

Lecture 3: Simple Linear Regression and OLS

Economics 326 — Introduction to Econometrics II

Vadim Marmer, UBC

April 5, 2026

Introduction and Definitions

Overview of Simple Linear Regression

- The **simple linear regression model** is used to study the relationship between **two variables**.
- It has many limitations, but nevertheless there are examples in the literature where the simple linear regression is applied (e.g., stock returns predictability).
- It is also a good starting point for learning the regression technique.

Key Definitions

Terminology for Simple Regression

y	x
Dependent variable	Independent variable
Explained variable	Explanatory variable
Response variable	Control variable
Predicted variable	Predictor variable
Regressand	Regressor

Sample vs Population

- The econometrician observes random data:

observation	dependent variable	regressor
1	Y_1	X_1
2	Y_2	X_2
\vdots	\vdots	\vdots
n	Y_n	X_n

- A pair X_i, Y_i is called an observation.
- **Sample:** $\{(X_i, Y_i) : i = 1, \dots, n\}$.
- The population is the **joint distribution** of the sample.

Simple Linear Regression Model

Conditional Expectation Function

- We model the relationship between Y and X using the **conditional expectation**:

$$E[Y_i | X_i] = \alpha + \beta X_i.$$

- **Intercept:** $\alpha = E[Y_i | X_i = 0]$.
- **Slope:** β measures the effect of a unit change in X on Y :

$$\begin{aligned}\beta &= E[Y_i | X_i = x + 1] - E[Y_i | X_i = x] \\ &= [\alpha + \beta(x + 1)] - [\alpha + \beta x].\end{aligned}$$

- **Marginal effect** of X on Y :

$$\beta = \frac{dE[Y_i | X_i]}{dX_i}.$$

- **The effect is the same for all x .**

Regression Error Term

- α and β in $E[Y_i | X_i] = \alpha + \beta X_i$ are **unknown**.
- **Residual (error):**

$$U_i = Y_i - E[Y_i | X_i] = Y_i - (\alpha + \beta X_i).$$

U_i 's are **unobservable**.

- **The model:**

$$\begin{aligned}Y_i &= \alpha + \beta X_i + U_i, \\ E[U_i | X_i] &= 0.\end{aligned}$$

Functional Form: Linearity

- We consider a model that is linear in the coefficients α, β : $Y_i = \alpha + \beta X_i + U_i$.
- The dependent variable and the regressor can be nonlinear functions of some other variables.
- The most popular function is log.

Log-Linear Specification

- Consider the following model:

$$\log Y_i = \alpha + \beta X_i + U_i.$$

- In this case,

$$\begin{aligned}\beta &= \frac{d(\log Y_i)}{dX_i} \\ &= \frac{dY_i/Y_i}{dX_i} = \frac{dY_i/dX_i}{Y_i}.\end{aligned}$$

- β measures **percentage** change in Y as a response to a unit change in X .
- In this model, it is assumed that the percentage change in Y is the same for all values of X (constant).
- In $\log(\text{Wage}_i) = \alpha + \beta \times \text{Education}_i + U_i$, β measures the **return** to education.

Log-Log Specification

- Consider the following model:

$$\log Y_i = \alpha + \beta \log X_i + U_i.$$

- In this model,

$$\begin{aligned} \beta &= \frac{d \log Y_i}{d \log X_i} \\ &= \frac{dY_i/Y_i}{dX_i/X_i} = \frac{dY_i}{dX_i} \frac{X_i}{Y_i}. \end{aligned}$$

- β measures **elasticity**: the percentage change in Y as a response to 1% change in X .
- Here, the elasticity is assumed to be the same for all values of X .
- Example: Cobb-Douglas production function:

$$Y = \alpha K^{\beta_1} L^{\beta_2} \implies \log Y = \log \alpha + \beta_1 \log K + \beta_2 \log L$$

(two regressors: log of capital and log of labour).

Assumption of “Orthogonality”

The model:

$$Y_i = \alpha + \beta X_i + U_i.$$

We assume that $E[U_i|X_i] = 0$.

- $E[U_i] = 0$.

$$E[U_i] \stackrel{\text{Law of Iterated Expectation}}{=} E[E[U_i|X_i]] = E[0] = 0.$$

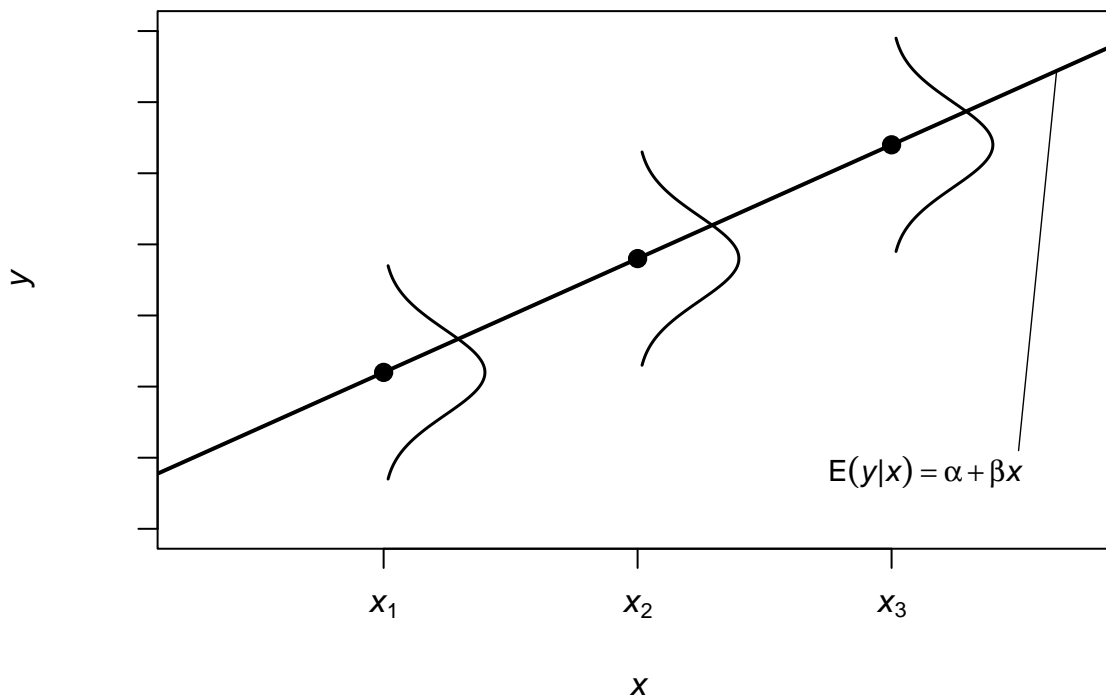
- $\text{Cov}(X_i, U_i) = E[X_i U_i] = 0$.

$$\begin{aligned} E[X_i U_i] &\stackrel{\text{Law of Iterated Expectation}}{=} E[E[X_i U_i|X_i]] \\ &= E[X_i E[U_i|X_i]] = E[X_i \cdot 0] = 0. \end{aligned}$$

Decomposition into Predicted and Error Components

$$Y_i = \underbrace{\alpha + \beta X_i}_{\text{Predicted by } X} + \underbrace{U_i}_{\text{“Orthogonal to” } X}$$

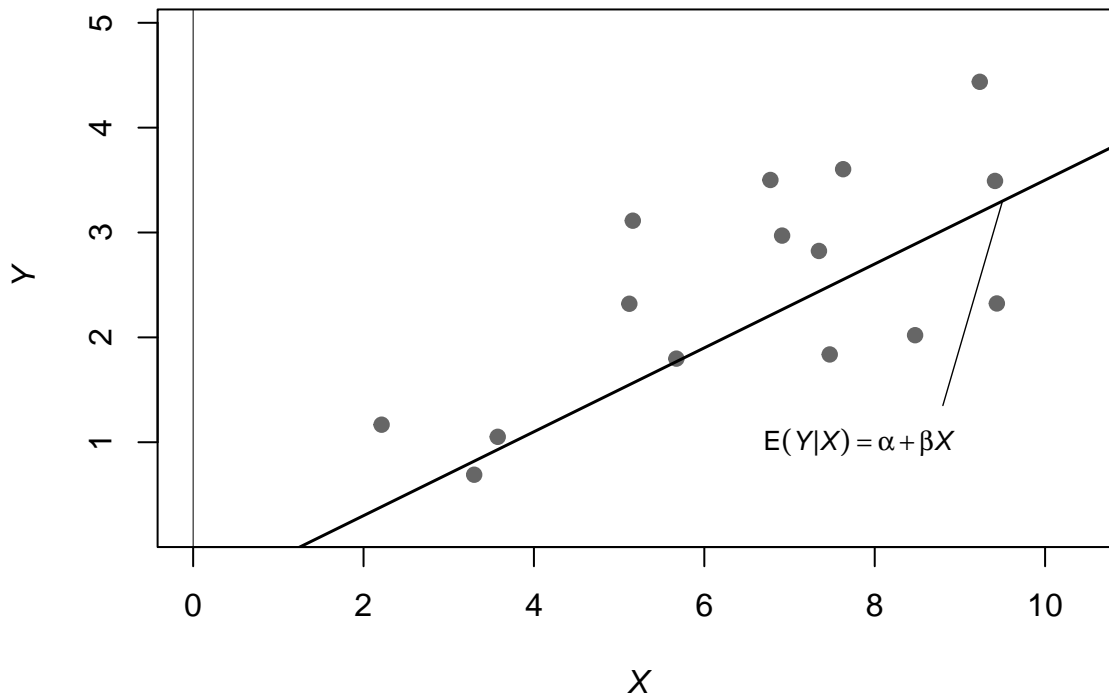
$E(y|x)$ as a linear function of x



Estimation of Parameters

Estimation Problem

Scatterplot of Y and X, and the population regression



Problem: estimate the unknown parameters α and β using the data (n observations) on Y and X .

Method of Moments Estimator

- Our assumptions imply:

$$E[U_i] = E[Y_i - \alpha - \beta X_i] = 0.$$

$$E[X_i U_i] = E[X_i (Y_i - \alpha - \beta X_i)] = 0.$$

- An **estimator** is a function of the observable data; it can depend only on observable X and Y . Let $\hat{\alpha}$ and $\hat{\beta}$ denote the estimators of α and β .
- **Method of moments:** replace expectations with averages. **Normal equations:**

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta} X_i) = 0.$$

$$\frac{1}{n} \sum_{i=1}^n X_i (Y_i - \hat{\alpha} - \hat{\beta} X_i) = 0.$$

Normal Equations

- Let $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ and $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ (averages).

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta} X_i) = 0$$

implies

$$\frac{1}{n} \sum_{i=1}^n Y_i - \frac{1}{n} \sum_{i=1}^n \hat{\alpha} - \hat{\beta} \frac{1}{n} \sum_{i=1}^n X_i = 0 \text{ or}$$
$$\bar{Y} - \hat{\alpha} - \hat{\beta} \bar{X} = 0.$$

The fitted regression line goes through the averages.

•

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}.$$

Slope Estimator Derivation

- $\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}$ and therefore

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n X_i (Y_i - \hat{\alpha} - \hat{\beta} X_i) \\ &= \sum_{i=1}^n X_i (Y_i - (\bar{Y} - \hat{\beta} \bar{X}) - \hat{\beta} X_i) \\ &= \sum_{i=1}^n X_i [(Y_i - \bar{Y}) - \hat{\beta} (X_i - \bar{X})] \\ &= \sum_{i=1}^n X_i (Y_i - \bar{Y}) - \hat{\beta} \sum_{i=1}^n X_i (X_i - \bar{X}). \end{aligned}$$

Final Formulas for OLS Estimators

•

$$0 = \sum_{i=1}^n X_i (Y_i - \bar{Y}) - \hat{\beta} \sum_{i=1}^n X_i (X_i - \bar{X}) \text{ or}$$
$$\hat{\beta} = \frac{\sum_{i=1}^n X_i (Y_i - \bar{Y})}{\sum_{i=1}^n X_i (X_i - \bar{X})}.$$

- Since

$$\begin{aligned} \sum_{i=1}^n X_i (Y_i - \bar{Y}) &= \sum_{i=1}^n (X_i - \bar{X}) (Y_i - \bar{Y}) = \sum_{i=1}^n (X_i - \bar{X}) Y_i \\ \sum_{i=1}^n X_i (X_i - \bar{X}) &= \sum_{i=1}^n (X_i - \bar{X}) (X_i - \bar{X}) = \sum_{i=1}^n (X_i - \bar{X})^2 \end{aligned}$$

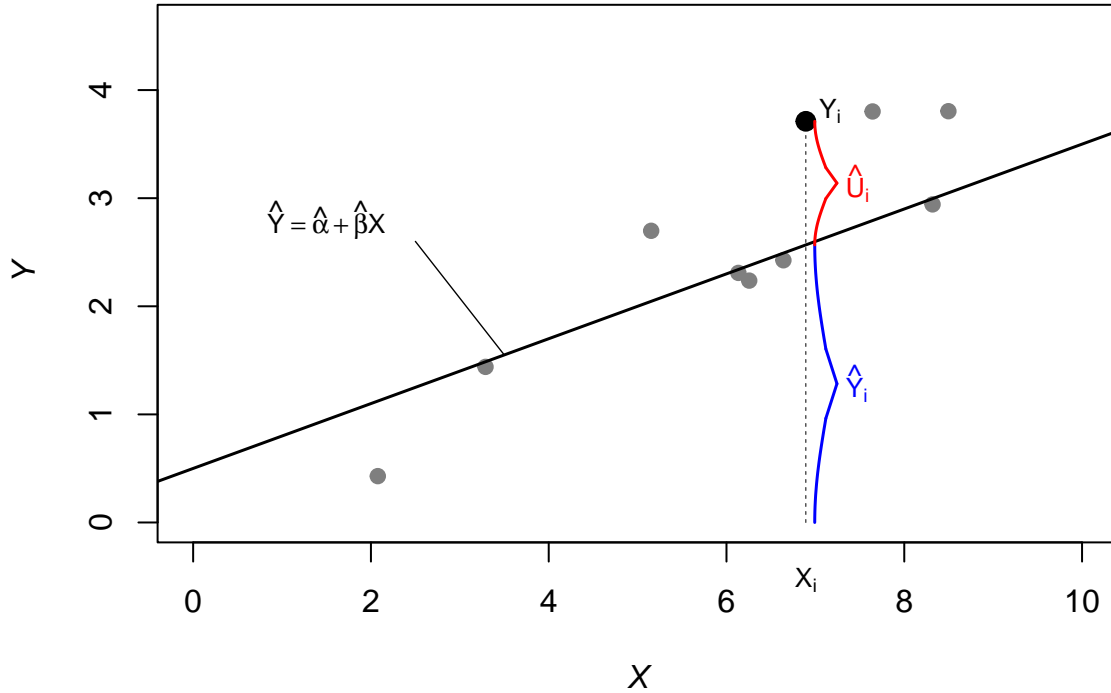
we can also write

$$\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X}) Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2}.$$

Fitted Values and Residuals

- **Fitted values:** $\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$.
- **Fitted residuals:** $\hat{U}_i = Y_i - \hat{\alpha} - \hat{\beta} X_i$.

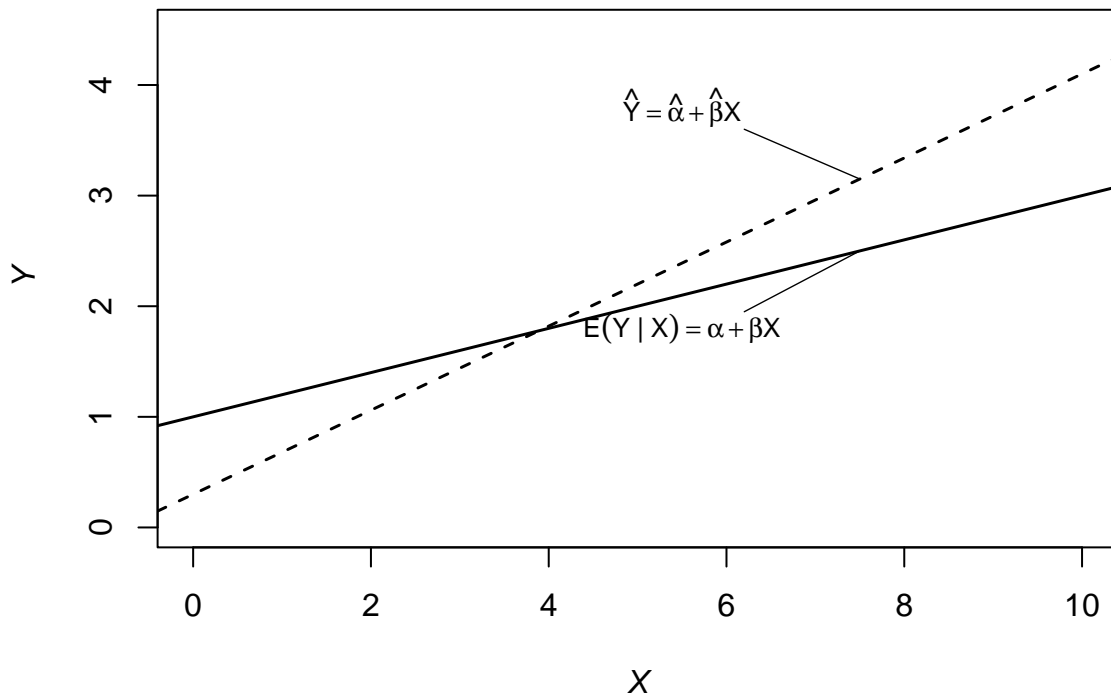
Fitted values and residuals



True vs Fitted Regression Lines

- **True:** $Y_i = \alpha + \beta X_i + U_i$, $E[U_i] = E[X_i U_i] = 0$.
- **Fitted:** $Y_i = \hat{\alpha} + \hat{\beta} X_i + \hat{U}_i$, $\sum_{i=1}^n \hat{U}_i = \sum_{i=1}^n X_i \hat{U}_i = 0$.

True vs fitted regression lines



Ordinary Least Squares (OLS) Minimization

- Minimize $Q(a, b) = \sum_{i=1}^n (Y_i - a - bX_i)^2$ with respect to a and b .

- Derivatives:

$$\frac{dQ(a, b)}{da} = -2 \sum_{i=1}^n (Y_i - a - bX_i).$$

$$\frac{dQ(a, b)}{db} = -2 \sum_{i=1}^n (Y_i - a - bX_i) X_i.$$

- First-order conditions:

$$0 = \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}X_i) = \sum_{i=1}^n \hat{U}_i.$$

$$0 = \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}X_i) X_i = \sum_{i=1}^n \hat{U}_i X_i.$$

- Method of moments = OLS:

$$\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X}) Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad \text{and} \quad \hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}.$$