Interpreting OLS Coefficients

Different functional forms imply different interpretations of coefficients. Understanding these distinctions is fundamental for empirical analysis.

Marco Rosso

1. Linear–Linear Model

$$y = \beta_0 + \beta_1 x + u$$

Interpretation:

- A one-unit increase in x changes y by β_1 units.

Example: If $\beta_1 = 2.5$, then increasing x by 1 increases y by 2.5 units.

2. Log–Linear Model (Semi-elasticity)

$$\ln(y) = \beta_0 + \beta_1 x + u$$

Interpretation:

- A one-unit increase in x changes y by approximately

$$100 \times \beta_1\%.$$

- This approximation is valid for $|\beta_1| < 0.1$ (roughly).
- For large β_1, use the exact change:

$$100(e^{\beta_1} - 1)\%.$$

Example: If $\beta_1 = 0.04$, then a 1-unit increase in x increases y by about 4%.

3. Linear–Log Model (Semi-elasticity)

$$y = \beta_0 + \beta_1 \ln(x) + u$$

Interpretation:

- A 1% increase in x changes y by:

$$0.01 \times \beta_1 \text{ units} \quad (\text{i.e., } \frac{\beta_1}{100} \text{ units}).$$

Example: If $\beta_1 = 8$, then a 1% increase in x increases y by $0.01 \times 8 = 0.08$ units.

4. Log–Log Model (Elasticity)

$$\ln(y) = \beta_0 + \beta_1 \ln(x) + u$$

Interpretation:

- β_1 is an **elasticity**.
- A 1% increase in x changes y by β_1%.

Example: If $\beta_1 = 0.7$, then increasing x by 1% increases y by 0.7%.

5. Dummy Variable in Linear Model

$$y = \beta_0 + \beta_1 D + u$$

where $D \in \{0, 1\}$.

Interpretation:

- β_1 is the difference in the mean of y between the two groups:

$$\beta_1 = E[y|D = 1] - E[y|D = 0].$$

Example: If $\beta_1 = 5$: when $D = 1$ (vs. $D = 0$), y is approximately 5 units higher on average.

6. Dummy Variable in Log–Linear Model

$$\ln(y) = \beta_0 + \beta_1 D + u$$

Interpretation:

- The percentage difference between $D = 1$ and $D = 0$ is:

$$100(e^{\beta_1} - 1)\%.$$

- For small β_1: approximately $100\beta_1\%$.

Example: If $\beta_1 = 0.2$: $100(e^{0.2} - 1) \approx 22.14\% \Rightarrow$
That is, when $D = 1$ (vs $D = 0$), y is approximately 22% higher.

7. Interaction of Continuous Variable and Dummy

$$y = \beta_0 + \beta_1 x + \beta_2 D + \beta_3 (x \cdot D) + u$$

Interpretation:

- Slope for group $D = 0$: β_1
- Slope for group $D = 1$: $\beta_1 + \beta_3$
- Difference in slopes: β_3

Used for: heterogeneous effects, gender differences, treatment interactions.

Summary Table: Interpretation of β

Model	Interpretation of β_1
y vs. x	$\Delta y = \beta_1 \Delta x$
$\ln y$ vs. x	1 unit $\uparrow$ in $x \rightarrow \approx 100\beta_1\%$ change in y
y vs. $\ln x$	$1\% \uparrow$ in $x \rightarrow 0.01\beta_1$ units change in y
$\ln y$ vs. $\ln x$	elasticity (1% in $x \rightarrow \beta_1\%$ in y)
y vs. dummy D	difference in means
$\ln y$ vs. dummy D	$100(e^{\beta_1} - 1)\%$ difference

Common Mistakes to Avoid

- Using $100\beta_1\%$ when β_1 is large (use exact formula).
- Treating dummy coefficients in log models as additive (they are multiplicative).
- Misinterpreting elasticities when logs are missing.
- Forgetting that log-linear models require $y > 0$.
- Forgetting heterogeneity when interactions are present.

Final Takeaways

- ✓ Always check the functional form before interpreting coefficients.
- ✓ Logs turn units into percentages or elasticities.
- ✓ Dummy variables shift levels or percentages depending on log/not-log.
- ✓ Interaction terms change slopes.
- ✓ When in doubt: compute marginal effects explicitly.