ECONOMETRICS BASICS

BLUE in Econometrics

One famous acronym. One very useful idea.

What does BLUE mean?

BLUE stands for Best Linear Unbiased Estimator. Under the right assumptions, the Ordinary Least Squares (OLS) estimator has the smallest variance among all linear unbiased estimators.

Break the word into four parts

B — BEST
Smallest variance among linear unbiased estimators.

L — LINEAR
The estimate is a linear combination of observed Y values.

U — UNBIASED
Across repeated samples, its expected value equals the true parameter.

E — ESTIMATOR
A rule or formula used to estimate an unknown value, such as a regression coefficient.

A quick memory trick

Think of several students estimating the same true value. Their estimates may vary from sample to sample. BLUE is the one with the tightest spread - but only within the class of linear, unbiased estimators.

THE GAUSS-MARKOV IDEA

When is OLS BLUE?

The Gauss-Markov theorem says that OLS is BLUE when the relevant regression assumptions hold. In simple words, the data and errors must behave in the right way.

1

1. Linear in parameters

The regression is written in a form such as Y = beta0 + beta1 X + u. The parameters enter linearly; explanatory variables can be transformed.

2

2. Appropriate sampling / data structure

The observations follow the sampling assumptions used by the model.

3

3. No perfect multicollinearity

No explanatory variable is an exact linear combination of the others.

4

4. Zero conditional mean

E(u | X) = 0. After accounting for X, the error has no systematic relationship with the explanatory variables.

5

5. Constant variance and no error correlation

The errors are homoskedastic and uncorrelated across observations (in the standard version of the theorem).

Important: normality is NOT required

The errors do not have to be normally distributed for OLS to be BLUE. Normality is often introduced for exact small-sample t and F tests, but it is not a Gauss-Markov requirement.

Textbooks may group or phrase assumptions slightly differently depending on whether the data are cross-sectional or time-series.

EXAMPLE + EXAM NOTES

A simple example

Suppose you want to estimate how education is related to wages:

$$\text{Wage} = \beta_0 + \beta_1 \text{Education} + u$$

OLS uses the observed data to estimate beta0 and beta1. If the Gauss-Markov assumptions hold, the OLS estimate of beta1 is unbiased and has the smallest variance among linear unbiased estimators of beta1.

What BLUE does - and does not - mean

It DOES mean

Under the assumptions, OLS is the most precise (minimum-variance) estimator in the class of linear unbiased estimators.

It does NOT mean

OLS is perfect in every sample, has the highest R-squared, or beats every possible nonlinear or biased estimator.

Exam-ready answer

According to the Gauss-Markov theorem, under the classical regression assumptions, the Ordinary Least Squares (OLS) estimator is BLUE: the Best Linear Unbiased Estimator. “Best” means that it has the minimum variance among all linear unbiased estimators of the regression coefficients.

Quick check

True or false: OLS must have normally distributed errors to be BLUE.

Answer: False.

Normality is not required for the Gauss-Markov result.

Tiny joke to remember it: BLUE is not a colour-coded regression line - it is OLS showing its best behaviour under the right assumptions.