ECONOMICS 174

Global Poverty and Impact Evaluation

LECTURE 3

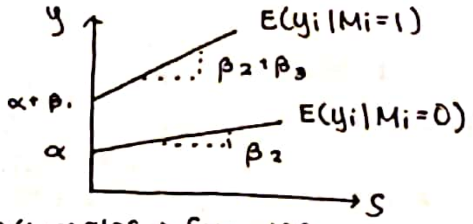
Linear Regression: Interaction terms

$y_i = \alpha + \beta_1 M_i + \beta_2 S_i + \beta_3 (M_i \cdot S_i)$

$E(y_i|M_i=1) = \alpha + \beta_1 + \beta_2 E(S_i|M_i=1) + \beta_3 E(S_i|M_i=1)$

$= \alpha + \beta_1 + (\beta_2 + \beta_3) E(S_i|M_i=1)$

$E(y_i|M_i=0) = \alpha + \beta_2 E(S_i|M_i=0)$



interpretation: β_1 is the diff in ave. wages when $S_i = 0$

β_3 is the diff in 2 slopes/returns to educ b/t males + females

why does education matter for development?

→ Millenium Development Goals (MDG): one is universal primary enrollment
→ Education matters for growth: production of goods + services
 ↳ perfect substitutes: educ workers are more efficient
 ↳ imperfect substitutes: educ workers necessary for producing certain goods
 ↳ tech adoption + productivity enhancing
→ Non-pecuniary returns
 ↳ intra-household effects
  → women's educ ⇒ lower fertility ⇒ improved child health + education
  → each additional year of schooling is associated w/5-10% drop in infant mort
  → inter-generational correlation for educ is 0.6
 ↳ political: improved political climate + more civic participation; less asymm info
 ↳ social: less violence + crime rates; more social mobility
 ↳ culture: openness to new ideas
 ↳ positive externalities
→ economic growth eq involves labor, capital, + tech
→ externalities are usually primary justification for gov't intervention

Modeling School Choice

General model
→ 2 periods where $t \in \{1,2\}$
→ instantaneous utility: V
→ discount factor: β
→ chooses to study or not (S)
→ opp cost of going to school: C
 ↳ includes direct costs
→ return to schooling: B
 ↳ receive w/probability p

$EU(S=0) = V + \beta V$

$EU(S=1) = V - C + \beta[p(V+B) + (1-p)V]$

$= V - C + \beta(V + pB)$

will go to school if $EU(S=0) \leq EU(S=1)$
so $\beta p B \geq C$ (benefits must outweigh the costs)

* decision depends on discount factor, probability of benefit, size of benefit, cost

But in developing countries, there is violence, corruption, no avail jobs, + social stratif

why are there diff levels of education?

1. Difference in ability
 → assume all individuals have the same C, β, + p
 → differ in ability = diff. in benefits: $B_i \in [0, B_{max}]$ *heterogeneity
 → if B^* is such that $\beta p B^* = C$ then when $B > B^*$, they will study
2. Difference in cost
 → individuals w/cost $C < C^*$ will go to school
3. Credit constraints
 → assume minimum standard of living: $\bar{V}$ where $V > \bar{V} > V - C$
  ↳ need credit market; can't afford cost w/o borrowing money
 → Borrowing:
  $t=1: \bar{V}$
  $t=2: \beta[V + pB - (C - (V - \bar{V}))]$
 → Not Borrowing
  $t=1: V$
  $t=2: \beta V$
 → individ. in extreme cannot invest in school even if returns are positive

* All models have assumption of optimally maximizing utility
 ↳ but this does not mean their decisions are optimal

Decision to go to school: Positive Externality

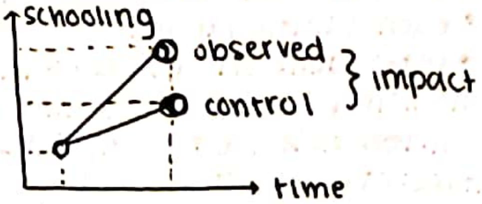
		Student 1 school	Student 1 no school	
Student 2	school	3,3	1,4	
	no school	4,1	2,2	← equilibrium but not efficient

What should policymakers do?
1. Reduce cost by subsidizing schooling
2. Reduce credit constraints for optimal investments
3. Incr. probability benefits are realized (improve labor market outcomes)

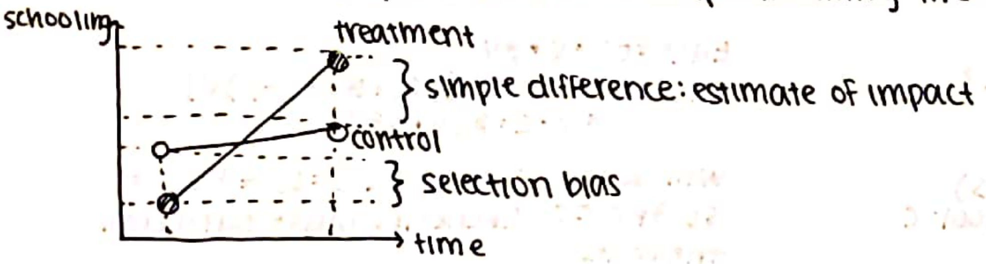
LECTURE 4

Observing Impact with the Counterfactual

How do we measure impact of a policy?
- → other factors like gov't intervention or weather may create impacts
- → can have prepost designs: data observed before + after an event (prices of stock)
  - ↳ can infer causality if ex: Apple announces new phone + prices ↑ w/n a few sec
- → need a counterfactual, what would have happened in the absence of a program
  - ↳ issue: unobserved
  - ↳ control group: data on similar group of non-beneficiaries



- → selection bias: diff between groups of participants + non-participants (background)
  - ↳ challenge: hard to randomize b/c gov't has incentive to give to baseline group so cannot randomize
  - ↳ can also have spillover effects (end up combining the two)



Potential Outcomes

We have Y_i^T - schooling of student i if he receives the tuition waiver

Y_i^C - ______________ if he does not receive the tuition waiver

We want to know $Y_i^T - Y_i^C$ or $E(Y_i^T - Y_i^C)$

We have $Y_i = (1-W_i)Y_i^C + W_i Y_i^T$ where W_i is an indicator of whether child i is assigned the tuition waiver
- → So we have Difference $= E(Y_i^T | W=1) - E(Y_i^C | W=0)$
- → By adding + subtracting $E(Y_i^C | W=1)$ we get:

$$D = [E(Y_i^T | W=1) - E(Y_i^C | W=1)] + [E(Y_i^C | W=1) - E(Y_i^C | W=0)]$$

- * $E(Y_i^C | W=1)$ are those assigned to receive treatment but was not treated
- * $E(Y_i^T | W=1) - E(Y_i^C | W=1)$: treatment effect on treated - treatment effect for those that participated in the program
- * $E(Y_i^C | W=1) - E(Y_i^C | W=0)$: selection bias
  - ↳ diff. b/t individ. who would have received waivers + those that did not in absence of the treatment

The Gold Standard

Randomization solves the selection bias; if random, Y_i will be independent of W_i

$D = [E(Y_i^T|W=1) - E(Y_i^C|W=1)] + [E(Y_i^C|W=1) - E(Y_i^C|W=0)]$
$D = TOT + [E(Y_i^C|W=1) - E(Y_i^C|W=0)]$
$D = TOT + [E(Y_i^C) - E(Y_i^C)]$
$D = TOT = ATE$

* TOT: treatment of treated
* ATE: average treatment effect

Gold standard: randomization assures us that in the absence of treatment, children that would have received the waivers are the same in expectation to those that would have not received it
→ solves the missing data problem: can compute ave. effect for the population w/o observing individ. counterfactuals

Regression Analysis:
outcome of children that did not receive treatment: $Y_i^C = \alpha + \epsilon_i$
→ α = average amount (parameter)
→ ϵ = amount that varies by person
If treatment increases schooling by δ: $Y_i^T = Y_i^C + \delta$
→ If $Y_i = (1-W_i)Y_i^C + W_i Y_i^T$ then $Y_i = (1-W_i)Y_i^C + W_i(Y_i^C + \delta)$
$Y_i = Y_i^C + \delta W_i$
$Y_i = \alpha + \delta W_i + \epsilon_i$

* just regress Y on W
* is $E(\epsilon_i|W) = 0$ ok?

Benefits to randomization: 1) no machinery or techniques; just compare means
2) everyone can understand the results
3) internal validity is high

Internal vs. External Validity
→ Internal validity: est impact of the program net of all other potential confounding factors, or that the comparison group represents the true impact of the program (random assignment guarantees this)
↳ make causal inference on pop. we have

→ External validity: the impact est. in the evaluation sample can be generalized to the pop.
↳ take out result + extrapolate to whole population
↳ evaluation sample must be representative of pop. of eligible units (random sampling guarantees this)

treatment group
control group
evaluation sample ← population of eligible units
*random selection preserves characteristics

Note: percentage points v. percent

LECTURE 5

Randomized Control Trials (RCT)

Progresa (Opportunidades): began in 1997 to present; began in rural areas + then extended to semi-urban areas; transfers to over 5.8 bil/mil households across Mexico
→ welfare program provides monthly stipend to poor households conditional on children's attendance
↳ conditional-cash programs believed to be silver bullet for develop. economics
→ but not costless so why?
↳ not costless but gives direct outcome policymakers want
↳ bargaining model: incr. power in household
→ more valuable commodity; invest more to go to school
↳ break cycle of poverty in short-run (stipend) + long run (educ)
→ distinguishing feature: included evaluation component at inception
↳ need political courage to have control group

Timeline:

1997: Baseline survey (6–8 hr about income, assets, migr. patterns, etc) to create welfare index to see eligibility
→ treatment assignment

1998: Wave 1

1999: Wave 2

2000: lost control group when incorp (political pressures)

* attrition: cannot survey person in t+1 period due to inability, migration, receiving treatment

Payments incr. by grade level + higher for females
→ higher female payments due to
↳ gender imbalances: bargaining power
↳ females get paid higher in labor force (but more males in labor force)

Eligibility
→ Two-step targetting procedure
1. Poor rural communities were selected on the basis of a marginality index
2. Within these comm, household census was conducted to identify which households were poor (welfare index of assets which was secret to prevent distortions)

→ 50,000 comm. selected + 78% of households w/n comm eligible
↳ household level

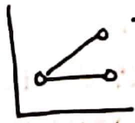
Research Design

Treatment	Control
non-poor	non-poor
poor (beneficiary)	Poor

welfare threshold

P=1 if poor
W=1 if treatment vil
T=1 if post 1997

1. $E(sch|P=1,W=1,T=1) - E(sch|P=1,W=1,T=0)$



} treatment

counterfactual: sch same in 1998 as 1997 (no ↑ or ↓ trend)
→ prepost design: assume no other factors affect enrollment
$sch = \alpha + \beta T + \varepsilon$ (sample: poor + treated)
reg sch T if p==1 + W==1

2. $E(sch|P=1,W=1,T=1) - E(sch|P=1,W=0,T=1)$
→ counterfactual: in absence of treatment, enrollment of poor in treatment = enrollment of poor in control
$sch = \alpha + \beta W + \varepsilon$ (sample: poor in 1998)
reg sch W if p==1 + T==1, cluster(village) } variation across treatment + control
→ cluster adjust SE (nonzero elements of off diagnol)
→ cluster units @ level of randomization if data collected in smaller unit
↳ household w/n villages are not indep / random

3. $[E(sch|P=1,W=1,T=1) - E(sch|P=1,W=1,T=0)] - [E(sch|P=1,W=0,T=1) - E(sch|P=1,W=0,T=0)]$
→ (enrollment of beneficiaries in 1998 less 1997) − (enroll. of control in 1998 less 1997)
↳ double diff across time
→ counterfactual: no diff b/t control + treatment group, no spillovers, no selection bias
↳ net out diff before program started
→ even if randomized in theory, may not see perf balance = initial diff
↳ sometimes have better SE

4. $[E(sch|P=1,T=1,W=1) - E(sch|P=1,W=0,T=1)] - [E(sch|P=0,W=1,T=1) - E(sch|P=0,W=0,T=1)]$
→ if everything works, 2=3=4
↳ reasons for not working:
→ spillovers in same comm
↳ nonpoor also being treated so no counterfactual

$sch = \alpha + \beta_1 P + \beta_2 W + \beta_3 (PW) + \varepsilon$ if T=1

$x = \frac{100}{1+r}$

$PV(F_t) = \Sigma_{t=0}^{n} \frac{F_t}{(1+r)^t}$

* 2000: people may Δ behavior when adding control
* Biggest impact in enrollment: 6th gr (hurdle into secondary); less observ going down
→ Schultz: Progresa rate of return = 8% per year
↳ incr schooling by 0.66 years + 12% additional return of schooling per year
→ Elasticity of demand low @ 0.2 (reduced cost by more ½ but ↑ edu. attainment by 10%)

LECTURE 6

Cost-Benefit Analysis

Present value of future value F_t obtained t years from now

$PV(F_t) = \sum_{t=0}^{n} \frac{F_t}{(1+r)^t}$ where r is the internal rate of return

Schultz est 8% return rate for Progresa
- → incr. schooling by 0.66 years; 12% return to educ. each additional year
- → elasticity of demand is low: 0.2
  - ↳ reduce cost of attending school by more than half but increased educ. attainment by only 10%
- → If externalities, Schultz may underest. effects

Spillovers

What outcomes do we test for?
- → school attendance of non-treated, crowding out, academic performance
- → income (more money means raising prices), incr. in prices
- → nutrition/health/consumption from exchange of food
  - ↳ risk sharing in villages: mutual exchange in bad times that allow consumption levels to be above sustenance

How do we test for any spillover effects?
- → within village:
  - ↳ $E(sc|T=1, P=0, W=1) - E(sc|T=0, P=0, \underline{W=1}) + E(sc|T=1, W=0, P=0) + E(sc|T=0, P=0, W=0)$
    - → test before + after but does not exploit randomization (first part)
  - ↳ $E(sc|T=1, P=0, W=1) - E(sc|T=1, P=0, W=0) = \delta$
    - → simple diff: unbiased causal est/measure in treatment + those w/o treatment

Post

	T	C
N	A	B
P		

Pre

	T	C
N	C	D
P		

  - ↳ Double difference: $\delta - [E(sc|T=0, P=0, W=1) - E(sc|T=0, P=0, W=0)]$
    - → subtracting pre-existing diff: $(A-B)-(C-D)$

How to test how it changes over time?
- → ask about graphs in slide 9 + chart in slide 10

Testing across villages
- → measure of D (distance to treatment village) + interact this term

  $SC = \alpha + \beta_1 W + \beta_2 D + \beta_3 WD + \varepsilon$

Bobonis + Finan (2008)
- → Welfare index: Δ across regions (Panel A + C is poor / B + D is nonpoor)
  - ↳ double diff computing diff in Panel B + D
  - ↳ nonpoor in post period, treatment enrollment rates is high
    - → concentrated in lower region of Welfare Index
- → Slide 10: effect on schooling on W (regression counterpart in 9)

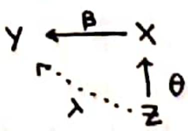
Instrumental Variables

If we have $y_i = \beta_0 + \beta_1 x_i + u_i$ but $corr(x, u) \neq 0$
- → instrumental variables: correlated w/ x but uncorr. w/ u

$$\hat{\beta}_{IV} = \frac{\sum z_i y_i}{\sum z_i x_i} = \frac{cov(z,y)}{cov(z,x)}$$

$$= \frac{cov(\alpha + \beta x + u, z)}{cov(x,z)}$$

$$= \frac{cov(\alpha,z) + \beta cov(x,z) + cov(u,z)}{cov(x,z)} = \beta$$



$E(\hat{\beta}_{IV}) = \frac{\beta\theta}{\theta} = \beta$

$E(\hat{\beta}_{IV}) = \frac{\theta\beta + \lambda}{\theta}$

* assume $cov(u,z) = 0$

Peer effects: behavior determined by average behavior of others in village, $\bar{y}_{-i}$

$y_i = \alpha + \beta \bar{y}_{-i} + \delta x_i + \epsilon_i$

imitation
↳ but other explanations: carpooling, transp. info on educ
- → issue of reference group
- → causally estimate β? (selection + omitted factors)

LECTURE 7

Spillover effects continued...

Research design of Progresa provides a clean test for spillovers by computing $E(sc|P=0, W=1, T=1) - E(sc|P=0, W=0, T=1)$
- → double diff: subtract above equation from pre-treatment difference
- → Bobonis + Finan: divergence in Panel D is evidence of spillover effects

Testing peer effects: $y_i = \alpha + \beta\bar{y}_{-i} + \delta x_i + \epsilon_i$
- → y_i = schooling outcomes of student i (nonpoor in treatment village)
- → $\bar{y}_i$ = average schooling of everyone in the village
  - ↳ assumption: everyone has the same reference group + averages affect others
    - → peer group can be just family or people outside the village
    - → can be measured w/a lot of error
  - ↳ mapping out social network: "who are your 5 closest friends?"
  - ↳ can we causally estimate β?
    - → selection: for ex, good students may hang out together
    - → omitted variable bias: parent's educ? income?
    - → simultaneity: both observing + doing the action
- → To identify β, we can use randomization + instrumental variables
  - ↳ instrument: z corr w/x but uncorr w/u (exclusion restriction: not testable)
    - → cannot run y on z + x b/c z + x are highly correlated
  - ↳ instrument: whether it is a treatment village
    - → uncorr w/error because of randomization
- → But other effects of Progresa (channels)?
  - ↳ perception of schooling, better transport methods, externalities in health, less teacher absenteeism, higher consumption

Conditional Cash Transfers vs. Unconditional Cash Transfers

CCT programs are transfers conditional on certain behaviors of recipient households
- → due to de jure + defacto conditionality, we assume CCT > UCT due to distortions of behavior
  1. income effect: household income goes up
  2. price effect: reduces price of activity (CCT) *paternalistic
- → marginal: does not affect welfare
- → do not internalize positive externalities (direct policy) + politically appealing

CT: welfare is higher w/lower admin cost — but lower program impact?

RCT: Zomba Project
- → background: unmarried school girls ages 13-22; monthly stipend of \$5
  - ↳ 2 treatment arms randomly assigned across 88 districts

$y = \alpha + \beta(CT) + \beta_2(CCT) + \epsilon_i$

mean of the control group (α); indicator for cash transfer (CT); indicator if person is assigned CCT (CCT)

- → in self reported estimates, looks like UCT > CCT
- → teacher reported estimates, CCT > UCT
  - ↳ less sample size because taken from a couple years removed
  - ↳ by 2009, all term diff are significant
- → UCT decreased marriage + teen pregnancy rates
  - ↳ marriage rates much higher in unenrolled
    - → CCT same as control b/c no money if you do not attend school
    - → wealthier you are, less likely to get married
- → Baird et al (2011): experimental evidence in support of CCT
  - ↳ to achieve same impact
    - → conditional: \$5
    - → unconditional: \$10

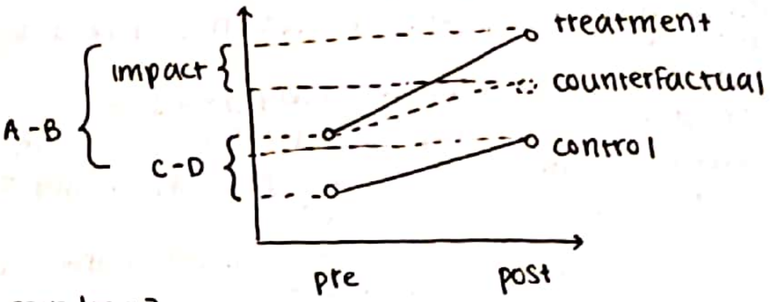
LECTURE 8

Difference-in-Differences

	T	C
post	A	B
pre	C	D

DiD: (A−C)−(B−D)

counterfactual: in absence of treatment, the Δ in outcomes in the treatment group (A−C) = Δ in outcomes in control (B−D)



Supply-Side Policies in Education

Policies that promote school enrollment may not promote learning
- → lack of imputs is an impediment to learning (ex: lower teaching days overest. by absenteeism)
- → do students perform poorly due to low spending?
  - ↳ large variation but may not be accurate due to corruption
  - ↳ diff pop + context — may need years of education to see Δ

Glewwe (2011): Meta-Study
- → looks at diff papers on test scores: mixed results
- → How should we interpret this?
  1. Inputs do not matter?
  2. Interventions are context specific.

* Valanities Rule: cannot have >40 students

Education Production Function: $H_{ij} = F(X_{ij}, X^*_{ij}, X_{-ij}, X^*_{-ij}, Z_j, Z_j^*)$
- → student characteristics: X_{ij}, X_{ij}^*
- → peer characteristics: X_{-ij}, X^*_{-ij}
- → school (classroom) characteristics: Z_j, Z_j^*

} * denotes unobservables

  - ↳ important as levers because easier to adjust
- → Linearity + additivity: $H_{ij} = \alpha_1 X_{ij} + \alpha_2 X_{-ij} + \alpha_3 Z_j + u_{ij}$
  - ↳ u_{ij}: all * (unobservable) characteristics
  - ↳ correl. does not imply causation because $E(u|X,Z)=0$ is hard to prove

Glewwe, Kremer, Moulin (2009): Textbook Experiment
- → in 1996, 100 schools (random) divide into 4 groups (25) based on alphabetical order
  - ↳ students grades 3-7 receive English textbooks; students grade 3,5,7 = math textb.
  - ↳ Kenyan textbook cost \$2-3 (GDP per capita in 1997 is \$330)
  - ↳ in 1997, another 25 schools given block grants = \$2.65

1995: Baseline → 1996: Group 1 receives textbooks → 1997: Group 2 schools receive grants → 1998: Group 3 schools receive grants → 2000: Group 4 schools receive grants

- → Roll-out system: political structure (give to everyone) or monetary constraints
- → quasi-random: those w/ last name in beginning of alpha tend to have more educ.
- → comparison 1: At end of 1996, compare Group 1 to Group 2,3,4
  - ↳ Was impact in yr 1 same as yr 2?
- → comparison 2: At end of 1997, compare Group 1 to Group 3,4
- → Comparison 3: Compare Group 1 + Group 2 measured at yr 2
  - ↳ difference b/t receiving grants + textbooks
- → To test if research design was truly random: (1) t-test or (2) regress one characteristic on treatment dummy: $y = \alpha + \beta T + \varepsilon \Rightarrow E(y|T=1) = \alpha + \beta \Rightarrow E(y|T=0) = \alpha$

Slide 23: make sure treatment worked (first-stage did what they wanted to do)
Normalized test score: in std deviation units

Attrition: subjects disappear from sample
- → If correlated w/ treatment, affects internal validity of research design

Results: impact only smartest students b/c textbooks in English
- → curriculum poorly matched to student needs

LECTURE 9

Primary Education in India

SSA (Education for all movement) 2000: provide educ to all children 6-14 by 2010
- → focus on inputs: school construction, investment in sch. infrastructure, more teachers, materials, etc.
- → cost since 2005-06: $18b US

Right to Education Act 2010: all children 6-14 need to be provided w/8 yrs of educ. compulsorily under force of law
- → Nearly US $80b to be spent over next 5 years

Net primary enrollment risen 80% to 95% in less than a decade
- → present an enrollment @ 96% gives opportunity to focus on quality instead of quantity
  - ↳ India is shining + drowning
    - → 5th largest number of children in world meeting advanced TIMMS benchmark
    - → largest number not meeting a low benchmark
    - } due to large pop.
    - → diff between top 5% + bottom 5% is 2nd highest in world

Motivation + effort-levels of gov't school teachers in India are a serious problem
- → high teacher absenteeism (25%)
  - ↳ 90% non-capital spending goes to teacher salaries
    - → teachers who are paid more (older, more educ) are more freq. absent
      - ↳ may signal to students that educ is unimportant
      - ↳ incentives — not wages — is the issue
- → 55% of teachers not teaching

Quality of education — not only quantity — matters
- → comparing quantities is hard because of diff educ. standards
- → strong evidence that cognitive skills (rather than school attainment which has not guaranteed improved economic conditions) are related to earnings, income equality, + economic growth
  - ↳ skill deficit due to more than diff in enrollment + attainment

Muralidharan + Sundararaman (2009)

input only ←——————————————→ incentive only

feedback + monitoring (100 schools)	extra contract teacher + diagnostic feedback (100 schools)	individ incentive + diagnostic feedback (100 schools)
block grants + diagnostic feedback (100 schools)		Group incentive + diagnostic feedback (100 schools)

— business as usual (100 sch) —

Feedback + monitoring: learning levels may be low b/c teachers don't know how to help
- → intervention: existing teachers provided w/detailed feedback on students + subject to low-stakes monitoring

Block grants: signific. amt of money committed under RTE
- → students provided cash grants for student inputs

Contract teachers: use is widespread but highly controversial
- → schools provided w/additional teacher (on contract)

Performance pay: teacher salaries largest component of educ spending but poor predictor of outcomes
- → teachers eligible for bonuses based on improved student performance

Intervention impressive due to scale (multiple treatment arms)
- → can compare using same sample
- → performance pay most notable (hard to implement)
  - ↳ randomized 2 performance incentives on individ + group level

Feedback + Monitoring: embedded in all treatments (interesting intervention + because there is a degree of monitoring in all interventions)

→ Treatment Schools
- ↳ took baseline tests
- ↳ detailed feedback on stud. perform., note how to read + use perform. reports, informed of an endline test
- ↳ 6 unannounced classroom monitoring visits/yr
- ↳ endline test

→ Comparison Schools
- ↳ no baseline test
- ↳ no initial contact w/schools + no feedback reports
- ↳ 1 unannounced visit
- ↳ surprise endline test

→ Results
1. Teachers in perform. monitoring schools performed better on measures of teaching activi
2. No diff. in test scores b/t students in treatment + comparison schools
   - ↳ suggests teachers temp. Δ behavior when observed but did not actively use feedback reports in teaching

Block Grant
→ schools spent most of grant on non-durables
- ↳ nearly half spent on child stationary + 40% on classroom materials/practice books
- ↳ small amt allocated to durable materials + sports goods

→ Results
1. Student learning improved 1st yr but not 2nd
2. Household spending fell significantly when grant was anticipated
   - ↳ undoing treatment: overall spending not increased
     - → grant used for materials parents likely to buy (substitution)
     - → should focus on items less substitutable by households (teachers or infrastr.)

Extra Contract Teachers
→ CTs hired by school committees + tend to be young females w/no formal teacher qualif, + from same village as school
- ↳ younger, less educ., closer to school, + paid signif. less

→ concerns:
1. untrained + less qualified CT's will not improve learning
2. Decentralized hiring will lead to local elite capture of teacher post
   - ↳ elite patron: i.e. hiring just to give a job?
3. CT's are exploited as a result of being paid significantly less

→ Results
1. Students in extra CT schools signific. outperform students in comparison schools
2. CTs have lower rates of absence + higher rates of teaching activity
   - ↳ intrinsic motivation, not far travel, shame, short-term contracts (perform well if value their job)

Performance for Pay Concerns
1. Reduction of intrinsic motivation
   - → recognize framing matters; program framed in terms of recognition + reward for outstanding teaching rather than accountability
2. Teaching to the Test
   - → test designed for deeper knowledge/understanding
   - → less concern given low levels of learning
   - → research shows process of taking test can enhance learning
3. Threshold effects / neglecting weak kids
   - → min by making bonus a func. of average improvement of all students
   - → drop outs assigned to low scores
4. Cheating/paper leaks
   - → testing independent of Azim Premji Foundation
     - ↳ no connection to school

→ Results
1. Bonus schools perform better across board
   ↳ better in all subgroups: grades, subjects, districts, level of question difficulty
   ↳ no signific. difference by most student demographic levels /variables (interaction terms)
      → i.e. household literacy, caste, gender, baseline score
   ↳ indicator of broad based gains
2. Not a single child was worse off relative to comparable child in control school
3. Broad-based impact
   ↳ true learning: perform better on conceptual—not only mechanical—questions
   ↳ spillovers: perform better on non-incentive subjects

→ Group incentives
   ↳ Net impact ambiguous.
      → may induce less effort due to free-riding
      → if gains to cooperation (i.e. complements), group incentives may yield better res
   ↳ Findings from the study
      1. Both group + individ. incentive programs had signif. positive impacts on test score in both
      2. First year: equally effective; second year: individ incentives do sign better
      3. Equally cost effective

→ Absence + active teaching was not improved under observation, but effort intensity went up (special prep: extra hw, classwork, classes, tests, etc.)
→ Incentives act as a force-multiplier to magnify impact of input
   ↳ qualifications combined w/incentives can impact learning outcome
   ↳ magnitude of incentives matter (higher paid teachers respond less)
   ↳ young teachers respond better new policy initiatives (more exp. teachers respond less)
→ Teacher opinion on perf. pay positive
   ↳ also support performance pay under an overall wage-neutral expect.
   ↳ signif. positive correlation b/t teacher performance + desired pay performance
      → suggest effective teachers know who they are
         ↳ likely to be sorting benefits from performance pay

Bonuses on performance had larger impact than input incentives
→ pure incentives most effective before mixed input-incentive
   ↳ also more cost effective (3X)

LECTURE 10

Difference in Differences

Looking at the change over time in the control group
1. Collect baseline data for both the Treatment + Control groups before the program
2. Collect final for both groups after the program
3. Compute the before-after Δ for each group
4. Subtract the differences

Assumption of parallel trend: same Δ over time would have happened in the treatment
→ method fails if comparison group on diff. trajectory

When to use DiD?
→ cannot always randomize (est. impact of "past" program)
→ need to identify treatment + control
→ can use natural experiments (unexpected Δ in policy, policy which affects certain ages. exploit variation of policies in time + space)
→ quality of comparison determines quality of eval.

we want to est. ave. treatment effect on treated (ATET) which is $E(Y_1^T|T) - E(Y_1^C|T)$
→ where Y_1^T is potential outcome if treated in period 1 (after treatment)
Y_1^C ——— untreated ———
Y_0^T ——— treated in period 0 (before treatment)
Y_0^C ——— untreated ———

$$E(Y_1^T|T) - E(Y_1^C|T)$$
$$\underbrace{E(Y_1^T|T) - E(Y_1^C|C)}_{\text{observable}} + E(Y_1^C|C) - E(Y_1^C|T)$$

if we add + subtract $E(Y_0^C|T) + E(Y_0^C|C)$ we get

$$ATET = \underbrace{[E(Y_1^T|T) - E(Y_1^C|C)] - [E(Y_0^C|T) - E(Y_0^C|C)]}_{\text{DiD}} + \underbrace{[E(Y_1^C|C) - E(Y_0^C|C)] - [E(Y_1^C|T) - E(Y_0^C|T)]}_{\text{Remainder (parallel trend)}}$$

in a regression framework: $Y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Program_i + \beta_3 (Post \times Program_i) + \varepsilon_{it}$

$E(Y_{i1}|Prog=1) = \beta_0 + \beta_1 + \beta_2 + \beta_3$
$E(Y_{i0}|Prog=1) = \beta_0 + \beta_2$
$E(Y_{i1}|Prog=0) = \beta_0 + \beta_1$
$E(Y_{i0}|Prog=0) = \beta_0$

DiD $= \beta_3$

Sensitivity Analysis for DiD
1. Perform "placebo" DD: use "fake treatment group"
   → ex: use previous years to see if parallel trends hold
   → ex: use treatment group you know was not affected
2. Use diff comparison group (two DD should give same est)
3. Use an outcome variable Y_2 which you know is not affected by intervention
   → using same comparison + treatment year
   → problem if DD est diff from zero

Instrumental Variables
IV regressions can eliminate bias when $E(u|x) \neq 0$
→ IV breaks x into 2 parts: correlated w/u + not
↳ instruments detect movements in x uncorrel. w/u to est β

Two conditions:
1. **Instrument relevance**: $corr(z_i, x_i) \neq 0$
2. **Instrument exogeneity**: $corr(z_i, u_i) = 0$

Two-Stage Least Squares
1. Regress $X_i = \pi_0 + \pi_1 z_i + v_i$ (uncorrel w/ u_i by assumption)
   → compute $\hat{x}_i = \hat{\pi}_0 + \hat{\pi}_1 z_i$
2. Replace x_i w/ $\hat{x}_i$ + estimate $Y_i = \beta_0 + \beta_1 \hat{X}_i + u_i$
   → $\hat{x}_i$ is uncorrel. w/ u_i + β_1 can be est. by OLS called $\hat{\beta}_1^{TSLS}$ (consistent est)

A note about control variables:
1. Do not want to control for aspects that treatment affects
2. Good controls are predetermined before treatment
3. Bad ——— are affected by variables we are interested in

LECTURE 11

When designing experiment want to reconcile good experimental procedure w/political pressure
- → restrict areas of need + randomize within

Supply-side Resources: School construction in Indonesia

Dasar INPRES Program: financed by oil revenues
- → 62000 schools (3 teachers + 120 pupils) constructed (more than one school per 500 children ages 5-14 in 1971)
- → idea: reduce the cost of school (less distance)
- → built in areas with more children + where initial enrollment is low

Research Design
- → 2 areas: high + low intensity treatment areas
  - ↳ just comparing these 2 variables may cause OVB
    - → 2 conditions for OVB
      1. Omitted variable is correlated w/ included regressor
      2. ________ a determinant of dependent variable
    - → True: $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$
      Est: $y = \alpha + \beta_1 x_1 + u_i$ where $u_i = \varepsilon_i + \beta_2 x_2$ $\Big\}$ $\hat{\beta}_1 = \beta_1 + \beta_2 \left[\frac{cov(x_1, x_2)}{var(x_1)}\right]$ * zero if randomized
    - → ex: wage = $\alpha + \beta_1$ educ $+ u_i$
      - ↳ ability (overest) + experience (underest)

- → Duflo (2001): Difference-in-Difference
  1. Ages 2-6 yrs old in 1974 (during school expansion)
  2. Ages 12-17 ________] placebo
  3. High intensity
  4. Low intensity] need this to capture diff. between cohorts (educ ↑ over time)
- → Counterfactual: in absence of program, incr. in educ attainment of young cohorts relative to older would not have been systematically diff. in low + high program regions
- → Data: 1995 intercensal survey of Indonesia (SUPAS)
  - ↳ men born b/t 1950 + 1972: 152,989 individ in sample (60633 work for wage)
  - ↳ list of schools scheduled for construction in each district for INPRES
  - ↳ individ match w/region of birth + program intensity
    - → assume most attain primary educ in place of birth

$S_{ijk} = C_1 + \alpha_{ij} + \beta_{ik} + (P_j \times T_i)\gamma_1 + (C_j \times T_i)\delta_1 + \varepsilon_{ijk}$

S_{ijk} = educ of i born in j in yr k
T_i = indicator of young cohort
P_j = intensity of program
C_j = vector of region-specific vars
α_{ij} = district of birth fixed effect
β_{ik} = cohort ________

* w/control variables, DID coefficient is still robust + statistically significant

* see individ age coeff (effects on young)

- → Estimating returns to educ using INPRES as an instrument
  - ↳ assume wages Δ through educ which is affected by program
  - ↳ issues: adding schools can affect wages through
    1. increased supply
    2. increased employment] upward pressure on wages
  - ↳ not much Δ (7-8% incr. in wages)

Conclusion
1. Program led to incr. of 0.25 to 0.40 yrs of educ.
2. Suggests 3-5.4% incr. in wages
3. Combined effects generates 2SLS est of economic returns to educ: 6.8-10.6

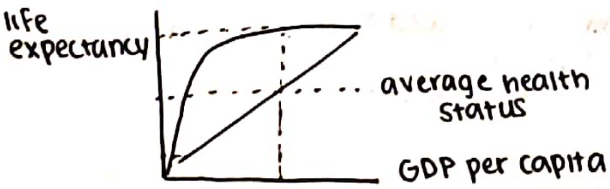
LECTURE 12

Overview of Health + Development

Epidemiological transition: a dramatic incr. in life expectancy due to several medical breakthroughs (e.g. penicillin + antibiotics)
- → DDT pesticide: eradicated malaria
- → WHO: info campaigns, vaccination drives; set std for healthcare
- → massive improvements but poor countries still lag behind (esp. w/diseases)
  - ↳ diseases incidence concentr. in poor countries

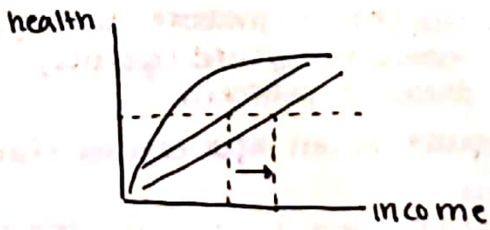
Measuring Health Outcomes
1. Self-reported health status
   - → issues: diff perspectives, unaware of illness, underest., nonclassical measurement error (systematic error correlated w/characteristics not controlled in reg like age, edu)
2. Self-reported symptoms
3. Activities of daily life
   - → usually more objective but can be overest. (gender)
   - → typically rated on a scale
4. Nutrient intake: malnutrition is a major source of health problems; ask food purchased or consumed in last month
   - → issue: memory, public good nature of food (household level consumption − intra-household consumption), hard to categorize (work food?)
5. Anthropometric measures
   - → height: measure of health stock/endowment
   - → weight: short term health
   - → BMI
6. Biomarker data: bio tests like bloodwork, hair sample, cortisone levels
   - → low measurement errors but causality issues
   - → hard to implement
7. Disability Adjusted Life Years (DALY): $DALY = N \times L + I \times D \times L$
   - → NL: number of years lost to fatalities from disease
     - ↳ (number of deaths from disease) (life expectancy at age of death)
   - → IDL: number of years lost through disability
     - ↳ (incidence of disease)(disability weight)(ave. duration of disease)
   - → standardize so makes disease incidence comparable

Preston Curve: concave relationship b/t GDP per capita + life expectancy



- → diminishing marginal returns
- → life expectancy has large implications on behavior
- → microdata to show income affects health
  - ↳ matters for income redistribution policies

- → but causality can go the other way
- → gradient in wage based on height in Brazil
- → policy: health redistribution
  - ↳ witch killings: kill off elders
    - → elders were a major drain to society



Instrumental variables

First stage: $x = \delta_0 + \delta_1 z + u$

Reduced form: $y = \gamma_0 + \gamma_1 z + \varepsilon$

IV estimator: $\hat{\beta}_{IV} = \dfrac{\text{reduced form}}{\text{first stage}}$

$= \dfrac{\sum z_i y_i}{\sum z_i x_i}$

$= \dfrac{\text{cov}(z, y)}{\text{cov}(z, x)}$

Intention-to-Treat Estimator (ITT)

→ needed when some in treatment group are not treated + control are (sample selection bias)
  ↳ causes allocation of treatment to be nonrandom
→ preserve original assignment where it was randomized
→ Ex: deworming medicine to children
  ↳ A = weight gain of children who were treated] both in treatment + control
  ↳ B = ______ not treated

  Impact of the program: A − B
  ↳ Ave outcome of treatment group: $Y(T) = A \times Pr(treat|T) + B \times (1 - Pr(treat|T))$
  ↳ ______ control: $Y(C) = A \times Pr(treat|C) + B \times (1 - Pr(treat|C))$

  measured outcome: $Y(T) - Y(C) = A(Pr(treat|T) - Pr(treat|C)) + B(Pr(treat|C) - Pr(treat|T))$

  $A - B = \frac{Y(T) - Y(C)}{Pr(treat|T) - Pr(treat|C)}$ → reduced form: diff b/t Y in treatment + control
  → first stage: propor. of treated conditional on treatment

  If we have $y_i = \alpha + \beta_1 P + \varepsilon$ where P = indicator of if they took a pill (not random if not only treatment took it)
  → need a N that affects P but not y → random assignment of pill (T)
  → First stage: $P = \delta_0 + \delta T + u \Rightarrow E(P|T=1) - E(P|T=0) = \delta_1$
  → Reduced form: $y = \alpha_0 + \alpha_1 T + v \Rightarrow E(y|T=1) - E(y|T=0) = \alpha_1$
    ↳ ave weight gain in treatment − ave weight gain in control
  → $\hat{\beta}_{IV} = \frac{\alpha_1}{\delta_1}$

This is a general technique when compliance is not perfect (ex: encouragement design)
→ But need treated amt in both groups to obtain Pr(treat|T) − Pr(treat|C)
  ↳ not easy
→ When not to use TOT
  ↳ when encouragement affects outcomes directly
    → not just getting people to take treatment (i.e. educational effects or signal that worms are bad from pill; so will take more)
  ↳ w/significant spillover externalities

LECTURE 13

Health + Economic Prosperity

Income → Health (improved diet, better sanitation + health services, used ↑ of health services)
Health → income (productivity, through concentr., strength, etc.)
confounds: age (β underest), gender (overest), educ (overest), genetics

Health Production Function: $H = H(N, L; A, B, D)$
→ N: vector of health inputs (choice)
→ L: labor supply (choice)
→ A: tech (hard to measure; usually residual)
→ B: family background (genetics)
→ D: disease environment

} policy levers: manipulate health inputs (nutrition, medicine) through RCTs or IVs
  → need exogenous, randomize variation

*manipulate health input does not effect how households invest in others

WISE Project

Goal: measure causal effect of iron deficiency on econ + social success
→ iron deficiency most common nutritional def. in world (2 bil)
  ↳ effects: reduced immune response, delayed cogn + physical development, fatigued + reduced work capacity

Intervention: 17000 respondents living in 4300 households
→ start in Aug 2002 + lasted 14 months

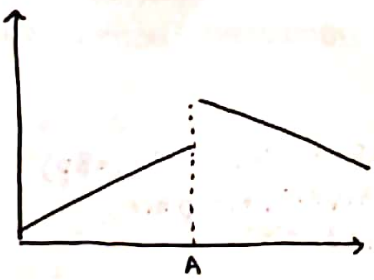
Data collection: pre-baseline survey (-4 months), baseline, + during (4 month, 8 month, + 16 month)
→ data intensive

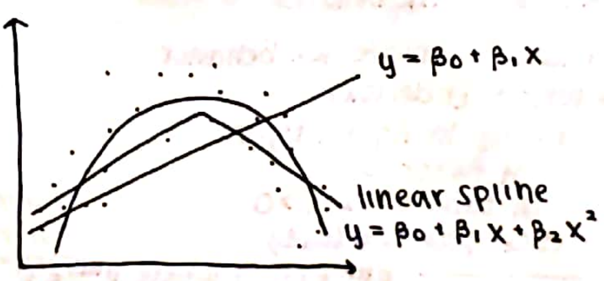
Research Design
→ treatment: 120 mg iron tablets (ferrous sulphate) + orange syrup for children 5 yr or less
→ control: placebo
→ Double blind: subject + researcher do not know allocation of treatment + control
↳ increases the validity of experiment
→ Potential issues w/o double blind
↳ John Henry Effect: control group responds to treatment
↳ Experimenter Effect: researcher projects desired outcome
↳ Hawthorne Effect: being observed can Δ your behavior
↳ Placebo Effect: believe receiving treatment
*had a facilitator
→ Household level assignment: avoid sharing within households (spillovers)
→ attrition is slow + not correlated to treatment + control status)

Results
→ high compliance rates (not self-reported; counted pills)
→ increased Hb in treatment (first-stage)
→ incr. physical endurance in those w/ low baseline
→ ↓ Pr(not working) + ↑ earnings
↳ use (wage)^1/4 instead of log to est. that relationship w/o losing data
→ $\log(0) = -\infty$
*↑ in social wellbeing

Conclusions
1. Cost of iron fortification is low (fish sauce: $6 per person per yr)
2. Earning over 4 mon. of low Hb self-employed males in WISE expected to incr. by ~RP 150,000 ≈ $40 over a year
3. Benefits > cost, so why not consume more iron-rich foods? (some free at health centers)
→ info? present-bias? no immediate benefits?

Linear Splines
we have $y_i = \beta_0 + \beta_1 x_i + \beta_2 (x_i > A) + \beta_3 (x_i (x_i > A_+)) + u_i$
→ where $(x_i > A)_+ = 1$ if $x > A$ + 0 otherwise
→ $x < A$ then $E(y_i|x_i) = \beta_0 + \beta_1 x_i$
→ $x > A$ then $E(y_i|x_i) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x_i$
@ A, there is a new intercept point (β_2) + new slope (β_3)
→ allow for kink/discontinuous jump in graph





Biggest learning curve?
Finding rewarding points when non-partisan analysis?
Culture? Working w/others?
Concentration or broad?
Results of test?
Favorite part of the job?

Migration + labor flows → need proper incentive w/o undue burden
→ human capital flows a natural valve for pressure (wage flexibility), spillovers
→ xenophobia, gov't deficit (anecdotle evidence)
→ humanitarian
→ dispell myths, social unrest (globalization backlash)

LECTURE 14

Facts About HIV/AIDs

- → 2003: out of 35 mil w/HIV, est 25 mil in Sub-Saharan Africa
  - ↳ in some countries (Botswana, Swaziland), over 30% are HIV+
  - ↳ life expectancy at birth in Sub Saharan Africa: 46 yrs
- → first doc case: Belgian soldier in 1959
- → 2% preg. woman carry HIV in Argentina, Brazil, Cambodia, India, Thailand
- → In Francistown (Botswana) + Harare (Zimbabwe), more than 1/4 preg women in ante-natal clinics HIV+
- → increasing dramatically in Asia
- → Spread of HIV
  - ↳ in Africa, higher infection rates in Southern + Eastern Africa
    - → wealth positively correlated w/HIV infection rates
    - → among middle class (high internal migration): soldiers, truckers, miners, etc.
  - ↳ main mechanism is heterosexual sex in developing countries
  - ↳ muslim countries have low rates (less extramarital sex or # of partners is smaller)
- → Treatment
  - ↳ symptoms: 5-10 yrs after contraction
  - ↳ untreated illness: 6-24 months before death
  - ↳ Drugs: ARVs + HAART — original cost $10k per person
    - → India: price wars between brand + generic forced pharma. companies to lower price of drugs
      - ↳ 2001: triple combo therapy available for $295 per person/yr
    - → 2013: ART was $115 per person/yr
    - → lower cost decr. incentive for HIV
- → Economic Development Implications
  - ↳ average human capital falls: sexually active ages = productive years
  - ↳ lower life expectancy = shorter time horizons (short discount rate)
    - → low saving/investment
    - → next gen human capital accumulation may be affected
  - ↳ labor productivity falls
    - → AIDS-related morbidity
    - → incr. job-turnover (less incentive for employers to promote job-specific HC accumul)
- → Why does it spread?
  - ↳ lack of info, risk behavior

Economic Model of Unsafe Sex Behavior

- → 2 time periods (1: decision)
- → instantaneous living utility: v
- → time discount factor: β
- → utility from unsafe sex: $S > 0$
- → HIV-survival prob: $p \in (0,1)$
- → HIV+ ———: $p^{HIV} \in (0,1)$ where $p^{HIV} < p$
- → Assumptions
  - ↳ unsafe sex, contract virus (+ vice versa)
  - ↳ max inter-temp. expected utility

$$EU(\text{safe sex}) = V + \beta(pV + (1-p)0) = V(1+\beta p)$$

$$EU(\text{unsafe}) = V + S + \beta(p^{HIV}V + (1-p^{HIV})0)$$
$$= S + V(1+\beta p^{HIV})$$

$EU(\text{unsafe}) > EU(\text{safe})$

→ decision rule: $\frac{S}{\beta V} > \underbrace{p - p^{HIV}}_{}$

risk premium: ↑ in developing, ↓ in developed

Comparative Statics

1. value of life (V) lower if individ is poor
2. lower β (↑ discount of future utility) may lead to more risky behavior
3. Difference in survival probabilities ($p - p^{HIV}$) may not always be high

	p	p^{HIV}	$p - p^{HIV}$
low income	med	low	small
middle income	high	low	large
high income	high	med	small

Extending the model: role of information
→ If do not know about infection status: likelihood $R \in (0,1)$ already being HIV+
↳ Assuming $\beta=1$: $EU(\text{safe sex}) = V + V[Rp^{HIV} + (1-R)p]$
$EU(\text{unsafe sex}) = S + V(1+p^{HIV})$
$EU(\text{unsafe}) - EU(\text{safe}) = S + V(p^{HIV} - p)(1-R) > 0$
↳ Nothing-to-lose effect: cost of unsafe sex smaller
→ If effect is important, incr. the R term can incr. risky behavior

What is missing from the model?
1. Altruism toward potential partners + members of society ($S = S(1-R)$)
2. Involuntary sexual choices (e.g. rape)
3. Social/cultural norms regarding acceptable sexual behavior
4. Pockets of poor info about HIV/AIDs in key pop (e.g. sex workers)
5. Other risky behaviors
6. Access to contraceptives

* anticipatory disutility

Public Policy: testing + informing change R to zero
→ can ↑ or ↓ unsafe sex depending on expectations of infectious status

Thorton (2008): Malawi
Randomized evaluation in rural Malawi
→ 120 villages in 3 regions
→ nurses also test for STIs
→ 91% of those offered accepted the HIV test (representative + higher than most DHS surveys)
→ 7% HIV infection for females, 5% for males
↳ couples informed separately
↳ all receive same sex counseling + educ
→ free treatment for STIs but not HIV (expensive + logistics)
Research Design
→ Two stages: 1) free door-to-door test
2) cash incentive offered at random to encourage results seeking
↳ randomized clinic distance (measure elasticity)
→ Researchers know true infection status of all in sample
↳ variation in knowledge driven by incentive (those who just go for cash ex)
→ exogenous variation in knowledge + distance
Results:
1. Small cash incentive effective @ boosting acquisition of info
→ ~50% point incr (may be affected by educ)
→ 34% people retrieve test results w/o incentive
→ causes less neg. relationship b/t receiving results + distance of clinic
2. Small costs impede people from getting test results
→ moderate stigma costs
3. HIV+ people who received results more likely to get condoms than HIV+ w/o test results
→ HIV- people not more or less likely to purchase condoms w/test results
4. Is the causal?
→ Getting results is own choice + cannot control or observe this
→ instrument: distance of clinic + cash incentive
↳ first stage: affects if you get results
↳ only affect condom purchase through this
→ but maybe cash incentives will allow you to buy more condoms?
5. condom purchase did not incr. a lot (from 1.5 to ~3)

LECTURE 15

Economic Consequences with Health

Health may be complementary w/other investments in an economy
- → education
- → skill-acquisition capabilities + cognitive development of individuals

Why might health capital raise return on investments in education?
1. Health is an imp. factor in school attendance
   - → healthier children more successful in school, learn more efficiency
2. School-age children deaths incr. cost of educ per worker
3. Longer life spans raise return to investments in educ
4. Healthier individ. more able to productively use educ at any point in life

Why might health/educ. capital raise return on investments in health?
1. Health prog./treatments rely on skills learned in school
2. Schools teach basic hygiene + sanitation
3. Educ needed for formation + training of health personnel

Linkages between Investment in Health + Education
1. Non-experimental studies find poor early children nutrition (height-for-age) is associated with: delayed primary school enrollment, reduced academic attainment, reduced academic achievement
   - → problems of causal identification: OVB, reverse causality

Miguel + Kremer (2004): Primary School Deworming

→ Busia District, Kenya
- ↳ agricultural area + fishing
- ↳ 1/3 children finish primary school (grades 1-8)
- ↳ Approx 20,000 school children screened for intestinal worms, age 10-18
  - → 90% infection rate + 33% serious infection rate

→ NGO: Primary School Deworming Project (PSDP)
- ↳ school-age children suffer highest intensity of worm infections + greatest morbidity

→ Randomized evaluation design
1. Schools divided in geographic zones + alphabetized
2. Divided into 3 groups

	1998	1999
Group 1 (25 schools) ↓	treatment	treatment
Group 2	comparison	treatment
Group 3	comparison	comparison

* 75 of 89 primary schools in area included in the sample

→ Randomization check
- ↳ sick often is stat. signific.: this is the outcome we want though
  - → may need to use DiD
  - → perhaps there will be mean-reversion

Program
1. Direct treatment effects: simple difference between treatment + comparison schools
   $E(Y_i|T_i=1) - E(Y_i|T_i=0) = b$

	Group 1	Group 2	G1-G2
rate of moderate-heavy infection	0.27	0.52	-0.25*
sick in past week	0.41	0.45	-0.04*
height for age	-1.13	-1.22	0.09*
hemoglobin	124.8	123.2	1.6

Why not zero? (in treated)
1. Only 75% compliance
2. Drugs are good but not perfect
3. Reinfection is still substantial

2. within-school externality impacts

	Group 1 (treat)	Group 1 (untr)	Group 2
rate of mod-heavy infection (1999)	0.24	0.34	0.52
(1998)	0.39	0.41	

* design not really structured for within school b/c not randomly assigned due to active choice
  - → could be due to spillover or selection effect (e.g. healthy people do not take pills)
    - ↳ panel data from previous years: selection effect not substantial but can still occur
  - → can modify w/double randomization: w/n + between schools

3. cross-school externality impacts
  - → large reductions in mod-heavy infection levels w/n 3 km (2 mi) of treatment schools in 1999, smaller positive reductions up to 6 km
    - ↳ compare children who live close to children assigned to a Group 1 school nearby, relative to children who do not have many neighbors attending a Group 1 school
      - → experiment works because Group 1 schools is randomized
      - → also control for density of children in area
        - ↳ areas of higher density, contagion may be more likely

$Y_{ijt} = a + bT_{ijt} + \sigma_d \gamma_d N^T_{dit} + \sigma_d \phi_d N_{dit} + \delta X_{ijt} + e_{ijt}$

Y_{ijt}: health outcome for each child in school j at time t

T_{ijt}: indicator equal to one if the school is Group 1

N^T_{dit}: number of children in Group 1 who lived w/n d kilometers of child (divided by 1000); number of treated
  - → control for nonrandom component of children group

N_{dit}: number of children who lived w/n d kilometers of child (divided by 1000)

b: average direct effect of receiving deworm treatment on health outcome

γ_d: effect of having one more treated child w/n d kilometers from you on your health outcome; est. of externality

→ $E(Y_{ijt} \mid T_{ijt}=1, N^T_{dit}=1, N_{dit}, \bar{X}) = a + b + \gamma_d(1) + \phi_d N_{dit} + \delta\bar{X}$
ave outcome for a child in Group 1 school who has 1000 children in Group 1 schools living w/n distance d from him

→ $E(Y_{ijt} \mid T_{ijt}=1, N^T_{dit}=2, N_{dit}, \bar{X}) = a + b + \gamma_d(2) + \phi_d N_{dit} + \delta\bar{X}$
ave outcome for a child living/in Group 1 school who has 2000 children in Group 1 schools living w/n distance d from him

→ $a + b + \gamma_d(2) + \phi_d N_{dit} + \delta\bar{X} - (a + b + \gamma_d(1) + \phi_d N_{dit} + \delta\bar{X}) = \gamma_d$
ave effect of having 1000 more children in Group 1 living around you (cross-school externality)

→ an ave. moderate-heavy infection reduction of approx 20% points in study area attributable to cross-school externality

Implications of Treatment Externalities
  - → standard public finance theory: goods w/positive externalities are underprovided
    - ↳ social benefits not accounted for
    - ↳ less than optimal deworming w/o subsidy
  - → prev studies show deworming in schools have small impacts
    - ↳ underest from externalities (simple diff may not give reliable est)
      - → 10% point diff b/t group 1 +
      - → 25% ——— b/t group 1 treated + group 2

Drugs or behavioral change (handing washing, etc.)
→ If deworming drug + worm prevention practices are substitutes, then taking drugs will make kids less conscientious about avoiding exposure

	Group 1	Group 2	G1-G2
wearing shoes	0.24	0.26	-0.02 (0.03)
clean hands, clothes	0.59	0.6	-0.01 (0.02)
days of contact w/ fresh water	2.4	2.2	0.2 (0.3)

for girls 13+, effect was higher

Educational Impact
→ "school participation" measured by unannounced primary school visits
↳ reduces school absenteeism w/externality effects

	Group 1 (T)	Group 2 (C)	Group 3 (C)
Younger girls + all boys	0.84	0.73	0.77
older girls (≥13 yrs)	0.86	0.80	0.81
preschool, grade 1-2	0.80	0.69	0.70 ← biggest effect
grades 6-8	0.93	0.86	0.89

→ academic performance as measured by standardized tests (grades 3-8)
↳ may affect through: incr. time in school + greater efficiency of learning in school
↳ average test gain is zero
1. incr. congestion in classroom
2. quality of learning is low
3. time lags
4. supply side ↓

deworming is a demand side reform

Cost-Benefit Analysis
→ Program: US \$1.46 per pupil per yr
↳ larger scale program: US \$0.49 per pupil per yr (mass deworming cheaper as drug prices have fallen)
→ deworming as human capital investment
↳ health gains = more schooling = higher adult wages
→ Deworming led to 7% gain in school participation
↳ prev study: each yr of school = 7% higher wages
↳ gains in wages (7% × 7%) over 40 yrs + discounted @ 5% a yr
↳ deworming benefits 3x bigger than treatment costs

LECTURE 16

Health Facts: Charge or Not Charge?

→ 9–10 mil children under 5 die every year
  ↳ 2/3 of these deaths could be averted using existing health products

→ Policy debate: subsidize/give or not?

Arguments For
1. Diseases are communicable (externality of health behavior onto others)
2. People need to experiment w/product to learn its effectiveness
3. Enable experimentation + learning (rationale for temp. subsidy)
4. People are too poor + don't have access to credit
5. People do not invest due to immediate costs vs. long-term benefit

Arguments Against
1. Subsidized products may not be used properly (value less)
2. selection effect: people w/ low WTP have low returns (low willingness to maintain or less vulnerable to health iss)
3. Sunk cost fallacy (psych effect): more we pay, more compelled to follow thru
4. May not invest once subsidy is lifted (first price as ref point)
5. i.e. using bed nets for fishing

Free Distribution vs. Cost Sharing

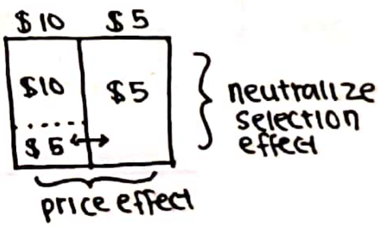
Providing bed nets for pregnant women in Kenya

Research Design

→ 16 randomly assigned health clinics w/discount varying b/t 90–100% of market price
  ↳ trade off between community + individual level
    → @ individ level, seeing others receiving the subsidy may influence decision
    → @ comm. level, usually have larger SEs ($\beta / \frac{\sigma}{\sqrt{n}}$)
      ↳ need large n
        → can limit treatment arms but doing so can lead to poor estimation (i.e. linear approx vs. quadratic)
      ↳ need the effect (β) to be large
      ↳ cost: must visit every community

→ Demand Curve has 2 effects:
  1. Selection effect: positive prices could weed out those who do not value the good
  2. Sunk cost effect: ——— induce more usage

→ 2 Stage Price Design: disentangle the effects
  ↳ in clinics w/a positive price, a subsample of those who decided to buy the net were surprised w/an additional discount
  ↳ double randomization



Results

1. Charging even small positive prices considerably decreased demand
   → ITN uptake dropped by 60% points when price incr. from 0 to 60¢
     ↳ 100 to 90% subsidy
2. Women who paid positive prices were no more likely to use ITNs than those who received for free
   → w/confidence interval, cannot reject diff b/t free + payment groups
3. Combo of lower uptake @ positive prices + no Δ in usage rate translates into lower absolute number of ITNs in use
4. Supports argument for free-distribution

MIDTERM REVIEW

[Econometrics]

Simple Difference $= E(Y_i^T|W_i=1) - E(Y_i^C|W_i=0)$

$= \underbrace{[E(Y_i^T|W_i=1) - E(Y_i^C|W_i=1)]}_{\text{average treatment effect on treated}} + \underbrace{[E(Y_i^C|W_i=1) - E(Y_i^C|W_i=0)]}_{\text{selection bias}}$

assign treatment but not treated

Gold Standard: Randomized Control Trials (RCT)

→ randomization assures no selection bias so that ATE = TOT

→ counterfactual: there are no systematic differences b/t treatment + control; in absence of treatment, outcome of both groups would be the same; $E(u_i|x)=0$

Difference-in-Difference (Double Difference)

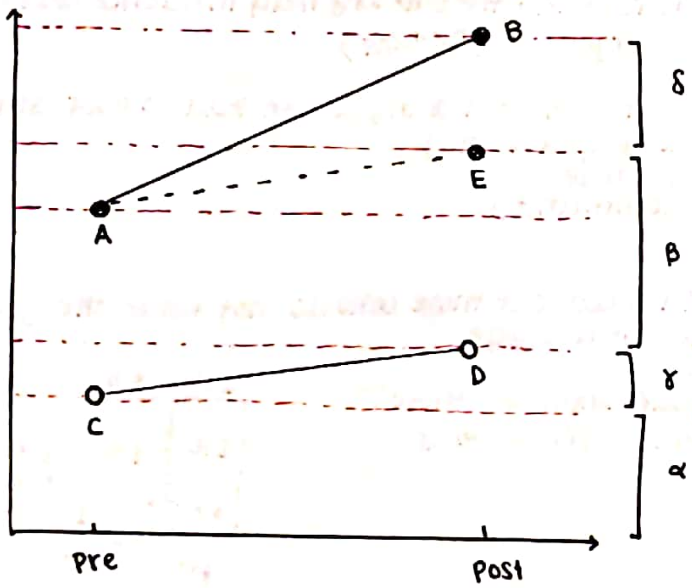
$DiD = \{[E(Y_{i1}^T|T_i=1) - E(Y_{i1}^C|T_i=0)] - [E(Y_{i0}^T|T_i=1) - E(Y_{i0}^C|T_i=0)]\}$ ← DiD

$+ \{[E(Y_{i1}^C|T_i=1) - E(Y_{i0}^C|T_i=1)] - [E(Y_{i1}^C|T_i=0) - E(Y_{i0}^C|T_i=0)]\}$ ← remainder

$y_{it} = \alpha + \beta T_i + \gamma Post_t + \delta(T_i * Post_t) + \varepsilon_{it}$

- $E(Y_{i1}^T|T_i=1) = \alpha+\beta+\gamma+\delta$
- $E(Y_{i1}^C|T_i=0) = \alpha+\gamma$
- $E(Y_{i0}^T|T_i=1) = \alpha+\beta$
- $E(Y_{i0}^C|T_i=0) = \alpha$

$DiD = (\alpha+\beta+\gamma+\delta-\alpha-\gamma) - (\alpha+\beta-\alpha)$
$= \beta+\delta-\beta = \delta$



counterfactual: parallel trends – in absence of the intervention, Δ between pre + post intervention periods for treatment group would have been the same for the control group

* can be used when you cannot randomize or when using a natural experiment (exploit variation of time + space)

* the 4-way treatment group is most optimal (think PROGRESA)

β: ave. difference in outcome b/t treatment + control arms for non-treated groups

γ: ______ in measured outcome b/t treatment + control in control arm (before the intervention)

δ: differential ave. treatment effect b/t treated + untreated groups

Omitted Variable Bias

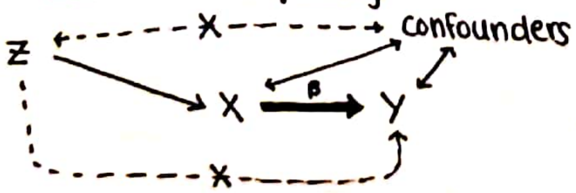
1. Determinant of y: $cov(y, x_2) \neq 0$ or $cov(x_2, u) \neq 0$
2. Correlated w/ x: $cov(x_1, x_2) \neq 0$

] $\hat{\beta_1} = \beta_1 + \beta_2\left(\frac{cov(x_1, x_2)}{var(x_1)}\right)$

Instrumental Variables:

Conditions: 1) Instrument Revelance: $cov(x_i, z_i) \neq 0$

2) Instrument Exogeneity (exclusion restriction): $cov(z_i, u_i) = 0$



$\hat{\beta}_{IV} = \frac{cov(z, y)}{cov(z, x)}$

Two stages least square

1. Regress $x_i = \pi_0 + \pi_1 z + \pi_2 W + v_i$ + compute $\hat{x_i}$
2. use $\hat{x_i}$ in the normal regression

*need at least one instrument for each endogenous variable

→ tradeoff: we give up precision (higher SE) when using IV typically because we only view part of the variation in x, but IV est. tend to be more consistent

Control Function approach

1. Calculate first stage regression + estimate $\hat{v_i}$ (residuals)
2. Use $\hat{v_i}$ as a control in structural regression

Intention-to-Treat

First-stage: $x_i = \gamma_0 + \gamma_1 z + u_i \Rightarrow \gamma_1 = E(X_i|Z_i=1) - E(X_i|Z_i=0)$

Reduced form: $y_i = \pi_0 + \pi_1 z_i + v_i \Rightarrow \pi_1 = E(Y_i|Z_i=1) - E(Y_i|Z_i=0)$

] $\beta_{IV} = \frac{\pi_1}{\gamma_1}$

* if x is an indicator of treatment take-up then

$$\beta_{IV} = \frac{E(Y_i|Z=1) - E(Y_i|Z_i=0)}{Pr(X_i=1|Z_i=1) - Pr(X_i=1|Z_i=0)} \quad \text{or impact} = \frac{Y(T)-Y(C)}{Pr(treat|T)-Pr(treat|C)}$$

Linear Splines: $y_i = \beta_0 + \beta_1 x_i + \beta_2 (x_i > A) + \beta_3 x_i (x_i > A) + u_i$

if $x < A$ then $E(y_i|x_i) = \beta_0 + \beta_1 x_i$

if $x > A$ then $E(y_i|x_i) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x_i$

} allow for discontinuous jumps

Experimental Procedures]

External validity: sample can be generalizable to whole population

Internal validity: net impact (minus confounding + control) is true impact ← allow for causal interpretation

population of eligible units → evaluation sample → treatment group / → control group

Golden Standard: RCTs benefits for external + internal validity

1. Just compare means
2. Easy to understand results
3. High internal validity

Confounding factors: extraneous variable correlated w/ x + y

counterfactual: world in which the treatment had not occurred

Conditional Cash Transfer: transfers conditional on certain behavior of recipient households
- → politically appealing
- → paternalistic
- → income effect + price effect

Unconditional Cash Transfer: transfers given regardless of behavior
- → income effect: incr. welfare
- → lower admin cost
- → lower impact on target

$$y = \alpha + \beta_1(UCT) + \beta_2(CCT) + \varepsilon_i$$

Attrition: systematic differences b/t treatment + control in withdrawals from study
→ affects internal validity when correlated w/ treatment var.

Double Blind experiment: subject + experimenter do not know allocation of treatment/control
→ increases validity + eliminates
1. John Henry Effect: control group responds to treatment
2. Experimenter Effect: researcher projects desired outcome
3. Hawthrone Effect: being observed Δs your behavior
4. Placebo Effect: believed to be receiving treatment

DiD Sensitivity Analysis:
1) Placebo or fake treatment group (see if parallel trend holds)
2) Different comparison group (same est)
3) outcome variable not affected by intervention (should be zero)

	T	C
C1	A	B
C2	C	D

Double randomization: used to disentangle 2 demand curve effects
1. Selection effect: positive prices indicate those who value the good
2. Sunk cost effect: positive prices induce more usage

Spillovers / Externalities: cause under/overest. of results
→ can see effects of control in treatment group
→ usually a factor of Distance (D)
→ may be that complementary behavior is reduced (treatment substitute)

Cost-Benefit Analysis: $PV(F_t) = \sum_{t=0}^{n} \frac{F_t}{(1+r)^t}$ where r = internal rate of return

Control variables: pre-determined before treatment (aspects that treatment does not affect)

[Education module]

Benefits of education for development
1. Perfect substitutes: more efficient workers
2. Imperfect substitutes: necessary for producing certain goods
3. Nonpecuinary: intrahousehold (women, intergen), political (civic participation, more info), social (less crime/violence, mobility), culture
4. Positive externalities: primary justification for intervention

Modeling School Choice

$EU(sc=0) = V + \beta V$

$EU(sc=1) = V - C + \beta[P(V+B) + (1-P)V] = V - C + \beta(V + pB)$

Decision rule: go to school if $\beta pB \geq C$

instantaneous utility: V
cost: C
benefit: B
Pr(benefit): P

Different Levels of Educ due to
1. Ability
2. Cost
3. Credit constraints

} assuming utility max

Policy
1. incr. realization of benefits: improve labor market outcomes
2. Reduce credit constraints for optimal investment
3. Reduce cost of schooling through subsidies
→ Demand-side

Schultz (2004): PROGRESA
↳ 50000 communities + poor ones selected based on marginality + welfare index in 1997; waves in 1998 + 1999
↳ DiD across time: by grade + sex
↳ counterfactual: parallel trends
↳ results: 8% return rate (incr. schooling by 0.66 yrs w/12% return per yr)
→ low elasticity of demand: 0.2
↳ limitations: externalities + attrition

Bobonis + Finan (2009): Peer Effects
↳ use simple diff. + IV + double diff.; interact w/distance
↳ counterfactual: no other mechanisms through which spillovers occur, parallel trends; simple — same outcome
↳ results: ineligible households in treatment group ↑ enrollment rate by 5% point
→ differential effects on welfare + grade level
→ 10% point incr. network enrollment rate
↳ limitations: other spillovers, OVB, reference group, behavior Δ

Baird, McIntosh, Ozler (2011): Cash or Condition?
↳ unmarried school girls ages 13-22 given monthly \$5 stipend (CCT + UCT) in 88 districts
↳ simple difference w/indicator variables
↳ counterfactual: outcome would have been the same
↳ findings: CCT had larger effect on schooling than UCT (↓ marriage + pregn rates)
→ CCT cost effective for target; UCT provide reg. income support
→ CCT of \$5 = UCT of \$10

→ Supply-side

Glewwe, Kremer, Moulin (2009): Textbooks + Test Scores
↳ 100 schools in 4 groups: treatment in yr 1 + block grants in next 2 yrs
→ students grade 3-7 receive textbooks in English, grade 3,5,7 math
↳ DID + IV (pre-test) on other test scores)
↳ counterfactual: parallel trends
↳ results: ones w/ higher pre-test scores impacted
→ curriculum poorly matched to student needs
↳ limitations: external validity (quasi-random), roll out system, IV validity

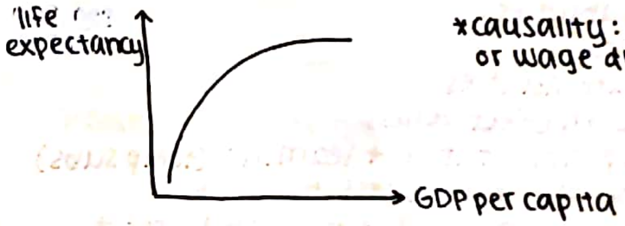
Muralidharan + Sundararaman (2011): Teacher Performance pay
↳ 600 sch. w/ 5 treatment + 1 control
↳ simple difference
↳ results: feedback had no Δ in test scores, student learning ↑ initially w/ block grants but lower due to substitution. more CTs caused better performance, individ incentive have broad-based gains w/ force multiplier, group incentive less effective through years (cost-effective)
↳ limitations: external validity (attrition), causality, linear relationship?

Duflo (2001): School Construction
↳ 62000 schools built for children 5-14 (esp. in ↑ pop. areas)
→ high/low intensity + ages 2-6/ages 12-17
↳ DID, instrument INPRES program
↳ results: incr. 0.25 - 0.4 yr of educ ≈ 3-5.4% incr. in wages
→ combined effect w/ TSLS, returns to educ 6.8-10.6
↳ limitations: attrition, educ location assump, progr. incr wages through ↑ supply + wages

[Health module]

Measuring Health Outcomes
1. Self-reported health status
2. ——— symptoms
3. Daily Activities
4. Nutrient intake
5. Anthropometric measures (BMI)
6. Biomarker data
7. DALY

Preston curve



*causality: health or wage distribution

Income → Health: diet, sanitation, health services] confounds: age, gender, educ
Health → Income: productivity

Health Production Function: H = H(vector inputs, labor supply, tech, genetics, environ)

Thomas (2006): Health on Labor Market Outcomes
↳ 17500 individ in 4300 households - iron supplements for 14 months
→ double blind, ITT due to noncompliance
↳ DID w/ linear splines: age + treatment, ITT
↳ ↑ Hb, Pr(working), earnings, physical endurance; low cost
↳ limitations:

HIV implications
→ on most productive yrs, low investment, fall in productivity
→ spread due to lack of info + risk behavior

$EU(\text{safe sex}) = V + \beta(pV + (1-p)0) = V(1+\beta p)$

$EU(\text{unsafe sex}) = V + S + \beta(p^{HIV}V + (1-p^{HIV})0) = S + V(1+\beta p^{HIV})$

V: instant utility
S: unsafe sex —
P: HIV-survival pr
P^{HIV}: HIV+ ———

Have unsafe sex if $\frac{S}{\beta V} > \underbrace{p - p^{HIV}}_{\text{risk premium}}$

risk premium → large for middle income
→ small for low + high income

→ Nothing-to-lose effect: cost of unsafe sex smaller when likelihood of already being HIV+ is greater than zero

Thornton (2008): Learning HIV Status

↳ 120 villages in 3 regions w/randomized clinic locations
  → free testing + cash incentives
↳ simple diff, IV (knowing results on cash incentive), interact knowledge on HIV status,
  → high compliance rates, low attrition
↳ results: cash effective at boosting info (but stigma is a high cost)
  → HIV+ + results = more likely to purchase a condom
↳ limitations: info spread, spillovers, behavior Δs, etc.

<u>Health on education investments</u>
1. School attendance + efficiency in learning
2. Children death ↑ cost of educ per worker
3. Longevity = more investment
4. Healthier = more productive use of investment

<u>Education on health investments</u>
1. Health prog/treat rely on skills
2. Teach basic sanitation
3. Formation + training of health personnel

Miguel + Kremer (2004): Worms

↳ school level (75 out of 89 chosen): high infection rates
  → divided by geozone + into 3 groups
↳ simple diff for direct treatment effects
↳ dummy + interaction for cross-school + within school externality
  → use distance as a measurement
↳ Results: small cash / less heavy infection rate + spillovers evident (up to 3 km)
  → More participation but no Δ in test scores
  → benefits = 3x cost
↳ limitations: incomplete compliance, reinfection, behavioral Δ (stat insignif)

<u>Full subsidies</u>

Pro:
1. Externalities
2. Learn effectiveness
3. Experimentation + learning (temp subs)
4. Too poor + no access to credit
5. No investment due to immed. cost + long-term benefit

Cons:
1. Valued less + not used properly
2. <u>Selection effect</u>
3. <u>Sunk cost fallacy</u>
4. Cause Δ in ref point (decreased investment)

Cohen + Dupas (2010): Free Distribution or Cost-Sharing?

↳ clinic level: double randomization of discounts for pregn. women
  → 16 clinics w/random discount 90-100%
  → separate selection + sunk cost effects
↳ simple diff b/t varying prices
↳ results: positive price dampen demand + no difference in usage
  → decr. absolute usage
↳ limitation: community vs. individ tradeoff (limited treatment arms + ↑ SEs), external validity (highly valued ITNs)

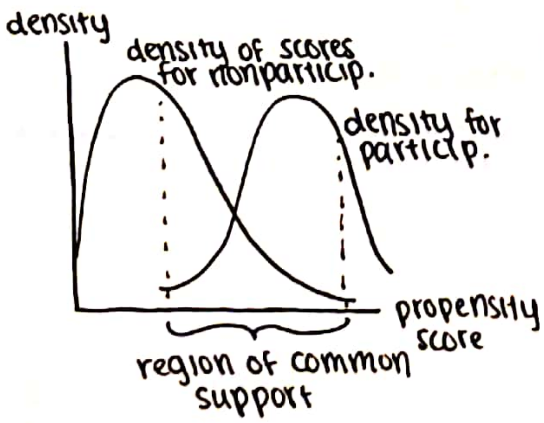
LECTURE 19

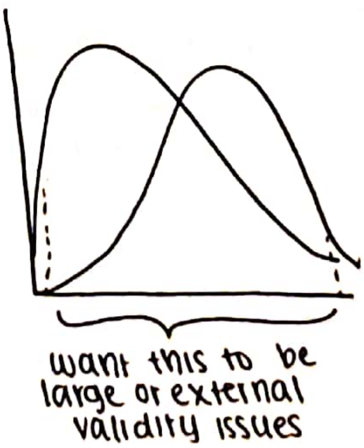
Propensity score matching (PSM)

- → when treatment cannot be randomized. mimic randomization (reproduce balance)
- → matching
  - ↳ develop a counterfactual or control that is similar to treatment in terms of observed characteristics
  - ↳ find large group of nonparticipants who are similar to participants in term of observed characteristics not affected by program
    - → each participant is matched w/similar participant (non)
    - → ave. difference in outcomes across 2 groups is compared for treatment effect
    - *causal → true treatment effect if differences in participation are based solely on differences in observed characteristics + if enough matching participants
  - ↳ issues: may be many important characteristics
    - → curse of dimensionality: too many bins (2^k)
    - → hard to find 2 individuals to match along characteristics } indexes can be useful
    - → Rosenbaum + Rubin solution: compute everyone's probability/propensity of participanting based on observ. characteristics
      - ↳ choose matches w/same probability
- → PSM: statistical comparison group by modeling the probability of participating in the program on the basis of observed characteristics unaffected by the program
  - ↳ match on basis of probability/propensity score
  - ↳ ATE = mean diff in outcome of the 2 groups
  - ↳ useful only when observed characteristics believed to affect outcome
    - → assumption: no selection bias on unobserved characteristics
  - ↳ Estimate the model: $P(X) = Pr(T=1|X)$
    - → capture effects of different observed controls (x) on participation in a single propensity score or index
    - → Rosenbaum + Rubin: under certain assumptions, as good as matching on x

PSM Assumptions

1. Conditional independence, unconfoundedness: given a set of observable covariates x that are not affected by treatment, potential outcomes Y are independent of treatment assignment T (when conditioned on X, T is treated as random)
   - → modeled as: $(Y_i^T, Y_i^C) \perp T_i | X_i$
   - → very strong assumption
     - ↳ violated if treatment is determined by unobserved characteristics
     - ↳ not testable but guaranteed w/randomization
   - → useful to have preprogram data (controls) + to understand assignment mechanism
2. Common Support: treatment observation have comparison observations nearby propensity score distribution — $0 < P(T_i=1|X_i) < 1$
   - → effectiveness also depends on large + roughly equal # of participants + non-participants observations so a significant region of common support
     - ↳ treatment units should be similar to non-treat units by obs. characteristics

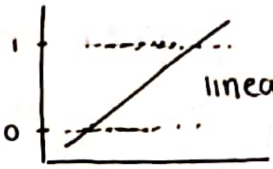




- → some non-treatments may have to be dropped due to noncomparability
- → sampling bias: nonrandom subset of treatment sample may have to be dropped if similar comparison units do not exist
  - ↳ examining characteristics may be useful in interpreting potential bias

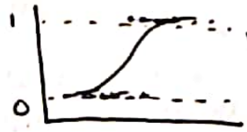
Estimating the Treatment Effect

1. **Estimate a Model of Program Participation**: $Pr(T=1|X)$, probability of participation T on a set of covariates X that are likely to determine participation
   - → **Linear model**: $T_i = \beta_0 + \beta_1 X_i + u_i$ where $E(u_i|X)=0$ (conditional independence)
     - ↳ $E(T_i|x_i) = \beta_0 + \beta_1 x_i$
     - ↳ If T_i is binary: $Pr(T_i=1|X_i) = \beta_0 + \beta_1 X_i$
     - ↳ Issue: functional form b/c not bound by Pr to 0 and 1
       - → but values go beyond bounds
       - → need nonlinear, asymptotic, S-shape form



linear prob model

   - → **Probit model**: $Pr(T_i=1|X_i) = \Phi(\beta_0 + \beta_1 X_i)$
   - → **Logit model**: $Pr(T_i=1|X_i) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_i)}}$



probit + logit model

   - → What covariates?
     - ↳ no guidance on selecting X
     - ↳ typically include variables correlated w/ program participation but not treatment
     - ↳ want flexible specification: interaction terms, higher polynomials
     - ↳ do not want to overfit, want good balance
       - → machine learning
2. **Defining the Region of Common Support + Balancing Tests**
   - → Find overlapping regions of distribution where propensity score for treatment + control
     - ↳ drop other observations
       - → check characteristics to see sampling bias
   - → conduct balance tests w/n each quantile of propensity score distribution
     - ↳ make sure X's are balanced b/t treatment + control: $\hat{P}(X|T=1) = \hat{P}(X|T=0)$
3. **Matching Participants to Nonparticipants**
   - → many different matching criteria on basis of propensity score
     - ↳ **Nearest-neighbor**: each treatment unit matched to comparison unit w/ closest propensity score; can choosen nearest neighbors (n=5 typically)
       - → can be done w/ or w/o replacement
       - → issue: 5 can still be far apart
     - ↳ **Caliper or radius**: threshold or tolerance on max propensity distance
       - → issue: high # of dropped participants likely (sampling bias)
     - ↳ **Stratification or interval matching**: partition data into strata + calculate program impact w/ each interval (mean difference)
       - → overall impact = weighted ave of interval impacts (weighted by share of participants)
     - ↳ **Kernel + local linear matching**: nonparametric technique uses a weighted ave of all nonparticipants to construct counterfactual match for each
   - → Regression methods
     - $Y_i = \beta_0 + \beta_1 T_i + \beta_2 \hat{P}(X) + \beta_3 (T_i * (\hat{P}(X) - E[(\hat{P}(X))]) + u_i$ → when taking $E(Y_i)$ goes to zero + allows interpretation
     - or $Y_i = \beta_0 + \beta_1 T_i + \delta X + u_i$ w/ weighted least squares
       - ↳ weights are 1 for participants + $\frac{\hat{P}(X)}{1-\hat{P}(X)}$ for non-participants
4. compute treatment effect

LECTURE 20

Global Disease Burden

major health problems in developing countries (esp sub-Saharn Africa):
- → Diarrhea
- → childhood diseases, respiratory infections, worms, measles
- → Malnutrition
- → HIV/AIDs

} Account for 70% of deaths among children under 5

Diarrhea causes dehydration
- → concentrated in early childhood (4 mil each year) usually from contaminated water
- → treatment: promotion of breastfeeding, oral rehydration therapy (19% poor + 50% rich)
- → prevention: access to clean water
- → 48% of poorest quintile did not seek medical treatment + 22% of richest

Jalan + Ravallion: Piped Water

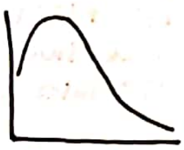
Cross-sectional survey in rural India 1993-1994
- → Est 1.5 mil child deaths per year (1/5 of rural have no access to safe water)
- → infrastructure projects hard to randomize so use PSM to evaluate, instrument
  - ↳ other ideas: using young vs. old cohort, find previous data + use DiD

Evaluation Design
- → Two groups: 1) households w/piped water: $D_i = 1$
  2) ——— w/o piped water: $D_i = 0$
- → Radius + near neighbor matching: nearest neighbor to the i^{th} participants is defined as non-participants that min $[p(x_i) - p(x_j)]^2$ over all j in set of participants w/5 nearest
  - ↳ $p(x_k)$: predicted odds ratio of observation k → $p(x_k) = \hat{p}(x_k)/(1-\hat{p}(x_k))$
  - ↳ matches accepted if $[p(x_i) - p(x_k)]^2 < 0.001$
- → 33000 rural households from 1765 households over 16 states in India
  - ↳ access to piped water: indicator whether households have piped water in/outside
  - ↳ outcome variable: prevalence under 5 ages + illness duration
- → Summary stats show heterogeneity between poor/nonpoor + educated

Propensity Score Matching
- → 650 treatment households dropped due to inability to find a good match
- → after match, mean of pscores for piped + control: 0.374 (SE = 0.189)

PS for piped

PS for nonpiped



* w/randomization, it should be uniform + identical

↳ there are characteristics which determine lack of piped water

Policy Implications + Results
1. Piped water has important impact on health
2. Average impact is deceptive indicator for inferring gains to children in poor families
   - → little/no impact on poor; high impact on rich
   - → w/better education, impact highest in poor areas
     - ↳ complementary of education
3. Policymakers trying to reach children of poor families will need to do more than rely on making facility placement pro-poor (i.e. locating interventions in poor area)
4. Highlight importance of combing public investment of infrastructure w/other interventions in education + income-poverty reduction

LECTURE 21

Microfinance Institutions (MFI)

- → marginal returns to capital for poor is presumably quite high
  - ↳ access to credit can lead to important welfare gains for the poor
  - ↳ but comercial banks have a difficult time lending to the poor (profitability)
    1. Asymmetric info: credit rationing (moral hazard + adverse selection)
    2. Contracts are difficult to enforce
    3. Lack of collateral
    4. Loan sizes tend to be small (high admin overhead)
- → Info Asymm: different levels of risk; risk may be correlated w/observable + unobservable characteristics
  - ↳ interest rates affect mix of interested clients
  - ↳ lenders do not want to raise rates lest they attract too many high-risk customers

A model of Credit Rationing

- → 2 potential customers: safe + risky
- → loan size: L
- → limited liability: borrower only repays if investment produces sufficient returns
- → safe always obtains secure loans: $R > L$
- → risky has probability of p earning R' + $R' > R$
- → net return safe: $R = (1+i)L$
- → ——— risky: $p(R' - (1+i)L) + (1-p)0$

Lender can

1. Auction off to highest interest rate
   - Safe: $i_1 = \frac{R}{L} - 1$
   - Risky: $i_2 = \frac{R'}{L} - 1$ $\Big\}\ i_2 > i_1$
   - → risky borrower willing to pay higher rate of interest than safe + IR is independent of probability of success
2. Randomly choose if $i = i_1$
   - $\pi_2 = p(1+i_2)L - L$
   - $\pi_1 = \frac{1}{2} i_1 L + \frac{1}{2}[p(1+i_2)L - L)$ $\Big\}\ \pi_1 > \pi_2$
   - if $p < \frac{R}{2R' - R}$ (low p = not raise IR to i_2)

Asymmetric Info

moral hazard: appears when contract provide incentives for borrower to engage behavior that goes against the interests of other (further risk)
- → don't observe actions of individ
- → insurance ex.

Adverse Selection: occurs when the contract attracts riskier types of borrowers
- → do not observe type of person
- → cars for lemons

Microfinance Revolution (1980s: Grameen Bank by Mohammed Yunus)

- → targets very poor households + lends to groups of borrowers
  - ↳ small average loan size ($100), no collateral, low interest rates (20%)
  - ↳ profitable + repayment rates are high
- → why success?
  - ↳ group liability: in event of default, no group member is allowed to borrow again
    - → reduce AS: chose own group; positive assortative matching (self-select)
    - → reduce MH: incentive to monitor peers
  - ↳ drawbacks: if one defaults, all members have an incentive to default
    - → lead to overly safe projects w/o sufficient risk

Miracle of Microfinance (Duflo, 2010)

- → In 2005, 52 of 104 neighborhoods in Hyderabad, India selected for opening of MFI (Spandana)
  - ↳ Group loan: 6-10 women per group w/25-45 groups forming a center
  - ↳ jointly responsible; first loan is Rs. 10,000 ($200)
    - → interest rate @ 12% w/50 weeks pay back — no strings attached
  - ↳ if all members pay, eligible for 2nd loan
- → Eligibility: women ages 18-59, resident for at least 1 yr (proof), valid ID, self-formed, + 80% must own a home (reduce transient population)
- → Experimental Design
  - ↳ 120 areas w/no pre-existing microfinancing, poor but not poorest, excluded large areas w/high amt of migrant workers
    - → stratified randomized design: 104 areas paired based on min distance of per capita consump. fraction w/debt, fraction who had business, + each pair randomly assigned to treatment
      - ↳ assure balance

→ 52 treatment areas b/t 2006 + 2007
  ↳ endline survey in 2007-2008 (oversampled households w/high propensity to borrow)
→ setting characteristics
  ↳ average household size: 5 ... see slide 24

Timeline
Jan 2005 - Feb 2006: baseline
Apr 2006 - Apr 2007: Spandana in treatment areas
Feb 2007 - Jan 2007: census
Jul 2007: endline sample frame selection
May 2008: moves into control areas

First-stage: statiscally significant relationship b/t treatment + credit

Results: no real significant effect on self-employment activities, income, consumption, + social effects (reduce informal credit use however)
→ Despite claims, demands for MFI is not universal
  ↳ 33% of households borrowed by end of 3-year study

Mobile Money
M-PSEA (in Kenya, 2007): SMS-based money transfer system
→ facilitates trade (easier for people to pay for + receive payments for g+s)
→ increase liquidity
→ save storage mechanism that promotes saving
→ increase remittance
→ increase risk sharing

LECTURE 22

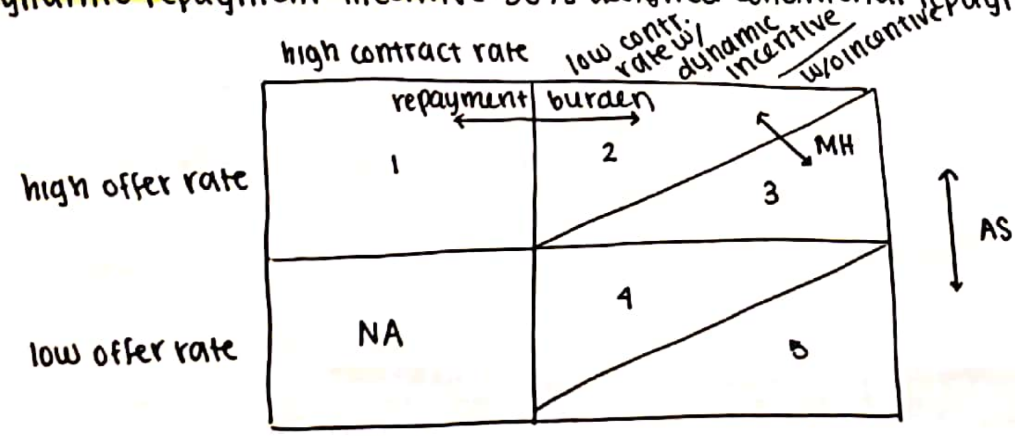
Asymmetric Info Experimental Design

→ Finding the relationship of high interest rate on default rate
- ↳ may be a positive relationship
  - → Adverse selection: attract risk-seeking people
  - → Moral hazard: repayment burden (higher end cost) or choosing higher risk projects once loan is received
- ↳ estimate regression w/OLS, logit, or probit (y = default rate between 0 → 1)
- ↳ causal relationship?
  - → reverse causation, confounders/OVB (person's time)
- ↳ Experimental Design
  - → Level: how many levels? Rate vary above or below market?
  - → control: those w/access to the market
    - ↳ will people go to area w/lowest interest
  - → self-selection: ruins randomization (adverse selection?)
- ↳ Distinguish b/t MH + AS
  - → give diff interest rates: examine adverse selection
  - → double randomize: moral hazard comparing w/n treatment

Karlan + Zinman (2007)

→ 57533 direct mail solicitations w/randomly different offer IR to former clients
- ↳ internal validity: repeat customers
- ↳ large sample b/c demand for credit is small
- ↳ 5028 clients go to branch + apply for the loan
  - → loan officer makes credit + loan supply decisions based on normal interest rates to blind experimental rates (4348 approved)
  - → client offered loan @ r, revision + of size + maturity allowed
    - ↳ finalized + told whether rate is good for one yr or one loan
  - → short survey, given loan, repayment behavior observed

→ offer rate: lender sends direct mail w/randomly assigned IR to prospective borrowers
→ contract rate: 40% randomly assigned a lower rate
→ dynamic repayment incentive: 50% assigned conditional repayment loan



→ conclusion
1. Weak AS on interest rate in full sample: stronger for women
2. Robust MH on dynamic incentive: may be driven by men
3. Info effects abate w/relationship
4. Economic significance: perhaps 20% of default due to asymmetries

LECTURE 23

Regression Discontinuity Design (RDD)

Many social programs select beneficiaries using an index or score
- → antipoverty: poverty index/income
- → pensions: age
- → education: high scores on standardized text (schloarship)
- → agriculture: fertilizers depend on # of hectares

Ex: election majority, Maimonides Rule

When to use RDD?
1. Beneficiaries/Non-beneficiaries can be ordered along a quantifiable dimension
2. This dimension can be used to compute a well-defined index or parameter
3. This index/parameter has a cut-off point for eligibility
4. ______ value is what drives assignment of a potential beneficiary to treatment

Intuitive explanation of method
- → potential beneficiaries (units) just above the cut-off point are similar to potential beneficiaries just below the cut-off
  - ↳ counterfactual: those above + below cut-off would have the same outcome in absence of the treatment
- → compare outcomes for units just above + below cutoff

s^* = cutoff point
--- = counterfactual
— = observation

*need a lot of data for best identification

bias vs. efficiency trade-off:
↑ data = ↑ interval = ↓ SE = ↑ bias

external validity issue: treatment effect for people @ discontinuity only

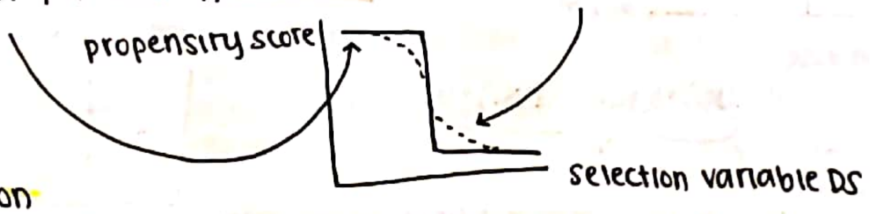
Types of Discontinuity

1. **Sharp discontinuity**: discontinuity precisely determines the treatment; equivalent to random assignment in a neighborhood
   - → ex: SS on age
   - → regression model: $y_i = \beta s_i + f(s_i) + u_i$

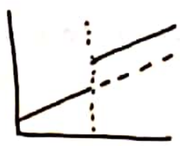
     $s_i = \begin{cases} 1 \text{ if } s_i \leq s^* \\ 0 \text{ otherwise} \end{cases}$
   - ↳ $f(s_i)$ is the functional form so if linear, $y_i = \alpha + \beta T_i + \delta s_i + u_i$

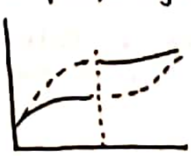
2. **Fuzzy discontinuity**: discontinuity highly correlated w/treatment
   - → ex: rules determine eligibility w/a margin of admin error
   - → regression model: use assignment as IV as program participation

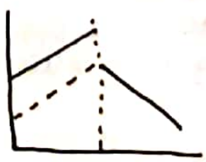
     $S_i = \alpha_0 + \alpha_1 1(s_i \leq s^*) + f(s_i) + \epsilon_i$
   - ↳ S_i is who should've taken treatm



Identification Assumption
1. Counterfactual of those below the threshold are those just above
2. Smoothness assumption: other factors that determine outcome are not discontinuous at threshold point; balance test
   - → cannot check unobservable but can plot for observable
3. Running variable has not been manipulated (i.e. ballot stuffing, asset reporting)
   - → compare distribution of index score w/typical or normal one
4. Functional form of running variable is properly specified

linear


quadratic


spline


LECTURE 24

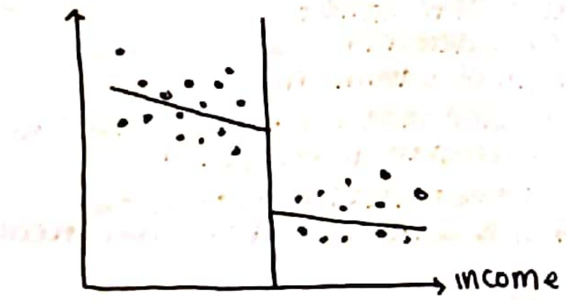
Monacorda, Miguel, Vigorito (20011) - PANES Program, Uruguay

→ examine impact of PANES (large anti-poverty cash transfer program) on gov't support
→ 2001-04: major economic crisis: 11% fall in real income
  ↳ left wing Frente Amplio coalition won in 2004, promising to help the poor
  ↳ Plan de Atención Nacional a la Emergencia Social
    → temporary by design 4/2005 - 12/2007
    → conditional on educ + health behaviors (de jure) but not enforced (de facto)
      ↳ income tested
    → Program components: (1) annual cash transfer = \$672, 50% of income
      (2) Food card for HH w/children = \$156-396 per yr
      (3) other: public works job, training, health care
    → Why should it affect gov't support? Voters chose based on past decisions + swing voters are targetted
    → Design of program
      ↳ 188,671 applic. households + 102,353 beneficiaries
        → 10% of all households in Uruguay or 0.41% of GDP
      ↳ assignment based on strict threshold of predicted poverty score
        → blind: households + officials not given formula
  ↳ PANES end in 12/2007 + replaced w/less generous Plan de Equidad (PE)
    → no difference in receipt of PE transfers between PANES + non-PANES household (55%)

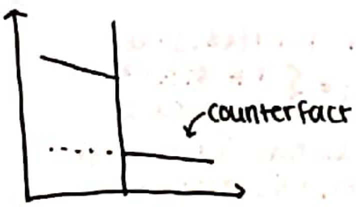
→ Program Assignment:

↳ bin scatters: create bins from running/continuous var + compute the ave. of x + y
↳ sharp design
↳ balance check: no discrenable jump in reported income, educ, household size, gender, age, etc.
↳ manipulation check:

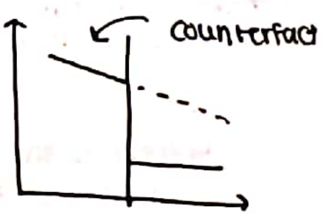
→ Graphical Results



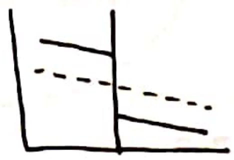
Those who receive program = more support?



Those who did not receive program = less support?



or both?



→ Regression Model

$$y_i = \beta_0 + \underbrace{\beta_1 \underset{\substack{\uparrow \\ \text{program assignment}}}{1(N_i<0)}}_{\text{constant}} + \underbrace{f_1(N_i) + 1(N_i<0) f_2(N_i)}_{\text{slope}} + u_i$$

* diff intercept + slope at discontin.
* spline model

Conclusion
1. PANES transfers boosted support for gov't + confidence by 10-15% points
   → robust w/RD validity checks (uptake high)
   → likely to translate into voting: pol + social views greatly affected by econ policies
2. Effects persist after end of program
   → pocketbook voting insufficient explanation
   → consistent w/simple rational model of learning about politician redistributive pref
3. Shortcoming: only stated pref.

LECTURE 25

Institutional Development

Difference in prosperity across countries
→ Standard economic answers
1. Physical capital differences: poor countries do not save enough → not enough investment
2. Human capital ———: ——— invest enough in educ + skills
3. Tech differences: poor countries do not einvest enough in R+D + tech adoption
   ↳ don't organize production efficiently
→ These are proximate causes of underdevelopment — but why does it occur?
→ Potential fundamental causes of underdevelopment
1. **Geography**: exogenous diff in environment (temp, land lock, disease)
   ↳ ex: poorest countries around equator
2. **Culture**: diff in beliefs, attitudes, + preferences (risk aversion + redistribution)
3. **Institutions**: humanly-devised rules shaping incentives
   ↳ North 1990: "rules of the game in a society or, formally, humanly devised constraints that shape human interaction"
      → key: humanly-devised, set constraints, shape incentives
   ↳ Two classes of institutions:
      1. Economic institutions: shape econ incentives, contracting possibilities
         → ex: property rights, contract enforcement, etc.
      2. Political institutions: shape political incentives + distribution of political power
         → ex: form of gov't, constraints on politicians + elites
   ↳ variation: enforcement of prop rights, legal systems, corruption, entry barriers, democracy vs. dictatorship, constraints on politicians + elites, electoral rules
   ↳ GDP per capita positively correlated w/ protection against risk expropriation, control of corruption, constraint on executive (checks-+-balances)

State capacity (gov't presence): ability of state to provide public goods, tax, + control violence
→ improve to impose "better" institutions? need good people
→ How to attract good people? (money, mission, benefits)
   ↳ **Personal traits**: matter for job performance (measure? screening? cost?)
      → Big Five: open, extraverted, conscientious, stable, agreeable
   ↳ **IQ: Raven Test** (educ proxy for harder to assess factors)
   ↳ **Intrinsic Motivation**: public service motivation index, pro-social behavior
      → dictator game + ultimatum game

Programa de Desarollo Regional

→ improve local gov't capacity in areas afflicted by drug-violence
   ↳ build network of comm. development agents
   ↳ assess demands of comm + report directly to fed gov't
   ↳ channel federal resources to local comm.

] selection of municipalities: random + not in border towns (high income + high violence)

→ Experiment
   ↳ attract younger, more educ recruits — went to universities
      → randomize wage based on location: \$5000 (87 percentile) + \$3750 (66 percentile)
→ Results
1. High wage offer attract those who had higher prev wage
2. ——— higher Raven score + less lower score (self-selection)
3. ——— those w/ more work exp, white collar jobs, conscientiousness, + emotional stability
4. Similar distribution on motivation b/t high + low wage (no crowd-out) → NGOs can pay ↑
5. High wage offers = higher acceptance rate (55% vs 42%)
6. Low wage: distance neg. correlated w/ acceptance (+ role of violence)
   ↳ higher wage offers = more willing to relocate

LECTURE 26

Economics of Corruption

Corruption: misuse of public funds for private gains
- → Pros: more to private sector, no welfare Δ (transfer of funds from Treasury to officials), most efficient firm can pay the highest bribe, grease/speed money
  - ↳ grease the wheel hypothesis: bribe to bypass inefficient systems or regulations
    - → may only be in the short-run; corruption ⟷ red tape
- → Cons: distortive, ↓ incentives for innovation, less competition (uneven playing field), invest in areas where corruption is less visible or easy (ex: construction)

Measuring Corruption
- → Challenges
  1. Corruption is illegal
  2. Gray Area: Mexico in 1990s gave 20% commercial loans to firms owned by bank's owners
  3. Traditional measures relied on perceptions: difficult to interpret
- → Old measures: control of corruption positively correlated w/GDP growth
- → **New approach**: better, objective measures of corruption
  - ↳ microevidence: investigate whether corruption affects channels + drivers of economic growth
  - ↳ ex: Ports of Maputo + Durban given agents to record corruption
    - → firms travel an additional 319 km (2x transportation cost) to avoid corruption
  - ↳ ex: corruption in educ in Brazil (↓ salary, no electricity, ↓ schooling) = ↓ test scores

Causes of Corruption
1. Incentives: asymm. info (moral hazard: exploit hidden action for gain)
2. Culture/norms: same incentive for diplomats not to pay parking ticket in NYC
   - → hard to Δ + a long process
   - → can overlay institutions or incentives to overlay dimension (ex: post 9/11 enforcement)
3. **Market/Organizational structure**: more centralized, less corruption
   - → Brazil: very decentralized + local gov't responsible for large provision of public goods (educ, health care, sanitation, infra)
     - ↳ receive 100s of millions from gov't
     - ↳ 2003: audit in Anti-Corruption Program through financial analysis, suppliers, on-site observ (schools, public housing, roads)
     - ↳ only 25% w/no evidence of corruption
     - ↳ incentives: first-term mayors stole $800 mill less + provided more public goods than second-term (unless equal likely to win re-election)

Fighting Corruption
1. **Higher wages**: more incentive to keep job
   - → key is transparency + accountability (politics)
     - ↳ **transparency**: decr. re-election rates w/higher # of corrupt violations
       - → randomized post + pre-election audits
       - → voters punish corruption: ↓ re-election rates w/higher vio in pre
       - → incr re-election rates after 4+: noise or Δ in campaign strategies
       - → prior: expect one violation (intersection of pre + post @ 1)
         - ↳ reward those w/no corruption
     - ↳ causal: randomized audit
     - ↳ accompaniment of public radio coverage enhanced effects
2. **Legal enforcement**
   - → Post 9/11 parking enforcement

* counterfactual: post election audits

* corruption is difficult to measure + more difficult to control
→ but there is progress!

FINAL REVIEW

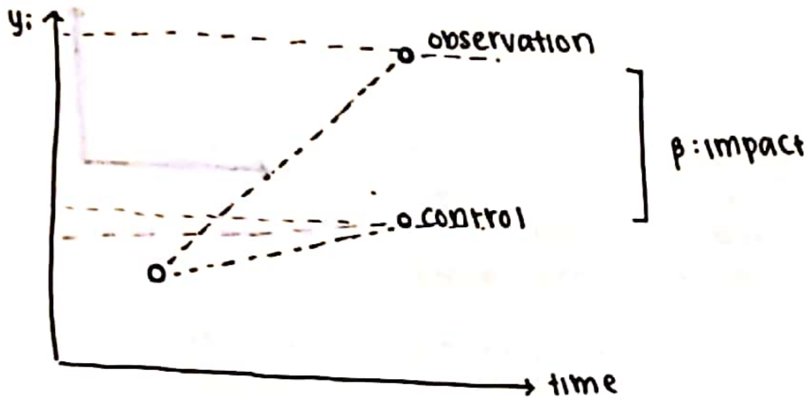
[Empirical Methods]

Simple Difference

1. Data set: RCT to assure no selection bias so that ATE = TOT
2. Counterfactual: in absence of treatment, outcome of both groups would be the same; no systematic difference between treatment + control; $E(u|x_i)=0$
3. Expectation Eq: $E(Y_i^T|W=1) - E(Y_i^C|W=0)$

$$\underbrace{[E(Y_i^T|W=1)-E(Y_i^C|W=1)]}_{\text{average treatment effect on treated}} + \underbrace{[E(\overbrace{Y_i^C|W=1}^{\text{assigned to treatment but not treated}})-E(Y_i^C|W=0)]}_{\text{selection bias}}$$

Regression: $y_i = \alpha + \beta T_i + \varepsilon_i$



Difference-in-Difference

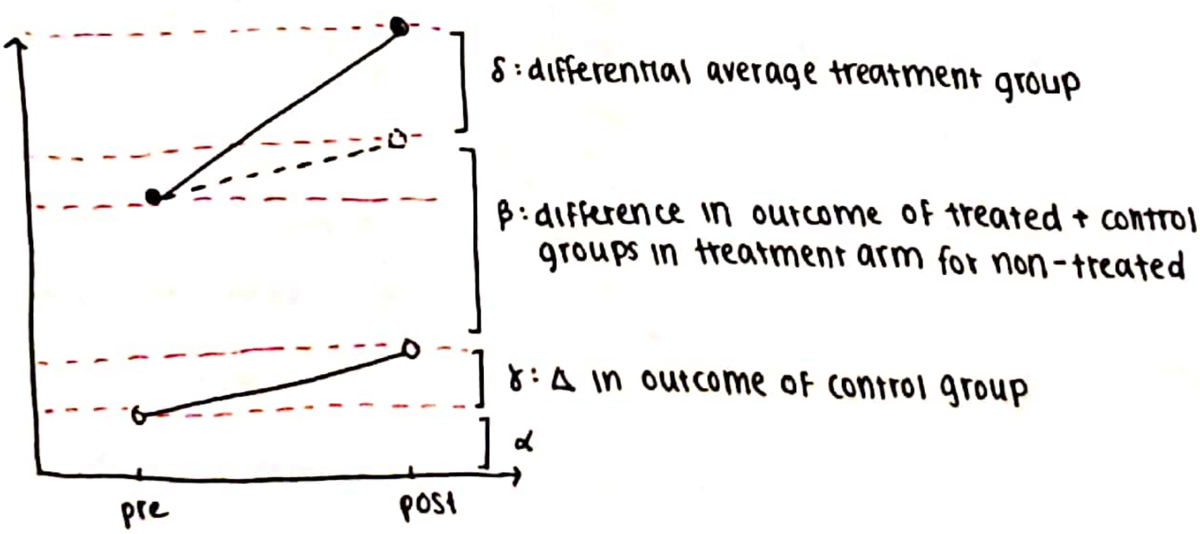
1. Data set: baseline data, panel data, treatment variable (natural exp); cannot randomize
2. Counterfactual: parallel trends (in absence of intervention, Δ between post + pre intervention for the treatment group would be the same as the control group
3. Expect. Eq: $\{[E(Y_{i1}^T|T_i=1) - E(Y_{i1}^C|T=0)] - [E(Y_{i0}^T|T_i=1) - E(Y_{i0}^C|T=0)]\}$ ← DiD

$+\{[E(Y_{i1}^C|T_i=1) - E(Y_{i0}^C|T_i=1)] - [E(Y_{i1}^C|T_i=0) - E(Y_{i0}^C|T_i=0)]\}$ ← remainder

Reg: $y_{it} = \alpha + \beta T_i + \gamma Post + \delta(T_i \times Post_t) + \varepsilon_{it}$

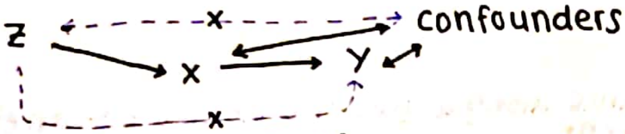
→ $E(Y_{i1}^T|T_i=1) = \alpha+\beta+\gamma+\delta$
→ $E(Y_{i0}^C|T_i=0) = \alpha+\gamma$
→ $E(Y_{i0}^T|T_i=1) = \alpha+\beta$
→ $E(Y_{i0}^C|T_i=0) = \alpha$

$DiD = (\alpha+\beta+\gamma+\delta-\alpha-\gamma) - (\alpha+\beta-\alpha)$
$= \beta+\delta-\beta = \delta$



Instrumental Variables

1. Data set: imperfect randomization, biased estimators
2. Identification assumption
   → Instrument Relevance: $cov(x_i, z_i) \neq 0$; inclusion restriction
   → Instrument Exogeneity: $cov(z_i, u_i) = 0$; exclusion restriction
3. Regression: Two Stage Least Squares
   → First Stage: $x_i = \pi_0 + \pi_1 z + \pi_2 w + v_i$ + compute $\hat{x_i}$] can use control function by est $\hat{v_i}$ + placing in 2nd st.
   → Second Stage: $y_i = \alpha_0 + \beta_i \hat{x_i} + \gamma w + \varepsilon_i$



* tradeoff: give up precision (higher SE) b/c only view partial variation in x (but more consistent)

$$\hat{\beta}_{IV} = \frac{cov(z, y)}{cov(z, x)}$$

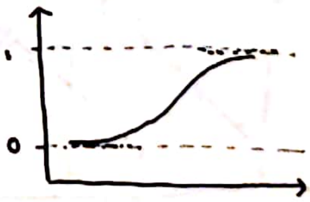
Intention-to-Treat

1. Data: nonrandom allocation of treatment + control; when some in treatment group are not treated but some in control are
2. Identification assumption: same as above
3. Regression:
   → First stage: $x_i = \gamma_0 + \gamma_1 z + u_i \Rightarrow \gamma_1 = E(x_i | z_i = 1) - E(x_i | z_i = 0)$
   → Reduced form: $y_i = \pi_0 + \pi_1 z_i + v_i \Rightarrow \pi_i = E(y_i | z_i = 1) - E(y_i | z_i = 0)$

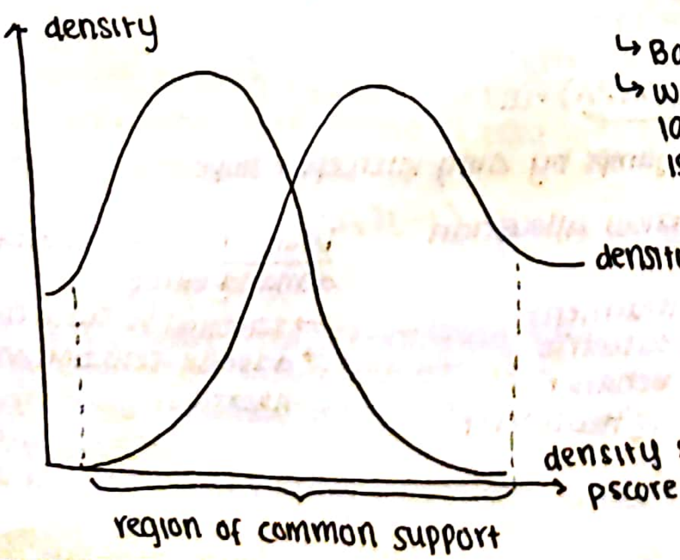
$$\beta_{IV} = \frac{Y(T) - Y(C)}{Pr(treat|T) - Pr(treat|C)} = \frac{E(y_i|z=1) - E(y_i|z=0)}{Pr(x_i=1|z_i=\ ) - Pr(x_i=1|z_i=0)} = \frac{\pi_i}{\gamma_1}$$

Propensity Score Matching

1. Data: when treatment cannot be randomized; observable characteristics which affect participation in the program but are unaffected by treatment
2. Identification Assumption
   → Conditional independence/unconfoundedness: $(Y_i^T, Y_i^C) \perp T_i | X_i$; outcome Y are independent of treatment T when conditioned on observable covariates x that are unaffected by the treatment (i.e. no OVB)
      ↳ strong assumption: may be unobservable characteristics, but guaranteed by randomization
   → Common Support: treatment observ have comparisons nearby propensity score distribution $[0 < P(T_i = 1 | X_i) < 1]$
   → Issues: sampling bias + choosing characteristics
3. Regression
   → Program Participation: Probit $[Pr(T_i = 1 | X_i) = \Phi(\beta_0 + \beta_1 x_i)]$
   Logit $[\ \ \ \ \ = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_i)}}]$



   → Region of Common Support:



↳ Balance test: $\hat{p}(x|T=1) = \hat{p}(x|T=0)$
↳ want common support region to be large or we run into external validity issues

→ matching based on propensity score

1. **Nearest neighbor**: usually n=5 but this can be far
2. **Caliper or radius**: threshold on max propensity difference, but this can result in many dropped participants due to sampling bias
3. **Stratification or interval**: weighted average of impact within intervals
4. **Kernel + local linear**: non parametric technique to construct counterfactual match for each

Regression Discontinuity Design

1. Data set: ordering of beneficiaries + non-beneficiaries along a quantifiable index w/ an eligibility cut-off + which drives assignment to treatment
2. Identification Assumption
   - → counterfactual: those above + below the threshold would have the same outcome in absence of the treatment
   - → smoothness: other factors that determine outcome are not discontinuous @ threshold
     - ↳ balance test to see manipulation
   - → running variable has not been manipulated
     - ↳ see if smooth or distribution is typical + not skewed to treatment
   - → functional form of running variable is properly specified
3. Regression:
   - → **Sharp Discontinuity**: discontinuity precisely determines treatment (like SS + age)

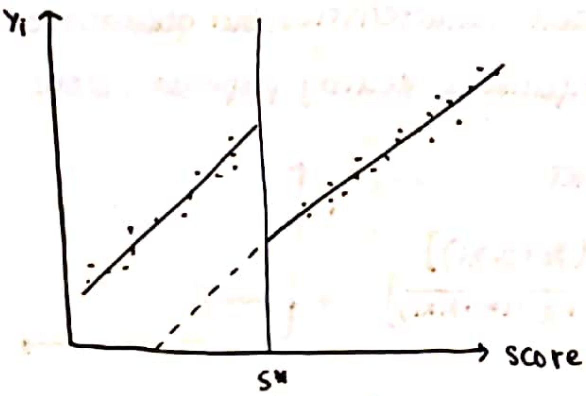
     $y_i = \beta s_i + f(S_i) + u_i$

     $s_i = \begin{cases} 1 & \text{if } S_i \leq S^* \\ 0 & \text{otherwise} \end{cases}$

     ex: linear $y_i = \alpha + BT_i + \delta S_i + u_i$

   - → **Fuzzy Discontinuity**: disconitn highly correl. w/ treatment; eligibility w/ margin of admin error
     - ↳ use assignment as IV for program participation

       $IV = T = \alpha_0 + \alpha_1 1(S_i \leq S^*) + f(S_i) + \varepsilon_i$



S^* = cut off point
--- = counterfactual
— = observation

→ need a lot of data for best identification but there is a bias vs. efficiency tradeoff (↑ data = ↑ interval = ↓ SE = ↑ bias)

Linear Spline: $y_i = \beta_0 + \beta_1 x_i + \underbrace{\beta_2 (x_i > A) + \beta_3 x_i (x_i > A)}_{\text{allow for discontinuous jumps by } \Delta\text{ing intercept + slope}} + u_i$

Double blind: subject + experimenter do not know allocation of treatment + control which eliminates

1. **John Henry Effect**: control responds to treatment
2. **Experimenter** — : exp. projects desired outcome
3. **Hawthorne** — : being observed Δes behavior
4. **Placebo** — : believe to be receiving treatment

Double random: disentangle 2 demand effects
- → selection vs sunk cost effect
- → adverse selection vs. moral hazard

[Concerns + Issues]

Omitted Variable Bias: confounding factors that bias the treatment variable

1. Determinant of y: $cov(y, x_2) \neq 0$ or $cov(y, u) \neq 0$
2. Correlated w/x: $cov(x_1, x_2) \neq 0$

$\Big\} \hat{\beta_1} = \beta_1 + \beta_2 \left[\frac{cov(x_1, x_2)}{var(x)}\right]$

* hard when this variable is unobservable

Selection Bias: selection of a sample that is not randomized (representative) of the sample; treatment + control differ in important + systematic ways

Sample Bias: sample collected in a way that some members of intended pop are less likely to be included than others (ex: PSM)

Attrition: withdrawals of the study w/systematic diff. b/t control + treatment (internal val)

Spillovers / Externalities: cost or benefit accruing to party / control group (usually factor of dist)
→ can under or overest. results depending if it is a complement or substitute

External validity: sample can be generalizable to the whole pop

Internal validity: net impact (minus confounding + control) is true impact
→ allow for causal interpretation

[Case Studies]

Health + Sanitation

Jalan + Ravallion (2003): piped water in India
→ PSM: radius + near neighbor matching
↳ heterogeneity bt poor/nonpoor + educ/noneduc
→ need public investment of infra w/other educ interventions (complement)

Credit

Microfinance institutions: access to credit can lead to important welfare gains for the poor but banks have a hard time lending to poor

1. Asymmetric info: levels of risk correlated w/observable + unobservable characteristics
   → moral hazard: incentive for borrowers to engage in risky behavior (ex: insurance) ↳ actions
   → adverse selection: contracts attract riskier borrowers (cars) ↳ type of person
2. Difficult to enforce contracts
3. Lack of collateral
4. Loan sizes tend to be

Model of credit rationing

→ Net return on safe loans: $R = (1+i)L \Rightarrow i_1 = \frac{R}{L} - 1$
where L = loan, R = return, i = interest
→ Net return on risky loans $= P(R' - (1+i)L) + (1-p)0 \Rightarrow i_2 = \frac{R'}{L} - 1$
where R' = high return w/ probability p

$\Big\} i_2 > i_1$

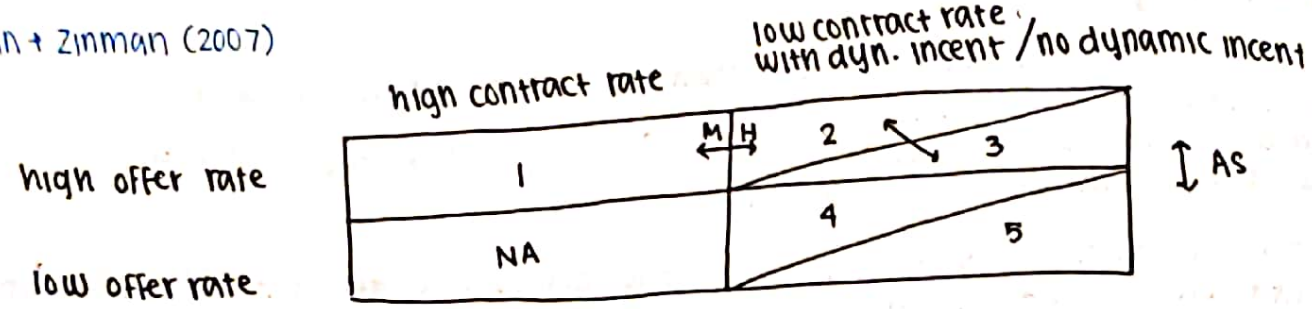
→ so riskier consumer willing to pay higher IR w/o regards to probability
→ what has higher pay off?

$\Pi_1 = \frac{1}{2} i_1 L + \frac{1}{2}[p(1+i_2)L - L]$
$\Pi_2 = p(1+i_2)L - L$

$\Big\} \Pi_1 > \Pi_2$ when $p < \frac{R}{2R' - R}$

· Duflo (2010): MFI in India
→ MFI in 120 areas w/no pre-existing MFI or high amt of migrant workers
→ PSM: stratified randomized design - 104 areas paired based on min distance of per capita consumption, fraction of households w/debt, fraction of households w/businesses
→ no real significant effect + demand is not universal

Karlan + Zinman (2007)



→ offer rate based on location; 40% given lower contract rate (of those, 50% w/dynamic repayment incentives)
→ weak evidence of AS on full sample (but stronger for women)
→ MH on dyn. incentives may be driven by men
→ 20% defaults due to asymmetries

*internal validity issues b/c repeat customers

Manacorda, Miguel, Vigorito (2011): PANES Program in Uruguay
→ regression model: $y_i = \underbrace{\beta_0 + \beta_1 \overbrace{I(N_i<0)}^{\text{program assign.}}}_{\text{constant}} + \underbrace{f(N_i) + I(N_i<0) f_2(N_i)}_{\text{slope}} + u_i$ (spline)

→ transfer boosted gov't support by 10-15% points w/persistent effects
↳ RD validity checks, no pocketbook voting
↳ but this is only stated pref

Institutions

Difference in economic prosperity:
1. Physical capital
2. Human
3. Tech

} Fundamental
1. Geography
2. Culture
3. Institutions

Bo, Finan, Rossi (2013): Programa de Desarollo Regional in Mexico
→ improve state capacity by recruiting comm. development agents
↳ better people defined in personality traits, IQ, + intrinsic motivation
↳ recruit young, educ students + randomize wage + municipality
→ High wage attract those w/high wage prev, higher Raven scores (IQ) — self selection, more exp / conscientious,
→ Wage diff. of high + low did not crowd out motivated (same distribution)
→ High wage = high acceptance
↳ low wage had neg. correlation w/distance + violence of muni

Corruption
→ Pro: more to private sector, wealfare Δ only, most efficient pay high bribe, grease/speed money
→ Cons: distortive, less comp, less innovation, invest only in areas where corruption is less visible, grease money can lead to more red tape
→ Causes: (1) incentives, (2) culture/norms, (3) market/org structure

Ferraz + Finan (2008): Corruption in Brazil
→ incr. transparency + accountability through random audits (pre + post election)
→ voters punish corruption through re-election rates
↳ accompaniment of public radio enhanced effects
↳ prior: at least one violation (reward none, punish large)
→ need higher wages + legal enforcement