Lecture Notes (3)

Modelling long-run relationship in finance

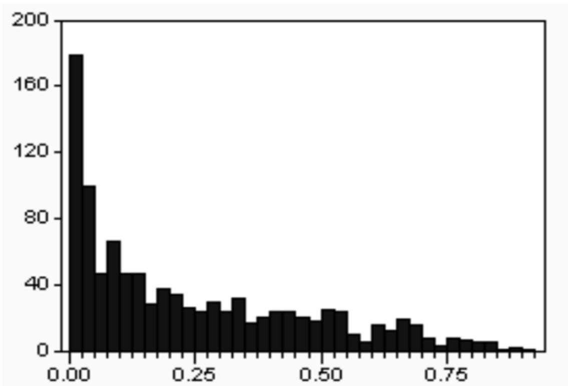
1

Stationarity and Unit Root Testing

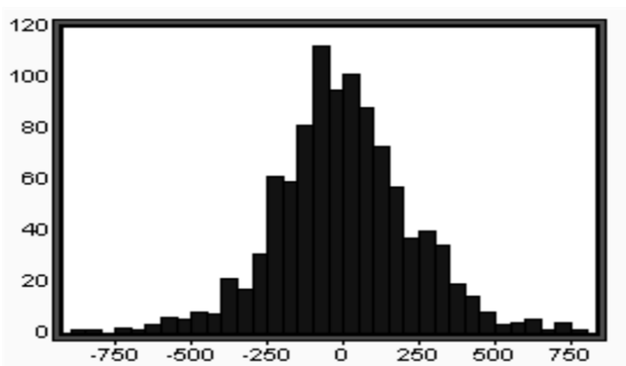
Why do we need to test for Non-Stationarity?

- The stationarity or otherwise of a series can strongly influence its behaviour and properties - e.g. persistence of shocks will be infinite for nonstationary 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 are totally unrelated
- If the variables in the regression model are not stationary, then it can be proved that the standard assumptions for asymptotic analysis will not be valid. In other words, 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:

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

and the deterministic trend process:

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

where u_t is *iid* in both cases.

Stochastic Non-Stationarity

- 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, the explosive 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:

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

7

The Impact of Shocks for Stationary and Non-stationary Series

- We have 3 cases:

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

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:

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

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.

8

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

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

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

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

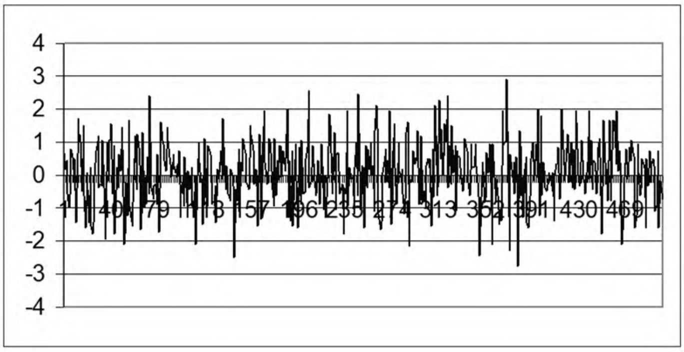
$$\Delta y_t = \mu + u_t$$

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

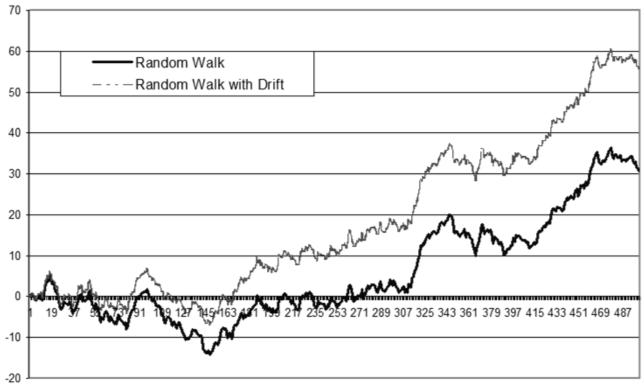
Detrending a Series: Using the Right Method

- Although trend-stationary and difference-stationary series are both "trending" over time, the correct approach needs to be used in each case.
- 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.
- 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 since deterministic non-stationarity does not adequately describe most series in economics or finance.

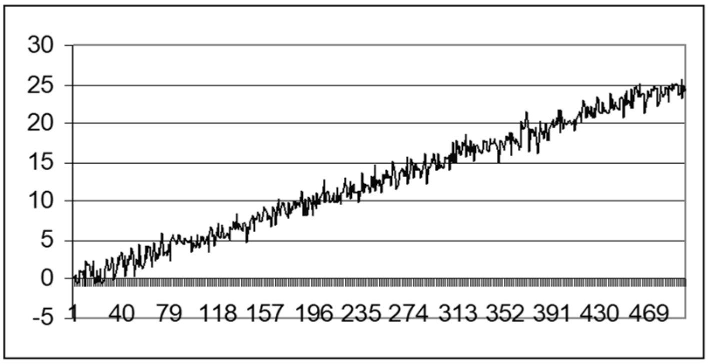
Sample Plots for various Stochastic Processes: A White Noise Process



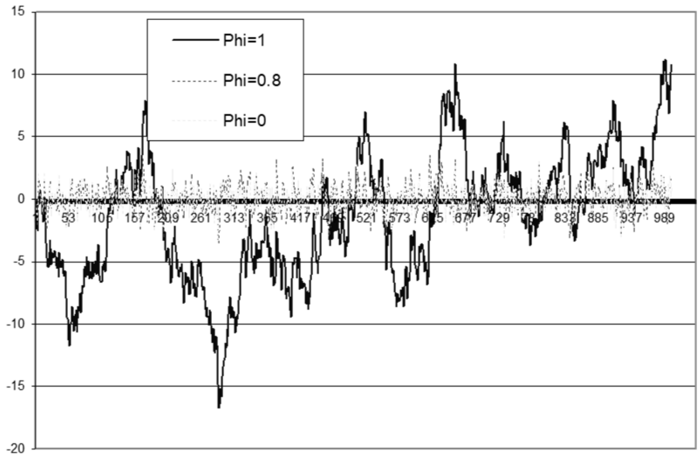
Sample Plots for various Stochastic Processes: A Random Walk and a Random Walk with Drift



Sample Plots for various Stochastic Processes: A Deterministic Trend Process



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



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.
- 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 (Dickey and Fuller 1979, Fuller 1976). 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 a unit root

vs. H_1: series is stationary.

- We usually use the regression:

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

so that a test of $\phi=1$ is equivalent to a test of $\psi=0$ (since $\phi-1=\psi$).

Different forms for the DF Test Regressions

- Dickey Fuller tests (1979, 1981) are also known as τ tests: τ, τ_μ, τ_τ.

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

i) H_0: $y_t = y_{t-1}+u_t$
H_1: $y_t = \phi y_{t-1}+u_t$, $\phi<1$

This is a test for a random walk against a stationary autoregressive process of order one (AR(1))

ii) H_0: $y_t = y_{t-1}+u_t$
H_1: $y_t = \phi y_{t-1}+\mu+u_t$, $\phi<1$

This is a test for a random walk against a stationary AR(1) with drift.

iii) H_0: $y_t = y_{t-1}+u_t$
H_1: $y_t = \phi y_{t-1}+\mu+\lambda t+u_t$, $\phi<1$

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 $\Delta y_t = u_t$ where $\Delta y_t = y_t - y_{t-1}$, 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 are derived from Monte Carlo experiments in, for example, Fuller (1976). Relevant examples of the distribution are shown in the table below:

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

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.

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

  - 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: $\psi=0$ vs. H_1: $\psi<0$.
- If H_0 is rejected we simply conclude that y_t does not contain a unit root.
- 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

$$H_0: y_t \sim I(2) \text{ vs. } H_1: y_t \sim I(1)$$

  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).

The Phillips-Perron Test

- Phillips and Perron have developed a more comprehensive theory of unit root nonstationarity. The tests are similar to ADF tests, but they incorporate an automatic correction to the DF procedure to allow for autocorrelated residuals.

- The tests usually give the same conclusions as the ADF tests, and the calculation of the test statistics is complex.

Criticism of Dickey-Fuller and Phillips-Perron-type tests

- 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 if
  $$\phi=1 \text{ or } \phi=0.95,$$
  especially with small sample sizes.

- If the true data generating process (dgp) is
  $$y_t = 0.95y_{t-1} + u_t$$
  then the null hypothesis of a unit root should be rejected.

- One way to get around this is to use a stationarity test as well as the unit root tests we have looked at.

Stationarity tests

- Stationarity tests have

  H_0: y_t is stationary

  versus H_1: y_t is non-stationary

  So that by default under the null the data will appear stationary.

- One such stationarity test is the KPSS test (Kwaitowski, Phillips, Schmidt and Shin, 1992).

- Thus we can compare the results of these tests with the ADF/PP procedure to see if we obtain the same conclusion.

Stationarity tests (cont'd)

- A Comparison

ADF / PP	KPSS
H_0: $y_t \sim I(1)$	H_0: $y_t \sim I(0)$
H_1: $y_t \sim I(0)$	H_1: $y_t \sim I(1)$

- 4 possible outcomes

Reject H_0	**and**	Do not reject H_0
Do not reject H_0	**and**	Reject H_0
Reject H_0	and	Reject H_0
Do not reject H_0	and	Do not reject H_0

(Box 8.1, on page 365 – Confirmatory data analysis)

Testing for Unit Roots in EViews

- This example uses the same data on UK house prices as employed in last lecture.

- ADF-test
- PP-test
- KPSS-test

- Discuss tables and outputs on pages 369~372.

Cointegration: An Introduction

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

- More generally, if we combine variables with differing orders of integration, the combination will have an order of integration equal to the largest. i.e.,

  if $X_{i,t} \sim \mathrm{I}(d_i)$ for $i = 1,2,3,...,k$

  so we have k variables each integrated of order d_i.

  Let $$z_t = \sum_{i=1}^{k} \alpha_i X_{i,t} \qquad (1)$$

  Then $z_t \sim \mathrm{I}(\max d_i)$

Linear Combinations of Non-stationary Variables

- Rearranging (1), we can write

$$X_{1,t} = \sum_{i=2}^{k} \beta_i X_{i,t} + z'_t$$

where $\beta_i = -\frac{\alpha_i}{\alpha_1}$, $z'_t = \frac{z_t}{\alpha_1}$, $i = 2,\ldots,k$

- This is just a regression equation.
- But the disturbances would have some very undesirable properties: z_t' is not stationary and is autocorrelated if all of the X_i are I(1).
- We want to ensure that the disturbances are I(0). Under what circumstances will this be the case?

Definition of Cointegration (Engle & Granger, 1987)

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

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

  ii) There is at least one vector of coefficients α such that $\alpha'\boldsymbol{z}_t \sim$ I(d-b)
- In practice, many financial variables/time series are non-stationary, contain one unit root, i.e. they are I(1), but "move together" over time.
- If variables are cointegrated, it means that a linear combination of them will be stationary (i.e. d=b=1, ➔I(0)).
- 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.

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

Equilibrium Correction or Error Correction Models

- 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 Error Correction Models

- 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.

Testing for Cointegration in Regression

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

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

- 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

$$\Delta \hat{u}_t = \psi \hat{u}_{t-1} + v_t \qquad \text{with } v_t \sim \text{iid.}$$

- 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.

- We can also use the Durbin Watson test statistic or the Phillips Perron approach to test for non-stationarity of $\hat{u}_t$.

- 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, i.e. $\hat{u}_t \sim I(1)$
  H_1 : residuals from cointegrating regression are stationary, i.e. $\hat{u}_t \sim I(0)$

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 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.

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

where $\hat{u}_{t-1} = y_{t-1} - \hat{\gamma} x_{t-1}$

The Engle-Granger Approach: Some Drawbacks

This method suffers from a number of problems:

1. Unit root and cointegration tests 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.