Seminar 7: Heteroskedasticity, Collinearity & Unit Roots

SLR Assumptions, Dickey-Fuller Tests, and Error Correction Models

Giulio Rossetti*

giuliorossetti94.github.io

March 6, 2026

* email: giulio.rossetti.1@wbs.ac.uk

Disclaimer

Full solutions are available on my.wbs. All exercises are examinable material, not just the ones we covered in the seminars.

Roadmap

Exercise 1: Heteroskedasticity in SLR

Exercise 2: Perfect Collinearity in MLR

Exercise 3: Dickey-Fuller Unit Root Test

Exercise 4: Error Correction Model

Exercise 5: ADF Test for U.S. Inflation

Exercise 6: ADF Test for Log U.S. Real GDP

Roadmap

Exercise 1: Heteroskedasticity in SLR

Exercise 2: Perfect Collinearity in MLR

Exercise 3: Dickey-Fuller Unit Root Test

Exercise 4: Error Correction Model

Exercise 5: ADF Test for U.S. Inflation

Exercise 6: ADF Test for Log U.S. Real GDP

Why Does This Matter?

Motivation

OLS relies on key assumptions about the error term. Understanding when these assumptions hold — and when they don't — is essential for valid inference.

In this exercise we explore a savings function where:

- The zero conditional mean assumption (SLR.4) is satisfied
- But the homoskedasticity assumption (SLR.5) is violated

Key idea: The error variance depends on income — richer people have more variable savings.

Exercise 1: The Savings Function

$var(aX|Y) = a^2 var(X|Y)$

Consider the savings function:

$$sav = \beta_0 + \beta_1 inc + u, \qquad u = \sqrt{inc} \cdot e$$

where e is a random variable with $E[e] = 0$ and $\text{Var}[e] = \sigma_e^2$. Assume e is independent of inc.

Part (a): Show that $E[u \mid inc] = 0$ (SLR.4 is satisfied).

$e \perp inc$

Part (b): Show that $\text{Var}(u \mid inc) = \sigma_e^2 \cdot inc$ (SLR.5 is violated).

$E[u|inc] = E[\sqrt{inc} \cdot e \mid inc] = \sqrt{inc}\, E[e|inc] = \sqrt{inc}\, E[e] = 0$

$var(u|inc) = var(\sqrt{inc} \cdot e \mid inc) = inc \cdot var(e|inc) = inc \cdot var(e)$

$= inc \cdot \sigma_e^2$

Solution: Zero Conditional Mean (SLR.4)

When we condition on inc, the term $\sqrt{inc}$ becomes a constant:

$$E[u \mid inc] = E[\sqrt{inc} \cdot e \mid inc] = \sqrt{inc} \cdot E[e \mid inc]$$

Since e is independent of inc, we have $E[e \mid inc] = E[e] = 0$, so:

$$\boxed{E[u \mid inc] = 0}$$

Intuition: Independence of e from inc ensures that the error has no systematic relationship with income. The $\sqrt{inc}$ scaling doesn't change this because it only affects the *spread*, not the *average*.

Solution: Heteroskedasticity (SLR.5 Violated)

Again, conditioning on inc makes $\sqrt{inc}$ a constant in the variance:

$$\text{Var}[u \mid inc] = \text{Var}[\sqrt{inc} \cdot e \mid inc] = inc \cdot \text{Var}[e \mid inc] = inc \cdot \sigma_e^2$$

$$\boxed{\text{Var}(u \mid inc) = \sigma_e^2 \cdot inc}$$

Intuition: The variance of savings increases with income. This makes sense: higher-income individuals have more discretion in how much they save, leading to more variable savings behaviour.

Consequence: OLS is still unbiased (SLR.4 holds), but the usual standard errors are incorrect. Use heteroskedasticity-robust standard errors.

Roadmap

Exercise 1: Heteroskedasticity in SLR

Exercise 2: Perfect Collinearity in MLR

Exercise 3: Dickey-Fuller Unit Root Test

Exercise 4: Error Correction Model

Exercise 5: ADF Test for U.S. Inflation

Exercise 6: ADF Test for Log U.S. Real GDP

Exercise 2: The GPA Model

Students report weekly hours in four activities: *study*, *sleep*, *work*, *leisure*. Consider:

$$\rightarrow GPA = \beta_0 + \beta_1 study + \beta_2 sleep + \beta_3 work + \beta_4 leisure + u$$

$$study + sleep + work + leisure = 168$$

Question: Does it make sense to hold *sleep*, *work*, and *leisure* fixed while changing *study*?

$$leisure = 168 - study - sleep - work$$

$$\hat{\beta} = \underbrace{(X'X)}^{-1} X'y$$

Solution: The Constraint Problem

No. By definition:

$$study + sleep + work + leisure = 168$$

If we change $study$, at least one other variable must change to maintain the sum.

Which assumption is violated?

- We can write: $study = 168 - sleep - work - leisure$
- This is a perfect linear function of the other regressors
- This holds for *every* observation $\Rightarrow$ violates MLR.3 (no perfect collinearity)

Solution: How to Fix It

Simply drop one variable, say $leisure$:

$$\rightarrow GPA = \beta_0 + \beta_1 study + \beta_2 sleep + \beta_3 work + u$$

β_1: ↑ study ↓ leisure

$$\rightarrow GPA = \beta_0 + \beta_1 study + \beta_2 sleep + \beta_3 leisure + u$$

β_1: ↑ study ↓ work

Interpretation of β_1:

- The change in GPA when $study$ increases by one hour, holding $sleep$, $work$, and u fixed
- If $sleep$ and $work$ are fixed but $study$ goes up by 1 hour, then $leisure$ must go down by 1 hour
- So β_1 measures the effect of trading one hour of leisure for one hour of study

Roadmap

Exercise 1: Heteroskedasticity in SLR

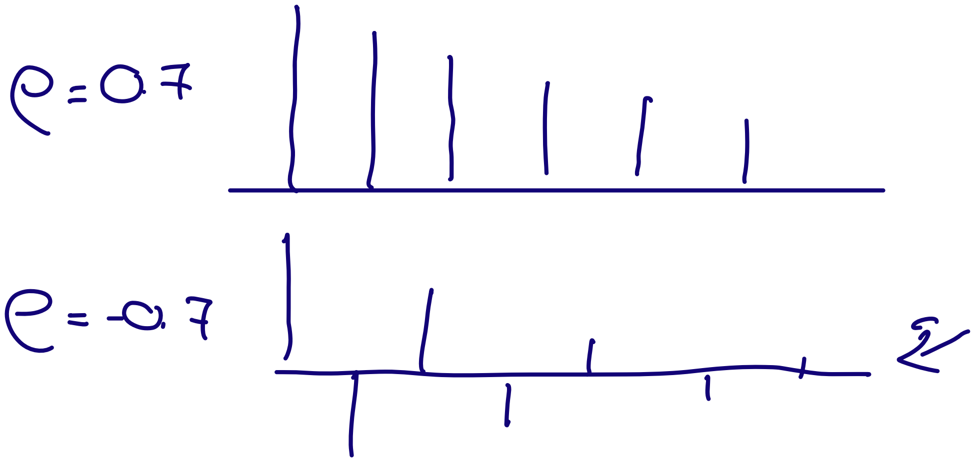
Exercise 2: Perfect Collinearity in MLR

Exercise 3: Dickey-Fuller Unit Root Test

Exercise 4: Error Correction Model

Exercise 5: ADF Test for U.S. Inflation

Exercise 6: ADF Test for Log U.S. Real GDP



RW: $X_t = \mu + X_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2_\varepsilon)$

I(1) $\quad X_t - X_{t-1} = \mu + \varepsilon_t \Rightarrow \Delta X_t = \quad + \varepsilon_t$

Exercise 3: Testing for a Unit Root in T-Bill Rates

We want to test whether three-month T-bill rates $r3_t$ contain a unit root.

Step 1: Start with an AR(1) model:

$$r3_t = \mu + \rho\, r3_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma^2)$$

Step 2: Subtract $r3_{t-1}$ from both sides:

$$\Delta r3_t = \mu + (\rho - 1)\, r3_{t-1} + \epsilon_t = \mu + \theta\, r3_{t-1} + \epsilon_t$$

where $\theta = \rho - 1$.

The Dickey-Fuller Test

Hypotheses:

- $H_0 : \theta = 0$ (unit root $\Leftrightarrow \rho = 1$)
- $H_1 : \theta < 0$ (stationary $\Leftrightarrow |\rho| < 1$)

Intuition: If $\theta = 0$, the series has no tendency to revert to any level — it "wanders" randomly. If $\theta < 0$, the series is pulled back toward its mean.

Can we use standard t-tests?

No! Under H_0 (unit root), the usual t-statistic does not follow a normal distribution, even asymptotically. We must use the Dickey-Fuller critical values instead.

Takeaway: Dickey-Fuller Test

Key Points

- The DF test reparameterises the AR(1) so that $H_0 : \theta = 0$ tests for a unit root
- Under H_0, standard t-tables are invalid — use DF critical values
- The test is one-sided (left-tail): reject H_0 if the test statistic is sufficiently negative
- Critical values depend on whether you include a constant, trend, or neither

Roadmap

Exercise 1: Heteroskedasticity in SLR

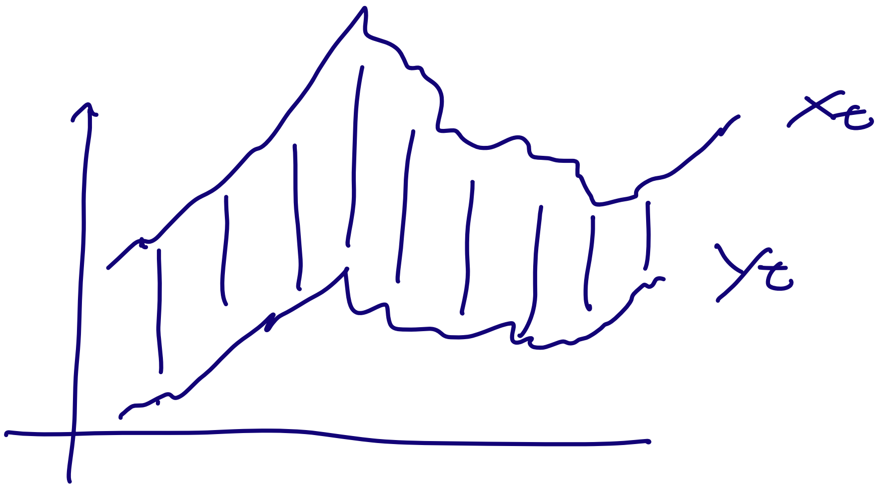
Exercise 2: Perfect Collinearity in MLR

Exercise 3: Dickey-Fuller Unit Root Test

Exercise 4: Error Correction Model

Exercise 5: ADF Test for U.S. Inflation

Exercise 6: ADF Test for Log U.S. Real GDP



Motivation

$y_t \qquad x_t$

$y_t = \beta x_t + z_t$

residuals $z_t = y_t - \beta x_t$

Definition (Cointegration).

Two variables x_t and y_t are cointegrated if

$$x_t, y_t \sim I(1) \quad \text{and} \quad \exists\, \beta \neq 0 \text{ such that } z_t = y_t - \beta x_t \sim I(0)$$

Implication.

The linear combination

$$y_t = \beta x_t + u_t, \qquad u_t \sim I(0)$$

is stationary. Deviations from the relation are therefore mean-reverting.

Interpretation.

the spread $y_t - \beta x_t$ cannot drift arbitrarily far and tends to return toward its mean. If x_t and y_t are cointegrated, they admit an Error Correction Model

Exercise 4: Setup

Suppose $\{(x_t, y_t)\}$ satisfies:

$$y_t = \beta x_t + u_t \qquad \text{and} \qquad \Delta x_t = \gamma \Delta x_{t-1} + \nu_t$$

where $E[u_t \mid I_{t-1}] = E[\nu_t \mid I_{t-1}] = 0$, $\beta \neq 0$, $|\gamma| < 1$.

Goal: Show these imply an error correction model:

ECM

$$\Delta y_t = \underbrace{\gamma_1 \Delta x_{t-1}}_{\text{SR}} + \underbrace{\delta(y_{t-1} - \beta x_{t-1})}_{\text{LR}} + e_t$$

where $\gamma_1 = \beta\gamma$, $\delta = -1$, $e_t = \beta\nu_t + u_t$.

Solution: Step 1 — Express Δy_t

Start from $y_t = \beta x_t + u_t$ and subtract y_{t-1}:

$$\Delta y_t = \beta x_t - y_{t-1} + u_t \quad + \beta x_{t-1} - \beta x_{t-1}$$

Step 2: Add and subtract βx_{t-1}:

$$\Delta y_t = \beta \Delta x_t - (y_{t-1} - \beta x_{t-1}) + u_t$$

This isolates the error correction term $(y_{t-1} - \beta x_{t-1})$, which measures last period's deviation from the long-run equilibrium.

Solution: Step 2 — Substitute Δx_t

Plug in $\Delta x_t = \gamma \Delta x_{t-1} + \nu_t$:

$$\begin{aligned}\Delta y_t &= \beta(\gamma \Delta x_{t-1} + \nu_t) - (y_{t-1} - \beta x_{t-1}) + u_t \\ &= \underbrace{\beta\gamma}_{\gamma_1} \Delta x_{t-1} \underbrace{-}_{\delta} (y_{t-1} - \beta x_{t-1}) + \underbrace{(\beta \nu_t + u_t)}_{e_t}\end{aligned}$$

Identifying the components:

ECM

$$\boxed{\Delta y_t = \underbrace{\gamma_1}_{\beta\gamma} \Delta x_{t-1} + \underbrace{\delta}_{-1} (y_{t-1} - \beta x_{t-1}) + \underbrace{e_t}_{\beta\nu_t + u_t}}$$

Intuition: y_t adjusts due to (i) lagged changes in x (short-run dynamics) and (ii) last period's equilibrium error (error correction, with $\delta = -1$ meaning full correction each period).

Interpreting the Error Correction Model

Short-run dynamics: $\gamma_1 \Delta x_{t-1}$

captures how recent changes in x_t affect current changes in y_t.

Example: if $\gamma_1 = 0.4$, a one-unit increase in Δx_{t-1} raises Δy_t by 0.4 units.

Error correction term: $\delta(y_{t-1} - \beta x_{t-1})$

The term $y_{t-1} - \beta x_{t-1}$ measures the deviation from the long-run relation.

The coefficient δ determines how quickly the system moves back toward equilibrium.

- If $\delta < 0$, deviations are corrected over time.
- The magnitude of δ measures the speed of adjustment.

Example: if $\delta = -0.3$ and the spread $y_{t-1} - \beta x_{t-1} = 10$, $\Delta y_t = -3$ so about 30% of the disequilibrium is corrected in one period.

Roadmap

Exercise 1: Heteroskedasticity in SLR

Exercise 2: Perfect Collinearity in MLR

Exercise 3: Dickey-Fuller Unit Root Test

Exercise 4: Error Correction Model

Exercise 5: ADF Test for U.S. Inflation

Exercise 6: ADF Test for Log U.S. Real GDP

Why Does This Matter?

Motivation

Is U.S. inflation stationary or does it contain a unit root? The answer matters for modelling and forecasting. The Augmented Dickey-Fuller (ADF) test handles serial correlation in the errors by including lagged differences.

DF $\Delta x_t = \mu + \boxed{\theta} x_{t-1} + \varepsilon_t$

ADF $\Delta x_t = \mu + \boxed{\theta} x_{t-1} + \gamma_1 \Delta x_{t-1} + \varepsilon_t$

Exercise 5: ADF Specification for Inflation

Using annual U.S. inflation data (1948–1996), estimate:

$$\Delta y_t = \alpha + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + e_t$$

where $|\gamma_1| < 1$.

Why "augmented"? The extra Δy_{t-1} term absorbs serial correlation in the errors, ensuring the test statistic has the correct distribution.

Hypotheses: Same as DF — $H_0 : \theta = 0$ (unit root) vs $H_1 : \theta < 0$ (stationary).

Exercise 5: Results and Interpretation

Conceptual approach:

1. Construct Δy_t, y_{t-1}, Δy_{t-1} from the inflation data
2. Run OLS on the ADF regression
3. Compute the t-statistic for $\hat{\theta}$
4. Compare with DF critical values (with constant, no trend)

Result: The DF 5% critical value is -2.86.

Conclusion: We reject the unit root hypothesis at the 5% level — inflation appears to be stationary.

Caveat: The sample is small ($T \approx 47$), which reduces the power of the test. Small samples make it harder to distinguish a unit root from a near-unit-root process.

Roadmap

Exercise 1: Heteroskedasticity in SLR

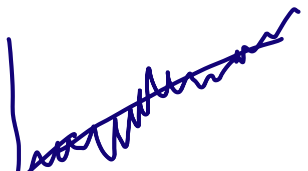
Exercise 2: Perfect Collinearity in MLR

Exercise 3: Dickey-Fuller Unit Root Test

Exercise 4: Error Correction Model

Exercise 5: ADF Test for U.S. Inflation

Exercise 6: ADF Test for Log U.S. Real GDP



Exercise 6: ADF with Trend for Log GDP

For log real GDP, include a time trend because GDP has a clear upward trajectory:

$$\Delta y_t = \alpha + \delta t + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + e_t$$

Why include a trend? Log GDP trends upward over time. Without the trend term, the test would confuse the deterministic trend with a stochastic trend (unit root).

Exercise 6: Results and Interpretation

Conceptual approach:

1. Take logs of real GDP, construct differences and lags
2. Run OLS on the ADF regression with trend
3. Compute t-statistic for $\hat{\theta}$
4. Compare with DF critical values (with constant *and* trend)

Result: The DF 5% critical value (with trend) is -3.41; the 10% value is -3.12.

Conclusion: We fail to reject the unit root hypothesis at both the 5% and 10% levels — log real GDP is consistent with having a unit root.

Caveat: Again, sample size is small ($T \approx 37$). The test has low power to distinguish between a unit root and a highly persistent but stationary process.

Summary

Key Takeaways

1. **Heteroskedasticity:** The error variance can depend on regressors even when $E[u \mid x] = 0$. Use robust standard errors.
2. **Perfect collinearity:** Exact linear relationships among regressors make OLS impossible. Drop one variable.
3. **Dickey-Fuller test:** Reparameterise AR(1) to test $\theta = 0$. Standard t-tables are invalid — use DF critical values.
4. **ECM:** Cointegrated variables adjust through both short-run dynamics and long-run error correction.
5. **ADF test:** Include lagged differences to handle serial correlation. Include a trend when the data has a deterministic trend.

Key Formulas

Concept	Formula
Heteroskedasticity	$\text{Var}(u \mid x) = \sigma_e^2 \cdot x$
DF test regression	$\Delta y_t = \mu + \theta y_{t-1} + \epsilon_t$
ADF (with lag)	$\Delta y_t = \alpha + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + e_t$
ADF (with trend)	$\Delta y_t = \alpha + \delta t + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + e_t$
ECM	$\Delta y_t = \gamma_1 \Delta x_{t-1} + \delta(y_{t-1} - \beta x_{t-1}) + e_t$

$|\rho| = 1$ rw

$|\rho| \neq 1$ $|\rho| > 1$ explo.

$\theta = (\rho - 1)$ $\rho = 1 \Rightarrow \theta = 0$

$\rho < 1$