Econometrics, Lecture 12A: Cointegration

Ron Smith
EMS, Birkbeck, University of London

Autumn 2020

Last time

- Looked at ARDL and ECM, they will feature this time as well.
- Looked at reparameterizations and restrictions.
- Concepts of reparameterisation and restriction useful beyond time series, e.g. Cobb-Douglas.
- Suggested a general to specific approach to testing but did not discuss the tests
- Did not discuss order of integration much.

Cointegration

- Suppose y_t and x_t are I(1) then in general any linear combination of them, $y_t - \beta x_t = z_t$, will also be I(1).
- If there is a linear combination that is I(0), they are said to cointegrate.
- I(1) variables have a stochastic trend like the random walk with drift: $\Delta y_t = \alpha + \varepsilon_t$

$$y_t = y_0 + \alpha t + \sum_{i=0}^{t-1} \varepsilon_{t-i}$$

- If they cointegrate, they have a common stochastic trend, $\sum \varepsilon_{t-i}$,, which is cancelled out by the linear combination; and this linear combination is called the cointegrating vector, $(1, -\beta)$
- The cointegrating relationship is often interpreted as an equilibrium relationship and z_t a measure of disequilibrium.
- Note I(1) is more general than a random walk, any ARIMA(p,1,q) y_t is I(1). It has to be differenced once to make it stationary.

Purchasing Power Parity, PPP

- Suppose $\mathbf{y}_t = (s_t, p_t, p_t^*)$ are the logs of the spot exchange rate, domestic and foreign price indexes and all are I(1).
- Purchasing Power Parity says that the real exchange rate $z_t = s_t - p_t + p_t^*$ is stationary. So $z_t = e + u_t$ where e is the equilibrium real exchange rate and u_t is a I(0), but not necessarily white noise, disequilibrium term.
- The cointegrating vector is then $\boldsymbol{\beta}' = (1, -1, 1) : \boldsymbol{\beta}'\mathbf{y}_t = z_t$

$$\begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} s_t \\ p_t \\ p_t^* \end{bmatrix} = z_t$$

- It may be that p_t, p_t^* are $I(2)$, but cointegrate to $I(1)$ so $d_t = p_t - p_t^*$ is $I(1)$ and $s_t - d_t$ is $I(0)$.

Great Ratios

- Ratios of non-stationary economic variables may be roughly stationary. These 'great ratios' include the real exchange rate, the savings ratio, the velocity of circulation of money, the capital-output ratio, the share of wages in output, the profit rate, etc.
- Linear combinations of the logs of the variables (weights plus and minus one) should be stationary. This is testable with unit root or cointegration tests.
- In the practicals we looked at the log of the payout ratio: dividends divided by earnings. This gave an ADF statistic of -5.7 with a critical value of -2.88, so we would clearly reject the unit root hypothesis and conclude that log payout ratio was I(0).

Cointegrating vector

- If for $I(1)$ y_t and x_t in

$$y_t = \alpha + \beta x_t + u_t \tag{1}$$

  u_t is stationary, the cointegrating vector is $(1, -\beta)$ since $(y_t - \beta x_t = \alpha + u_t)$ is I(0).
- If y_t and x_t are I(1) and do not cointegrate, say they are independent unrelated random walks, $\Delta y_t = \varepsilon_{1t}$, $\Delta x_t = \varepsilon_{2t}$, $E(\varepsilon_{1t}\varepsilon_{2t}) = 0$, u_t in (1) will be I(1) and this will be a 'spurious' regression.
- As $T \to \infty$, the R^2 of this regression will go to unity and the t ratio for $\widehat{\beta}$ will go to a non-zero random variable. Thus even if there is no relationship, the regression would indicate a close relationship. Therefore it is important to test for cointegration,
- This can be better done in the context of an ECM which nests $\Delta y_t = \varepsilon_{1t}$, which a levels relationship like (1) does not.

Consistency (revision)

- $\widehat{\theta}$ is consistent if $P\lim(\widehat{\theta}) = \theta$, if for any $\epsilon > 0$

$$\lim_{T \to \infty} \Pr(\mid \widehat{\theta}_T - \theta \mid > \epsilon) = 0.$$

- As T gets large $\widehat{\theta}_T$ converges to θ, they become the same, both the bias and the variance $\rightarrow 0$ so we cannot compare the variances of consistent estimators, all have variance zero.
- $\overline{X}$ has variance σ_x^2/T which $\rightarrow 0$ as $T \rightarrow \infty$. But so does the variance of $\widetilde{X}$ with $V(\widetilde{X}) = 2\sigma_x^2/T$.
- To deal with this problem we scale the difference and look at $\sqrt{T}(\widehat{\theta} - \theta)$ as $T \rightarrow \infty$, so the asymptotic variance of the mean is $AV(\overline{X}) = \sigma_x^2$ not $\sigma_x^2/T = 0$.
- When the variance falls at rate T then $\mid \widehat{\theta}_T - \theta \mid \rightarrow 0$ at rate $\sqrt{T}$ and $\widehat{\theta}$ is said to be $\sqrt{T}$ consistent.

Consistent estimators in the LRM

- In the LRM using deviations from means,

$$y_t = \beta x_t + u_t$$

with $E(u_t^2) = \sigma_u^2$, $E(u_t u_{t-i}) = 0$; $Var(x_t) = E(x_t^2) = \sigma_x^2$.

- The OLS estimator $\widehat{\beta} = \sum x_t y_t / \sum x_t^2$ has variance

$$V\left(\widehat{\beta}\right) = E(\widehat{\beta} - \beta)^2 = E\left(\frac{\sum x_t u_t}{\sum x_t^2}\right)^2 = \frac{\sigma_u^2}{\sum x_t^2} = \frac{\sigma_u^2}{T\sigma_x^2}$$

- So $\widehat{\beta}$ is also $\sqrt{T}$ consistent.
- If each individual x_t has variance $E(x_t^2) = \sigma_x^2$, then the variance of the sum is

$$E(\sum_{t=1}^{T} x_t^2) = T\sigma_x^2.$$

Super-consistency

- Suppose x_t is a random walk, $x_t = x_{t-1} + \varepsilon_t$ with $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$ where $E(x_0) = 0$, so $x_t = x_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}$.
- The variance of $x_T : \sigma_x^2 = E(x_T^2) = T\sigma_\varepsilon^2$
- Then the variance of $\widehat{\beta}$ is

$$V\left(\widehat{\beta}\right) = \frac{\sigma_u^2}{\sum x_t^2} = \frac{\sigma_u^2}{T\sigma_x^2} = \frac{\sigma_u^2}{T(T\sigma_\varepsilon^2)} = \frac{\sigma_u^2}{T^2\sigma_\varepsilon^2}.$$

- So when x_t is a random walk (or more generally $I(1)$) $\widehat{\beta}$ converges to its true value at rate T rather than $\sqrt{T}$, it is said to be T consistent or super-consistent.
- Of course, even if it converges rapidly to its true value with T, it may have a large bias in small samples.
- In a regression on a time trend $y_t = a + bt + u_t$, $\widehat{b}$ is $T^{3/2}$ consistent, since $\sum t^2$ is in the denominator of variance.

Testing for cointegration

- Just as there are many tests for unit roots, there are many tests for the existence of a singe cointegrating vector that differ as to such things as
  - the null: cointegration or no contegration
  - the treatment of serial correlation: parameteric or non-parametric
- They include the Engle-Granger procedure, fully modified OLS, dynamic OLS, canonical cointegrating relations.
- ARDL/ECM is easy, effective and flexible. Does not require knowing order of integration/cointegration.

Cointegration and the ARDL/ECM

- Do not estimate time series equations in levels. Add lags. ARDL/ECM handles I(0), I(1) or cointegrated variables.
- ECM

$$
\begin{aligned}
\Delta y_t &= a_0 + b_0 \Delta x + a_1 y_{t-1t} + b_1 x_{t-1} + u_t \\
\Delta y_t &= a_0 + b_0 \Delta x + \lambda (y_{t-1} - \theta_x x_{t-1}) + u_t
\end{aligned}
$$

  has LHS Δy_t I(0) and RHS 2 $I(1)$ terms y_{t-1} and x_{t-1}.
- It only balances if y_t and x_t cointegrate to $I(0)$; $(y_t - \theta_x x_t)$ is I(0) with cointegrating vector, CV, $(1, -\theta_x)$ or equivalently $a_1 y_{t-1t} + b_1 x_{t-1}$ is $I(0)$ with CV (a_1, b_1).
- Normalisation does not matter.

Cointegration and the ARDL/ECM

- In
$$\Delta y_t = a_0 + b_0 \Delta x + \lambda (y_{t-1} - \theta_x x_{t-1}) + u_t$$
- If y_t and x_t are I(1) and cointegrate λ must be non-zero and negative: this is the feedback that keeps y_t and x_t from diverging.
- The estimate of the long run parameter $\widehat{\theta}_x$ is T consistent, the estimates of the short run parameters $\widehat{a}_0$, $\widehat{b}_0$, and $\widehat{\lambda}$ are $\sqrt{T}$ consistent.

Unknown order of integration

- To examine more cases, consider the ECM for 2 exogenous variables

$$\Delta y_t = a_0 + a_1 y_{t-1} + b_0 \Delta x_t + b_1 x_{t-1} + c_0 \Delta z_t + c_1 z_{t-1} + \varepsilon_t.$$

- The long run relationship, LRR, exists if $a_1 \neq 0$, and is

$$\begin{aligned} y_t^* &= -\frac{a_0}{a_1} - \frac{b_1}{a_1} x_t - \frac{c_1}{a_1} z_t \\ y_t^* &= \theta_0 + \theta_x x_t + \theta_z z_t. \end{aligned}$$

- The ECM is robust in that OLS provides consistent estimates for a range of orders of integration.
- The concern with spurious regression arose from the, then common, practice of estimating levels relationships without dynamics, which could not capture the data generating process.

Various cases

With dynamics the ECM can nest and consistently estimate a number of special cases, which will all balance, in the sense that the two sides are the same order of integration.

$$\Delta y_t = a_0 + a_1 y_{t-1} + b_0 \Delta x_t + b_1 x_{t-1} + c_0 \Delta z_t + c_1 z_{t-1} + \varepsilon_t.$$

- If all the variables are $I(0)$, ECM is a standard LRM and LRR is a standard long-run relationship between $I(0)$ variables, though it is not a cointegrating relationship since the variables are not $I(1)$.
- If all the variables are $I(1)$ and cointegrated, then LRR is the cointegrating relationship.
- If all the variables are $I(1)$ and not cointegrated then $a_1 = b_1 = c_1 = 0$ and there is no long run relationship and OLS will estimate ECM as a first difference equation.

More cases

$$\Delta y_t = a_0 + a_1 y_{t-1} + b_0 \Delta x_t + b_1 x_{t-1} + c_0 \Delta z_t + c_1 z_{t-1} + \varepsilon_t.$$

- If y_t is $I(0)$, x_t and z_t are $I(1)$ and cointegrated such that $(b_1 x_{t-1} + c_1 z_{t-1})$ is $I(0)$ the equation is balanced.
- If y_t is $I(0)$, x_t and z_t are $I(1)$ and $b_1 = c_1 = 0$ the equation is balanced, since Δx_t and Δz_t are $I(0)$.
- If y_t is $I(1)$, x_t and z_t are $I(0)$ and $a_1 = 0$ the equation is balanced, since Δy_t is $I(0)$.
- If y_t is $I(2)$, x_t and z_t are $I(1)$ and $a_1 = 0$ the equation is balanced, since Δy_t is $I(1)$.

In all these cases the OLS estimates can make it balance.

I(2) cases

- With 2 lags similar arguments can be applied to include $I(2)$ cases.
- If the variables are $I(2)$ and cointegrate to $I(1)$ then both sides of the equation below are $I(0)$

$$\Delta^2 y_t = a_0 + b_0 \Delta^2 x_t + \lambda(\Delta y_{t-1} - \theta_x \Delta x_{t-1}) + u_t$$

- This equation is a special case of the ARDL(2,2) with 2 restrictions:

$$\begin{aligned} y_t &= a_0 + (2+\lambda) y_{t-1} - (1+\lambda) y_{t-2} \\ &\quad + b x_t - (2b + \lambda\theta_x) x_{t-1} + (b + \lambda\theta_x) x_{t-2} + u_t . \end{aligned}$$

- Thus OLS on an ARDL(2,2) will be able to capture the relationship.

Testing

- Although the estimation is standard, testing is not. The critical values for testing for a long run relationship are different depending on whether the variables are $I(0)$ or $I(1)$.
- Testing $a_1 = 0$ in

$$\Delta y_t = a_0 + a_1 y_{t-1} + b_0 \Delta x_t + b_1 x_{t-1} + c_0 \Delta z_t + c_1 z_{t-1} + \varepsilon_t.$$

is like a Dickey Fuller test so the test statistic is non standard.

- The PSS Bounds Test, provides $I(0)$ and $I(1)$ non-standard critical values for testing for a long run relationship.
- Pesaran, Shin and R.J. Smith, Journal of Applied Econometrics, 2001, p289-326, Bounds Testing Approaches to the Analysis of Level Relationships.

Bounds test

- 1. Use a t statistic and the non-standard critical values to test $a_1 = 0$. No long run relationship is defined if $a_1 = 0$.
- 2. If $a_1 \neq 0$, calculate the F statistic for the no levels relationship hypothesis: $a_1 = b_1 = c_1 = 0$ in the ECM.
- 3. If the F statistic is below the $I(0)$ critical value, there is no long run relation. If it is above the $I(1)$ critical value, there is a long run relation. If it is in between it depends on the order of integration of the variables.
- The test is for a long run levels relationship which is a more general category than a cointegrating relationship. While there can be long run levels relationships whether the variables are I(0) or I(1), cointegrating relationships only apply if the variables are I(1). If the variables cointegrate there is a long run relationship, but there can also be a long run relationship without cointegration if the variables are I(0).

Do not use the Engle-Granger, EG, procedure:

- Engle-Granger suggested estimating the levels equation

$$y_t = \widehat{\alpha} + \widehat{\beta} x_t + \widehat{u}_t \tag{2}$$

- Testing whether the residuals are I(1), using an ADF test on

$$\begin{aligned} \Delta \widehat{u}_t &= b \widehat{u}_{t-1} + v_t && (3) \\ \Delta y_t - \widehat{\beta} \Delta x_t &= b \left(y_{t-1} - \widehat{\alpha} - \widehat{\beta} x_{t-1} \right) + v_t && (4) \end{aligned}$$

  with appropriate critical values for the t stat on $\widehat{b}$, which are different from those for a variable.
- Although the estimates of $\widehat{\beta}$ are 'super-consistent' (T rather than $\sqrt{T}$), (2) is misspecified by omitting dynamics and the estimates can be badly biased in small samples.
- (3) restricts the short-run dynamics compared to an ECM:

$$\Delta y_t = \alpha_0 + \beta_0 \Delta x_t + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + u_t$$

Multiple cointegrating vectors

- With 2 I(1) variables there can only be one cointegrating vector, CV, but with more variables there can be more CVs and any linear combination of these CVs will also be a CV.
- Suppose that we have data on domestic and foreign interest rates and inflation $(r_t, r_t^*, \Delta p_t, \Delta p_t^*)$ and all are I(1) (this implies that p_t is I(2)).
- If real interest rates $(r_t - \Delta p_t$ and $r_t^* - \Delta p_t^*)$ are I(0) with CVs $(1, 0, -1, 0)$ and $(0, 1, 0, -1)$; then the real interest rate differential $(r_t - \Delta p_t) - (r_t^* - \Delta p_t^*)$ would also be I(0), with CV $(1, -1, -1, 1)$.
- If the CV is known a priori (as with the real exchange rate or real interest rate examples above) we can form the hypothesised I(0) linear combination (the log of the real exchange rate or the real interest rates) and use an ADF test to determine whether it is in fact I(0).

Next time

- Often we do not know how many CVs there are or the coefficients in the multiple unknown CVs.
- There are a variety of procedures for multiple CVs.
- The most commonly used is the Johansen procedure for testing for the number of CVs and estimating them.
- This procedure operates in the context of a Vector autoregression VAR, which we consider first.
- We have assumed that we know which is the exogenous variable. With the Johansen procedure we can test.