Econ 103: Introduction to Econometrics

Lecture 12 — Functional Forms

Ryan Longmuir

UCLA

Summer Session C, 2026

Reading: Hill, Griffiths & Lim (5th ed.), §2.8, 4.3–4.6; Stock & Watson (4th ed.), §8.1–8.2.

Where we are

So far the regression line has been *straight*: $\mathbb{E}(y \mid x) = \beta_1 + \beta_2 x$. But the world rarely is.

The key realization: "linear regression" means linear in the parameters β — *not* in the variables. We are free to feed OLS *transformed* variables:

$$y = \beta_1 + \beta_2 \underbrace{f(x)}_{\text{e.g. } x^2,\ \ln x} + e, \qquad \underbrace{g(y)}_{\text{e.g. } \ln y} = \beta_1 + \beta_2 x + e.$$

It is still OLS, still BLUE — only the interpretation of β_2 changes.

Today: the menu of functional forms — quadratics, logs, elasticities — when to use each, and how to read a transformed coefficient.

Today's plan

1. Linear in parameters, not variables
2. Polynomials
3. Logarithms: the three cases
4. Choosing and interpreting

Part 1

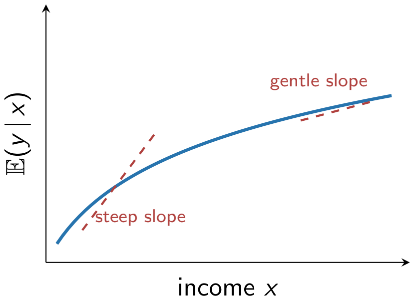
Linear in parameters, not variables

Why a straight line often won't do

Economic theory frequently predicts a changing slope:

- Food spending rises with income, but **at a decreasing rate** — the marginal propensity to spend falls as you get richer.
- Test scores rise with district income, steeply at first, then flatten.
- Cost curves are U-shaped; total product is S-shaped.

A straight line forces *one* slope everywhere. Transforming x or y lets the slope — and the elasticity — change from point to point.



"Increasing at a decreasing rate."

Two tools, many shapes

Almost everything in this lecture is built from just two transformations:

Powers

x^2, x^3, $1/x, \ldots$ — quadratics (U or ∩ shapes), cubics (S-shapes). Captures turning points and acceleration.

Natural logarithm

$\ln(x)$, $\ln(y)$ — converts changes into percentage changes. Tames right-skewed money variables (income, wages, prices) and delivers **elasticities**.

The one habit that prevents every mistake

Whenever you transform a variable, the slope and elasticity formulas change. Before interpreting any coefficient, ask: *is each side in levels, or in logs, or squared?* The answer dictates the wording.

Part 2

Polynomials

Quadratics: a slope that grows

The simple quadratic regression puts a *squared* regressor on the right:

$$y = \beta_1 + \beta_2 x^2 + e \qquad \Longrightarrow \qquad \text{slope} = \frac{dy}{dx} = 2\beta_2 x.$$

The slope is no longer constant — it grows (in magnitude) with x. If $\beta_2 > 0$ the curve sweeps upward ever more steeply.

Example (House prices (HGL, Baton Rouge))

$\widehat{PRICE} = 55{,}776 + 0.0154\, SQFT^2$. *The estimated price of one more square foot is* $2(0.0154)\, SQFT$*:*

$$\textit{at } 2000 \textit{ sqft} : \$61.69, \qquad \textit{at } 4000 \textit{ sqft} : \$123.37.$$

Bigger homes command a higher price per added square foot — exactly the curvature a single straight line cannot express.

Cubics, and a note on the general polynomial

A cubic term captures S-shapes (total cost, total product) or growth that accelerates:

$$y = \beta_1 + \beta_2 x^3 + e, \qquad \text{slope} = 3\beta_2 x^2.$$

Example (Wheat yield over time (HGL))

A straight line in TIME left U-shaped residuals (it missed the acceleration). Using $TIMECUBE = (TIME/100)^3$*:*

$$\widehat{YIELD} = 0.874 + 9.682\, TIMECUBE, \quad R^2 : 0.649 \to 0.751.$$

Here a polynomial uses *one* transformed regressor, so it fits in simple regression. The richer form $y = \beta_1 + \beta_2 x + \beta_3 x^2$ has *two* regressors (x and x^2) — that needs **multiple regression** (Lecture 13), where we can also *test* whether the curve is needed.

Reading a polynomial: plot it, or evaluate the slope

A polynomial coefficient has no standalone interpretation — "β_2 holding x^2 fixed" is meaningless. Instead:

- Plot the fitted curve over the data and describe its shape.
- Evaluate the slope dy/dx at a few interesting values of x (low, median, high) and report those marginal effects — as we did with house prices at 2000 vs. 4000 sqft.

Marginal effects are local now

The whole point of a nonlinear form is that "the effect of x on y" is no longer a single number. Always quote it *at a stated value of* x.

Part 3

Logarithms: the three cases

The magic of logs: changes become percentages

The one fact behind every log interpretation: for a *small* change,

$$\ln(x + \Delta x) - \ln(x) \approx \frac{\Delta x}{x} = \%\Delta x \,/\, 100.$$

A *difference in logs* is (approximately) a percentage change.

Why economists reach for logs constantly:

- many relationships are naturally proportional (“a 1% price rise cuts quantity by η%”);
- money variables (income, wages, prices, sales) are right-skewed — logging them pulls in the long tail toward normality (helps SR6);
- logs deliver unit-free elasticities you can compare across studies.

Case 1 — log-linear: $\ln(y) = \beta_1 + \beta_2 x$

Only the *dependent* variable is logged. A one-unit change in x is associated with a $100\beta_2\%$ change in y — a *constant growth rate* interpretation. (Requires $y > 0$.)

Example (Returns to education (HGL, `cps5_small`))

$$\widehat{\ln(WAGE)} = 1.597 + 0.0988\, EDUC.$$

Each extra year of education raises the wage by about $100(0.0988) \approx 9.9\%$ *(95% CI* $\approx 8.9\%$ *to* 10.9%*). A percentage return — not a fixed dollar amount — which matches how the labor market actually works.*

Sign of β_2 sets the direction; the curve rises (or falls) at a changing rate in the *levels* of y.

Case 2 — linear-log: $y = \beta_1 + \beta_2 \ln(x)$

Only the *regressor* is logged. A 1% change in x is associated with a $\beta_2/100$-unit change in y. (Requires $x > 0$.) The curve rises at a *decreasing* rate when $\beta_2 > 0$ — ideal for the food example.

Example (Food expenditure, linear-log (HGL))

$$\widehat{FOOD\_EXP} = -97.19 + 132.17 \ln(INCOME), \quad R^2 = 0.357.$$

A 1% income rise adds about $132.17/100 = \$1.32$ *to weekly food spending. The marginal effect of \$100 more income shrinks with income (\$13.22 at \$1,000/wk vs. \$6.61 at \$2,000/wk) — the declining MPC theory predicted.*

Case 3 — log-log: $\ln(y) = \beta_1 + \beta_2 \ln(x)$

Both sides logged. Now β_2 *is* the **elasticity** of y with respect to x — and it is constant along the whole curve. (Requires $x, y > 0$.)

$$\beta_2 = \frac{\%\Delta y}{\%\Delta x} = \text{elasticity of } y \text{ w.r.t. } x.$$

- A 1% change in x $\Rightarrow$ a β_2% change in y.
- This is why log-log is the workhorse for demand curves (price elasticity) and production functions (constant-returns checks) — the elasticity *is* the parameter.
- Example (SW, test scores): $\ln(\widehat{\text{TestScore}}) = 6.336 + 0.0554 \ln(\text{Income})$ — a 1% income rise $\rightarrow$ 0.0554% higher scores.

The master table

Name	Model	Slope dy/dx	Interpretation of β_2
Linear	$y = \beta_1 + \beta_2 x$	β_2	1-unit $\Delta x \to \beta_2$-unit Δy
Quadratic	$y = \beta_1 + \beta_2 x^2$	$2\beta_2 x$	slope changes with x
Cubic	$y = \beta_1 + \beta_2 x^3$	$3\beta_2 x^2$	slope changes with x
Log-linear	$\ln y = \beta_1 + \beta_2 x$	$\beta_2 y$	1-unit $\Delta x \to 100\beta_2\%\ \Delta y$
Linear-log	$y = \beta_1 + \beta_2 \ln x$	β_2/x	$1\%\ \Delta x \to \beta_2/100$-unit Δy
Log-log	$\ln y = \beta_1 + \beta_2 \ln x$	$\beta_2\, y/x$	$1\%\ \Delta x \to \beta_2\%\ \Delta y$ (elasticity)

Decode any specification by location of the logs

In on the left $\Rightarrow$ effect is in *percent of* y. In on the right $\Rightarrow$ cause is a *percent of* x. Logs on both $\Rightarrow$ elasticity. Keep this map and you can read any of them.

Part 4

Choosing and interpreting

Choosing a functional form

Three guideposts (HGL):

1. Theory first. Pick a shape consistent with the economics — declining MPC, constant elasticity, U-shaped cost. Decide *before* looking at the data whether the slope should vary, and how.
2. Flexibility. The form must be able to bend the way the data bend (residual plots reveal a missed curve — more in Lecture 18).
3. Assumptions. Prefer a form under which SR1–SR6 look reasonable (e.g. logging a skewed y often tames heteroskedasticity and non-normal errors).

We never know the “true” form — every choice is an approximation. The goal is a form that is theoretically sensible, fits, and respects the assumptions.

Comparing fit: a sharp warning about R^2

You may compare R^2 across models only when they share the same dependent variable.

Fair comparison

Linear y vs. linear-log y vs. quadratic in x — all have y on the left. Food: linear $R^2 = 0.385$ vs. linear-log 0.357 — comparable (and nearly tied).

Invalid comparison

y-model vs. $\ln(y)$-model. Different dependent variables explain different "total variation," so their R^2's are not on the same scale. Choose between them with theory, not R^2.

When you must summarize fit for a logged-y model, use the generalized $R_g^2 = [\mathrm{corr}(y, \hat{y})]^2$ computed on the *original* y scale.

A bonus: logs and growth rates

Log-linear models fall straight out of **compound interest**. If y grows at a constant rate g per period, $y_t = y_0(1+g)^t$, then taking logs:

$$\ln(y_t) = \underbrace{\ln(y_0)}_{\beta_1} + \underbrace{\ln(1+g)}_{\beta_2}\, t, \qquad \beta_2 = \ln(1+g) \approx g.$$

So the slope on t in a log-linear time trend is the growth rate.

Example (Wheat-yield growth (HGL))

$\widehat{\ln(YIELD)} = -0.343 + 0.0178\, t \Rightarrow \hat{g} \approx 1.78\%$ *per year from technological progress.*

One caveat for later: to predict y itself from a log-linear model, $\exp(b_1 + b_2 x)$ slightly *under*-predicts; a correction factor $e^{\hat{\sigma}^2/2}$ fixes it (HGL §4.5.1). Mind the level-vs-log scale when forecasting.

Recap

The big idea

- "linear" = linear in β, not in x
- transform with powers and logs; still OLS/BLUE
- slope & elasticity now vary point to point

Polynomials

- $y = \beta_1 + \beta_2 x^2$: slope $2\beta_2 x$ (house prices)
- interpret by plotting / slope at chosen x

The three log cases

- log-linear: 1-unit $\Delta x \rightarrow 100\beta_2\%\ \Delta y$ (wage $\approx$9.9%/yr)
- linear-log: 1% $\Delta x \rightarrow \beta_2/100$ unit Δy (food)
- log-log: β_2 = elasticity (demand, production)

Choosing

- theory + flexibility + assumptions
- R^2 comparable only if same dependent variable

Next time (Lecture 13): the multiple regression model

One regressor is rarely enough. We add $X_2, X_3, \ldots$ — to control for confounders and finally give *ceteris paribus* real teeth. Enter Big Andy's Burgers (SALES, PRICE, ADVERT).

Questions?