Interactions and Group Comparisons

26 Oct 2026 | Chapter 8

ECO 351 – Introduction to Econometrics

Shi Qiu – Fall 2026

Last lecture: a common return to schooling

In Lecture 18's constructed wage example, S was years of schooling and W hourly wage:

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

1. What percentage change does one more year of schooling imply, approximately?
2. Does this equation allow that percentage change to differ across groups of workers?

Explain what feature of the equation determines your answer.

Training premiums in four groups

In this constructed example, W is hourly wage in AED. $T = 1$ denotes training and $L = 1$ low prior experience. A two-indicator model is

$$\widehat{W}(T, L) = b_0 + b_1 T + b_2 L + b_3 TL,$$

where the b's are fitted coefficients.

	No training	Training
More experienced, $L = 0$	50	58
Less experienced, $L = 1$	45	?

Suppose the fitted training premium is 5 AED larger for less experienced workers.

Complete the missing cell and find all four coefficients. Explain what b_3 measures.

The product of two indicators captures the extra gap

The four fitted values can be written as

$$\widehat{W}(T, L) = 50 + 8T - 5L + 5TL,$$

where W is hourly wage in AED.

$$L = 0: \quad \widehat{W}(1, 0) - \widehat{W}(0, 0) = 8,$$
$$L = 1: \quad \widehat{W}(1, 1) - \widehat{W}(0, 1) = 8 + 5 = 13.$$

The interaction coefficient is $13 - 8 = 5$: the difference in the two training gaps.

The coefficient 8 is the training gap only in the reference group $L = 0$.

A group indicator shifts the line, but not its slope

For the school data, let Y be score points, $X = STR$ students per teacher, and $D = 1$ for more than 10% English learners ($D = 0$ otherwise).

$$Y = \beta_0 + \beta_1 X + \beta_2 D + u.$$

The β's are population coefficients and u contains other score determinants. Under $E(u \mid X, D) = 0$, the regression function is

$$f(x, 0) = \beta_0 + \beta_1 x,$$
$$f(x, 1) = (\beta_0 + \beta_2) + \beta_1 x.$$

The group indicator shifts the intercept. Both groups still have slope β_1.

The product term relaxes the common-slope restriction

Add the **interaction** XD: it is zero when $D = 0$ and equals X when $D = 1$.

$$Y = \beta_0 + \beta_1 X + \beta_2 D + \beta_3 XD + u.$$

Substitution now gives

$$f(x, 0) = \beta_0 + \beta_1 x,$$
$$f(x, 1) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)x.$$

	$D = 0$	$D = 1$
Slope	β_1	$\beta_1 + \beta_3$

β_3 is the second group's slope minus the first group's slope.

Stata estimates both group lines in one regression

```
. regress score i.high_el##c.str, vce(robust)
Linear regression                               Number of obs     =        420
------------------------------------------------------------------------------
             |               Robust
       score | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      high_el |
    EL > 10%  |   5.639141   19.51456     0.29   0.773    -32.72029    43.99857
         str  |  -.9684601   .5891016    -1.64   0.101    -2.126447    .1895268
high_el#c.str |
    EL > 10%  |  -1.276613   .9669194    -1.32   0.187     -3.17727    .6240436
        _cons |   682.2458   11.86781    57.49   0.000     658.9175    705.5742
------------------------------------------------------------------------------
```

includes the main terms and their product.

Check: Reconstruct the intercept and STR slope for each group. Compare with a partner.

Two feasible class-size plans

The fitted lines are

$$\widehat{f}(x,0) = 682.246 - 0.968x, \qquad \widehat{f}(x,1) = 687.885 - 2.245x.$$

Consider two hypothetical districts, one from each group, both starting at STR 22. Suppose either package is feasible:

Package	Low-EL STR change	High-EL STR change
A	-1	-1
B	0	-2

1. Find each district's fitted gain under each package.
2. Compare the average gain across the two districts, weighting them equally.
3. What would you want to know before recommending a package?

Debrief: targeting the larger fitted response

Using the displayed rounded slopes:

Package	Low-EL gain	High-EL gain	Average gain
A	0.968	2.245	1.607
B	0	4.490	2.245

All gains are in score points. Package B has a larger fitted average gain by about 0.639 points.

The point estimates favor B under the stated objective. Before making a recommendation, we still need to assess uncertainty and whether the regression identifies a causal response.

Which claim does this output support?

At a common STR x, subtraction gives the fitted group gap

$$\widehat{f}(x,1) - \widehat{f}(x,0) = \widehat{\beta}_2 + x\widehat{\beta}_3.$$

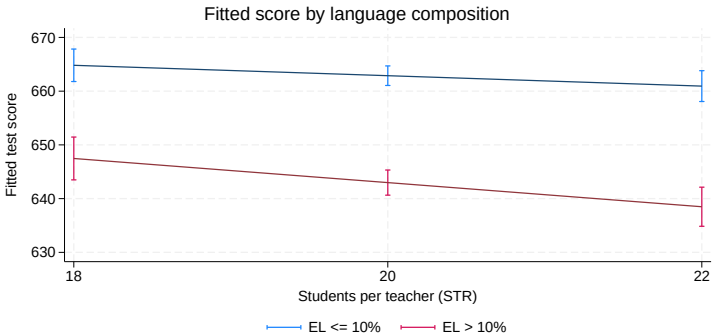
The excerpt below evaluates the groups at STR 20 (row 2).

```
. margins high_el, at(str=(18 20 22))
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
 _at#high_el |
2#EL <= 10%  |   662.8766   .9248105   716.77   0.000     661.0588    664.6945
 2#EL > 10%  |   642.9835    1.18965   540.48   0.000      640.645     645.322
------------------------------------------------------------------------------
```

1. At STR 20, the high-EL fitted score is about 19.89 points lower.
2. Reducing STR by two in the high-EL group raises its fitted score by 19.89 points.

The graph separates a vertical gap from a slope difference

`marginsplot` displays the preceding fitted values and 95% intervals.



The vertical gap at STR 20 is about 19.89 points. A two-unit STR reduction along the high-EL line gives about 4.49 points.

A large group gap does not establish different slopes

```
. test 1.high_el#c.str
 ( 1)  1.high_el#c.str = 0
       F(  1,   416) =    1.74
            Prob > F =    0.1875
```

This tests equal population STR slopes: $H_0 : \beta_3 = 0$.

With $p = 0.1875$, we do not reject equal slopes at 5%. We have not proved that the slopes are equal.

Check: Would you now describe Package B's larger fitted gain as an established policy advantage? Explain.

Keep English-learner share continuous

The 10% cutoff discards variation within each group. Instead, let $Z = EL$, measured in **percentage points**, and $X = STR$.

$$Score = \gamma_0 + \gamma_1 X + \gamma_2 Z + \gamma_3 XZ + u.$$

The γ's are population coefficients. Under $E(u \mid X, Z) = 0$, hold $Z = z$ fixed and subtract the regression functions at $x + h$ and x:

$$f(x + h, z) - f(x, z) = (\gamma_1 + \gamma_3 z)h.$$

Thus the STR slope is $\gamma_1 + \gamma_3 z$.

An interaction can let a slope vary with a measured quantity, not just with group membership.

Stata: estimate the continuous-by-continuous interaction

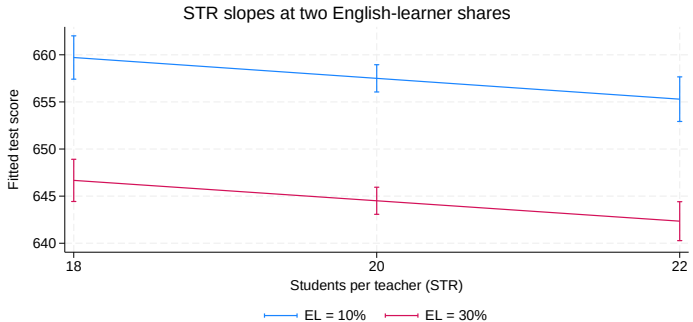
```
. regress score c.str##c.el, vce(robust)
Linear regression                               Number of obs     =        420
------------------------------------------------------------------------------
             |               Robust
       score | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         str |  -1.117018   .5875135    -1.90   0.058    -2.271884     .037847
          el |  -.6729114   .3741231    -1.80   0.073    -1.408319     .062496
  c.str#c.el |   .0011618   .0185357     0.06   0.950    -.0352736    .0375971
       _cons |   686.3385   11.75935    58.37   0.000     663.2234    709.4537
------------------------------------------------------------------------------
```

At EL share z, the fitted STR slope is $-1.117018 + 0.0011618z$.

Check: Calculate the fitted gain from reducing STR by two at EL 10 and EL 30. Does $p = 0.950$ support claiming different responses?

Nearly parallel lines can still have a large vertical gap

```
margins, at(str=(18 20 22) el=(10 30))
marginsplot, xdimension(str) plotdimension(el)
```



A two-unit STR reduction gives about 2.211 points at EL 10 and 2.164 at EL 30.

Training and experience

In this constructed fitted model, P is hourly pay in AED, H is training hours and E is years of prior experience:

$$\widehat{P}(H, E) = 40 + 1.5H + 2E + 0.1HE.$$

1. Find the fitted pay gain from one extra training hour at $E = 0$ and at $E = 10$.
2. Explain what the interaction coefficient 0.1 measures. Include units.

Use two before-and-after predictions if you are unsure which terms cancel.