MW82: Time Series Analysis, Tutorial VII

Dr. Giulia Sabbadini

DICE

2023/24

Recap: Tutorial VI

We went through:

- Stationary VAR(1) and VAR(p) theory and estimation
- Granger causality, IRFs

Today we will go through:

- Cointegration (Error Correction Models, Engle-Granger and Johansen procedure)

Cointegration

- Many economic time series are non-stationary, i.e. integrated, I(1).
- The linear combination of the variables can be stationary
- This occurs when the variables share the same stochastic trend
- The effect of the common stochastic trend may be contained when the variables are combined
- In these cases, we say that the variables are cointegrated

Cointegration

- If both $y_{1,t}$ and $y_{2,t}$ are I(1), they are **non-stationary**, i.e. have a stochastic trend
- Consider a linear combination s.t.: $y_{1,t} = \beta_2 y_{2,t} + u_t$.
- We need to distinguish two situations:
    1. If u_t is I(0) $\rightarrow$ the linear combination is stationary. $y_{1,t}$ and $y_{2,t}$ are **cointegrated**.
    2. If u_t is I(I) $\rightarrow$ **spurious regression**.

Cointegration

- Linear combination of $y_{1,t}$ and $y_{2,t}$ is stationary, whereas each series is non-stationary.
- This implies that processes move together $\rightarrow$ Long-run eqm.
- Possible examples (economics):
    - consumption and income
    - short and long-term interest rates
    - stock prices and earnings

Example: Cointegration in T-Bills

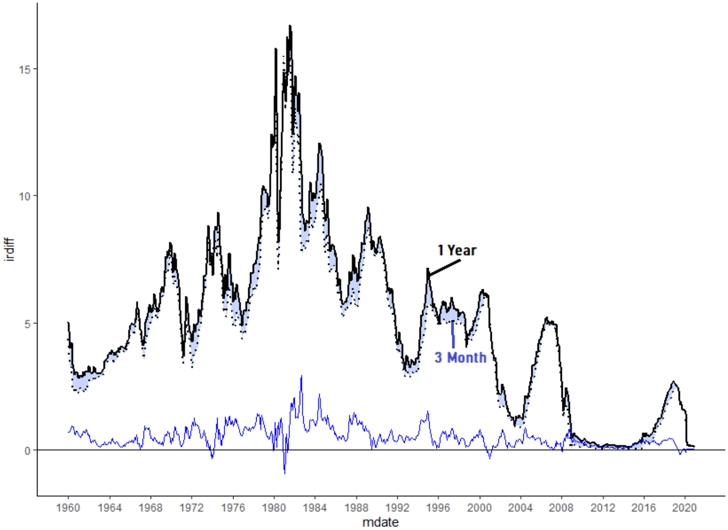


Figure: US Treasury Bill 1 Year (solid) and Treasury Bill 3 Month (dashed) yield curves. Difference in blue.

Spurious Regressions

- $y_{1,t}$ and $y_{2,t}$ are non-stationary, and have no relationship.
- Example: Two independent random walks.
- Problem (spurious regression): Regressing $y_{1,t}$ on $y_{2,t}$ will often yield highly significant results for $\hat{\beta}_2$ with a high R^2.
- Meaningless regression (truth: $\beta_2 = 0$).

Example: Spurious Regression

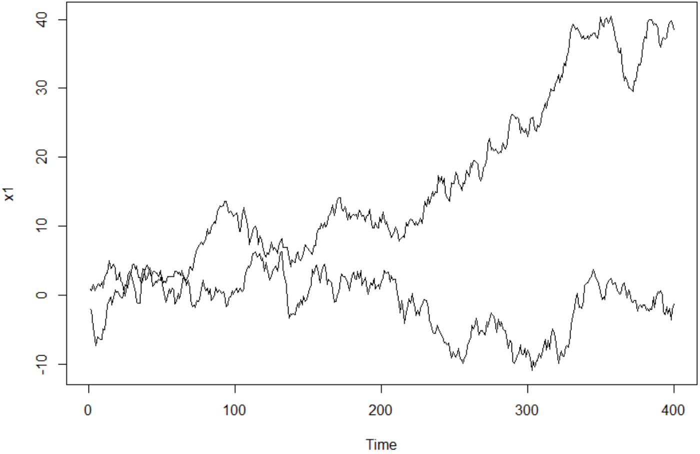


Figure: Two independent random walks.

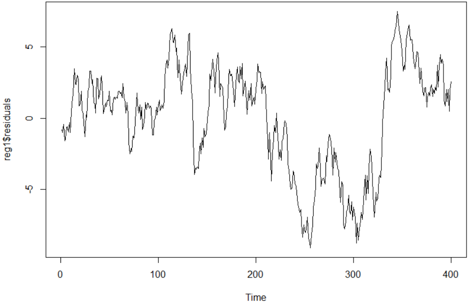
Results of Regression

Results:

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1.39120    0.29475   4.720 3.27e-06 ***
x2          -0.13667    0.01451  -9.418  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.686 on 398 degrees of freedom
Multiple R-squared:  0.1822,	Adjusted R-squared:  0.1802 
F-statistic:  88.7 on 1 and 398 DF,  p-value: < 2.2e-16
```

Residuals are a random walk too:



Cointegration Defined

- Given $\mathbf{y}_t = (y_{1,t}, y_{2,t})'$, a (2×1) vector of $I(1)$ variables, and $\beta = (1, -\beta_2)'$, it follows that:

$$\beta' \mathbf{y}_t = y_{1,t} - \beta_2 y_{2,t}.$$

- The system is cointegrated when $\beta' \mathbf{y}_t \sim I(0)$.
- The vector β is termed a cointegrating vector.

Cointegration and Equilibrium

- The *equilibrium* in a cointegration model refers to the existence of a long-run relationship.
- This can only occur if the two variables share a common equilibrium path.
- These variables will periodically move away from the equilibrium path, but the effect of this will not be permanent (i.e., the errors are stationary).
- The residuals from the cointegrated model are then described as *equilibrium errors*.
- A cointegrated system has an error correction representation
- The error correction model depicts the dynamics of a variable as a function of the deviations from long-run equilibrium

Error Correction Models

For example, if P_1 and P_2 are cointegrated share prices, assume:

- The gap between the prices is relatively large compared to the long-run equilibrium values (i.e., some dis-equilibria).
- The low-priced share P_2 must rise relative to the high-priced share P_1.

This may be accomplished by:

- An increase ($\uparrow$) in P_2 or a decrease ($\downarrow$) in P_1.
- An increase ($\uparrow$) in P_1 with a larger increase ($\uparrow$) in P_2.
- A decrease ($\downarrow$) in P_1 with a smaller decrease ($\downarrow$) in P_2.

OLS Regression and Stationary Errors

- The OLS regression then takes the form:

$$P_{1,t} = \beta_2 P_{2,t} + u_t$$

- When the errors are stationary:

$$u_t = \phi_1 u_{t-1} + \varepsilon_t \quad \text{with} \quad |\phi_1| < 1$$

- Combining the two:

$$u_t = P_{1,t} - \beta_2 P_{2,t}$$

$$P_{1,t} - \beta_2 P_{2,t} = \phi_1 (P_{1,t-1} - \beta_2 P_{2,t-1}) + \varepsilon_t$$

$$P_{1,t} = \beta_2 P_{2,t} + \phi_1 (P_{1,t-1} - \beta_2 P_{2,t-1}) + \varepsilon_t$$

Adjustment Mechanism with Stationary Errors

- Adding and subtracting $P_{1,t-1}$ and $\beta_2 P_{2,t-1}$:

$$\Delta P_{1,t} = -(1-\phi_1)(P_{1,t-1} - \beta_2 P_{2,t-1}) + (\beta_2 \Delta P_{2,t} + \varepsilon_{1,t})$$

$$= \alpha(P_{1,t-1} - \beta_2 P_{2,t-1}) + \varepsilon_{1,t}$$

- where $\alpha = -(1-\phi_1)$, $\Delta P_{2,t}$ is stationary and $\varepsilon_{1,t} = (\beta_2 \Delta P_{2,t} + \varepsilon_{1,t})$.
- The parameter α is the speed of adjustment that describes how changes to share prices react to past deviations from the equilibrium path in the respective share prices
- Note that large persistence in the autoregressive error would imply a slow speed of adjustment.

Adjustment Mechanism with Stationary Errors

- The error correction model depicts the dynamics of a variable as a function of the deviations from long-run equilibrium
- It can therefore be generalised to include lagged changes of both equations

$$\Delta y_{1t} = c_1 + \alpha_1(y_{1,t-1} - \beta_2 y_{2,t-1}) + \sum_j \psi_{11}^j \Delta y_{1,t-j} + \sum_j \psi_{12}^j \Delta y_{2,t-j} + \epsilon_{1t}$$

$$\Delta y_{2t} = c_2 + \alpha_2(y_{1,t-1} - \beta_2 y_{2,t-1}) + \sum_j \psi_{21}^j \Delta y_{1,t-j} + \sum_j \psi_{22}^j \Delta y_{2,t-j} + \epsilon_{2t}$$

- If both α are equal zero there is:
  - no equilibrium relationship
  - no error-correction
  - no cointegration

Adjustment Mechanism with Stationary Errors

Term	Description	Intuition
$y_{1,t-1} - \beta_2 y_{2,t-1}$	Cointegrated long-run equilibrium	Because this is an equilibrium relationship, it plays a role in dynamic paths of both y_{1t} and y_{2t}.
α_1, α_2	Adjustment coefficients	Captures the reactions of y_{1t} and y_{2t} to disequilibrium.
$\sum_j \psi_{11}^j \Delta y_{1,t-j} + \sum_j \psi_{12}^j \Delta y_{2,t-j}$	Autoregressive distributed lags	Captures additional dynamics.

Figure: Source: https://www.aptech.com/blog/a-guide-to-conducting-cointegration-tests/

Engle-Granger Procedure (ECM for n=2)

1. Plot both time series and test if both I(1)
2. (Does intuition or economic theory suggest a cointegrating relationship?)
3. Estimate $y_{1t} = \beta_0 + \beta_2 y_2 t + u_t$
4. These vars are cointegrated if u_t is stationary $\rightarrow$ Test if $\hat{u}_t$ is stationary with an ADF test [1]

$$\Delta \hat{u}_t = \hat{\pi}_1 \hat{u}_{t-1} + k \sum_{j=1}^{\infty} \gamma_j \Delta \hat{u}_{t-j} + \epsilon_t$$

   where $\hat{\pi}_1 = (1 - \phi)$.

5. If stationary $\rightarrow$ estimate an error correction model by substituting $(y_{1,t-1} - \beta_2 y_{2,t-1})$ with $\hat{u}_t$

[1] use critical values from Engle and Yoo (1987) or MacKinnon (2010, Tbl. 1)

Summary of the model

In the previous bivariate case:

$$\begin{bmatrix}\Delta y_{1,t}\\ \Delta y_{2,t}\end{bmatrix} = \begin{bmatrix}c_1\\ c_2\end{bmatrix} + \begin{bmatrix}\alpha_1 & \alpha_2\end{bmatrix} \begin{bmatrix}\beta_1 & \beta_2\end{bmatrix} \begin{bmatrix}y_{1,t-1}\\ y_{2,t-1}\end{bmatrix} + \begin{bmatrix}u_{1,t}\\ u_{2,t}\end{bmatrix}$$

which can be written as

$$\Delta y_{1,t} = c_1 + \alpha_1(\beta_1 y_{1,t-1} + \beta_2 y_{2,t-1}) + u_{1,t}$$

$$\Delta y_{2,t} = c_2 + \alpha_2(\beta_1 y_{1,t-1} + \beta_2 y_{2,t-1}) + u_{2,t}$$

and the cointegration relationship $\beta' y_t$ is given by,

$$\beta' y_t = \beta_1 y_{1,t} + \beta_2 y_{2,t} \sim I(0)$$

VECM model

In the three-variable case, we have $n = 3$ and $r = 2$:

$$\begin{bmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \\ \Delta y_{3,t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix}$$

With $r = 2$, there are two cointegration (linear) relationships:

$$\beta_1' y_t = \beta_{11} y_{1,t} + \beta_{21} y_{2,t} + \beta_{31} y_{3,t} \sim I(0)$$

$$\beta_2' y_t = \beta_{12} y_{1,t} + \beta_{22} y_{2,t} + \beta_{32} y_{3,t} \sim I(0)$$

Johansen procedure

- Cointegration test for $n \geq 2$ time series.
- Possibility: There can be $r < m$ cointegration relationships $(r = 0, 1, \ldots, m - 1)$.
  1. Choose lag order of VAR in levels
  2. Apply Johansen test (trace or eigenvalue method) accounting for constant, trend, both or no terms in the cointegration relationship
  3. Create VECM with r cointegration relationships
- in R: `urca::ca.jo` or `tsDyn::rank.test` .

Exercise I

- Load the *yields.csv* dataset. It contains quarterly data on yields from US treasury bonds (10 year, 3 month) from 1962Q1.
- Plot the time series; discuss stationarity.
- Test for cointegration and apply the Engle-Granger method for estimating an ECM, if applicable.
- Interpret the regression output.