Recitation 2

UC Berkeley gender case.

- Berkeley sued for bias against women in 1973.
- Evidence:
    - 44% men admitted
    - 35% women admitted

  Convincing?

Breakdown by department:

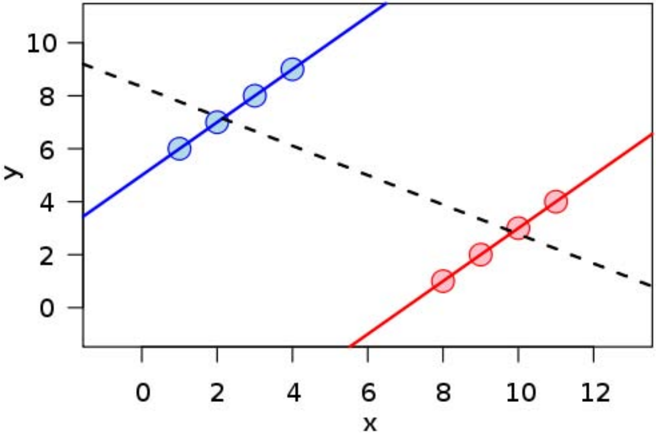
Dept	Male		Female	
	App	Admit	App	Admit
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	272	6%	341	7%

Breakdown by department:

Dept	Male		Female	
	App	Admit	App	Admit
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	272	6%	341	7%

Fixed Effects

More Motivation



Fixed Effects

Bottom line, we might care about *within* variation.

We have a *panel* data set, which consists of n individuals observed for T periods.

$$y_{it} = \alpha_i + \gamma_t + x_{it}\beta_0 + \epsilon_{it}.$$

Interpretations:

- What is α_i?
- What is γ_t?
- What is β_0?

Fixed Effects

$$y_{it} = \alpha_i + x_{it}\beta_0 + \epsilon_{it}.$$

Let us take an average.

$$\underbrace{n^{-1}\sum_t y_{it}}_{:=\bar{y}_i} = \underbrace{n^{-1}\sum_t \alpha_i}_{:=??} + \underbrace{n^{-1}\sum_t x_{it}}_{:=\bar{x}_i}\beta_0 + \underbrace{n^{-1}\sum_t \epsilon_{it}}_{:=\bar{\epsilon}_i}.$$

S
o we have

$$\bar{y}_i = \alpha_i + \bar{x}_i\beta_0 + \bar{\epsilon}_i.$$

Fixed Effects

Target equation:

$$y_{it} = \alpha_i + x_{it}\beta_0 + \epsilon_{it}.$$

Average equation:

$$\bar{y}_i = \alpha_i + \bar{x}_i\beta_0 + \bar{\epsilon}_i.$$

Let us subtract the second from the first:

$$\underbrace{y_{it} - \bar{y}_i}_{interp??} = \underbrace{\alpha_i - \alpha_i}_{=??} + \underbrace{x_{it}\beta_0 - \bar{x}_i\beta_0}_{interp??} + \underbrace{\epsilon_{it} - \bar{\epsilon}_i}_{interp??}.$$

Fixed Effects

Bottom line:

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\,\beta_0 + (\epsilon_{it} - \bar{\epsilon}_i)\,.$$

Reminder - Potential Outcomes

Two universes:

- $Y_i(0)$: outcome for person i without treatment
- $Y_i(1)$: outcome for person i with treatment

W
hat do we want to study?

$$Y_i(1) - Y_i(0).$$

POTENTIAL OUTCOMES

Let $W_i \in \{0,1\}$ be whether the treatment was received. What is the observed outcome?

$$Y_i = (1 - W_i) \cdot Y_i(0) + W_i \cdot Y_i(1).$$

Covariates

- We also have access to variables X_i and Z_i which have not been affected by the treatment.
- In particular, X_i will be important for the RD design.
- We observe:
$$(Y_i, W_i, X_i, Z_i)\,.$$
- Basic idea is that the assignment to the treatment is going to be determined fully or partially by the value of a predictor (the covariate X_i).

- We call X the forcing variable or the treatment-determining variable:

$$W_i = \mathbf{1}\{X_i \geq c\}.$$

- Interpret

$$\tau_{SRD} = \mathrm{E}\left[Y_i(1) - Y_i(0) | X_i = c\right]$$

- We can estimate

$$\lim_{x \downarrow c} \mathrm{E}\left[Y_i | X_i = x\right] - \lim_{x \uparrow c} \mathrm{E}\left[Y_i | X_i = x\right].$$

SHARP RD - ASSUMPTIONS

(1) $Y_i(0), Y_i(1) \perp W_i | X_i$ - does it hold?

- Trivially, but X fully determines W and there is no variation.

(2) Continuity of conditional expectations in x

$$\mathrm{E}\left[Y(0)|X=x\right] \text{ and } \mathrm{E}\left[Y(1)|X=x\right].$$

- Notice, this method requires extrapolation. Why?

Sharp RD - Assumptions

- Observe that

$$\lim_{x\downarrow c} \mathrm{E}\left[Y(0)|X=x\right] = \mathrm{E}\left[Y(0)|X=c\right] = \mathrm{E}\left[Y(1)|X=x\right] = \lim_{x\uparrow c} \mathrm{E}\left[Y(1)|X=x\right]$$

- Therefore the average treatment effect at c, τ_{SRD}, is

$$\tau_{SRD} = \lim_{x\downarrow c} \mathrm{E}\left[Y|X=x\right] - \lim_{x\uparrow c} \mathrm{E}\left[Y|X=x\right].$$

Fuzzy RD

- Imagine the probability of treatment changes as we cross c.
    - Examples?
- Idea:
    - Variation in $X \implies$ variation in $W \implies$ variation in Y.
    - Looks familiar?

Fuzzy RD

- We then estimate

$$\tau_{FRD} = \frac{\lim_{x\downarrow c} \mathrm{E}\left[Y|X=x\right] - \lim_{x\uparrow c} \mathrm{E}\left[Y|X=x\right]}{\lim_{x\downarrow c} \mathrm{E}\left[W|X=x\right] - \lim_{x\uparrow c} \mathrm{E}\left[W|X=x\right]}$$

- Assumptions:
  - $W_i(x)$ is non-increasing in x at $x = c$.
  - Compliers:

$$\lim_{x\downarrow X_i} W_i(x) = 0 \text{ and } \lim_{x\uparrow X_i} W_i(x) = 1$$

  - Nevertakers:

$$\lim_{x\downarrow X_i} W_i(x) = 0 \text{ and } \lim_{x\uparrow X_i} W_i(x) = 0$$

  - Alwaystakers:

$$\lim_{x\downarrow X_i} W_i(x) = 1 \text{ and } \lim_{x\uparrow X_i} W_i(x) = 1$$

Then

$$\tau_{FRD} = \mathrm{E}\left[Y_i(1) - Y_i(0)|i \text{ is a complier and } X_i = c\right].$$

14.75 Political Economy and Economic Development
Fall 2012