Time Series Analysis and Forecasting

Chapter 7: Cointegration and VECM

Daniel Traian PELE

Bucharest University of Economic Studies
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Romanian Academy, Institute for Economic Forecasting
MSCA Digital Finance

Learning Objectives

By the end of this chapter, you will be able to:

- Understand the problem of spurious regression with non-stationary data
- Test for cointegration using Engle-Granger and Johansen methods
- Estimate Vector Error Correction Models (VECM)
- Interpret error correction mechanisms and adjustment speeds

Outline

Why Cointegration Matters

The Challenge

- Many economic/financial time series are **non-stationary** (I(1))
- GDP, stock prices, exchange rates, interest rates all have unit roots
- Standard regression with I(1) variables $\Rightarrow$ **spurious results**
- Differencing removes non-stationarity but loses **long-run information**

The Solution: Cointegration

Some non-stationary series share a **common stochastic trend**—they move together in the long run.

Nobel Prize 2003

Granger & Engle received the Nobel Prize for "methods for analyzing economic time series with common trends."

Real-World Applications

Finance

- **Pairs Trading**: Cointegrated stocks
- **Term Structure**: Interest rates
- **Spot-Futures**: Arbitrage

Macroeconomics

- **Consumption & Income**
- **Money & Prices**
- **PPP**: Exchange rates

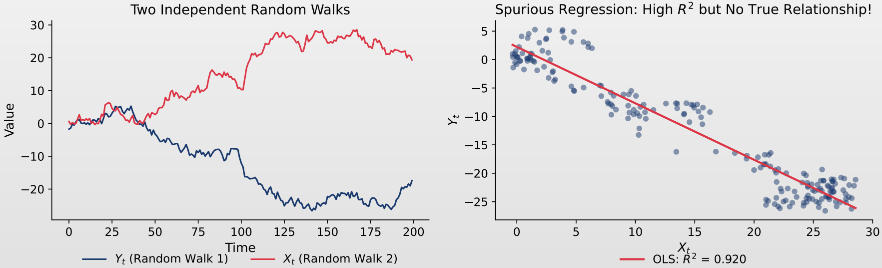
Policy Analysis

- **Fiscal**: Spending & taxes
- **Monetary**: Rate pass-through
- **Labor**: Wages & productivity

Spurious Regression: Visual Example

Warning

- **Result**: Two completely independent random walks show high correlation ($R^2 > 0.8$) purely by chance — in fact, $R^2 \to 1$ as $T \to \infty$ (Phillips, 1986). This is why we need cointegration analysis



TSA_ch7_spurious_regression

The Spurious Regression Problem

Granger & Newbold (1974)

Regressing one random walk on another **independent** random walk: $Y_t = \alpha + \beta X_t + u_t$ where Y_t and X_t are independent I(1) processes.

Symptoms of Spurious Regression

- High R^2 (often > 0.9) even though variables are **unrelated**
- Highly significant t-statistics (reject $H_0 : \beta = 0$)
- Very low Durbin-Watson statistic ($DW \approx 0$)
- Residuals are non-stationary (have unit root)

Rule of Thumb

If $R^2 > DW$, suspect spurious regression.

Spurious Correlations in the Real World

Data Mining Can Produce Meaningless Correlations

With enough variables and long time series, purely coincidental patterns emerge:

- Distance between Neptune and Uranus $\leftrightarrow$ SAP SE stock price (2002–2023)
- GMO corn use in South Dakota $\leftrightarrow$ Google searches for "i cant even" (2004–2023)
- *Two and a Half Men* season ratings $\leftrightarrow$ Jet fuel used in Serbia (2006–2015)
- "Its Wednesday my dudes" meme popularity $\leftrightarrow$ Boeing stock price (2006–2023)

Lesson

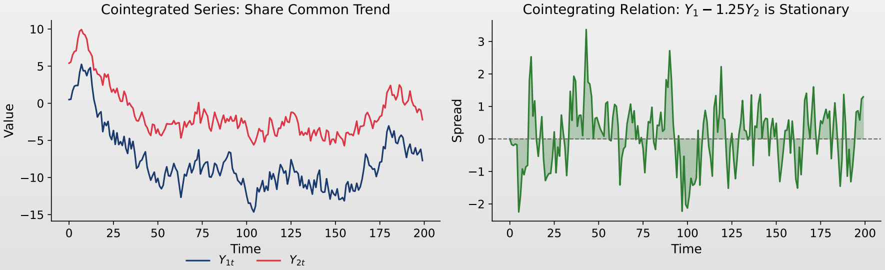
- High correlation $\neq$ causation
- Non-stationary series with common trends produce high R^2 by construction
- Always test for stationarity and cointegration before interpreting regression results.

Explore more examples: `tylervigen.com/spurious-correlations`

Cointegration: Visual Example

Key Insight

- **Cointegration**: Both series are I(1) and trend together, but their linear combination (spread) is stationary — this is cointegration.



TSA_ch7_cointegrated_series

Definition of Cointegration

Definition 1 (Cointegration (Engle & Granger, 1987))

Variables $Y_{1t}, Y_{2t}, \ldots, Y_{kt}$ are **cointegrated of order** (d, b), written $CI(d, b)$, if:

1. All variables are integrated of order d: $Y_{it} \sim I(d)$
2. There exists a linear combination $\boldsymbol{\beta}'\mathbf{Y}_t = \beta_1 Y_{1t} + \cdots + \beta_k Y_{kt}$ that is integrated of order $(d - b)$, where $b > 0$

Most Common Case: $CI(1, 1)$

- Variables are $I(1)$ (have unit roots)
- Linear combination is $I(0)$ (stationary)
- Vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)'$ is the **cointegrating vector**

The cointegrating vector is unique only up to scalar multiplication. Usually normalized: $\beta_{\mathbf{1}} = 1$.

Intuition: Common Stochastic Trends

Why Does Cointegration Occur?

Cointegrated variables share **common stochastic trends**: $Y_{1t} = \gamma_1 \tau_t + S_{1t}$, $Y_{2t} = \gamma_2 \tau_t + S_{2t}$ where τ_t is a common random walk and S_{it} are stationary.

Linear Combination Eliminates the Trend

$\gamma_2 Y_{1t} - \gamma_1 Y_{2t} = \gamma_2 S_{1t} - \gamma_1 S_{2t} \sim I(0)$

Economic Interpretation

- Cointegration = **long-run equilibrium relationship**
- Variables may deviate in the short run, but are *pulled back*
- The cointegrating vector defines the equilibrium

Cointegrating Rank

How Many Cointegrating Relationships?

For k variables that are $I(1)$:

- Maximum possible cointegrating relationships: $r = k - 1$
- If $r = 0$: No cointegration (variables drift apart)
- If $r = k$: All variables are $I(0)$ (inconsistent with the $I(1)$ assumption; use VAR in levels)

Example: 3 Variables

- $r = 0$: No cointegration
- $r = 1$: One cointegrating relationship
- $r = 2$: Two cointegrating relationships (only 1 common trend)

The number of common stochastic trends $= k - r$

Engle-Granger Two-Step Method

Step 1: Estimate Cointegrating Regression

Run OLS: $Y_t = \alpha + \beta X_t + e_t$. **Save residuals:** $\hat{e}_t = Y_t - \hat{\alpha} - \hat{\beta} X_t$

Step 2: Test Residuals for Stationarity

Test if $\hat{e}_t$ **is** $I(0)$ **using ADF:** $\Delta \hat{e}_t = \rho \hat{e}_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta \hat{e}_{t-j} + v_t$

- H_0: $\rho = 0$ (unit root $\Rightarrow$ no cointegration)
- H_1: $\rho < 0$ (stationary $\Rightarrow$ cointegration)

Important

Use **Engle-Granger critical values**, not standard ADF! (More negative because residuals are estimated)

Engle-Granger Critical Values

Critical Values for Cointegration Test

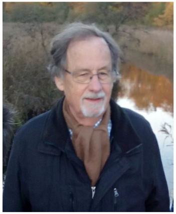
Variables	1%	5%	10%
2	−3.90	−3.34	−3.04
3	−4.29	−3.74	−3.45
4	−4.64	−4.10	−3.81
5	−4.96	−4.42	−4.13

MacKinnon (1991), $T = 100$

Limitations of Engle-Granger

- Only tests for **one** cointegrating relationship
- Results depend on choice of dependent variable
- Small sample bias; cannot test hypotheses on cointegrating vector

Researcher Spotlight: Søren Johansen



*1939

W Wikipedia

Biography

- Danish statistician and econometrician, Professor Emeritus at University of Copenhagen
- Known for his rigorous mathematical approach to econometrics
- Fellow of the Econometric Society; recipient of numerous honors in statistical science

Key Contributions

- **Johansen cointegration test** (1988, 1991) — maximum likelihood approach to testing for multiple cointegrating vectors
- **Trace and maximum eigenvalue** statistics for determining cointegration rank
- **VECM estimation** — linking cointegration with error correction models
- Standard framework for multivariate cointegration analysis in economics and finance

Johansen Cointegration Test

Advantages over Engle-Granger

- Tests for **multiple** cointegrating relationships
- Maximum likelihood estimation (more efficient)
- Can test restrictions on cointegrating vectors
- Does not require choosing a dependent variable

Starting Point: VAR in Levels

$$Y_t = c + A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + \varepsilon_t$$

Rewrite in **Vector Error Correction** form...

VECM Representation

Vector Error Correction Model

$\Delta \mathrm{Y}_t = \mathrm{c} + \boldsymbol{\Pi}\mathrm{Y}_{t-1} + \sum_{j=1}^{p-1} \boldsymbol{\Gamma}_j \Delta \mathrm{Y}_{t-j} + \varepsilon_t$

- $\boldsymbol{\Pi} = \sum_{i=1}^{p} \mathrm{A}_i - \mathrm{I}$ (long-run impact); $\boldsymbol{\Gamma}_j$ (short-run dynamics)

Key Insight: Rank of $\boldsymbol{\Pi}$

The **rank of $\boldsymbol{\Pi}$** determines cointegration:

- $\text{rank}(\boldsymbol{\Pi}) = 0$: No cointegration (VAR in differences)
- $\text{rank}(\boldsymbol{\Pi}) = k$: All variables are $I(0)$ (VAR in levels)
- $0 < \text{rank}(\boldsymbol{\Pi}) = r < k$: r cointegrating vectors

Derivation: From VAR to VECM

Starting Point: VAR(p) in Levels

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + \varepsilon_t$$

Step 1: Subtract Y_{t-1} from Both Sides

$$Y_t - Y_{t-1} = A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} - Y_{t-1} + \varepsilon_t$$

$$\Delta Y_t = (A_1 - I) Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + \varepsilon_t$$

Goal

Rewrite so that all terms are either in **levels** (Y_{t-1}) or **differences** (ΔY_{t-j}).

Derivation: From VAR to VECM (cont.)

Step 2: Add and Subtract Terms Strategically

Add $A_2 Y_{t-1}$ and subtract $A_2 Y_{t-1}$: $\Delta Y_t = (A_1 + A_2 - I)Y_{t-1} - A_2(Y_{t-1} - Y_{t-2}) + A_3 Y_{t-3} + \cdots + \varepsilon_t$
Continue adding $A_3 Y_{t-1}$, etc., until all lagged **levels** are collected in one term.

Step 3: General Pattern

After algebraic manipulation: $\Delta Y_t = \Pi Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + \varepsilon_t$

The Key Matrices

$$\boxed{\Pi = \sum_{i=1}^{p} A_i - I = -(I - A_1 - A_2 - \cdots - A_p)}$$

$$\boxed{\Gamma_j = -\sum_{i=j+1}^{p} A_i \text{ for } j = 1, \ldots, p-1}$$

Derivation: Verifying the $\boldsymbol{\Gamma}_j$ Formula

Example: VAR(2)

Starting from: $Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \varepsilon_t$

Subtract Y_{t-1}:

$$\Delta Y_t = (A_1 - I)Y_{t-1} + A_2 Y_{t-2} + \varepsilon_t$$

Add and subtract $A_2 Y_{t-1}$:

$$\Delta Y_t = (A_1 + A_2 - I)Y_{t-1} + A_2(Y_{t-2} - Y_{t-1}) + \varepsilon_t$$

$$\Delta Y_t = \underbrace{(A_1 + A_2 - I)}_{\boldsymbol{\Pi}} Y_{t-1} \underbrace{-A_2}_{\boldsymbol{\Gamma}_1} \Delta Y_{t-1} + \varepsilon_t$$

Verification

For VAR(2): $\boldsymbol{\Pi} = A_1 + A_2 - I$ and $\boldsymbol{\Gamma}_1 = -A_2$

Using our formula: $\boldsymbol{\Gamma}_1 = -\sum_{i=2}^{2} A_i = -A_2 \quad \checkmark$

Economic Interpretation of Error Correction

The VECM with Cointegration

When $\text{rank}(\mathbf{\Pi}) = r$, we write $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$: $\Delta \mathrm{Y}_t = \boldsymbol{\alpha} \underbrace{(\boldsymbol{\beta}'\mathrm{Y}_{t-1})}_{\text{equilibrium error}} + \sum_{j=1}^{p-1} \mathbf{\Gamma}_j \Delta \mathrm{Y}_{t-j} + \boldsymbol{\varepsilon}_t$

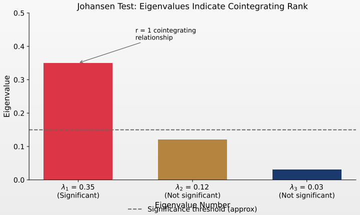
Economic Interpretation

- $\boldsymbol{\beta}'\mathrm{Y}_{t-1}$ = **equilibrium error**: deviation from long-run relationship
- $\boldsymbol{\alpha}$ = **adjustment speeds**: how fast variables correct deviations
- $\mathbf{\Gamma}_j$ = **short-run dynamics**: transitory effects

Error Correction Mechanism

If $\boldsymbol{\beta}'\mathrm{Y}_{t-1} > 0$ (above equilibrium) and $\alpha_i < 0$, then ΔY_{it} decreases. **The system self-corrects toward equilibrium.**

Johansen Test Statistics



When rank($\mathbf{\Pi}$) $= r < k$

$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ where $\boldsymbol{\beta}$ $(k \times r)$ = **cointegrating vectors**, $\boldsymbol{\alpha}$ $(k \times r)$ = **adjustment coefficients**

Interpretation

- $\boldsymbol{\beta}'\mathbf{Y}_{t-1}$ = deviations from equilibrium (error correction terms)
- $\boldsymbol{\alpha}$ = speed of adjustment; rows show each variable's response

VECM: $\Delta \mathbf{Y}_t = \mathbf{c} + \boldsymbol{\alpha}(\boldsymbol{\beta}'\mathbf{Y}_{t-1}) + \sum_{j=1}^{p-1} \mathbf{\Gamma}_j \Delta \mathbf{Y}_{t-j} + \boldsymbol{\varepsilon}_t$

TSA_ch7_johansen_eigenvalues

Testing Procedure

Sequential Testing (Trace Test)

1. Test H_0: $r = 0$. If rejected $\Rightarrow$ continue
2. Test H_0: $r \leq 1$. If not rejected $\Rightarrow r = 1$
3. Continue until H_0 is not rejected

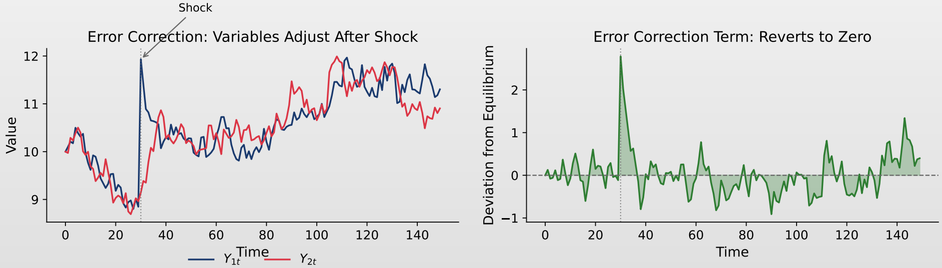
Deterministic Components

- No constant, no trend (rarely used)
- Constant in cointegrating relation only
- **Constant in both** (most common)
- Constant + trend in cointegrating relation

Error Correction Mechanism: Visual

Interpretation

- **Error correction**: When series deviate from equilibrium (shaded regions), the adjustment mechanism pulls them back. Positive deviations lead to downward adjustment, negative deviations lead to upward adjustment



TSA_ch7_error_correction

VECM Structure

Full VECM Specification

For $k = 2$ variables with $r = 1$ cointegrating relation:

$$\Delta Y_{1t} = c_1 + \alpha_1(Y_{1,t-1} - \beta Y_{2,t-1}) + \gamma_{11}\Delta Y_{1,t-1} + \gamma_{12}\Delta Y_{2,t-1} + \varepsilon_{1t}$$
$$\Delta Y_{2t} = c_2 + \alpha_2(Y_{1,t-1} - \beta Y_{2,t-1}) + \gamma_{21}\Delta Y_{1,t-1} + \gamma_{22}\Delta Y_{2,t-1} + \varepsilon_{2t}$$

Components

- $(Y_{1,t-1} - \beta Y_{2,t-1})$ = error correction term (deviation from equilibrium)
- α_1, α_2 = adjustment speeds (signs must ensure error correction toward equilibrium; typically opposite in bivariate systems)
- γ_{ij} = short-run dynamics
- ε_{it} = innovations

Interpreting Adjustment Coefficients

The α Coefficients

If the cointegrating relation is $Y_1 - \beta Y_2 = 0$ (equilibrium):

- $\alpha_1 < 0$: Y_1 adjusts downward when above equilibrium
- $\alpha_2 > 0$: Y_2 adjusts upward when Y_1 is above equilibrium

Weak Exogeneity

If $\alpha_i = 0$, variable Y_i does **not** respond to disequilibrium.

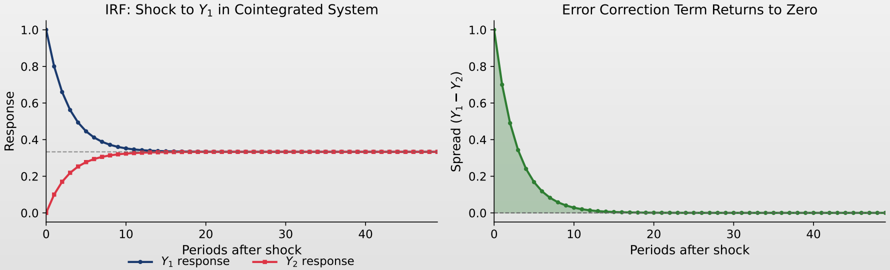
- Y_i is **weakly exogenous** for the long-run parameters
- The other variable does all the adjusting
- Can simplify estimation (single-equation approach)

Test weak exogeneity: $H_0 : \alpha_i = 0$ using likelihood ratio test.

VECM Impulse Response Functions

IRF Interpretation

- **Permanent effects**: In a cointegrated system, shocks have permanent effects on levels, but the system returns to equilibrium — they converge to a new long-run value



TSA_ch7_vecm_irf

VECM vs VAR in Differences

When Variables are Cointegrated

	VAR in Differences	VECM
Long-run info	Lost	Preserved
Short-run dynamics	Yes	Yes
Error correction	No	Yes
Forecasting	Poor (long-run)	Better
IRF interpretation	Short-run only	Both

Granger Representation Theorem

- If variables are cointegrated, there **must** exist an error correction representation
- Ignoring cointegration = model misspecification.

Practical Workflow

Step-by-Step Procedure

1. **Unit Root Tests**: Verify all variables are $I(1)$
   - ADF, KPSS on levels and first differences
2. **Lag Length Selection**: Choose p for VAR in levels
   - Use AIC, BIC, or sequential LR tests
3. **Cointegration Test**: Johansen trace/max-eigenvalue tests
   - Determine cointegrating rank r
4. **Estimate VECM**: If $0 < r < k$
   - Estimate $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\Gamma}_j$
5. **Diagnostics**: Check residuals for autocorrelation, normality
6. **Analysis**: IRF, FEVD, hypothesis tests

Common Pitfalls

Things to Watch Out For

- **Structural breaks**: Cause spurious unit roots or cointegration
- **Near-unit-root**: Tests have low power
- **Lag selection**: Too many/few lags bias results
- **Small samples**: Johansen test oversized

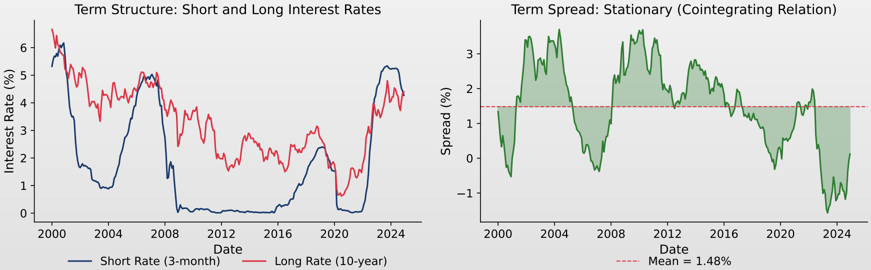
Recommendation

Always check: residual diagnostics, stability of cointegrating relationship, sensitivity to specification

Example 1: Term Structure of Interest Rates

Expectations Hypothesis

- **Conclusion**: Short and long rates share a common trend. The spread (term premium) is stationary — evidence of cointegration.

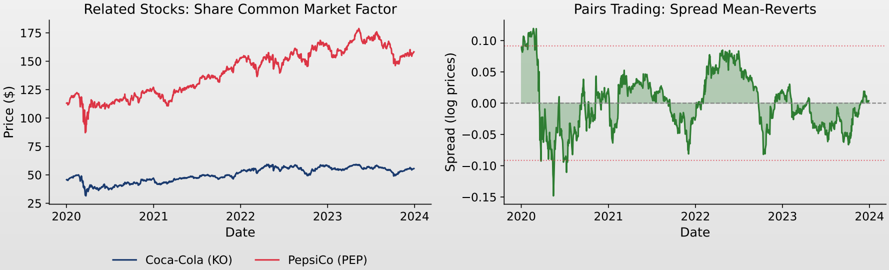


TSA_ch7_interest_rates_coint

Example 2: Pairs Trading in Finance

Strategy

- **Pairs trading**: Find cointegrated stock pairs (e.g., Coca-Cola & Pepsi). When the spread deviates from the mean, trade expecting mean reversion



TSA_ch7_pairs_trading

Interest Rates: Economic Theory

Expectations Hypothesis of the Term Structure

- **Formula**: Long-term rate as average of expected future rates
  - $R_t^{(n)} = \frac{1}{n}\sum_{i=0}^{n-1} E_t[r_{t+i}] + \text{term premium}$
- **Implication**: If the term premium is constant, r_t and R_t are cointegrated
  - Cointegrating vector: $(1, -1)$

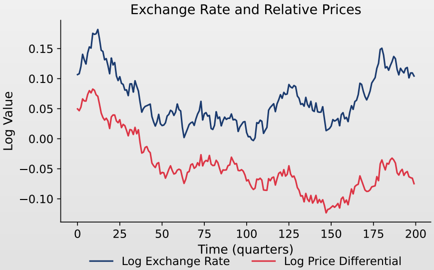
Empirical Results

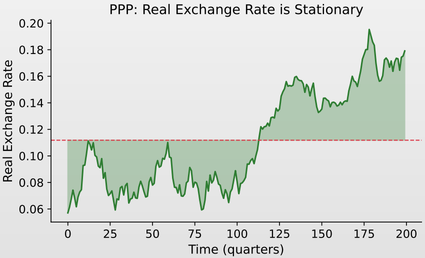
- **Unit root tests**: Both rates are $I(1)$
  - One cointegrating relationship (Johansen test)
- **Cointegrating vector**: $\approx (1, -1)$, the spread is stationary
  - The short rate adjusts to disequilibrium (the long rate is weakly exogenous)

Example 3: Purchasing Power Parity (PPP)

PPP Theory

- **Formula**: $e_t = p_t - p_t^*$ (log exchange rate equals price differential). The real exchange rate should be stationary in the long run





TSA_ch7_ppp_cointegration

VECM Results for Interest Rates

Typical Results

- **Integration**: Both rates are $I(1)$, one cointegrating relationship identified
  - Cointegrating vector close to $(1, -1)$: the spread is stationary
- **Adjustment**: The short rate adjusts to the long rate
  - The long rate does not adjust (weakly exogenous)

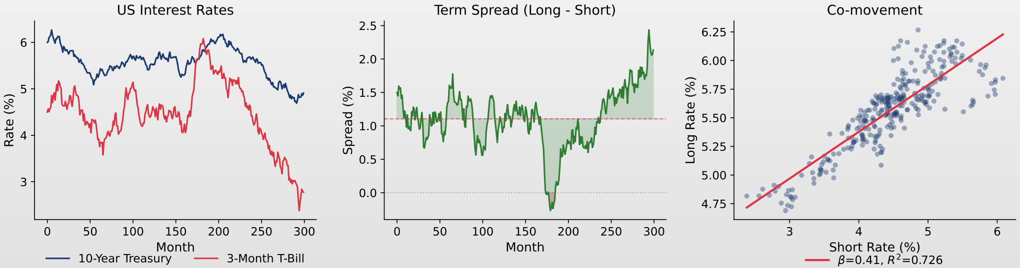
VECM Equations (Stylized)

- **Estimated system**:
  - $\Delta r_t = 0.02 - 0.15(r_{t-1} - R_{t-1}) + \text{lags} + \varepsilon_{1t}$
  - $\Delta R_t = 0.01 - 0.02(r_{t-1} - R_{t-1}) + \text{lags} + \varepsilon_{2t}$
- **Interpretation**: The short rate adjusts faster ($\alpha_1 = -0.15$)
  - The long rate is nearly weakly exogenous ($\alpha_2 \approx 0$)

Case Study: Cointegration of Interest Rates

Data

- **US Interest Rates**: Long-term (10 years) and short-term (3 months)
- **Observation**: Both series are I(1), but the spread appears stationary

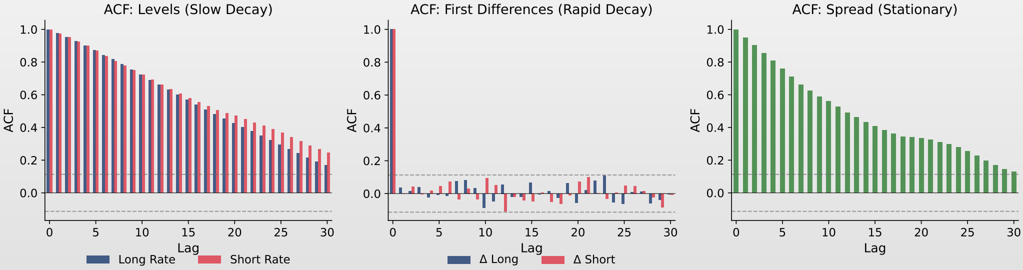


TSA_ch7_case_raw_data

Step 1: Unit Root Tests

Results

- **ACF levels**: Slow decay — non-stationarity; after differencing: rapid decay — I(1)
- **ACF spread**: Stationary — possible cointegration.

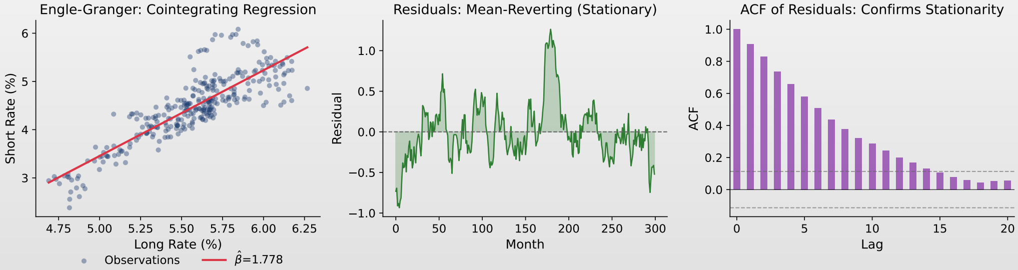


TSA_ch7_case_unit_root

Step 2: Engle-Granger Cointegration Test

Results

- **Engle-Granger regression**: Short rate $= \alpha + \beta \times$ Long rate $+ \varepsilon_t$
- **Conclusion**: The series are cointegrated — a long-run equilibrium relationship exists

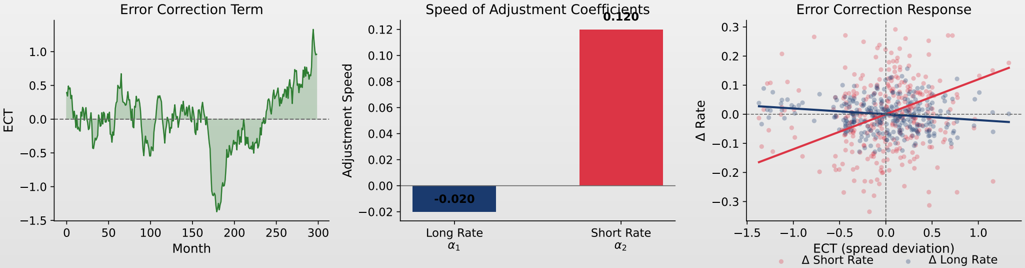


TSA_ch7_case_cointegration

Step 3: VECM Estimation

Model

- **VECM(2)**: Cointegration rank = 1
- **Adjustment**: The α coefficients indicate the speed of return to equilibrium

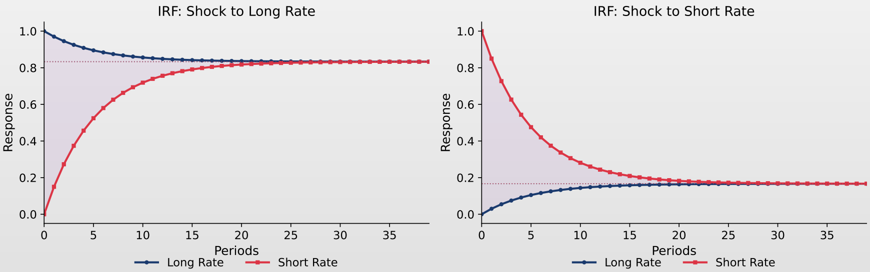


TSA_ch7_case_vecm

Step 4: Impulse Response Functions

Interpretation

- **Permanent effects**: Shocks to the long rate persistently affect both rates
- **Cointegration**: Effects do not converge to zero — characteristic of cointegrated series

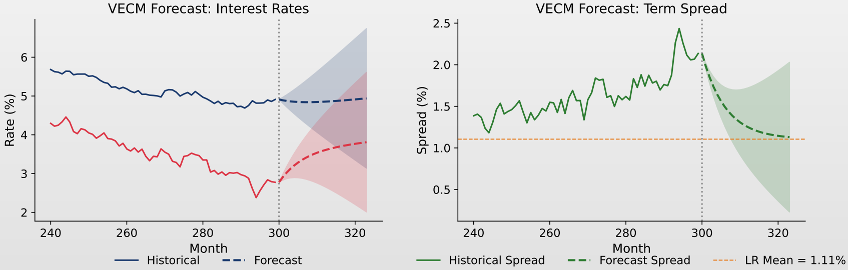


TSA_ch7_case_irf

Step 5: VECM Forecast

Forecast

- **Horizon**: 24 months for both rates simultaneously
- **Advantage**: VECM maintains the cointegrating relationship in the forecast



TSA_ch7_case_forecast

AI Exercise: Critical Thinking

Prompt to test in ChatGPT / Claude / Copilot

"Using yfinance, download daily closing prices for gold (GC=F) and silver (SI=F) from 2019-01-01 to 2024-12-31 (approx. 1,500 observations). Test whether each series is I(1), test for cointegration using both Engle-Granger and Johansen methods, and estimate a VECM. Analyze the speed of adjustment parameters. Give me complete Python code."

Exercise:

1. Run the prompt in an LLM of your choice and critically analyze the response.
2. Does it verify that each series is I(1) before testing for cointegration?
3. Does it use both Engle-Granger and Johansen tests? What are the trade-offs?
4. How does it determine the cointegration rank? Trace vs max-eigenvalue statistics?
5. Does it correctly interpret the α (adjustment) coefficients?

Warning: AI-generated code may run without errors and look professional. *That does not mean it is correct.*

Key Takeaways

Main Concepts

- **Cointegration**: $I(1)$ variables with stationary linear combination
- **Spurious regression**: High R^2 with unrelated $I(1)$ variables
- **VECM**: VAR with error correction for cointegrated systems

Testing Methods

- **Engle-Granger**: Simple, one vector only
- **Johansen**: Multiple vectors, MLE-based

Remember

- Tests have low power in small samples
- Theory should guide specification

What's Next?

Extensions and Related Topics

- **Structural VECM**: Identifying structural shocks
- **Threshold cointegration**: Nonlinear adjustment
- **Panel cointegration**: Multiple cross-sections
- **Fractional cointegration**: Long memory
- **Time-varying cointegration**: Regime changes

- **Questions?**

Key Formulas – Summary

Cointegration

- **Definition**: $Y_t - \beta X_t = u_t \sim I(0)$
- **Interpretation**: Long-run equilibrium

Engle-Granger Test

- **Step 1**: $Y_t = \alpha + \beta X_t + u_t$
- **Step 2**: ADF test on $\hat{u}_t$
- **Note**: Special critical values

Cointegration Rank

- **Rank** r: $0 \leq r \leq K - 1$ relationships

VECM Model

- **Equation**: $\Delta \mathbf{Y}_t = \mathbf{\Pi} \mathbf{Y}_{t-1} + \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta \mathbf{Y}_{t-i} + \boldsymbol{\varepsilon}_t$
- **Factorization**: $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$

Interpretation of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$

- $\boldsymbol{\beta}$: Cointegrating vectors
- $\boldsymbol{\alpha}$: Speed of adjustment

Johansen Test

- **Trace**: $\lambda_{trace} = -T \sum_{i=r+1}^{K} \ln(1 - \hat{\lambda}_i)$
- **Max-Eigen**: $\lambda_{max} = -T \ln(1 - \hat{\lambda}_{r+1})$

Question 1

Question

□ Analyze the two pairs of I(1) series below. Which pair is cointegrated?

Answer Choices

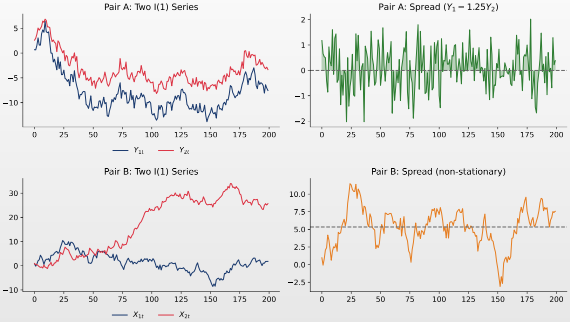
(A) Pair A, because the series have the same trend

(B) Pair B, because the series are uncorrelated

(C) Pair A, because their spread is stationary

(D) Both pairs are cointegrated

Question 1: Answer



Answer: (C)

- Cointegration = stationary linear combination, not just correlation
- Pair B's spread is non-stationary ⇒ not cointegrated

TSA_ch7_quiz1_cointegration_concept

Question 2

Question

- Both regressions below have high R^2. How can you distinguish a spurious regression from a genuine one?

Answer Choices

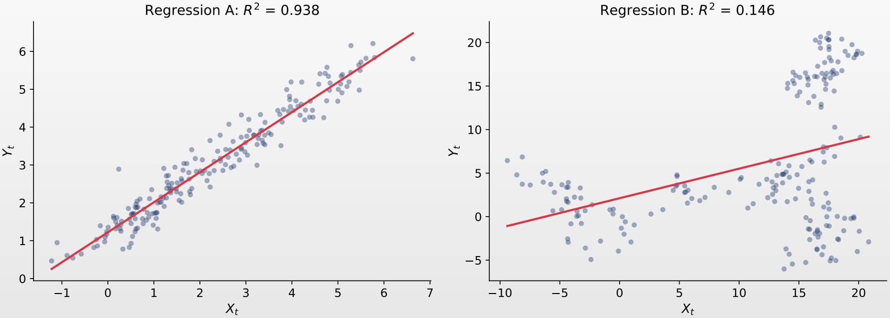
(A) Cannot distinguish – both have high R^2

(B) Test the residuals: stationary residuals = genuine cointegration

(C) Check the significance of the β coefficient

(D) Compare R^2 values: higher = more real relationship

Question 2: Answer



Answer: (B)

- Engle-Granger test: if OLS residuals are stationary (ADF), the relationship is genuine
- High R^2 does NOT imply a real relationship between $I(1)$ variables.

TSA_ch7_quiz2_spurious_regression

Question 3

Question

- In which scenario is the long rate weakly exogenous (does not adjust to disequilibrium)?

Answer Choices

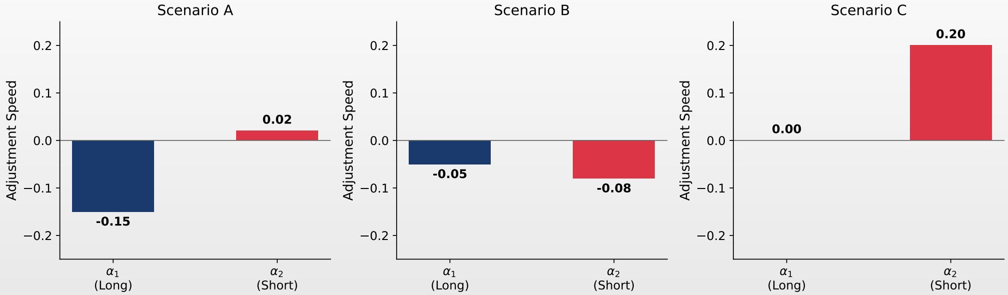
(A) Scenario A: $\alpha_1 = -0.15$, $\alpha_2 = 0.02$

(B) Scenario B: $\alpha_1 = -0.05$, $\alpha_2 = -0.08$

(C) Scenario C: $\alpha_1 = 0.00$, $\alpha_2 = 0.20$

(D) No scenario – both variables must adjust

Question 3: Answer



Answer: (C)

- $\alpha_1 = 0$: the long rate does not respond to disequilibrium (weakly exogenous)
- All adjustment is done by the short rate ($\alpha_2 = 0.20$)

TSA_ch7_quiz3_error_correction

Question 4

Question

- Given the Johansen Trace test results for $K = 3$ variables, what is the cointegrating rank?

Answer Choices

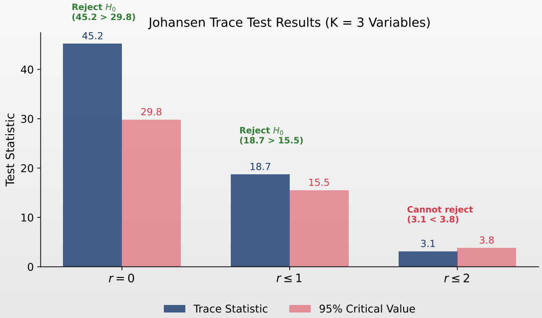
(A) $r = 0$ (no cointegrating relationships)

(B) $r = 1$ (one cointegrating relationship)

(C) $r = 2$ (two cointegrating relationships)

(D) $r = 3$ (fully stationary system)

Question 4: Answer



Answer: (C)

- Reject $H_0 : r = 0$ (45.2 > 29.8) and $H_0 : r \leq 1$ (18.7 > 15.5)
- Cannot reject $H_0 : r \leq 2$ (3.1 < 3.8) $\Rightarrow$ rank is $r = 2$

TSA_ch7_quiz4_johansen_rank

Question 5

Question

☐ What is the main advantage of VECM over VAR in differences for forecasting?

Answer Choices

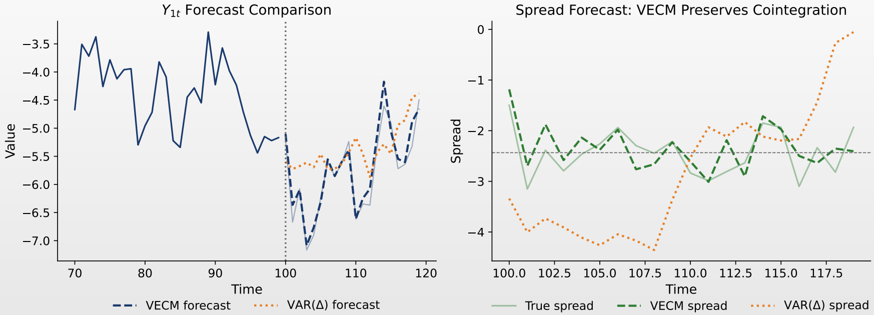
(A) VECM has fewer parameters to estimate

(B) VECM preserves the cointegrating relationship in long-run forecasts

(C) VAR in differences cannot produce forecasts

(D) No advantage – both are equivalent

Question 5: Answer



Answer: (B)

- VAR(Δ) loses the level relationship ⇒ spread diverges
- VECM incorporates long-run equilibrium ⇒ forecast stays coherent

TSA_ch7_quiz5_vecm_forecast

Bibliography I

Fundamental Cointegration Papers

- Engle, R.F., & Granger, C.W.J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55(2), 251–276.
- Johansen, S. (1988). Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, 12(2-3), 231–254.
- Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, 59(6), 1551–1580.

VECM and Cointegration Textbooks

- Juselius, K. (2006). *The Cointegrated VAR Model: Methodology and Applications*, Oxford University Press.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer.

Bibliography II

Tests and Applications

- Phillips, P.C.B., & Ouliaris, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica*, 58(1), 165–193.
- Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press.
- Banerjee, A., Dolado, J.J., Galbraith, J.W., & Hendry, D.F. (1993). *Co-Integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press.

Online Resources and Code

- **Quantlet**: `https://quantlet.com` – Code platform for quantitative methods
- **Quantinar**: `https://quantinar.com` – Learning platform for quantitative methods
- **GitHub TSA**: `https://github.com/QuantLet/TSA/tree/main/TSA_ch7` – Python code for this chapter

Thank You!

Questions?

Course materials available at: https://danpele.github.io/Time-Series-Analysis/

Quantlet Quantinar