Notes on Econometrics

Lecture 4 addition

Chayanee Chawanote

OLS regression

- OLS >> linear regression: $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi} + \varepsilon_i$
  - β_0: intercept, $E[y_i|$holding all x_i constant$]$
  - β_{pi}: how x_{pi} can explain y_i; when x_{pi} changes 1 unit, y_i will change β_{pi}, holding other x (factors) fixed. It could also say that "β is the slope parameter in the relationship between y and x holding the other factors (both other x and ε) fixed."
  - ε: error term, assume to be random; or we call "unobservable random variable"
  - Let $\hat{y}_i$ be predicted dependent variable after the estimation. $e_i = y_i - \hat{y}_i$ is a residual
- For all β to be unbiased estimator, we need zero conditional mean (exogeneity):
  - $E[\varepsilon_i|x_i] = 0$, or $cov(x_i, \varepsilon_i) = 0$: x is independent from ε
  - This also implies $E(\varepsilon_i) = 0$
  - This assumption is often violated in OLS regression. The problems could be from
    - x has some relationship with the unobservable random variable (ε)
    - omitted variables from the regression
    - x has a reverse relationship with y (y is also determine x)
    - x is endogenous, determined by some factors

- For development economics, we want to evaluate public policy or want to know the effect of x on y, or what causes y to happen. Our goal is to infer that one variable (such as education) has a causal effect on another variable (such as worker productivity).
  - Key condition to ensure that x causes y is to have the exogeneity assumption valid.
  - We could do so by controlling for all possible factors that can explain y so that when we say "holding other factors fixed", β_p that we are interested has a clean effect of x_p on y