COINTEGRATION

Ranjit Kumar Paul
I.A.S.R.I, Library Avenue, New Delhi- 110012
ranjitstat@gmail.com, ranjitstat@iasri.res.in

1. Introduction

Cointegration is a statistical property possessed by some time series data that is defined by the concepts of stationarity and the order of integration of the series. A stationary series is one with a mean value which will not vary with the sampling period. For instance, the mean of a subset of a series does not differ significantly from the mean of any other subset of the same series. Further, the series will constantly return to its mean value as fluctuations occur. In contrast, a non-stationary series will exhibit a time varying mean. The order of integration of a series is given by the number of times the series must be differenced in order to produce a stationary series. A series generated by the first difference is integrated of order 1 denoted as I(1). Thus, if a time series, is I(0), it is stationary, if it is I(1) then its change is stationary and its level is non-stationary.

In econometrics, cointegration analysis is used to estimate and test stationary linear relations, or cointegration relations, between non-stationary time series variables such as consumption and income, interest rates at different maturities, and stock prices. The vector autoregressive (VAR) model framework has been widely applied to model cointegration system. In the modeling of cointegrated systems, the determination of the number of cointegrating relations, or the cointegration rank, is the most important decision. Cointegration is said to exist between two or more non-stationary time series if they possess the same order of integration and a linear combination (weighted average) of these series is stationary. Thus, if x_t and y_t are non-stationary and are of the same order, there may exist a number b such that, the residual series, g_t, $(y_t - bx_t)$ is stationary. In this case x_t and y_t are said to be cointegrated with a cointegrating factor of b.

The significance of cointegration analysis is its intuitive appeal for dealing with difficulties that arise when using non-stationary series, particularly those that are assumed to have a long-run equilibrium relationship. For instance, when non-stationary series are used in regression analysis, one as a dependent variable and the other as an independent variable, statistical inference becomes problematic [Granger and Newbold, 1974]. Cointegration analysis has also become important for the estimation of error correction models (ECM). The concept of error correction refers to the adjustment process between short-run disequilibrium and a desired long run position. As Engle and Granger (1987) have shown, if two variables are cointegrated, then there exists an error correction data generating mechanism, and vice versa. Since, two variables that are cointegrated, would on average, not drift apart over time, this concept provides insight into the long-run relationship between the two variables and testing for the cointegration between two variables. With regard to testing procedures for the error correction model, once cointegration is ascertained, then the residuals from the cointegrating test, lagged one period, are used in a vector autoregression involving the appropriate differencing of the series (to ensure stationarity) forming the hypothesized relationship.

The following steps are followed in the application of cointegration. First, the order of integration of the time series data are tested. Next, if these series are integrated of the same order, then a cointegrating regression is estimated and the null hypothesis that the residuals of that regression are non-stationary is tested. Only if non-cointegration is rejected would the estimation of an ECM be attempted.

The most frequently used test for the cointegration rank is the likelihood ratio (LR) test (Johansen, 1988). Its popularity stems from the fact that it is conceptually simple since it an LR test and the test statistic is easy to compute by reduced rank regression. However, many studies have shown that the small sample distribution is not well approximated by the limiting distribution (Johansen, 2002). The distribution of the likelihood ratio statistic in finite samples depends on the parameters. A good account of understanding on cointegration can be obtained from Lütkepohl (2005), Greene (2000), Harris and Sollis (2003), to name a few. The likelihood ratio test of cointegration rank has been extensively used with financial time series.

2. Cointegration: Some basic concepts

***Vector Autoregressive (VAR) process*:**

Let us consider a univariate time series y_t, t=1,2,…,T arising from the model

$$y_t=\upsilon+\emptyset_1 y_{t-1}+\emptyset_2 y_{t-2}+\ldots+\emptyset_k y_{t-k}+u_t\,, \quad u_t\sim IN(0,\sigma) \tag{2.1}$$

where, u_t is a sequence of uncorrelated error terms and $\emptyset_j$, j=1,…,k are the constant parameters. This is a sequentially defined model; y_t is generated as a function of its own past values. This is a standard autoregressive framework or AR(k), where k is the order of the autoregression.

If a multiple time series $\mathbf{y_t}$ of n endogenous variables is considered, the extension of (2.1) will give the VAR(k) model (VAR model of order k), i.e. it is possible to specify the following data generating procedure (d.g.p.) and model $\mathbf{y_t}$ as an unrestricted VAR involving upto k lags of $\mathbf{y_t}$,

$$\mathbf{y_t}=\upsilon+\mathbf{A_1 y_{t-1}}+\ldots+\mathbf{A_k y_{t-k}}+\mathbf{u_t}, \quad \mathbf{u_t}\sim IN(\mathbf{0},\mathbf{\Sigma}) \tag{2.2}$$

where, $\mathbf{y_t}=(y_{1t}, y_{2t},\ldots, y_{nt})'$ is (n×1) random vector, each of the $\mathbf{A_i}$ is an (n×n) matrix of parameters, $\boldsymbol{\upsilon}$ is a fixed (n×1) vector of intercept terms. Finally, $\mathbf{u_t}=(u_{1t},u_{2t},\ldots,u_{nt})$ is a n-dimensional white noise or innovation process, i.e., $E(\mathbf{u_t})=\mathbf{0}$, $E(\mathbf{u_t u'_t})=\mathbf{\Sigma}$ and $E(\mathbf{u_t u'_s})=\mathbf{0}$ for $s\neq t$. The covariance matrix $\mathbf{\Sigma}$ is assumed to be non-singular. Using lag operator (L) (2.2) can be written as, $(\mathbf{I_n}-\mathbf{A_1}L-\ldots-\mathbf{A_k}L^k)\,\mathbf{y_t}=\upsilon+\mathbf{u_t}$.

The process $\mathbf{y_t}$ is said to be stable if the roots of the polynomial, $|\mathbf{I_n}-\mathbf{A_1}L-\ldots-\mathbf{A_k}L^k|=\mathbf{0}$ lie outside the complex unit circle i.e. have modulus greater than one.

2.1 Cointegrated systems

When the data is non-stationary purely due to unit roots (integrated once, denoted by I(1)), they could be brought back to stationarity by the linear transformation of differencing. If a series must be differenced d times before it becomes stationary, then it contains d unit roots and is said to be integrated of order d, denoted by I(d). Let $\mathbf{y_t}$ be and n×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×1 vector $\boldsymbol{\alpha}_\mathbf{i}$ such that $\boldsymbol{\alpha}_\mathbf{i}'\mathbf{y_t}$ is I(0), $\boldsymbol{\alpha}_\mathbf{i}\neq\mathbf{0}$, then it is said that the variables in $\mathbf{y_t}$ are cointegrated of order one, denoted CI(1) and $\boldsymbol{\alpha}_\mathbf{i}$ is a cointegrating vector. It is to be mentioned that if $\boldsymbol{\alpha}_\mathbf{i}$ is a cointegrating vector then so is $k\boldsymbol{\alpha}_\mathbf{i}$ for any $k\neq0$ since $k\boldsymbol{\alpha}_\mathbf{i}'\mathbf{y_t}\sim I(0)$.

Definition:

If, $\mathbf{y_t}\sim I(d)$ and $k\boldsymbol{\alpha}_\mathbf{i}'\mathbf{y_t}\sim I(d\text{-}b)$, $\boldsymbol{\alpha}_\mathbf{i}\neq0$ then,

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

There can be r different cointegrating vectors, where $0 \leq r < n$, i.e. r must be less than the number of variables n. In such a case, we can distinguish between a long-run relationship between the variables contained in $\mathbf{y_t}$, that is, the manner in which the variables drift upward together, and the short-run dynamics, that is the relationship between deviations of each variable from their corresponding long-run trend. The implication that non-stationary variables can lead to spurious regressions unless at least one cointegration vector is present means that some form of testing for cointegration is almost mandatory.

2.2 Error correction models (ECM)

Before the concept of cointegration was introduced, the closely related error correction models were discussed in the econometrics literature. In an error correction model, the changes in a variable depend on the deviations from some equilibrium relation. Suppose, for instance, that y_t represents the price of a commodity in a particular market and x_t is the corresponding price of the same commodity in another market. Assume further more that the equilibrium relation between the two variables is given by $y_t=\beta x_t$ and the changes in y_t (i.e. $\Delta y_t = y_t - y_{t-1}$) depend on the deviation from this equilibrium in period (t-1).

$$\Delta y_t=\alpha_1 \left(y_{t-1}-\beta_1 x_{t-1}\right)+u_{yt} \tag{2.3}$$

A similar relation may hold for x_t,

$$\Delta x_t=\alpha_2 \left(y_{t-1}-\beta_1 x_{t-1}\right)+u_{xt} \tag{2.4}$$

In a more general error correction model, the Δy_t and Δx_t may in addition depend on previous changes in both variables as, for instance, in the following model:

$$\Delta y_t=\alpha_1 \left(y_{t-1}-\beta_1 x_{t-1}\right)+\gamma_{11}\Delta y_{t-1}+\gamma_{12}\Delta x_{t-1}+u_{yt} \tag{2.5}$$

$$\Delta x_t=\alpha_2 \left(y_{t-1}-\beta_1 x_{t-1}\right)+\gamma_{21}\Delta y_{t-1}+\gamma_{22}\Delta y_{t-1}+u_{xt} \tag{2.6}$$

In vector and matrix notation the ECM can be written as,

$$\mathbf{\Delta y_t=\alpha\beta' y_{t-1}+\Gamma_1\Delta y_{t-1}+u_t} \tag{2.7}$$

Where, $\boldsymbol{\alpha}=[\alpha_1, \alpha_2]'$, $\boldsymbol{\beta'}=[1, -\beta_1]$ and, $\mathbf{\Gamma_1}=\begin{bmatrix}\gamma_{11} & \gamma_{12}\\ \gamma_{21} & \gamma_{22}\end{bmatrix}$

Vector Error Correction Model (VECM)

Equation (2.2) can be reformulated into a vector error correction model (VECM) form:

$$\mathbf{\Delta y_t= \Pi y_{t-1}+\textstyle\sum_{j=1}^{k-1}\Gamma_j\Delta y_{t-j} + u_t},\ \ t=k+1,\ldots,T \tag{2.8}$$

where, $\mathbf{\Gamma_i=-(A_{i+1}+\ldots+A_k)}$, $i=1,\ldots,k-1$, and $\mathbf{\Pi=-(I-A_1-\ldots-A_k)}$.This way of specifying the system contains information on both the short-run and long run adjustments to changes in $\mathbf{y_t}$, via the estimates of $\mathbf{\hat{\Gamma}_i}$ and $\mathbf{\hat{\Pi}}$ respectively. As will be discussed in the later sections, $\mathbf{\Pi=\alpha\beta'}$, where $\boldsymbol{\alpha}$ represents the speed of adjustments to disequilibrium and $\boldsymbol{\beta}$ is a matrix of long run coefficients such that the term $\mathbf{\beta' y_{t-1}}$ embedded in (2.8) represents up to (n-1) cointegration relationships in the multivariate model.

2.3 Rank of cointegration

Assuming $\mathbf{y_t}$ is a vector of non-stationary I(1) variables, then all the terms in (2.8) involving $\Delta\mathbf{y}_{t-i}$ are I(0) while $\mathbf{\Pi y_{t-1}}$ must also be stationary for $\mathbf{u_t}$ ~I(0) to be white noise. There are three instances when the requirement that $\mathbf{\Pi y}_{t-1}$~I(0) is met; first, when all the variables in $\mathbf{y_t}$ are in fact stationary, which is an uninteresting case in the present context since it implies there is no problem of spurious regression and the appropriate modeling strategy is to estimate the standard VAR in levels. The second instance is when there is no cointegration at all, implying that there are no linear combinations of the $\mathbf{y_t}$ that are I(0), and consequently $\mathbf{\Pi}$ is an (n×n) matrix of zeroes. In this case, the appropriate model is a VAR in fist-differences involving no long-run elements. The third way for $\mathbf{\Pi y_{t-1}}$ to be I(0) is when there exists up to (n-1) cointegration relationships: $\mathbf{\Pi' y_{t-1}}$~I(0). The matrix $\mathbf{\Pi}$ produces linear combinations of the variables in $\mathbf{y_t}$. But, as mentioned earlier, not all linear combinations can be cointegrated. The number of such independent linear combinations is r<n. Therefore, although there must be VAR representation of the model, cointegration implies a restriction in the rank of $\mathbf{\Pi}$**.** It cannot have full rank, its rank is r<n. In this instance r ≤ (n-1) cointegration vectors exists in $\mathbf{\Pi}$ (i.e., r columns of $\mathbf{\Pi}$ form r linearly independent combinations of the variables in $\mathbf{y_t}$, each of which is stationary), together with (n-r) non-stationary vectors. Here, r is known as the 'Rank of cointegration'. Only the cointegration vectors in $\boldsymbol{\beta}$ enters (2.8), otherwise $\mathbf{\Pi y_{t-1}}$ would not be I(0), which implies that the last (n-r) columns of $\boldsymbol{\alpha}$ are insignificantly small i.e. effectively zero. Consequently, testing for cointegration amounts to a consideration of rank of $\mathbf{\Pi}$ (i.e. finding the number, r, of linearly independent columns in $\mathbf{\Pi}$).

If $\mathbf{\Pi}$ has full rank, then the variables in $\mathbf{y_t}$ are I(0), while if the rank of $\mathbf{\Pi}$ is zero, then there are no cointegration relationships. Neither of these two cases is of interest. More usually, $\mathbf{\Pi}$ has a reduced rank. Next the test for the reduced rank of $\mathbf{\Pi}$ is discussed.

3. Test for the Rank of Cointegration

A natural first step in the analysis of cointegration is to establish that it is indeed a characteristic of the data. Two broad approaches for testing for cointegration have been developed. The Engle and Granger (1987) method is based on assessing whether the single equation estimates of the equilibrium error appear to be stationary. The second approach, due to Johansen (1988), is based on the VAR approach. As noted earlier, if a set of variables is truly cointegrated, then it is possible to detect the implied restriction in an otherwise unrestricted VAR. In the next section, the Johansen's approach is described.

3.1 Johansen's method of reduced rank regression

In the VECM (2.8), $\mathbf{\Pi}$ has a reduced rank of r, so, $\mathbf{\Pi}$ can always be written as the product of two n×r matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ of full column rank as $\mathbf{\Pi}=\boldsymbol{\alpha\beta}'$. Let us consider an n-dimensional vector autoregressive (VAR) model in error correction form with a deterministic term, (2.8) can be written as:

$$\Delta \mathbf{y_t} = \boldsymbol{\mu}_t + \mathbf{\Pi y_{t-1}} + \sum_{j=1}^{k-1} \mathbf{\Gamma_j} \Delta \mathbf{y_{t-j}} + \mathbf{u_t} \tag{3.1}$$

where, the initial observations are assumed to be fixed. Here, $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ denotes the deterministic part of the model. The matrix $\mathbf{\Pi}=\boldsymbol{\alpha\beta}'$ has a reduced rank r, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are n×r matrices of rank r<n. The deterministic part can be written as $\boldsymbol{\mu}_t=\boldsymbol{\mu}_0+\boldsymbol{\alpha\rho}_1 t$, where $\boldsymbol{\rho}_1$ is r×1. From the discussion in previous section, imposing the null hypothesis of r cointegrating relations, (3.1) can be written as,

$$\Delta \mathbf{y_t} = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}(\boldsymbol{\rho}_1 t + \boldsymbol{\beta}' \mathbf{y_{t-1}}) + \sum_{j=1}^{k-1} \mathbf{\Gamma_j} \Delta \mathbf{y_{t-j}} + \mathbf{u_t} \quad t=k+1,\ldots,T \tag{3.2}$$

To ensure that $\mathbf{y}_t$ is an I(1) process the it is assumed that the roots of $|\mathbf{\Pi}(z)|=0$, where, $|\mathbf{\Pi}(z)|=(1-z)\mathbf{I}_n-\boldsymbol{\alpha}\boldsymbol{\beta}'z-\mathbf{\Gamma}_1(1-z)z-\ldots-\mathbf{\Gamma}_{k-1}(1-z)z^{k-1}$ and, the matrix $\boldsymbol{\alpha}'_{\perp}\mathbf{\Gamma}\boldsymbol{\beta}_{\perp}$ has full rank, where $\mathbf{\Gamma}=\mathbf{I}_n-\sum_{j=1}^{k-1}\mathbf{\Gamma}_j$ and $\boldsymbol{\alpha}\perp$ and $\boldsymbol{\beta}\perp$ are the orthogonal complements of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively.

It is possible to correct for short-run dynamics by regressing $\Delta\mathbf{y}_t$ and $\tilde{\mathbf{y}}_t=(\mathbf{y}'_t,t)'$ separately on the lagged differences $\Delta\mathbf{y}_{t-1},\ldots,\Delta\mathbf{y}_{t-k+1}$ and a constant. The residual vectors $\mathbf{R}_{0t}$ and $\mathbf{R}_{1t}$ are obtained as,

$$\mathbf{R}_{0t}=\text{Residual from the regression of }\Delta\mathbf{y}_t\text{ on }\Delta\mathbf{y}_{t-1}\ldots,\Delta\mathbf{y}_{t-k+1}$$

$$\mathbf{R}_{1t}=\text{Residual from the regression of }\tilde{\mathbf{y}}_t\text{ on }\Delta\mathbf{y}_{t-1}\ldots,\Delta\mathbf{y}_{t-k+1}$$

which then can be used to form the product moment matrices of the residuals:

$$\mathbf{S}_{ij}=T^{-1}\sum_{t=k+1}^{T}\mathbf{R}_{it}\mathbf{R}'_{jt}\,,\qquad i,j=0,1$$

The estimate of $\boldsymbol{\beta}$ is obtained as the eigenvectors corresponding to the r largest eigenvalues from solving the equation:

$$\left|\lambda\mathbf{S}_{11}-\mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\right|=0 \tag{3.3}$$

Which gives the n eigenvalues $\hat{\lambda}_1>\hat{\lambda}_2>\ldots>\hat{\lambda}_n>0,\ \hat{\lambda}_{n+1}=0$ and the corresponding eigenvectors $\hat{\mathbf{V}}=\hat{\mathbf{v}}_1>\hat{\mathbf{v}}_2>\ldots>\hat{\mathbf{v}}_n$. Those r elements in $\hat{\mathbf{V}}$ that determine the linear combination of stationary relationships can be denoted by $\hat{\boldsymbol{\beta}}=(\hat{\mathbf{v}}_1,\ldots\hat{\mathbf{v}}_r)$ (i.e. they are the cointegration vectors). This is because the largest eigenvalues are the largest squared canonical correlations between the 'levels' residuals $\mathbf{R}_{1t}$ and the 'difference' residuals $\mathbf{R}_{0t}$; that is, estimates of all the distinct $\hat{\mathbf{v}}'_i\mathbf{y}_t$, (i=1,2,…,r) combinations of the I(1) levels of $\mathbf{y}_t$ that produce high correlations with the stationary $\Delta\mathbf{y}_t\sim I(0)$ elements in (3.1), such combinations being the cointegration vectors by virtue of the fact that they must themselves be I(0) to achieve a high correlation. Thus the magnitude of $\hat{\Box}_{\Box}$ is a measure of how strongly the cointegrating relations $\hat{\mathbf{v}}'_i\mathbf{y}_t$ are correlated with the stationary part of the model. The last (n-r) combinations i.e. $\hat{\mathbf{v}}'_i\mathbf{y}_t$, (i=r+1,…,n) indicate the non-stationary combinations, and theoretically these are uncorrelated with the stationary elements in (3.1). Consequently, for the eigenvectors corresponding to the non-stationary part of the model, $\hat{\lambda}_i=0$ for i= r+1,…,n.

3.2 Likelihood Ratio (LR) test of cointegration rank

A test for r cointegrating vectors can be based on the above maximum likelihood approach. The statistical problem is to derive a test procedure to discriminate between the λ_i, i=1,…,r, which are large enough to correspond to stationary $\boldsymbol{\upsilon}_i'\mathbf{y}_t$, and those λ_i, i=r+1,…,n, which are small enough to correspond to non-stationary eigenvectors. The rank r is determined by a likelihood-ratio test procedure for the hypotheses:

H_0: rank=r<n, i.e., there exist r cointegration relations against the alternative,
H_1: rank=r_1>r.

Thus to test the null hypothesis that there are r cointegration vectors amounts to:

$$H_0:\lambda_i=0\qquad i=r+1,\ldots,n$$

where, only the first r eigenvalues are non-zero. This restriction can be imposed for different values of r. It is possible to test the null hypothesis by using what is called the trace statistic, given by:

$$Q(r)=-T \sum_{i=r+1}^{n} \log(1-\lambda_i) \qquad r=0,1,2,\ldots,n-2,n-1 \tag{3.4}$$

Let, $\tilde{Q}$ denote the realized value of the likelihood ratio statistic Q. If $\lambda_{r+1}=\ldots=\lambda_n=0$, the test statistic should be small (close to zero), which delivers the critical value under the null. The test is based on non-standard asymptotic distribution. Asymptotic critical values are provided in Doornik (1998). The asymptotic distributions depend on whether there is a constant and/or a trend in the model. In many economic applications, the size of the sample is often quite small and the tabulated asymptotic distributions can be rather poor approximations as has been demonstrated in many papers. Another reason of concern is that when using correct small sample distributions for the trace test, the size of the test is correct, but the power can be low, sometimes even of the same magnitude as the size. In such cases, a 5% test procedure will reject a unit root incorrectly 5% of the time, but accept a unit root incorrectly 95% of the time.

If dummy variables enter the deterministic part of the multivariate model, then these critical values are only indicative. Similarly, if there is only a small sample of observations on $\mathbf{y_t}$, then there are likely to be problems with the power and size properties of the above test when using asymptotic critical values. Many studies has shown that in such a situation the Johansen procedure over-rejects when the null is true. Thus he suggests taking account of the number of parameters to be estimated in the models and making an adjustment for degrees of freedom.

5. SAS syntax for cointegration test

Consider the quarterly data for the years 1954 to 1987 for the US econlomy (Lütkepohl 1993, Table E.3.). The program is taken from SAS help and documentation.

```
title 'Analysis of U.S. Economic Variables';
data us_money;
  date=intnx( 'qtr', '01jan54'd, _n_-1 );
  format date yyq. ;
  input y1 y2 y3 y4 @@;
  y1=log(y1);
  y2=log(y2);
  label y1='log(real money stock M1)'
      y2='log(GNP in bil. of 1982 dollars)'
      y3='Discount rate on 91-day T-bills'
      y4='Yield on 20-year Treasury bonds';
datalines;
450.9 1406.8  0.010800000 0.026133333
453.0 1401.2 0.0081333333 0.025233333

  ... more lines ...
run;
proc varmax data=us_money;
  id date interval=qtr;
  model y1-y4 / p=2 lagmax=6 dftest
           print=(iarr(3) estimates diagnose)
           cointtest=(johansen=(iorder=2))
           ecm=(rank=1 normalize=y1);
  cointeg rank=1 normalize=y1 exogeneity;
run;
```

Output: Unit Root Tests

Following table shows the output for Dickey-Fuller tests for the nonstationarity of each series. The null hypotheses is to test a unit root. All series have a unit root.

Table1: Unit Root Test					
Variable	**Type**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**
y1	**Zero Mean**	0.05	0.6934	1.14	0.9343
	Single Mean	-2.97	0.6572	-0.76	0.8260
	Trend	-5.91	0.7454	-1.34	0.8725
y2	**Zero Mean**	0.13	0.7124	5.14	0.9999
	Single Mean	-0.43	0.9309	-0.79	0.8176
	Trend	-9.21	0.4787	-2.16	0.5063
y3	**Zero Mean**	-1.28	0.4255	-0.69	0.4182
	Single Mean	-8.86	0.1700	-2.27	0.1842
	Trend	-18.97	0.0742	-2.86	0.1803
y4	**Zero Mean**	0.40	0.7803	0.45	0.8100
	Single Mean	-2.79	0.6790	-1.29	0.6328
	Trend	-12.12	0.2923	-2.33	0.4170

The Johansen cointegration rank test shows whether the series is integrated order either 1 or 2 as shown in table2. The last two columns in table 2 explain the cointegration rank test with integrated order 1. The results indicate that there is the cointegrated relationship with the cointegration rank 1 with respect to the 0.05 significance level because the test statistic of 20.6542 is smaller than the critical value of 29.38. Now, look at the row associated with r=1. Compare the test statistic value and critical value pairs such as (219.62395, 29.38), (89.21508, 15.34), and (27.32609, 3.84). There is no evidence that the series are integrated order 2 at the 0.05 significance level.

Output: Cointegration Rank Test

Table2: Cointegration Rank Test for I(2)						
r\k-r-s	**4**	**3**	**2**	**1**	**Trace of I(1)**	**5% CV of I(1)**
0	384.60903	214.37904	107.93782	37.02523	55.9633	47.21
1		219.62395	89.21508	27.32609	20.6542	29.38
2			73.61779	22.13279	2.6477	15.34

Table2: Cointegration Rank Test for I(2)						
r\k-r-s	**4**	**3**	**2**	**1**	**Trace of I(1)**	**5% CV of I(1)**
3				38.29435	0.0149	3.84
5% CV I(2)	47.21000	29.38000	15.34000	3.84000		

References

Doornik, A.J. (1998). Approximations to the asymptotic distributions of cointegration tests. *Journal of Economic Surveys*, **12**, 573-593.

Engle, R., and Granger, C. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica*, **35**, 251-276.

Greene, W. H. (2000). *Econometric Analysis*, Prentice Hall, Inc., Upper Saddle River, New Jersey.

Harris, R. and Sollis, R. (2003). *Applied Time Series Modelling and Forecasting*, John Wiley & Sons, West Sussex, England.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, **12**, 231-254.

Johansen, S. (2002). A small sample correction for the test of cointegration rank in the vector autoregressive model. *Econometrica*, **70**, 1929-1961.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin, Hiedelberg.