Experimental Evaluation Design

USAID MENA Advanced MEL Workshop

2024-09-17

Welcome!

- Who we are
- What we do
- How we hope to help you

Objectives of impact evaluation sessions

- Understand the need for impact estimation of USAID activities
- Understand how impact estimation fits into the Agency performance management framework
- Gain practical knowledge about impact evaluation to help USAID staff better manage and support IEs

Benchmarks for success

- Understand selection bias as a fundamental difficulty in identifying a valid counterfactual, and randomization as the most effective means of identifying the counterfactual
- Recognize the difference-in-means and difference-in-differences designs
- Understand the key validity threats to experimental evaluations, and where those threats may be addressed in the management cycle

Level Set

Measuring social benefit

We want to know the causal effect of an activity on its beneficiaries

- Job training on earnings and employment
- Teacher qualifications on student outcomes
- Humanitarian assistance on food security

Identifying a treatment assignment

- We established indicators for treatment assignment D_i and an outcome of interest Y_i
- We established the switching equation $Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$ mapping a potential treatment assignment to a realized outcome
- We re-wrote the switching equation to $Y_i = Y_i^0 + (Y_i^1 - Y_i^0) D_i$ in order to highlight the individual treatment effect term $\delta_i = Y_i^1 - Y_i^0$

The treatment assignment mechanism

- We stressed a distinction between treatment *assignment* and the treatment assignment *mechanism*
- Why is this distinction so important?
- Because many activities we evaluate target specific sub-samples of a broader population
  - The poor / marginalized / conflict-affected
- We can't just compare these participants to a randomly selected member of the population!

Capturing the treatment assignment mechanism

- We start with the individual treatment effect $\delta_i = Y_i^1 - Y_i^0$
- We take the average of all individual treatment effects

$$E[\delta] = E[Y^1 - Y^0]$$

- We take the difference in averages, rather than the average of the differences

$$E[Y^1 - Y^0] = E[Y^1] - E[Y^0]$$

Averages of effects

unit	y0	y1	y1 - y0
1	12	18	6
2	15	13	-2
3	9	22	13
4	20	19	-1
Average	14	18	4

$$E[\delta] = E[Y^1 - Y^0] = E[6, -2, 13, -1] = 4$$

$$E[\delta] = E[Y^1] - E[Y^0] = 18 - 14 = 4$$

Difference-in-means estimator

- Finally, we apply the switching equation to know who is treated and who is not

$$E[\delta] = E[Y^1 - Y^0]$$
$$E[Y^1 - Y^0] = E[Y^1] - E[Y^0]$$
$$E[Y^1] - E[Y^0] = E[Y^1|D = 1] - E[Y^0|D = 0]$$

- This is the *difference-in-means estimator*

Real world data!

unit	y0	y1	y1 - y0
1	?	18	?
2	15	?	?
3	?	22	?
4	20	?	?
Average	17.5	20	2.5

$$E[\delta] = E[Y^1 - Y^0]$$

$$E[Y^1 - Y^0] = E[Y^1] - E[Y^0]$$

$$E[Y^1] - E[Y^0] = E[Y^1|D=1] - E[Y^0|D=0]$$

$$= 20 - 17.5 = 2.5$$

Decomposing the difference in means

- What do we know about the treatment assignment mechanism for these groups?
- To find out, we decompose the *difference-in-means* estimator into the following:

Average Treatment Effect on the Treated (ATT)

$+$

Difference between treatment and control group, before treatment (selection bias)

Decomposing the difference in means

$$E[Y^1|D=1] - E[Y^0|D=0]$$

Average Treatment Effect on the Treated (ATT)

$$+$$

$$E[Y^0|D=1] - E[Y^0|D=0]$$

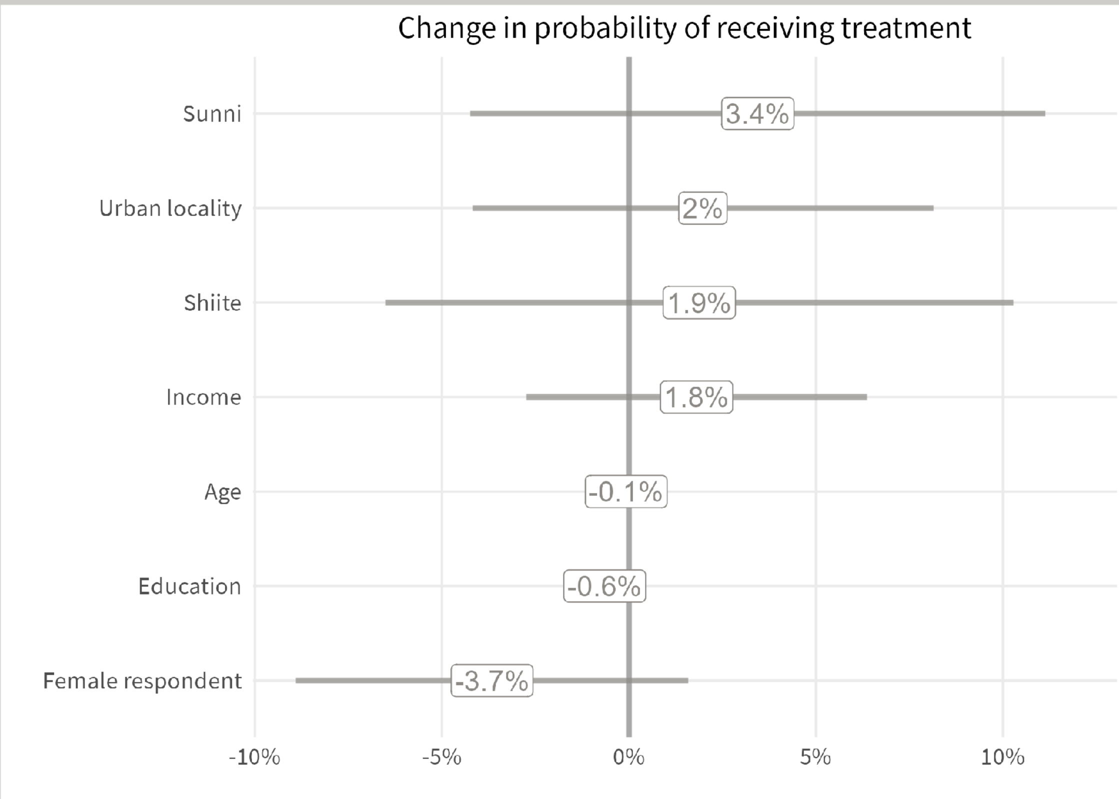
Difference between treatment and control group, before treatment (Selection bias)

Decomposing the difference-in-means

- Why did we just do this?
- To demonstrate that we need to know the treatment assignment *mechanism* in order to know how each of these terms are affected
- Under randomization, treatment assignment D_i is independent of the potential outcomes Y_i^0 and Y_i^1
- This means that the selection bias term is zero - there are no differences between the treatment and control groups before treatment

Observable characteristics under randomization

- Does $E[Y^0|D=1] - E[Y^0|D=0] = 0$?

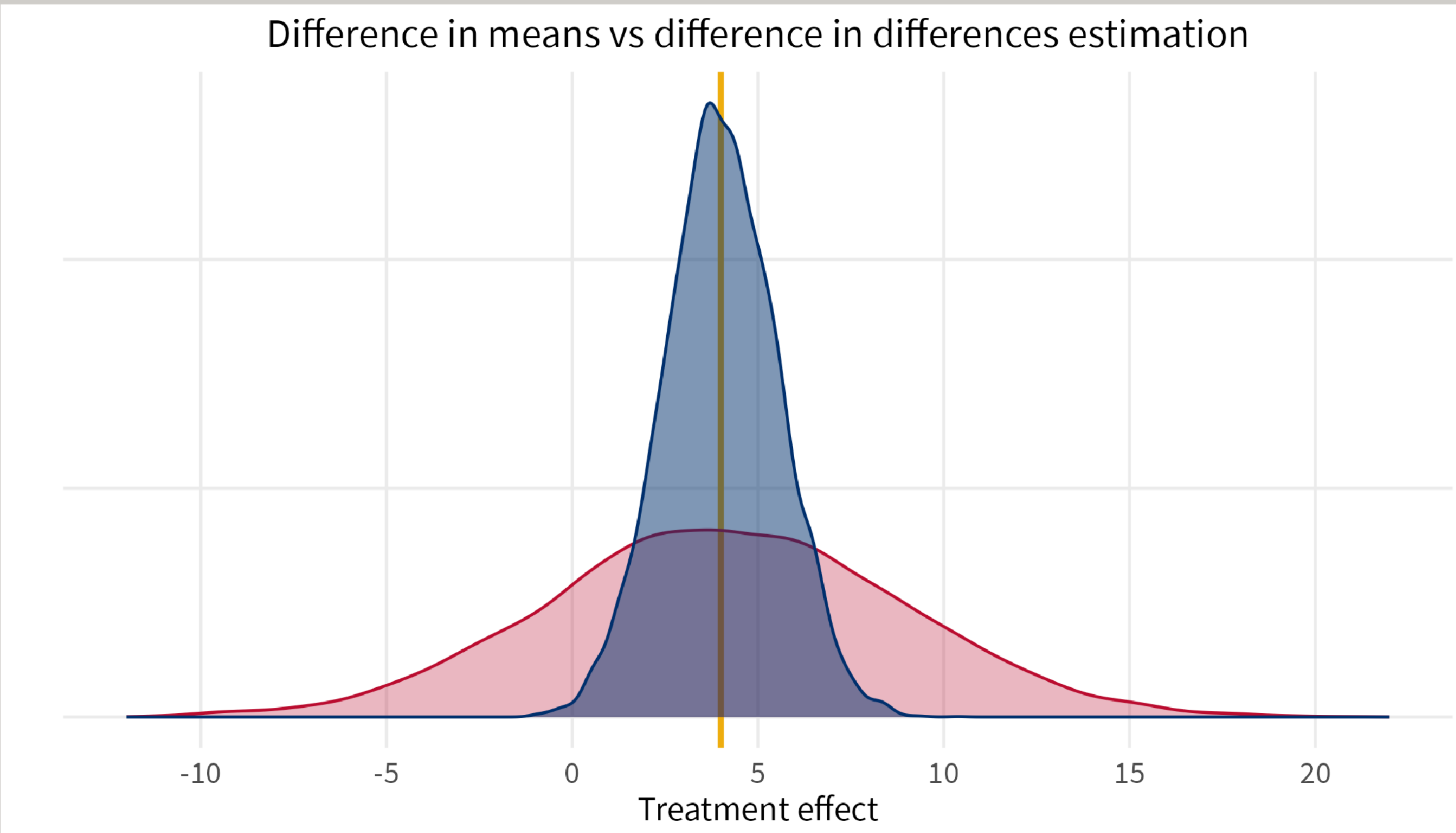


From difference-in-mean to difference-in-differences

- The difference-in-means estimator is easy - all we have to do is compare the post-treatment outcomes
- Is a better design available? Yes - the difference-in-differences (DD) design
- Under difference-in-differences design, we include a pre-treatment measurement of the outcome in addition to a post-treatment measurement

Difference-in-differences

- Including a pre-treatment measure of the outcome improves precision



Difference-in-differences

- Difference in differences also helps detect heterogeneous treatment effects
- Using the DD design, we would be able to discover insights such as whether the intervention is working for those already better prepared to benefit
- Enables stakeholders to determine whether they want the strongest outcomes, or prefer more modest outcomes for beneficiaries who need the intervention the most!

Threats to validity in experimental designs

- Randomization is the preferred method for identifying the counterfactual
- But randomization is vulnerable to threats that could introduce selection bias
- The primary threats to validity of randomization are compliance and attrition
- Other threats to validity include spillovers and measurement error

Threats to validity - compliance

Recall what happens when we apply the switching equation to potential outcomes:

$$E[Y^1] - E[Y^0] = E[Y^1|D=1] - E[Y^0|D=0]$$

- Note that this equation assumes perfect compliance!
- What happens if our treatment assignment is not followed?

Issues with compliance

$E[Y^0|D=1]$ –> Assigned treatment, but not treated

$E[Y^1|D=0]$ –> Not assigned treatment, but treated!

- Most common way to deal with compliance failure is to ignore it
- This is referred to as Intent to Treat (ITT) analysis
- ITT analysis is a conservative estimate of the treatment effect, but is also more policy relevant

Threats to validity - attrition

- If we collect pre-treatment measures of the outcome, what happens if we lose track of some participants at endline?
- If attrition is unrelated to treatment or outcomes, then it adds noise but not bias
- If attrition is correlated to the assignment of treatment or the outcome, we have both noise and bias
- For example, the outcome is household income, and we lose track of households due to shocks to household income

Management points for experimental evaluations

- Conduct logic modeling sessions to diagram the data generating process
- Pay attention to process of randomization. Verify integrity of randomization process.
- Think about ways the assignment of treatment might ‘leak’ via spillover or contamination
- Examine statistical tests of randomization, compliance, and spillover

What if we can't randomize?

- If we can't randomize the assignment of treatment, the next best alternative is to look for ways to approximate 'as-if' randomization
- Stay tuned for quasi-experimental evaluation design!

THANK YOU!