Cointegration Analysis

Richard G. Pierse

1 Introduction

Let $\mathbf{y}_t$ be an $n \times 1$ set of $I(1)$ variables. In general, any linear combination

$$\mathbf{a}'\mathbf{y}_t$$

will also be $I(1)$ for arbitrary $\mathbf{a} \neq \mathbf{0}$. However, suppose there exists an $n \times 1$ vector $\boldsymbol{\alpha}_i$ such that

$$\boldsymbol{\alpha}_i'\,\mathbf{y}_t \text{ is } I(0) \quad , \quad \boldsymbol{\alpha}_i \neq \mathbf{0}\ .$$

Then we say that the variables $\mathbf{y}_t$ are *cointegrated* and $\boldsymbol{\alpha}_i$ is a *cointegrating vector.*

Note that if $\boldsymbol{\alpha}_i$ is a cointegrating vector, then so is $k\boldsymbol{\alpha}_i$ for any $k \neq 0$ since $k\boldsymbol{\alpha}_i'\,\mathbf{y}_t \sim I(0)$.

Definition 1 *If*

$$\mathbf{y}_t \sim I(d) \quad \textit{and} \quad \boldsymbol{\alpha}_i'\,\mathbf{y}_t \ \sim \ I(d-b) \quad , \quad \boldsymbol{\alpha}_i \neq \mathbf{0}$$

then

$$\mathbf{y}_t \sim CI(d,b) \quad , \quad d \geq b > 0\,.$$

There can be r different cointegrating vectors, where $0 \leq r < n$. Note that r must be less than the number of variables n. If a test for r produces the result that $r = n$ then this is *incompatible* with the assumption that $y_t \sim I(1)$ and suggests some problem in the analysis.

Let

$$\boldsymbol{\alpha} = \left[\begin{array}{ccccc} \boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_i & \cdots & \boldsymbol{\alpha}_r \end{array}\right]$$

denote the $n \times r$ matrix of rank r, comprising all the cointegrating vectors. Then the $r \times 1$ vector

$$\boldsymbol{\alpha}'\mathbf{y}_t \ \sim \ I(0)$$

and, for *any nonsingular* $r \times r$ matrix $\mathbf{K}$, it also follows that

$$\mathbf{K}\boldsymbol{\alpha}'\mathbf{y}_t \ \sim \ I(0)\,.$$

In order to uniquely identify the cointegrating vectors, it is necessary to impose r^2 restrictions to pin down $\mathbf{K}$.

2 *VECM* Representation

Let $\mathbf{y}_t \sim I(1)$ be the pth order *VAR* model

$$\mathbf{y}_t = \mathbf{\Phi}_1\mathbf{y}_{t-1} + \mathbf{\Phi}_2\mathbf{y}_{t-2} + \cdots + \mathbf{\Phi}_p\mathbf{y}_{t-p} + \mathbf{u}_t\,.$$

If and only if the $\mathbf{y}$'s are cointegrated, with cointegrating vectors $\boldsymbol{\alpha}$, then the reparameterisation

$$\Delta\mathbf{y}_t = \mathbf{A}_1\Delta\mathbf{y}_{t-1} + \mathbf{A}_2\Delta\mathbf{y}_{t-2} + \cdots + \mathbf{A}_{p-1}\Delta\mathbf{y}_{t-p+1} + \boldsymbol{\gamma}\,\boldsymbol{\alpha}'\mathbf{y}_{t-1} + \mathbf{u}_t$$

will consist entirely of $I(0)$ variables. This result is called the *Granger Representation Theorem*, and the parameterisation is known as the *Vector Error Correction Mechanism* or *VECM.*

3 Estimating a Single Cointegrating Vector

Consider estimating a *single* cointegrating vector, $\boldsymbol{\alpha}_1'\,\mathbf{y}_t \sim I(0)$ in the *VAR* model

$$\mathbf{\Phi}(\mathbf{L})\mathbf{y}_t = \mathbf{u}_t$$

3.1 Static Regression

Partition $\mathbf{y}_t$ and $\boldsymbol{\alpha}_1$conformably as

$$\mathbf{y}_t' = \left[\begin{array}{ccc} y_{1t} & : & \mathbf{y}_{2t}' \end{array}\right]$$

and

$$\boldsymbol{\alpha}_1' = \left[\begin{array}{ccc} 1 & : & -\boldsymbol{\alpha}^{*\prime} \end{array}\right]\,.$$

This is an (*arbitrary*) normalising restriction. Then consider estimating the *static regression*

$$y_{1t} = \boldsymbol{\beta}'\mathbf{y}_{2t} + w_t\,.$$

From the definition of cointegration we know that for $\boldsymbol{\beta} = \boldsymbol{\alpha}^*$, $w_t \sim I(0)$, but for all other values of $\boldsymbol{\beta}$, then $w_t \sim I(1)$. Since *OLS* estimation minimises the *mean square error*, it is intuitively obvious that

$$\operatorname*{plim}_{T\to\infty} \widehat{\boldsymbol{\beta}} = \boldsymbol{\alpha}^*$$

and in fact it can be shown that the order of convergence is $O(T)$ as opposed to $O(\sqrt{T})$ in conventional models with $I(0)$ variables. This property of *OLS* with $I(1)$ variables is known as *super consistency.*

3.2 Testing Cointegration

Static regression provides a framework for testing cointegration, based on the *OLS* residuals $\widehat{w}_t$. Any of the standard unit root tests can be used, but the critical values will be different because $\widehat{w}_t$ is based on *estimated* parameters. The null hypothesis in the test is that $\widehat{w}_t \sim I(1)$, i.e. *zero cointegrating vectors*, against the alternative that $\widehat{w}_t \sim I(0)$, i.e. *one cointegrating vector*. Critical values for the *ADF* test, based on fitting response surfaces to simulation results, are given in MacKinnon (1991). A test based on the Durbin-Watson statistic from the static regression is described in Sargan and Bhargava (1983).

3.3 Engle-Granger Two-Step Procedure

Engle and Granger (1987) propose a two-step procedure for estimation.

Step 1: Estimate $\boldsymbol{\alpha}^*$ from the static regression

Step 2: Estimate the dynamics from the *VECM*

$$b_1(L)'\Delta\, y_{1t} = b_2(L)'\Delta\, y_{2t} + \gamma \widehat{w}_{t-1} + u_t$$

3.4 Problems with Static Regression

The static regression approach is simple and easy-to-use. However, it has certain drawbacks:

1. *It ignores dynamics*
2. *It ignores simultaneity*
3. *It is based on an artibrary normalisation*
4. If $r > 1$, then the static regression will find a *linear combination* of the r cointegrating vectors.

Although OLS estimates of $\boldsymbol{\alpha}^*$ are *super consistent*, they can still be *heavily biased* in finite samples, as has been found in simulation studies.

3.5 The Fully-Modified LS Estimator

Because of the problems of bias in the static regression, Phillips and Hanson (1990) have suggested a *non-parametric* correction for bias. This corrected OLS static regression is called the fully-modified LS estimator.

4 Estimating Several Cointegrating Vectors

The Johansen (1988, 1991) procedure is based on the maximum likelihood estimation of the *VECM* model

$$\Delta \mathbf{y}_t = \sum_{i=1}^{p-1} \mathbf{A}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\gamma}\, \boldsymbol{\alpha}' \mathbf{y}_{t-p} + \boldsymbol{\mu} + \boldsymbol{\delta}' \mathbf{x}_t + \mathbf{u}_t$$

where the *VAR* model has been generalised to include an intercept term $\boldsymbol{\mu}$ and a set of $I(0)$ exogenous variables $\mathbf{x}_t$. Note that the cointegration term has been redated at $t-p$ rather than $t-1$. (The dating of the cointegration term makes no essential difference to the analysis).

The log-likelihood function of this model, after concentrating out the nuisance parameters $\mathbf{A}_i$, $\boldsymbol{\mu}$, and $\boldsymbol{\delta}$, can be written as

$$L(\boldsymbol{\alpha}) = c - \frac{T}{2} \sum_{i=1}^{p} \log(1 - \lambda_i)$$

where λ_i are *generalised eigenvalues* that are the solution to the problem

$$\left| \boldsymbol{\lambda} \mathbf{S}_{kk} - \mathbf{S}_{k0} \mathbf{S}_{00}^{-1} \mathbf{S}_{0k} \right| = 0$$

where $\mathbf{S}_{ij} = T^{-1} \sum_{t=1}^{T} \mathbf{R}_{it} \mathbf{R}_{jt}'$, $i, j = 0, k$, and $\mathbf{R}_{0t}$ and $\mathbf{R}_{kt}$ are the vectors of residuals from regressing $\Delta \mathbf{y}_t$ and $\mathbf{y}_{t-p}$ respectively, on $\{\Delta \mathbf{y}_{t-1}, \cdots, \Delta \mathbf{y}_{t-p+1}, \boldsymbol{\mu}, \mathbf{x}_t\}$. The number of cointegrating vectors, r, is equal to the number of *non-zero* eigenvalues, λ_i.

4.1 Tests of the order of r

Let the eigenvalues λ_i, $i = 1, \cdots, n$ be ordered from largest to smallest. Then a test of the null hypothesis of r cointegrating vectors against the alternative of *more than* r can be based on either the *trace statistic*

$$H_0 : -T \sum_{i=r+1}^{n} \log(1 - \widehat{\lambda}_i) = 0$$

or the *maximal eigenvalue* statistic

$$H_0 : -T \ \log(1 - \widehat{\lambda}_{r+1}) = 0$$

4.2 The Distribution of the Test Statistics

The distribution of the trace and maximal eigenvalue statistics are *non-standard* and have been tabulated by Johansen (1995) and Osterwald-Lenum (1992). Unfortunately, as with the Dickey-Fuller statistic, the distribution depends on the nuisance parameter $\boldsymbol{\mu}$. Several models can be considered:

1. no intercept: $\boldsymbol{\mu} = 0$
2. restricted intercept (intercept only in error correction term)

$$\Delta \mathbf{y}_t = \sum_{i=1}^{p-1} \mathbf{A}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\gamma} \left(\boldsymbol{\alpha}' \mathbf{y}_{t-p} + \boldsymbol{\alpha}_0 \right) + \boldsymbol{\delta}' \mathbf{x}_t + \mathbf{u}_t$$

   where $\boldsymbol{\gamma}\boldsymbol{\alpha}_0 = \boldsymbol{\mu}$.

3. unrestricted intercept

It is also possible to consider the case where the data is generated by model 2 but model 3 is estimated. The three models are nested and it is possible to test the restricted models against the less restricted, *provided that the number of cointegrating vectors is known.*

4.3 Identification in the Johansen Procedure

In order to identify $\boldsymbol{\alpha}$, r^2 restrictions need to be imposed on the *VECM*. Johansen imposes the statistical restrictions

$$\boldsymbol{\alpha}_i' \mathbf{S}_{kk} \boldsymbol{\alpha}_i = 1 \quad \text{and} \quad \boldsymbol{\alpha}_i' \mathbf{S}_{kk} \boldsymbol{\alpha}_j = 0 \quad , \quad \forall i, j \quad j \neq i$$

However, several alternative identification restrictions have been proposed in the literature.

4.3.1 Phillips Triangular Form

Phillips (1991) proposed the triangular form identification restriction

$$\alpha = \begin{bmatrix} \mathbf{I}_r \\ -\overline{\boldsymbol{\alpha}} \end{bmatrix}$$

where $\overline{\boldsymbol{\alpha}}$ is $(n-r) \times 1$ and is unrestricted. This corresponds to a partitioning of the variables

$$\mathbf{y}_t' = \begin{bmatrix} \mathbf{y}_{1t}' & \mathbf{y}_{2t}' \end{bmatrix}$$

such that

$$\boldsymbol{\alpha}' \mathbf{y}_t = \mathbf{y}_{1t} - \overline{\boldsymbol{\alpha}}' \mathbf{y}_{2t} \sim I(0)$$

or

$$\mathbf{y}_{1t} = \overline{\boldsymbol{\alpha}}' \mathbf{y}_{2t} + \mathbf{v}_t$$

where the $n-r$ variables, $\mathbf{y}_{2t}$, are not themselves cointegrated.

4.3.2 Pesaran and Shin

Pesaran and Shin (1994) propose imposing the r^2 identifying restrictions on the basis of *a priori* economic theory. This is like imposing structural restrictions on the *VAR*. This procedure is available in *MicroFit Version 4*.

4.4 Hypothesis Testing

It is possible to test *overidentifying* restrictions on the cointegrating vectors. The Johansen identification restrictions make this a little awkward, however. For example, a set of *homogeneous* restrictions can be tested by

$$H_0 : \mathbf{R}_i' \boldsymbol{\alpha}_i = 0 \quad , \quad i = 1, \cdots, r$$

where $\mathbf{R}_i$ is an $n \times s$ matrix of known constants. The test statistic will be *asymptotically* distributed as χ^2 with $r(n-s)$ degrees of freedom.

5 Further reading

Good textbook accounts are given by Johansen (1995), Chapters 19 and 20 of Hamilton (1994) and Banerjee *et al.* (1993). Engle and Granger (1991) is a collection of readings that contains many of the classic papers in the field.

References

[1] Banerjee, A., J. Dolado, J.W. Galbraith and D.F. Hendry (1993), *Cointegration, Error-Correction, and the Analysis of Non-stationary Data*, Oxford University Press, Oxford.

[2] Engle, R.F. and C.W. J. Granger (1987), 'Cointegration and error correction: representation, estimation and testing', *Econometrica*, 55, 251–276.

[3] Engle, R.F. and C.W.J. Granger (1991), *Long Run Economic Relationships: Readings in Cointegration*, Oxford University Press, Oxford.

[4] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, NJ.

[5] Johansen, S. (1988), 'Statistical analysis of cointegrating vectors', *Journal of Economic Dynamics and Control*, 12, 231–54.

[6] Johansen, S. (1991), 'Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models', *Econometrica*, 59, 1551–80.

[7] Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

[8] MacKinnon, J.G. (1991), 'Critical values of cointegration tests', in R.F. Engle and C.W.J. Granger (eds.) *Long Run Economic Relationships: Readings in Cointegration*, Oxford University Press, Oxford.

[9] Osterwald-Lenum, M. (1992), 'A note with fractiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics: four cases', *Oxford Bulletin of Economics and Statistics*, 54, 461–472.

[10] Pesaran, M.H. and Y. Shin (1994), 'Long-run structural modelling', *mimeo*.

[11] Phillips, P.C.B. (1991), 'Optimal inference in cointegrated systems', *Econometrica*, 59, 282–306.

[12] Phillips, P.C.B. and B.E. Hanson (1990), 'Statistical inference in Instrumental Variables regression with I(1) processes', *Review of Economic Studies*, 57, 99–125.

[13] Sargan, J.D. and A.S. Bhargava (1983), 'Testing residuals from least squares regression for being generated by the Gaussian random walk', *Econometrica*, 51, 153–174.