API-202B Empirical Methods II

Session #10: Randomized Experiments

miguel_santos@hks.harvard.edu
@miguelsantos12

Today´s class: Randomized Experiments

- Introduction
- Policy question: Does deworming raise students' performance?
- Trichuris trichiura, growth and students' performance: Evidence from Jamaica
- Takeaways
- In preparation for Monday´s class with Lant Pritchett
- Vocabulary

Introduction to Randomized Experiments

- Randomly assigning subjects to control and treatment groups:
  - Ensures that treatment is not correlated with individual characteristics (even unobservable characteristics)
  - Ensures that any differences in outcomes between the treatment and control groups are attributable to the treatment, and not to other differences between the two groups

Introduction to Randomized Experiments

- We have two ways of estimate effects of randomized experiments:

	Difference in Means	Regression Analysis
Estimation	Impact = $\overline{Y}_T - \overline{Y}_C$	$Y = \beta_0 + \beta_1 TREAT + u$ TREAT is a dummy variable indicating person is in treatment group Impact : β_1
Hypothesis Testing	Conduct a t-test for the difference in means: H_0: $u_T - u_C = 0$	Conduct a t test to determine whether β_1 is significantly different from zero: H_0: $\beta_1 = 0$
Additional Benefits of Method:		Can include additional controls to improve the precision of the estimate: $Y = \beta_0 + \beta_1 TREAT + \beta_2 X_2 + \ldots + \beta_k X_k + u$ where $X_2, \ldots. X_k$ are several characteristics at baseline

Policy question: Does deworming raise students' performance

- Each year, approximately 1.3 billion people suffer from hookworm and roundworm infections
- There is concern that poor health due to such infections may reduce students' performance in developing countries
- Policy Question:
  - What are the effects of implementing a deworming program on students' performance?

Trichuris trichiura, growth and students' performance Evidence from Jamaica: Description

- Simeon DT, Grantham-McGregor SM, Callender JE, Wong MS (1995): "*Treatment of Trichuris trichiura infections improves growth, spelling scores and school attendance in some children*," Journal of Nutrition, 125, 1875-83.
- Study design:
  - 407 children in ages 6-12 in grades 2 to 5 from 14 public schools
  - Schools chosen on the advice of local health professionals: children already infected with *Trichuris trichiura infections* were recruited for study
  - Individual students were then randomly assigned to:
    - Treatment: Deworming drug given
    - Control: Placebo drug given
- Length of study: 6 months
- School attendance data were collected before and during the intervention (only available for 264 children due to loss of class registers in some schools/grades)

Trichuris trichiura, growth and students' performance Evidence from Jamaica: Description

- What do researchers mean by "residualized gain scores for each outcome variable"

They run a univariate regression between each outcome measure post-RCT (Y_t) and their corresponding levels pre-RCT (Y_{t-1}) and use residuals as their outcome variables. When they do this, they are stripping away the post-RCT result from all its correlation with pre-RCT levels on those same measures

Treatment effects were determined using multiple regression analyses. First, residualized gain scores were calculated for each outcome variable. To do this, linear regression analyses were conducted with the post-test measures as the dependent variable and the pretest measures as the independent variable. The residuals from these analyses were saved for use as the dependent variables in subsequent analyses. In these multiple regression analyses, treatment group was entered as a dummy variable and other potentially confounding variables were entered as covariates. The latter included gender, age, socioeconomic status, the intensity of *Ascaris* infections (log transformed) and school attended (as dummy variables). The analyses were repeated with treatment-by-stunting and treatment-by-intensity interactions as additional independent variables.

Trichuris trichiura, growth and school performance
Evidence from Jamaica: Team Work #1

TABLE 2

Mean school achievement test scores, school attendance and anthropometry of Jamaican school children at baseline and at post-test[1,2]

	Group		
	Treatment (n = 206)	Placebo (n = 201)	Significance[3] P
Arithmetic			
Baseline	19.2 ± 7.1	19.6 ± 6.5	0.54
Post-test	21.4 ± 6.2	21.5 ± 6.2	0.92
Spelling			
Baseline	14.3 ± 8.7	15.2 ± 8.6	0.29
Post-test	16.9 ± 10.2	17.6 ± 10.0	0.43
Reading			
Baseline	23.7 ± 13.4	24.7 ± 12.6	0.45
Post-test	27.7 ± 15.2	28.4 ± 14.0	0.65
School attendance (%)			
Baseline	62.6 ± 20.4	66.3 ± 20.8	0.15
Post-test	67.3 ± 18.4	69.3 ± 17.5	0.40
Height-for-age, *Z-score*			
Baseline	–0.48 ± 0.95	–0.39 ± 0.90	0.35
Post-test	–0.48 ± 0.97	–0.41 ± 0.89	0.46
Body mass index, kg/m^2			
Baseline	15.3 ± 1.3	15.5 ± 1.3	0.07
Post-test	15.6 ± 1.3	15.8 ± 1.4	0.25

[1]Values are means ± SD.
[2]Fifteen children (10 from the treatment group and five from the placebo group) did not complete the post tests because they changed school during the study period.
[3]t tests.

Trichuris trichiura, growth and school performance Evidence from Jamaica: Team Work #1

- Looking at Table 2, would you say there is significant evidence that deworming in these schools in Jamaica is associated with:
  - Improvements in students' performance?

    There is no significant evidence of any relationship between deworming and students' performance. Difference in post-RCT means score in arithmetic (-0.1), spelling (-0.7) and reading (-0.7) are not statistically different from zero (and besides, have the opposite sign: treatment lower than placebo!)
  - Increases in school attendance?

    There is no significant evidence either supporting the hypothesis that deworming is associated with school attendance. Difference in post-RCT mean attendance (-2.0) are not statistically different from zero (and have the opposite sign: treatment lower than placebo!)
  - Better anthropometric measures (height-for-age and body/mass index)?

    There is no significant evidence associating deworming to height-for-age or body mass index. In the case of body-mass index there is even a difference between baseline BMI (higher for placebo group) that is statistically significant at 10% level. In any case, post-RCT BMI are not significantly different from zero.

Trichuris trichiura, growth and school performance Evidence from Jamaica: Team Work #2

- When researchers switch from comparing means to regression analysis, what is the sample regression function with interactions they are running?

Regression #1 and #3 (only changes the Yr_i, residual spelling scores in the first one, residual body mass index in the third one)

$$Yr_i = \beta_0 + \beta_1 Treatment\ +\ \beta_2 AscarisInt +\ \beta_3 Treatment * AscarisInt + \beta_4 age + \beta_5 Female + \beta_6 socio + \beta_7 school + e_i$$

Regression #2 (Yr_i, residual school attendance)

$$Yr_i = \beta_0 + \beta_1 Treatment\ +\ \beta_2 Stunting +\ \beta_3 Treatment * Stunting + \beta_4 age + \beta_5 Female + \beta_6 socio + \beta_7 school + e_i$$

Trichuris trichiura, growth and school performance Evidence from Jamaica: Team Work #2

- What additional information do the interactions provide in this specification?

The interactions allows to determine if treatment was more effective (or effective at all) for children that already infected or stunted. The results (Table 3) suggest:

Spelling residual scores and intensity of infection:

- Deworming treatment and intensity of infection seem to be unrelated to spelling scores, although there is some evidence that treatment is more effective among the most intensely infected…

Body mass index and intensity of infection:

- Deworming treatment is positively and significantly associated to body mass index, although the impact is higher at low levels of infection intensity

School attendance:

- Deworming treatment has significant and positive impacts on school attendance for stunted children, the large the degree of stunting, the higher the impacts

TABLE 3

Results of the multiple regression analyses showing the significant treatment-by-stunting and treatment-by-intensity interactions for Jamaican school children, after controlling for the covariates[1]

	b^2	P
Spelling		
Treatment	−0.1	0.62
Intensity	−0.3	0.57
Treatment × intensity	1.6	<0.05
School attendance		
Treatment	−3.2	0.08
Stunting[3]	−6.8	0.01
Treatment × stunting	9.9	0.007
Body mass index		
Treatment	0.12	0.003
Intensity[4]	0.13	0.15
Treatment × intensity	−0.35	0.009

[1]Covariates included age, gender, socioeconomic status, intensity of *Ascaris* infection and school attended.
[2]*b* is the unstandardized regression coefficient.
[3]Defined as height-for-age Z-score <−1.
[4]Defined as *Trichuris* infection intensities above and below 7000 eggs per gram of stool.

Trichuris trichiura, growth and school performance
Evidence from Jamaica: Threats to Internal Validity #1

- **Threat #1: Unbalanced Groups**
  - If the treatment and controls group are different to begin with, the estimated treatment effect is likely to be biased
  - Are the treatment and control groups balanced here (Table 1)?

TABLE 1

Baseline characteristics of the treatment and placebo groups of Jamaican school children

	Group		
	Treatment (*n* = 206)	Placebo (*n* = 201)	Significance *P*
Age,[1] *y*	9.2 ± 1.2	9.2 ± 1.3	0.63
Gender, % *boys*	52	47	0.35
Residence, % *Kingston*	35	33	0.77
Ascaris infection			
% Infected	42	50	0.13
Light	11	10	
Moderate	25	29	
Heavy	7	11	
Socioeconomic index[1]	5.9 ± 2.2	6.1 ± 2.2	0.28
Trichuris intensity,[2,3] *epg*	2421 (1200–25,458)	2667 (1200–41,733)	0.06
Anemia,[4] % *hemoglobin* <*110 g/L*	10	16	0.20
Low ferritin,[4] % <*12* µ*g/L*	6	9	0.39

[1]Means ± SD.
[2]epg = eggs per gram of stool.
[3]Values are medians, with the range in parentheses.
[4]Blood samples were obtained from only 264 children.

Trichuris trichiura, growth and school performance Evidence from Jamaica: Threats to Internal Validity #2

- **Threat # 2: Imperfect compliance**
- In practice, the researcher may not be able to enforce that the treatment protocol is followed
- There are two forms of imperfect compliance:
  - **No-shows**: Not all the treatment group members get treated (i.e. take-up rate <100%).
  - **Crossovers**: Some control group members get treated.

Trichuris trichiura, growth and school performance Evidence from Jamaica: Threats to Internal Validity #2

- Suppose some children assigned to the treatment group never took the deworming drug. Will the mean difference in table 2 be larger or smaller than the true effect of taking deworming drugs?

  Not giving the drug to students on the treatment group will result in underestimating the impacts of deworming.

- Suppose some children assigned to the control group ended up taking deworming drugs. Will the mean difference in table 2 be larger or smaller than the true effect of taking deworming drugs?

  Having people in the control group taking the deworming drugs will result in underestimating the impacts of deworming.

- Summary: no-shows and crossovers lead to **under** estimation of the true treatment effects.
  - Jamaica study: no direct evidence of non-shows and crossovers.
  - We will learn statistical solution (instrumental variables) to estimate the treatment effect even in the presence of no-shows and crossovers.

Trichuris trichiura, growth and school performance Evidence from Jamaica: Threats to Internal Validity #3

- **Threat # 3: Attrition**
- Attrition – or non-response – means that some individuals disappear from the final sample. This may lead to an attrition bias.
  - If attrition is not correlated with treatment status, this does not create bias (but smaller sample size reduces precision)
  - If attrition is correlated with treatment status (i.e. those who left treatment group differ from those who left control group), this creates a bias i.e. attrition bias

Trichuris trichiura, growth and school performance Evidence from Jamaica: Threats to Internal Validity #3

- In the Jamaica study, 10 children from the treatment group and 5 children from the control group changed schools and thus were not in the final sample. Is this a threat to the internal validity of the experiment? Under what conditions is the answer yes?

  Only if attrition is correlated with treatment status will it pose an internal validity threat.

  If those who left treatment in the treatment group were somehow more likely to suffer from worms (i.e. they left the school because they live in rural areas and had to walk a lot), this would lead to a situation where the impacts of deworming will be overestimated.

Trichuris trichiura, growth and school performance Evidence from Jamaica: Threats to Internal Validity #3

- **Threat #4: Spillover Effects**
- If the treatment affects outcomes of control group members indirectly, there are spillover effects.
- We will **underestimate** estimate the true treatment effect if some control group members are affected indirectly by the treatment.
- Could there be spillover effects in the Jamaica deworming experiment?

This particularly type of infections seem to be contagious, that poses a problem because: a) we miss externality benefits from treatment-group to control group, and b) lower contagion within the treatment group might leads to underestimating the impact of deworming

Other papers correct for this possibility by randomizing at the school level (not at the student level as in Jamaica).

WORMS: IDENTIFYING IMPACTS ON EDUCATION AND HEALTH IN THE PRESENCE OF TREATMENT EXTERNALITIES

BY EDWARD MIGUEL AND MICHAEL KREMER[1]

Intestinal helminths—including hookworm, roundworm, whipworm, and schistosomiasis—infect more than one-quarter of the world's population. Studies in which medical treatment is randomized at the individual level potentially doubly underestimate the benefits of treatment, missing externality benefits to the comparison group from reduced disease transmission, and therefore also underestimating benefits for the treatment group. We evaluate a Kenyan project in which school-based mass treatment with deworming drugs was randomly phased into schools, rather than to individuals, allowing estimation of overall program effects. The program reduced school absenteeism in treatment schools by one-quarter, and was far cheaper than alternative ways of boosting school participation. Deworming substantially improved health and school participation among untreated children in both treatment schools and neighboring schools, and these externalities are large enough to justify fully subsidizing treatment. Yet we do not find evidence that deworming improved academic test scores.

KEYWORDS: Health, education, Africa, externalities, randomized evaluation, worms.

Takeaways

- To assess the validity of any study, we should focus on both internal and external validity
- Randomization is a particularly useful way of improving internal validity
- We should still consider whether any of the problems discussed today (unbalanced groups, imperfect compliance, attrition, spillovers) might create bias in our estimates
- In particular, we should think hard about whether the experimental design used correctly mimics the public policy measures that would be enacted in reality

In preparation for Monday´s class with Lant Pritchett

- Randomized experiments were once considered the "gold standard" for estimating causal effects and all World Bank programs were subject to a "specific program evaluation" (RCT)
- RCTs absorbed much of the funds available for development research and compromised funding for other disciplines addressing more comprehensive questions (with potentially larger impacts)
- Context matters: In trying to adapt every policy/program to the specifics of each context, we give studies lots of internal validity but compromise entirely external validity

Vocabulary

- Randomized experiments
- Selection bias
- Internal and external validity (review)
- Threats to internal validity:
  - Unbalanced groups
- Imperfect compliance
  - No shows
  - Cross-overs
- Attrition
- Attrition bias
- Spillover effects (in the context of experiments)