Econometrics V Lecture10
Introduction to Cointegration

Readings

- 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
- Granger, C. W.J. (2003): Time Series, Cointegration and Applications, Nobel lecture, December 8, 2003, downloadable from http://ideas.repec.org/p/cdl/ucsdec/2004-02.html

Harris – Using Cointegration Analysis in Econometric Modelling, 1995
  Useful applied econometrics textbook focused solely on cointegration

Almost all textbooks cover the introduction to cointegration

- Engle-Granger procedure (single equation procedure),
- Johansen multivariate framework

Sequence

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

Concept of **cointegration**

- evidence of long-run or equilibrium relationships

With cointegration the residuals from a regression are stationary.

Tested informally and formally for cointegration

Formal Tests include

(1) Cointegrating Regression Durbin Watson (CRDW) test

(2) Cointegrating Regression Dickey Fuller (CRDF) test

Cointegration in Single Equations: Lecture 14

Summary of Lecture

(1) Introduce Granger Representation Theorem.
- relates cointegration to Error Correction Models

(2) Suggest different ways of estimating long run coefficients and short run models

(3) Multivariate regressions and testing for cointegration.

Cointegration: The usefulness of ECMs

Error correction mechanisms are useful for representing the short run relationships between variables.

The level relationship does not hold at all times.
Shocks can move the relationship off track.

Another way of saying **we are not always at equilibrium**.

Nevertheless there is a tendency to move towards equilibrium.

The error correction model allows us to return to zero
i.e. corrects for deviations from equilibrium.

It relates deviations from equilibrium to changes in the dependent variable i.e. the means of correcting for errors.

The estimation of two variable ECMs

However, are we certain an ECM relationship exists for variables? Does cointegration help?

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

Granger Representation Theorem
Provided two time series are cointegrated, the short-term disequilibrium relationship between them can always be expressed in the error correction form.

Cointegration and ECMs

Granger Representation Theorem suggests that if we have cointegration then an ECM exists

$$y_t = \beta_0 + \beta_1 x_t + u_t$$
$$\Delta y_t = \text{lagged } (\Delta y_t, \Delta x_t) - \lambda u_{t-1} + \varepsilon_t$$

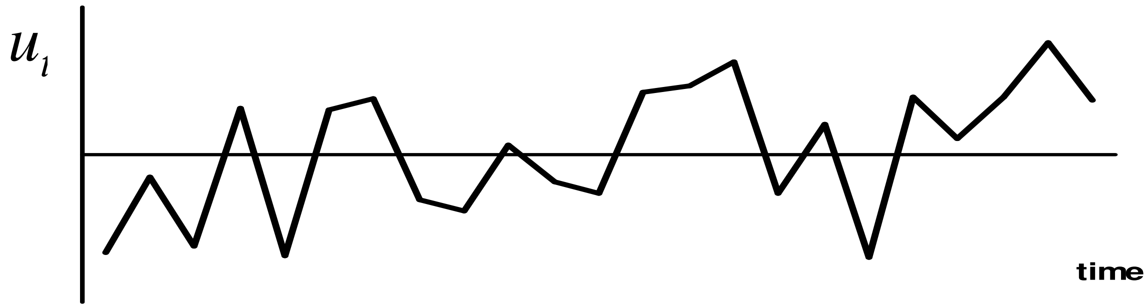
u_{t-1} is the disequilibrium error

λ is the short-run adjustment parameter

This is an important result since it is justification for using ECM.

Cointegration and ECMs

If y_t and x_t are cointegrated then the disequilibrium errors u_t will be stationary.



This means there is a force pulling the residual errors towards zero.

Previous departures from equilibrium are being corrected.

This is exactly what is implied by the error correction model.

Cointegration and ECMs

$$\Delta y_t = \text{lagged}\ (\Delta y_t, \Delta x_t) - \lambda u_{t-1} + \varepsilon_t$$

Notice that all first differenced variables are I(0).

Disequilibrium errors (u_{t-1}) also need to be I(0).

This is the case when y_t and x_t are cointegrated.

Exact lags are not specified by the Granger Representation Theorem.
Specification determined by general to specific approach.

Since Δx_{t-1} is an I(0) variable so is Δx_t
Hence it is possible to incorporate unlagged values of Δx_t
(but may then need to use Instrumental Variables regression).

Estimating ECMs using Cointegration

How do we obtain an error correction model?

Engle-Granger Two-Step approach

(1) Estimate long run relationship between y_t and x_t

(2) Incorporate residuals in a short run model

Estimating ECMs using Cointegration

Engle-Granger Two-Step approach

(1) Estimate long run relationship between y_t and x_t

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

- When there is cointegration we can be confident that β_0 and β_1 will not be biased (in large samples).

As Stock suggested β_0 and β_1 are consistent.

Also superconsistent. We can ignore dynamic terms.

Now use the residuals from the 'cointegrating regression' to test for cointegration (i.e. the existence of a long run equilibrium relationship)

Use the residuals of the estimated long run relationship, test (using DF/ADF statistics) whether or not u is STATIONARY

$$\text{DF}: \Delta\hat{u}_t = \beta\hat{u}_{t-1} + \varepsilon_t$$

$$\text{ADF}: \Delta\hat{u}_t = \beta\hat{u}_{t-1} + \sum_{i=1}^{K} \Delta\hat{u}_{t-i} + \varepsilon_t$$

Note: must use **special tabulated critical values** for CRDF/CRADF tests. If the residuals are stationary, then we can conclude that the series are COINTEGRATED

Estimating ECMs using Cointegration

Engle-Granger Two-Step approach

(2) Incorporate residuals in a short run model

We take the residuals from the estimated static equation u_{t-1} and incorporate them into the short run model.

$$\Delta y_t = \text{lagged } (\Delta y_t, \Delta x_t) - \lambda u_{t-1} + \varepsilon_t$$

We consequently estimate this regression.

We can do so by OLS since all the variables are stationary.

We should obtain the estimated coefficient λ

Engle-Granger Two Step

Problems with the Engle-Granger Two Step

These are concentrated on the first step.
- estimating the static OLS model.

We suggested that OLS estimates of cointegrating regressions will be unbiased in large samples (consistent).

However there may be bias in small samples (the samples we use).

If there is bias in the first step, this will spillover on to the second step.

Typically residuals are only used to test cointegration.

Cointegration and ECMs

One suggestion is that long run parameters should be estimated using methods unbiased in small samples, the implied residuals derived and then the short run model estimated.

Engle Granger Approach becomes

(1) Use AutoRegressive Distributed Lag (ARDL) method to estimate parameters

i.e. within a dynamic model

(2) Derive the residuals errors from the long run model

$$u_t = y_t - \beta_0 - \beta_1 x_t$$

(3) Incorporate residuals in the error correction model

$$\Delta y_t = \text{lagged}\ (\Delta y_t,\ \Delta x_t) - \lambda\, u_{t-1} + \varepsilon_t$$

Cointegration and ECMs

Alternative suggestion is that short run and long run parameters should be estimated in a single step to avoid bias estimates (in small samples).

Banerjee, Dolado, Hendry and Smith (1986) method

$$\Delta y_t = \text{lagged}\ (\Delta y_t,\ \Delta x_t) - \lambda u_{t-1} + \varepsilon_t$$

$$\Delta y_t = \lambda\beta_0 + \text{lagged}\ (\Delta y_t,\ \Delta x_t) - \lambda\, y_{t-1} + \lambda\, \beta_1 x_{t-1} + \varepsilon_t$$

where $u_t = y_t - \beta_0 - \beta_1 x_t$

Simulation studies of the properties of this estimator, suggest that in small samples Banerjee et al. approach performs better than Engle-Granger method.

Cointegration and ECMs

Banerjee, et al. (1986) approach

$$\Delta y_t = \lambda\beta_0 + \text{lagged}\,(\Delta y_t, \Delta x_t) - \lambda y_{t-1} + \lambda \beta_1 x_{t-1} + \varepsilon_t$$

$$\textit{where} \quad u_t = y_t - \beta_0 - \beta_1 x_t$$

Can check cointegration by testing the residuals ε_t for stationarity

Although there are two I(1) variables in this equation a linear combination should cointegrate to produce a stationary relationship.

Consequently all variables (or a combination of variables) will be I(0) and inference can proceed as normal.

Multivariate Cointegration Tests

Johansen Approach

We have concentrated on the bivariate case y_t and x_t.

There can only be one cointegrating relationship between these variables.

Is this the case when there are three variables?

It may be the case that there is more than one relationship.

Where we have variables y_t , x_t and z_t.

Johansen approach not only examines if y_t , x_t and z_t cointegrated.

But also if y_t cointegrates with x_t ***on its own*** and y_t cointegrates with z_t ***on its own***.

Multivariate Cointegration Tests

Single Equation Approach

$$\Delta y_t = \text{lagged}\ (\Delta y_t, \Delta x_t) - \lambda u_{t-1} + \varepsilon_t$$

Soren Johansen Approach

Can test for the number of cointegrating relationships.

Assuming y_t cointegrates with x_t $= LR1$

y_t cointegrates with z_t $= LR2$

Short run model becomes

$$\Delta y_t = \text{lagged}\ (\Delta y_t\ \Delta x_t\ \Delta z_t) - \lambda_{11} LR1_{t-1} - \lambda_{12} LR2_{t-1} + \varepsilon_{1t}$$
$$\Delta x_t = \text{lagged}\ (\Delta y_t\ \Delta x_t\ \Delta z_t) - \lambda_{21} LR1_{t-1} - \lambda_{22} LR2_{t-1} + \varepsilon_{2t}$$
$$\Delta z_t = \text{lagged}\ (\Delta y_t\ \Delta x_t\ \Delta z_t) - \lambda_{31} LR1_{t-1} - \lambda_{32} LR2_{t-1} + \varepsilon_{3t}$$

Testing for, and estimating, a cointegrating relationship

- Pretest the variables for their order of integration
- Estimate the Cointegrating Regression
- Check whether there is a cointegrating (i.e. long run equilibrium) relationship
- If so, estimate the dynamic error correction model
- Assess model adequacy

Pretest the variables for their order of integration

- By definition cointegration necessitates that the variables be integrated of the same order

- Use DF or ADF tests to determine the order of integration

•If variables are I(0) - Standard Time Series Methods

•If the variables are integrated of different order (one I(0), one I(1) or I(2) etc) than it is possible to conclude that the two variables are not cointegrated

•If the variables are I(1), or are integrated of the same order, go on

Assess model adequacy and obtain a parsimonious final specification

Assess if the ECM model you have estimated is misspecified using standard diagnostic tests

If the model is not misspecified, use a general –to – specific modelling approach to obtain a parsimonious final model.

Cointegration to recap

If $X_t \sim I(1)$, and $Y_t \sim I(1)$, but $Z_t = Y_t - \beta X_t \sim I(0)$, then X_t and Y_t are cointegrated.
There is a long-run equilibrium relation.
The relationship is stationary and converges; any divergence is temporary.
Empirical examples: PPP, Govn't debt ratio, … 1,451 papers in EconLit

No cointegration

If X_t and Y_t are not cointegrated,

$$Z_t = Y_t - \beta X_t \sim I(1).$$

Let $\Delta Z_t = u_t$.

$Y_t - \beta X_t = (Y_o - \beta X_o)$ + sum of u_j over j=1,..,t.

initial disequil. This term gets bigger as t increases.

Some concepts

Z_t : Equil. Error
Regression: Long-run relationship
β: cointegrating vector

Some results

If X_t and Y_t are cointegrated, so are X_{t-k} and Y_t.
If $X_t \sim I(1)$, X_{t-k} and X_t are cointegrated.
If X_t and Y_t are cointegrated, there must be "Granger causality" at least in one direction.
If $Y_t \sim I(0)$, and $X_t \sim I(1)$, $Y_t = \beta X_t + u_t$ is a nonsense regression.
If one is a rational forecast of another, they are cointegrated.
A pair of jointly efficient market prices cannot be cointegrated.

ECM

Granger Representation theorem

"Cointegration implies Error Correction Model (ECM)."

Consider

$$\alpha(L)Y_t = \beta(L)X_t + u_t$$

BN decomposition implies $\alpha(L) = \alpha(1)L + (1-L)\alpha^*(L)$, and $\beta(L) = \beta(1)L + (1-L)\beta^*(L)$

$$\alpha^*(L)\ \Delta Y_t = \beta^*(L)\ \Delta X_t - \alpha(1)[Y_{t-1} - \beta(1)/\alpha(1)\ X_{t-1}] + u_t$$

Here,

$$-\alpha(1)[Y_{t-1} - \beta(1)/\alpha(1)\ X_{t-1}] = \gamma\,[Y_{t-1} - \beta X_{t-1}] = \gamma\, Z_{t-1}$$

Now, ECM is,

$$\alpha^*(L)\ \Delta Y_t = \beta^*(L)\ \Delta X_t + \gamma\, Z_{t-1} + u_t$$

ECM

If X_t and Y_t are not cointegrated, γ should be zero. Why?
If $\gamma = 0$, the above model is VAR in difference.

If $\gamma \neq 0$, VAR in difference is mis-specified.

ECM in vector form

$$A^*(L)\Delta X_t = \gamma^* Z_{t-1} + e_t$$

where $X_t = (X_{1t}\ X_{2t})'$,
and $\alpha' X_t = Z_t$

Long run variance and cointegration

Consider a vector time series, $\Delta X_t = e_t$

Denote the l.r. variance (spectral density at freq. 0) as Ω.

Cointegration implies that Ω is singular.

Some cointegration tests examine if Ω is singular.

Let $\alpha = (1, -\beta)'$. Then, $\alpha'\Omega = 0$

We can show $\beta = \Omega_{21}'\Omega_{22}^{-1}$.

Identification issue

If $\alpha' X_t = Z_t \sim I(0)$, so is $b\alpha' X_t \sim I(0)$.

In the ECM term, $\gamma^* Z_{t-1} = \gamma^* \alpha' X_{t-1}$, γ^* and α are not separately identified.

So, we need normalization.

Min $\alpha'(X'X)\alpha$ such that $\alpha'\alpha = I$

$$L = \alpha'(X'X)\alpha - 2\lambda(\alpha'\alpha - I)$$

FOC. $2(X'X)\alpha - 2\lambda\alpha = 0$ or $(X'X - \lambda I)\alpha = 0$

λ = eigen value, α = eigen vector of $X'X$

($\lambda = 1$ implies a unit root) (α: cointegration vector)

Testing for cointegration

- Residual Based Tests
  - Test for a unit root on residuals
    - ADF, PP type
    - Cointegration tests
    - Structural breaks?
- Tests using eigen values
  - Stock & Watson (1988)
  - Johansen tests

Johansen tests

Hypothesis

H_0: r(or 0) cointegration

H_a: r+m (or 1) cointegration

LR = -T Σ i =r+1 to r+m log(1 - λ_i)

where λ_i is the eigen value of the residual moment matrix. Critical values are provided in Hamilton (1994).

All parameters are jointly estimated.

Cointegration regression

Reduced form models of Johansen type are limited.

Fully Modified (FM) estimation

Hansen & Phillips (1991); Transform the data using the estimate of the l.r. variance, and use OLS.

Canonical Cointegration Regression (CCR)

Park (1992, Econ); similar

Lead and lags

Saikkonen (1991, ET)

Further developments

Spectral based regression

Hannan's efficient estimator

Testing the null of Cointegration

Cointegration with structural breaks

Testing Granger causality with cointegration

Toda & Phillips (1991, Econ)

Further developments

Cointegration in the system of equations
Cointegration in the simultaneous equation
Mixed regression
Fractional Cointegration
Nonlinear cointegration