

EC203 – Applied Econometrics

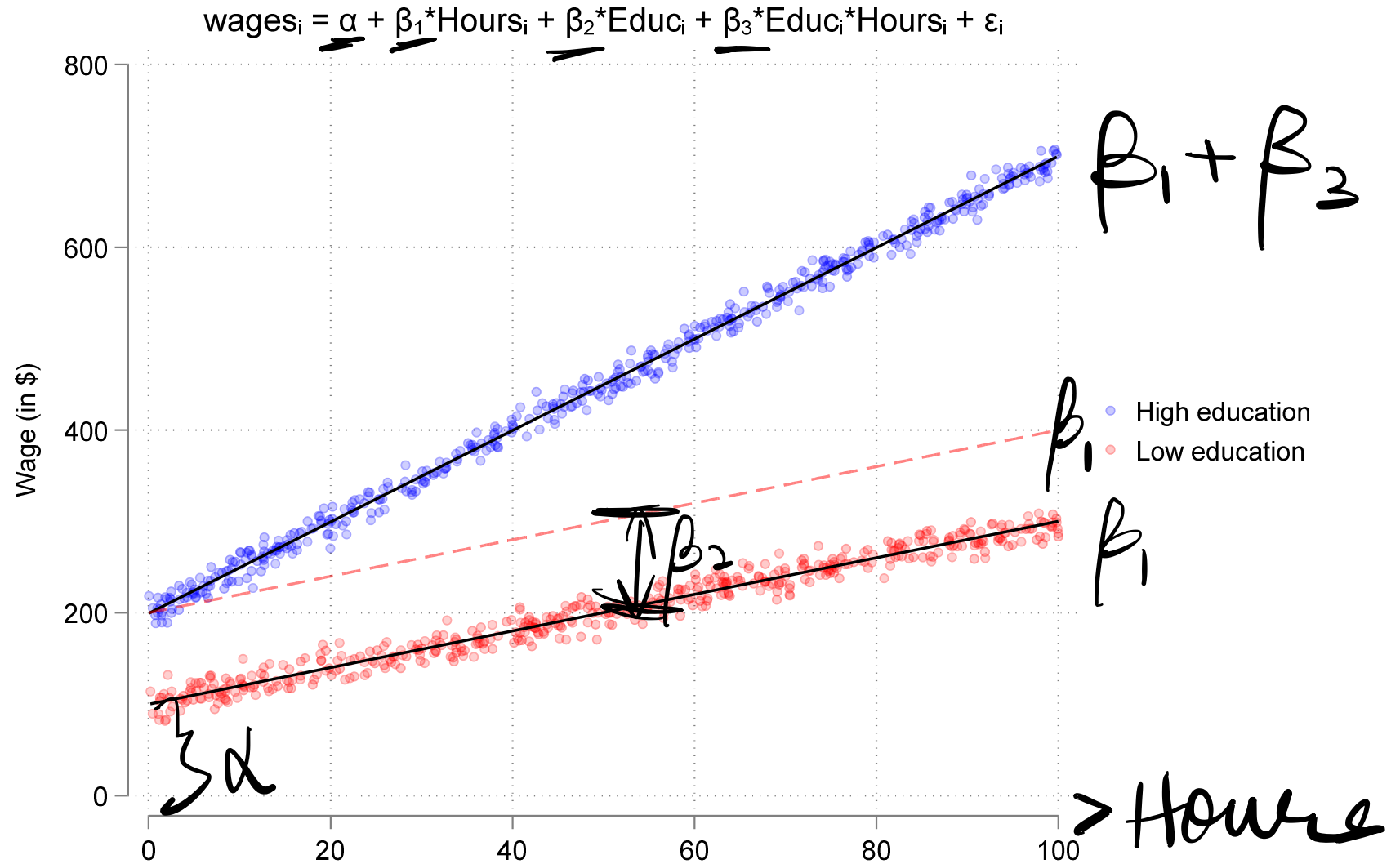
Term 2, Week 6

Sushil Mathew

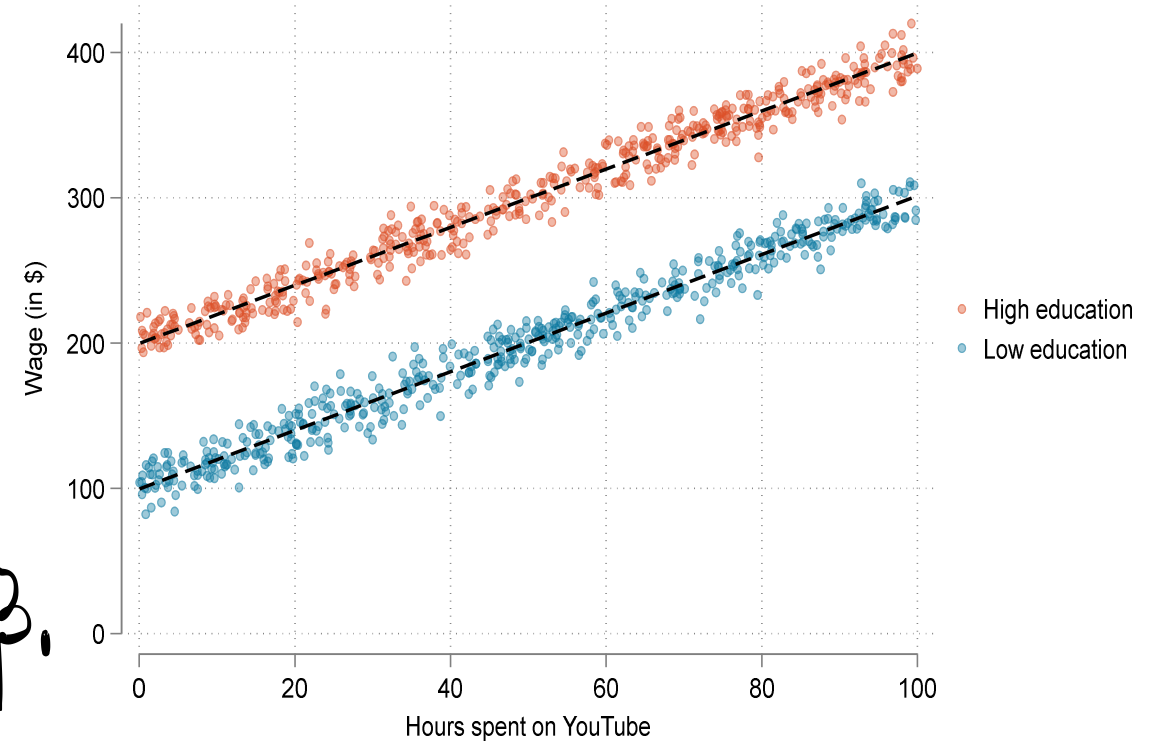
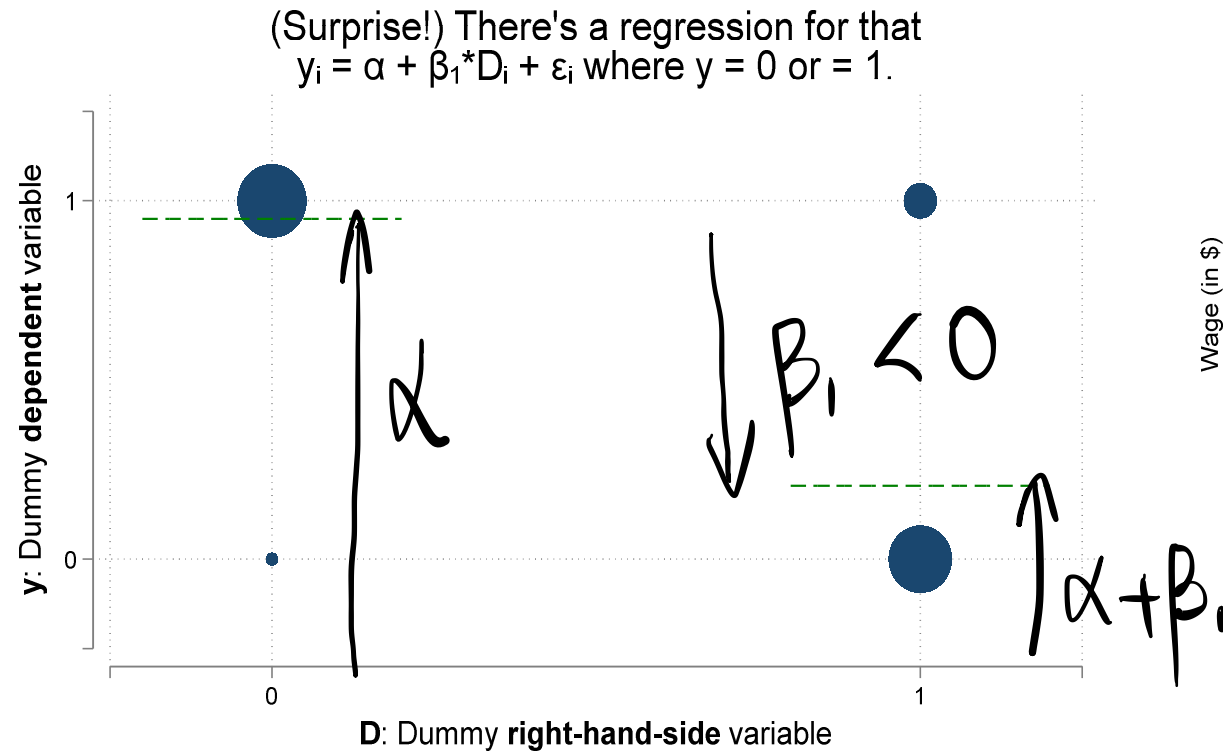


So far...

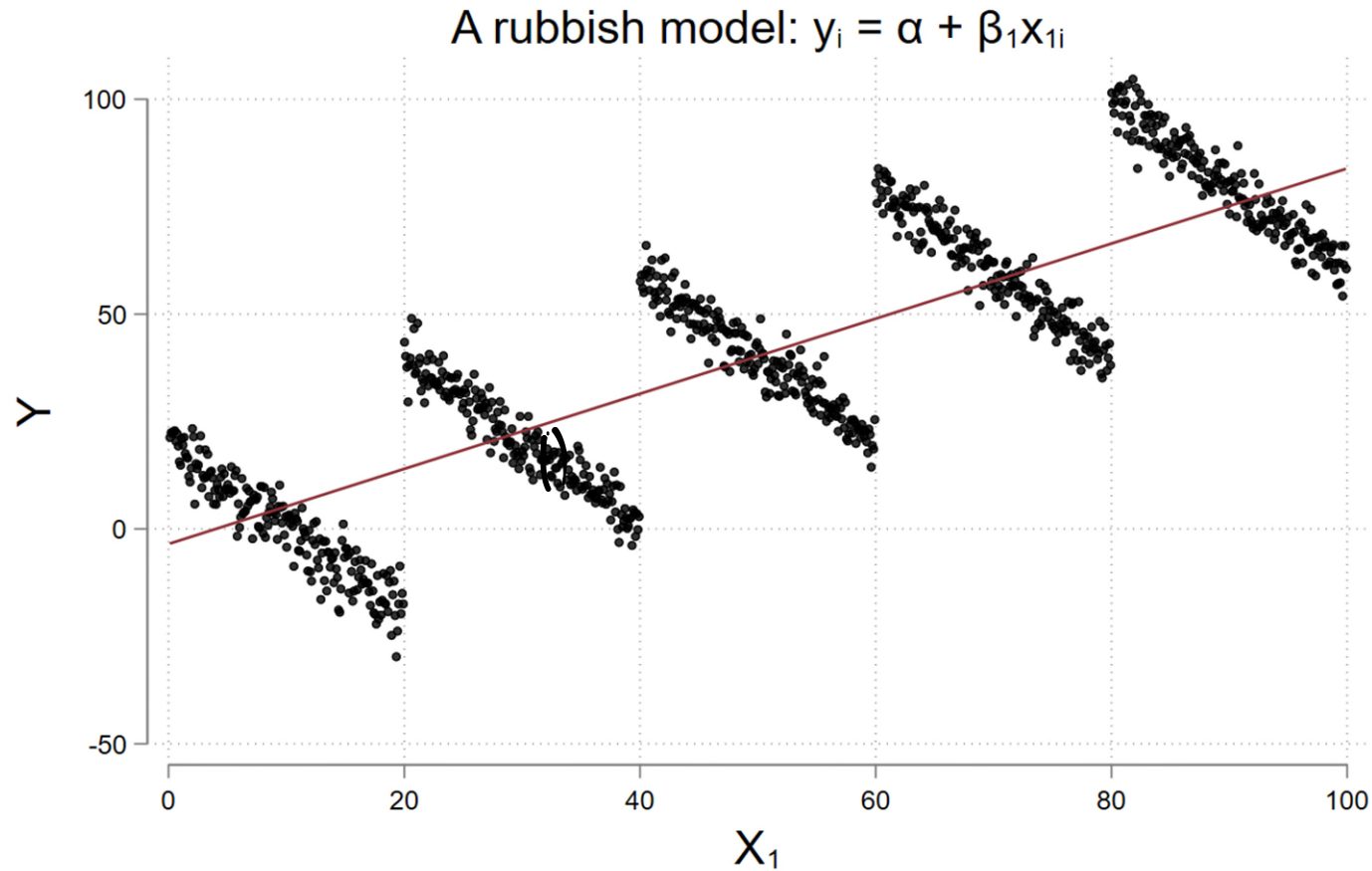
Visual interpretation of coefficients



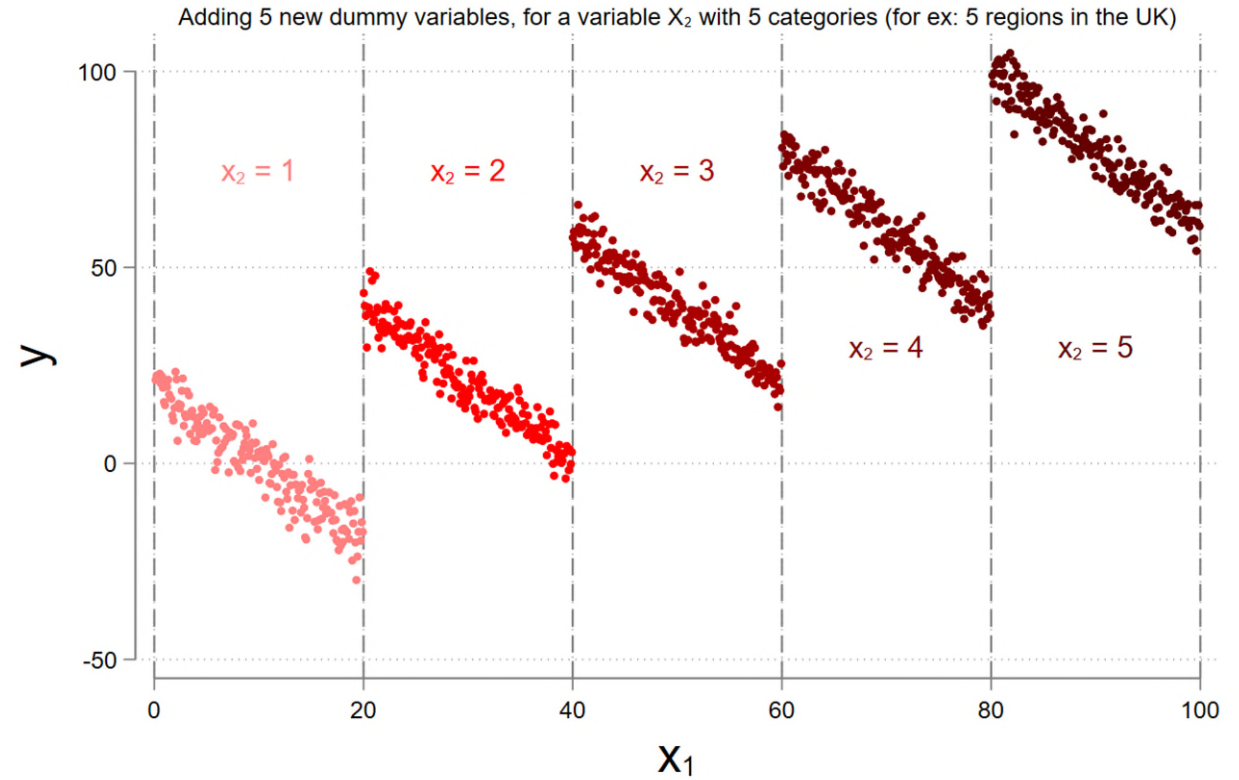
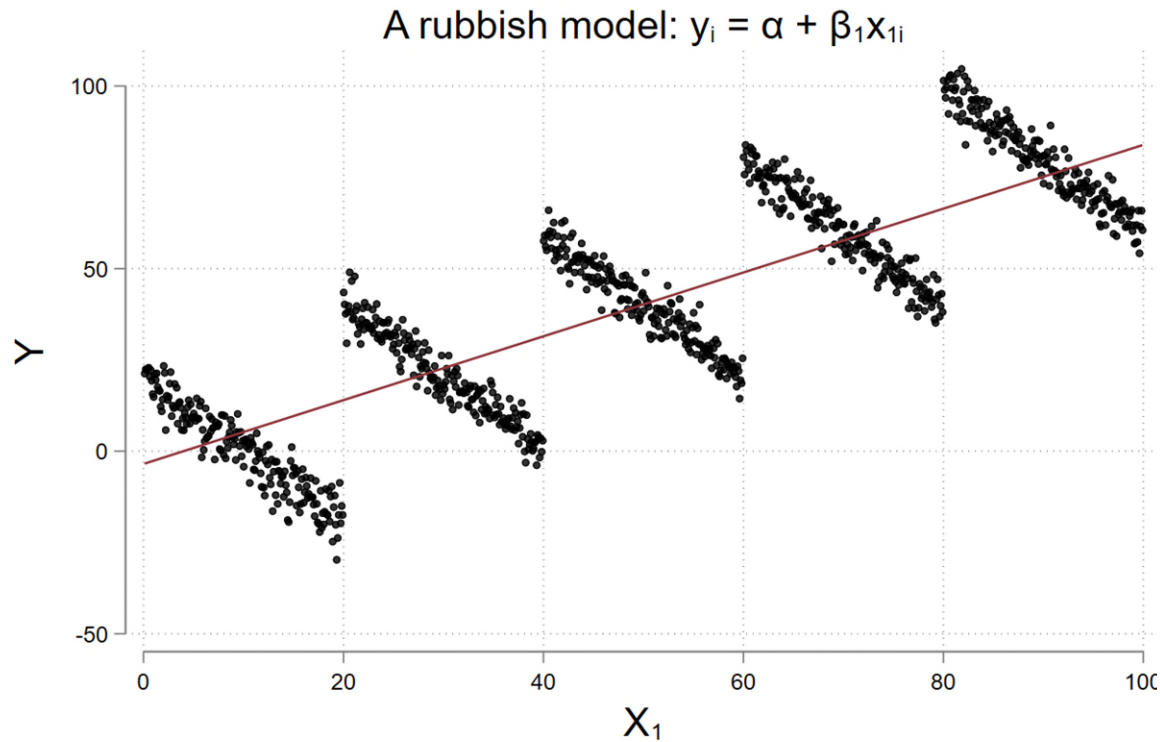
Regressions are not always a line



Adding variables to a regression is good



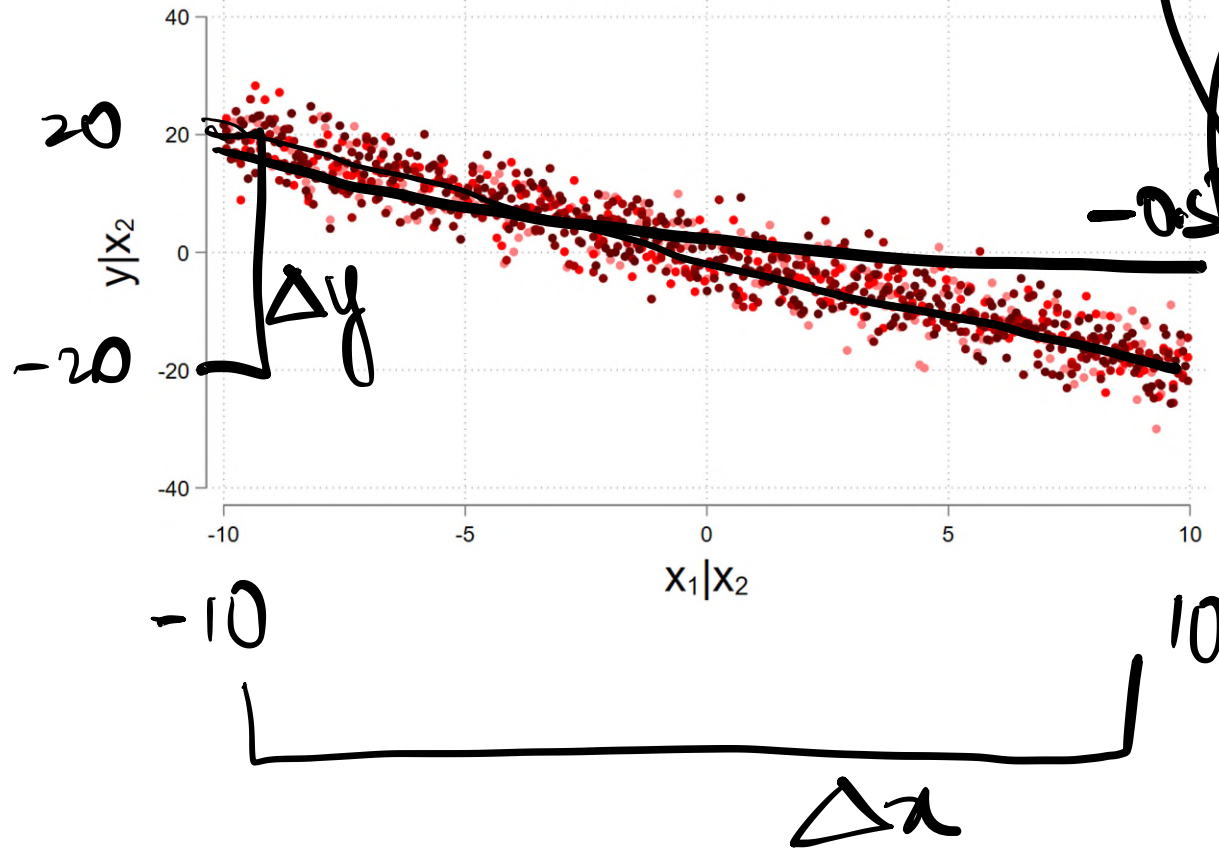
Adding variables to a regression is good



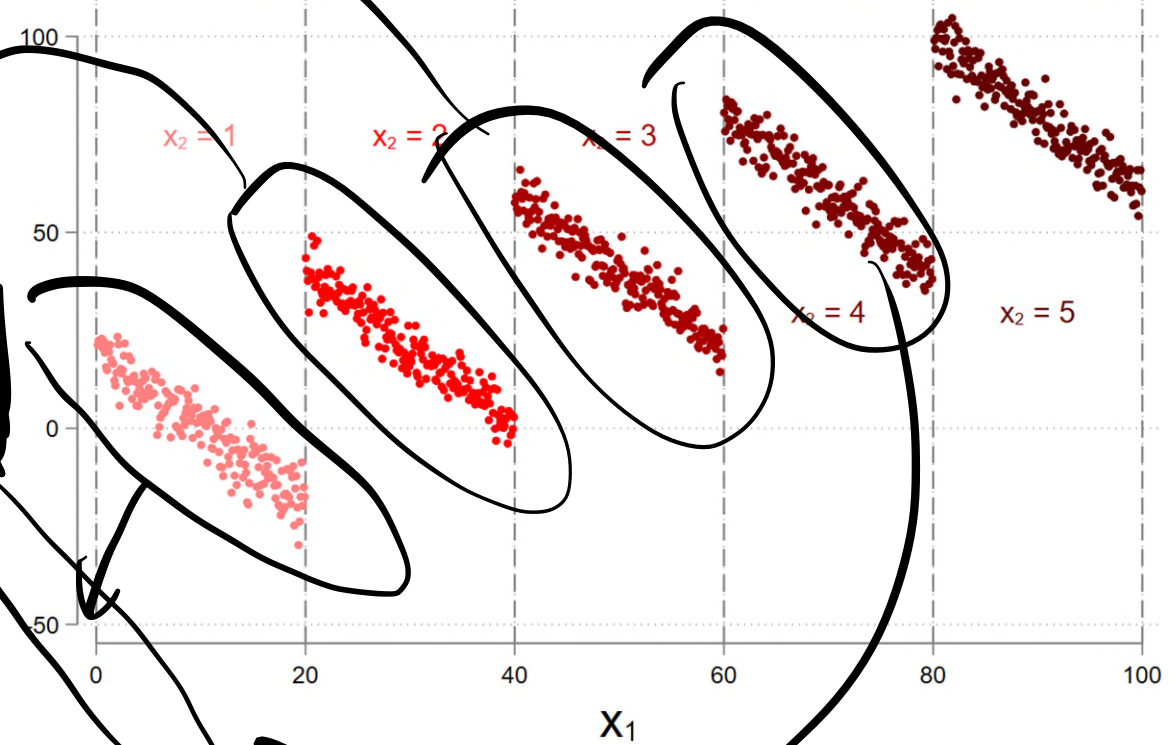
Adding variables to a regression is good

What controlling for a variable looks like visually

Notice the axes scale and colour of dots.
Compare this to Fig2 colours



Adding 5 new dummy variables, for a variable x_2 with 5 categories (for ex: 5 regions in the UK)



$$\frac{\Delta y}{\Delta x} = -2$$

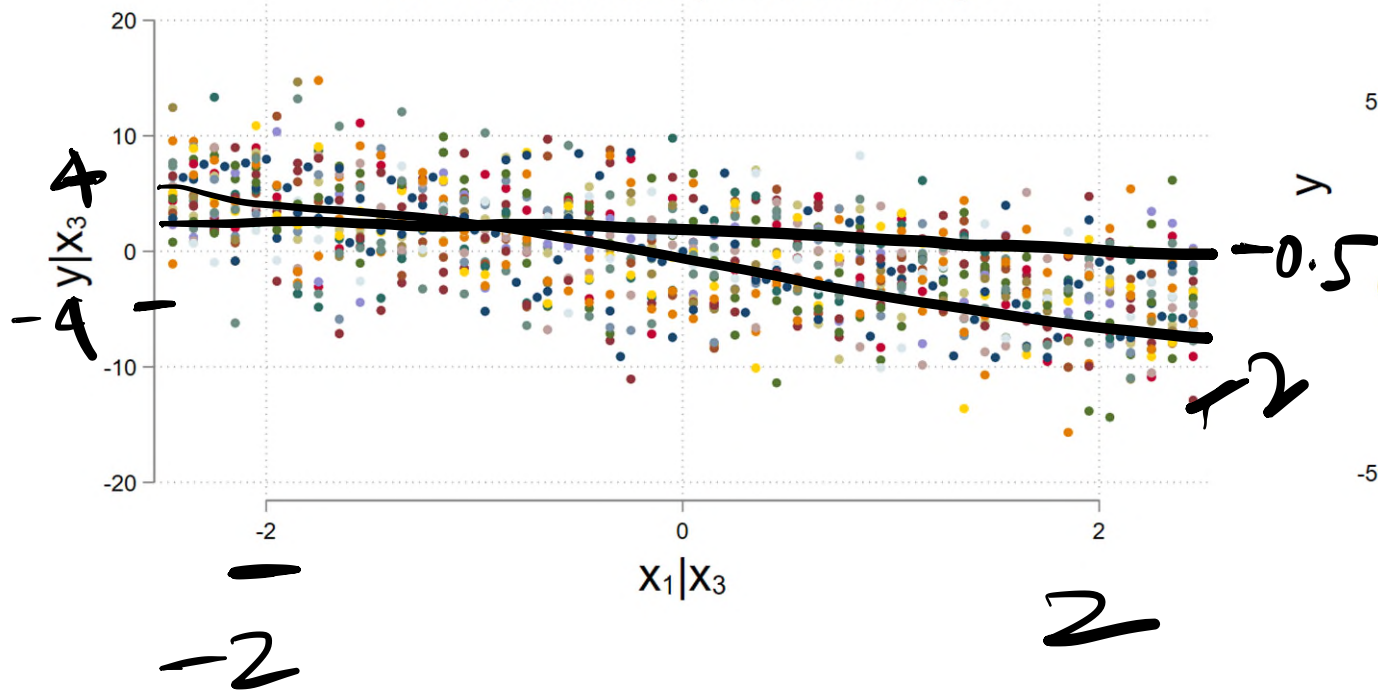
Well...not always

↑ STD. ERRORS

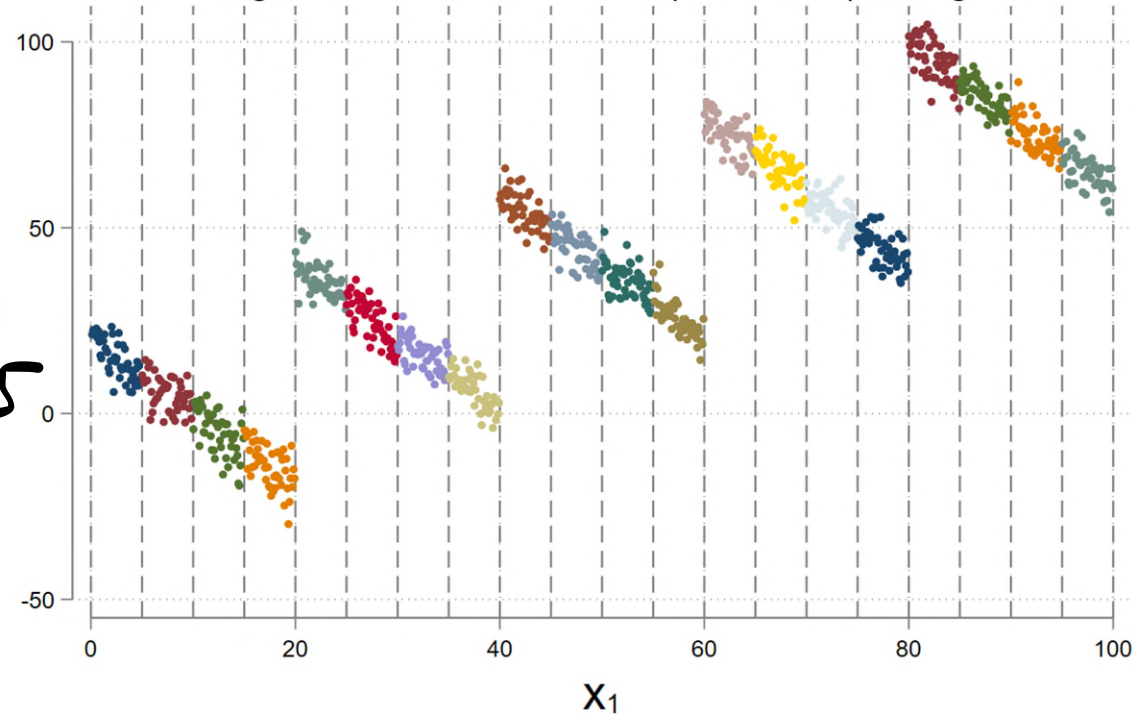


What controlling for an irrelevant variable looks like visually

The slope is still the same
But the uncertainty in the slope has increased
Also compare scale of y and x axis with fig3



Adding a variable X_3 with 20 (irrelevant) categories



From a causal inference perspective

- Single linear regression models are generally bad (**omitted relevant variables or omitted variable bias**)
- Multiple linear regressions are better (fixes omitted variable bias)
- Multiple linear regressions are also bad (addition of irrelevant variables)
- You can't randomize education to study the effect of education on wage.

From a causal inference perspective

- Single linear regression models are generally bad (**omitted relevant variables or omitted variable bias**)
- Multiple linear regressions are better (**fixes omitted variable bias**)
- Multiple linear regressions are also bad (addition of irrelevant variables)
- You can't randomize education to study the effect of education on wage.

From a causal inference perspective

- Single linear regression models are generally bad (**omitted relevant variables or omitted variable bias**)
- Multiple linear regressions are better (**fixes omitted variable bias**)
- Multiple linear regressions are also bad (**addition of irrelevant variables**)
- You can't randomize education to study the effect of education on wage.

From a causal inference perspective

- Single linear regression models are generally bad (**omitted relevant variables or omitted variable bias**)
- Multiple linear regressions are better (**fixes omitted variable bias**)
- Multiple linear regressions are also bad (**addition of irrelevant variables**)
- You can't randomize education to study the effect of education on wage.

From a causal inference perspective

- Single linear regression models are generally bad (**omitted relevant variables or omitted variable bias**)
- Multiple linear regressions are better (**fixes omitted variable bias**)
- Multiple linear regressions are also bad (**addition of irrelevant variables**)
- You can't randomize education to study the effect of education on wage.



From a causal inference perspective

- Single linear regression models are generally bad (**omitted relevant variables or omitted variable bias**)
- Multiple linear regressions are better (**fixes omitted variable bias**)
- Multiple linear regressions are also bad (**addition of irrelevant variables**)
- You can't randomize education to study the effect of education on wage.



Fortunately, we live in a strange world