Potential Outcomes

USAID MENA Advanced MEL Workshop
Session 1

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

By the end of this session, participants will be able to:

- Explain the fundamental problem of causal inference

- Explain how impact estimation can be seen as a problem of missing data

- Speculate as to whether causal inference is a two-body problem or a three-body problem

- Relive unpleasant childhood memories of having to learn algebra

The Fundamental Problem

Measuring social benefit

We want to know the causal effect of a project on its beneficiaries

- Job training on earnings and employment

- Teacher qualifications on student outcomes

- Humanitarian assistance on food security

Identifying a treatment assignment

Consider an indicator for a *potential* beneficiary, D_i

D tells us whether there is an activity, or a "treatment"

The subscript i denotes a single individual who is either treated or not treated

D_i = 1 means participation in an activity

D_i = 0 means no participation in an activity

Identifying an outcome

Now consider an indicator for the outcome of a potential beneficiary, Y_i, where i denotes each person or unit under study.

Y_i^1 is the outcome after activity participation ($D_i = 1$)

Y_i^0 is the outcome without the activity ($D_i = 0$)

Note that Y_1 and Y_0 denote *possibilities* for the *same person*, unit i!

Switching across treated and untreated outcomes

We use what is called a 'switching equation' to connect a treatment assignment to a realized outcome

$$Y_i = D_i Y_i^1 + (1 - D_i)Y_i^0$$

(Plug in $D_i = 1$ and $D_i = 0$ and see what you end up with)

Switching across treated and untreated outcomes

We can also write this same equation this way:

$$Y_i = Y_i^0 + (Y_i^1 - Y_i^0)D_i$$

This equation makes it easier to notice our treatment effect, which is $Y_i^1 - Y_i^0$, or the difference between the treated and untreated outcome

We call the difference $Y_i^1 - Y_i^0\ delta$, or δ_i

Again notice that the treatment effect δ_i refers to the *same* individual!

What a conundrum

To recap: the effect of the activity (treatment effect) on person i is the difference between the two potential outcomes

Treatment effect = $Y_i^1 - Y_i^0$, or δ_i

This is the difference in potential outcomes for the *same person*

A person participates in an activity, and then goes back in time and does not participate in the activity

You ask the impossible

But how can one person be both treated and untreated?

In the real world, person i experiences one of the potential outcomes, but not both

If $D_i = 1$, the potential outcome of Y_i becomes Y_i^1 in fact and the potential outcome of Y_i^0 is unobserved

If $D_i = 0$, the potential outcome of Y_i becomes Y_i^0 in fact and the potential outcome of Y_i^1 is unobserved

The fundamental problem of causal inference

This is the fundamental problem of causal inference:

We observe only one outcome, but we need both outcomes to describe the effect of the project

We refer to the outcome that didn't happen as the *counterfactual*, or what would have happened in the absence of the project

The Missing Data Problem

Something is missing

Group	Y_{i1}	Y_{i0}
Treatment	Observed	Counterfactual
Control	Counterfactual	Observed

- Researchers sometimes refer to impact evaluation as a "missing data problem"

- We are missing two pieces of information about what happens with or without the treatment

What do we do now?

How do we estimate the effect of a project, if we cannot observe the same person go through both potential outcomes?

We must compare a person who was treated with a person who was not treated

But, what are the differences between those two people? How do we know that project participation is the only difference between them?

CLIFFHANGER

Tune into the next session for a resolution of the Fundamental Problem of Causal Inference!

Teaser:

- Experimental impact evaluation

- Quasi-experimental impact evaluation

- Prediction via machine learning

- Getting an answer through Artificial General Intelligence (AGI)

End credits

Post-credit scene

The Three-Body Problem