Logarithms and Nonlinear Interpretation

21 Oct 2026 | Chapter 8

ECO 351 – Introduction to Econometrics

Shi Qiu – Fall 2026

Equal dollar increases and equal percentage increases

In Lecture 17, the quadratic gave about 2.96 score points for income \$10,000 to \$11,000, but only 0.42 for \$40,000 to \$41,000.

A short return to that example

1. Both add \$1,000. What is the percentage increase in each case?
2. What ending income would give a 10% increase from \$40,000?

Today we ask whether proportional income changes provide a useful way to describe the curve.

Log differences measure proportional changes

For positive starting and ending values x_0, x_1, the natural logarithm satisfies

$$\ln x_1 - \ln x_0 = \ln(x_1/x_0).$$

Income change	Percentage change	Log difference
10 to 11	10%	$\ln(1.10) = 0.09531$
40 to 41	2.5%	$\ln(1.025) = 0.02469$
40 to 44	10%	$\ln(1.10) = 0.09531$

For a small change, $\ln(x_1/x_0) \approx (x_1 - x_0)/x_0$.

ln means base e. Its inverse is exp: $\exp(\ln x) = x$.

Scores in points, income in logarithms

For district i, let $Score_i$ be test-score points and $Income_i$ income in thousands of dollars. Consider

$$Score_i = \beta_0 + \beta_1 \ln(Income_i) + u_i.$$

This is a **level-log** model: the outcome stays in levels, while the regressor is logged.

- Estimate the intercept and log-income slope by OLS.
- A 1% income increase is associated with approximately $\beta_1/100$ score points.

Income must be positive. Adding a constant just to make logs possible changes the specification.

Stata: estimate the level-log model

```
. assert income > 0 & !missing(income)
. generate logincome = ln(income)
. regress score logincome, vce(robust)
Linear regression                               Number of obs     =        420
                                                R-squared         =     0.5625
------------------------------------------------------------------------------
             |               Robust
       score | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   logincome |   36.41968   1.396943    26.07   0.000     33.67378    39.16559
       _cons |   557.8323    3.83994   145.27   0.000     550.2843    565.3803
------------------------------------------------------------------------------
```

The fitted equation, rounded for hand calculations, is

$$\widehat{Score}(x) = 557.8 + 36.42 \ln x.$$

Exact fitted changes use the log ratio

Let $x_0, x_1 > 0$ be starting and ending incomes in thousands of dollars. Subtract the two fitted scores:

$$\begin{aligned}\Delta\widehat{Score} &= [557.8 + 36.42 \ln x_1] - [557.8 + 36.42 \ln x_0] \\ &= 36.42 \ln(x_1/x_0).\end{aligned}$$

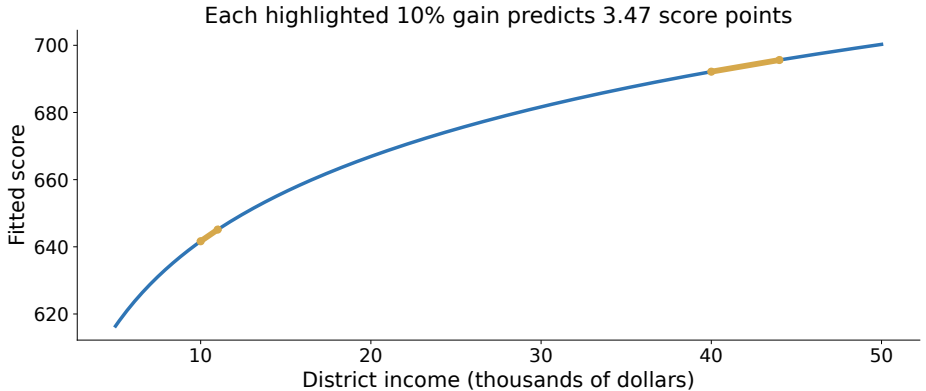
For \$10,000 to \$11,000,

$$36.42 \ln(1.10) = 3.4712 \text{ score points.}$$

Using the small-change approximation gives $36.42(0.10) = 3.642$ points.

The log-ratio calculation is exact for the displayed fitted equation. The percentage shortcut is an approximation.

Equal percentage increases on the fitted curve



The highlighted changes are 10 to 11 and 40 to 44: both raise income by 10%. This curve evaluates the rounded equation, not additional district observations.

Practice: the same dollar increase at two incomes

Use the fitted school model

$$\widehat{Score}(x) = 557.8 + 36.42 \ln x,$$

where income x is in thousands of dollars.

Pairs – 5 minutes

1. Find the exact fitted gain for \$20,000 to \$24,000.
2. Find the exact fitted gain for \$40,000 to \$44,000.
3. Both add \$4,000. Explain which gain is larger and why.

State the answer's units and distinguish a fitted association from a causal effect.

Debrief: the ratio determines the fitted gain

Income change	Ratio	Fitted score gain
20 to 24	1.20	$36.42 \ln(1.20) = 6.6402$
40 to 44	1.10	$36.42 \ln(1.10) = 3.4712$

The same \$4,000 is a 20% increase at the first starting income, but only 10% at the second.

The approximate gains are $36.42(0.20) = 7.284$ and $36.42(0.10) = 3.642$ points.

First form the income ratio. Then take its log and multiply by the slope.

Stata: uncertainty for the 20-to-24 comparison

```
. local logchange20 = ln(24/20)
. lincom `logchange20'*logincome
 ( 1)  .1823216*logincome = 0
------------------------------------------------------------------------------
       score | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   6.640093   .2546928    26.07   0.000     6.139455    7.140732
------------------------------------------------------------------------------
```

The multiplier is the exact log difference. The interval describes uncertainty about the fitted score difference, in points.

Small differences from 6.6402 come from using full-precision coefficient estimates.

A logged outcome answers a percentage question

Suppose we describe wages rather than test scores. In this constructed example, S is years of schooling and $W > 0$ is hourly wage:

$$\widehat{\ln W}(S) = 1.8 + 0.08S.$$

This is a **log-level** model. One extra year changes the fitted log wage by 0.08.

Define $\widetilde{W}(S) = \exp[\widehat{\ln W}(S)]$, the exponentiated fitted log wage. Then

$$\frac{\widetilde{W}(S+1)}{\widetilde{W}(S)} = \exp(0.08).$$

The change is proportional on this wage scale, rather than a fixed number of dollars.

Converting a log-wage change into a percentage

For a schooling increase of d years, the fitted log change is $0.08d$:

$$\%\Delta\widehat{W} = 100[\exp(0.08d) - 1].$$

Schooling increase	Exact fitted change	Approximation
1 year	$100(e^{0.08} - 1) = 8.33\%$	8%
2 years	$100(e^{0.16} - 1) = 17.35\%$	16%

With a logged outcome, exponentiate the log change before subtracting one. Multiplying by 100 alone is the small-change approximation.

An average log and an average level are different

At a fixed schooling level, suppose two wages are equally likely: \$1 and \$9.

$$\text{Arithmetic mean wage} = \frac{1+9}{2} = \$5.$$

The average log wage, exponentiated, gives the geometric mean:

$$\exp\left(\frac{\ln 1 + \ln 9}{2}\right) = \$3.$$

Exponentiating a fitted average log wage does not automatically recover the average wage in dollars. Our percentage formulas describe the exponentiated fitted log scale.

Logging both variables gives a constant elasticity

In a constructed demand example, $P > 0$ is price per unit and $Q > 0$ quantity per week:

$$\widehat{\ln Q}(P) = 4 - 1.2 \ln P.$$

This is a **log-log** model. The coefficient -1.2 is an elasticity on the fitted scale: a 1% price increase corresponds to approximately a 1.2% quantity decrease.

With $r = P_1/P_0$ and $\widetilde{Q} = \exp[\widehat{\ln Q}]$,

$$\frac{\widetilde{Q}(P_1)}{\widetilde{Q}(P_0)} = r^{-1.2}, \qquad \%\Delta\widetilde{Q} = 100(r^{-1.2} - 1).$$

Practice: a price rise and a price cut

Use the same fitted elasticity, -1.2. Let r be the new price divided by the old price.

Pairs – 4 minutes

1. For a 10% price rise, calculate the approximate and exact fitted percentage change in quantity.
2. Repeat for a 10% price cut. Which ratio belongs in the formula?
3. Must the exact quantity changes have equal magnitudes? Explain.

Use $100(r^{-1.2} - 1)$ for the exact percentage change.

Debrief: finite changes need the full ratio

Price change	Approximate change in Q	Exact fitted change in Q
$+10\%$	-12%	$100(1.10^{-1.2} - 1) = -10.81\%$
-10%	$+12\%$	$100(0.90^{-1.2} - 1) = +13.48\%$

The elasticity is the same in both calculations. The exact finite changes differ in magnitude.

A 10% price cut uses $r = 0.90$, not $r = -0.10$.

Elasticity times percentage change is a local approximation. Use the ratio-power calculation for a finite change.

Which side is logged determines the units

Form	Fitted equation	Small-change interpretation
Level-log	$\widehat{Y} = a + b \ln X$	1% increase in X: about $b/100$ units
Log-level	$\widehat{\ln Y} = a + bX$	1-unit increase in X: about $100b\%$
Log-log	$\widehat{\ln Y} = a + b \ln X$	1% increase in X: about $b\%$

Here a, b are fitted coefficients. Units refer to Y; percentage changes refer to $\widetilde{Y} = \exp[\widehat{\ln Y}]$.

These are approximations for small changes. Use log ratios or exponentiation for exact finite comparisons.

A coefficient means different things in different forms

Individual check – 4 minutes

Both examples are constructed.

1. $\widehat{Rent} = a - 20\ln(D)$, where rent is dollars per month and $D > 0$ is distance in kilometers. Find the exact fitted rent change when distance rises by 10%.
2. $\widehat{\ln Sales} = c + 0.04A$, where sales are positive and A is advertising spending in thousands of dollars. Find the exact fitted percentage change for \$2,000 more advertising.

For each answer, identify which side is logged and state the units.

Debrief: state the scale before the number

Rent: only the regressor is logged.

$$\Delta\widehat{Rent} = -20\ln(1.10) \approx -\$1.91 \text{ per month.}$$

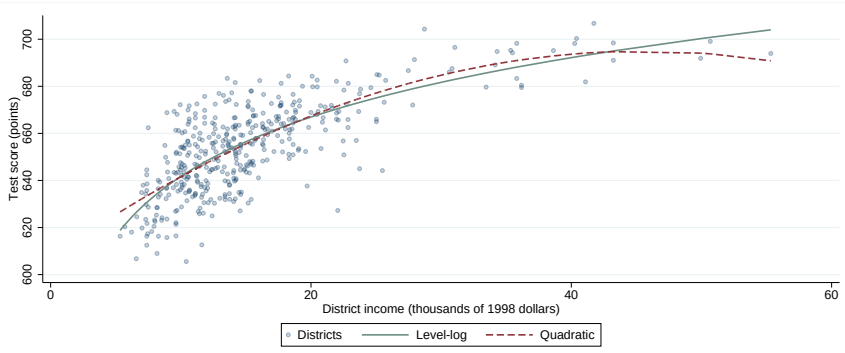
Sales: only the outcome is logged. An extra \$2,000 means $\Delta A = 2$.

$$\%\Delta\widetilde{Sales} = 100[e^{0.04(2)} - 1] \approx 8.33\%.$$

The approximate answers are a \$2 rent decrease and an 8% sales increase.

A percentage change in a regressor need not produce a percentage answer. The dependent variable's scale determines the answer's units.

Log and quadratic fits on the same score scale



Same 420 districts, same score outcome. R^2 comparisons require the same outcome scale.

Next: interactions between regressors.

Optional B1: the retransformation factor

Let $\ln Y = m(X) + u$, where $m(X) = E(\ln Y \mid X)$ and $E(u \mid X) = 0$.

$$E(Y \mid X) = e^{m(X)}\, E(e^{u} \mid X).$$

Define $c(X) = E(e^{u} \mid X)$. The ratio of arithmetic conditional means is

$$\frac{E(Y \mid X = x_1)}{E(Y \mid X = x_0)} = e^{m(x_1) - m(x_0)} \frac{c(x_1)}{c(x_0)}.$$

If $c(x_1) = c(x_0)$, the factor cancels. Otherwise the exponentiated log prediction alone does not give the mean-level ratio.

Optional B2: exact finite-change reference

For positive X_0, X_1, let $r = X_1/X_0$ and $d = X_1 - X_0$. In each row, b is the fitted slope.

Fitted specification	Exact finite comparison
$\widehat{Y} = a + b \ln X$	$\Delta \widehat{Y} = b \ln r$
$\widehat{\ln Y} = a + bX$	$\%\Delta \widetilde{Y} = 100(e^{bd} - 1)$
$\widehat{\ln Y} = a + b \ln X$	$\%\Delta \widetilde{Y} = 100(r^b - 1)$

Here $\widetilde{Y} = \exp[\widehat{\ln Y}]$. Retain the same controls when comparing predictions in a multiple regression.

Logs require positive values. A percentage-point change in an already-percentage variable is not automatically a percentage change in its level.