ECON 640 Serial Correlation/Autocorrelation

Dr. Yang Liang

1 Serial Correlation

1.1 Definition

Serial correlation (also known as autocorrelation) occurs when the error term in a regression model is correlated across time (or space) within a cluster. This violates the i.i.d. assumption of OLS.

The general model is:

$$y_{it} = X_{it}\beta + u_{it}$$

where:

- i: individual dimension (e.g., firm, state).
- t: time dimension.
- s: state or spatial dimension.

1.2 Structure of Serial Correlation

An example is the AR(1) process:

$$u_{it} = \rho u_{i,t-1} + \epsilon_{it}, \quad |\rho| < 1, \quad \epsilon_{it} \sim \text{i.i.d.}$$

In this case:

- Across time:
$$u_{it} = \rho u_{i,t-1} + \epsilon_{it}$$
- Across space:
$$u_{it} = \rho u_{i,s-1} + \epsilon_{it}$$

Serial correlation is very common in panel data.

1.3 Consequences of Serial Correlation

When serial correlation is present, it leads to:

- Incorrect standard errors ($Var(\hat{\beta})$).
- Misleading inference and reduced precision.
- Inefficient OLS estimates.

1.4 Estimation Adjustments for Serial Correlation

1.4.1 Newey-West (HAC) Standard Errors

The Newey-West estimator (1987) adjusts standard errors for heteroskedasticity and autocorrelation. The variance-covariance matrix is:

$$Var(\hat{\beta})_{NW} = (X'X)^{-1} \left(\sum_{t=0}^{q} w_t \sum_{s=1}^{n} (X_t u_t)(X_{t+s} u_{t+s})' \right) (X'X)^{-1}$$

where:

- q: lag length.
- $w_t = 1 - \frac{t}{q+1}$: weight function that reduces the influence of longer lags.

1.4.2 Cluster-Robust Standard Errors

Cluster-robust standard errors allow for arbitrary correlation within clusters. For example, individuals, firms, or states across time.

The cluster-robust variance formula is:

$$Var(\hat{\beta})_{Cluster} = (X'X)^{-1} \left(\sum_{s=1}^{S} X_s' \hat{u}_s \hat{u}_s' X_s \right) (X'X)^{-1}$$

where:

- S: number of clusters.
- $\hat{u}_s$: residuals for cluster s (vector).

2 Comparison: White vs. Cluster-Robust Standard Errors

White standard errors correct for heteroskedasticity but assume that errors are independent across observations. This assumption breaks down when there is within-cluster correlation.

Cluster-robust standard errors, on the other hand, allow for arbitrary correlation of errors within clusters, making them a stronger and more general estimator in clustered data settings (e.g., panel data).

2.1 White Standard Errors

The variance-covariance matrix is:

$$Var(\hat{\beta})_{White} = (X'X)^{-1} \left(\sum_{i=1}^{n} \hat{u}_i^2 x_i x_i' \right) (X'X)^{-1}$$

- Corrects for heteroskedasticity.
- Assumes no correlation between errors across observations.
- Suitable for data with heteroskedasticity but no clustering.

2.2 Cluster-Robust Standard Errors

The variance-covariance matrix is:

$$Var(\hat{\beta})_{Cluster} = (X'X)^{-1} \left(\sum_{s=1}^{S} X_s' \hat{u}_s \hat{u}_s' X_s \right) (X'X)^{-1}$$

- Corrects for both heteroskedasticity and within-cluster correlation.
- Accounts for dependencies in errors within clusters.
- More general and robust in settings with clustered data.

2.3 Key Difference

- White standard errors only adjust for heteroskedasticity and assume error independence.
- Cluster-robust standard errors adjust for heteroskedasticity and also allow for arbitrary correlation of errors within clusters.

2.4 Important Example: White vs. Cluster-Robust Standard Errors

Prepare the Environment:

Consider a dataset with $n = 2$ clusters and $t = 2$ time periods. The data structure is:

$$y = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \end{bmatrix}, \quad X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix}, \quad u = \begin{bmatrix} u_{11} \\ u_{12} \\ u_{21} \\ u_{22} \end{bmatrix}$$

Using OLS, the residuals are:

$$\hat{u} = y - X\hat{\beta}, \quad \hat{u} = \begin{bmatrix} \hat{u}_{11} \\ \hat{u}_{12} \\ \hat{u}_{21} \\ \hat{u}_{22} \end{bmatrix}$$

For White-robust standard errors, only the diagonal element is accounted for:

$$Var(\hat{\beta})_{White} = (X'X)^{-1} \left(\sum_{i=1}^{n} \hat{u}_i^2 x_i x_i' \right) (X'X)^{-1}$$

Expanding the summation for a small example with two clusters and two observations per cluster:

$$Var(\hat{\beta})_{White} = (X'X)^{-1} \left(\hat{u}_{11}^2 x_{11} x_{11}' + \hat{u}_{12}^2 x_{12} x_{12}' + \hat{u}_{21}^2 x_{21} x_{21}' + \hat{u}_{22}^2 x_{22} x_{22}' \right) (X'X)^{-1}$$

Where:

- x_{ij}: the covariates for the i-th cluster and j-th observation.
- $\hat{u}_{ij}^2$: the squared residual for the i-th cluster and j-th observation.

For cluster-robust standard errors, all within-cluster correlation is accounted for:

$$Var(\hat{\beta})_{Cluster} \rightarrow \text{Allows within-cluster correlation.}$$

Cluster Residuals

$$\hat{u}_1 = \begin{bmatrix} \hat{u}_{11} \\ \hat{u}_{12} \end{bmatrix}, \quad \hat{u}_2 = \begin{bmatrix} \hat{u}_{21} \\ \hat{u}_{22} \end{bmatrix}$$

Cluster Covariates

$$X_1 = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, \quad X_2 = \begin{bmatrix} x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix}$$

$$Var(\hat{\beta})_{Cluster} = (X'X)^{-1} \left(X_1'\hat{u}_1\hat{u}_1'X_1 + X_2'\hat{u}_2\hat{u}_2'X_2 \right) (X'X)^{-1}$$

Where:

$$\hat{u}_1 = \begin{bmatrix} \hat{u}_{11} \\ \hat{u}_{12} \end{bmatrix}, \quad X_1 = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$

$$\hat{u}_2 = \begin{bmatrix} \hat{u}_{21} \\ \hat{u}_{22} \end{bmatrix}, \quad X_2 = \begin{bmatrix} x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix}$$

Concluding remarks

- Serial correlation violates OLS assumptions and requires correction for valid inference.
- Newey-West and cluster-robust standard errors are widely used adjustments.
- Clustered data structures (e.g., panel data) necessitate careful attention to within-cluster correlation.