More on Cointegration

D.E. Allen

Department of Accounting, Finance and Economics
Edith Cowan University

ECF4116 Econometrics

Outline

In this lecture I shall draw from several sources.
The Section on Cointegration from the GRETL Manual - Chapter 22 Cointegration
David A. Dickey, Dennis W. Jansen and Daniel L. Thornton, “A Primer On Cointegration with an Application to Money and Income”
Juan J. Dolado , Jesús Gonzalo, and Francesc Marmol “Cointegration”.

Cointegration

A substantial part of economic theory generally deals with long-run equilibrium relationships generated by market forces and behavioral rules. Correspondingly, most empirical econometric studies entailing time series can be interpreted as attempts to evaluate such relationships in a dynamic framework.

For some time now, macroeconomists have been aware that many macroeconomic time series are not stationary in their levels and that many time series are most adequately represented by first differences. In the parlance of time-series analysis, such variables are said to be integrated of order one and are denoted 1(1). The level of such variables can become arbitrarily large or small so there is no tendency for them to revert to their mean level. Indeed, neither the mean nor the variance is a meaningful concept for such variables.

Non-stationarity

- Those problems were somehow ignored in applied work until important papers by Granger and Newbold (1974) and Nelson and Plosser (1982) alerted many to the econometric implications of non-stationarity and the dangers of running nonsense or spurious regressions. In particular, most of the attention focussed on the implications of dealing with integrated variables which are a specific class of non-stationary variables with important economic and statistical properties. These are derived from the presence of unit roots which give rise to stochastic trends, as opposed to pure deterministic trends, with innovations to an integrated process being permanent rather than transitory.The presence of, at least, a unit root in economic time series is implied in many economic models. Among them, there are those based on the rational use of available information or the existence of very high

Differencing

- Statisticians, in turn, following the influential approach by Box and Jenkins (1970), had advocated transforming integrated time series into stationary ones by successive differencing of the series before modelization. Therefore, from their viewpoint, removing unit roots through differencing ought to be a pre-requisite for regression analysis. However, some authors, notably Sargan (1964), Hendry and Mizon (1978) and Davidson et al. (1978), inter alia, started to criticized on a number of grounds the specification of dynamic models in terms of differenced variables only, especially because of the difficulties in inferring the long-run equilibrium from the estimated model.

Granger's insight.

- Granger (1981), resting upon the previous ideas, solved the puzzle by pointing out that a vector of variables, all which achieve stationarity after differencing, could have linear combinations which are stationary in levels. Later, Engle and Granger (1987) were the first to formalize the idea of integrated variables sharing an equilibrium relation which turned out to be either stationary or have a lower degree of integration than the original series. They denoted this property by cointegration, signifying co-movements among trending variables which could be exploited to test for the existence of equilibrium relationships within a fully dynamic specification framework..

Cointegration and equilibrium relationships

- the basic concept of cointegration applies in a variety of economic models including the relationships between capital and output, real wages and labor productivity, nominal exchange rates and relative prices, consumption and disposable income, long and short term interest rates, money velocity and interest rates, price of shares and dividends, production and sales, etc. In particular, Campbell and Shiller (1987) have pointed out that a pair of integrated variables that are related through a Present Value Model, as it is often the case in macroeconomics and finance, must be cointegrated.

Unit roots and cointegration

- A well-known result in time series analysis is Wold's (1938) decomposition theorem which states that a stationary time series process, after removal of any deterministic components, has an infinite moving average (MA) representation which, in turn, can be represented by a finite autoregressive moving average (ARMA) process. However, as mentioned in the Introduction, many time series need to be appropriately differenced in order to achieve stationarity. From this comes the definition of integration : a time series is said to be integrated of order d, in short, I(d), if it has a stationary, invertible, non-deterministic ARMA representation after differencing d times.

Cointegration the basic idea

Consider two time series y_{1t} and y_{2t} which are both $I(d)$. In general any linear combination of y_{1t} and y_{2t} will also be $I(d)$. However, if there exists a vector $(1, -\beta)'$ such that the linear combination

$$z_t = y_{1t} - \alpha - \beta y_{2t} \text{ is } I(d-b),\ d \geq b \geq 0, \tag{1}$$

then, following Engle and Granger (1987), y_{1t} and y_{2t} are defined as cointegrated of order (d, b)
denoted $y_t = (y_{1t}, y_{t2})' \sim CI(d, b)$, with $(1, -\beta)'$ called the cointegrating vector.

Cointegration - some features

Some features of eqn (1) are noteworthy.

1. Cointegration refers to a linear combination of non-stationary variables. (Its also feasible that a non-linear combination may be stationary too)
2. The cointegrating vector is not uniquely defined. since for any non zero value of λ, $(\lambda, -\lambda\beta)'$is also a cointegrating vector. Thus, a normalization rule needs to be used. $\lambda = 1$has been chosen in (1).
3. Third, all variables need to be integrated to the same order to be candidates to form a cointegrating relationship.
4. Most of the cointegration literature focusses on the case where the variables contain a single unit root, since few economic variables prove to be integrated to a higher order.

Engle Granger two step procedure was discussed last week

System based approaches to cointegration

In general if $y_t$now represents a vector of $n\,I(1)$variables its Wold representation (assuming no deterministic terms) is given by
$\Delta y_t = C(L)\varepsilon_t,$ (2)
where now $\varepsilon_t \sim nid(0,\sum)$,$\sum$being the covariance matrix of ε_t and $C(L)\,an\,(nxn)$invertible matrix of polynomial lags where the term 'invertible' means that $\mid C(L) = 0 \mid$has all its roots larger than unity in absolute value. If there is a cointegrating vector $(nx1)$vector, $\beta' = (\beta_{11},....,\beta_{nm})$then premultiplying (2) by β'yields

System based approaches to cointegration

$\beta' = \Delta y_t = \beta' \left[C(1) + \tilde{C}(L)\Delta \right] \varepsilon_t$ (3)

where $C(L)$has been expanded around $L = 1$using a first order Taylor series expansion and $C\tilde{(L)}$can be shown to be an invertible lag matrix. Since the cointergartion property implies that $\beta' y_t$is $I(0)$, then it must be the case that $\beta' C(1) = 0$ and hence $\Delta(= 1 - L)$will cancel out on both sides of (3). Then, given that $C(L)$ is invertible, then y_t has a vector autoregressive representation, such that

$A(L)y_t = \varepsilon_t$ (4)

where $A(L)C(L) = \Delta I_n$, I_n being the (nxn)identity matrix. Hence we must have that $A(1)C(1) = 0$, implying that $A(1)$ can be written as a linear combination of the elements of β

System based approaches to cointegration

Johansen (1995) developed an MLE procedure based on the reduced rank regression method. This has some advantages:

1. It relaxes the assumption that the cointegrating vector is unique.
2. It takes into account the short run dynamics of the system when estimating the cointegrating vectors.

The intuition can be explained as follows: Assume that y_t has a $VAR(1)$ representation, that is $A(L)$ in (4) such that $A(L) = I_n - A_1 L$. Hence the $VAR(1)$ process can be reparameterised in the VECM representation as

$$\Delta y_t = (A_1 - I_n) y_{t-1} + \varepsilon_t \qquad (5)$$

if $A_1 - I_n = -A(1) = 0$, then y_t is $I(1)$ and there are no cointegrating relationships (r=0)

Testing the rank of a system

On the otherhand if $rank\,(A_1 - I_n) = n$, there are n cointegrating relationships between the n series and hence $y_t \sim I(0)$. Thus testing the null hypothesis that the number of of cointegrating vectors (r) is equivalent to testing whether $rank\,(A_1 - I_n) = r$. Johansen (1995) deals with the more general case where $y_t$follows a $VAR\,(p)$ process of the form

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \varepsilon_t \qquad (6)$$

which can be written in the ECM representation

$$\Delta y_t = D_1 \Delta y_{t-1} + D_2 \Delta y_{t-2} + ... + \Delta_{p-1} \Delta y_{t-p+1} + D y_{t-1} + \varepsilon_t \qquad (7)$$

where $D_i = -(A_{i+1} + ... + A_p)$, i = 1,2,....,p-1, and $D = (A_1 + ... + A_p - I_n) = -A(1) = -B\Gamma'$.

Johansen estimation

To estimate $B\, and\ \Gamma$we need to estimate Dsubject to some indentification restriction since otherwise $B\, and\ \Gamma$ could not be separately identified. MLE of D goes along the same principles as a partitioned regression model; the regressor and the regressand of interest $(\Delta y_t\, and\, y_{t-1})$ are regressed by OLS on the remaining set of regressors $(\Delta y_{t-1}, ..., \Delta y_{t-p+1})$ giving rise to two matrices of residuals denotes as $\hat{e_0}\, and\, \hat{e_1}$ and the regression model $\hat{e}_{0t} = \hat{D}\hat{e}_{1t} + residuals$. Johansen (1995) shows that testing for the rank of $\hat{D}$ is equivalent to testing for the number of canonical correlations between $\hat{e_0}$and $\hat{e_1}$that are different from zero.

Johansen estimation

This can be done using either of the following two test statistics
$\lambda_{tr}(r) = -T\sum_{i=r+1}^{n} ln(1-\widehat{\lambda}_i)$ (8)
$\lambda_{max}(r, r+1) = -Tln(1-\hat{\lambda}_{r+1}$ (9)
where the $\hat{\lambda}_i$s are the eigenvalues of the matrix $S_{10}S_{00}^{-1}S_{01}$with respect to the matrix S_{11}ordered in decreasing order $(1 \geq \hat{\lambda}_1 > .... > \hat{\lambda}_n > 0)$, where $S_{ij} = T^{-1}\sum_{t=1}^{T} \hat{e}_{it}\hat{e}_{jt}^{'}$,i,j = 0,1. These eigenvalues can be obtained as the solution of the determinantal equation.
The statistic in (9) known as the Trace statistic, tests the null hypthesis that the number of cointegrating vectors is less than or equal to r against a general alternative.

Johansen estimation

Note that since $ln(1) = 0$ and $ln(0) \uparrow -\infty$, it is clear that the trace statistic equals zero when all the λs are zero, whereas the further the eigenvalues are from zero, the more negative is $ln(1 - \hat{\lambda}_1$ and the larger is the statistic. Likewise the statistic in (9), known as the maximum eigenvalue statistic, tests a null of r cointegrating vectors against the specific alternative of r+1. *if* $\hat{\lambda}_{r+1}$is close to zero, the statistic will be small. Further if the null hypothesis is not rejected, the r cintegrating vectors contained in the matrix Γ can be estimated as the first r columns of matrix $\hat{V} = (\hat{v}_1, ..., \hat{v}_n)$which contains the eigenvectors associated to the eigenvalues in (9) computed as:
$(\lambda_i S_{11} - S_{10} S_{00}^{-1} S_{01})\hat{v}_i = 0, \quad i = 1, 2, ..., n$ Subject to the length normalization rule $\hat{V}' S_{11} V \hat{=} I_n$. Once Γhas been estimated, estimates of the B, $D_i$and $\sum$matrices in (7) can be obtained by inserting $\hat{\Gamma}$in their corresponding OLS formulae which will be functions of Γ.

Conclusion

I have presented the bare bones of Johansen estimation.
Sometimes it will be appropriate to test restrictions on the model as guided by theory. for example, if one is considering the relation between short-term and longterm interest rates, it may be wise to impose the restriction that the processes for both interest rates do not have linear trends and that the drift terms are restricted to appear in the cointegrating relationship interpreted as the "term structure".
Anyway we will look at some practical examples.