

UNIVERSITY OF ILLINOIS

AT URBANA-CHAMPAIGN

**Final thoughts**



[illinois.edu](http://illinois.edu)

## Overview

- Internal vs. external validity
- Summary of methods
- Creating a control group ex post
- Reflection exercise



UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# External validity



[illinois.edu](http://illinois.edu)

## Internal vs. external validity

- This week we've emphasized internal validity
- For policy, external validity is crucial
  - Rodrik 2008
  - Better policy to charge for malaria bednets or give them out for free? Study in W. Kenya
  - Find those who received them for free resulted in higher prevalence and a higher usage rate

Does this answer the question?



## Bed net study

- Area where social marketers had been very active  
—spread of information
- Bednets offered to women seeking prenatal care  
in clinics
- Experiment supplied bednets to the clinics
- Difference between subsidized price and free was  
very small



## External validity

- We only know that this experiment worked in this setting
- We don't know which contextual factors are relevant
- Conduct experiments in other areas
- Put together systematic reviews of the evidence



## What have we learned this week?

- What is impact assessment?
- Methods
  - Randomized Control Trials
  - Matching
  - Difference-in-Differences
  - Instrumental Variables

Econometric techniques for panel data

Different treatment effects



## Planning an evaluation

- What is the experiment I would run if I could?
- Can I do an experiment?
- Can I get a counterfactual?
- Do I have more than one round of data for both treatment and control?
- Can I match? Can I assume selection was not driven by unobservables?
- Was there some rule or random aspect by which the program was allocated?



# Randomized Control Trials

- Advantages
  - Solve selection problem through random assignment
  - Can be related to economic theory
  - Flexible
- Disadvantages
  - Must be planned in advance
  - Can be expensive



# Randomized Control Trials

- Challenges for the researcher
  - Requires funding for baseline and followup
  - Achieving sufficient power
  - Addressing spillover effects
  - Getting NGO/government agency to randomize
  - And that they actually do randomize!



# Matching

- Advantages
  - Can use existing data
  - Straightforward to do in Stata
  - Can be used with diff-in-diff
  - Can be used to match units for RCT
- Disadvantages
  - Usually does not solve selection problem by itself
  - Assumes that selection is on observables only



## Matching--Challenges for researcher

- Think about how you want to match—within or across villages
- Results can be sensitive to choice of the bandwidth



# Difference-in-Differences

- Advantages
  - Deals with selection problem when related to time-invariant error term
  - Can be done as a t-test or in a regression context
  - Can be combined with matching
- Disadvantages
  - Need a baseline
  - Need a control group
  - Biased if treatment correlated with time varying error term



## DiD—Challenges for the researcher

- Parallel trends assumption
- Need to establish that treatment and control groups were similar and following the same trends
- Can't do this with a statistical test—need to make an argument
  - Data a period before the baseline
  - Graphing, etc.



## Instrumental variables

- Advantages
  - Corrects for selection on time invariant and time variant variables
  - Can do this with just a cross section
- Disadvantages
  - Very difficult to find a valid instrument
  - Need to make a convincing argument about validity
  - IV has to be correlated with treatment, but not with the error term in the outcome regression



## IV—Challenges for researchers

- Finding a convincing instrument
- Was there a random aspect to the policy implementation?
- For CGIAR center researchers especially—you can't let availability of a cool IV drive the work that you do.



## What if you have no comparison group?

- White (2014)—You can use observational data to do impact evaluation
  - Unobservables can be observed with better measurement of trust, risk aversion, social capital
  - Matching can be used to construct a control group
  - Recreate a baseline ex post



## Creating an ex post baseline/control

- White (2014) is optimistic
- Objections to recall data are “hugely overstated”
- Focus on control variables that change little over time (gender, year of birth, race, education)
- Be realistic about what people can remember
- Refer to major events, ask in chronological order
- Ask treatment group same recall questions for validation



## Creating an ex post baseline/control

- Are there data from other sources that can be used that are accurate?
- Satellite data—deforestation
- Government data
  - Tax authority
  - Social programs



## Take a moment to pair and share

- Please pair with somebody who is not from your center, perhaps somebody you haven't talked to yet this week.
- Give one example of how you will apply something that you learned this week when you return to your job
- Give at least one example of how we can build a community of impact evaluation practitioners (Facebook group? Linked in?)



## Final thoughts

- Think about the audience for the impact assessment
- Plan for evaluation from the beginning of the project
  - Need to budget for control group, evaluation
- Qualitative data collection can be very useful
- Impact evaluation is an art as well as a science
  - Choosing appropriate method
  - Demonstrating that assumptions hold

