RANDOMIZED TRIALS

Technical Track Session II

Phillippe Leite
The World Bank

These slides were developed by Christel Vermeersch and modified by Phillippe Leite for the purpose of this workshop

Randomized Trials

- How do researchers learn about counterfactual states of the world in practice?
- In many fields, evidence about counterfactuals is generated by randomized trials or experiments. (E.g. medical research)
- Under certain conditions, randomized trials ensure that outcomes in the comparison group really do capture the **counterfactual** for a treatment group.

Randomization for causal inference

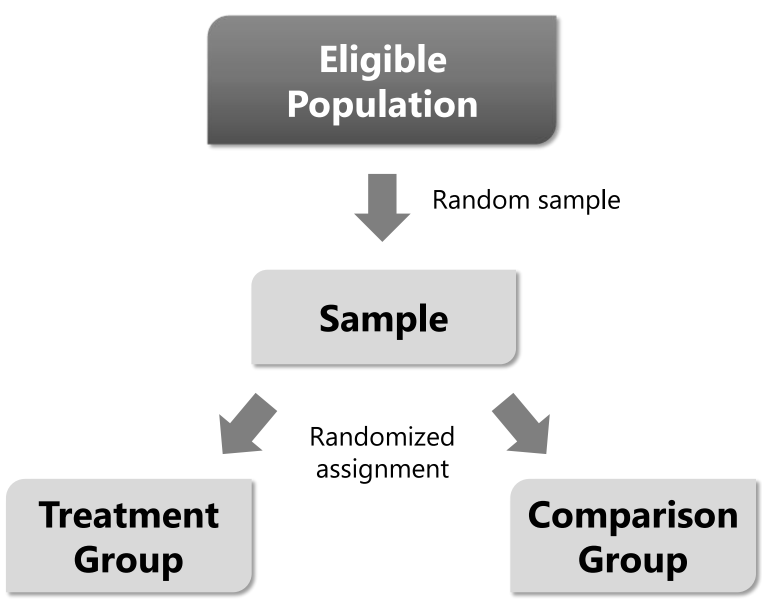
Statisticians recommend a formal two-stage randomization model:

First stage

A random sample of units is selected from a defined population.

Second stage

This sample of units is randomly assigned to treatment and comparison groups.



Why two stages of randomization?

First stage — For external validity

I.e. ensure that the results in the sample will represent the results in the population within a defined level of sampling error

Second stage — For internal validity

I.e. ensure that the observed effect on the dependent variable is due to the treatment rather than to other confounding factors

- Remember the **two conditions?**

Two-Stage Randomized Trials

In large samples, two-stage randomized trials ensure that:

$$[\bar{Y}_1 | D=1] = [\bar{Y}_1 | D=0]$$

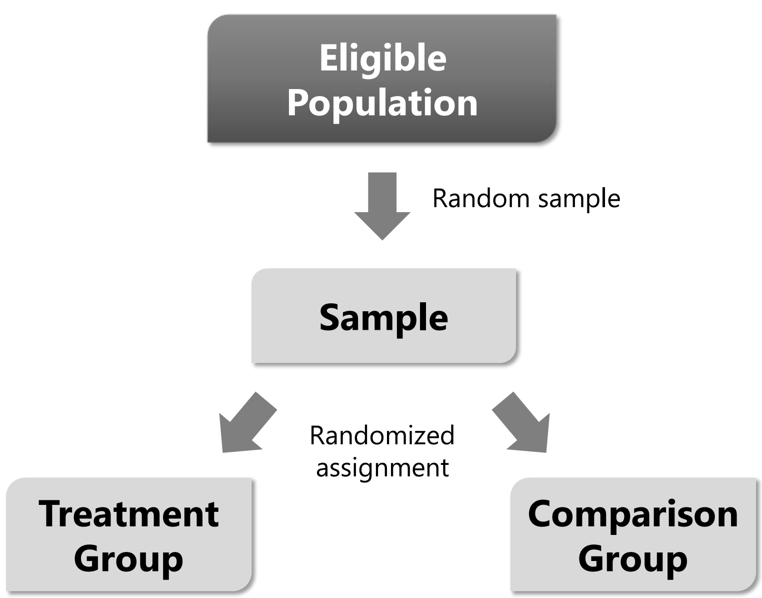
and

$$[\bar{Y}_0 | D=1] = [\bar{Y}_0 | D=0]$$

Why is this true?

The law of large numbers

- When you randomly draw a sample of units from a large population, and
- when the number of units you draw gets large,
- the average of any characteristic of your sample will tend to become closer to the expected value.
- If the number of units in your sample grows, on average the sample will look like its original population.



Two-Stage Randomized Trials

In large samples, two-stage randomized trials ensure that:

$[\bar{Y}_1 | D=1]=[\bar{Y}_1 | D=0]$ and $[\bar{Y}_0 | D=1]=[\bar{Y}_0 | D=0]$

Thus, the estimator:

$$\hat{\delta} = [\ \hat{\bar{Y}}_1 \mid D = 1] - [\ \hat{\bar{Y}}_0 \mid D = 0\ ]$$

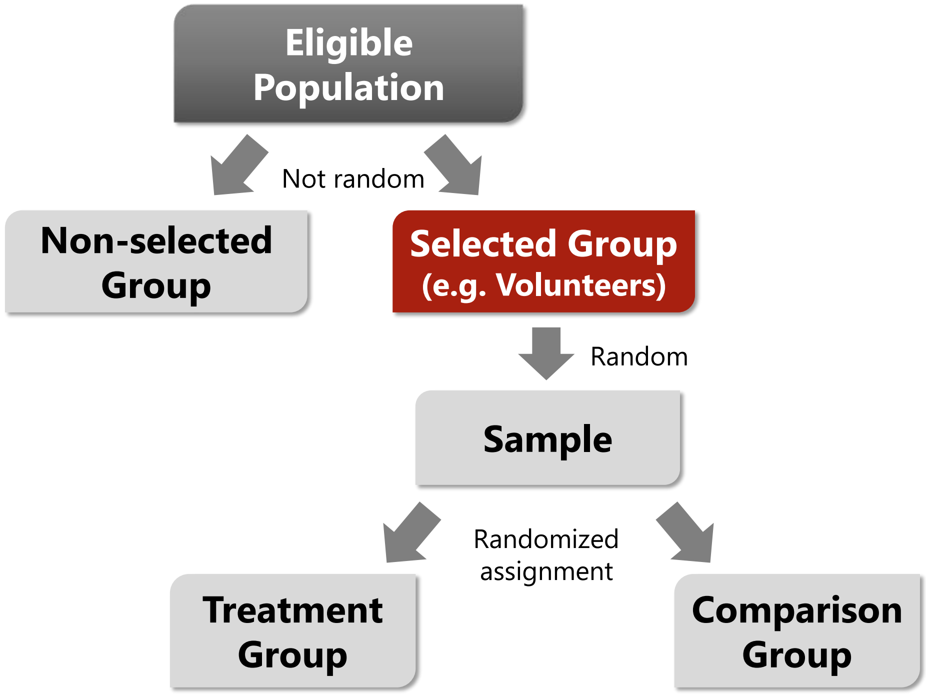
consistently estimates the Average Treatment Effect *(ATE)*

Population vs. Selected group?

If the randomization takes place on a selected group of units

we'll be estimating what??

The treatment effect on that selected group of units!



Randomized assignment: How?

- If you want to assign 50% of the sample to treatment and comparison: flip a coin for each person.
- If you want to assign 40% of the sample to the treatment group, then roll a dice for each person. A 1 or a 2 is treatment; a 3, 4, 5 or a 6 is comparison.
- Other percentages: Let Excel give each unit a random number. Decide how many units will be in the treatment group *(call this X)*. Assign the X units that get the highest numbers to the treatment group.

Example

Computers for Education

Barrera and Linden (2009)

Randomized Trial of in Colombia, Latin-America

- Program activities:
  - Re-furbish computers donated by private firms and installs them in public schools.
  - Train teachers in the pedagogic uses of computers with the help of a local university.
- 2006: 97 schools were subject to a randomization,
  - 48 of them received computers.
  - 49 did not received computers.
- 2008: Follow-up survey

Step 1: Evaluation question

What is the impact of intervention ***D*** on variable of interest ***Y***?

Intervention package *D*

$$D = \begin{cases} 1 & \text{If school selected for installation of computers + teacher training} \\ 0 & \text{Otherwise} \end{cases}$$

Variable of interest *Y*

- Student learning (final outcome)
- Classroom practice (intermediate outcome)
- Number of teachers trained (output)
- Number of working computers in schools (output)

HUMAN DEVELOPMENT NETWORK

Step 2: Verify balance

Goal

Check that characteristics are balanced between treatment and comparison groups.

Regression

$$Y_{ij} = \beta_0 + \beta_1 T_j + \varepsilon_{ij}$$

i	Student i
j	School j
T_j	Dummy variable for whether school j was assigned to treatment in the randomization process

Method

OLS with standard errors clustered at the school level.

HUMAN DEVELOPMENT NETWORK

Verify balance of test scores

Normalized Test Scores	Treatment Average	Comparison Average	Difference
Language	0.08 (0.10)	0.00 (0.08)	0.08 (0.13)
Math	0.04 (0.08)	0.00 (0.08)	0.04 (0.11)
Total	**0.08** **(0.11)**	**0.00** **(0.10)**	**0.08** **(0.15)**

Verify balance of demographic characteristics

Demographic Characteristics	Treatment Average	Comparison Average	Difference
Gender	0.50 (0.02)	0.52 (0.02)	-0.027 (0.03)
Age	11.79 (0.27)	11.55 (0.36)	0.24 (0.45)
Nr parents in household	1.60 (0.02)	1.63 (0.02)	-0.033 (0.03)
Nr siblings	3.71 (0.22)	3.99 (0.20)	-0.286 (0.30)
Receives allowance	0.76 (0.02)	0.73 (0.03)	0.03 (0.03)
Nr friends	17.91 (1.91)	16.12 (1.15)	1.79 (2.22)
Hours of work	5.95 (0.40)	7.09 (0.73)	-1.136 (0.83)

Step 3: Estimate impact

In Words

Compare the average *Y* for the treatment group with the average *Y* for the comparison group.

In a Table

% of questions responded correctly	Treatment Average	Comparison Average	Simple Difference
Spanish	0.42	0.40	0.02
Math	0.24	0.23	0.01

In a Regression

$$Y_{ij} = \beta_0 + \beta_1 T_j + \varepsilon_{ij}$$

- Cluster at the school level.
- This is where you get the standard errors from.

Step 3: Findings

% of questions responded correctly	Treatment Average	Comparison Average	Simple Difference
Spanish	0.42 (0.01)	0.40 (0.01)	0.02 (0.02)
Math	0.24 (0.02)	0.23 (0.01)	0.01 (0.02)

Step 4: Estimate impact, multivariate

In Words

- Compare the average *Y* for the treatment and comparison groups, and
- add controls for baseline characteristics.

In a Multivariate Regression

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 X_{ij} + \varepsilon_{ij}$$

X_{ij} is a vector of baseline characteristics for student *i* in school *j*

Step 5: Findings

% of questions responded correctly	Treatment Average	Comparison Average	Simple Difference	Difference with controls (multivariate)
Spanish	0.42 (0.01)	0.40 (0.01)	0.02 (0.02)	0.02 (0.02)
Math	0.24 (0.02)	0.23 (0.01)	0.01 (0.02)	0.01 (0.02)

Output variable	Treatment Average	Control Average	Difference
Nr of Computers at School	13.38 (1.28)	5.10 (0.75)	8.281*** (1.49)
% of Teachers trained	0.95 (0.03)	0.08 (0.04)	0.865*** (0.05)
Schools treated	0.96 (0.03)	0.04 (0.03)	0.918*** (0.04)

Step 5: Findings

- Little effect on students' test scores
  Across grade levels, subjects, and gender.
- Why?
  Computers not incorporated into educational process.

More thoughts on Randomized Trials

Potential Caveats

- Non-compliance
  - Not all units assigned to the treatment will actually receive the treatment (non-compliance)
  - Some units assigned to comparison may still receive treatment (non-compliance)
  - For limited non-compliance: use instrumental variables (see session IV)
- Attrition
  - We may not be able to observe what happens to all units.
  - Careful for differential attrition!

More Potential Caveats

- Hawthorne effect
  - Just observing units makes them behave differently…
  - Tends to disappear over time.
- John Henry effect

  The "comparisons" work harder to compensate.

Level of randomization

- District, village, family, child?
- Level of randomization determines the power of the trial:

  More units in the treatment and comparison groups → Estimate of difference between T and C becomes more precise.

- The lower the level:
  - the more potential for contamination –children in a class, families in a village- and
  - the harder to administer the program.
- In General: Choose the lowest level that is still administratively feasible.

HUMAN DEVELOPMENT NETWORK

Randomized vs. Non-Randomized Trials

- Randomized experiments
  - Assumptions play a minor role
  - Or no role at all when testing the hypothesis of "no treatment effect"
  - In the absence of difficulties such as noncompliance or attrition…
- Non-randomized methods
  - Requires strong assumptions in order to validate the counterfactual
  - Attrition is equally a problem as in randomized trials.

HUMAN DEVELOPMENT NETWORK

References

- Rosenbaum, Paul (2002): *Observational Studies*, Springer. Chapter 2.
- Cochran, W. G. (1965): *"The planning of observational studies of human populations"*, Journal of the Royal Statistics Association Series A 128, pp. 134-155, with discussion.
- Angrist, J., E. Bettinger, E. Bloom, E. King and M. Kremer (2002): *"Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment"*, American Economic Review, 92, pp. 1535-58.
- Angrist, J. and V. Lavy (2002): *"The Effect of High School Matriculation Awards: Evidence from Randomized Trials"*, NBER Working Paper.
- Barrera-Osorio, F. and L. Linden (2009): *"The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia"*, World Bank Policy Research Working Paper WPS4836.

HUMAN DEVELOPMENT NETWORK

HUMAN DEVELOPMENT NETWORK

SIEF

Thank You

Q & A