# Lecture 6-3: Violations of Homoskedasticity

**Goal:** Diagnose and correct violations of homoskedasticity in regression, focusing on heteroskedasticity, clustering, and serial correlation.

- Violations of homoskedasticity
- Heteroskedasticity
  - Definition
  - Consequences for inference
  - Visual diagnostics and residual plots
  - Formal tests (Breusch-Pagan, White)
  - Remedies: transformations, WLS, robust errors
  - White's heteroskedasticity-robust SE
- Clustering
  - Concept of cluster dependence
  - Error variance matrix structure
  - Estimation of cluster-robust SE
- Serial correlation
  - Time dependence in errors
  - AR(1) error structure
  - Variance and covariance of AR errors
  - Detection: plots, Durbin-Watson
  - Corrections: Newey-West HAC, AR adjustments
- Summary
  - Three key violations: heteroskedasticity, clustering, serial correlation
  - Detection through plots and tests
  - Use of robust or cluster-robust standard errors

## Worksheet

1. The classic variance estimator is

$$\widehat{V}_0 = \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1}.$$

The White estimator is

$$\widehat{V}_{HC0} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\text{diag}(\hat{u}_1^2, \ldots, \hat{u}_n^2)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}.$$

Assume (1) $E[u_i \mid \mathbf{X}] = 0$; (2) $\text{Var}(u_i \mid \mathbf{X}) = \sigma^2$; (3) $\frac{1}{n}\mathbf{X}'\mathbf{X} \to Q \succ 0$; and (4) standard regularity conditions. Show that

$$\text{plim}(n\,\widehat{V}_0) = \text{plim}(n\,\widehat{V}_{HC0}) = \sigma^2 Q^{-1}.$$

**Hint:**

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' \hat{u}_i^2 = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' u_i^2 + o_p(1) \xrightarrow{p} \sigma^2 E[\mathbf{x}_i \mathbf{x}_i'] = \sigma^2 Q.$$

2. Consider the linear regression model

$$y_{ij} \;=\; \beta_0 + \beta_1 x_{ij} + \varepsilon_{ij} \;=\; \beta_0 + \beta_1 x_{ij} + v_j + u_{ij},$$

where units $i$ are nested in clusters $j$ and

$$v_j \overset{\text{i.i.d.}}{\sim} N(0, \rho\sigma^2), \qquad u_{ij} \overset{\text{i.i.d.}}{\sim} N(0, (1-\rho)\sigma^2),$$

with $v_j$ and $u_{ij}$ mutually independent and independent of $x_{ij}$. The parameter $\rho \in [0, 1]$ is called the within-cluster correlation.

Suppose we have $M = 2$ clusters and $N = 6$ units: cluster 1 contains units $i = 1, 2, 3$ and cluster 2 contains units $i = 4, 5, 6$. Stack the errors in the order

$$\boldsymbol{\varepsilon} = (\varepsilon_{1,1}, \varepsilon_{2,1}, \varepsilon_{3,1}, \varepsilon_{4,2}, \varepsilon_{5,2}, \varepsilon_{6,2})'.$$

(a) Show that $V[\varepsilon_{ij}] = \sigma^2$ for all $i, j$.

(b) For two units $i \neq \ell$ in the *same* cluster $j$, compute $\text{Cov}(\varepsilon_{ij}, \varepsilon_{\ell j})$.

(c) For units in *different* clusters $j \neq k$, compute $\text{Cov}(\varepsilon_{ij}, \varepsilon_{\ell k})$.

(d) Using your answers above, write out the $6 \times 6$ covariance matrix $\Sigma = V[\boldsymbol{\varepsilon}]$ for the given ordering $(\varepsilon_{1,1}, \varepsilon_{2,1}, \varepsilon_{3,1}, \varepsilon_{4,2}, \varepsilon_{5,2}, \varepsilon_{6,2})'$.

(e) Compute the correlation $\text{Cor}(\varepsilon_{ij}, \varepsilon_{\ell j})$ for $i \neq \ell$ in the same cluster $j$ and verify that it equals $\rho$.

3. The general cluster-robust variance estimator can be written as

$$\widehat{V}_{CR}[\hat{\boldsymbol{\beta}}] = (\mathbf{X}'\mathbf{X})^{-1} \left( \sum_{j=1}^{M} \mathbf{X}_j' \hat{\mathbf{u}}_j \hat{\mathbf{u}}_j' \mathbf{X}_j \right) (\mathbf{X}'\mathbf{X})^{-1},$$

where $\hat{\mathbf{u}}_j$ is the vector of residuals in cluster $j$. **Explain** how this generalizes the White estimator.