

# COINTEGRATION

## 1 Introduction

Cointegration relates to the existence of stationary linear combinations of integrated variables.

For example, log consumption & income:  
Plausible that  $c_t, y_t \sim I(1)$ , but  $y_t - c_t \sim I(0)$ .

## 2 Univariate and Multivariate ADF Representations

Stationarity of  $AR(p)$   $\phi(L)y_t = \delta + \varepsilon_t$  depends on roots of  $\phi(z)$ .

If  $y_t \sim I(1)$  or  $y_t \sim I(0)$ , then either:

- all  $p$  roots of  $\phi(z)$  lie outside the unit circle, or
- $\phi(z)$  has a single unit root with all remaining roots outside the unit circle.

In ADF representation

$$\phi^*(L)\Delta y_t = \alpha y_{t-1} + \delta + \varepsilon_t$$

either

$$\alpha = 0 \iff y_t \sim I(1)$$

$$\alpha < 0 \iff y_t \sim I(0).$$

Now consider  $\mathbf{y}_t$ ,  $(k \times 1)$  vector.  
VAR( $P$ ):

$$\Phi(L)\mathbf{y}_t = \boldsymbol{\delta} + \boldsymbol{\varepsilon}_t.$$

We assume all  $y_{it} \sim I(0)$  or all  $y_{it} \sim I(1)$ ,  $i = 1, \dots, k$ .

Then, either:

- all roots of  $|\Phi(z)|$  are greater than 1  $\Rightarrow$  all  $y_{it} \sim I(0)$ ,  
or
- $|\Phi(z)|$  has a single unit root with all remaining roots  
greater than 1  $\Rightarrow$  all  $y_{it} \sim I(1)$ .

Latter case allows the possibility of cointegration.

VAR process can always be written in multivariate  
ADF form

$$\Phi^*(L)\Delta\mathbf{y}_t = \Pi\mathbf{y}_{t-1} + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_t$$

where

$$\Pi = -\Phi(1).$$

Unit root and cointegration properties depend on  $\Pi$   
( $k \times k$ ).

We concentrate on VAR(1) case with no intercept:

$$\mathbf{y}_t = \mathbf{\Phi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t,$$

where

$$\mathbf{\Phi}(L) = \mathbf{I}_k - \mathbf{\Phi}L.$$

$\Rightarrow$

$$\Delta \mathbf{y}_t = \mathbf{\Pi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

$$\mathbf{\Pi} = \mathbf{\Phi} - \mathbf{I}_k.$$

Note:

$$\mathbf{\Pi} = -\mathbf{\Phi}(1)$$

$\Rightarrow$

$$||\mathbf{\Phi}(1)|| = ||\mathbf{\Pi}||$$

where  $||.||$  is absolute value of the determinant.

## 2.1 Cointegration and Rank of $\Pi$

For VAR(1) in  $k$  variables,

$|\Phi(z)| = |\mathbf{I}_k - \Phi z|$  is  $k^{th}$  order polynomial in  $z$ .

Factorise as:

$$|\Phi(z)| = (1 - \nu_1 z)(1 - \nu_2 z) \dots (1 - \nu_k z).$$

Unit root in  $|\Phi(z)| \Rightarrow |\Phi(1)| = 0 \Rightarrow |\Pi| = 0$

Therefore,  $\mathbf{y}_t \sim I(1) \Rightarrow \Pi$  singular.

A specific singular matrix is  $\Pi = \mathbf{0}$ .

Aside: The rank of a matrix.

Consider  $(k \times k)$   $C$ .

If nonsingular, then  $|C| \neq 0$  and inverse  $C^{-1}$  exists.

$C$  can be nonsingular only if its rows are *linearly independent*.

Linear independence  $\Rightarrow$  no row (column) of  $C$  can be written as a linear combination of the other rows (or columns).

The number of linearly independent rows (or columns) is the *rank*.

$\Rightarrow$  If  $C$  is nonsingular,  $\text{rank}(C) = k$ ; if  $C$  is singular,  $\text{rank}(C) < k$ .

If  $\text{rank}(C) = 0$ , then  $C = 0$ .

See Chris D. Orme, *Lecture Notes in Linear Algebra*, Chapters 3 & 4.

Three possibilities for VAR(1):

1.  $\Pi = 0$ , or  $\text{rank}(\Pi) = 0$ .

As  $\Pi = \Phi - \mathbf{I}_k \Rightarrow \Phi = \mathbf{I}_k$ .

VAR is

$$\Delta \mathbf{y}_t = \boldsymbol{\varepsilon}_t.$$

$k$  distinct unit root processes, one corresponding to each element  $y_{it}$ ; no cointegration.

2.  $\Pi$  nonsingular, or  $\text{rank}(\Pi) = k \Rightarrow |\Phi(1)| \neq 0$ .

No unit root, so VAR is stationary  $\Rightarrow$  all  $y_{it} \sim I(0)$ .

3.  $\Pi$  is nonzero but singular, so  $\Pi \neq 0$ , with  $|\Pi| = 0 \Rightarrow \text{rank}(\Pi) = r$ , where  $0 < r < k$ .

Here  $|\Phi(1)| = 0$  &  $y_{it} \sim I(1)$ .

Case 3 implies cointegration:

$y_{it} \sim I(1)$ , but  $\Delta \mathbf{y}_t = \Pi \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$  with LHS stationary, so nonzero  $\Pi$  must remove the nonstationarity.

$$\Delta \mathbf{y}_t = \mathbf{\Pi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

Since  $k \times k$  matrix  $\mathbf{\Pi}$  has rank  $r < k$ , it is of *reduced rank*.

Rank  $\mathbf{\Pi} = r \Rightarrow$  there exist  $r$  distinct stationary linear combinations of  $y_{it}$  ( $i = 1, \dots, k$ ).

The coefficients of these linear combinations define  $r$  linearly independent cointegrating vectors.

Because  $r < k$ , there are at most  $k - 1$  cointegrating vectors between the  $k$  variables  $y_{it}$ .

## 2.2 Example

Bivariate VAR

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} .5 & 1.5 \\ -.1 & 1.3 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

with

$$\begin{aligned} |\mathbf{I}_2 - \Phi z| &= \begin{vmatrix} 1 - .5z & -1.5z \\ .1z & 1 - 1.3z \end{vmatrix} \\ &= 1 - 1.8z + .8z^2 = (1 - .8z)(1 - z) \end{aligned}$$

$$\Rightarrow y_{1t}, y_{2t} \sim I(1).$$

In multivariate ADF form,

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -.5 & 1.5 \\ -.1 & .3 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

& hence

$$\mathbf{\Pi} = \begin{bmatrix} -.5 & 1.5 \\ -.1 & .3 \end{bmatrix}.$$

$\mathbf{\Pi}$  is singular & with  $\text{rank}(\mathbf{\Pi}) = 1$ :

either row (or column) is a multiple of the other.

As  $\text{rank}(\mathbf{\Pi}) = r = 1$ , there is one cointegrating relationship between the two variables.



### 3 The Error Correction Model

With cointegration, there are  $r < k$  cointegrating relationships.

Define  $(k \times r)$  matrix  $\mathbf{B}$ , with columns containing the  $r$  cointegrating vectors.

By construction,  $\text{rank}(\mathbf{B}) = r$ .

Then in

$$\mathbf{z}_t = \mathbf{B}'\mathbf{y}_t,$$

$$z_{jt} \sim I(0), j = 1, \dots, r.$$

Further, it follows that

$$\mathbf{\Pi} = \mathbf{A}\mathbf{B}'$$

where  $\mathbf{A}$  is  $(k \times r)$  &  $\text{rank}(\mathbf{A}) = r$ .

$\Rightarrow$

$$\begin{aligned}\Delta\mathbf{y}_t &= \mathbf{\Pi}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= \mathbf{A}\mathbf{B}'\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= \mathbf{A}\mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t.\end{aligned}$$

For a VAR( $P$ ), this becomes

$$\begin{aligned}\Phi^*(L)\Delta\mathbf{y}_t &= \Pi\mathbf{y}_{t-1} + \varepsilon_t \\ &= \mathbf{A}\mathbf{B}'\mathbf{y}_{t-1} + \varepsilon_t \\ &= \mathbf{A}\mathbf{z}_{t-1} + \varepsilon_t.\end{aligned}$$

This is the *ECM representation*.

ECM is abbreviation for error-correction mechanism or (sometimes) equilibrium-correction mechanism.

Cointegrating relationships are long-run equilibrium relationships among  $\mathbf{y}_t$ ;

$\mathbf{z}_t$  is the disequilibrium in period  $t$ .

The *adjustment* matrix  $\mathbf{A}$  shows how each variable adjusts to achieve the longrun relationships.

## 3.1 Examples

### 3.1.1 The VAR(1) Again

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} .5 & 1.5 \\ -.1 & 1.3 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

One cointegrating relationship, given by a row of  $\Pi$ .  
To make it unique, scale coefficient of  $y_1$  to unity  
 $\Rightarrow y_{1t} = 3y_{2t}$  is the longrun relationship,  
while

$$z_t = y_{1t} - 3y_{2t}$$

measures disequilibrium.

Cointegrating vector is  $\mathbf{B}' = [1, -3]$  and

$$\begin{aligned} \mathbf{\Pi} &= \begin{bmatrix} -.5 & 1.5 \\ -.1 & .3 \end{bmatrix} = \begin{bmatrix} -.5 \\ -.1 \end{bmatrix} [1 \quad -3] \\ \Rightarrow \mathbf{A} &= \begin{bmatrix} -.5 \\ -.1 \end{bmatrix}. \end{aligned}$$

ECM equations

$$\Delta \mathbf{y}_t = \mathbf{A} \mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t.$$

are

$$\Delta y_{1t} = -.5(y_{1,t-1} - 3y_{2,t-1}) + \varepsilon_{1t}$$

$$\Delta y_{2t} = -.1(y_{1,t-1} - 3y_{2,t-1}) + \varepsilon_{2t}$$

$$\begin{aligned}\Delta y_{1t} &= -.5(y_{1,t-1} - 3y_{2,t-1}) + \varepsilon_{1t} \\ \Delta y_{2t} &= -.1(y_{1,t-1} - 3y_{2,t-1}) + \varepsilon_{2t}\end{aligned}$$

Coefficient  $-.5 \Rightarrow y_1$  changes to remove half of last period's disequilibrium;

coefficient  $-.1 \Rightarrow y_2$  moves further from equilibrium.

But overall, system moves to correct disequilibrium from earlier periods.

Suppose: at  $t = 0$ ,  $y_{1,0} = 10$ ,  $y_{2,0} = 4 \Rightarrow z_0 = -2$ .

At  $t = 1$  (no disturbances):  $y_{1,1} = y_{1,0} - .5 \times -2 = 11$   
&  $y_{2,1} = y_{2,0} - .1 \times -2 = 4.2$ .

$\Rightarrow z_1 = 11 - 3 \times 4.2 = -1.6$ , reducing the disequilibrium.

### 3.1.2 Three Equation Case

Suppose  $\mathbf{y}_t$  is  $3 \times 1$  with each element  $I(1)$ . In VAR(1) with  $\delta = 0$ , ADF representation is

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}.$$

If  $r = 1$ , then  $z_t$  is scalar:

$$z_t = \mathbf{B}'\mathbf{y}_t = y_{1t} + \beta_2 y_{2t} + \beta_3 y_{3t}$$

where we normalise on  $y_{1t}$ .

Matrix  $\mathbf{A}$  is  $3 \times 1 \Rightarrow \text{ECM}$

$$\begin{aligned} \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} &= \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} z_{t-1} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \\ &= \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} 1 & \beta_2 & \beta_3 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}. \end{aligned}$$

Two elements in  $\mathbf{B}$  & three in  $\mathbf{A}$ , but nine in  $\mathbf{\Pi}$ ;

$\Rightarrow 4 (= 9 - 5)$  restrictions between elements of  $\mathbf{\Pi}$  implied by the 1 cointegrating relationship.

If  $r = 2$  cointegrating relationships, we can write

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} 1 & \beta_{12} & \beta_{13} \\ \beta_{21} & 1 & \beta_{23} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix}$$

& ECM is

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}.$$

Thus, either or both  $z_{1,t-1}$  and  $z_{2,t-1}$  can enter each of the three equations of the ECM.

Note: The normalisation  $\beta_{11} = \beta_{22} = 1$  is arbitrary; we could (for example), normalise as  $\beta_{11} = \beta_{21} = 1$ .

In normalising, set a nonzero coefficient in each equation to unity.

In

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

each of  $z_{1,t-1}$  and  $z_{2,t-1}$  must enter *at least one* ECM equation.

If  $a_{12} = a_{22} = a_{32} = 0$   
second cointegrating relationship cannot apply,  
as no variable adjusts to enforce the equilibrium.

Thus, at least one  $a_{ij} \neq 0$  ( $i = 1, \dots, k$ ) checks  
existence of  $j$ th cointegrating relationship;

$\Rightarrow$  require such a nonzero coefficient for each  $z_{j,t-1}$ .

## 4 Cointegration Modelling

Engle-Granger approach is residual-based:  
estimate presumed cointegrating relationship & use residuals to check for the existence of cointegration.

This implicitly assumes that there is (at most) one cointegrating relationship.

For  $k = 2$ , works well.

For  $k > 2$   $I(1)$  variables, may be  $k - 1$  cointegrating relationships;

Residual-based approach does not separately identify these;

Hence not generally applicable with  $k > 2$ .

Most common multivariate approach is due to Soren Johansen.



## 4.1 Testing Number of Cointegrating Vectors

Every  $k \times k$  square matrix has  $k$  eigenvalues;  
rank = number of non-zero eigenvalues.

Denote the  $k$  eigenvalues of  $\Pi$  as  $\lambda_i (i = 1, \dots, k)$ ,  
these satisfy  $0 \leq \lambda_i < 1$ ;  
order so that  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ .

Johansen method how many  $\lambda_i$  are positive.

$\text{rank}(\Pi) = r \Rightarrow \lambda_i > 0 \ i = 1, \dots, r$   
&  $\lambda_i = 0 \ i = r + 1, \dots, k$ .

Johansen approach sequentially tests:

$$\begin{array}{ll} H_{00} : r = 0 & H_{A0} : r > 0 \\ H_{01} : r = 1 & H_{A1} : r > 1 \\ & \vdots \\ H_{0,k-1} : r = k - 1 & H_{A,k-1} : r = k \end{array}$$

$H_{00}$  is considered first. Rejection  $\Rightarrow$  at least one cointegrating relationship.

Consider  $H_{01}$ . If not rejected, conclude  $r = 1$ .

If  $H_{01}$  rejected, next consider  $H_{02} : r = 2$ , etc.

General procedure:

Sequentially test  $H_{00}, H_{01}, \dots$  until  $H_{0r}$  is not rejected.

Non-rejection of  $H_{0r}$  yields estimated  $r$ .

Johansen developed two test statistics:

- *Trace* statistic

$$\hat{\eta}_r = -T \sum_{i=r+1}^k \log(1 - \hat{\lambda}_i), \quad r = 0, 1, \dots, k - 1.$$

Reject  $H_{0r}$  if  $\hat{\eta}_r$  is significantly larger than 0.

- *Maximal eigenvalue* statistic

$$\hat{\varsigma}_{r+1} = -T \log(1 - \lambda_{r+1}), \quad r = 0, 1, \dots, k - 1.$$

Reject  $H_{0r}$  if  $\hat{\varsigma}_{r+1}$  is significantly larger than 0.

In practice, for estimated eigenvalues, these do not necessarily yield same estimated  $r$ .

Cannot use conventional hypothesis tests;  
Anologous problem to univariate ADF tests;  
Critical values tabulated by Johansen and others.

Note: If all null hypotheses are rejected, conclude  $\text{rank}(\Pi) = k$  and  $\mathbf{y}_t \sim I(0)$ .

But impossible if variables in  $\mathbf{y}_t$  are  $I(1)$ .

## 4.2 Estimating the ECM

Estimating  $r$  is the most difficult problem in cointegration modelling.

Once  $r$  is obtained, super-consistent estimate of  $\mathbf{B}$  is a side-product;

$\Rightarrow \mathbf{z}_t = \mathbf{B}'\mathbf{y}_t$  estimated.

ECM for VAR(1):

$$\Delta \mathbf{y}_t = \mathbf{A}\mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t.$$

Due to super-consistency,  $\mathbf{z}_t$  can be treated as observed;

$\Rightarrow$  coefficients in  $\mathbf{A}$  can be estimated by OLS, equation by equation.

Example: With  $k = 3$  and  $r = 2$ , ECM equations

$$\Delta y_{1t} = a_{11}z_{1,t-1} + a_{12}z_{2,t-1} + \varepsilon_{1t}$$

$$\Delta y_{2t} = a_{21}z_{1,t-1} + a_{22}z_{2,t-1} + \varepsilon_{2t}$$

$$\Delta y_{3t} = a_{31}z_{1,t-1} + a_{32}z_{2,t-1} + \varepsilon_{3t}$$

can be estimated by OLS.

In more general ECM

$$\Delta \mathbf{y}_t = \mathbf{A} \mathbf{z}_{t-1} + \Phi_1^* \Delta \mathbf{y}_{t-1} + \dots \\ + \Phi_{p-1}^* \Delta \mathbf{y}_{t-(p-1)} + \boldsymbol{\varepsilon}_t.$$

OLS estimation is valid for elements of  $\mathbf{A}$ ,  $\Phi_1^*$ , ...,  $\Phi_{p-1}^*$ .

### 4.3 Hypothesis Testing in ECM

Again due to super-consistency for cointegrating relationships, usual tests apply (in large samples) for adjustment coefficients in  $\mathbf{A}$ .

Thus, conventional  $t$  and  $F$ —tests can be applied for  $a_{ij}$ .

Recall:  $\text{rank}(\mathbf{A}) = r \Rightarrow$  each disequilibrium  $z_{j,t-1}$  ( $j = 1, \dots, r$ ) must appear in at least one equation.

The adjustment coefficients indicate how variables adjust to long-run disequilibria, namely  $\mathbf{z}_{t-1}$ .  
In

$$\Delta y_{1t} = a_{11}z_{1,t-1} + a_{12}z_{2,t-1} + \varepsilon_{1t}$$

$$\Delta y_{2t} = a_{21}z_{1,t-1} + a_{22}z_{2,t-1} + \varepsilon_{2t}$$

$$\Delta y_{3t} = a_{31}z_{1,t-1} + a_{32}z_{2,t-1} + \varepsilon_{3t}$$

If  $a_{31} = a_{32} = 0$ ,  $y_{3t}$  is unaffected by disequilibria  $z_{1,t-1}$  or  $z_{2,t-1}$ .

$\Rightarrow y_{3t}$  is said to be *weakly exogenous* to the system.

Test of weak exogeneity of  $y_{3t}$  is joint test of  $a_{31} = a_{32} = 0$ , usually via  $F$ -test.

For VAR( $P$ ), ECM is

$$\Delta \mathbf{y}_t = \mathbf{A} \mathbf{z}_{t-1} + \Phi_1^* \Delta \mathbf{y}_{t-1} + \dots + \Phi_{p-1}^* \Delta \mathbf{y}_{t-(p-1)} + \boldsymbol{\varepsilon}_t.$$

Conduct hypothesis testing for  $\mathbf{A}$  &  $\Phi_1^*, \dots, \Phi_{p-1}^*$ , in the conventional way.

Test of  $H_0 : \Phi_{p-1}^* = \mathbf{0}$  uses Likelihood ratio statistic, asymptotically  $\chi^2$  with  $k^2$  degrees of freedom under  $H_0$ .

If  $a_{i1} = a_{i2} = \dots = a_{ir} = 0$ :

$y_{it}$  does not adjust to disequilibria;

$y_{it}$  weakly exogenous to the system,

But  $\Delta y_{it}$  may still be influenced by lagged  $\Delta y_j$  ( $j \neq i$ ).

## 4.4 Intercepts and Cointegration

For univariate  $y_t \sim I(1)$

$$\phi^*(L)\Delta y_t = \delta + \varepsilon_t$$

$\delta \neq 0 \Rightarrow y_t$  trends over time.

Nonzero intercepts in a VAR can also allow trends in  $I(1)$  variables.

However, cointegration raises some specific issues.

VAR(1) with intercepts is

$$\Delta \mathbf{y}_t = \mathbf{\Pi} \mathbf{y}_{t-1} + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_t,$$

Assume all  $y_{it} \sim I(1)$ , with  $0 < r < k$  cointegrating relationships.

Thus

$$\Delta \mathbf{y}_t = \mathbf{A} \mathbf{B}' \mathbf{y}_{t-1} + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_t.$$

Two possibilities for  $\boldsymbol{\delta}$ :

1. Restrict  $E(y_{it})$  to be constant over time;
2. allow  $E(y_{it})$  to trend.

Critical values for cointegration differ for these.

#### 4.4.1 Restricted Intercepts

Intercepts arise only from cointegrating relationships.

Then include intercepts in the cointegrating relationships,  
write

$$\mathbf{z}_t = \mathbf{B}'\mathbf{y}_t + \boldsymbol{\delta}_0$$

with  $E(\mathbf{z}_t) = \mathbf{0}$ .

Since  $\mathbf{z}_t$  is  $r \times 1$ ,  $\boldsymbol{\delta}_0$  is also  $r \times 1$ .

$\Rightarrow$

$$\begin{aligned}\Delta\mathbf{y}_t &= \mathbf{A}\mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= \mathbf{A}(\mathbf{B}'\mathbf{y}_{t-1} + \boldsymbol{\delta}_0) + \boldsymbol{\varepsilon}_t \\ &= \mathbf{A}\mathbf{B}'\mathbf{y}_{t-1} + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_t\end{aligned}$$

where

$$\boldsymbol{\delta} = \mathbf{A}\boldsymbol{\delta}_0.$$

Then intercepts are *restricted to the cointegrating space*.

Restrictions result from  $k$  elements in  $\boldsymbol{\delta}$ , but only  $r < k$  elements in  $\boldsymbol{\delta}_0$ .



Using  $E(\mathbf{z}_t) = \mathbf{0}$  in

$$\Delta \mathbf{y}_t = \mathbf{A} \mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t$$

$\Rightarrow$

$$E(\Delta \mathbf{y}_t) = \mathbf{A} E(\mathbf{z}_{t-1}) = \mathbf{0}.$$

Hence

$$E(\mathbf{y}_t) = E(\mathbf{y}_{t-1})$$

$\Rightarrow$  no trend is permitted in any element of  $\mathbf{y}_t$ .

Thus, restricted intercepts do not permit a trend in any  $y_{it}$ .

$\Rightarrow$  cointegration with restricted intercepts is inappropriate when any variable(s) trend over time.

#### 4.4.2 Unrestricted Intercepts

This case allows intercepts short-run dynamics, as well as in cointegrating relationships.

Again write

$$\mathbf{z}_t = \mathbf{B}'\mathbf{y}_t + \delta_0$$

with  $E(\mathbf{z}_t) = \mathbf{0}$ .

But in

$$\Delta\mathbf{y}_t = \mathbf{A}\mathbf{B}'\mathbf{y}_{t-1} + \delta + \varepsilon_t.$$

$\delta$  can take any value;  
 $\delta = \mathbf{A}\delta_0$  is not required.

Effect is to allow

$$E(\Delta\mathbf{y}_t) \neq \mathbf{0}$$

& hence

$$E(\mathbf{y}_t) \neq E(\mathbf{y}_{t-1}).$$

$\Rightarrow$  allows trends in  $\mathbf{y}_t$ .

If cointegration analysis is applied to  $I(1)$  variables where any trend(s), unrestricted intercepts should be employed.

### 4.4.3 Example

Say  $k = 2$  with  $r = 1$  cointegrating relationship

$$z_t = y_{1t} + \beta y_{2t} + \delta_0$$

where  $E(z_t) = 0$ .

If intercept is restricted, ECM equations are

$$\begin{aligned}\Delta y_{1t} &= \alpha_1 z_{t-1} + \varepsilon_{1t} \\ \Delta y_{2t} &= \alpha_2 z_{t-1} + \varepsilon_{2t}.\end{aligned}$$

Since  $E(z_t) = E(z_{t-1}) = 0$ ,  
 $E(y_{1t}) = E(y_{1,t-1})$  &  $E(y_{2t}) = E(y_{2,t-1})$   
 $\Rightarrow$  neither  $y_{1t}$  nor  $y_{2t}$  trends.

If intercept is unrestricted,

$$\Delta y_{1t} = \alpha_1 z_{t-1} + \delta_1 + \varepsilon_{1t}$$

$$\Delta y_{2t} = \alpha_2 z_{t-1} + \delta_2 + \varepsilon_{2t}$$

&

$$E(y_{1t}) = E(y_{1,t-1}) + \delta_1$$

$$E(y_{2t}) = E(y_{2,t-1}) + \delta_2.$$

So that both may trend over time.

Also note values of  $\delta_1$  and  $\delta_2$  are linked.  
As  $E(z_t) = 0$ , then  $E(\Delta z_t) = 0$ .

Cointegrating relationship

$$z_t = y_{1t} + \beta y_{2t} + \delta_0$$

$\Rightarrow$

$$E(\Delta y_{1t}) + \beta E(\Delta y_{2t}) = 0$$

$\Rightarrow$

$$E(\Delta y_{1t}) = -\beta E(\Delta y_{2t}).$$

But  $E(\Delta y_{1t}) = \delta_1$  and  $E(\Delta y_{2t}) = \delta_2$ . Therefore,

$$\delta_1 = -\beta \delta_2$$

must be true.