



Never fear, control variable to the rescue!

$$\hat{\beta}_1 \rightarrow \beta_1 + \rho_{xu} \left( \frac{\sigma_u}{\sigma_x} \right) !!$$

Help us! He's making our  $u_i$  correlated!!

Mwah ha ha.  
Omit me and your  
OLS is ruined!!

Suppose you have the following OLS model:

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i$$

But you think there is an important variable,  $Z$ , still not included.

$$\text{For example: } \text{TestScore}_i = \beta_0 + \beta_1 \text{STR}_i + u_i$$

But you think that  $\text{Parents'Income}_i$  is probably an important determinant of  $\text{Testscore}_i$  too. Unfortunately, you don't have a variable for  $\text{Parents'Income}$ .

Who	What it says	In our example
The Villain: Z <i>An omitted variable</i>	“You’ll never catch me!!” or “Can’t find this variable, and/or it’s hard to measure AND it is contained in $u_i$ and correlated with $X_1$ ”	Parent’s Income (i.e. can’t collect private information, and correlated with STR)
The Evil Henchman: $u_i$ <i>Correlated error</i>	“They’re really screwed now!” Or “ $\hat{\beta}_1$ converges in probability to $\beta_1 + \rho_{xu}(\frac{\sigma_u}{\sigma_X})$ ”	How will the $\hat{\beta}_1$ be biased if Parent’s Income is omitted?! Too high? Too low? Figure it out!
The Innocents: $\hat{\beta}_1$ <i>Effect of <math>X_1</math> you want to estimate</i>	“Help us!” Or “Our error is correlated so our effect is going to be measured with bias!”	The regression will not measure the true effect of STR on Test Score.
The hero: W <i>A control variable</i>	“Blam! Kapow!” Or “Before: $E(u_i   X_{1i}) \neq 0$ , After: $E(u_i   X_{1i}, W_i) = E(u_i   W_i)$ ” “Given the W, the mean of $u_i$ no longer depends on the $X_1$ ”  AND  “No need to thank me citizens!” Or “Don’t try to interpret my $\beta$ as a causal coefficient!”	A variable on the student using the Free Lunch Program (which you only qualify for under certain income limits).