OLS Assumptions: Simple vs. Multiple Regression

EC 320 – Econometrics Summary Sheet

Assumption	Simple Linear Regression	Multiple Linear Regression
1. Linearity in Parameters	$Y = \beta_0 + \beta_1 X + u$	$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + u$
2. Variation in Regressors	X must vary in the sample	Each X_j must vary (no perfect collinearity)
3. Zero Conditional Mean	$\mathbb{E}[u_i] = 0$	$\mathbb{E}[u_i \mid X_1, \ldots, X_k] = 0$
4. Homoskedasticity	$\text{Var}(u_i) = \sigma^2$	$\text{Var}(u_i \mid X_1, \ldots, X_k) = \sigma^2$
5. No Autocorrelation	$\text{Cov}(u_i, u_j) = 0$ for $i \neq j$	Same: u_i and u_j are independent across $i \neq j$
6. Normality of Errors	$u_i \sim \mathcal{N}(0, \sigma^2)$	$u_i \sim \mathcal{N}(0, \sigma^2)$, given all X_j

Notes:

- Assumptions 1–5 are required for OLS to be unbiased and efficient (BLUE).
- Assumption 6 (Normality) is not required for unbiasedness, but enables hypothesis testing.

Unbiasedness and Efficiency of OLS

Summary Sheet – EC 320

Unbiasedness of $\hat{\beta}_j$

Under the classical assumptions (especially A3: $\mathbb{E}[u_i] = 0$), we can express each OLS coefficient as:

$$\hat{\beta}_j = \beta_j + \sum_i a_{ij} u_i$$

where a_{ij} are functions of the data interacted with each u_i. You may want to refer to your Chapter 2 notes to see how we got to this point.[1]

Taking expectations:

$$\mathbb{E}[\hat{\beta}_j] = \beta_j + \sum_i a_{ij}\mathbb{E}[u_i] = \beta_j + \sum_i a_{ij} \cdot 0 = \beta_j$$

Conclusion: OLS is an **unbiased estimator** of β_j.

Efficiency of OLS

Under the Gauss-Markov Theorem, OLS is the **Best Linear Unbiased Estimator (BLUE)** – i.e., it has the lowest variance among all linear unbiased estimators.

The table below summarizes how we estimate residual variance and standard error in both simple and multiple regression. We won't derive these in class, but you are welcome to explore them on your own (especially via matrix algebra).

Concept	**Simple Regression**	**Multiple Regression**
Estimator of Residual Variance	$\hat{\sigma}^2 = \frac{1}{n-2}\sum_i \hat{u}_i^2$	$\hat{\sigma}^2 = \frac{1}{n-k-1}\sum_i \hat{u}_i^2$
Standard Error of $\hat{\beta}_j$	$\text{SE}(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{\sum_i (X_i - \bar{X})^2}}$	$\text{SE}(\hat{\beta}_j) = \sqrt{\frac{\hat{\sigma}^2}{\sum_i (x_{ji} - \bar{x}_j)^2 \cdot (1 - R_j^2)}}$

Prepared for EC 320 — Fall 2025 *Instructor: David M. Hall*

[1] **Hint:** Substitute Y_i and $\bar{Y}$ into your basic OLS formula.