UNIVERSITY OF ILLINOIS

AT URBANA-CHAMPAIGN

# Survey Challenges and Challenges to Identification


illinois.edu

# Overview

- Internal vs. external validity
- Sampling methods
- Summary of methods
- Reflection exercise

# Internal vs. External Validity

# Internal vs. External Validity

- This week we've emphasized internal validity
- For policy, external validity is crucial
  - Rodrik, 2009. "We Shall Experiment, but How Shall We Learn?"
  - Is it better policy to charge for malaria bed nets 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
- Bed nets offered to women seeking prenatal care in clinics
- Experiment supplied bed nets 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
- How do we build external validity?

# Sampling Methods

# What Do We Mean by Representative?

- The donor wants to know the adoption rate of a new technology.
- You have been asked to conduct a nationally representative survey to determine the adoption rate.
- What qualifies as nationally representative?

# Example: Groundnut Adoption in Tanzania

- Need to measure adoption of new groundnut varieties by surveying a nationally representative sub-sample of the population.
- Do we sample from all households in Tanzania?
- Do we sample from all rural households?
- Do we sample from all rural households in areas that are suitable for adoption?

# Example: Groundnut Adoption in Tanzania

- We may want to weight our sample to ensure it is representative.
- Do we weight based on population?
  - Total people?
  - Percentage of people in a region?
- Do we weight on arable land?
  - Total arable land?
  - Percentage of arable land in a region?
- Do we weight at national or sub-national level?

# Example: Groundnut Adoption in Tanzania

| Region | Villages by groundnut area | Villages by % land to groundnut | Villages by groundnut growers |
|---|---|---|---|
| Shinyanga Region | 27 | 40 | 22 |
| Tabora Region | 25 | 9 | 21 |
| Dodoma Region | 20 | 14 | 14 |
| Rukwa Region | 9 | 11 | 7 |
| Mbeya Region | 9 | 4 | 12 |
| Mtwara Region | 7 | 12 | 9 |
| Mwanza Region | 5 | 14 | 6 |

# Sample design

- E.g. A: 10 households per 300 villages?
  B: 30 households per 100 villages?

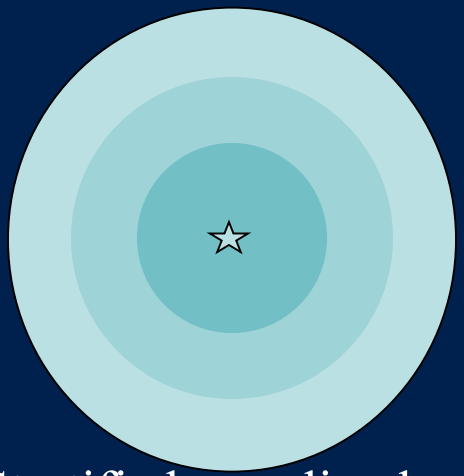
- Option A: good geographic distribution, but cannot say anything about village average
- Option B: lower geographic distribution, but can say something about village averages

...Depends on the question

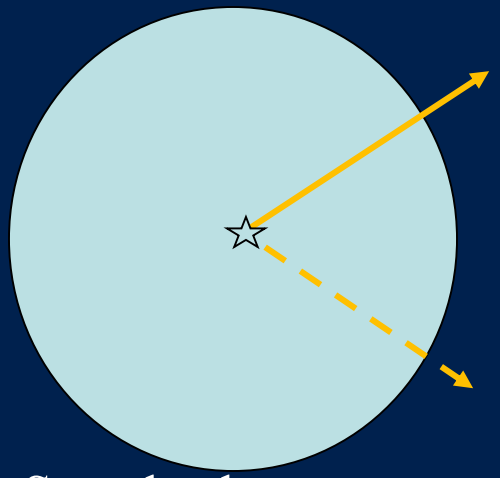
# Sample the population or sample the geography

e.g. Research question: how does distance to agro-dealer affect improved seed adoption?

May not want to randomly sample over households (may not have sufficient spatial variation)



Stratified sampling by distance to market



Sample along a transect

# Sampling Weights

| Name | CG | Asia | Gender (F=1) | Total Research Staff | Sampling Weights |
|---|---|---|---|---|---|
| Nelissa | CropTrust | 0 | 1 | 30 | 30/675 |
| Taj | IFPRI | 1 | 0 | 600 | 600/675 |
| Seng | WorldFish | 1 | 0 | 400 | 400/675 |
| Sven | WorldFish | 0 | 0 | 400 | 400/675 |
| Anurag | CIMMYT | 1 | 0 | 1250 | 1250/675 |
| Silvia | ILRI | 0 | 1 | 650 | 650/675 |
| Dilli | CIMMYT | 1 | 0 | 1250 | 1250/675 |
| Jeff | ICRISAT | 0 | 0 | 820 | 820/675 |
| Average | | 50% | 25% | 675 | |
| Weighted Average | | 65% | 12.6% | | |

# Sampling Weights

- The probability of an observation coming from a specific group within the population to extrapolate from sample to population.
- Easy if only one grouping variable
- May care about multiple dimensions: education level within each CGIAR center
- Can split into population in multiple subsets

# Summary of Methods

# 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
  - Regression discontinuity design

# 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

# RCTs — Challenges for the researcher

- Requires funding for baseline and follow-up
- 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 the 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.

# Regression Discontinuity

- Advantages
  - Unbiased estimate at discontinuity
  - Can deal with selection
  - Eligibles not excluded from treatment
- Disadvantages
  - LATE estimator—may or may not be of interest
  - Power issues—need many observations at threshold
  - Can be sensitive to functional form

# RD — Challenges for researchers

- You need a threshold for eligibility
- Sample size issues around threshold point
- Cutting edge method that is evolving rapidly—need to keep informed about the state of the art
- Agricultural applications—spillovers, lack of strict eligibility requirements for programs
- Eligibility requirements known—people respond to criteria

# 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

# 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