AREC 345: Global Poverty & Economic Development

Lecture 11:

The Randomization Revolution

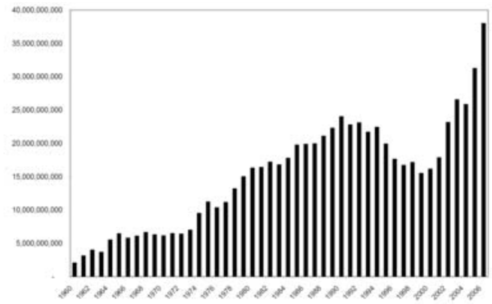
Professor: Pamela Jakiela

Department of Agricultural and Resource Economics
University of Maryland, College Park

Impact Evaluation Terminology

Does Aid Work?

Aid to Africa in Constant 2006 Dollars:



The fundamental problem of causal inference:
Can't know the counterfactual — what would have happened without aid

What Is an Impact Evaluation?

Goal: measure **causal** impacts of policy on participants

- We did *A*; as a result, *B* happened
- *A* is a policy or intervention
- *B* is what we hope to impact
- Examples:
  - We gave out insecticide-treated bednets, and fewer children got sick with malaria as a result
  - We distributed free lunches in elementary schools, and school attendance went up as a result

Establishing Causality

Goal: measure **causal** impacts of policy on participants (or eligibles)

- We want to be able to say B happened because of A
  - We need to rule out other possible causes of B
- If we can say this, then we can also say: if we did A again, we think that B would happen there as well

In an ideal world (research-wise), we could clone each program participant and observe the impacts of our program on their lives

 vs. 

Establishing Causality

In an ideal world (research-wise), we could clone each program participant and observe the impacts of our program on their lives

 vs. 

What is the impact of giving Lisa a book on her test score?

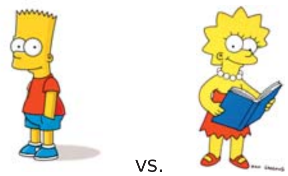
$$\text{Impact} = \text{Lisa's score with a book} - \text{Lisa's score without a book}$$
$$= E\left[Y_{1i}|P_i = 1\right] - E\left[Y_{0i}|P_i = 0\right]$$

In the real world, we either observe Lisa with a book or without

- We never observe the **counterfactual**

Establishing Causality

To measure the causal impact of giving Lisa a book on her test score, we need to find a comparison group that did not receive a book



Our estimate of the impact of the book is then the difference in test scores between the **treatment group** and the **comparison group**

- Impact = Lisa's score with a book - Bart's score without a book

As this example illustrates, finding a good comparison group is hard

Defining the Counterfactual

To do impact evaluation, we need to know what would have happened to every participant i if he/she hadn't participated in the program

- We call this the **counterfactual**

Of course, we can't actually clone our participants and see what happens to the clones if they don't participate in the program

- Instead, we estimate the counterfactual using a **comparison group**

The comparison group needs to:

- Look identical to the treatment group prior to the program
- Not be impacted by the program in anyway

YOU CANNOT HAVE A GOOD IMPACT EVALUATION WITHOUT A CREDIBLE, CONVINCING COMPARISON GROUP

The Moving Parts of an Impact Evaluation

A policy or program of interest (aka the "**treatment**")

- $P_i = 1$ if individual/community/unit i participated in the program
- $P_i = 0$ otherwise

The **evaluation sample** comprises all the individuals/units that we study

- The **treatment group**: a group of people for whom $P_i = 1$
- The **comparison group**: a group of people for whom $P_i = 0$

The Moving Parts of an Impact Evaluation

The **outcome of interest**: the dependent variable in our analysis

- Something that we care about
- Something that we expect to be impacted by the treatment

An impact evaluation compares values of the outcome of interest in the treatment group to values in the comparison group

- We attribute the difference to the impact of treatment

Estimated impact of the program: $E\left[Y|P_i = 1\right] - \left[Y|P_i = 0\right]$

The Moving Parts of an Impact Evaluation

Two potential outcomes for each individual, community, etc:

$$\text{Potential outcome} = \begin{cases} Y_{0i} & P_i = 0 \\ Y_{1i} & P_i = 1 \end{cases}$$

The problem: we only observe one of Y_{1i} and Y_{0i}

- Each individual either participates in the program or not
- The causal impact of program (P) on i is: $Y_{1i} - Y_{0i}$

We observe i's actual outcome:

$$Y_i = Y_{0i} + \underbrace{(Y_{1i} - Y_{0i})}_{\text{impact}} \cdot P_i$$

False Counterfactuals

False Counterfactuals

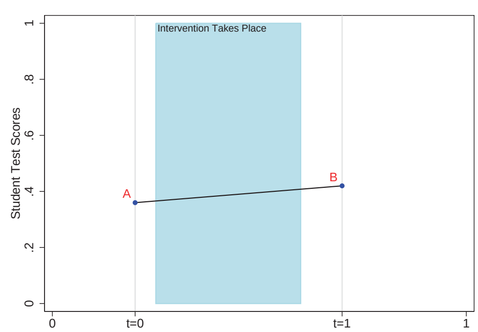
Two types of **false counterfactuals**:

- Before vs. After Comparisons
- Participant vs. Non-Participant Comparisons

Extremely strong (read: completely unreasonable) assumptions are required to make either of these impact evaluation approaches credible

Before vs. After Comparisons

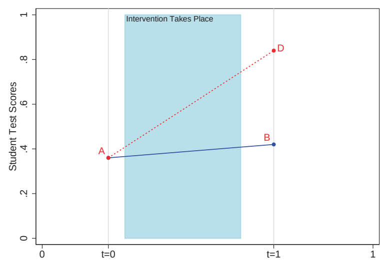
Impact of the program: $B - A$?

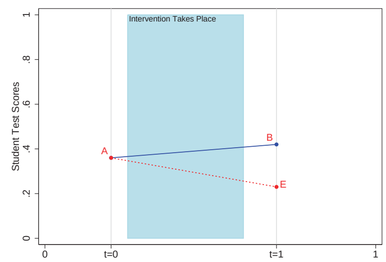


Before vs. after analysis assumes test scores would not have changed between $t = 0$ (pre) and $t = 1$ (post) in the absence of the program

Before vs. After Comparisons

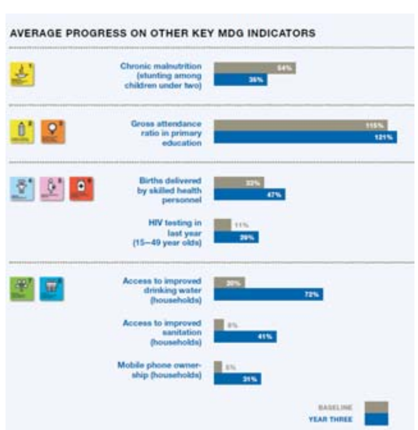
What if things change in the absence of the intervention?



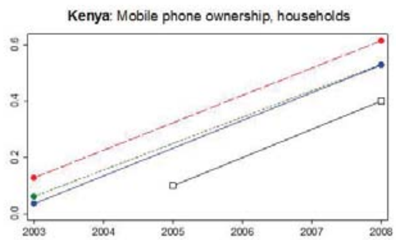


Before vs. After Comparisons: Example

The perils of pre vs. post analysis should be obvious. . .
But sometimes pre vs. post analysis still happens to smart people



Before vs. After Comparisons: Example

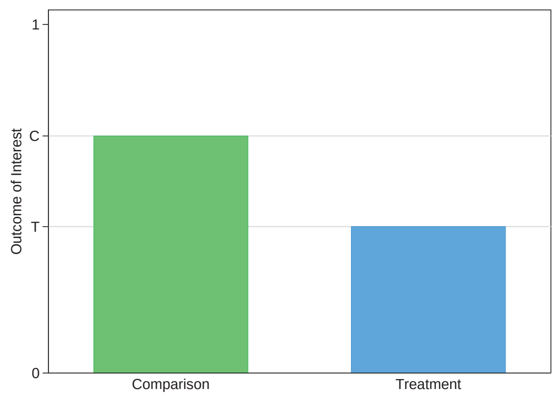


Clemens and Demombynes (2010) compare changes in mobile phone ownership in Bar Sauri (rectangles) to trends in Kenya (red), rural Kenya (green), and rural areas in Nyanza Province (blue)

- The problem is obvious: before vs. after analysis assumes that there is no time trend in mobile phone ownership in Kenya

Participants vs. Non-Participants

What if we compare outcomes post-intervention?



Can we estimate the impact of the program by calculating $T - C$?

Participants vs. Non-Participants

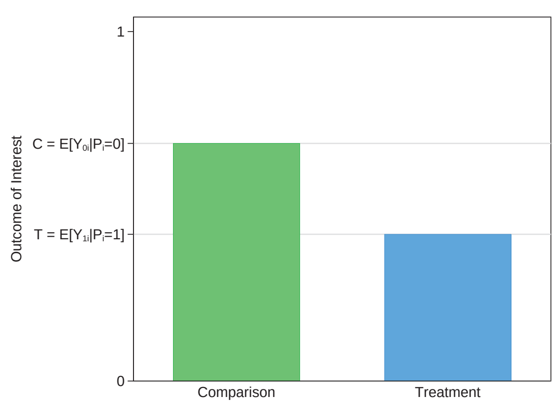
Wait!! Why didn't the comparison group participate?

- Maybe they were already better off, and didn't need help
- Maybe they were worse off; maybe they didn't know bout the program, or could not get organized to apply for the program
- Those who aren't eligible and those who choose not to participate may have different outcomes in the absence of the program
- This is **selection bias**

Remember: the true causal impact of program on i is: $Y_{1i} - Y_{0i}$

- When we compare participants to non-participants, we are assuming that, in the absence of the program, outcomes in program schools would have looked like outcomes observed in the comparison schools

Participants vs. Non-Participants



Participants vs. Non-Participants

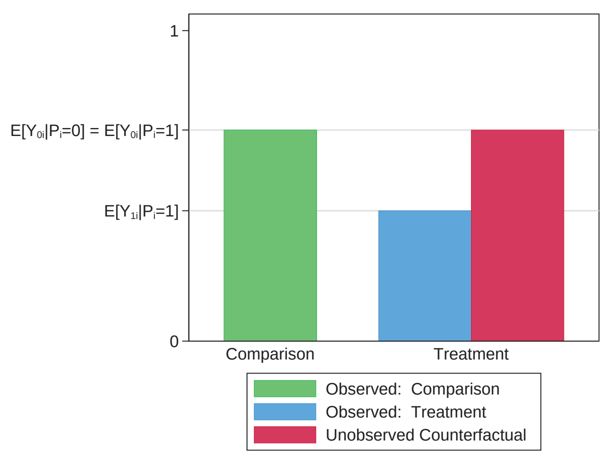
Our estimate of the impact of treatment is:

$$\begin{aligned}
\text{Impact} &= E\left[Y_i | P_i = 1\right] - E\left[Y_i | P_i = 0\right] \\
&= E\left[Y_i | P_{1i} = 1\right] - E\left[Y_{0i} | P_i = 0\right] \\
&= \underbrace{E\left[Y_{1i} | P_i = 1\right] - E\left[Y_{0i} | P_i = 1\right]}_{\text{program impact}} + \underbrace{E\left[Y_{0i} | P_i = 1\right] - E\left[Y_{0i} | P_i = 0\right]}_{\text{selection bias}}
\end{aligned}$$

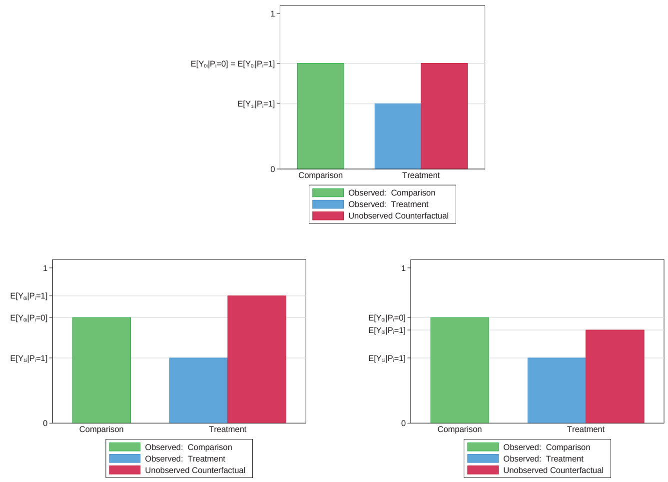
When $E\left[Y_{0i} | P_i = 1\right] - E\left[Y_{0i} | P_i = 0\right] \neq 0$, we have a problem.

- The treatment and comparison groups would not have looked the same in the absence of the program. Why might this occur?

Participants vs. Non-Participants



Participants vs. Non-Participants



Summary: False Counterfactuals

Before vs. After Comparisons:

- **Compares:** same individuals/communities before and after program
- **Drawback:** things (besides the program) may happen over time

Participant vs. Non-Participant Comparisons:

- **Compares:** participants to those not in the program
- **Drawback:** selection bias — why aren't they in the program?

Randomization to the Rescue

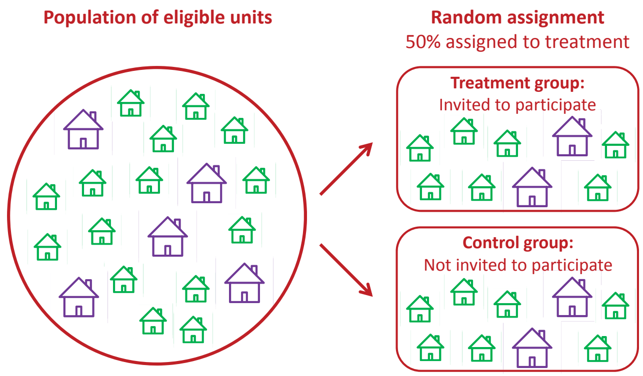
Randomization Solves the Selection Problem

In randomized evaluations, treatment and comparison groups are random samples of the eligible population of potential participants

- Every eligible unit (i.e. household, school, clinic, village, etc.) has an equal chance of being randomly assigned to the program
- The treatment group is chosen at random (by random chance), often quite literally through some sort of a public lottery or drawing
- The treatment group is a random sample of the evaluation sample
  - We can expect both the treatment group and the control group to look like the underlying population in terms of all characteristics

In the absence of the program, the treatment group and the comparison group would have looked the same in terms of all possible outcomes

Randomization Solves the Selection Problem



Randomization Solves the Selection Problem

Cross-sectional analysis of observational data leads to selection bias:

$$\text{Impact} = E\left[Y_i | P_i = 1\right] - E\left[Y_i | P_i = 0\right]$$

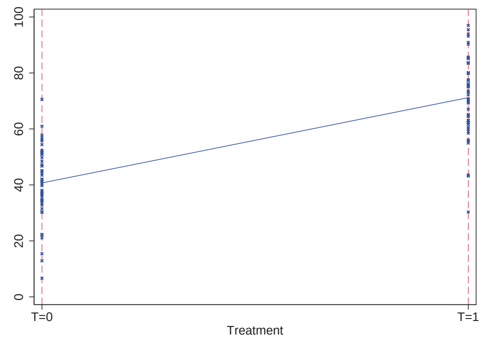
$$= \underbrace{E\left[Y_{1i} | P_i = 1\right] - E\left[Y_{0i} | P_i = 1\right]}_{\text{program impact}} + \underbrace{E\left[Y_{0i} | P_i = 1\right] - E\left[Y_{0i} | P_i = 0\right]}_{\text{selection bias}}$$

The key feature of of randomization is that (by construction) it breaks the correlation between P_i and Y_{0i} (and everything else)

$$\Rightarrow \quad E\left[Y_{0i} | P_i = 1\right] = E\left[Y_{0i} | P_i = 0\right] = E\left[Y_{0i}\right]$$

- Clearly, when this is true, selection bias disappears
- Any difference in mean outcomes can be attributed to the program, because the program is the only difference between groups

Regression Analysis of Randomized Evaluations



Independent variable can only take on two values:

So, all data points fall on one of two vertical lines: $T_i = 0$ or $T_i = 1$

Regression Analysis of Randomized Evaluations

The regression gives as a prediction of Y (the "**expected value**")

$$E[Y_i] = a + b \cdot T_i$$

We can also think of this as the average Y for a given value of T

Since T is a dummy variable and can only take on two values, there are only two possible predicted values of Y — what are they?

Regression Analysis of Randomized Evaluations

$E[Y_i] = a + b \cdot T_i$, so there are two possible predicted values:

- $E[Y \mid T_i = 0] =$
- $E[Y \mid T_i = 1] =$

$\Rightarrow b = E[Y \mid T_i = 1] - E[Y \mid T_i = 0]$

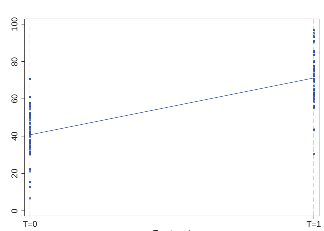
Interpretation of results when independent variable is a dummy:

- Estimate of *a* tells us average of dependent variable when $C = 0$
- Estimate of *b* tells us difference between average of dependent variable when $T_i = 1$ and average of dependent variable when $T_i = 0$

Regression Analysis of Randomized Evaluations

Mean Values of the Outcome Variable

Treatment group average	$a + b$
Control group average	a
Difference	b



$\leftrightarrow$

Dep. Var. = Log GDP	OLS (1)
Treatment	b
	(S.E. of b)
Constant	a
	(S.E. of a)

Study Guide: Key Terms

- impact evaluation
- counterfactual
- treatment group
- comparison group
- evaluation sample
- outcome of interest
- false counterfactuals
- selection bias
- randomization