Interpreting Regression Coefficients

450A Political Methodology I, handout

This handout explains how to interpret regression coefficients. Section 1 says what a good interpretation reports. Section 2 gives a few mechanical rules. Section 3 works through each combination of levels, logs and shares, with an example from real data and a numerical check.

The examples use three data sets from the lectures. The first covers the 50 U.S. states, with Trump's share of the two-party presidential vote in 2024 and the share of adults 25 and older with a bachelor's degree. The second covers 3,034 U.S. counties, with the same two shares, population, population density and votes cast in 2024. The third covers up to 168 countries in 2024, with V-Dem's political corruption index times 100 (0 is least corrupt, 100 most), women's share of seats in the legislature, GDP per capita in dollars adjusted for purchasing power, and whether the country is a democracy (V-Dem's electoral democracy index above 0.5). The vote shares are MIT Election Data and Science Lab returns. The education shares are from the Census Bureau's American Community Survey, 2019–2023. The country data are from V-Dem (version 16) and the World Bank's World Development Indicators. None of these regressions estimates a causal effect. Each one describes how two variables move together.

1 What to report

Always give the reader an interpretation of the key coefficients in your regression table. It has three parts.

1. **Meaning in units.** Say what the coefficient means in the units in which X and Y are measured.
2. **Statistical significance.** Say how precisely the coefficient is estimated: report a standard error, a p-value or a confidence interval.
3. **Substantive significance.** Say whether the association is large or small, and why. Relate the coefficient to the distributions of X and Y in your sample: pick a plausible, realistic change in X and say how much Y is expected to change with it.

Here are three ways to report the same estimate. Regressing Trump's share of the two-party vote on the share of adults with a bachelor's degree (BA share) across the 50 states gives a slope of -1.45 (standard error 0.17).

- Across states, a BA share 1 percentage point higher is associated with a Trump share 1.45 percentage points lower. The estimate is fairly precise: the 95% confidence interval runs from

-1.80 to -1.10. Trump won 53.4% of the two-party vote in the average state, so a gap of 1.45 points for each point of BA share is substantial.

- A BA share 1 percentage point higher is associated with a Trump share 1.45 percentage points lower (standard error 0.17). To see the size, compare Iowa, at the 25th percentile of the BA share (30.9%), with Illinois, at the 75th (37.2%). The gap of 6.3 points in BA share predicts a Trump share 9.2 points lower in Illinois. The actual gap was 12.3 points (56.7% against 44.4%).

- A BA share 1 percentage point higher is associated with a Trump share 1.45 percentage points lower ($p < 0.001$). The standard deviation across states is 5.4 points for the BA share and 10.2 points for the Trump share. So a state one standard deviation higher in BA share is predicted to have a Trump share 7.9 points, or 0.77 standard deviations, lower.

Two cautions when you judge substantive significance.

- **Stay within the data.** Choose a change in X that is realistic and inside the range of your data. Across countries, a women's share of the legislature 1 percentage point higher is associated with a corruption score 0.58 points lower. Taken literally, the slope says a legislature of only women would score 57.8 points lower than one with no women. But the share runs from 0 (Oman, Yemen) to 63.8% (Rwanda), so no country is anywhere near 100%. Compare realistic cases instead: going from the 25th percentile (North Korea, 17.6%) to the 75th (France, 36.1%) predicts a corruption score 10.7 points lower.

- **Watch low baselines.** When the baseline is small, percent changes can sound large even when the difference is not. Suppose (hypothetically) a study reports that autocracies have a 100% higher predicted probability of civil war in a given year than democracies. That sounds impressive. But if the probability is 0.002 for democracies, a 100% increase takes it only to 0.004: twice as high, yet still very unlikely.

You do not need a long discussion of substantive significance. You do need one realistic scenario that lets the reader judge the size of the slope.

Avoid calling your estimates "effects," or using other causal language, unless you can justify the strong assumptions a causal claim needs (for example, the zero conditional mean assumption). A randomized experiment is one such justification. A research design that rules out the confounding effects of omitted variables is another. Without one, write "is associated with," "predicts," "expected difference" or "correlated with." Nobody should read the women's-share example above as saying that electing more women reduces corruption.

2 Mechanical rules

- A *share* is a variable measured as a fraction from 0 to 1, such as a vote share or the share of adults with a degree. Describe a change in a share in *percentage points*.
- A change of 1 percentage point is a change of 0.01 in the share.
- A percentage-point change is the difference between two percentages. A percent change is the change relative to the starting value. A vote share going from 40% to 44% is a rise of 4 percentage points, or 10 percent.
- A 0–1 index that is not a share (V-Dem's electoral democracy index, say) has no percentage points. Describe changes in it in its own units, such as "0.1 on the index."
- When in doubt about scales or units, plug in simple numbers to check your interpretation. Every case below ends with such a check.
- The rule that a change in logs approximates a percent change works only for small changes. It often fails when Y is in logs and X is binary (Section 3.2).

Throughout, log is the natural logarithm, and Δ means a change. For small changes, $\Delta \log Y \approx \Delta Y/Y$, the proportional change in Y.

3 The main cases

3.1 Level-level: $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$

Y and X both in units (neither is a share)

A 1-unit increase in X is associated with a change of $\hat{\beta}_1$ units in Y.

Example. Across 168 countries, $\widehat{\text{corruption}} = 66.84 - 0.761 \times \text{GDP}$, with corruption on its 0–100 scale and GDP per capita in thousands of dollars. A country with \$1,000 more GDP per capita has a corruption score 0.76 points lower on average.

Check. At \$20,000 the fitted score is 51.63, and at \$21,000 it is 50.87. The difference is -0.76.

Y and X both shares

A 1-percentage-point increase in X is associated with a change of $\hat{\beta}_1$ percentage points in Y.

Example. Across the 50 states, $\widehat{\text{Trump share}} = 1.028 - 1.454 \times \text{BA share}$, both as fractions. A BA share 1 percentage point higher (0.01) is associated with a Trump share 1.45 percentage points lower (0.0145).

Check. At a BA share of 0.30 the fitted Trump share is 0.5923 (59.23%), and at 0.31 it is 0.5777 (57.77%). The difference is -0.0145, or 1.45 percentage points lower.

Do not read this slope as "a 1 percent increase in the BA share is associated with a 1.45 percent decrease in the Trump share." A 1 percent increase from 0.30 is $0.30 \times 1.01 = 0.303$. The fitted Trump share at 0.303 is 0.5879, which is 0.74 percent below 0.5923, not 1.45 percent.

Y a share, X in units

A 1-unit increase in X is associated with a change of $100\,\hat{\beta}_1$ percentage points in Y. The two variables are in different units, so do not forget to multiply by 100.

Example. Across 160 countries, $\widehat{\text{women's share}} = 0.2424 + 0.000888 \times \text{GDP}$, with the share as a fraction and GDP per capita in thousands of dollars. A country with \$1,000 more GDP per capita has a women's share of the legislature 0.089 percentage points higher on average.

Check. At \$20,000 the fitted share is 0.26020 (26.02%), and at \$21,000 it is 0.26109 (26.11%). The difference is 0.00089, or 0.089 percentage points. Even \$10,000 more GDP per capita predicts a share only 0.89 points higher, small next to the average share of 26.7%. The slope is statistically significant ($p = 0.017$) but substantively small.

Y in units, X a share

A 1-percentage-point increase in X is associated with a change of $\hat{\beta}_1/100$ units in Y. Do not forget to divide by 100.

Example. Across 168 countries, $\widehat{\text{corruption}} = 62.11 - 57.80 \times \text{women's share}$. A women's share 1 percentage point higher is associated with a corruption score 0.58 points lower.

Check. At a share of 0.25 the fitted score is 47.66, and at 0.26 it is 47.08. The difference is -0.58.

3.2 Log-level: $\widehat{\log Y} = \hat{\beta}_0 + \hat{\beta}_1 X$

X in units

A 1-unit increase in X is associated with a change of about $100\,\hat{\beta}_1$ percent in Y.

Example. Across 168 countries, $\widehat{\log \text{GDP}} = 10.80 - 0.0256 \times \text{corruption}$, with GDP per capita in dollars. A corruption score 1 point higher is associated with a GDP per capita about 2.6 percent lower.

Why. $\hat{\beta}_1 = \Delta \log \text{GDP} / \Delta \text{corruption}$, and $\Delta \log \text{GDP} \approx \Delta \text{GDP}/\text{GDP}$. So a 1-point change in corruption goes with $\Delta \text{GDP}/\text{GDP} \approx \hat{\beta}_1 \times 1 = -0.0256$, or -2.56 percent.

Check. At a score of 50 the fitted GDP per capita, $\exp(\widehat{\log \text{GDP}})$, is \$13,714, and at 51 it is \$13,367. That is 2.5 percent lower, close to the approximation.

X binary (a dummy coded 0 or 1)

Here a 1-unit change means going from 0 to 1, which is a large change. The approximation that a log change is a percent change can be badly off when $\hat{\beta}_1$ is not small. Do not say that going from 0 to 1 is associated with a $100\,\hat{\beta}_1$ percent change in Y.

Compute the exact percent change instead. Write $\hat{Y}_1$ and $\hat{Y}_0$ for $\exp(\widehat{\log Y})$ at $D = 1$ and $D = 0$. Then $\log \hat{Y}_1 - \log \hat{Y}_0 = \hat{\beta}_1$, so $\hat{Y}_1/\hat{Y}_0 = e^{\hat{\beta}_1}$ and

$$100 \times \frac{\hat{Y}_1 - \hat{Y}_0}{\hat{Y}_0} = 100 \times \left(\frac{\hat{Y}_1}{\hat{Y}_0} - 1 \right) = 100 \times \left(e^{\hat{\beta}_1} - 1 \right).$$

Say: going from $D = 0$ to $D = 1$ is associated with a $100(e^{\hat{\beta}_1} - 1)$ percent change in Y.

Example. Across 168 countries, $\widehat{\log \text{GDP}} = 9.139 + 0.919 \times \text{Democracy}$. The approximation would say democracies are about 92 percent richer. The exact figure is $100(e^{0.919} - 1) = 151$ percent.

Check. The fitted GDP per capita is \$9,314 for the 83 autocracies and \$23,339 for the 85 democracies. The ratio is 2.51, so democracies are 151 percent richer, not 92 percent.

These fitted values are the geometric means of GDP per capita in each group (the exponential of the mean log), so the exact figure compares typical values. It also equals the percent difference in average GDP per capita only if the errors have the same distribution in both groups (with normal errors, the same variance). The error of the approximation grows with $\hat{\beta}_1$. At $\hat{\beta}_1 = 0.2$ the exact change is already 22.1 percent, not 20.

X a share

A 1-percentage-point increase in X is associated with a change of about $\hat{\beta}_1$ percent in Y. Do not multiply $\hat{\beta}_1$ by 100.

Example. Across 3,034 counties, $\widehat{\log \text{density}} = 1.949 + 7.673 \times \text{BA share}$, with population density in people per square mile. A BA share 1 percentage point higher is associated with a density about 7.7 percent higher.

Why. A 1-percentage-point change is $\Delta\text{BA share} = 0.01$, so $\Delta\text{density}/\text{density} \approx \hat{\beta}_1 \times 0.01$, and in percent $100 \times \hat{\beta}_1 \times 0.01 = \hat{\beta}_1$.

Check. At a BA share of 0.20 the fitted density is 32.6 people per square mile, and at 0.21 it is 35.2. That is 8.0 percent higher. The approximation is off by 0.3 points even here, because the log change, 0.077, is not that small.

3.3 Level-log: $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \log X$

Y in units

A 1 percent increase in X is associated with a change of about $\hat{\beta}_1/100$ units in Y. Divide the coefficient by 100.

Example. Across 168 countries, $\widehat{\text{corruption}} = 207.9 - 16.75 \times \log \text{GDP}$. A GDP per capita 1 percent higher is associated with a corruption score 0.17 points lower.

Why. $\hat{\beta}_1 = \Delta\text{corruption}/\Delta \log \text{GDP}$ and $\Delta \log \text{GDP} \approx \Delta\text{GDP}/\text{GDP}$. A 1 percent change in GDP means $\Delta\text{GDP}/\text{GDP} = 0.01$, so $\Delta\text{corruption} \approx 0.01\,\hat{\beta}_1 = -0.1675$.

Check. At \$20,000 the fitted score is 41.94, and at \$20,200 (1 percent more) it is 41.77. The difference is -0.17. For larger changes, use the exact form $\hat{\beta}_1 \log(X_1/X_0)$: doubling GDP per capita predicts a score $16.75 \times \log 2 = 11.6$ points lower.

Y a share

A 1 percent increase in X is associated with a change of about $\hat{\beta}_1$ percentage points in Y. Do not divide $\hat{\beta}_1$ by 100.

Example. Across 3,034 counties, $\widehat{\text{Trump share}} = 0.8581 - 0.0471 \times \log \text{density}$. A county 1 percent denser has a Trump share about 0.047 percentage points lower.

Why. A 1 percent change in density changes the fitted share by about $0.01\,\hat{\beta}_1 = -0.000471$, which is -0.047 percentage points (1 percentage point is 0.01).

Check. At 100 people per square mile the fitted share is 0.641258, and at 101 it is 0.640790. The difference is -0.000469, or 0.047 percentage points lower. A 1 percent change is tiny, so describe a larger one: a county ten times as dense is predicted to give Trump a share $0.0471 \times \log 10 = 0.108$, or 10.8 percentage points, lower.

3.4 Log-log: $\widehat{\log Y} = \hat{\beta}_0 + \hat{\beta}_1 \log X$

A 1 percent increase in X is associated with a change of about $\hat{\beta}_1$ percent in Y. $\hat{\beta}_1$ is the *elasticity* of Y with respect to X.

Example. Across 3,034 counties, $\widehat{\log \text{votes}} = -0.552 + 0.983 \times \log \text{population}$, with votes cast in 2024. A county with 1 percent more residents cast about 0.98 percent more votes. An elasticity close to 1 means votes rise almost in proportion to population.

Why. $\hat{\beta}_1 = \Delta \log \text{votes} / \Delta \log \text{pop} \approx (\Delta \text{votes}/\text{votes})/(\Delta \text{pop}/\text{pop})$. Rearranged,

$$\hat{\beta}_1 \approx \frac{\Delta \text{votes}}{\Delta \text{pop}} \times \frac{\text{pop}}{\text{votes}},$$

which is the usual definition of an elasticity: $(\partial Y/\partial X) \times (X/Y)$.

Check. At a population of 50,000 the fitted number of votes is 24,032, and at 50,500 it is 24,268. That is 0.98 percent more.

3.5 Summary

Model	Change in X	Associated change in Y
Level-level	+1 unit	$\hat{\beta}_1$ units
	+1 unit (Y a share)	$100\,\hat{\beta}_1$ percentage points
	+1 percentage point (X a share)	$\hat{\beta}_1/100$ units
	+1 percentage point (both shares)	$\hat{\beta}_1$ percentage points
Log-level	+1 unit	about $100\,\hat{\beta}_1$ percent
	+1 percentage point (X a share)	about $\hat{\beta}_1$ percent
	$0 \to 1$ (X binary)	exactly $100(e^{\hat{\beta}_1} - 1)$ percent
Level-log	+1 percent	about $\hat{\beta}_1/100$ units
	+1 percent (Y a share)	about $\hat{\beta}_1$ percentage points
Log-log	+1 percent	about $\hat{\beta}_1$ percent