BANK OF ENGLAND

Economic modelling and forecasting

2-6 February 2015

Cointegration and error-correction

Ole Rummel

Adviser, CCBS at the Bank of England

ole.rummel@bankofengland.co.uk

BANK OF ENGLAND

Outline

- Unit roots and cointegration: motivation
- The problem of spurious regressions
- Cointegration: the long-run
- Estimating and testing single cointegrating relationships
- What happens if there is more than one cointegrating relationship?
- Vector autoregressions and vector error-correction mechanisms
- The Johansen methodology
- Hypothesis testing in cointegrated systems
- Weak exogeneity
- Summary

Unit roots and cointegration: motivation (1)

- Macroeconomic data often contain one or more **unit roots**
- It turns out that models of data with unit roots require special treatment
- In particular, regressions involving variables containing unit roots or stochastic trends can give rise to **spurious regressions** (Granger and Newbold (1974))
- Regression models with integrated data need to be **cointegrated** to be properly specified – in other words, regressions with I(1) data only make sense when the data are cointegrated
- An excellent source for this material is Maddala and Kim (1998)
- Many of the original papers on unit roots and cointegration can be found in Engle and Granger (1991)

Unit roots and cointegration: motivation (2)

- An important implication of the analysis of stochastic trends (unit roots) and unit-root tests is that nonstationary time series can be rendered stationary by differencing the series
- This use of the differencing operator represents a **univariate** approach to achieving stationarity
- In the case of $n > 1$ nonstationary series $Y_t = (y_{1t}, y_{2t}, \ldots , y_{nt})'$, an alternative method of achieving stationarity is to form **linear combinations** of the series
- The ability to find stationary linear combinations of nonstationary time series is known as **cointegration** (Granger (1981), Engle and Granger (1987))

Unit roots and cointegration: motivation (3)

- As we will see, the existence of cointegration among a set of nonstationary time-series variables has three important implications:
  - cointegration implies a set of dynamic long-run equilibria, where the weights used to achieve stationarity in the linear combination represent the parameters of the equilibrium relationship;
  - modelling a system of cointegrated variables allows for the specification of both long- and short-run dynamics in terms of a vector error-correction model (VECM); and
  - the estimates of the weights to achieve stationarity (the long-run parameter estimates) converge to their population values at a super-consistent rate of T compared to the usual $\sqrt{T}$ rate of convergence for stationary variables

The concept of cointegration (1)

- An n-dimensional vector of time-series variables, $Y_t = (y_{1t}, y_{2t}, \ldots, y_{nt})' = (y_{1t}, Y_{2t}')'$, is integrated of order d, denoted I(d) or $y_t \sim$ I(d), if at least one of its components is I(d)
- Furthermore, linear combinations involving I(d) variables also have d unit roots, i.e., are distributed as I(d)
- In general, regressing two (or more) I(d) variables, where $d > 0$, leads to the problem of a spurious regression
- Consider the following regression:

$$y_{1t} = \beta_0 + \sum_{j=2}^{n} \beta_j y_{jt} + \varepsilon_t = \beta_0 + \beta_2 y_{2t} + \ldots + \beta_n y_{nt} + \varepsilon_t = \beta_0 + \beta' Y_{2t} + \varepsilon_t \quad (1)$$

  where all the variables y_{jt} ($j = 1, 2, \ldots, n$) are I(1)
- This seems a prime candidate for a **spurious regression**

The concept of cointegration (2)

- Let $Y_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$ denote an $(n \times 1)$ vector of I(1) time series
- Y_t is **cointegrated** if there exists an $(n \times 1)$ vector $\beta = (\beta_1, \dots, \beta_n)'$ such that:

$$\beta' Y_t = \beta_1 y_{1t} + \dots + \beta_n y_{nt} \sim I(0) \qquad (2)$$

- In words, the nonstationary time series in Y_t are cointegrated if there is a linear combination of them that is stationary or I(0) (if some elements of β are equal to zero, then only the **subset** of the time series in y_t with non-zero coefficients is cointegrated)

The concept of cointegration (3)

- The linear combination $\beta' Y_t$ is often motivated by economic theory and referred to as a **long-run equilibrium relationship**:
  - the intuition is that I(1) time series with a long-run equilibrium relationship cannot drift too far apart from equilibrium because economic forces will act to restore the equilibrium relationship
- Even though cointegration is a purely statistical concept, it nonetheless provides a basis for interpreting (and validating) a number of models in macroeconomics and finance in terms of long-run relationships

The concept of cointegration (4)

- But note that the cointegrating vector β in (2) is not unique since for any scalar c the linear combination $c\beta' Y_t = \beta^{*\prime} Y_t$ is also I(0)
- Hence, some **normalisation** assumption is required to uniquely identify β
- A typical normalisation is:

$$\beta = (1, -\beta_2, \dots, -\beta_n)'$$

so that the cointegration relationship may be expressed as:

$$\beta' Y_t = y_{1t} - \beta_2 y_{2t} - \dots - \beta_n y_{nt} \sim I(0) \qquad (3)$$

or

$$y_{1t} = \beta_2 y_{2t} + \dots + \beta_n y_{nt} + u_t \qquad (4)$$

where the error term $u_t \sim I(0)$ – u_t is often referred to as the **disequilibrium error** or the **cointegrating residual**

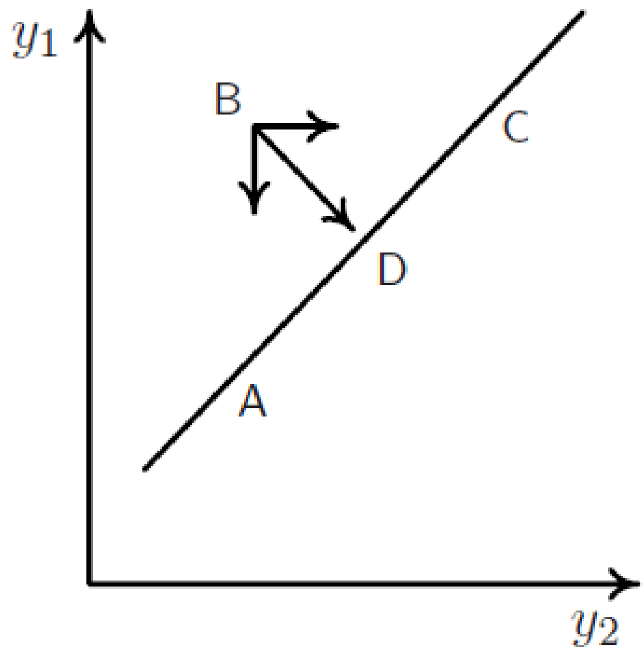
Equilibrium adjustment (1)

- Assume that we have two variables, y_{1t} and y_{2t}, who share a long-run equilibrium relationship given by:

$$y_{1t} = \mu + \beta y_{2t} + \varepsilon_t \qquad (5)$$

where ε_t is a mean-zero disturbance term and although the equation is normalised with respect to y_{1t}, the notation is deliberately chosen to reflect the fact that both variables may be endogenously determined

Equilibrium adjustment (2)



- Phase diagram illustrating the equilibrium adjustment if two variables are cointegrated

Equilibrium adjustment (3)

- For $\beta > 0$, the two-variable system is in long-run equilibrium anywhere along the line *ADC*
- Now suppose there is a shock to the system such that $y_{1,t-1} > \mu + \beta y_{2,t-1}$ or, equivalently, $\varepsilon_{t-1} > 0$ and the system is displaced to point *B*
- A long-run equilibrium relationship necessarily implies that any shock to the system will result in an adjustment taking place in such a way that equilibrium will be restored
- There are three cases

Equilibrium adjustment (4)

- The adjustment is done by y_{1t}:

$$\Delta y_{1t} = \alpha_1(y_{1,t-1} - \mu - \beta y_{2,t-1}) + \varepsilon_{1t} \qquad (6)$$

- Since $y_{1,t-1} - \mu - \beta y_{2,t-1} > 0$, inspection of (6) reveals that Δy_{1t} should be negative, which in turn suggests the restriction $\alpha_1 < 0$
- The adjustment is done by y_{2t}:

$$\Delta y_{2t} = \alpha_2(y_{1,t-1} - \mu - \beta y_{2,t-1}) + \varepsilon_{2t} \qquad (7)$$

- Since $y_{1,t-1} - \mu - \beta y_{2,t-1} > 0$, inspection of (7) reveals that Δy_{2t} should be positive, which in turn suggests the restriction $\alpha_2 > 0$
- Both y_{1t} and y_{2t} adjust, in which case both equations (6) and (7) operate with y_{1t} decreasing and y_{2t} increasing – the strength of the movements in the two variables is determined by the relative magnitudes of the parameters α_1 and α_2

Looking for cointegration (1)

- The only way to test for cointegration is through a careful statistical analysis
- What would be empirical evidence of cointegration?
- An obvious candidate for further analysis are the **cointegrating residuals**
- These are given by:

$$u_t = y_{1t} - \hat{\beta}_0 - \sum_{j=2}^{n} \hat{\beta}_j y_{jt} \tag{8}$$

and if they are stationary, then the underlying (long-run) relationship is cointegrated

- Any unit-root test can be used for this purpose

Looking for cointegration (2)

- An important implication of the definition of cointegration is that if we have **two** variables which are integrated of **different** orders then these two series cannot possibly be cointegrated
- It is, however, possible to have a mixture of different order series when there are **three or more** series under consideration
- In this case, a subset of the higher-order series may cointegrate to the order of the lower-order series:
  - suppose y_{1t} ~ I(1), y_{2t} ~ I(2) and y_{3t} ~ I(2);
  - if y_{2t} and y_{3t} cointegrate, then $v_t = ay_{2t} + by_{3t}$ will be I(1), meaning that v_t is now a potential candidate to cointegrate with the remaining I(1) series, y_{1t};
  - if so, then $z_t = cv_t + dy_{1t}$ will be I(0), such that y_{2t}, y_{3t} ~ CI(2,1), v_t, y_{1t} ~ CI(1,1) and z_t ~ I(0)

Estimating single cointegrating equations

- When looking at single equations it is easy to test for cointegration:
  - Engle and Granger’s (1987) two-step residual-based procedure for testing the null of no cointegration (extended to more than two variables by Engle and Yoo (1991));
  - Kremers *et al.*‘s (1992) test on the speed of the adjustment parameter in an error-correction model;
  - Saikkonen (1991) and Stock and Watson’s (1993) dynamic OLS (DOLS) estimation;
  - Phillips and Hansen’s (1990) fully-modified OLS (FMOLS) estimation; and
  - Pesaran *et al.*’s (2001) autoregressive distributed lag (ARDL) bounds testing approach

OLS as a cointegration estimator

- OLS is a good tool for estimating a cointegrating vector
- It is **superconsistent** because estimates of the cointegrating vector with nonstationary I(1) variables converge in distribution to the true value faster than for nonintegrated I(0) regressors, as shown by Stock (1987)
- This is just the same sort of property of the regressors which yield spurious regressions...
- ...in particular, the regressors tend to their asymptotic distributions faster
- A consistent estimator β converges to its true value at rate $\sqrt{T}$, while a superconsistent estimator converges at the rate T
- Note, however, that the OLS estimate will be biased and inefficient

The Engle and Granger two-step procedure (1)

- Static regressions provide a framework for testing cointegration, based on the OLS residuals
- As OLS is superconsistent, Engle and Granger (1987) suggest a two-step estimation procedure for cointegrated models
- We first verify that y_{1t} and y_{2t} are both I(1) and then estimate a bivariate version of (1), i.e., j = 2, by OLS, take the estimated residuals, u_t, and run an ADF test for stationarity ($\rho \leq 1$) using:

$$\Delta u_t = (\rho - 1) u_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta u_{t-i} + \varepsilon_t \qquad (9)$$

- This is a **residual-based** test for cointegration
- In performing the test remember we want to **reject** a unit root (ρ = 1) in the residuals for cointegration to be present

The Engle and Granger two-step procedure (2)

- When performing the unit-root test do not include a constant and/or trend in the first (estimation) step and then **also** in the ADF test (the second step)
- Only do this in **one** of the two steps:
  - since the u_t are residuals from a regression equation that generally includes a constant, there is no need to include an intercept term in the unit-root test; and
  - since Hansen (1992) demonstrated that including a time trend in the ADF test results in a loss of power, it is better to leave it out in the ADF test
- In other words, we generally do **not** include any deterministic variables (constant and time trend) in the second-stage ADF test

Critical values (1)

- Any of the standard unit-root tests can be used, but the critical values will be **different**
- If we did not need to estimate the parameters, and hence knew the cointegrating vector exactly, we could simply use the usual Dickey-Fuller table for the unit-root test
- But if the cointegrating vector is unknown then it must be estimated prior to testing for a unit root in the residuals
- The test is therefore based on **estimated** residuals – the procedure is prejudiced toward finding a stationary error process
- This preliminary step makes it necessary to use different critical values for the subsequent unit-root test

Critical values (2)

- Moreover – and rather unexpectedly – the test now depends on the number of regressors in much the same way that we have to be careful about the use of constants and trends in unit-root tests
- We need to take account of this in determining critical values and therefore correct for the number of regressors
- Engle and Yoo (1991) and MacKinnon (1991) suggested alternative critical values which depend on the number of regressors
- MacKinnon (1991), in particular, gives critical values for cointegration tests which take account of whether the test has been run using a constant and/or a trend and the number of regressors, n ($1 \leq n \leq 6$)

The Engle and Granger two-step procedure (3)

- Despite its popularity, the Engle-Granger two-step procedure has several weaknesses:
  - the step-wise procedure implies a possible compounding of errors;
  - any of the variables can be selected as the left-hand side variable;
  - the procedure rules out multiple cointegration between more than two variables; and
  - due to the unreliable standard errors, we cannot perform statistical tests on the estimated cointegrating vector (Phillips and Durlauf (1986))
- In fact, the single-equation approach is really only applicable when there is a single unique cointegrating vector and when all the right-hand side variables are weakly exogenous

The Engle and Granger two-step procedure (4)

- The estimation of the long-run equilibrium regression requires that the researcher place **one** variable on the left-hand side and use the other(s) as regressor(s)
- For example, in the case of two variables y_{1t} and y_{2t}, it is possible to run the Engle and Granger test for cointegration by using the estimated residuals from either of the two following 'equilibrium' regressions:

$$y_{1t} = \beta_{10} + \beta_{12} y_{2t} + u_{1t} \quad (10)$$

$$y_{2t} = \beta_{20} + \beta_{22} y_{1t} + u_{2t} \quad (11)$$

The Engle and Granger two-step procedure (5)

- As the sample size grows infinitely large, asymptotic theory indicates that the test for a unit root in the $\{u_{1t}\}$ sequence becomes equivalent to the test for a unit root in the $\{u_{2t}\}$ sequence
- In practice, it is possible to find that one regression indicates that the variables are cointegrated, whereas reversing the order indicates no cointegration

Engle and Granger with more than two variables (1)

- When there are only two variables it is possible to show that the cointegrating vector is unique
- When one moves beyond the two-variable case the possibility immediately arises of there being more than one cointegration relation
- A cointegration relation is a linear combination of I(1) variables that is I(0)
- With k (> 2) variables, all of which are I(1), there may be 0, 1, 2, … , k-1 cointegration relations…
- …and there will thus be problems with the (bivariate) Engle-Granger regression approach

Engle and Granger with more than two variables (2)

- Assume three I(1) variables: y_{1t}, y_{2t} and y_{3t}
- Technically, in order to test for cointegration between y_{1t}, y_{2t} and y_{3t}, we have to run the following **three** regressions:

$$y_{1t} = \alpha_1 + \alpha_2 y_{2t} + \alpha_3 y_{3t} + u_{1t} \tag{12}$$

$$y_{2t} = \beta_1 + \beta_2 y_{1t} + \beta_3 y_{3t} + u_{2t} \tag{13}$$

$$y_{3t} = \gamma_1 + \gamma_2 y_{1t} + \gamma_3 y_{2t} + u_{3t} \tag{14}$$

and determine whether the **three** estimated residual series u_{1t}, u_{2t} and u_{3t} from the equilibrium regressions are stationary

Engle and Granger with more than two variables (3)

- In many circumstances, only a subset of variables is cointegrated
- Suppose that y_{1t}, y_{2t} and y_{3t} are three I(1) variables and that y_{1t} and y_{2t} are cointegrated such that $y_{1t} – \alpha_2 y_{2t}$ is stationary;
  - a regression of y_{1t} on the other two variables – as in (12) – should yield the stationary relationship $y_{1t} = \alpha_1 + \alpha_2 y_{2t} + (0)y_{3t}$;
  - similarly, a regression of y_{2t} on the other two variables y_{1t} and y_{3t}, as in (13), should yield the stationary relationship $y_{2t} = \beta_1 + \beta_2 y_{1t} + (0)y_{3t} = y_{2t} = -(\alpha_1/\alpha_2) + (1/\alpha_2)y_{1t} + (0)y_{3t}$; but
  - a regression of y_{3t} on y_{1t} and y_{2t} – as in (14) – cannot reveal the cointegrating relationship (as it does not involve a transform of α_2)

Engle and Granger with more than two variables (4)

- In performing the ADF test on the estimated residuals, there is no presumption that any one of the residual series is preferable to any of the others
- But not all three equilibrium regressions may yield the same conclusion
- We should therefore be wary of a result indicating that the variables are cointegrated using one variable for normalisation, but are not cointegrated using another variable for normalisation

Alternative testing methodologies (1)

- The single two-step estimator of Engle and Granger is **not** asymptotically efficient
- The econometric problems with the Engle-Granger two-step procedure arise from the potential endogeneity of y_{2t} and autocorrelation in the disturbances u_t when simply estimating:

$$y_{1t} = \mu + \beta y_{2t} + u_t \qquad (15)$$

  by ordinary least squares
- Thus, while it is not necessary to take into account the short-run dynamics from the error-correction model to obtain superconsistent estimates of the long-run parameters, it is necessary to model the short-run dynamics to obtain an efficient estimator with *t*-statistics that have the standard distribution

Alternative testing methodologies (2)

- The ordinary least squares estimator of β is superconsistent but inefficient
- Solutions to the efficiency problem and bias introduced by possible endogeneity of the right-hand side variable and serial correlation have been addressed within single-equation frameworks
- Other **efficient** estimators are the three-step estimator by Engle and Yoo (1991), the dynamic OLS (DOLS) estimator due to Saikkonen (1991) and Stock and Watson (1993), the fully-modified OLS (FMOLS) estimator of Phillips and Hansen (1990), the nonlinear single-equation least-squares estimator due to Phillips and Loretan (1991) and Park's (1992) canonical cointegrating regression

Alternative testing methodologies (3)

- In addition, Johansen's (1988, 1991) maximum-likelihood procedure – to be discussed below – also falls in this category
- For all these **efficient** estimators, the *t*- and Wald-statistics for hypotheses about the cointegrating coefficient, β, correctly rescaled if necessary, have **standard** asymptotic distributions

Alternative testing methodologies (4)

- A very general specification that controls for the possible endogeneity of the right-hand side variable and serial correlation contains lag polynomials, allowing for different lag orders and also leads as well as lags
- In other words, this general specification will allow for both the leads and the lags of the cointegrating relationship, $(y_{1t} - \beta y_{2t})$, and leads and lags of Δy_{2t}
- A reduced-form version of this general specification is:

$$y_{1t} = \beta y_{2t} + \sum_{i=-q, i \neq 0}^{r} \gamma_i (y_{1,t-i} - \beta y_{2,t-i}) + \sum_{j=-s}^{t} \delta_i \Delta y_{2,t-j} + \eta_t \quad (16)$$

- The intuition behind this approach is that improved estimates of the long-run parameters, β, can be obtained by using information in the short-run dynamics

Phillips and Loretan (1991)

- The Phillips and Loretan (1991) estimator excludes the leads of the cointegrating vector from (16)
- The equation is:

$$y_{1t} = \beta y_{2t} + \sum_{i=1}^{r} \gamma_i (y_{1,t-i} - \beta y_{2,t-i}) + \sum_{j=-s}^{t} \delta_i \Delta y_{2,t-j} + \eta_t \qquad (17)$$

which needs to be estimated by **nonlinear** least squares

- This procedure yields (super) consistent and asymptotically efficient estimates of the cointegrating vector if all the (exclusion) restrictions are satisfied

Dynamic ordinary least squares

- The dynamic ordinary least squares (DOLS) due to Saikkonen (1991) and Stock and Watson (1993) excludes the lags and leads of the cointegrating vector from (16)
- The equation is:

$$y_{1t} = \beta y_{2t} + \sum_{j=-s}^{t} \delta_i \Delta y_{2,t-j} + \eta_t \qquad (18)$$

which can be quite easily estimated by ordinary least squares

- This procedure yields (super) consistent and asymptotically efficient estimates of the cointegrating vector if all the (exclusion) restrictions are satisfied
- This approach is available in EViews 7

Fully modified ordinary least squares

- The fully modified ordinary least squares (FMOLS) due to Phillips and Hansen (1990) excludes the lags and leads of the cointegrating vector from (16) and limits the terms in Δy_{2t} to the contemporaneous difference with coefficient δ
- The equation is:

$$y_{1t} = \beta y_{2t} + \delta \Delta y_{2t} + \eta_t \qquad (19)$$

- The FMOLS methodology is implemented in three steps described in the original paper
- The approach is available in EViews 7

Engle and Yoo (1991)

- The Engle and Yoo (1991) estimator starts by formulating the error-correction version of (15) by adding and subtracting $y_{1,t\text{-}1}$ and adding and subtracting $\beta y_{2,t\text{-}1}$ from the right-hand side and rearranging to yield:

$$\Delta y_{1t} = -(y_{1,t\text{-}1} - \beta y_{2t}) + (\beta + \delta)\Delta y_{2t} + \eta_t \qquad (20)$$

- Given an estimate $\hat{\beta}$, a reduced-form version of (20) is:

$$\Delta y_{1t} = -\rho(y_{1,t-1} - \hat{\beta} y_{2,t-1}) + \alpha \Delta y_{2t} + w_t \qquad (21)$$

where:

$$w_t = \alpha\rho y_{2,t\text{-}1} + \eta_t \quad \alpha = \beta - \hat{\beta}$$

- The Engle and Yoo (1991) estimator is implemented in three steps described in the original paper

Cointegration with more than two variables (1)

- Return to our *n*-dimensional vector of potentially endogenous time-series variables, $Y_t = (y_{1t}, y_{2t}, \dots , y_{nt})'$
- Assume that all variables in y_t are integrated of the same order, and that this order of integration is either I(0) or at most I(1)
- We frequently model the vector Y_t by assuming the following (unrestricted) VAR(p) relationship in levels:

$$Y_t = A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \Phi D_t + \varepsilon_t \qquad \varepsilon_t \sim \mathrm{N}(0,\Sigma) \qquad (22)$$

  where D_t contains deterministic terms
- VARs constitute a useful class of model for modelling:
  - they are a natural generalisation of the AR model as a method for representing multiple time series; and
  - VARs are often used as a means of imposing minimal restrictions (cointegration being one of those) whilst saying something useful

Cointegration with more than two variables (2)

- A **vector error-correction mechanism** (VECM) is a more appropriate model as it distinguishes between stationary variables with transitory (temporary) effects and nonstationary variables with persistent (permanent) effects
- Based on the Granger representation theorem, we can re-parameterise the VAR(p) in (22) as a VECM(p-1):

$$Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + \Phi D_t + \varepsilon_t \qquad (23)$$

- We can re-write the above as a VECM (see the next two slides):

$$\Delta Y_t = \Pi Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + \Gamma_2 \Delta Y_{t-2} + \ldots + \Gamma_p \Delta Y_{t-p+1} + \Phi D_t + \varepsilon_t \qquad (24)$$

where:

$$\Pi = -\mathrm{I}_n + A_1 + \ldots + A_p$$

$$\Gamma_i = -(A_{i+1} + \ldots + A_p),\ i = 1, \ldots, p - 1$$

From a VAR(*p*) to a VECM(*p*-1) (1)

- Start by subtracting Y_{t-1} from both sides of the VAR:

$$\Delta Y_t = -(\mathrm{I}_n - A_1)Y_{t-1} + A_2 Y_{t-2} + \ldots + A_p Y_{t-p} + \Phi D_t + \varepsilon_t \qquad (23')$$

- Add and subtract $A_2 Y_{t-1}$ on the right-hand side:

$$\Delta y_t = -(\mathrm{I}_n - A_1 - A_2)Y_{t-1} - A_2 \Delta Y_{t-1} + A_3 Y_{t-3} + \ldots + A_p Y_{t-p} + \Phi D_t + \varepsilon_t \qquad (23'')$$

- Now add and subtract $A_3 Y_{t-1}$ as well as $A_3 Y_{t-2}$ from the right-hand side:

$$\Delta Y_t = -(\mathrm{I}_n - A_1 - A_2 - A_3)Y_{t-1} - (A_2 + A_3)\Delta Y_{t-1} + A_3 \Delta Y_{\mathrm{t-2}} + A_4 Y_{t-4} + \ldots + A_p Y_{t-p} + \Phi D_t + \varepsilon_t \qquad (23'')$$

- In the next step, we add and subtract $A_4 Y_{t-1}$, $A_4 Y_{t-2}$ and $A_4 Y_{t-3}$ from the right-hand side

From a VAR(*p*) to a VECM(*p*-1) (2)

- By repeating this procedure until *p*-1, we end up with the following specification:

$$\Delta Y_t = \Pi Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + \Gamma_2 \Delta Y_{t-2} + \ldots + \Gamma_p \Delta Y_{t-p+1} + \Phi D_t + \varepsilon_t \quad (24)$$

$$= \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \Phi D_t + \varepsilon_t \quad (24')$$

where:

$$\Pi = -I_n + \sum_{j=1}^{p} A_j$$

$$\Gamma_i = -\sum_{j=i+1}^{p} A_j$$

are the **long-run impact matrix** and the **short-run impact matrices** respectively

A bivariate cointegrated VAR(2) model (1)

- Consider the bivariate VAR(2) model for $Y_t = (y_{1t}, y_{2t})'$ without an explicit term for deterministic variables:

$$Y_t = A_1 Y_{t\text{-}1} + A_2 Y_{t\text{-}2} + \varepsilon_t \qquad (25)$$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} a_{11,1} & a_{12,1} \\ a_{21,1} & a_{22,1} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} a_{11,2} & a_{12,2} \\ a_{21,2} & a_{22,2} \end{pmatrix} \begin{pmatrix} y_{1,t-2} \\ y_{2,t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

- The VECM is:

$$\Delta Y_t = (-\mathrm{I}_2 + A_1 + A_2)Y_{t\text{-}1} - A_2 \Delta Y_{t\text{-}1} + \varepsilon_\mathrm{t} \qquad (26)$$

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} -1 + a_{11,1} + a_{11,2} & a_{12,1} + a_{12,2} \\ a_{21,1} + a_{21,2} & -1 + a_{22,1} + a_{22,2} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} - \begin{pmatrix} a_{11,2} & a_{12,2} \\ a_{21,2} & a_{22,2} \end{pmatrix} \begin{pmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

A bivariate cointegrated VAR(2) model (2)

- An equivalent representation of the VECM is given by:

$$\Delta Y_t = \Pi Y_{t\text{-}1} + \Gamma_1 \Delta Y_{t\text{-}1} + \varepsilon_\mathrm{t} \tag{27}$$

where $\Pi = -\mathrm{I}_2 + A_1 + A_2$ and $\Gamma_1 = -A_2$

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

A bivariate cointegrated VAR(2) model (3)

- Assuming Y_t is cointegrated, there exists a (2×1) cointegrating vector $\beta = (\beta_1, \beta_2)'$ such that $\beta' Y_t = \beta_1 y_{1t} + \beta_2 y_{2t}$ is I(0)
- Using the normalisation $\beta_1 = 1$ and $\beta_2 = -\beta_2$ the cointegration relation becomes $\beta' Y_t = y_{1t} - \beta_2 y_{2t}$
- This normalisation suggests the stochastic long-run equilibrium relation:

$$y_{1t} = \beta_2 y_{2t} + \varepsilon_t$$

where ε_t is I(0) and represents the stochastic deviations from the long-run equilibrium $y_{1t} = \beta_2 y_{2t}$

A bivariate cointegrated VAR(2) model (4)

- Imposing the cointegration restriction, we can re-write the VECM equation-by-equation:

$$\Delta y_{1t} = \alpha_{11}(y_{1,t\text{-}1} - \beta_2 y_{2,t\text{-}1}) + \gamma_{11}\Delta y_{1,t\text{-}1} + \gamma_{12}\Delta y_{2,t\text{-}1} + \varepsilon_{1t} \qquad (28)$$

$$\Delta y_{2t} = \alpha_{21}(y_{1,t\text{-}1} - \beta_2 y_{2,t\text{-}1}) + \gamma_{21}\Delta y_{1,t\text{-}1} + \gamma_{22}\Delta y_{2,t\text{-}1} + \varepsilon_{2t} \qquad (29)$$

- The first equation relates the change in y_{1t} to the lagged disequilibrium error ($y_{1,t\text{-}1} - \beta_2 y_{2,t\text{-}1}$) and the second equation relates the change in y_{2t} to the lagged disequilibrium error as well
- Notice that the reactions of y_{1t} and y_{2t} to the disequilibrium errors are captured by the adjustment coefficients α_{11} and α_{21}
- In matrix form, (28) and (29) correspond to:

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & -\alpha_{11}\beta_2 \\ \alpha_{21} & -\alpha_{21}\beta_2 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

A bivariate cointegrated VAR(2) model (5)

- Comparing this to (27), we find that:

$$\Pi = \begin{bmatrix} \alpha_{11} & -\alpha_{11}\beta_2 \\ \alpha_{21} & -\alpha_{21}\beta_2 \end{bmatrix}$$

which is of rank 1, since a multiple of the first column yields the second column

- Since the matrix Π has reduced rank (rank(Π) = 1), Π can be decomposed as the outer product of a weighting vector, α, and the cointegrating vector, β, where α and β are (2×1) matrices with rank 1 (i.e., vectors)
- The cointegrated VECM for ΔY_t may therefore be re-written as:

$$\Delta Y_t = \alpha\beta' Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + \varepsilon_t \qquad (30)$$

A bivariate cointegrated VAR(2) model (6)

- Thus:

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} = \alpha\beta' = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} \begin{pmatrix} 1 & -\beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_{11} & -\alpha_{11}\beta_2 \\ \alpha_{21} & -\alpha_{21}\beta_2 \end{pmatrix} \quad (31)$$

- β is a vector of cointegrating parameters, so that $\beta' Y_{t-1}$ denotes the I(0) cointegrating relationship or error-correction mechanism
- α is the vector of adjustment or loading coefficients (weights) with which the cointegrating vector enters the two equations of the VAR – the elements of α measure how the deviations from equilibrium feed back on the changes ΔY_t

A bivariate cointegrated VAR(2) model (7)

- Equation (31) can now be re-written as:

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} \begin{pmatrix} 1 & -\beta_2 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} - \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} (y_{1,t-1} - \beta_2 y_{2,t-1}) - \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} ecm_{t-1} - \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \qquad (32)$$

where $ecm_{t-1} = y_{1,t-1} - \beta_2 y_{2,t-1}$ is called the **error-correction term** and (32) is the vector-error correction model (VECM)

- A VAR in first differences would be misspecified as it omits the error-correction term

The Johansen methodology (1)

- We established that the long-run properties of the system are described by the matrix Π
- The methodology which originated this type of analysis and is the one most frequently used is due to Johansen (1988, 1991)
- The methodology has four basic steps:
  - specify and estimate a VAR(*p*) model for Y_t;
  - construct likelihood ratio tests for the rank of Π to determine the number of cointegrating vectors;
  - if necessary, impose normalisation and identifying restrictions on the cointegrating vectors; and
  - given the normalised cointegrating vectors, estimate the resulting cointegrated VECM by maximum likelihood
- Details can be found in Johansen (1995)

The Johansen methodology (2)

1. Rank(Π) = n

   Π is of full rank, which means that all variables in Y_t are stationary (in other words, our initial assumption that all variables in $Y_t \sim I(1)$ cannot be valid); there is no problem of a spurious regression; any linear combination of Y_t is stationary and the appropriate modelling strategy is to estimate a normal VAR in levels:

$$Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + \Phi D_t + \varepsilon_t$$

2. Rank(Π) = 0

   This implies that Π = 0 which means that the term ΠY_{t-1} in (15) disappears and there is no cointegration at all, implying that there are no linear combinations of the Y_t that are stationary; the VECM reduces to a VAR(p-1) and we estimate a normal VAR(p-1) in first differences:

$$\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \ldots + \Gamma_{p-1} \Delta Y_{t-p+1} + \Phi D_t + \varepsilon_t$$

The Johansen methodology (3)

3. Rank(Π) = 0 < r < n

The system is nonstationary, but Y_t is I(1) with r cointegrating relationships (equivalently, $n - r$ common stochastic trends) among the variables in Y_t that are stationary (r rows are linearly independent, thus r linearly independent combinations of Y_t are stationary)

Since Π has rank r it can be written as $\Pi = \alpha\beta'$, where α is an $(n \times r)$ matrix of weights, showing the amount of changes in the variables required to bring the system back to equilibrium, and β is an $(n \times r)$ matrix of parameters determining the long-run equilibrium or cointegrating relationships between the levels of the variables in Y_t

A VAR in levels would be consistent but inefficient (as it ignores the cointegrating relations) and the VAR in differences is misspecified (the lagged cointegrating vector will be a missing regressor)

Hypothesis testing: overview

- Inference is straightforward in the Johansen procedure
- We are primarily interested in testing hypotheses on:
  - the number of cointegrating relationships, r (this is the **most** important hypothesis as any other depends on the number of cointegrating relationships)
- Furthermore, given r, it is easy to formulate simultaneous linear restrictions on α and β:
  - the α matrix of adjustment coefficients; and
  - the β matrix of long-run, or cointegrating, relationships
- As such, we would like to test for linear hypotheses on cointegrating relations and for unique cointegrating vectors

The number of cointegrating vectors, *r*

- Since the rank of the long-run impact matrix Π gives the number of cointegrating relationships in Y_t, Johansen formulates likelihood ratio (LR) statistics for the number of cointegrating relationships, i.e., for determining the rank of Π
- These **sequential** tests are based on the estimated eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_n$ of the matrix Π
- The Johansen procedure is based on the fact that the rank of the matrix Π is equal to the number of eigenvalues (significantly) different from zero
- If rank(Π) = r, the λ_i should be approximately zero for $i > r$

Hypothesis testing: testing the order of *r* (1)

- A test of the null hypothesis of **at most** *n* cointegrating vectors:

  $$H_0(r_0): r = r_0 \quad \text{vs.} \quad H_1(r_0): r > r_0$$

  can be based on the **trace statistic**:

  $$LR_{trace}(r_0 \mid n) = -T \sum_{i=r_0+1}^{n} \log(1 - \hat{\lambda}_i) = 0 \qquad (33)$$

Hypothesis testing: testing the order of *r* (2)

- A test of the null hypothesis of r_0 against the alternative of r_0+1 cointegrating vectors:

  $$H_0(r_0): r = r_0 \quad \text{vs.} \quad H_1(r_0): r > r_0 + 1$$

  can alternatively be based on the **maximum eigenvalue statistic**:

  $$LR_{\max}(r_0, r_0 + 1) = -T \log(1 - \hat{\lambda}_{r_0+1}) = 0 \qquad (34)$$

  where the eigenvalues λ_i, $i = 1, 2, \ldots, n$ of the Π matrix are ordered from largest to smallest

Hypothesis testing: testing the order of *r* (3)

- The cointegrating rank specified in the first null hypothesis which **cannot be rejected** is then chosen as the estimate for the true cointegrating rank, *r*
- Note that:
  - if $H_0(0)$, the first null hypothesis in this sequence, cannot be rejected, a VAR process in first differences is considered; and
  - if all the null hypotheses can be rejected including $H_0(n\text{-}1)$, the process is treated as I(0) and a VAR model in levels is specified
- Reinsel and Ahn (1992) and Reimers (1992) have suggested that the likelihood-ratio tests perform better in small samples if the factor $(T - np)$ is used instead of T in the construction of the LR tests (where n is the number of variables in Y_t and p is the number of lags in the VAR)

Deterministic terms in the model (1)

- Johansen (1994, 1995) derives estimators for models with a number of different specifications concerning both the underlying VAR and the cointegrating relationships
- For example, there can be trends in the VAR (as well as other exogenous variables) and constants and trends in the cointegrating relationship, i.e.:

$$\Delta Y_t = \alpha(\beta' Y_{t-1} - \beta_0 - \beta_1 t) - \gamma_0 - \gamma_1 t + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + \varepsilon_t \qquad (35)$$

- Each model requires special treatment, with different significance levels for the cointegration tests (Johansen (1994, 1995))

Deterministic terms in the model (2)

- Each model requires special treatment, with different significance levels for the cointegration tests (Johansen (1994, 1995))
- Following Johansen (1995), the deterministic terms in (15) are restricted to the form:

$$\Phi D_t = \mu_t = \mu_0 + \mu_1 t \qquad (36)$$

Deterministic terms in the model (3)

- The trend behaviour of Y_t can be classified into **five** cases:
  - no constant or trend ($\mu_t = 0$) and the restricted VECM is:

$$\Delta Y_t = \alpha\beta' Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + ... + \Gamma_{p-1}\Delta Y_{t-p+1} + \varepsilon_t$$

  - restricted constant ($\mu_t = \mu_0 = \alpha\rho_0$) and the restricted VECM is:

$$\Delta Y_t = \alpha(\beta' Y_{t-1} + \rho_0) + \Gamma_1 \Delta Y_{t-1} + ... + \Gamma_{p-1}\Delta Y_{t-p+1} + \varepsilon_t$$

  - unrestricted constant ($\mu_t = \mu_0$) and the restricted VECM is:

$$\Delta Y_t = \mu_0 + \alpha\beta' Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + ... + \Gamma_{p-1}\Delta Y_{t-p+1} + \varepsilon_t$$

  - restricted trend ($\mu_t = \mu_0 + \alpha\rho_1 t$) and the restricted VECM is:

$$\Delta Y_t = \mu_0 + \alpha(\beta' Y_{t-1} + \rho_1 t) + \Gamma_1 \Delta Y_{t-1} + ... + \Gamma_{p-1}\Delta Y_{t-p+1} + \varepsilon_t$$

  - unrestricted constant and trend ($\mu_t = \mu_0 + \mu_1 t$) and the unrestricted VECM is:

$$\Delta Y_t = \mu_0 + \mu_1 t + \alpha\beta' Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + ... + \Gamma_{p-1}\Delta Y_{t-p+1} + \varepsilon_t$$

The Johansen methodology: example (1)

- Assume a VECM(1) with three variables $Y_t = (y_{1t}, y_{2t}, y_{3t})'$ (n = 3) and two cointegrating vectors (r = 2):

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \Gamma_1 \begin{bmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

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

$$= \begin{bmatrix} \alpha_{11} ecm_{1,t-1} + \alpha_{12} ecm_{2,t-1} \\ \alpha_{21} ecm_{1,t-1} + \alpha_{22} ecm_{2,t-1} \\ \alpha_{31} ecm_{1,t-1} + \alpha_{32} ecm_{2,t-1} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{bmatrix} \begin{bmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \qquad (37)$$

The Johansen methodology: example (2)

- The two cointegrating vectors, ecm_{1t} and ecm_{2t}, are given by:

$$ecm_{1,t-1} = \beta_{11}y_{1,t-1} + \beta_{21}y_{2,t-1} + \beta_{31}y_{3,t-1} \quad (38)$$

and:

$$ecm_{2,t-1} = \beta_{12}y_{1,t-1} + \beta_{22}y_{2,t-1} + \beta_{32}y_{3,t-1} \quad (39)$$

or, after normalisation:

$$ecm_{1,t-1} = y_{1,t-1} + \beta_{21}/\beta_{11}y_{2,t-1} + \beta_{31}/\beta_{11}y_{3,t-1} \quad (38')$$

and:

$$ecm_{2,t-1} = \beta_{12}/\beta_{22}y_{1,t-1} + y_{2,t-1} + \beta_{32}/\beta_{22}y_{3,t-1} \quad (39')$$

The Johansen methodology: example (3)

- As we can see from (37), the ECM terms $\beta' Y_{t-1}$ may enter more than one equation:

$$\begin{bmatrix} \alpha_{11} ecm_{1,t-1} + \alpha_{12} ecm_{2,t-1} \\ \alpha_{21} ecm_{1,t-1} + \alpha_{22} ecm_{2,t-1} \\ \alpha_{31} ecm_{1,t-1} + \alpha_{32} ecm_{2,t-1} \end{bmatrix}$$

The Johansen methodology: example (4)

- Under the assumption of **two** cointegrating vectors ($r = 2$), estimation of a single-equation ECM, say, the first one with Δy_{1t} on the left-hand side, yields:

$$\Delta y_{1t} = (\alpha_{11} ecm_{1,t-1} + \alpha_{12} ecm_{2,t-1}) + \gamma_{11}\Delta y_{1,t-1} + \gamma_{12}\Delta y_{2,t-1} + \gamma_{13}\Delta y_{3,t-1} + \varepsilon_{1t}$$

- Thus, by estimating a single-equation ECM with Δy_{1t} on the left-hand side we would not be able to recover the two cointegrating vectors; instead, we get a linear combination of the two cointegrating vectors...
- ...which is again a cointegrating vector
- This result applies whichever element of $Y_t = (y_{1t}, y_{2t}, y_{3t})'$ is used as the left-hand side variable in the single-equation model

The Johansen methodology: example (5)

- On the other hand, when there really is only one cointegrating relationship ($\beta_{11}y_{1,t\text{-}1} + \beta_{21}y_{2,t\text{-}1} + \beta_{31}y_{3,t\text{-}1}$) rather than two entering into all three ECMs with differing speeds of adjustment (α_{11}, α_{21}, α_{31})′, then using the single-equation approach will obtain an estimate of the single cointegrating vector
- Writing out just the error-correction part of, say, the first equation, gives:

$$\alpha_{11}(\beta_{11}y_{1,t\text{-}1} + \beta_{21}y_{2,t\text{-}1} + \beta_{31}y_{3,t\text{-}1}) \quad (40)$$

- But even if there is only one cointegrating vector, there is information to be gained from estimating the other two equations in the system, since α_{21} and α_{31} are not zero – more efficient estimates of the cointegrating vector β can be obtained by using all the information the model has to offer (Johansen (1992))

The Johansen methodology: example (6)

- Only when, say, $\alpha_{21} = \alpha_{31} = 0$, will a single-equation estimator of the unique cointegrating vector be efficient
- Then the cointegration relationship does not enter the other two equations (i.e., Δy_{2t} and Δy_{3t} do not depend on the disequilibrium changes represented by $(\beta_{11}y_{1,t\text{-}1} + \beta_{21}y_{2,t\text{-}1} + \beta_{31}y_{3,t\text{-}1})$)
- This means that when estimating the parameters of the model (i.e., Γ_1, Π, α, β) there is no loss of information from **not** modelling the determinants of Δy_{2t} and Δy_{3t}; so, these variables can enter on the right-hand side of a single-equation ECM
- This therefore also means that $\alpha_{21} = \alpha_{31} = 0$ amounts to y_{2t} and y_{3t} being **weakly exogenous**

Weak exogeneity (1)

- Engle *et al.* (1983) refer to a variable as exogenous **with respect to a particular parameter** if knowledge of the process generating the exogenous variable contains no information about that parameter
- We are interested in **weak exogeneity**, which is required for efficient inference (i.e., estimation and hypothesis testing)
- The loading coefficients in the matrix α can be used to assess whether the cointegrating relations enter a specific equation significantly – thereby allowing us to model partial systems with exogenous variables
- Weak, strong and super exogeneity are the relevant concepts for estimation, forecasting and policy analysis respectively (Ericsson *et al.* (1998))

Weak exogeneity (2)

- If there are $r \leq (n - 1)$ cointegrating vectors in β, then this implies that the last $(n - r)$ columns of α are zero
- Thus the typical problem faced, that of determining how many $r \leq (n - 1)$ cointegrating vectors exist in β, amounts to equivalently testing which columns of α are zero
- Another natural question to ask about the adjustment matrix is whether the coefficients in α are zero for a certain subset of equations – these hypotheses can be formulated as linear restrictions on the rows of α
- This hypothesis means that the subset of variables is **weakly exogenous** for the long-run parameters and the remaining adjustment parameters, i.e., disequilibrium in the cointegrating relationship, does not feed back directly onto the corresponding variables

Testing weak exogeneity (1)

- Suppose that $Y_t = (y_{1t}, y_{2t}, y_{3t})'$ ($n = 3$) and $r = 1$, which means that $\alpha = (\alpha_{11}, \alpha_{21}, \alpha_{31})'$:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \Gamma_1 \begin{bmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix}$$

- The first term in α represents the speed at which Δy_{1t}, the dependent variable in the first equation of the VECM, adjusts towards the single long-run cointegrating relationship ($\beta_{11}y_{1,t\text{-}1} + \beta_{21}y_{2,t\text{-}1} + \beta_{31}y_{3,t\text{-}1}$), while α_{21} represents the speed at which Δy_{2t} adjusts and α_{31} shows how fast Δy_{3t} responds to the disequilibrium changes represented by the cointegrating vector

Testing weak exogeneity (2)

- More generally, each of the r non-zero columns of α contain information on which cointegrating vector enters which short-run equation and on the speed of adjustment
- Taking things a step further, the presence of **all** zeros in row i of α_{ij}, $j = 1, \dots , r$, indicates that the cointegrating vectors in β do not enter the equation determining Δy_{it}
- A variable is weakly exogenous for the cointegrating parameters if none of the cointegration relations enter the equation for that variable
- s linear restrictions on the loading matrix α of the type $\alpha = G\psi$, with G a given fitted $n \times s$ matrix and ψ an $s \times r$ parameter matrix with $s \geq r$, can be easily imposed

Testing weak exogeneity (3)

- For instance, in our three-variable system (n = 3) with two cointegration relations (r = 2), we may wish to consider the hypothesis that the third variable, y_{3t}, is weakly exogenous:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \Gamma_1 \begin{bmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix} \quad (41)$$

- Testing the weak exogeneity of y_{3t} amounts to testing $\alpha_{31} = \alpha_{32} = 0$:

$$\alpha = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = G\psi \quad (42)$$

Testing weak exogeneity (4)

- If the restriction cannot be rejected:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \Gamma_1 \begin{bmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{11} ecm_{1,t-1} + \alpha_{12} ecm_{2,t-1} \\ \alpha_{21} ecm_{1,t-1} + \alpha_{22} ecm_{2,t-1} \\ 0 \end{bmatrix} + \Gamma_1 \begin{bmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix} \quad (43)$$

Testing weak exogeneity (5)

- If $\alpha_{3j} = 0$ ($j = 1, 2$), then the equation for Δy_{3t} contains no information about the long-run β since the cointegration relationships do not enter into this equation, and it is therefore valid to condition on the weakly exogenous variable y_{3t} and proceed with the following partial version of the VECM:

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{pmatrix} + \Gamma_1 \begin{pmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{pmatrix} + \begin{pmatrix} \gamma_{10} \\ \gamma_{20} \end{pmatrix} \Delta y_{3t} + \varepsilon_t$$

- Note that the weakly exogenous variable, y_{3t}, remains in the long-run model (i.e., the cointegrating vectors) although its short-run behaviour is not modelled because of its exclusion from the vector on the left-hand side of the equation

Summary (1)

- Cointegrating relationships resemble long-run relationships and can be used to **validate** economic hypotheses – when carefully applied, it allows for the analysis of long-run economic relationships
- It is therefore important to be able to test for the existence of cointegration and the number of cointegrating vectors
- Depending on the number of variables, it is important to figure out whether this is done in an ECM or VECM
- Except in the case when only two variables are involved, the analysis of cointegration should really begin by using a multivariate framework and not by using a single-equation approach

Summary (2)

- Modified single-equation approaches can help to improve inference in cointegrating regressions…
- …although the limitation of these approaches remains that the dimension of the cointegration space is always limited to unity
- The Johansen methodology is only one of several for dealing with more than one cointegrating relationship, but it is the most natural extension to the simple VAR – as such, it has become an essential tool for applied economists who wish to estimate time-series models
- It is a very flexible tool for modelling systems where we expect more than one cointegrating vector

Summary (3)

- It can be used as a tool for both estimating the number of cointegrating relationships and the relationships themselves
- But note that it is quite a ‘fragile’ method and needs to be used with care

References and further reading (1)

Dolado, J J, Jenkinson, T and Sosvilla-Rivero, S (1990), 'Cointegration and unit roots', *Journal of Economic Surveys*, Vol. 4. No. 3, pages 249-73.

Engle, R F and Granger, C W J (1987), 'Co-integration and error correction: representation, estimation and testing', *Econometrica*, Vol. 55, No. 2, pages 251-76.

Engle, R F and Granger, C W J (1991), *Long-run economic relationships: readings in cointegration*, Oxford, Oxford University Press.

Engle, R F, Hendry, D F and Richard, J-F (1983), 'Exogeneity', *Econometrica*, Vol. 51, No. 2, pages 277-304.

References and further reading (2)

Engle, R F and Yoo, B S (1987), 'Forecasting and testing in cointegrated systems', *Journal of Econometrics*, Vol. 35, No. 1, pages 143-59.

Engle, R F and Yoo, B S (1991), 'Cointegrated economic time series: an overview with new results', Chapter 12 in Engle, R F and Granger, C W J (eds), *Long-run economic relationships: readings in cointegration*, Oxford, Oxford University Press, pages 237-66.

Ericsson, N R, Hendry, D F and Mizon, G E (1998), 'Exogeneity, cointegration and economic policy analysis', *Journal of Business & Economic Statistics*, Vol. 16, No. 4, pages 370-87.

References and further reading (3)

Granger, C W J (1981), 'Some properties of time series data and their use in econometric model specification', *Journal of Econometrics*, Vol. 16, No. 1, pages 121-30.

Granger, C W J and Newbold, P (1974), 'Spurious regressions in econometrics', *Journal of Econometrics*, Vol. 2, No. 2, pages 111-20.

Hansen, B E (1992), 'Efficient estimation and testing of cointegrating vectors in the presence of deterministic trends', *Journal of Econometrics*, Vol. 53, No. 1-3, pages 87-121.

Hendry, D F and Juselius, K (2000), 'Explaining cointegration analysis: part II', *The Energy Journal* , Vol. 21, No. 1 ,pages 1-42.

References and further reading (4)

Johansen, S (1988), 'Statistical analysis of cointegrating vectors', *Journal of Economic Dynamics and Control*, Vol. 12, No. 2/3, pages 231-54.

Johansen, S (1991), 'Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models', *Econometrica*, Vol. 59, No. 6, pages 1551-80.

Johansen, S (1992), 'Cointegration in partial systems and the efficiency of single-equation analysis', *Journal of Econometrics*, Vol. 52, No. 3, pages 389-402.

Johansen, S (1994), 'The role of the constant and linear terms in cointegration analysis of nonstationary variables', *Econometric Reviews*, Vol. 13, No. 2, pages 205-29.

References and further reading (5)

Johansen, S (1995), *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford, Oxford University Press.

Juselius, K (1995), 'Do purchasing power parity and uncovered interest rate parity hold in the long-run? An example of likelihood inference in a multivariate time-series model', *Journal of Econometrics*, Vol. 69, No. 1, pages 211-40.

Kremers, J J M, Ericsson, N R and Dolado, J J (1992), 'The power of cointegration tests', *Oxford Bulletin of Economics and Statistics*, Vol. 54, No. 3, pages 325-48.

MacKinnon, J G (1991), 'Critical values for cointegration tests', Chapter 13 in Engle, R F and Granger, C W J (eds), *Long-run economic relationships: readings in cointegration*, Oxford, Oxford University Press, pages 267-76.

References and further reading (6)

Maddala, G S and Kim, I-M (1998), *Unit roots, cointegration and structural change*, Cambridge, Cambridge University Press.

Park, J Y (1992), 'Canonical cointegrating regressions', *Econometrica*, Vol. 60, No. 1, pages 119-43.

Pesaran, M H, Shin, Y and Smith, R J (2001), 'Bounds testing approaches to the analysis of level relationships', *Journal of Applied Econometrics*, Vol. 16, No. 3, pages 289-326.

Phillips, P C B and Durlauf, S N (1986), 'Multiple time series regression with integrated processes', *Review of Economic Studies*, Vol. 53, No. 4, pages 473-95.

Phillips, P C B and Hansen, B E (1990), 'Statistical inference in instrumental variables regression with I(1) variables', *Review of Economic Studies*, Vol. 57, No. 1, pages 99-125.

References and further reading (7)

Phillips, P B C and Loretan, M (1991), ‘Estimating long-run economic equilibria’, *Review of Economic Studies*, Vol. 59, No. 3, pages 407-36.

Reimers, H-E (1992), ‘Comparisons of tests for multivariate cointegration’, *Statistical Papers*, Vol. 33, No. 1, pages 335-59.

Reinsel, G C and Ahn, S K (1992), ‘Vector autoregressive models with unit roots and reduced rank structure: estimation, likelihood ratio test and forecasting’, *Journal of Time Series Analysis*, Vol. 13, No. 4, pages 353-75.

Saikkonen, P (1991), ‘Asymptotically efficient estimation of cointegration regressions’, *Econometric Theory*, Vol. 7, No. 1, pages 1-21.

References and further reading (8)

Stock, J H (1987), ‘Asymptotic properties of least squares estimators of cointegrating vectors’, *Econometrica*, Vol. 55, No. 5, pages 1035-56.

Stock, J H and Watson, M W (1993), ‘A simple estimator of cointegrating vectors in higher order integrated systems’, *Econometrica*, Vol. 61, No. 4, pages 783-820.