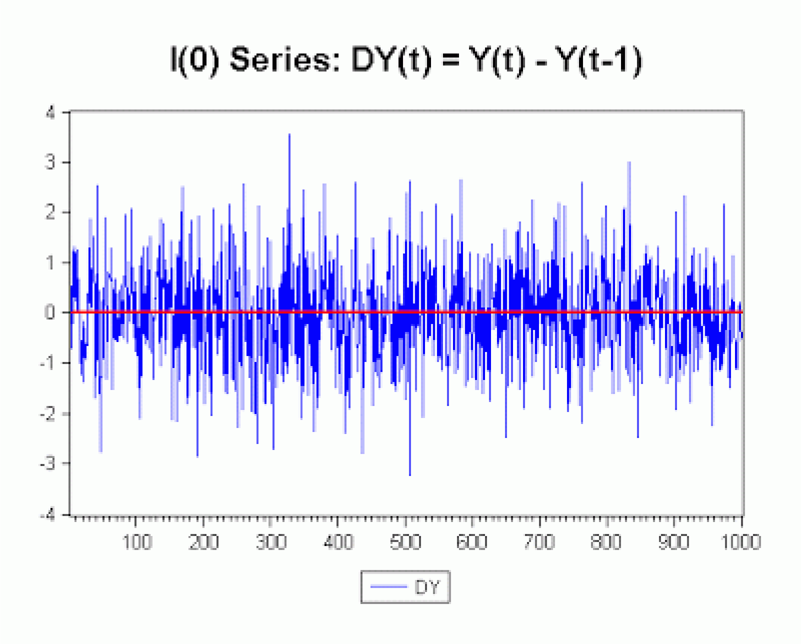
Chapter 8

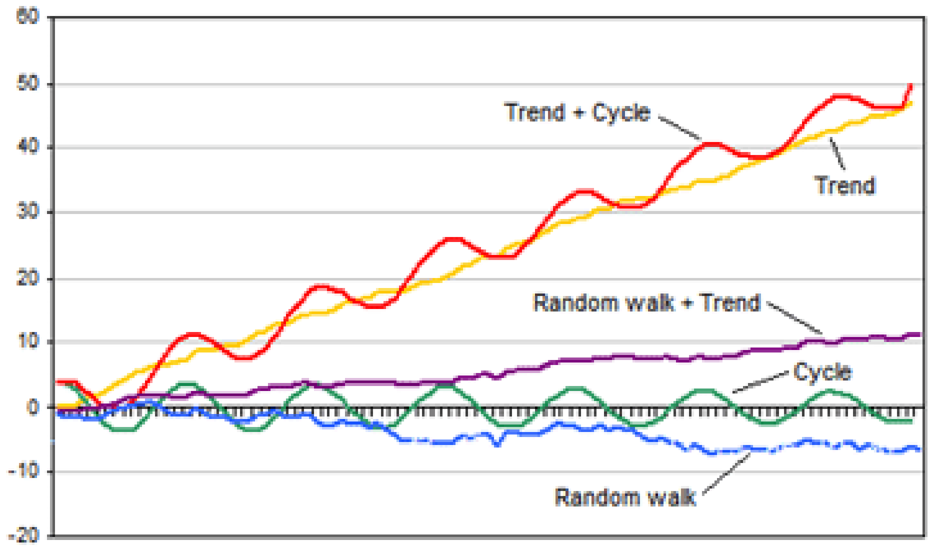
Modeling long-run relationship in finance

Stationary process



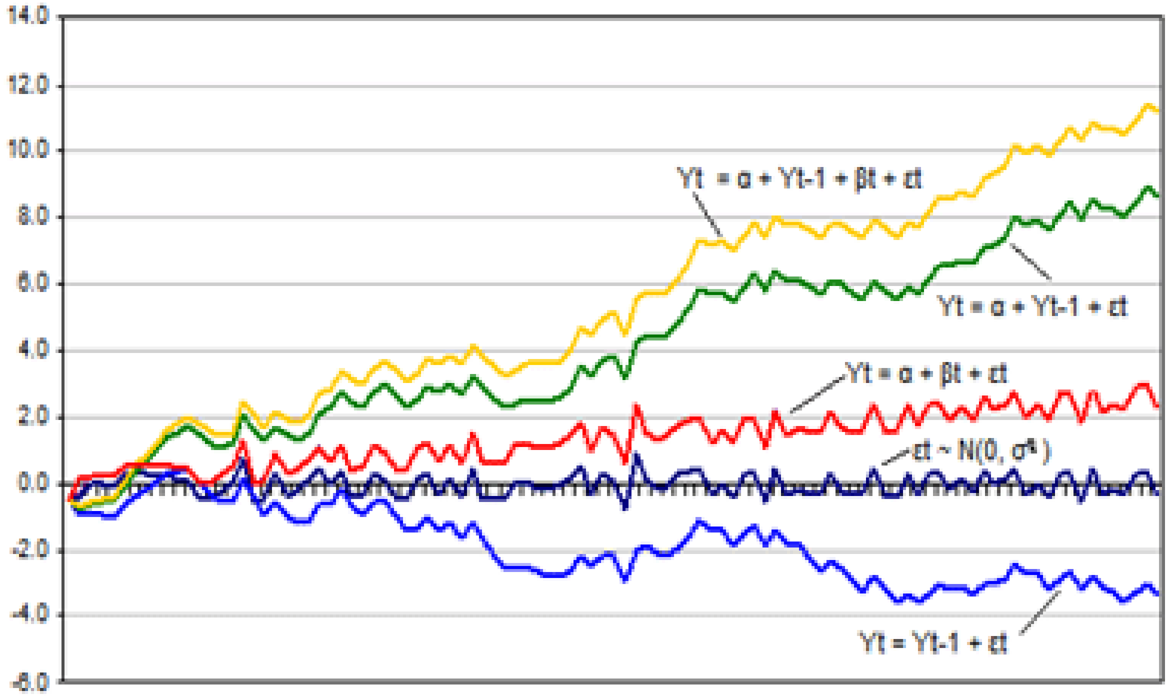
Non-stationary process

Table 1 Non-stationary behavior



Non-stationary process

Table 2 Non-stationary processes



1. Stationarity and Unit Root Testing

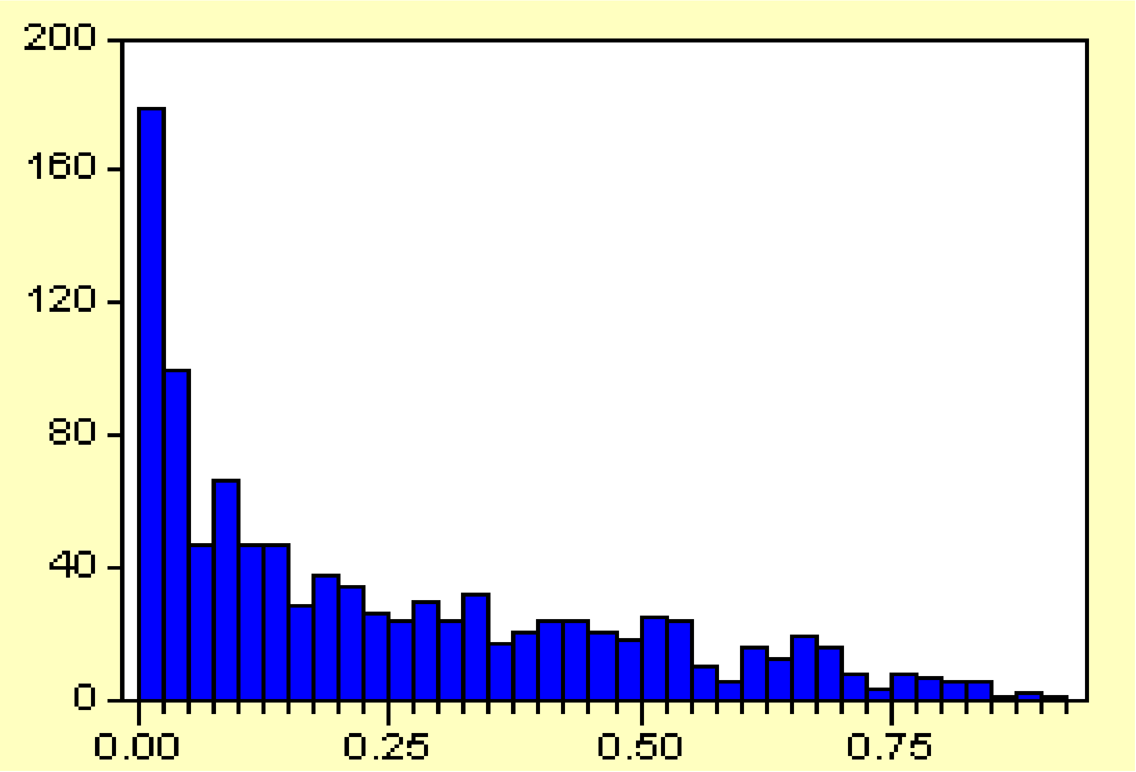
Why do we need to test for Non-Stationarity?

- The **stationarity** or otherwise of a series can strongly **influence its behaviour and properties** - e.g. persistence of shocks will be infinite for non-stationary series

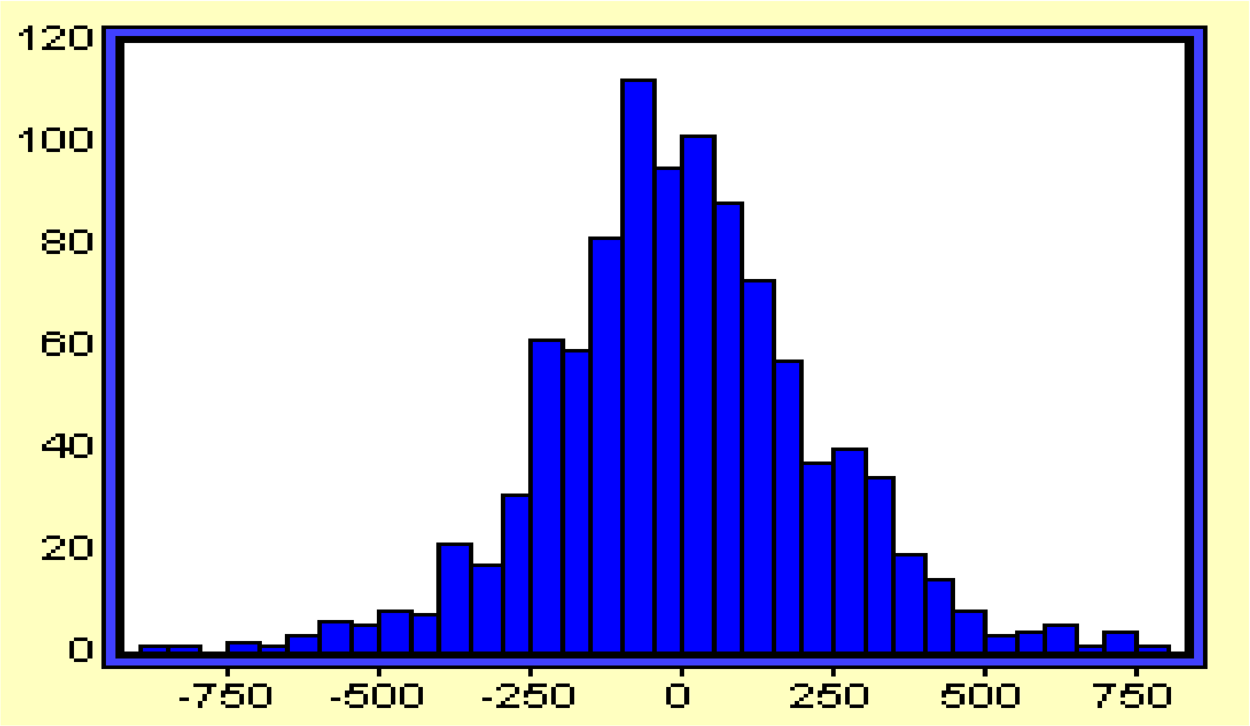
- **Spurious regressions**. If two variables are trending over time, a regression of one on the other could have a **high R^2 even if the two variables are totally unrelated**

- If the variables in the regression model are **not stationary**, then the standard assumptions for asymptotic analysis will not be valid. The **usual "*t*-ratios" will not follow a *t*-distribution**, so we cannot validly undertake hypothesis tests about the regression parameters.

Value of R^2 for 1000 Sets of Regressions of a Non-stationary Variable on another Independent Non-stationary Variable



Value of *t*-ratio on Slope Coefficient for 1000 Sets of Regressions of a Non-stationary Variable on another Independent Non-stationary Variable



Two types of Non-Stationarity

- Various definitions of non-stationarity exist
- In this chapter, we are really referring to the weak form or covariance stationarity
- There are two models which have been frequently used to characterise non-stationarity: the **random walk model with drift** (stochastic non-stationarity):

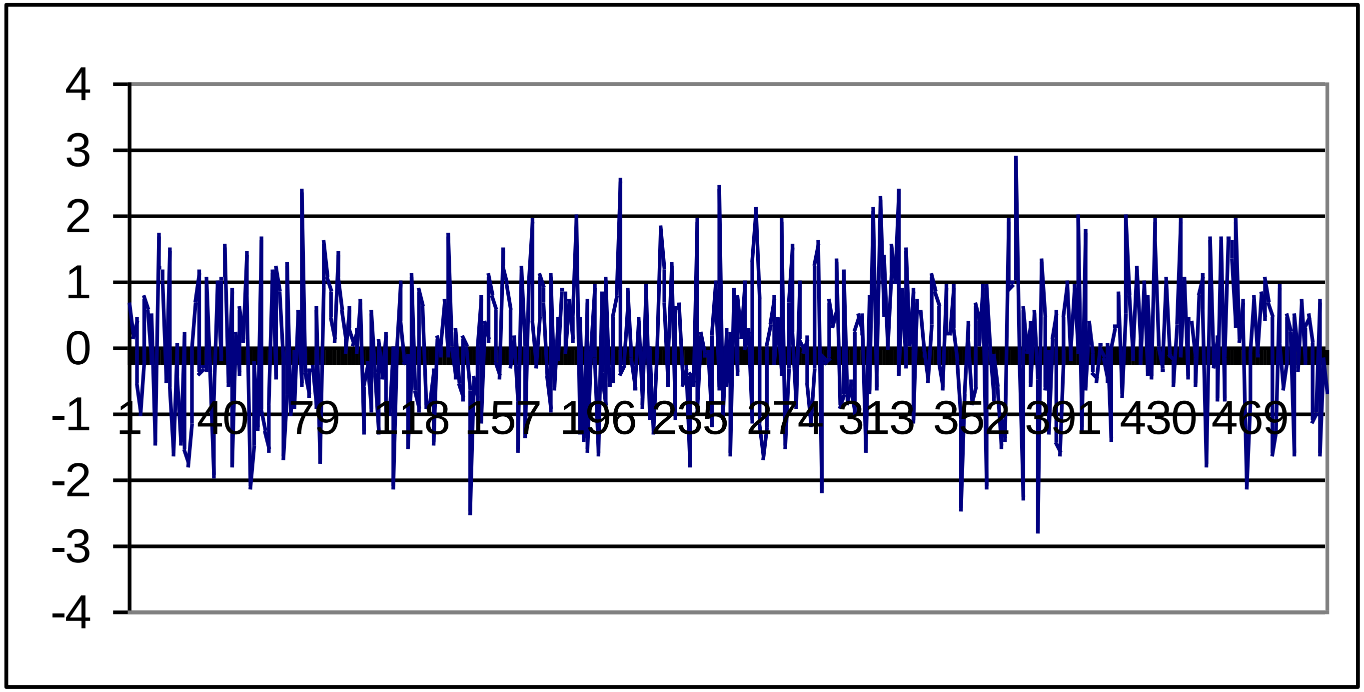
$$y_t = \mu + y_{t-1} + u_t \qquad (1)$$

and the **deterministic trend process** (deterministic non-stationarity):

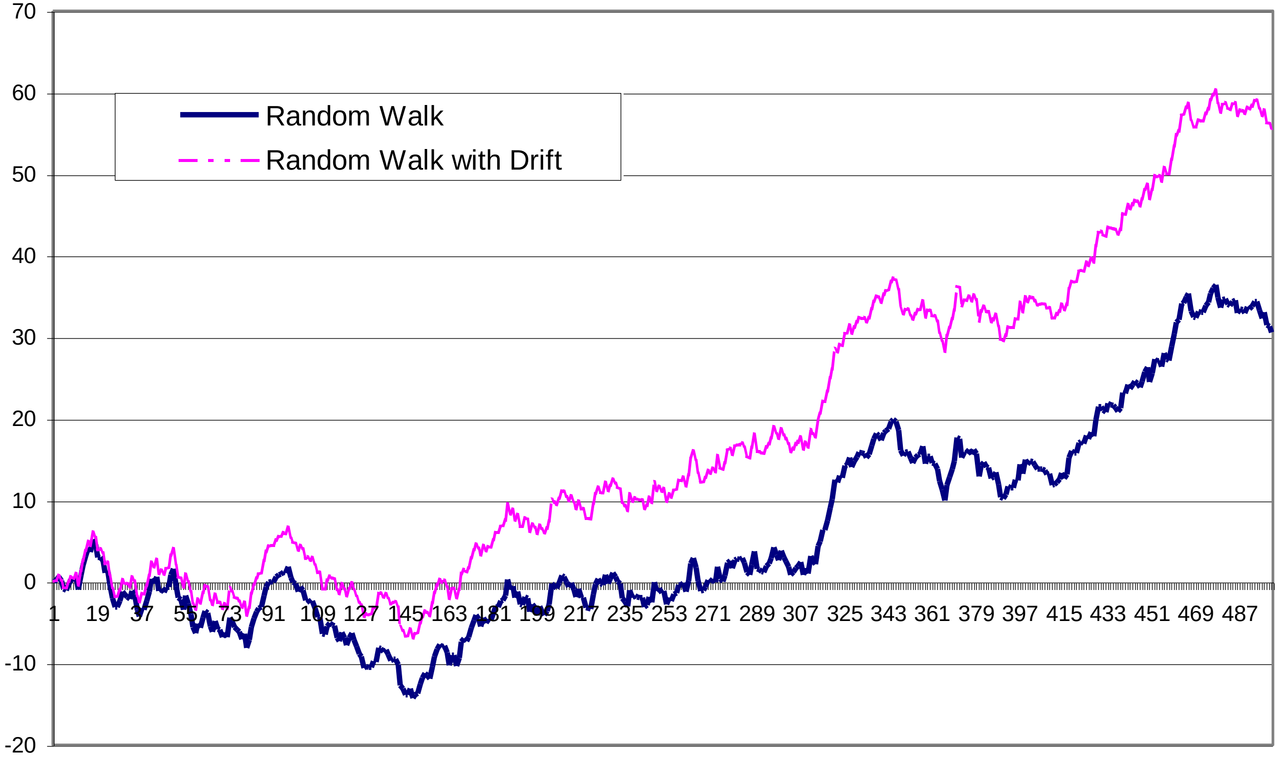
$$y_t = \alpha + \beta t + u_t \qquad (2)$$

where u_t is a white noise disturbance term.

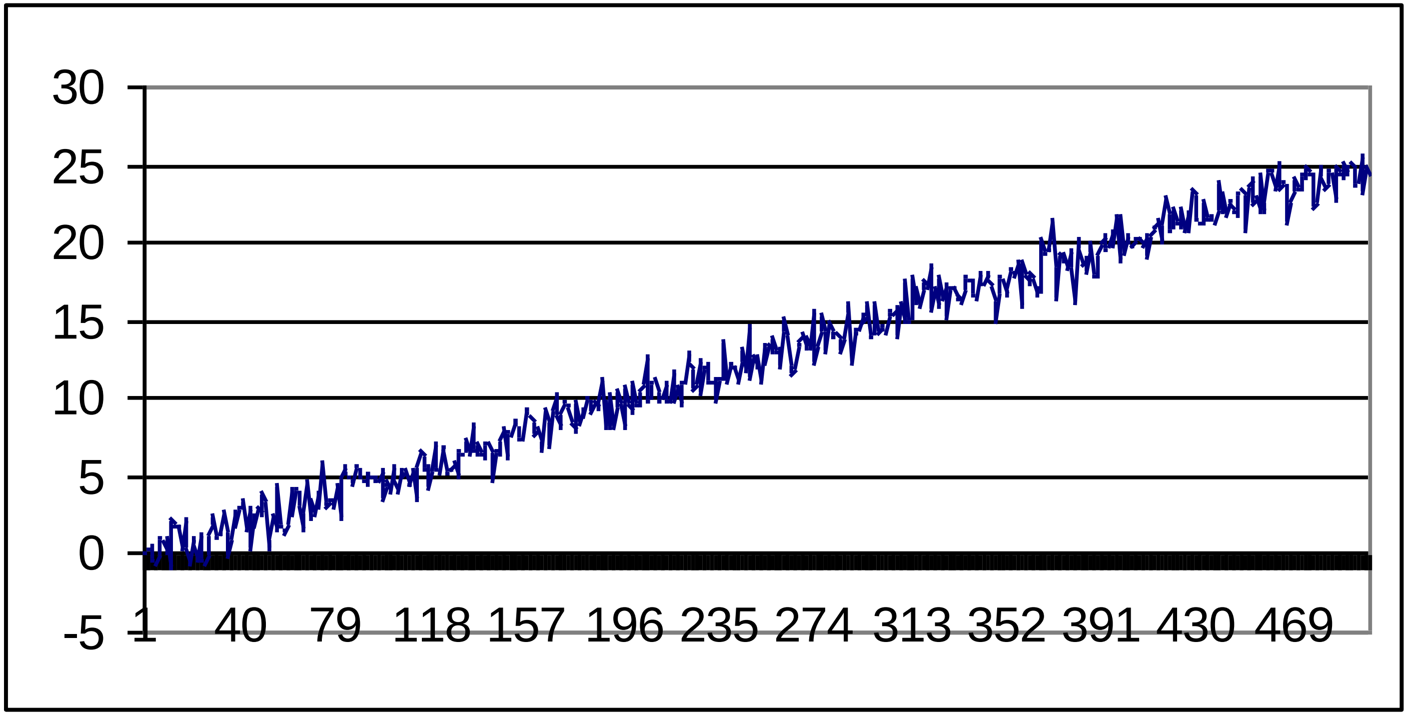
Sample Plots for various Stochastic Processes: A White Noise Process



Sample Plots for various Stochastic Processes: A Random Walk and a Random Walk with Drift



Sample Plots for various Stochastic Processes: A Deterministic Trend Process



Stochastic Non-Stationarity (Random walk)

- Note that the model (1) could be generalised to the case where y_t is an **explosive process**:

$$y_t = \mu + \phi y_{t-1} + u_t$$

where $\phi > 1$.

- Typically, this case is ignored and we use $\phi = 1$ to characterise the non-stationarity because
  - $\phi > 1$ does not describe many data series in economics and finance.
  - $\phi > 1$ has an intuitively unappealing property: shocks to the system are not only persistent through time, they are propagated so that a given shock will have an increasingly large influence.

Stochastic Non-stationarity: The Impact of Shocks

- To see this, consider the general case of an **AR(1) with no drift**:

$$y_t = \phi y_{t-1} + u_t \qquad (3)$$

Let ϕ take any value for now.

- We can write: $y_{t-1} = \phi y_{t-2} + u_{t-1}$

$$y_{t-2} = \phi y_{t-3} + u_{t-2}$$

- Substituting into (3) yields: $y_t = \phi(\phi y_{t-2} + u_{t-1}) + u_t$

$$= \phi^2 y_{t-2} + \phi u_{t-1} + u_t$$

- Substituting again for y_{t-2}: $y_t = \phi^2(\phi y_{t-3} + u_{t-2}) + \phi u_{t-1} + u_t$

$$= \phi^3 y_{t-3} + \phi^2 u_{t-2} + \phi u_{t-1} + u_t$$

- Successive substitutions of this type lead to:

$$y_t = \phi^T y_0 + \phi u_{t-1} + \phi^2 u_{t-2} + \phi^3 u_{t-3} + \ldots + \phi^T u_0 + u_t$$

The Impact of Shocks for Stationary and Non-stationary Series

- We have **3 cases:**

  1. $\phi<1 \Rightarrow \phi^T \rightarrow 0$ as $T \rightarrow \infty$

  So the **shocks to the system gradually die away**.

  2. $\phi=1 \Rightarrow \phi^T = 1 \forall\ T$

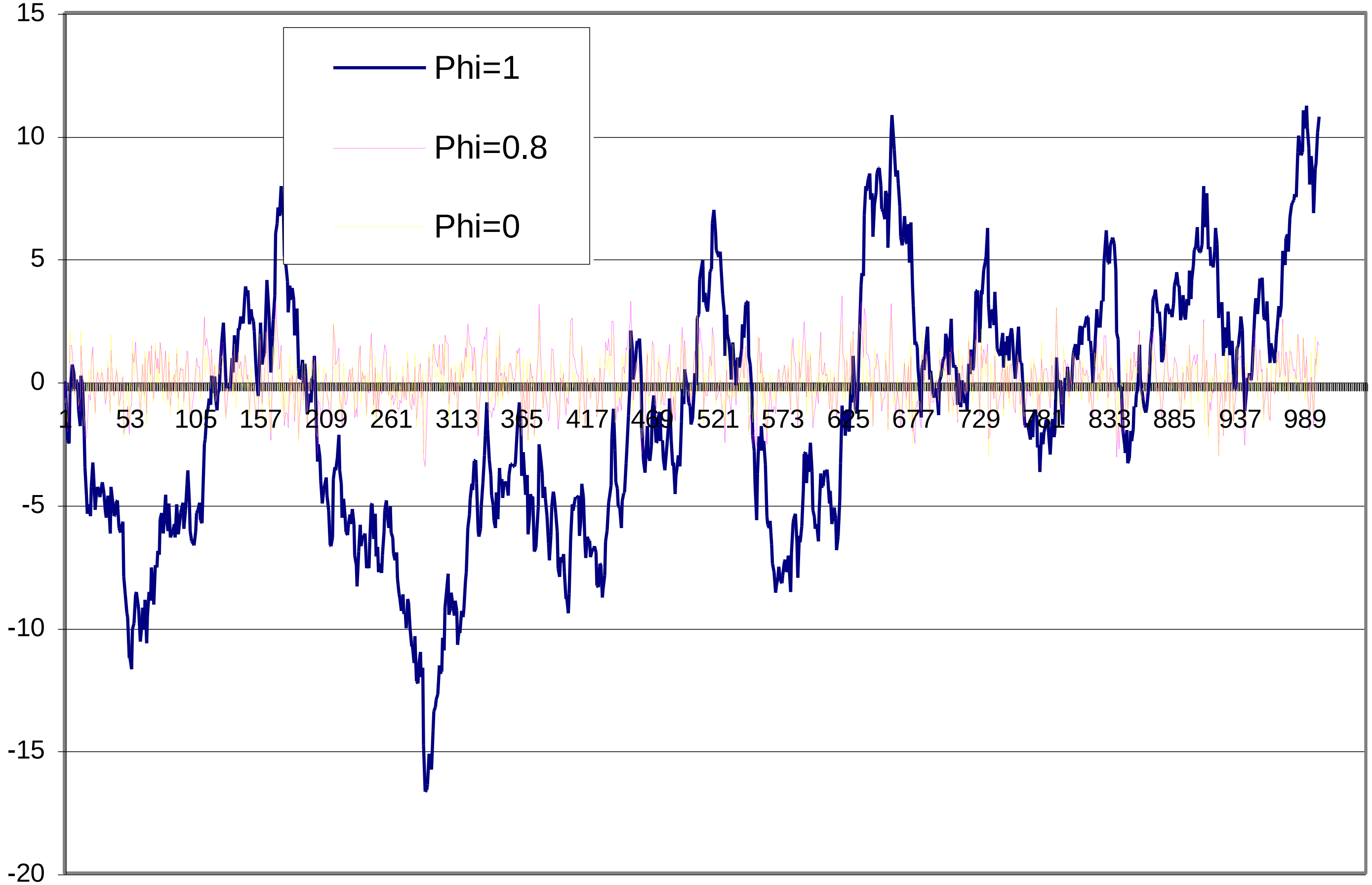
  So **shocks persist in the system and never die away**. We obtain:

$$y_t = y_0 + \sum_{i=0}^{\infty} u_t \quad \text{as } T \rightarrow \infty$$

  So just an infinite sum of past shocks plus some starting value of y_0.

  3. $\phi>1$. Now given **shocks become <u>more</u> influential as time goes on**, since if $\phi>1$, $\phi^3>\phi^2>\phi$ etc.

Autoregressive Processes with differing values of ϕ (0, 0.8, 1)



Detrending a Stochastically Non-stationary Series

- Going back to our 2 characterisations of non-stationarity, the r.w. with drift:

$$y_t = \mu + y_{t-1} + u_t \qquad (1)$$

and the trend-stationary process

$$y_t = \alpha + \beta t + u_t \qquad (2)$$

- The two will **require different treatments to induce stationarity**. The second case is known as **deterministic non-stationarity and what is required is detrending**. The first case is known as **stochastic non-stationarity**, if we let

$$\Delta y_t = y_t - y_{t-1}$$

and $\qquad L\, y_t = y_{t-1}$

so $\qquad (1-L)\, y_t = y_t - L\, y_t = y_t - y_{t-1}$

If we take (1) and subtract y_{t-1} from both sides:

$$y_t - y_{t-1} = \mu + u_t$$

$$\Delta y_t = \mu + u_t$$

We say that we have induced stationarity by "**differencing once**".

Detrending a Series: Using the Right Method

- Although trend-stationary and difference-stationary series are both “trending” over time, the **correct approach needs to be used in each case**.

- Note: **If we first difference the trend-stationary series**, it would “remove” the non-stationarity, but at the expense on introducing an MA(1) structure into the errors.

- Note: **Conversely if we try to detrend a series which has stochastic trend**, then we will not remove the non-stationarity.

- We will now concentrate on the **stochastic non-stationarity model** which is more common to describe most series in economics or finance.

Definition of Non-Stationarity

- Consider again the simplest stochastic trend model:

$$y_t = y_{t-1} + u_t$$

or

$$\Delta y_t = u_t$$

- We can generalise this concept to consider the case where the series contains more than one "unit root". That is, we would need to apply the first difference operator, Δ, more than once to induce stationarity.

<u>Definition</u>

If a non-stationary series, y_t must be differenced d times before it becomes stationary, then it is said to be **integrated of order *d***. We write $y_t \sim I(d)$.

So if $y_t \sim I(d)$ then $\Delta^d y_t \sim I(0)$.

An **I(0) series is a stationary series**

An **I(1) series contains one unit root**,

e.g. $y_t = y_{t-1} + u_t$

Characteristics of I(0), I(1) and I(2) Series

- An **I(2) series contains two unit roots** and so would require differencing twice to induce stationarity.
- I(1) and I(2) series can wander a long way from their mean value and cross this mean value rarely.
- I(0) series should cross the mean frequently.
- **Note**: **The majority of economic and financial series contain a single unit root**, although some are stationary and consumer prices have been argued to have 2 unit roots.

How do we test for a unit root?

- The early and pioneering work on testing for a unit root in time series was done by **Dickey and Fuller** .
- The basic objective of the test is to test the null hypothesis that $\phi = 1$ in:

$$y_t = \phi y_{t-1} + u_t$$

against the one-sided alternative $\phi < 1$. So we have

H_0: series contains at least a unit root (ϕ=1)

vs. **H_1: series is stationary (ϕ<1)**

- We usually use the regression:

$$\Delta y_t = \psi y_{t-1} + u_t$$

so that a test of ϕ=1 is equivalent to a test of ψ=0 (since ϕ-1=ψ).

Different forms for the DF Test Regressions

- **Dickey Fuller tests are also known as τ tests: τ, τ_{μ}, τ_{τ}.** (page 623)

- The null (H_0) and alternative (H_1) models in each case are

i) H_0: $y_t = \phi y_{t-1} + u_t$ with $\phi=1$
H_1: $y_t = \phi y_{t-1} + u_t$ with $\phi<1$
This is a **test for a random walk against a stationary autoregressive process of order one (AR(1))**

ii) H_0: $y_t = \phi y_{t-1} + \mu + u_t$ with $\phi=1$
H_1: $y_t = \phi y_{t-1} + \mu + u_t$ with $\phi<1$
This is a **test for a random walk against a stationary AR(1) with drift.**

iii) H_0: $y_t = \phi y_{t-1} + \mu + \lambda t + u_t$ with $\phi=1$
H_1: $y_t = \phi y_{t-1} + \mu + \lambda t + u_t$ with $\phi<1$
This is a **test for a random walk against a stationary AR(1) with drift and a time trend.**

Computing the DF Test Statistic

- We can write

  $$\Delta y_t = \mu + \lambda t + u_t$$

  and the alternatives may be expressed as

  $$\Delta y_t = \psi y_{t-1} + \mu + \lambda t + u_t$$

  with $\mu = \lambda = 0$ in case i), and $\lambda = 0$ in case ii) and $\psi = \phi - 1$. In each case, the tests are based on the *t*-ratio on the y_{t-1} term in the estimated regression of Δy_t on y_{t-1}, plus a constant in case ii) and a constant and trend in case iii). The test statistics are defined as

  $$\text{test statistic} = \frac{\hat{\psi}}{\hat{SE}(\hat{\psi})}$$

- The test statistic does not follow the usual *t*-distribution under the null, since the null is one of non-stationarity, but rather follows a non-standard distribution.

Critical Values for the DF Test

Significance level	10%	5%	1%
C.V. for constant but no trend	-2.57	-2.86	-3.43
C.V. for constant and trend	-3.12	-3.41	-3.96

Table 4.1: Critical Values for DF and ADF Tests (Fuller, 1976, p373).

The **null hypothesis of a unit root is rejected** in favour of the stationary alternative in each case **if the test statistic is more negative than the critical value (left-tailed test)**

The Augmented Dickey Fuller (ADF) Test

- The tests above are only valid if u_t is white noise. In particular, u_t will be autocorrelated if there was autocorrelation in the dependent variable of the regression (Δy_t) which we have not modelled. The solution is to "augment" the test using p lags of the dependent variable. The alternative model in case (i) is now written:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t$$

- The same critical values from the DF tables are used as before. A problem now arises in determining the optimal number of lags of the dependent variable.

  There are 2 ways:

  - use the frequency of the data to decide (12 lags if monthly, 4 lags if quarterly)

  - use information criteria

Testing for Higher Orders of Integration

- Consider the simple regression:

$$\Delta y_t = \psi y_{t-1} + u_t$$

  We test **H$_0$: ψ=0** (at least 1 unit root) vs. **H$_1$: ψ<0** (stationary).
- If H$_0$ is rejected we simply conclude that y_t is stationary.
- But what do we conclude if H$_0$ is not rejected? The series contains a unit root, but is that it? No! What if y_t~I(2)? We would still not have rejected. So we now need to test

  H$_0$: y_t~I(2) (at least 2 unit roots) vs. **H$_1$: y_t~I(1)** (1 unit root)

  We would continue to test for a further unit root until we rejected H$_0$.
- We now regress $\Delta^2 y_t$ on Δy_{t-1} (plus lags of $\Delta^2 y_t$ if necessary).
- Now we test H$_0$: Δy_t~I(1) which is equivalent to H$_0$: y_t~I(2).
- So in this case, if we do not reject (unlikely), we conclude that y_t is at least I(2).

Criticism of Dickey-Fuller test

- Main criticism is that the **power of the tests is low if the process is stationary but with a root close to the non-stationary boundary**.

  e.g. the tests are poor at deciding whether

  $$\phi=1 \text{ or } \boldsymbol{\phi=0.95},$$

  especially with small sample sizes.

3. Cointegration: An Introduction

- **In most cases**, if we combine two variables which are I(1), then the combination will also be I(1).

- **Sometimes**, the combination will be I(0) → **cointegration**.

Definition of Cointegration (Engle & Granger, 1987)

- Many time series are non-stationary but "**move together**" over time.

- Let $\boldsymbol{z}_t$ be a $k\times 1$ vector of variables, then the **components of $\boldsymbol{z}_t$ are cointegrated of order (1,1)** if

  i) All components of $\boldsymbol{z}_t$ are **I(1)**

  ii) There is at least one vector of coefficients α such that $\alpha' z_t \sim$ **I(0)**

- If variables are cointegrated, it means that a linear combination of them will be stationary.

- There may be up to r linearly independent cointegrating relationships (where $r \leq k$-1), also known as **cointegrating vectors**. r is also known as the **cointegrating rank** of z_t.

- A **cointegrating relationship** may also be seen as a **long term relationship**.

Cointegration and Equilibrium

- **Examples** of possible Cointegrating Relationships in finance:
  - spot and futures prices
  - ratio of relative prices and an exchange rate
  - equity prices and dividends

- Market forces arising from **no arbitrage conditions** should ensure an equilibrium relationship.

- *Note: No cointegration implies that series could wander apart without bound in the long run.*

4. Equilibrium Correction Models or Error Correction Models (ECM)

- When the concept of non-stationarity was first considered, a usual response was to independently take the first differences of a series of I(1) variables.

- The problem with this approach is that pure first difference models have no long run solution.

  e.g. Consider y_t and x_t both I(1).

  The model we may want to estimate is

  $$\Delta y_t = \beta \Delta x_t + u_t$$

  But this collapses to nothing in the long run.

- The definition of the long run that we use is where

  $$y_t = y_{t-1} = y; \; x_t = x_{t-1} = x.$$

- Hence all the difference terms will be zero, i.e. $\Delta y_t = 0; \; \Delta x_t = 0$.

Specifying an ECM

- One way to get around this problem is to use both first difference and levels terms, e.g.

$$\Delta y_t = \beta_1 \Delta x_t + \beta_2 (y_{t-1} - \gamma x_{t-1}) + u_t \qquad (2)$$

- $y_{t-1} - \gamma x_{t-1}$ is known as the **error correction term**.

- Providing that y_t and x_t are cointegrated with cointegrating coefficient γ, then $(y_{t-1} - \gamma x_{t-1})$ will be I(0) even though the constituents are I(1).

- We can thus validly use OLS on (2).

- The **Granger representation theorem** shows that any cointegrating relationship can be expressed as an equilibrium correction model.

5. Testing for Cointegration in Regression: Residual Based Approach

- The model for the **equilibrium correction term** can be generalised to include more than two variables:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \ldots + \beta_k x_{kt} + u_t \qquad (3)$$

- **u_t should be I(0) if the variables y_t, x_{2t}, ... x_{kt} are cointegrated.**

- **So what we want to test is the residuals of equation (3) to see if they are non-stationary or stationary**. We can use the **DF / ADF test on u_t**.

  So we have the regression

$$\Delta\hat{u}_t = \psi\hat{u}_{t-1} + v_t$$

- However, since this is a test on the residuals of an actual model, $\hat{u}_t$, then the critical values are changed.

Testing for Cointegration in Regression: Conclusions

- Engle and Granger (1987) have tabulated a new set of critical values and hence the test is known as the **Engle Granger (E.G.) test**.

- What are the null and alternative hypotheses for a test on the residuals of a potentially cointegrating regression?

H_0: unit root in cointegrating regression's residuals (not cointegrated)

H_1: residuals from cointegrating regression are stationary (cointegrated)

6. Methods of Parameter Estimation in Cointegrated Systems: The Engle-Granger Approach

- There are (at least) 3 methods we could use: Engle Granger, Engle and Yoo, and Johansen.
- **The Engle Granger 2 Step Method**

  This is a single equation technique which is conducted as follows:

  Step 1:

  - Make sure that all the individual variables are I(1).

  - Then estimate the cointegrating regression (3) using OLS.

  - Save the residuals of the cointegrating regression, $\hat{u}_t$.

  - Test these residuals to ensure that they are I(0).

  Step 2:

  - Use the step 1 residuals as one variable in the error correction model e.g.

$$\Delta y_t = \beta_1 \Delta x_t + \beta_2 (\hat{u}_{t-1}) + u_t$$

and we can test hypothesis about the parameters, forecast the values of y

The Engle-Granger Approach: Some Drawbacks

This method suffers from a number of problems:

1. Unit root and cointegration tests can have **low power in finite samples**

2. We are forced to **treat the variables asymmetrically** and to specify one as the dependent and the other as independent variables.

3. **Cannot perform any hypothesis** tests about the actual cointegrating relationship estimated at stage 1.

- Problem 1 is a small sample problem that should disappear **asymptotically**.
- Problem 2 is addressed by the **Johansen approach**.
- Problem 3 is addressed by the Engle and Yoo approach or the **Johansen approach**.

8. Testing for and Estimating Cointegrating Systems Using the Johansen Technique Based on VARs

- To use Johansen's method, we need to turn the VAR of the form

$$\underset{g\times 1}{y_t} = \underset{g\times g}{\beta_1}\ \underset{g\times 1}{y_{t-1}} + \underset{g\times g}{\beta_2}\ \underset{g\times 1}{y_{t-2}} + \ldots + \underset{g\times g}{\beta_k}\ \underset{g\times 1}{y_{t-k}} + \underset{g\times 1}{u_t}$$

(where y is a set of g variables that are I(1) which might be co-integrated), into a VECM, which can be written as

$$\Delta y_t = \Pi\, y_{t\text{-k}} + \Gamma_1\, \Delta y_{t-1} + \Gamma_2\, \Delta y_{t-2} + \ldots + \Gamma_{k-1}\, \Delta y_{t-(k-1)} + u_t$$

where $\Pi = (\sum_{j=1}^{k} \beta_i) - I_g$ and $\Gamma_i = (\sum_{j=1}^{i} \beta_j) - I_g$

Π is a **long run coefficient matrix** since all the $\Delta y_{t-i} = 0$.

Review of Matrix Algebra necessary for the Johansen Test

- Let Π denote a $g \times g$ square matrix and let c denote a $g \times 1$ non-zero vector, and let λ denote a set of scalars.

- λ is called a **characteristic root** or set of roots of Π if we can write

$$\underset{g\times g}{\Pi}\ \underset{g\times 1}{c} = \lambda\ \underset{g\times 1}{c}$$

- We can also write

$$\Pi\, c = \lambda\, I_p\, c$$

and hence

$$(\Pi - \lambda I_g)\, c = 0$$

where I_g is an identity matrix.

Review of Matrix Algebra (cont'd)

- Since $c \neq 0$ by definition, then for this system to have zero solution, we require the matrix ($\Pi - \lambda I_g$) to be singular (i.e. to have zero determinant).

$$|\Pi - \lambda I_g| = 0$$

- For example, let Π be the 2×2 matrix $\Pi = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix}$

- Then the characteristic equation is

$$|\Pi - \lambda I_g| = \left\| \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\| = 0$$

$$= \begin{vmatrix} 5-\lambda & 1 \\ 2 & 4-\lambda \end{vmatrix} = (5-\lambda)(4-\lambda) - 2 = \lambda^2 - 9\lambda + 18$$

Review of Matrix Algebra (cont'd)

- This gives the solutions $\lambda = 6$ and $\lambda = 3$.

- The characteristic roots are also known as **eigenvalues**.

- The rank of a matrix is equal to the number of linearly independent rows or columns in the matrix

  = the order of the largest square matrix we can obtain from Π which has a non-zero determinant

  = number of its characteristic roots (eigenvalues) that are different from zero.

-

- For example, the determinant of Π above $\neq 0$, therefore it has rank 2.

The Johansen Test and Eigenvalues

- Some properties of the eigenvalues of any square matrix A:

  1. the sum of the eigenvalues is the trace (sum of elements on main diagonal)
  2. the product of the eigenvalues is the determinant
  3. the number of non-zero eigenvalues is the rank

- Returning to Johansen's test, the VECM representation of the VAR was

$$\Delta y_t = \Pi\, y_{t-1} + \Gamma_1\, \Delta y_{t-1} + \Gamma_2\, \Delta y_{t-2} + \ldots + \Gamma_{k-1}\, \Delta y_{t-(k-1)} + u_t$$

- The test for cointegration between the y's is calculated by looking at the rank of the Π matrix via its eigenvalues.

The Johansen Test and Eigenvalues (cont'd)

- The eigenvalues denoted by λ_i are put in order:

$$1 > \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_g > 0$$

Note: If $\lambda_i = 0$, $\ln(1-\lambda_i) = 0$

If λ_i is non-zero and <1, then $\ln(1-\lambda_i) < 0$.

The Johansen Test Statistics

- The test statistics for cointegration are formulated as

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{g} \ln(1 - \hat{\lambda}_i)$$

and

$$\lambda_{\max}(r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1})$$

where $\hat{\lambda}_i$ is the estimated value for the ith ordered eigenvalue from the Π matrix, r is the number of cointegrating vectors under H_o

λ_{trace} tests the null that the number of cointegrating vectors is less than or equal to r against an alternative of more than r.

$\lambda_{trace} = 0$ when all the $\lambda_i = 0$, so it is a joint test.

$\lambda_{\max}$ tests the null that the number of cointegrating vectors is r against an alternative of $r+1$.

The Johansen Testing Sequence

- **If the test statistic is greater than the critical value from Johansen's tables, reject the null hypothesis.**
- The testing sequence under the null is $r = 0, 1, ..., g-1$
  so that the hypotheses for λ_{trace} are

H_0: $r = 0$	vs	H_1: $0 < r \leq g$
H_0: $r = 1$	vs	H_1: $1 < r \leq g$
H_0: $r = 2$	vs	H_1: $2 < r \leq g$
...	...	...
H_0: $r = g-1$	vs	H_1: $r = g$

- **We keep increasing the value of *r* until we no longer reject the null**.

Interpretation of Johansen Test Results

- But how does this correspond to a test of the rank of the Π matrix?

- ***r* is the rank of Π**.

- Π cannot be of full rank (g) since this would correspond to the original y_t being stationary.

- **If Π has zero rank**, there is no cointegration.

- **For 1 < rank (Π) < *g*** , there are multiple cointegrating vectors.

Decomposition of the Π Matrix

- For any $1 < r < g$, Π is defined as the product of two matrices:

  $$\underset{g\times g}{\Pi} = \underset{g\times r}{\alpha}\ \underset{r\times g}{\beta'}$$

- β contains the **cointegrating vectors** while α gives the amount of each cointegrating vector in each equation, called "**adjustment parameter**"
- For example, if g=4 and r=1, α and β will be 4×1, and Πy_{t-k} will be given by:

$$\Pi = \begin{pmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{14} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}_{t-k} \quad \text{or} \quad \Pi = \begin{pmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{14} \end{pmatrix} \begin{pmatrix} \beta_{11}y_1 & \beta_{12}y_2 & \beta_{13}y_3 & \beta_{14}y_4 \end{pmatrix}_{t-k}$$

Hypothesis Testing Using Johansen

- Engle-Granger did not allow us to do hypothesis tests on the cointegrating relationship itself, but the Johansen approach does.

- If there exist *r* cointegrating vectors, only these linear combinations will be stationary.

- **Idea:** We can test a hypothesis about one or more coefficients in the cointegrating relationship by viewing the hypothesis as a restriction on the Π matrix.

Hypothesis Testing Using Johansen (cont'd)

- A **test statistic** to test this hypothesis is given by

$$-T\sum_{i=1}^{r}[\ln(1-\lambda_i)-\ln(1-\lambda_i^*)] \quad \sim \chi^2(m)$$

where,

λ_i^* are the characteristic roots of the restricted model

λ_i are the characteristic roots of the unrestricted model

r is the number of non-zero characteristic roots in the unrestricted model, and m is the number of restrictions.

Cointegration Tests using Johansen: Three Examples

Example 1: Hamilton(1994, pp.647)

- Does the PPP relationship hold for the US / Italian exchange rate - price system?
- A VAR was estimated with 12 lags on 189 observations. The **Johansen test statistics** were:

r	λ_{max}	critical value
0	22.12	20.8
1	10.19	14.0

- Conclusion: there is **one cointegrating relationship**.

Example 3: Are International Bond Markets Cointegrated?

- Mills & Mills (1991)

- If financial markets are cointegrated, this implies that they have a "common stochastic trend".

Data:

- Daily closing observations on redemption yields on government bonds for 4 bond markets: US, UK, West Germany, Japan.

- For cointegration, a necessary but not sufficient condition is that the yields are nonstationary. All 4 yields series are I(1).

Testing for Cointegration Between the Yields

- The Johansen procedure is used. There can be at most 3 linearly independent cointegrating vectors.

- Mills & Mills use the trace test statistic: $\lambda_{trace}(r) = -T\sum_{i=r+1}^{g}\ln(1-\hat{\lambda}_i)$ where λ_i are the ordered eigenvalues.

Johansen Tests for Cointegration between International Bond Yields

r (number of cointegrating vectors under the null hypothesis)	Test statistic	Critical Values 10%	5%
0	22.06	35.6	38.6
1	10.58	21.2	23.8
2	2.52	10.3	12.0
3	0.12	2.9	4.2

Source: Mills and Mills (1991). Reprinted with the permission of Blackwell Publishers.

Testing for Cointegration Between the Yields (cont'd)

- Conclusion: No cointegrating vectors.

- The paper then goes on to estimate a VAR for the first differences of the yields, which is of the form

$$\Delta X_t = \sum_{i=1}^{k} \Gamma_i \Delta X_{t-i} + \upsilon_t$$

where

$$X_t = \begin{bmatrix} X(US)_t \\ X(UK)_t \\ X(WG)_t \\ X(JAP)_t \end{bmatrix}, \Gamma_i = \begin{bmatrix} \Gamma_{11i} & \Gamma_{12i} & \Gamma_{13i} & \Gamma_{14i} \\ \Gamma_{21i} & \Gamma_{22i} & \Gamma_{23i} & \Gamma_{24i} \\ \Gamma_{31i} & \Gamma_{32i} & \Gamma_{33i} & \Gamma_{34i} \\ \Gamma_{41i} & \Gamma_{42i} & \Gamma_{43i} & \Gamma_{44i} \end{bmatrix}, \upsilon_t = \begin{bmatrix} \upsilon_{1t} \\ \upsilon_{2t} \\ \upsilon_{3t} \\ \upsilon_{4t} \end{bmatrix}$$

They set $k = 8$.