Time Series Analysis and Forecasting

Chapter 6: Cointegration & VECM

Long-Run Equilibrium Relationships

Lecture 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. This long-run relationship can be modeled!

Nobel Prize 2003

Clive Granger received the Nobel Prize in Economics (with Robert Engle) for developing cointegration analysis—"methods for analyzing economic time series with common trends."

Real-World Applications

Finance

- **Pairs Trading**: Trade the spread between cointegrated stocks
- **Term Structure**: Short and long interest rates
- **Spot-Futures**: Arbitrage relationships

Macroeconomics

- **Consumption & Income**: Permanent income hypothesis
- **Money & Prices**: Quantity theory of money
- **PPP**: Exchange rates and price levels

Policy Analysis

- **Fiscal Policy**: Government spending and tax revenues
- **Monetary Policy**: Interest rate pass-through
- **Labor Markets**: Wages and productivity

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.

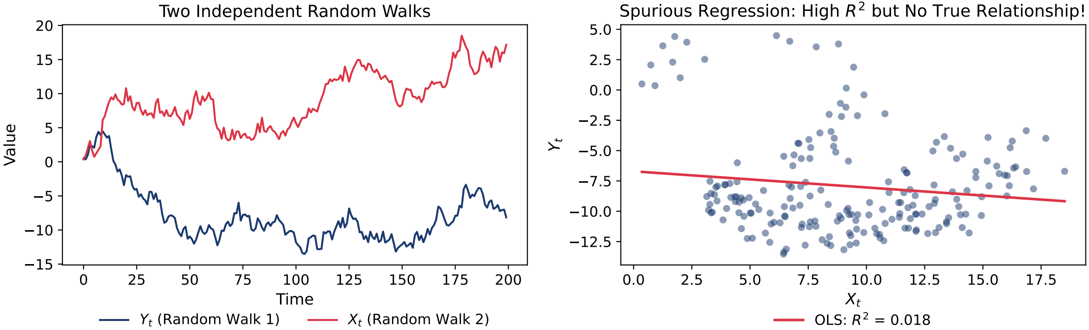
Symptoms of Spurious Regression

- 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 (Granger)

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

Spurious Regression: Visual Example



Warning: Two completely independent random walks show high correlation ($R^2 > 0.8$) purely by chance! This is why we need cointegration analysis.

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:

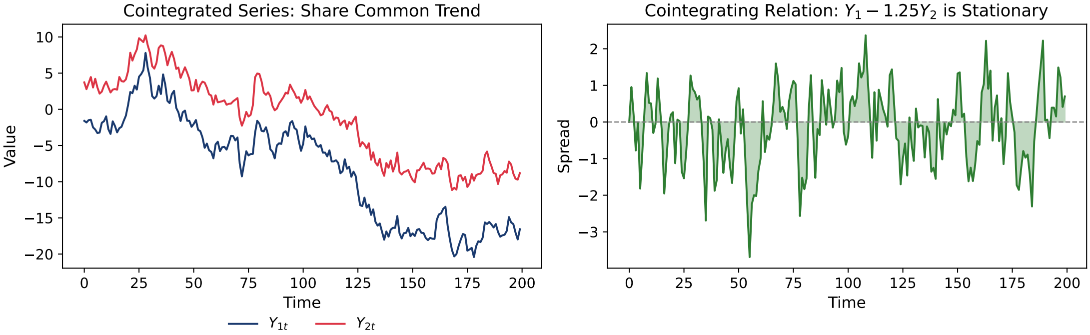
1. All variables are integrated of order d: $Y_{it} \sim I(d)$
2. There exists a linear combination $\beta' \mathbf{Y}_t = \beta_1 Y_{1t} + \cdots + \beta_k Y_{kt}$ that is integrated of order $(d - b)$, where $b > 0$

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: $\beta_1 = 1$.

Cointegration: Visual Example



Key insight: Both series are I(1) and trend together, but their linear combination (spread) is stationary—this is cointegration!

Intuition: Common Stochastic Trends

Why Does Cointegration Occur?

Cointegrated variables share **common stochastic trends**:

$$Y_{1t} = \gamma_1 \tau_t + S_{1t}, \qquad Y_{2t} = \gamma_2 \tau_t + S_{2t}$$

where τ_t is a common random walk and S_{it} are stationary components.

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 represents a **long-run equilibrium relationship**
- Variables may deviate in the short run
- But they are "pulled back" to equilibrium over time
- 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 $I(0)$ (contradiction)

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 regression (assuming Y_t is the dependent variable):

$$Y_t = \alpha + \beta X_t + e_t$$

Save the 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 test:

$$\Delta \hat{e}_t = \rho \hat{e}_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta \hat{e}_{t-j} + v_t$$

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

Important

Engle-Granger Critical Values

Critical Values for Cointegration Test

Number of 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

Based on MacKinnon (1991) response surface estimates, $T = 100$

Limitations of Engle-Granger

- Only tests for **one** cointegrating relationship
- Results depend on which variable is chosen as dependent
- Small sample bias in estimated cointegrating vector
- Cannot test hypotheses on the cointegrating vector

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

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{A}_1\mathbf{Y}_{t-1} + \mathbf{A}_2\mathbf{Y}_{t-2} + \cdots + \mathbf{A}_p\mathbf{Y}_{t-p} + \varepsilon_t$$

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

VECM Representation

Vector Error Correction Model

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

where:

- $\mathbf{\Pi} = \mathbf{A}_1 + \mathbf{A}_2 + \cdots + \mathbf{A}_p - \mathbf{I}$ (long-run impact matrix)
- $\mathbf{\Gamma}_j = -(\mathbf{A}_{j+1} + \cdots + \mathbf{A}_p)$ (short-run dynamics)

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

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

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

Decomposition of $\mathbf{\Pi}$

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

The matrix $\mathbf{\Pi}$ can be decomposed as:

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$$

where:

- $\boldsymbol{\beta}$ is $k \times r$ matrix of **cointegrating vectors**
- $\boldsymbol{\alpha}$ is $k \times r$ matrix of **adjustment coefficients**

Interpretation

- $\boldsymbol{\beta}'\mathbf{Y}_{t-1}$ = deviations from long-run equilibrium (error correction terms)
- $\boldsymbol{\alpha}$ = speed of adjustment to equilibrium
- Each row of $\boldsymbol{\alpha}$ shows how each variable responds to disequilibrium

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

Johansen Test Statistics

Two Test Statistics

Based on eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_k$ of a certain matrix:

Trace Test:

$$\lambda_{\text{trace}}(r) = -T \sum_{i=r+1}^{k} \ln(1 - \hat{\lambda}_i)$$

Tests H_0: rank $\leq r$ vs H_1: rank $> r$

Maximum Eigenvalue Test:

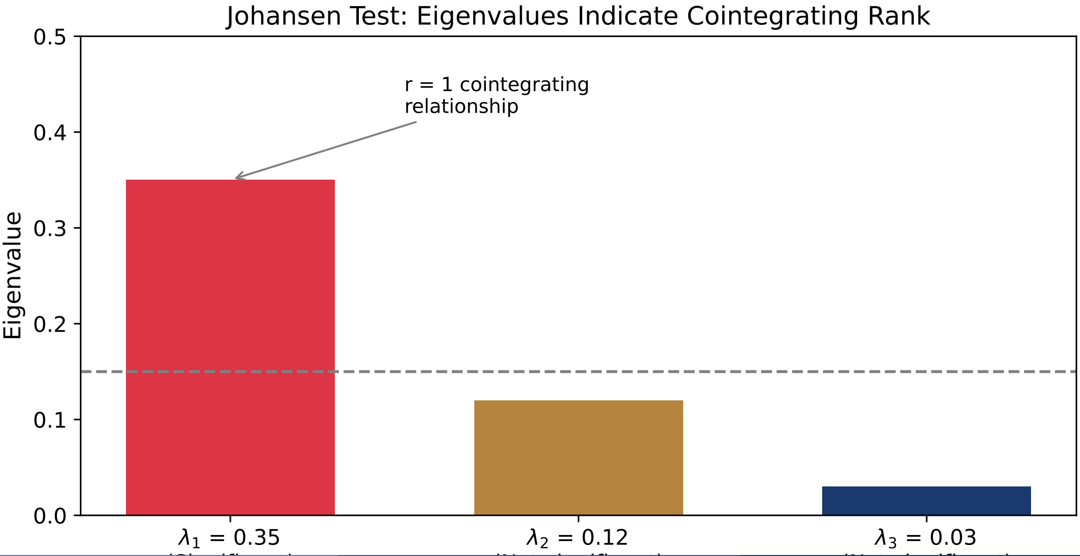
$$\lambda_{\max}(r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1})$$

Tests H_0: rank $= r$ vs H_1: rank $= r + 1$

Critical values from Johansen & Juselius (1990), depend on:

- Number of variables k
- Deterministic components (constant, trend)

Johansen Test: Visual Interpretation



Testing Procedure

Sequential Testing (Trace Test)

1. Test H_0: $r = 0$ vs H_1: $r > 0$
   - If not rejected: No cointegration. Stop.
   - If rejected: At least one cointegrating vector. Continue.
2. Test H_0: $r \leq 1$ vs H_1: $r > 1$
   - If not rejected: $r = 1$. Stop.
   - If rejected: At least two cointegrating vectors. Continue.
3. Continue until H_0 is not rejected...

Deterministic Components

Choose specification carefully:

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

VECM Structure

Full VECM Specification

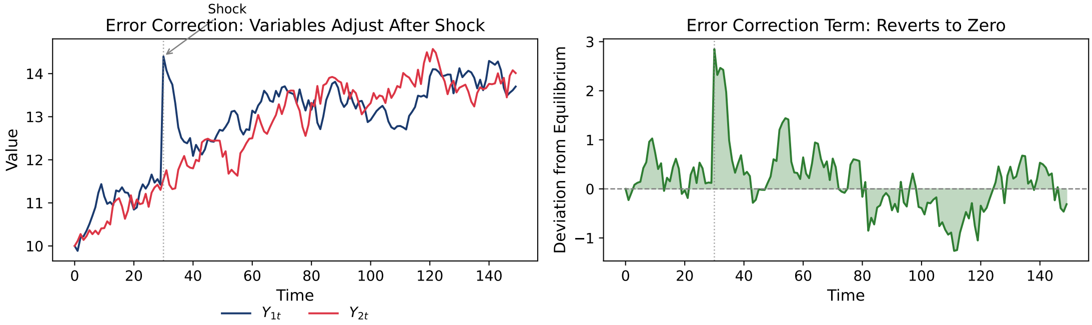
For $k = 2$ variables with $r = 1$ cointegrating relation:

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

Components

- $(Y_{1,t-1} - \beta Y_{2,t-1})$ = error correction term (deviation from equilibrium)
- α_1, α_2 = adjustment speeds (should have opposite signs)
- γ_{ij} = short-run dynamics
- ε_{it} = innovations

Error Correction Mechanism: Visual



Error correction in action: 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.

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 vs VAR in Differences

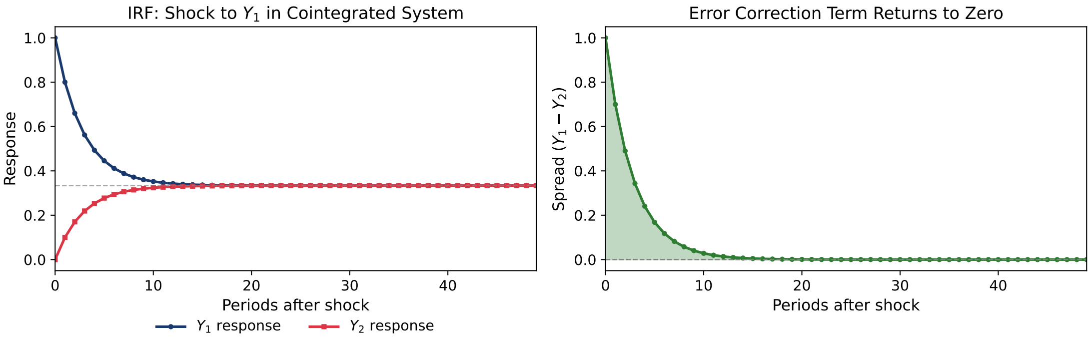
When Variables are Cointegrated

	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!

VECM Impulse Response Functions



IRF interpretation: In a cointegrated system, shocks have **permanent effects** on levels but the system returns to equilibrium. Unlike stationary VAR, effects don't decay to zero—they converge to a new long-run value.

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

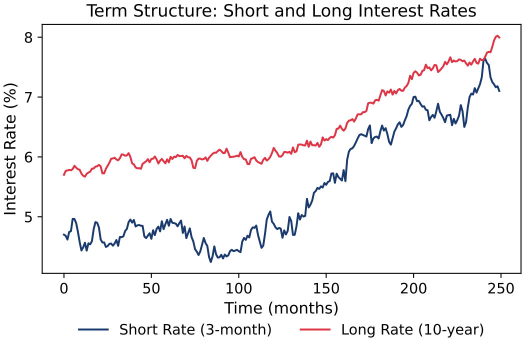
- **Structural breaks**: Can cause spurious unit roots or cointegration
- **Near-unit-root processes**: Tests have low power
- **Too many lags**: Over-parameterization, loss of efficiency
- **Too few lags**: Residual autocorrelation, biased estimates
- **Wrong deterministic specification**: Affects critical values
- **Small samples**: Johansen test oversized in small samples

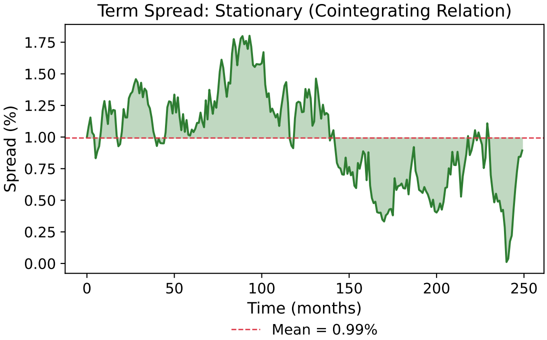
Recommendation

Always check:

- Residual diagnostics (Portmanteau test, normality)
- Stability of estimated cointegrating relationship over time
- Sensitivity to lag length and deterministic specification

Example 1: Term Structure of Interest Rates





Expectations Hypothesis: Short and long rates share common trend. The spread (term premium) is stationary—evidence of cointegration!

Interest Rates: Economic Theory

Expectations Hypothesis of Term Structure

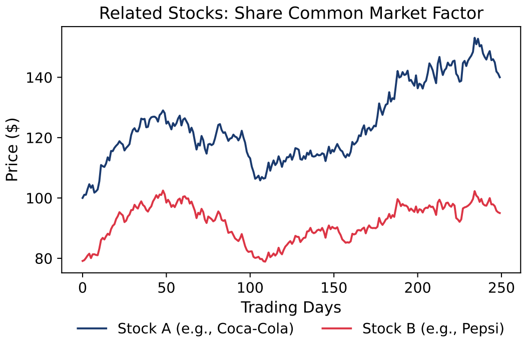
$$R_t^{(n)} = \frac{1}{n}\sum_{i=0}^{n-1} E_t[r_{t+i}] + \text{term premium}$$

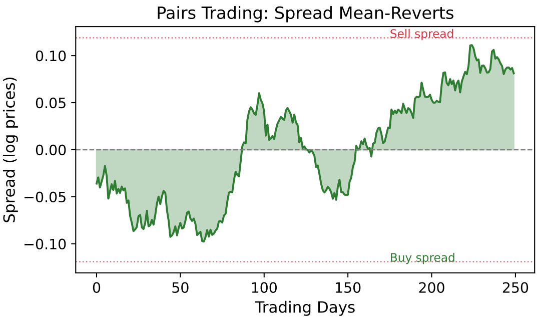
If term premium is constant, short rate r_t and long rate R_t should be cointegrated with vector $(1, -1)$.

Empirical Findings

1. Both rates are $I(1)$ (unit root tests)
2. One cointegrating relationship (Johansen test)
3. Cointegrating vector $\approx (1, -1)$: spread is stationary
4. Short rate adjusts to disequilibrium (long rate is weakly exogenous)

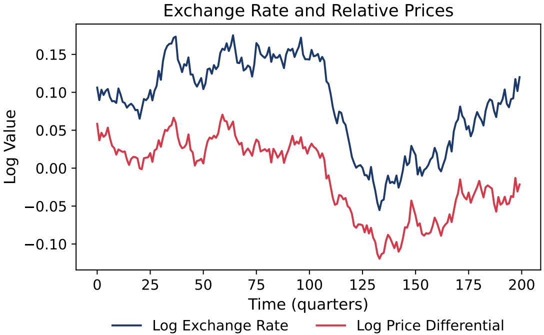
Example 2: Pairs Trading in Finance

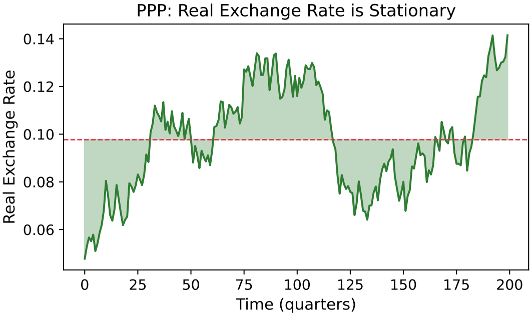




Strategy: Find cointegrated stock pairs (e.g., Coca-Cola & Pepsi). When spread deviates from mean, trade expecting mean reversion. Sell spread when high, buy when low.

Example 3: Purchasing Power Parity (PPP)





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

Interest Rate VECM Results

Typical Findings

- Both rates are $I(1)$
- One cointegrating relationship found
- Cointegrating vector close to $(1, -1)$: spread is stationary
- Short rate adjusts to long rate (not vice versa)

VECM Equations (stylized)

$$\Delta r_t = 0.02 - 0.15(r_{t-1} - R_{t-1}) + \text{lags} + \varepsilon_{1t}$$
$$\Delta R_t = 0.01 - 0.02(r_{t-1} - R_{t-1}) + \text{lags} + \varepsilon_{2t}$$

- Short rate adjusts faster ($\alpha_1 = -0.15$)
- Long rate nearly weakly exogenous ($\alpha_2 \approx 0$)

Key Takeaways

Main Concepts

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

Testing Methods

- **Engle-Granger**: Simple, but only one cointegrating vector
- **Johansen**: Multiple vectors, more powerful, MLE-based

Remember

Cointegration tests have low power in small samples. Economic theory should guide specification. Always check diagnostics!

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?

Quick Quiz

1. What does it mean for two $I(1)$ variables to be cointegrated?
2. What is the "spurious regression" problem?
3. In a VECM, what do the α coefficients represent?
4. What is the main advantage of Johansen over Engle-Granger?
5. If $\alpha_i = 0$ for variable Y_i, what does this imply?

Quiz Answers

1. **Cointegration**: A linear combination of the variables is $I(0)$ (stationary). They share a common stochastic trend.
2. **Spurious regression**: Regressing one $I(1)$ variable on another unrelated $I(1)$ variable gives high R^2 and significant coefficients even though there's no true relationship.
3. α **coefficients**: Speed of adjustment—how quickly each variable responds to deviations from long-run equilibrium.
4. **Johansen advantage**: Can test for multiple cointegrating relationships, uses MLE (more efficient), doesn't require choosing dependent variable.
5. $\alpha_i = 0$: Variable Y_i is weakly exogenous—it doesn't respond to disequilibrium. Other variables do all the adjusting.