d - b$, Y_t , $Xt \sim CI(d, b)$
- For example, money supply and price level are typically integrated of order one $(Y_t, X_t \sim I(1))$, but their difference should be stationary $(I(0))$ in the long run, as money supply and price level cannot according to economic theory diverge in the long run.

Introduction cont.

- Assume two time series Y_t , and X_t , are integrated of order $d(Y_t, X_t \sim I(d))$
- If there exists β such that $Y_t - \beta * X_t = u_t$, where u_t is integrated of order less than d (say $d - b$), we say that Y_t and X_t are cointegrated of order $d - b$, Y_t , $Xt \sim CI(d, b)$
- For example, money supply and price level are typically integrated of order one $(Y_t, X_t \sim I(1))$, but their difference should be stationary $(I(0))$ in the long run, as money supply and price level cannot according to economic theory diverge in the long run.

Introduction cont.

- Assume two time series Y_t , and X_t , are integrated of order $d(Y_t, X_t \sim I(d))$
- If there exists β such that $Y_t - \beta * X_t = u_t$, where u_t is integrated of order less than d (say $d - b$), we say that Y_t and X_t are cointegrated of order $d - b$, Y_t , $Xt \sim CI(d, b)$
- For example, money supply and price level are typically integrated of order one $(Y_t, X_t \sim I(1))$, but their difference should be stationary $(I(0))$ in the long run, as money supply and price level cannot according to economic theory diverge in the long run.

Introduction cont.

The prices of goods expressed in common currency should be identical, so $S_t * P_{t,foreign} = P_t$

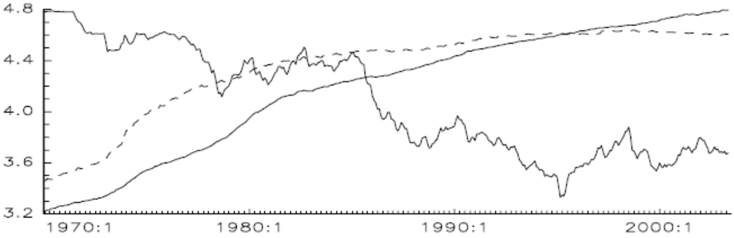

Figure 1.1: Logarithm (rescaled) of the Japanese yen/US dollar exchange rate (decreasing solid line), logarithm of seasonally adjusted US consumer price index (increasing solid line) and logarithm of seasonally adjusted Japanese consumer price index (increasing dashed line), 1970:1 − 2003:5, monthly observations

Introduction cont.

If $S_t, P_{t,foreign}$ and P_t are $I(1)$ and cointegrated, its linear combination is $I(0)$.

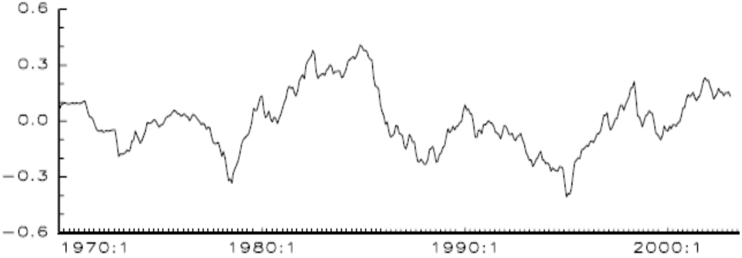

Figure: Regression residuals

Introduction cont.

- If Y_t and X_t are integrated of order one and are cointegrated, you do not have to difference the data and may (simply by OLS) estimate $Y_t = \rho + \beta * X_t + u_t$
- β is superconsistent in this case, converge to its true counterpart at a faster rate than the usual OLS estimator with $I(0)$ variables ΔY_t and ΔX_t, however standard errors not consistent, not worth reporting
- Note that, if you want to difference Y_t and X_t, you will not have unit root in variables Y_t and X_t but unit root will arise in the error term $u_t = e_t - e_{t-1}$ (overdifferenced data)

Introduction cont.

- If Y_t and X_t are integrated of order one and are cointegrated, you do not have to difference the data and may (simply by OLS) estimate $Y_t = \rho + \beta * X_t + u_t$
- β is superconsistent in this case, converge to its true counterpart at a faster rate than the usual OLS estimator with $I(0)$ variables ΔY_t and ΔX_t, however standard errors not consistent, not worth reporting
- Note that, if you want to difference Y_t and X_t, you will not have unit root in variables Y_t and X_t but unit root will arise in the error term $u_t = e_t - e_{t-1}$ (overdifferenced data)

Introduction cont.

- If Y_t and X_t are integrated of order one and are cointegrated, you do not have to difference the data and may (simply by OLS) estimate $Y_t = \rho + \beta * X_t + u_t$
- β is superconsistent in this case, converge to its true counterpart at a faster rate than the usual OLS estimator with $I(0)$ variables ΔY_t and ΔX_t, however standard errors not consistent, not worth reporting
- Note that, if you want to difference Y_t and X_t, you will not have unit root in variables Y_t and X_t but unit root will arise in the error term $u_t = e_t - e_{t-1}$ (overdifferenced data)

Introduction cont.

Note

If say $X \sim I(0)$ and $Y \sim I(1)$, surely no cointegration (no long run relationship), X is more or less constant over time, while Y increases over time

Cointegration

- If you difference I(1) data, you lose long run information and estimate only short run model
  - This is, with differenced data you know what is the effect of the change of x on change of y, not the level effect
- Alternative is to use error-correction model (ECM), great advantage is that you may model both short run and long run relationship jointly (if variables cointegrated)
- Granger representation theorem: for any set of I(1) variables, error correction and cointegration are the equivalent representations (*'same'*)

Cointegration

- If you difference I(1) data, you lose long run information and estimate only short run model
  - This is, with differenced data you know what is the effect of the change of x on change of y, not the level effect
- Alternative is to use error-correction model (ECM), great advantage is that you may model both short run and long run relationship jointly (if variables cointegrated)
- Granger representation theorem: for any set of I(1) variables, error correction and cointegration are the equivalent representations (*'same'*)

Cointegration

- If you difference I(1) data, you lose long run information and estimate only short run model
  - This is, with differenced data you know what is the effect of the change of x on change of y, not the level effect
- Alternative is to use error-correction model (ECM), great advantage is that you may model both short run and long run relationship jointly (if variables cointegrated)
- Granger representation theorem: for any set of I(1) variables, error correction and cointegration are the equivalent representations (*'same'*)

Deriving ECM

- Assume the model $C_t = \alpha_0 + \alpha_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t$ and assume C_t and Y_t both $\sim I(1)$
- Subtract C_{t-1} from both sides of equation and get

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t \quad (1)$$

- Now add: $-\beta_0 Y_{t-1} + \beta_0 Y_{t-1}$ and get:

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 \Delta Y_t + \theta_1 Y_{t-1} + u_t \quad (2)$$

- Now, LHS stationary, ΔY_t stationary, if C_t and Y_t cointegrated (thus together I(0)), then u_t must be I(0) as well
- May generalize to more variables and time trend as well

Deriving ECM

- Assume the model $C_t = \alpha_0 + \alpha_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t$ and assume C_t and Y_t both $\sim I(1)$
- Subtract C_{t-1} from both sides of equation and get

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t \quad (1)$$

- Now add: $-\beta_0 Y_{t-1} + \beta_0 Y_{t-1}$ and get:

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 \Delta Y_t + \theta_1 Y_{t-1} + u_t \quad (2)$$

- Now, LHS stationary, ΔY_t stationary, if C_t and Y_t cointegrated (thus together I(0)), then u_t must be I(0) as well
- May generalize to more variables and time trend as well

Deriving ECM

- Assume the model $C_t = \alpha_0 + \alpha_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t$ and assume C_t and Y_t both $\sim I(1)$
- Subtract C_{t-1} from both sides of equation and get

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t \quad (1)$$

- Now add: $-\beta_0 Y_{t-1} + \beta_0 Y_{t-1}$ and get:

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 \Delta Y_t + \theta_1 Y_{t-1} + u_t \quad (2)$$

- Now, LHS stationary, ΔY_t stationary, if C_t and Y_t cointegrated (thus together I(0)), then u_t must be I(0) as well
- May generalize to more variables and time trend as well

Deriving ECM

- Assume the model $C_t = \alpha_0 + \alpha_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t$ and assume C_t and Y_t both $\sim I(1)$
- Subtract C_{t-1} from both sides of equation and get

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t \tag{1}$$

- Now add: $-\beta_0 Y_{t-1} + \beta_0 Y_{t-1}$ and get:

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 \Delta Y_t + \theta_1 Y_{t-1} + u_t \tag{2}$$

- Now, LHS stationary, ΔY_t stationary, if C_t and Y_t cointegrated (thus together I(0)), then u_t must be I(0) as well
- May generalize to more variables and time trend as well

Deriving ECM

- Assume the model $C_t = \alpha_0 + \alpha_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t$ and assume C_t and Y_t both $\sim I(1)$
- Subtract C_{t-1} from both sides of equation and get

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + u_t \quad (1)$$

- Now add: $-\beta_0 Y_{t-1} + \beta_0 Y_{t-1}$ and get:

$$\Delta C_t = \alpha_0 + \rho_1 C_{t-1} + \beta_0 \Delta Y_t + \theta_1 Y_{t-1} + u_t \quad (2)$$

- Now, LHS stationary, ΔY_t stationary, if C_t and Y_t cointegrated (thus together I(0)), then u_t must be I(0) as well
- May generalize to more variables and time trend as well

Testing for Cointegration

Test the residuals for a unit root (ADF test)

- No constant required (if constant already included in original regression)

$$\Delta\hat{u}_t = \beta\hat{u}_{t-1} + \delta_1\Delta\hat{u}_{t-1} + \cdots + \delta_n\Delta\hat{u}_{t-n} + \nu_t \quad (3)$$

- Test $H_0 : \beta = 0$
- $\beta = 0 \Rightarrow$ Unit root $\Rightarrow$ non-stationary $\Rightarrow$ Y_t and X_t not cointegrated
- $\beta \neq 0 \Rightarrow$ No Unit root $\Rightarrow$ stationary $\Rightarrow$ Y_t and X_t cointegrated
- Note DF critical values for CI are not the same as for I, critical values from Engle and Yoo (1987)

Testing for Cointegration

Test the residuals for a unit root (ADF test)

- No constant required (if constant already included in original regression)

$$\Delta \hat{u}_t = \beta \hat{u}_{t-1} + \delta_1 \Delta \hat{u}_{t-1} + \cdots + \delta_n \Delta \hat{u}_{t-n} + \nu_t \quad (3)$$

- Test $H_0 : \beta = 0$
- $\beta = 0 \Rightarrow$ Unit root $\Rightarrow$ non-stationary $\Rightarrow$ Y_t and X_t not cointegrated
- $\beta \neq 0 \Rightarrow$ No Unit root $\Rightarrow$ stationary $\Rightarrow$ Y_t and X_t cointegrated
- Note DF critical values for CI are not the same as for I, critical values from Engle and Yoo (1987)

Testing for Cointegration

Test the residuals for a unit root (ADF test)

- No constant required (if constant already included in original regression)

$$\Delta \hat{u}_t = \beta \hat{u}_{t-1} + \delta_1 \Delta \hat{u}_{t-1} + \cdots + \delta_n \Delta \hat{u}_{t-n} + \nu_t \quad (3)$$

- Test $H_0 : \beta = 0$
- $\beta = 0 \Rightarrow$ Unit root $\Rightarrow$ non-stationary $\Rightarrow$ Y_t and X_t not cointegrated
- $\beta \neq 0 \Rightarrow$ No Unit root $\Rightarrow$ stationary $\Rightarrow$ Y_t and X_t cointegrated
- Note DF critical values for CI are not the same as for I, critical values from Engle and Yoo (1987)

Testing for Cointegration

Test the residuals for a unit root (ADF test)

- No constant required (if constant already included in original regression)

$$\Delta \hat{u}_t = \beta \hat{u}_{t-1} + \delta_1 \Delta \hat{u}_{t-1} + \cdots + \delta_n \Delta \hat{u}_{t-n} + \nu_t \qquad (3)$$

- Test $H_0 : \beta = 0$
- $\beta = 0 \Rightarrow$ Unit root $\Rightarrow$ non-stationary $\Rightarrow$ Y_t and X_t not cointegrated
- $\beta \neq 0 \Rightarrow$ No Unit root $\Rightarrow$ stationary $\Rightarrow$ Y_t and X_t cointegrated
- Note DF critical values for CI are not the same as for I, critical values from Engle and Yoo (1987)

Testing for Cointegration

Test the residuals for a unit root (ADF test)

- No constant required (if constant already included in original regression)

$$\Delta \hat{u}_t = \beta \hat{u}_{t-1} + \delta_1 \Delta \hat{u}_{t-1} + \cdots + \delta_n \Delta \hat{u}_{t-n} + \nu_t \qquad (3)$$

- Test $H_0 : \beta = 0$
- $\beta = 0 \Rightarrow$ Unit root $\Rightarrow$ non-stationary $\Rightarrow$ Y_t and X_t not cointegrated
- $\beta \neq 0 \Rightarrow$ No Unit root $\Rightarrow$ stationary $\Rightarrow$ Y_t and X_t cointegrated
- Note DF critical values for CI are not the same as for I, critical values from Engle and Yoo (1987)

Testing for Cointegration (cont.)

- Alternatively, use Durbin-Watson (DW)
- DW roughly equal to $2(1-\rho)$, where ρ is measure of autocorrelation
- Null hypothesis: No CI, $\rho = 1$, DW=0
- Alternative: CI, $-1 < \rho < 1$, DW>0
- Developed by Sargan and Bhargava, 1983, but applicable only if the residual follows 1-st order autoregression (not so widely used)

Testing for Cointegration (cont.)

- Alternatively, use Durbin-Watson (DW)
- DW roughly equal to $2(1-\rho)$, where ρ is measure of autocorrelation
- Null hypothesis: No CI, $\rho = 1$, DW=0
- Alternative: CI, $-1 < \rho < 1$, DW>0
- Developed by Sargan and Bhargava, 1983, but applicable only if the residual follows 1-st order autoregression (not so widely used)

Testing for Cointegration (cont.)

- Alternatively, use Durbin-Watson (DW)
- DW roughly equal to $2(1-\rho)$, where ρ is measure of autocorrelation
- Null hypothesis: No CI, $\rho = 1$, DW=0
- Alternative: CI, $-1 < \rho < 1$, DW>0
- Developed by Sargan and Bhargava, 1983, but applicable only if the residual follows 1-st order autoregression (not so widely used)

Testing for Cointegration (cont.)

- Alternatively, use Durbin-Watson (DW)
- DW roughly equal to $2(1-\rho)$, where ρ is measure of autocorrelation
- Null hypothesis: No CI, $\rho = 1$, DW=0
- Alternative: CI, $-1 < \rho < 1$, DW>0
- Developed by Sargan and Bhargava, 1983, but applicable only if the residual follows 1-st order autoregression (not so widely used)

Testing for Cointegration (cont.)

- Alternatively, use Durbin-Watson (DW)
- DW roughly equal to $2(1-\rho)$, where ρ is measure of autocorrelation
- Null hypothesis: No CI, $\rho = 1$, DW=0
- Alternative: CI, $-1 < \rho < 1$, DW>0
- Developed by Sargan and Bhargava, 1983, but applicable only if the residual follows 1-st order autoregression (not so widely used)

ECM Estimation

If you find evidence of cointegration, then specify the corresponding ECM
Estimate the ECM using the lagged residuals (u_{t-1})

- as the EC Mechanism
  $\Delta Y_t = \beta_0 + \beta_1 \Delta X_t - \beta_2 (Y_{t-1} - C - \beta X_{t-1})$

EC Mechanism

$$(Y_{t-1} - C - \beta X_{t-1}) = u_{t-1} \tag{4}$$

- In the cointegrating regression
  $Y_t = C + \beta X_t + u_t$

$$u_t = Y_t - C - \beta X_t \Rightarrow u_{t-1} = Y_{t-1} - C - \beta X_{t-1} \tag{5}$$

NOTE

$(4) \equiv (5) \Rightarrow u_{t-1} \equiv$ EC Mechanism

Engle-Granger procedure

1. Test the order of integration for all variables by unit root test such as ADF or PP test
2. Estimate (by OLS) $C_t = \alpha_0 + \beta_0 Y_t + u_t$,
3. Test for cointegration
4. Estimate error-correction model $\Delta C_t = \alpha_0 + \beta \Delta Y_t + \rho u_{t-1} + e_t$, you may include lags of ΔC_t and ΔY_t in the RHS, if needed

NOTE

that u_{t-1} is from the equation in the step 2

5. Evaluate the model adequacy (note that the estimated parameter ρ should be negative and can be interpreted as the speed of adjustment)

Engle-Granger procedure

1. Test the order of integration for all variables by unit root test such as ADF or PP test
2. Estimate (by OLS) $C_t = \alpha_0 + \beta_0 Y_t + u_t$,
3. Test for cointegration
4. Estimate error-correction model $\Delta C_t = \alpha_0 + \beta \Delta Y_t + \rho u_{t-1} + e_t$, you may include lags of ΔC_t and ΔY_t in the RHS, if needed

NOTE

that u_{t-1} is from the equation in the step 2

5. Evaluate the model adequacy (note that the estimated parameter ρ should be negative and can be interpreted as the speed of adjustment)

Engle-Granger procedure

1. Test the order of integration for all variables by unit root test such as ADF or PP test
2. Estimate (by OLS) $C_t = \alpha_0 + \beta_0 Y_t + u_t$,
3. Test for cointegration
4. Estimate error-correction model $\Delta C_t = \alpha_0 + \beta \Delta Y_t + \rho u_{t-1} + e_t$, you may include lags of ΔC_t and ΔY_t in the RHS, if needed

NOTE

that u_{t-1} is from the equation in the step 2

5. Evaluate the model adequacy (note that the estimated parameter ρ should be negative and can be interpreted as the speed of adjustment)

Engle-Granger procedure

1. Test the order of integration for all variables by unit root test such as ADF or PP test
2. Estimate (by OLS) $C_t = \alpha_0 + \beta_0 Y_t + u_t$,
3. Test for cointegration
4. Estimate error-correction model $\Delta C_t = \alpha_0 + \beta \Delta Y_t + \rho u_{t-1} + e_t$, you may include lags of ΔC_t and ΔY_t in the RHS, if needed

NOTE

that u_{t-1} is from the equation in the step 2

5. Evaluate the model adequacy (note that the estimated parameter ρ should be negative and can be interpreted as the speed of adjustment)

Engle-Granger procedure

1. Test the order of integration for all variables by unit root test such as ADF or PP test
2. Estimate (by OLS) $C_t = \alpha_0 + \beta_0 Y_t + u_t$,
3. Test for cointegration
4. Estimate error-correction model $\Delta C_t = \alpha_0 + \beta \Delta Y_t + \rho u_{t-1} + e_t$, you may include lags of ΔC_t and ΔY_t in the RHS, if needed

NOTE

that u_{t-1} is from the equation in the step 2

5. Evaluate the model adequacy (note that the estimated parameter ρ should be negative and can be interpreted as the speed of adjustment)

Drawback of Engle-Granger approach

- Single equation model
- There can be more than one cointegrating relationships (if there are more than 2 variables)
  - For example, 2 cointegration relationships likely for demand and supply of credit
- The drawback tackled by Johansen procedure

Drawback of Engle-Granger approach

- Single equation model
- There can be more than one cointegrating relationships (if there are more than 2 variables)
  - For example, 2 cointegration relationships likely for demand and supply of credit
- The drawback tackled by Johansen procedure

Drawback of Engle-Granger approach

- Single equation model
- There can be more than one cointegrating relationships (if there are more than 2 variables)
  - For example, 2 cointegration relationships likely for demand and supply of credit
- The drawback tackled by Johansen procedure

Questions

Thank you for your Attention !