Universidad Nacional de La Plata

Departamento de Economía
Facultad de Ciencias Económicas
Universidad Nacional de La Plata

Econometría

A Primer on Unit-Roots and Cointegration

Walter Sosa Escudero

Trabajo Docente Nro. 3
Marzo 2000

A Primer on Unit-Roots and Cointegration*

Walter Sosa Escudero
(wsosa@feedback.net.ar)
Universidad Nacional de La Plata

May 3, 2000

The analysis of unit-root processes and cointegrated systems has played a prominent role in econometrics and macroeconomics in the last decade, with applications to diverse fields such as macroeconomics, finance, economic history, international economics, etc. The reasons for such a rapid expansion of the subject are its strong intuitive appeal and its highly involved technical complexity. After a decade of the publication of the seminal work by Engle and Granger (1987) cointegration and unit-root techniques have become standard elements in every applied econometrician's toolkit. These notes present an introductory and informal approach to the topic at the level of any undergratuate book in econometrics. The subject well deserves a detailed and rigurous study for which there are many specialized textbooks discussed at the end of these notes.

1 Time Series. Stationarity

We are assuming some familiarity with the basic theory of stationary ARMA process. In this section we review some basic concepts. A *time series* process is a sequence of random variables ordered in time. The notation Y_t stands for the random variable variable Y at period t.

A process Y_t is *stationary* if the following conditions hold:

1. $E(Y_t) = \mu < \infty$ (constant mean)

*These notes where written to be used in classes taught at University of Illinois at Urbana-Champaign, University of San Andres, and National University of La Plata. Comments welcome.

2. $Cov(Y_t, Y_{t-s}) = \gamma_s < \infty$ (depends on s but not on t)

Actually this is *weak* stationarity[1] or *covariance* stationarity, which means that the first and second order moment structure of Y_t is constant over time. For $s = 0$ the second condition implies that a stationary process has constant variance, $Var(Y_t) = \sigma^2$.

1.1 Some Examples:

1. *White Noise:* This is the simplest form a time series process can take. The white note process is a zero mean, constant variance collection of random variables which are uncorrelated over time. More specifically, Y_t is a white noise process if $Y_t = \varepsilon_t$

   where:

   (a) $E(\varepsilon_t) = 0$

   (b) $Var(\varepsilon_t) = \sigma^2$

   (c) $Cov(\varepsilon_t, \varepsilon_{t-s}) = 0$ for all $s, t, s \neq 0$

   Just by checking the properties of ε_t we see that the white noise process is stationary.

2. *The zero mean first order autorregresive process AR(1):* Y_t is a AR(1) process with zero mean if:

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

   where ε_t is a white noise process as defined above.

   This process will be stationary if and only if $|\phi| < 1$. It can be easily checked that when this is so:

   1. $E(Y_t) = 0$
   2. $V(Y_t) = \sigma^2/(1 - \phi^2)$
   3. $Cov(Y_t, Y_{t-j}) = \phi^j \sigma^2/(1 - \phi^2)$

[1] In this notes we will restrict our atention to weak (as opposed to *strong*) stationarity and will use 'stationary' to mean weak stationarity, which suffices for the purposes of this notes. See Hamilton (1994) for details

These properties can be verified by reexpresing Y_t using repeated substitutions in the AR(1) model in the following way:

$$\begin{array}{rcl} Y_t & = & \phi Y_{t-1} + \varepsilon_t \\ & = & \phi(\phi Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ & = & \phi(\phi(Y_{t-3} + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t \\ \cdots & \cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ Y_t & = & \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \phi^3\varepsilon_{t-3} + \cdots \\ Y_t & = & \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i} \end{array}$$

Then, from the last expression we get:

(a) $E(Y_t) = \sum_{i=0}^{\infty} \phi^i E(\varepsilon_{t-i}) = 0$

(b) $V(Y_t) = \sum_{i=0}^{\infty} \phi^{2i} V(\varepsilon_{t-i}) = \sum_{i=1}^{\infty} \phi^{2i}\sigma^2 = \sigma^2/(1-\phi^2)$
since $Cov(\varepsilon_t, \varepsilon_{t-s}) = 0$ for all s, t

(c) $Cov(Y_t, Y_{t-s}) =$

The general AR(1) process is specified as $Y_t = \mu + \phi Y_{t-1} + \varepsilon_t$ and it can be easily verified that $E(Y_t) = \mu/(1-\phi)$ and that the variance and covariances are exactly the same as before.

2 Non-stationary processes

So far we have assumed that the series were stationary, which is a somewhat unrealistic situation in most macroeconomic variables. Trivially, a non-stationary process arises when one of the conditions for stationarity does not hold.

2.1 Some examples:

1. The *determistic trend* process corresponds to:

$$Y_t = a + dt + u_t$$

where: a and d are parameters, t is a time index and u_t is any zero mean stationary process with variance σ^2. The deterministic trend process presents stationary fluctuations around a linear trend. The process is obviously non-stationary since its mean changes with time:

$$E(Y_t) = a + dt$$

Nevertheless, its variance is constant:

$$V(Y_t) = V(u_t) = \sigma^2$$

2. The *random walk:*

$$Y_t = Y_{t-1} + \varepsilon_t$$

where ε_t is a white noise process with variance σ^2. Note that this is the zero mean $AR(1)$ process with $\phi = 1$. It can be easily checked that $E(Y_t) = 0$ and $V(Y_t) = t\sigma^2$

In order to obtain these results, assume that the process starts at $t = 0$ and also assume that $Y_0 = 0$. Then:

$$\begin{array}{lcl} Y_0 & = & 0 \\ Y_1 & = & Y_0 + \varepsilon_1 = \varepsilon_1 \\ Y_2 & = & Y_1 + \varepsilon_2 = \varepsilon_1 + \varepsilon_2 \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ Y_t & = & Y_{t-1} + \varepsilon_t = \sum_{i=1}^{t} \varepsilon_i \end{array}$$

and then take expectations and variance to obtain the results.

In this case even though the mean of the process is constant, its variance is not, it grows unboundedly over time, so the process is not stationary.

3. The *random walk with drift:*

$$Y_t = m + Y_{t-1} + \varepsilon_t$$

where m is a parameter known as the 'drift' and ε_t is a white noise process.

Using repeated substitutions as in the previous case:

$$\begin{array}{lcl} Y_0 & = & 0 \\ Y_1 & = & m + Y_0 + \varepsilon_1 = m + \varepsilon_1 \\ Y_2 & = & m + Y_1 + \varepsilon_2 = 2m + \varepsilon_1 + \varepsilon_2 \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ Y_t & = & m + Y_{t-1} + \varepsilon_t = tm + \sum_{i=1}^{t} \varepsilon_i \end{array}$$

we can easily see that in this case both the mean and the variance are time dependent:

$$E(Y_t) = tm$$

$$V(Y_t) = t\sigma^2$$

Note that now both the mean and the variance grow over time.

3 Unit Roots

The problem of testing for unit uoots was likely to be the one of the most important and controversial topics in econometrics in the last decade. In our simple framework testing for unit-roots means testing the hypothesis:

$$H_0 : \phi = 1 \qquad \text{vs.} \qquad H_A : |\phi| < 1$$

in the following general model:

$$Y_t = m + \phi Y_{t-1} + dt + \varepsilon_t \tag{1}$$

where ε_t is a white noise process.

It is easy to see that this general model contains all the previous examples as special cases:

Case	Process	Parameters	Hypothesis about ϕ
1	AR(1)	$\lvert\phi\rvert < 1, d = 0$	Alternative
2	Deterministic Trend	$\lvert\phi\rvert < 1, d \neq 0$	Alternative
3	Random Walk	$\phi = 1, d = m = 0$	Null
4	Random Walk with drift	$\phi = 1, d = 0$	Null

The origin of the term 'unit-root' can be seen by rewriting (1) in terms of the lag operator L:

$$Y_t(1 - \phi L) = m + dt + \varepsilon_t$$

The polynomial $(1 - \phi Z)$, which is obtained by taking L as variable in $(1 - \phi L)$, has a single root equal to $1/\phi$, which is equal to one if $\phi = 1$. Then, processes like 3 or 4 in the table are called *unit-root processes.*

3.1 Why is it important to test for unit roots?

1. Stationary processes posess many interesting features that make them desirable. Unit root tests can provide a test of whether a series is stationary of not when used under $d = 0$. For example, if we test Model 3 against Model 1, then the model is non-stationary under the null hypothesis (a random walk) and stationary under the alternative (an AR(1)).

2. Even though the realizations of a random walk with drift (model 4) look much like the deterministic trend process, they have a completely different statistical structure. First, in the deterministic trend process what is causing the non-stationarity is the presence of the time trend, so by 'removing' it we restore stationarity. In the random walk with drift we have that both the mean and the variance vary over time. In our simple case, if Y_t is a random walk with drift:

   $$Y_t = m + Y_{t-1} + \varepsilon_t$$

   then substracting Y_{t-1} in both sides we get:

   $$\Delta Y_t = m + \varepsilon_t$$

   which is a stationary process (Δ is the difference operator). Then, if Y_t is a random walk with drift its first difference is a stationary process, so the non-stationarity of Y_t is 'removed' by considering its first difference.

   Second, consider the effect of a shock at time t (ε_t) s periods after it took place. Using the same logic as before, lets write the random walk with drift using repeated substitutions but this time forward:

   $$\begin{aligned} Y_t &= m + Y_{t-1} + \varepsilon_t \\ Y_{t+1} &= m + Y_t + \varepsilon_{t+1} \\ Y_{t+2} &= m + Y_{t+1} + \varepsilon_{t+2} = 2m + Y_t + \varepsilon_{t+1} + \varepsilon_t \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Y_{t+s} &= m + Y_{t+s-1} + \varepsilon_{t+s} = sm + Y_t + \varepsilon_{t+s} + \varepsilon_{t+s-1} + \cdots + \varepsilon_t \end{aligned}$$

   In order to obtain a similar result for the deterministic trend process, consider the case where u_t follows a zero-mean AR(1) process: $u_t =$

ϕu_{t-1}. In this case, similar substitutions for the deterministic trend process give:

$$
\begin{array}{lcl}
Y_t & = & a + dt + u_t \\
Y_{t+1} & = & a + d(t+1) + u_{t+1} = a + d(t+1) + \phi u_t + \varepsilon_t \\
Y_{t+2} & = & a + d(t+2) + u_{t+2} = a + d(t+2) + \phi \varepsilon_{t+1} + \phi^2 \varepsilon_t \\
\ldots & \ldots & \ldots \\
Y_{t+s} & = & a + d(t+s) + u_{t+s} = a + d(t+s) + \varepsilon_{t+s} + \phi \varepsilon_{t+s-1} + \cdots + \phi^s \varepsilon_t
\end{array}
$$

Now consider the effect of a shock ε_t in Y_t s periods ahead, that is, $\partial Y_{t+s}/\partial \varepsilon_t$. In the case of the random walk process this derivative is equal to 1 while in the deterministic trend process this derivative is ϕ^s. When ε_t is stationary ($|\phi| < 0$), the effect of ε_t on Y_{t+s} tends to zero when s tends to infinity whereas in the case of the random walk with drift this effect remains constant. Then, in the case of the deterministic trend process, the effect of a shock tends to wear out (is *transitory*) while in the case of the random walk is *permanent*.

The main differences between these processes are summarized in the following table:

Deterministic Trend	Random Walk with Drift
1. Transformations to achieve stationarity	
Detrend	Difference
2. Effects of shocks	
Wears out	Permanent
3. Variance	
Bounded	Unbounded

A test for unit roots may help determine which is the source of non-stationarity. For example, if we test Model 4 against Model 2, then the process is non-stationary under both the null and the alternative hypothesis. But under the later the process has a deterministic trend while under the former it has a stochastic trend. For example, in the case of the GDP, economic theory suggest that is this series is a growing process so testing for stationarity is not an issue. The point is to determine what is the source of the non-stationarity.

3. *Spurious regressions:* Suppose Y_t and X_t are unit-root processes. In a very influential paper, Granger and Newbold (1974) showed that the regression between such series is very likely to be *spurious* in the sense that it will present a high R^2 and significant t-statistics even when the series are completely unrelated. It was common practice to either detrend or difference the variables to avoid the danger of spurious regression. Unit root tests may help determine which procedure to apply to achieve stationarity. It is important to remark that not every regression between unit root processes is spurious. In fact, cointegration analysis studies cases when this is not the case.

4 Testing for Unit-Roots

We will consider two simple cases frequently used in practice:

- *Case 1:* Test $Ho : \phi = 1$ against $H_A : |\phi| < 1$ in:

$$Y_t = \phi Y_{t-1} + u_t \tag{2}$$

In our examples, this corresponds to testing the random walk model against the alterniative that the process follows a stationary zero mean AR(1) process. Substracting Y_{t-1} in both sides we get:

$$\Delta Y_t = (\phi - 1) Y_{t-1} + u_t \tag{3}$$

Then, testing $H_o : \phi = 1$ in (2) is equivalent to testing $H_o : (\phi - 1) = 0$ in (3).

It is tempting to estimate g in:

$$\Delta Y_t = g Y_{t-1} + u_t$$

by OLS (note that we do not include a constant in the regression) and use a t-statistic to test $H_0 : g = 0$. We would reject the null hypothesis of the presence of a unit root when values of this statistic are significatively different from zero. The problem is that even when the u_t's are taken to be independently normally distributed, under the null hypothesis the t-statistic of such a regression *does not* have the usual t-distribution neither it has an asymptotic normal distribution. The distribution of this statistic for the case where the u_t's are uncorrelated was initially approximated by simulation by Dickey and Fuller (1976) and later obtained analytically by Phillips (1987). For practical purposes we can use Dickey-Fuller tables available in several econometrics books.

> *Example 1* (Hamilton p.489): Interest rate
> The goal is to test for a unit-root in the series of nominal interest rate I_t. The following OLS regression was estimated:
>
> $$\Delta I_t = \underset{(-0.29)}{-0.0031 I_{t-1}}$$
>
> The t-statistic lies above the 5% critical value according to the (one-tailed) Dickey-Fuller table (-1.95), so the null hypothesis that the interest rate follows a randow walk is not rejected by this test.

The inclusion of a constant (but not a trend) in this model serves several purposes. Under the null hypothesis, the process can be a random walk with drift (if the constant is different from zero) or a random walk (when the constant is zero). Under the alternative the model is always a stationary AR(1) with zero mean (constant equal to zero) or with mean different from zero.

- *Case 2:* Test $Ho : \phi = 1$ against $H_A : |\phi| < 1$ in:

$$Y_t = m + \phi Y_{t-1} + dt + \varepsilon_t \tag{4}$$

In our simple models, this corresponds to testing model 4 (random walk with drift) against the deterministic trend model. That the model under the null hypothesis corresponds to a deteriministic trend can be verified as follows.

> Rewrite (3) in terms of the lag operator:
>
> $$Y_t(1 - \phi L) = m + dt + \varepsilon_t$$
>
> Since $|\phi| < 1$, the operator $1/(1 - \phi L)$ is well defined, so:

$$Y_t = \frac{m}{1-\phi} + \frac{dt}{1-\phi L} + \frac{\varepsilon_t}{1-\phi L}$$

Now,

$$t/(1-\phi L) = \sum_{i=0}^{\infty} \phi^i (t-i) = \frac{t}{1-\phi} - \frac{\phi}{(1-\phi)^2}$$

replacing above:

$$Y_t = \mu^* + \frac{dt}{1-\phi} + \frac{\varepsilon_t}{1-\phi L}$$

where $\mu^* = [m(1-\phi) - d\phi]/(1-\phi)^2$. Given $|\phi| < 1$, $\varepsilon_t/(1-\phi L)$ is a stationary $MA(\infty)$ process, then Y_t has the specification of a deterministic trend.

As mentioned before, we are not testing for stationarity here since the model is non-stationary under both the null and the alternative hypothesis. Instead we are trying to determine the *source* of the non-stationarity. Substracting Y_{t-1} in both sides we get:

$$\Delta Y_t = m + (1-\phi)Y_{t-1} + dt + \varepsilon_t \tag{5}$$

and the testing procedure consists in testing $H_0 : g = 1$ in the OLS estimation of:

$$\Delta Y_t = m + gY_{t-1} + dt + \varepsilon_t$$

which now includes a constant and a time-trend.

4.1 The Augmented Dickey Fuller test

The original version of the Dickey-Fuller test assumes that the error terms are uncorrelated. When serial correlation is present the 'augmented Dickey-Fuller' version of the test proposes to include in the regression several lags of the difference of the series to account for the serial correlation. The test procedure is the same as described before but we include in the regression lags of the variable to account for possible serial correlation. A practical questions is how many lags to include. There are many strategies but one that is frequently used in practice is to start with several lags and estimate the Dickey-Fuller auxiliar regression and check the 't-statistic' of the estimated coefficients of the lagged differences and use this criterion to discard irrelevant lags in a 'general-to-particular' fashion.

Example 3: GDP (Hamilton p. 489)

Here the idea is to test for a unit-root in the GDP (Y_t). Economic theory suggest that this should be a growing non-stationary process a test of a unit root is a test to detect the source of non-stationarity. The following OLS regression was estimated with four lags of the GDP to account for possible serial correlation:

$$\begin{aligned}\Delta Y_t &= -.329\Delta Y_{t-1} + .209\Delta Y_{t-2} - .084\Delta Y_{t-3} \\ &\quad - .075\Delta Y_{t-4} + 35.92 - .051Y_{t-1} + .0378t\end{aligned}$$

The t-statistic corresponding to the variable Y_{t-1} is .051/ .0193 = -.26 which is greater than the Dickey-Fuller critical value at (one-tailed) 5% significante (-3.44). According to this, we do not reject the null hypothesis that the GPD follows a random walk.

5 Cointegration

5.1 Basic Concepts

A series Y_t is *integrated of order d* (denoted $I(d)$) if it *must* be differenced at least d times in order to make it stationary.

Examples:

1. The random walk is $I(1)$ since

$$Y_t = Y_{t-1} + \varepsilon_t \Longrightarrow \Delta Y_t = e_t$$

which is stationary.

2. Any stationary process is trivially $I(0)$

Two series Y_t and X_t are said to be *cointegrated* if:

1. Both are $I(d)$, $d \neq 0$ and the same for both series.

2. There is a linear combination of them that is $I(0)$, that is, there exist $a = (a_1, a_2)$ non zero such that:

$$a_1 Y_t + a_2 X_t \text{ is } I(0)$$

the vector a is called the cointegrating vector.

5.2 Discussion

- Cointegration refers to a relationship between non-stationary, unit-root processes. The existence of a cointegration relationship between two variables has the following economic intuition. When two series are cointegrated it suggests that even though both processes are non-stationary, there is some long-run equilibrium relationship linking both series so that relationship is stationary. This long run relationship is represented by the linear combination implicit in the cointegration relationship. Economic theory often suggests the existence of such relationships: the consumption function, the PPP, the theory of the demand for money, etc.

- There is a basic non-uniqueness problem. If a is a cointegrating vector, then ba is also a contegrating vector, where b is any non-zero scalar. Then it is customary to 'normalize' the cointegrating vector by imposign the condition $a_1 = 1$. The 'normalized' cointegrating vector will be $(1, a2)$ and in this setup a_2 is sometimes refferd to as the cointegration coefficient.

- Going back to the *spurious regression* problem, with the concepts introduced in this section it is more clear to see why regressions between two non-stationary unit-root processes are likely to produce spurious results. When two series Y_t and X_t are $I(1)$ and we impose the classical linear model structure between them:

  $$Y_t = \alpha + \beta X_t + u_t$$

  where u_t is a stationary $I(0)$ random term, this makes sense only if the series are cointegrated, that is, when there exists a linear combination of the $I(1)$ series that is $I(0)$.

- Graphical example here -

5.3 Testing for cointegration in the bivariate case

We will explore what is called the 'Engle-Granger residual approach'. Here the idea is to get an estimate of the cointegration coefficient and then test whether the linear combination formed with this coefficient is $I(0)$.

We proceed in three steps:

1. *Step 1:* Test that both variables have the same order of integration, say, that they are both $I(1)$. This can be performed with the unit-root tests described before.

2. *Step 2:* Estimate a 'long-run relationship':

$$Y_t = \hat{a} + \hat{b} X_t$$

   by OLS. Remember that we concluded that the classical analysis of a series of this sort is invalid unless the series are cointegrated. Engle and Granger (1987) showed that when this is the case, OLS in this equation yields a consistent estimate of the cointegrating vector. Note that if the series are cointegrated and that if $\hat{a}$, and $\hat{b}$ are 'good' estimates of the cointegrating coefficients, then $Y_t - \hat{a} - \hat{b} X_t$ should be $I(0)$. The latter are, by definition, the residuals of the OLS regression of Y_t on X_t and a constant, then we can use these residuals to test for cointegration.

3. Extract the residuals of this regression (e_t) and test for a unit-root in this series:

$$e_t \text{ has a unit root} \Longrightarrow \text{reject cointegration}$$

   Note that according to the way this test is specified, this is a test of *no-cointegration*: acceptance of a unit root in the residuals suggest that the residual term is non-stationary, which implies rejection of cointegration. This may cause some confusion since *acceptance* of the null of a unit root in the residuals suggest *rejection* of cointegration.

5.4 Discussion

1. Try to compare what we are doing here with the specification tests in classical econometrics. Even though the estimation stage proceeds in the same way (OLS) we are interested in different results. In the case of cointegration analysis we are testing for stationary residuals while in the context of the classical linear model we test for white noise residuals.

2. We do not observe the cointegrating relationship directly. Instead we use an estimate of it (the residuals of the estimation). Then, the procedure of testing for a unit root in this series remains the same but the usual Dickey-Fuller table is no longer valid. We have to use a specific table for this case. The same considerations apply regarding the use of the augmented version of the Dickey-Fuller tests.

3. To obtain an estimate of the cointegrating vector we regressed one variable on the other, but we could have inverted the order of the regression and have obtained a different result (unless, of course, $R^2 = 1$). There are some cases where this can produce contradictory results in terms of the test for cointegration. Fortunately, a multivariate approach handles this problem.

5.5 The multivariate case

We extend the analysis to more than two varibles (like in the money demand example).

A vector Y_t of n time series is said to be cointegrated if each of the series taken individually is $I(d)$ while some linear combination of the series $a'Y_t$ is $I(0)$ for some non-zero vector a. This vector is called the cointegrating vector.

Now the problem of non-uniqueness persists after the normalization. It can be shown that there might be at most $n-1$ linearly independent cointegrating vectors. Then, any linear combination of these vectors is also a cointegrating vectors. Hence, now the problem is to find a basis of the space of cointegrating vectors.

There are several interesting aspects to test in the multivariate case. The Johansen procedure (1990) is a multivariate extension of the Dickey Fuller test. A relevant question is not only if there is cointegration but how many cointegrating relationships there are.

6 Suggested Bibliography

The literature on the topic is enourmous and constantly updated. Here are some recent references.

The corresponding chapters of books like Greene (2000), Davidson and MacKinnon (1993) or Hendry (1995) can give a quick overview of the topic.

Enders (1995) is a general introduction at an accesible level with abundant empirical examples.

Banerjee, Dolado, Galbraith and Hendry (1993), Maddala and Kim (1999) and Hamilton (1994) provide a detailed book-length treatement at a graduate level.

The survey articles by Stock (1994) and Watson (1994) are useful sources for advanced reading. Francisco Cribari (1995) has a good overview of current research on the subject.

A humourous (and very illustrative) example is Murray (1994).

References

Banerjee, A., Dolado, J., Galbraith, J. and Hendry, D. (1993) 'Co-integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data', Oxford University Press.

Cribari-Neto, F. (1995), *On Time Series Econometrics*, mimeo, University of Illinois.

Davidson, R. and MacKinnon, J. (1993), *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.

Enders, W. (1993), *Applied Econometric Time Series*, Wiley & Sons Inc., New York.

Greene, W. (2000), *Econometric Analysis*, 4th edition, Prentice Hall, New York.

Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press, Princeton NJ.

Hendry, D. (1995), *Dynamic Econometrics*, Oxford University Press, Oxford.

Maddala, G.S, and Kim, I. (1999), *Unit Roots, Cointegration, and Structural Change*, Cambridge University Press, Cambridge.

Murray, M. (1994), A Drunk and Her Dog: An Illustration of Cointegration and Error Correction, *The American Statistician*, February, Vol.48, No. 1.

Stock, J. (1994) Unit Roots, Structural Breaks and Trends, in Engle, R. and McFadden, D. (eds.) *Handbook of Econometrics* Vol.4, North Holland.

Watson, M. (1994) Vector Autorregressions and Cointegration, in Engle, R. and McFadden, D. (eds.) *Handbook of Econometrics* Vol.4, North Holland.