Predictive Analysis for Decision Making

Weeks 8 and 9

Modelling Non Stationary Time Series

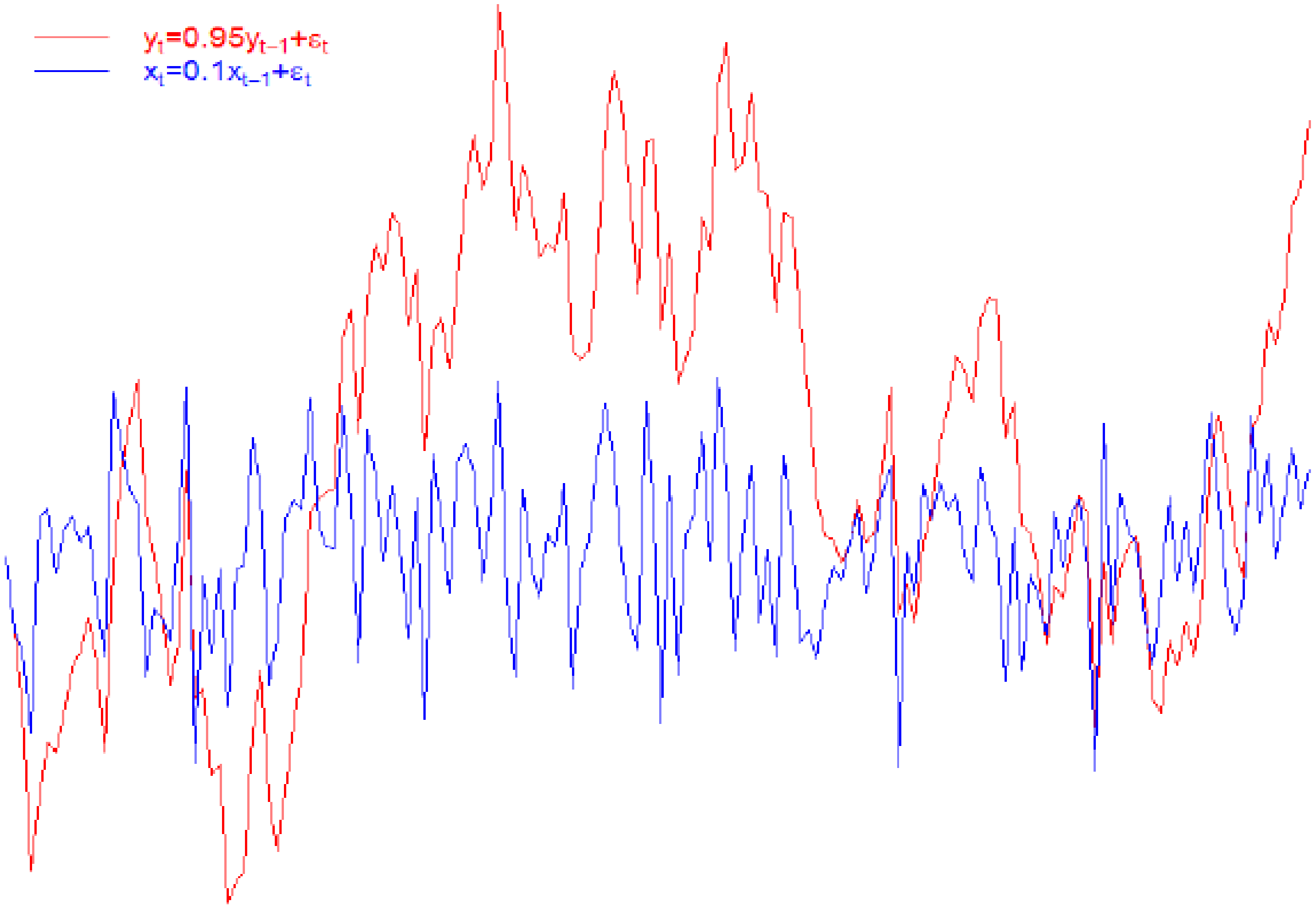
Dr Issam Malki
School of Finance and Accounting
Westminster Business School

November 2021

Non-Stationary Time Series

- We established the following types of time series data
  - Strictly Stationary Time Series
  - Covariance Stationary
  - Random Walk processes
  - Time Varying Volatility
- Implications:
  - Spurious regressions
  - The absence of long run relationships
  - Misleading predictions
  - Incorrect inference
- Plan:
  - Focus on single equation models
  - Introducing unit root tests
  - Introducing cointegration tests

RECAP: RANDOM WALK PROCESSES

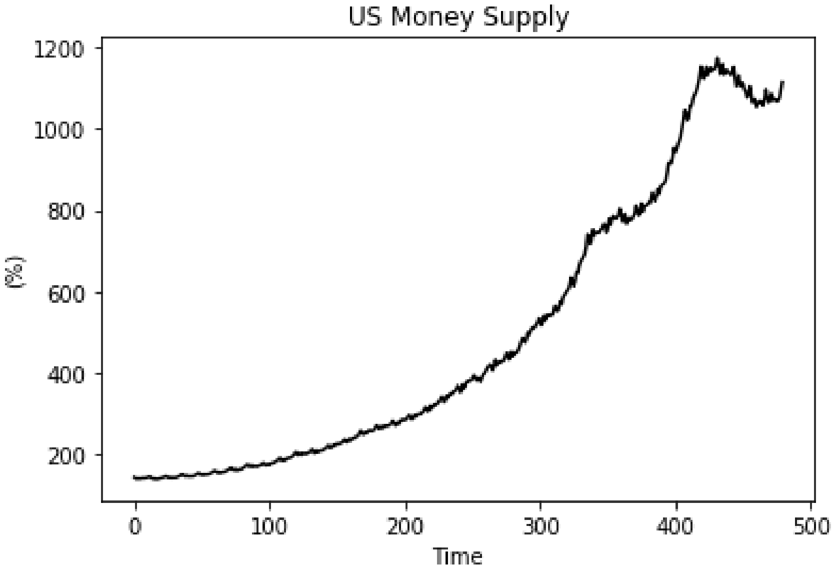


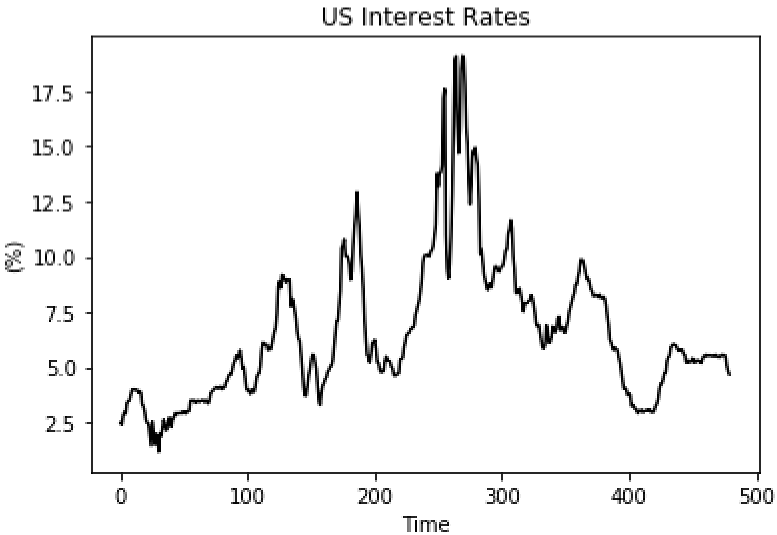
Implications: Spurious Regression

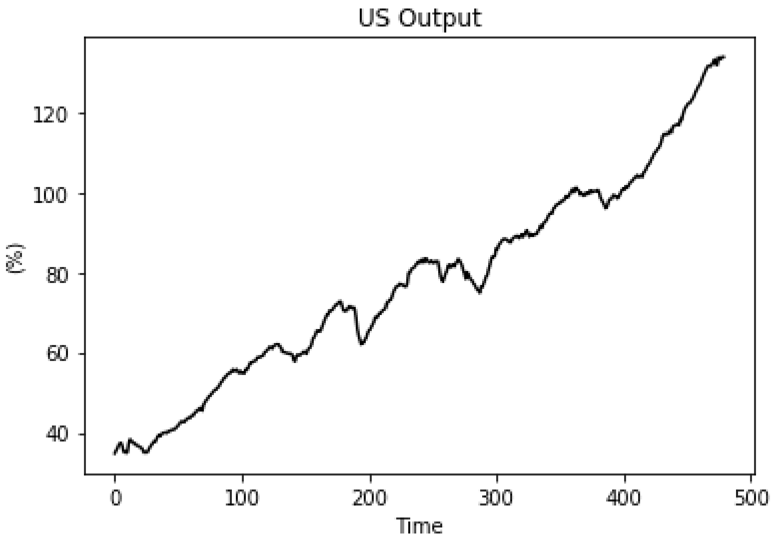
- Spurious Regression
  - Refers to a regression that provide misleading statistical evidence of linear relationship between two or more variables.
  - The regression often reports high t and F statistics, R-squared (near 1).
  - The issue is that these indicators and statistics often hold for variables that are not related or those non-stationary.
  - **Rule of thumb:** R-squared > D.W statistics.
  - If we have a spurious regression, we need to test the stationarity (or non stationarity) of the data.

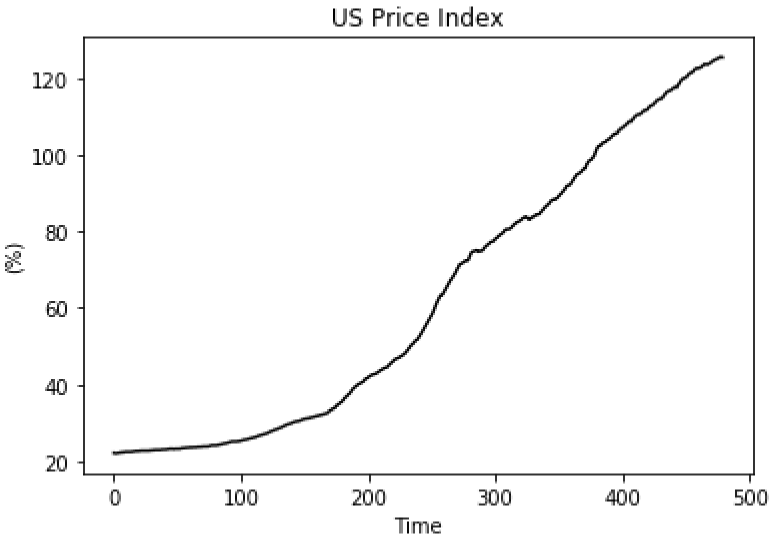
- Recall some of the examples from past lecture
  - The long run static relationship is not valid.
  - When variables suffer from non-stationarity, then we can retrieve the long run relationship using the Error Correction Model, ECM.

Example (1)


US Money Supply
1200
1000
800
600
400
200
(%)
0
100
200
300
400
500
Time


US Interest Rates
17.5
15.0
12.5
10.0
7.5
5.0
2.5
(%)
0
100
200
300
400
500
Time


US Output
120
100
80
60
40
(%)
0
100
200
300
400
500
Time


US Price Index
120
100
80
60
40
20
(%)
0
100
200
300
400
500
Time

Example (1)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                     ot   R-squared:                       0.932
Model:                            OLS   Adj. R-squared:                  0.931
Method:                 Least Squares   F-statistic:                     2171.
Date:                Wed, 17 Mar 2021   Prob (F-statistic):          3.38e-277
Time:                        13:42:54   Log-Likelihood:                -1591.9
No. Observations:                 480   AIC:                             3192.
Df Residuals:                     476   BIC:                             3208.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     31.3661      0.823     38.106      0.000      29.749      32.984
rt             1.4015      0.130     10.808      0.000       1.147       1.656
mt             0.0476      0.006      7.745      0.000       0.035       0.060
pt             0.2206      0.060      3.651      0.000       0.102       0.339
==============================================================================
Omnibus:                       21.864   Durbin-Watson:                   0.022
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               14.793
Skew:                           0.306   Prob(JB):                     0.000613
Kurtosis:                       2.397   Cond. No.                     1.63e+03
==============================================================================
```

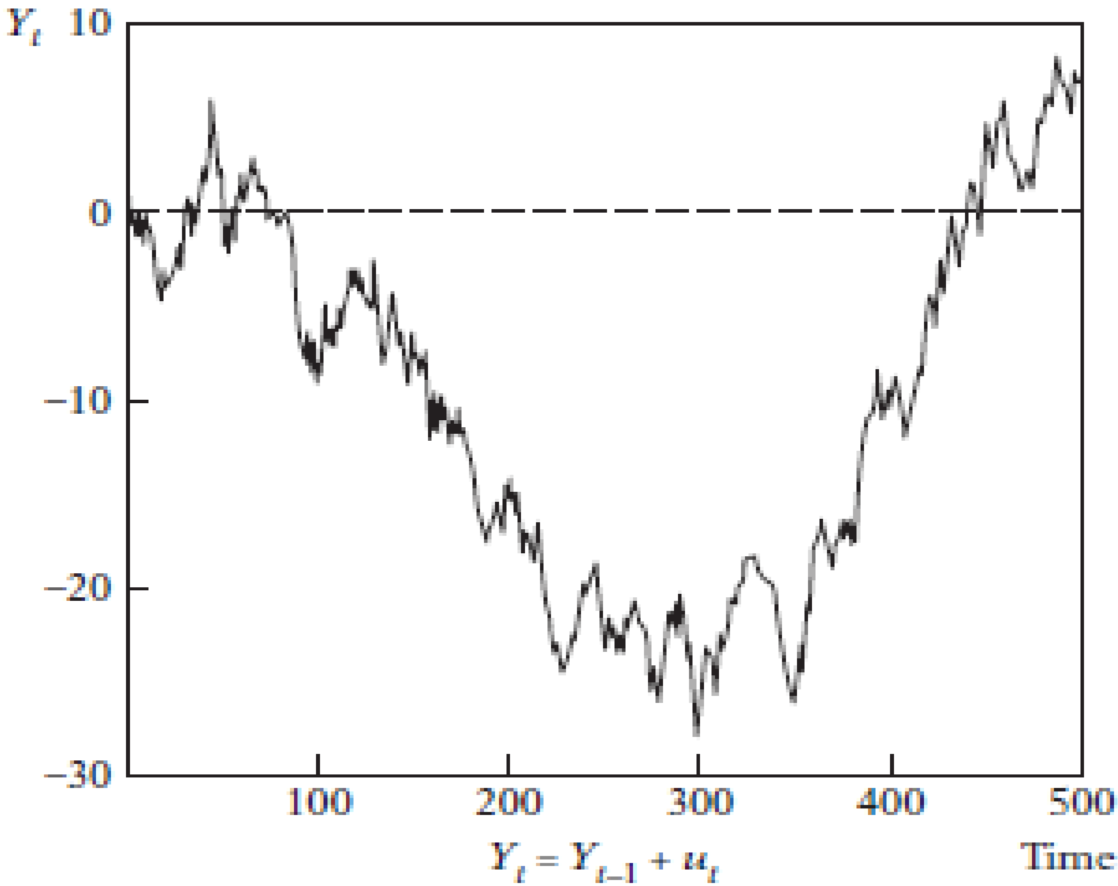
Unit Root Tests
Introduction

- Consider the following Data Generating Process, DGP, for a process y_t:

$$y_t = \rho y_{t-1} + u_t \qquad \text{with} \qquad -1 \leq \rho \leq 1$$

  - u_t is white noise process (a process with zero mean, constant variance and uncorrelated terms)
  - The coefficient ρ is the persistence term. If $\rho = 1$, the process y_t becomes a Random Walk process.
  - In fact, if $\rho = 1$ then y_t contains a **unit root**.
  - **The main implication:** in theory, the process y_t will never converge to the long run level, or equilibrium.
  - Many Economic applications can be found:
    - Income convergence
    - CO2 Convergence
    - Efficient Market Hypothesis
- Therefore, one way to test for the non-stationarity of the data is to test for the presence of a unit root.

Unit Root Tests
Random Walk Process



Unit Root Tests
Dickey-Fuller Test

- Dickey-Fuller test proposes the following procedure to test for the presence of unit root:

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

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

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

where $\Delta y_t = y_t - y_{t-1}$ and $\delta = \rho - 1$

- Remarks
  - The model above the basic DGP that captures the structure of a random walk process.
  - If $\rho = 1$, then $\delta = 0$ and $\Delta y_t = u_t$. This has two implications: i). The process y_t contains a unit root and ii). The change in the process y_t, Δy_t, is stationary since it is a function of a white noise process.
  - Furthermore, if $\delta = 0$ the estimated *t* statistics of the estimated value of δ does not follow normal distribution even in large samples.

Unit Root Tests
Dickey-Fuller Test

- Dickey-Fuller test proposes also the following:
  - Using three forms (DGPs)

y_t is Random Walk: $\Delta y_t = \delta y_{t-1} + u_t$

y_t is Random Walk with drift: $\Delta y_t = \beta_1 + \delta y_{t-1} + u_t$

y_t is Random Walk with drift around stochastic trend

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + u_t$$

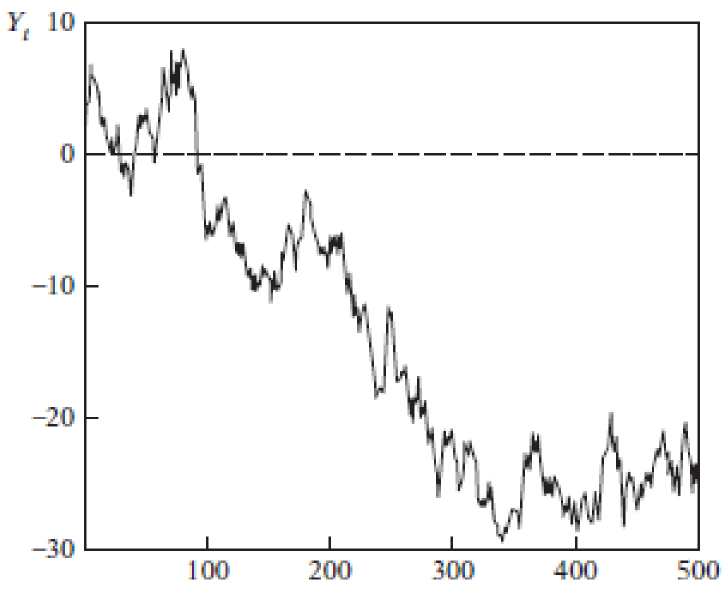
  - The test statistic is called **tau**, which follows different distribution with own critical values.

  - The null and alternative hypotheses are defined as follows:

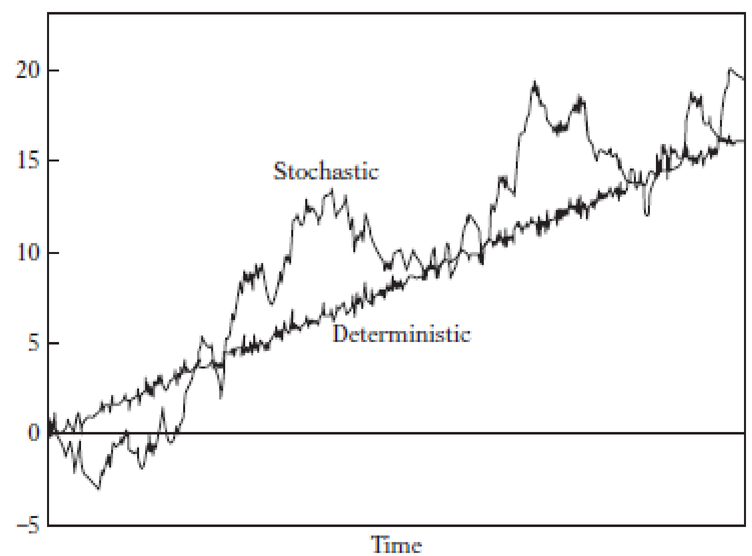
$$H_0\text{: } \delta = 0 \text{ there is a unit root}$$

$$H_1\text{: } \delta < 0 \text{ there is no unit root}$$

Unit Root Tests
Dickey-Fuller Test



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



Unit Root Tests
Augmented Dickey-Fuller Test

- Augmented Dickey-Fuller test, ADF, extends the DF test to correct for the presence of serial correlation. The ADF has the same asymptotic distribution as DF test. Thus, the only difference is the DGPs forms, which are modified to include lagged dependent variables:

y_t is Random Walk: $\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{k} \theta_i \Delta y_{t-i} + u_t$

y_t is Random Walk with drift: $\Delta y_t = \beta_1 + \delta y_{t-1} + \sum_{i=1}^{k} \theta_i \Delta y_{t-i} + u_t$

y_t is Random Walk with drift around stochastic trend

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^{k} \theta_i \Delta y_{t-i} + u_t$$

- The lag length, *k*, is determined using information criteria.

Unit Root Tests in Python

- Use the package 'arch'

- Install in Anaconda using
  - In Anaconda prompt type: conda install arch-py -c conda-forge

- The package offers a wide range of unit root tests

- Example using the variable *'ot'*

```
from arch.unitroot import DFGLS, ADF, KPSS, PhillipsPerron
ot_test=ADF(data1['ot'], lags=10)
print(ot_test.summary())
```

Example (2)

```
  Augmented Dickey-Fuller Results
=====================================
Test Statistic                  0.316
P-value                         0.978
Lags                               10
-------------------------------------

Trend: Constant
Critical Values: -3.44 (1%), -2.87 (5%), -2.57 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

Unit Root Tests
Order of Integration

- All variables tested contain a unit root.
- The natural question is what to do next?
  - We take the first difference of the data.
  - Run unit root tests on the differenced data.
  - If the first differenced data contain no unit root (i.e. rejecting the null), then we conclude that the data are first-difference stationary.
  - The process that becomes stationary after taking its first difference is known as an **integrated process** of order 1.
  - The order of the integration is equal to the number of times the process is differenced to become stationary.
- Remarks
  - When applying unit root tests on the first differenced data, do not choose trend.
  - Notation: *I(0)* variable refers to a variable that does not contain a unit root. It is read as 'Integrated of order 0'. *I(1):* an integrated process of order 1. This means the variable is stationary when taking its first difference. In general, *I(d)* refers to a process that is integrated of order *d* (i.e. the variable contain *d* unit roots and stationary only when taking the *d* th difference).

Spurious Regression - Again

- From above
  - The presence of unit roots in the data give a rise to spurious regression
  - The question we need to answer: Can we estimate the long run relationship between macroeconomic variables?
  - The answer: we may be able to estimate the long run if
    - Theory suggests the presence of this relationship.
    - The statistical properties for such a long run suggest to exist.
- Properties of integrated processes
  - If $X_t \sim I(0)$ and $Y_t \sim I(0)$, then $Y_t = \beta_0 + \beta_1 X_t \sim I(0)$
  - If $X_t \sim I(d)$ and $Y_t \sim I(0)$, then $Y_t = \beta_0 + \beta_1 X_t \sim I(d)$
  - If $X_t \sim I(d)$ and $Y_t \sim I(d)$, then $Y_t = \beta_0 + \beta_1 X_t \sim I(d)$
  - This means that if unit roots are the source of spurious regression, then we are faced with either $X_t \sim I(1)$ and $Y_t \sim I(0)$ or $X_t \sim I(1)$ and $Y_t \sim I(1)$.

Cointegration Test

- Granger Representation Theorem
  - If two variables $X_t \sim I(1)$ and $Y_t \sim I(1)$, then it is possible to find a linear combination $U_t \sim I(0)$.
  - This means:
    - There is a common trend between X_t and Y_t.
    - There is a long run error correction equilibrium.

- Engle-Granger Cointegration Test
  - Estimate the long run relationship: $Y_t = \beta_0 + \beta_1 X_t + U_t$
  - Save residuals: $\widehat{U_t}$
  - Run a unit root test on $\widehat{U_t}$
  - The null and alternative hypotheses

$$H_0: There\ is\ unit\ root\ (no\ cointegration)$$
$$H_1: There\ is\ no\ unit\ root\ (cointegration)$$

Error Correction Model

- One way to retrieve the long run equilibrium is to estimate an Error Correction Model
- The ECM model
  - Baseline model

$$\Delta Y_t = \alpha_0 + \alpha_1 \Delta X_t + \rho \hat{u}_{t-1} + v_t$$

  - Extended Model (Granger Causality)

$$\Delta Y_t = \alpha_0 + \alpha_1 \Delta X_t + \rho \hat{u}_{t-1} + \sum_{i=1}^{k} \theta_i \Delta X_{t-i} + \sum_{i=1}^{k} \delta_i \Delta X_{t-i} + v_t$$

- The interpretation of the ECM
  - α_1: The short run relationship.
  - ρ: The long run speed of adjustment.
  - ρ: it has to be negative and less than 1 in absolute values.

Example: Error Correction Model

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   dlot   R-squared:                       0.022
Model:                            OLS   Adj. R-squared:                  0.014
Method:                 Least Squares   F-statistic:                     2.636
Date:                Wed, 24 Mar 2021   Prob (F-statistic):             0.0335
Time:                        00:10:16   Log-Likelihood:                 2253.9
No. Observations:                 479   AIC:                            -4498.
Df Residuals:                     474   BIC:                            -4477.
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      0.0009      0.000      6.168      0.000       0.001       0.001
dlpt          -0.2771      0.116     -2.386      0.017      -0.505      -0.049
dlmt          -0.0043      0.039     -0.110      0.913      -0.081       0.072
dlrt           0.0013      0.001      1.486      0.138      -0.000       0.003
lresid         0.0025      0.001      1.828      0.068      -0.000       0.005
==============================================================================
Omnibus:                      104.298   Durbin-Watson:                   1.303
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1735.152
Skew:                           0.403   Prob(JB):                         0.00
Kurtosis:                      12.289   Cond. No.                     1.16e+03
==============================================================================
```

THANK YOU