ECO 372: Introduction to Econometrics

Spring 2025

Introduction to Time Series: Multivariate Modelling

Sakib Bin Amin, Ph.D.

Associate Professor in Economics

Director, Accreditation Project Team (APT)

Outline

Our objectives for this lecture will be to learn:

- Non-Stationary Data
- Spurious Regression
- Unit Root Detection
- Cointegration or Long-Run Equilibrium
- Estimation of Long-Run Equilibrium Models

Non-Stationary Data

- So far we have assumed that our data is stationary.
- In stationary time series variables, mean, variance, and correlation of past and future values are constant.
- Given these characteristics, stationary variables do not portray any sudden fluctuations and tend to stay around the mean value.
- Even if any shock distorts the variables, after a certain time they will come back to the old state.
- On the other hand, non-stationary variables do not have constant mean and variance.
- The mean and variance of the non-stationary variables are time-dependent.
- When a variable is non-stationary, we say that the variable has a unit root.

Non-Stationary Data

- In general stationary series will follow a theoretical ACF and will have quicker geometrical decay as the lag-length increases.
- The theoretical ACF of a non-stationary time series will not die out for increasing lag- length.
- However, this method is bound to be imperfect/imprecise because a near-unit-root process will have the same shape of the autocorrelation function (ACF) as that of a real unit-root process.
- Thus, what might appear as a unit root for one researcher may appear as a stationary process for another.
- Therefore, we need specific tests to check non-stationarity.

Non-Stationary Data

- Consider the AR(1) model:

$$y_t = a_1 y_{t-1} + \varepsilon_t \; ; \; \varepsilon_t \sim N(0, \sigma^2)$$

- Where ε_t is a white noise process and the stationarity condition is $|a_1| < 1$
- In general we can have three possible cases:
  - Case 1: $|a_1| < 1$ therefore, the series is stationary.
  - Case 2: $|a_1| > 1$ where in this case, the series explodes.
  - Case 3: $|a_1| = 1$ where in this case, the series contains a unit root
- If $|a_1| = 1$, y_t contains a unit root (unity, hence unit root).
- Subtracting y_{t-1} from both sides of a non-stationary AR(1) equation we get:

$$y_t - y_{t-1} = y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \varepsilon_t$$

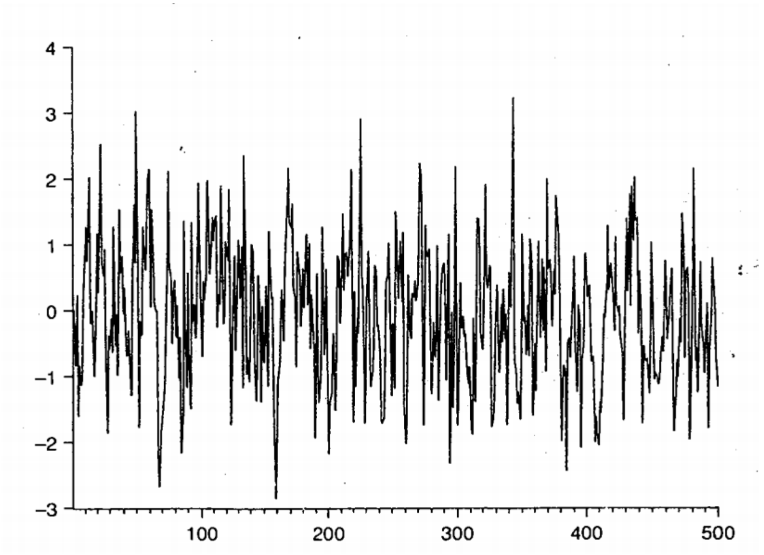
- Because ε_t is a white-noise process then we have that Δy_t is a stationary series. So differencing makes a series stationary.

Non-Stationary Data

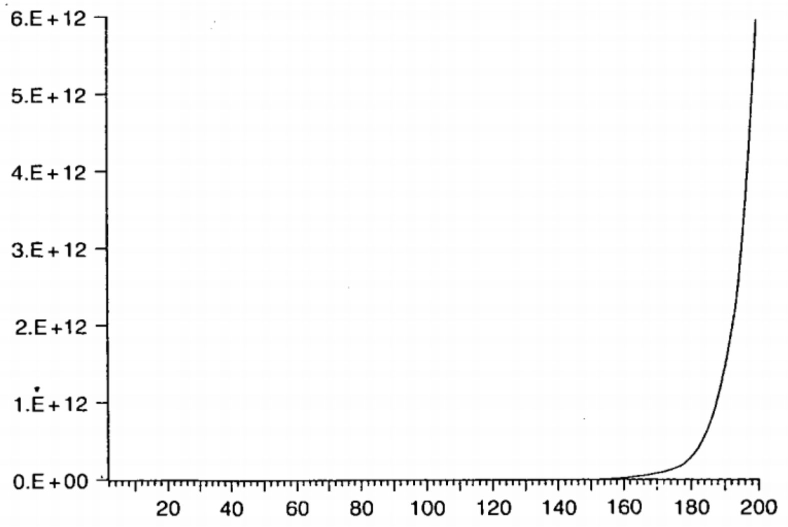
- When a series is stationary after a single differencing, we refer it as integrated of order one or I (1).
- When a series is stationary without first difference (or at level), we refer it as integrated of order zero or I (0).
- In general a non-stationary time series y_t might need to be differenced more than once before it becomes stationary.
- We can summarize the above information under a general rule:

$$(Order\ of\ Integration) \rightarrow (Number\ of\ difference\ needed) \rightarrow (Number\ of\ unit\ roots)$$

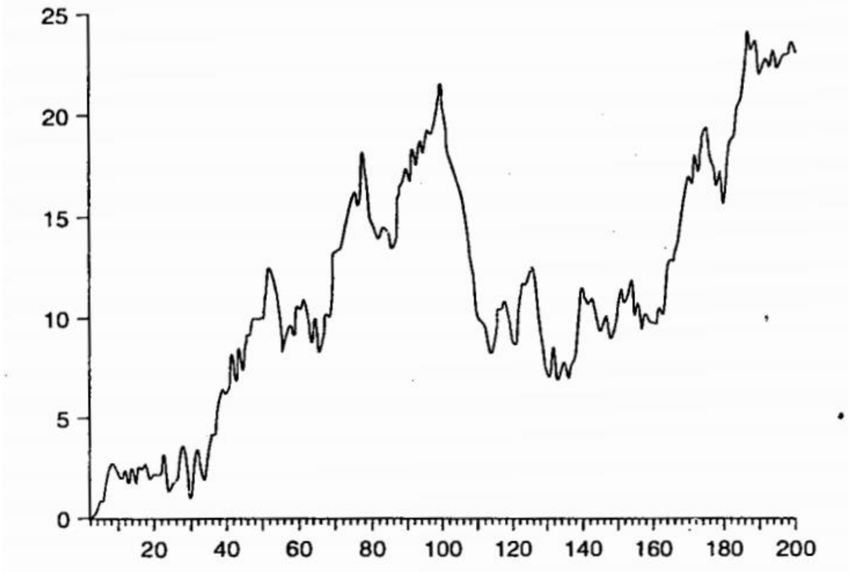
Non-Stationary Data



Case 1: $|a_1| < 1$



Case 2: $|a_1| > 1$



Case 3: $|a_1| = 1$

Unit Root and Spurious Regression

- Most macroeconomic time series are trended and therefore in most cases are non-stationary (e.g., GDP, Money Supply, and CPI).
- The problem with non-stationary or trended data is that the standard OLS regression procedures can easily lead to incorrect conclusions or known as spurious regressions.
- If we consider two completely unrelated series which are both non-stationary.
- If we then performed a regression of one series on the other we would then find either a significant positive relationship if they are going in the same direction or a significant negative one if they are going in opposite directions even though they are both unrelated.
- A spurious regression usually has a very high R^2, t-statistics that appear to provide significant estimates, but the results may have no economic meaning whatsoever.
- Therefore, econometricians should be very careful when working with trended variables.

Testing for Unit Roots or Non-Stationarity

- There are many tests to check the unit root. Dickey and Fuller or DF (1979, 1981), Phillips and Perron or PP (1988), Kwiatkowski-Phillips-Schmidt-Shinor or KPSS (1991), etc.
- We will focus on DF tests. Supposes, we have an AR (1) model:

$$y_t = a_1 y_{t-1} + \varepsilon_t \; ; \; \varepsilon_t \sim N(0, \sigma^2)$$

- What we need to examine here is whether a_1 is equal to 1. The null hypothesis is $H_0: a_1$=1 and the alternative hypothesis is $H_0: a_1$ <1.
- Rather than using the above framework, we can work with a more convenient framework to run the test.

$$y_t - y_{t-1} = a_1 y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = (a_1 - 1) y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

- Where of course $\gamma = (a_1 - 1)$. Now the null hypothesis is H_0: γ =0 and the alternative hypothesis is H_0: γ <0. when $\gamma = 0$ its non-stationary and pure random walk series.

Testing for Unit Roots or Non-Stationarity

- Dickey and Fuller (1979) also proposed two alternative regression equations that can be used for testing for the presence of a unit root.

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \alpha_2 t + \varepsilon_t$$

- The difference between the three regressions concerns the presence of the deterministic elements α_0 and $\alpha_2 t$. Where, α_0 is drift or constant. $\alpha_2 t$ is a linear time trend.
- When $\gamma = 0$ in the presence of α_0 or α_0 and $\alpha_2 t$, then the series is called random walk with drift and random walk with drift and trend, respectively.
- Dickey and Fuller test uses t-statistics: $\gamma/\mathrm{SE}(\gamma)$.
- Critical values are not usual in this case and depend on whether an intercept and/or time trend is included in the regression equation.
- MacKinnon (1991) tabulated appropriate critical values for each of the three specifications, which all software use.

Testing for Unit Roots or Non-Stationarity

- Dickey and Fuller extended their test procedure suggesting an augmented version of the test which includes extra lagged terms of the dependent variable to eliminate autocorrelation.
- Formally the test is known as the Augmented Dickey-Fuller (ADF) Test.
- The lag-length on these extra terms is either determined by the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC).
- The three possible forms of the ADF test are given by the following equations:

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \alpha_2 t + \varepsilon_t$$

Cointegration

- We know that most macroeconomic variables are trended and therefore the spurious regression problem is highly likely to be present in most macroeconometric models.
- One way of resolving this is to differentiate the series successively until stationarity is achieved and then use the stationary series for regression analysis.
- However, this solution is not ideal.
- The model may no longer give a unique long-run solution due to loss of information (the smaller sample size), resulting from differencing.
- If the model is correctly specified as a relationship between y and x, and we differentiate both variables, then implicitly we are also differencing the error process in the regression.
  - This would then produce a non-invertible moving average error process and would present serious estimation problems.

Cointegration

- If the two variables are non-stationary then we can represent the error as a combination of two cumulated error processes.
- Stochastic trends, found in the cumulated error processes are usually not related, which produces another non-stationary process.
- However, in the case of cointegration, X and Y are moved together and so the two stochastic trends would be very similar/related to each other.
- When we combine them, it should be possible to find a combination of them that eliminates the non-stationarity.
- In theory, this should only happen when there exist true relationships between two variables.
- So, cointegration becomes a very powerful way of detecting the presence of economic structures.

Cointegration

- If there real genuine long-run relationship between y_t and x_t, then although the variables will rise and decline over time, there will be a common trend/underlying features or patterns that link them together.
- For the long-run relationship to exist what we require, then, is a linear combination of y_t and x_t that is a stationary variable (an I (0) variable).
- A linear combination of y_t and x_t can be directly taken from estimating the following regression and taking the residuals:

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$$

$$\hat{\varepsilon}_t = y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t$$

- If $\hat{\varepsilon}_t \sim I(0)$, then the variables y_t and x_t are said to be cointegrated.
- This will only happen when y_t and x_t are integrated in same order.
- In literature, the long-run relationship is widely referred as the long-run equilibrium.

Cointegration

- ➢ Engle and Granger (1987) further formalized this concept by introducing a very simple test for the existence of long-run equilibrium.
- ➢ This involves several steps.

<u>Step 1 : Test the variables for their order of integration</u>

- By definition, cointegration necessitates that the variables be integrated of the same order.
- Thus, the first step is to test each variable to determine its order of integration.
- The ADF or other unit root tests an be applied in order to infer the number of unit roots.
- If both variables are integrated of the same order (but not I(0)) then we proceed with step two.

<u>Step 2: Estimate the long-run (possible cointegrating) equilibrium</u>

- Estimate the long-run equilibrium relationship of the following form and obtain the residuals of this equation.

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$$

- If there is no cointegration, the results obtained will be spurious.
- However, if the variables are cointegrated, then we can find long-run coefficients.

Cointegration

Step 3 : Check for (cointegration) the order of integration of the residuals

- Perform a ADF test on the residual series to determine their order of integration. There is no need of adding drift or time trend.

$$\Delta\hat{\varepsilon}_t = \gamma\hat{\varepsilon}_{t-1} + \sum_{i=1}^{p} \delta_i \Delta\hat{\varepsilon}_{t-i} + u_t$$

- If we find that $\hat{\varepsilon}_t \sim I(0)$, we can reject the null that the variables y_t and x_t are not cointegrated.

Cointegration

- Advantages of the Engle-Granger test:
  - One of the best features of the EG approach is that it is both very easy to understand and to implement.
- Disadvantages of the Engle-Granger test:
  - One very important issue has to do with the order of the variables. When estimating the long-run equilibrium, one has to place one variable on the left-hand side and use the others as regressors. The test does not say anything about which of the variables can be used as a regressor and why.
  - A second problem is that when there are more than two variables, there may be more than one cointegrating relationship, and the Engle-Granger procedure using residuals from a single relationship cannot treat this possibility.
  - A third and final problem is that it relies on a two-step estimator. The first step is to estimate the residual series, and the second step is to estimate a regression for this series in order to see if the series is stationary or not. Hence, any error introduced in the first step is carried into the second step.

Cointegration

- It was mentioned before that if we have more than two variables in the model, then there is a possibility of having more than one equilibrium.
- In general, for *n* number of variables we can have only up to *n- 1* cointegrating equilibriums
- Therefore, an alternative to the Engle-Granger approach is needed and this is the Johansen (1991) approach for multiple equations.
- The Johansen approach uses a VAR framework to find out the rank of the system equation matrix.
- It tests the null hypothesis that rank=1 (1 cointegration) against the alternate hypothesis that rank=r+1 (more than 1 cointegration).
- The hypothesis is tested by two tests: the Maximal Eigen Value statistic and the Trace Statistic.
- Before testing the hypothesis, one has to ensure that all the variables are integrated in the same order.

Cointegration

- In the early 2000s, Pesaran et al. (2001), introduced a cointegration testing approach called the ARDL Bounds test.
- This approach became popular as it breaks the traditional restriction of cointegration tests in that the tested variables must be non-stationary and all the variables are integrated in the same order.
- Some researchers favour this approach as many of the applications involve economic variables of a mixed or unknown order of integration.
- Nevertheless, the dependent variable has to be I(1) to run the ARDL Bounds test.
- The ARDL Bounds test uses F-Statistics to test the null hypothesis of no cointegration against the alternate hypothesis of cointegration.
- The updated version of the ARDL Bounds test proposed by Sam et al. (2019) does not require the assumption of an I(1) dependent variable.

Estimation of Long-Run Models: Distributed Lag and Autoregressive Distributed Lag

- When variables are found to be cointegrated, we can estimate long-run models.
- Existing literature shows numerous approaches for estimating long-run models.
- For instance, the Distributed Lag or DL approach and the Autoregressive Distributed Lag or ARDL approach.
- An DL (q) and ARDL (p, q) approach can be shown as follows:

$$y_t = a_0 + \beta_1 x_t + \beta_2 x_{t-1} + \cdots + \beta_z x_{t-z} + \varepsilon_t;\ \varepsilon_t \sim N(0, \sigma^2)$$

$$y_t = a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_z x_{t-q} + \varepsilon_t;\ \varepsilon_t \sim N(0, \sigma^2)$$

- The ARDL approach can be further redesigned to not only yield long-run coefficients but also short-run coefficients. This redesigned approach is known as the ARDL Error Correction Model (ECM).
- Note, in ARDL, variables can have a mixed order of integration.

Estimation of Long-Run Models: Distributed Lag and Autoregressive Distributed Lag

- Consider the ARDL (1, 1):

$$y_t = a_0 + \lambda y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$$

- With some algebraic calculations, this model can be written equivalently as:

$$\Delta y_t = a_0 + \beta_0 \Delta x_t - (1-\lambda)(y_{t-1} - \theta x_{t-1}) + \varepsilon_t$$

- Where, $\theta = (\beta_0 + \beta_1)/(1-\lambda)$ is the slope coeicient in the long-run relationship between y_t and x_t.
- The term Δx_t is referred to as the derivative effect or short-run effect and $(y_{t-1} - \theta x_{t-1})$ as the ECM.
- Stability of the ARDL ECM depends on $(1-\lambda)$. If $0 < (1-\lambda) < 1$, the process is mean reverting (residual will be I(0)) and converges to a long-run equilibrium. If $(1-\lambda) < 0$, then it is regarded as an explosive model.
- The coefficient of $-(1-\lambda)$ tell us the speed of adjustment. For instance, if $-(1-\lambda) = -0.25$, it indicates the model will converge to long-run equilibrium from short-run by 25% by each period.

Estimation of Long-Run Models: Dynamic OLS and Fully Modified OLS

- The Dynamic Ordinary Least Square (DOLS) is an improved version of the OLS approach.
- This is a very effective approach for estimating long-run coefficients.
- DOLS can deal with small sample sizes and dynamic sources of bias, which is also one of the major advantages of this technique.
- Moreover, it is a robust single-equity approach that corrects the regressor's endogeneity by incorporating lags and leads on the first differenced regressors.

$$y_t = \alpha_0 + \alpha x_t + \beta z_t + \sum_{i=1,0}^{q=n} \alpha \Delta x_{t\pm i} + \sum_{i=1,0}^{q=n} \beta \Delta z_{t\pm i} + \varepsilon_t$$

- It should be noted that a mixed order of integration is not suitable for DOLS estimation.
- Different trend patterns can be added to the framework to better suit the data.

Long-Run Models: Fully Modified OLS

- The FMOLS estimation is a modified version of OLS to take into account the serial correlation in the residuals due to the augmentation of differenced regressors.
- To adjust for these factors, nonparametric adjustments are made to the dependent variable and then to the estimated long-run parameters obtained from regressing the adjusted dependent variable on the regressors.

$$\dot{y}_t = \alpha_0 + \alpha x_t + \beta z_t + \alpha \Delta x_t + \beta \Delta z_t + \varepsilon_t$$

- $\dot{y}_t$ is the adjusted dependent variable. $\dot{y}_t = y_t - \delta$. Where δ is the adjustment that eliminates residual correlation (i.e., estimated error correlation).
- Similar to DOLS, it also accounts for the endogeneity bias caused by the causal influence from the endogenous to the exogenous variables.
- The empirical literature has shown that DOLS performs well than FMOLS, specifically by accounting for correlations among regressors.
- It should be noted that a mixed order of integration is not suitable for FMOLS estimation.

Estimation of Long-Run Models: VAR

- When we have cointegrated variables, we can also use the VAR framework to estimate a system of equations for analysing long-run shocks and causalities.

- Similar to the ARDL ECM, VAR can be used for deriving both short-run and long-run inferences. This model is known as the Vector Error Correction Model (VECM).

$$\Delta y_t = a_1 + \sum_{i=1}^{n} a_i \,\Delta y_{t-i} + \sum_{j=1}^{m} \pi_i \,\Delta x_{t-j} + b_1(y_{t-1} - \theta x_{t-1}) + e_{1t}$$

$$\Delta\, x_t = a_2 + \sum_{i=1}^{n} \theta_i \,\Delta y_{t-i} + \sum_{j=1}^{m} \mu_i \,\Delta x_{t-j} + b_2(y_{t-1} - \theta x_{t-1}) + e_{2t}$$

- $(y_{t-1} - \theta x_{t-1})$ is the ECM .

Example from Literature

Table 3
Unit root tests results.

Variable	Level		First Difference	
	Intercept	Intercept and Trend	Intercept	Intercept and Trend
ADF				
YGR	−1.50	−2.30	−4.86***	−4.96***
I/Y	−1.00	−3.58	−4.16***	−4.07***
LGR	−0.81	−1.48	−4.48**	−4.42***
EGR	−2.95	−3.33	−4.65***	−4.71***
PP				
YGR	−2.13	−2.21	−15.52***	−14.68***
I/Y	0.09	−1.91	−3.54***	−3.60**
LGR	−2.39	−3.39*	−7.04***	−6.86***
EGR	−2.34	−3.32	−7.41***	−7.25***
KPSS				
YGR	0.47**	0.20**	0.15	0.13
I/Y	0.90***	0.12**	0.10	0.09
LGR	0.40*	0.20**	0.19	0.16
EGR	0.46**	0.12*	0.17	0.11
DF-GLS				
YGR	0.82	−2.22	−2.30**	−3.83***
I/Y	0.47	−2.25	−2.60**	−4.31***
LGR	−0.72	−2.01	−1.60*	−3.10**
EGR	−1.12	−2.37	−3.87***	−4.04***

Note: ***, **, and * show significance at 1%, 5%, and 10% respectively. ADF, PP and DF-GLS perform under null hypothesis where series is non-stationary against alternative hypothesis of series is stationary. KPSS performs under null hypothesis where series is stationary against alternative hypothesis of series is not stationary.

Table 4
Johansen cointegration test results.

Hypothesised No. of CE (s)	Trace Test	Probability	Max-Eigen Test	Probability
None	65.18	0.00	35.24	0.00
At most 1	29.93	0.04	20.67	0.10
At most 2	9.26	0.34	8.90	0.29
At most 3	0.36	0.55	0.36	0.55

Note: The test is run on intercept but no trend configuration.

Table 7
Sensitivity analysis of the long-run coefficients of two-sector model.

Variables	DOLS	FMOLS	DARDL
I/Y	0.21*** (0.02)	0.20***(0.01)	0.25**(0.05)
LGR	0.78*** (0.06)	0.50***(0.04)	0.60* (0.16)
EGR	0.10*** (0.01)	0.06***(0.01)	0.10**(0.02)
Adj-R^2	0.97	0.80	0.90
J-B	1.26	2.60	0.16
AC	4.59	1.24	1.89

Note: Standard errors are in parenthesis. ***, **, and * show significance at 1%, 5%, and 10% respectively. Error Correction Term of DARDL is −0.42 which is significant at 5%. J-B and AC refer Jarque-Bera and Autocorrelation Tests. Both tests have been done in the residuals of the regressions. DARDL has been obtained after 5000 simulations.

Source: Amin, S. B., Al Kabir, F., & Khan, F. (2020). Energy-output nexus in Bangladesh: a two-sector model analysis. *Energy Strategy Reviews*, 32, 100566.