J-PAL Africa Executive Education Course

How to Randomise? (Randomisation Design)

Lecture 4, January 24, 2012

Martin Abel

University of Cape Town
Roadmap to Randomised Evaluations

1. Environment / Context
   - partner
   - sufficient time
   - interesting policy question / theory
   - sufficient resources

2. Theory of Change
   - mechanism of change (log frame)
   - state assumptions
   - identify research hypothesis
   - identify target population
   - identify indicators
   - identify threats to validity

3. Randomization Design
   - Intervention
     - multiple interventions
     - simple program
     - packages
   - Unit of randomization
     - individual
     - cluster design
     - block random.
   - Randomization mechanism
     - encouragement
     - gradual rollout
     - simple lottery
     - rotation design

4. Sufficient Sample Size
   - statistical validity
   - cluster correlation

5. Strategy to Manage Threats
   - Spillovers
   - Discouragement
   - Attrition
   - Political interference

Check on
Revise
**Basic setup of a randomized evaluation**

**Target Population**
- Random Sampling
- External Validity

**Evaluation Sample**
- Baseline Survey
- Random Assignment
- Internal Validity
- Balance Check

**Not in evaluation**

- Treatment group
- Control group
- Endline Survey
- Measure Impact

**Example:** Effect of Extra Text Books

- 5000 Primary Schools in Western Cape
- Random Sample of 200 Schools
- *Measure class size*
- *Randomly assign*
- *Compare class sizes between C and T groups → are they balanced?*
- Treatment schools receive extra text books
- Measure Test Outcomes
Outline

• Intervention Variations
• Unit of randomisation
• Method of randomisation
• Stratification
Basic setup of a randomized evaluation

- **Target Population**
  - Random Sampling

- **Evaluation Sample**
  - Not in evaluation

- **Random Assignment**
  - Baseline Survey

- **Intervention Variations**
  - Treatment group
  - Control group

- **Endline Survey**

- **Unit of Randomization**
- **Randomization Mechanism**
- **Stratification**
Outline

• Intervention Variations
  – Multiple Treatments
  – Cross-cutting Treatments
  – Varying Levels of Treatment

• Unit of randomization
• Method of randomization
• Constraints that effect design
• Stratification
1. Multiple treatments

• What questions does your organization want to have answered?

• Example:
  – Microfinance vs. Financial literacy training
  – Extra Teacher vs. Free text books
  → You can randomise these programmes to multiple treatment groups
1. Multiple Treatment

- **Target Population**
- **Not in evaluation**
- **Evaluation Sample**
- **Random Assignment**
  - Financial Literacy
  - Microfinance
  - Control group
Multiple treatments

- Microfinance
- Financial Lit
- Control
2. Cross-cutting treatments

- Test different components of treatment in different combinations
- Test whether components serve as substitutes or compliments

Example:
- Vouchers to go to school and returns to education
- Microfinance and financial literacy training
- What is most cost-effective combination?

→ Why not always cross-cutting? When should we do it?
2. Cross-cutting treatments

Target Population

Evaluation Sample

Not in evaluation

Random Assignment

Financial Literacy AND Microfinance

Financial Literacy

Microfinance

Control group
3. Varying levels of treatment

• We may want to decide the level / intensity of the program

• Example:
  – Once we know e.g. that financial literacy is more effective we may want to test for how long we should train people

  WHY?
3. Varying levels of treatment

Target Population

Not in evaluation

Evaluation Sample

Random Assignment

1 month Financial Literacy

6 month Financial Literacy

Control group
Sophisticated designs – Example 1: Spillover Design

• Example: We want to test the effect of an job search programme that coaches job applicants

• Finding: in the treatment group 100 found employment, in the control group only 20

→ What can we conclude?

• What can we find out by varying the coverage, i.e. one area 75% get programme, in different area 25%?
Sophisticated designs – Example 1: Spillover Design

- Target Population
- Not in evaluation
- Evaluation Sample
- Random Assignment
- 75% in Treatment
- 25% in Treatment
- 75% in Control
- 25% in Control
- Control group

Compare
What can we find out by varying the coverage, i.e. one area 75% get programme, in different area 25%?

- **General equilibrium / displacement effect**: Was there a net increase in jobs?
- Why is this important?
- → External validity important for scaling up intervention
Sophisticated designs – Example 2: 2 staged randomization

Evaluation Sample 150 Clinics

Random Assignment

50 Clinics: 100 Kwacha

50 Clinics: 10 Kwacha

50 Clinics: 200 Kwacha

Offered for free once they show up

Pay 100

Bed net study in Kenya: Subsidy to reduce price (200 K)

→ What is the demand for bednets?

→ Does the usage depend on how much people value bednets and have to pay for them?
Outline

- Intervention: Variations on simple treatment-control
- Unit of randomisation
  - Randomisation at the individual level
  - Randomisation at the group / cluster level
- Method of randomization
- Constraints that effect design
- Stratification
Unit of Randomisation

- Unit of Randomisation:
  - The unit for which we ‘flip the coin’

- Cluster: Examples

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditional cash transfers</td>
<td>Villages</td>
</tr>
<tr>
<td>Bed net distribution</td>
<td>Health clinics</td>
</tr>
<tr>
<td>Community management</td>
<td>Schools</td>
</tr>
<tr>
<td>Social support</td>
<td>Family</td>
</tr>
</tbody>
</table>
Unit of Randomization: Individual?
Unit of Randomization: Individual?
Unit of Randomization: Clusters?

“Groups of individuals”: Cluster randomised Trial
Unit of Randomization: Class?
Unit of randomisation: Class?
Unit of randomisation: School?
Unit of randomisation: School?
How to Choose the Level

• Nature of the Treatment
  – How is the intervention administered?
  – How wide is the potential impact? If the intervention affects an entire group we need a cluster design.

• Generally, best to randomise at the level at which the treatment is administered.
  – Nurse-training: unit should be nurse and not hospital
  – Flip-charts: unit should be the class room, not school

• BUT: Practical concerns: Example – Randomly assign students to receive free books
  – Contamination: can we prevent students from sharing books?
  – Fairness: Do school principals / teachers / parents agree to our research design?

• Caveat: outcomes for individuals within a cluster may be correlated
  – Implications for power calculation (Lecture 5)
  – Implication for analysis (Lecture 6)
Outline

• Intervention: Variations on simple treatment-control
• Unit of randomisation
• Method of Randomisation
  • Lottery
  • Randomisation in the bubble
  • Phase-in design
  • Encouragement design
  • Rotation Design
• Stratification
Basic setup of a randomized evaluation

Target Population

Random Sampling

Not in evaluation

Evaluation Sample

Baseline Survey

Random Assignment

Balance Check

Treatment group

Control group

Endline Survey

Measure Impact
Random assignment

Income per person, per month, rupees

Treat | Compare
---|---
1457 | 1449
1000 | |
500  | |
0    | |

Note: The diagram illustrates a comparison between two groups labeled 'Treat' and 'Compare' with income data presented in rupees per person per month.
I. Lottery

- Suppose there are 2000 (eligible) applicants for a public service project, but resources only for 1000 participants.
- Randomisation can help selecting in a fair way and *and* help us evaluate.

- Randomisation mechanisms:
  - Pull out of a hat/bucket
  - Use random number generator in EXCEL or STATA and order observations randomly
II. Randomisation in “the bubble”

- Sometimes a partner may not be willing to randomise among eligible people.

- Partner might be willing to randomise in “the bubble” (=borderline in terms of eligibility)
II. Example – Effect of Credit Loans

Lender computes credit score and randomises marginal loan applicants:

- 1 – 40: Auto reject
- 41 – 45: Randomly approve 60%
- 46 – 100: Auto approve

→ Why do we randomise only ‘in the middle’?
→ What are implications for interpretation?
III. Phase-in design

- Everyone gets program eventually
- Natural approach when expanding program faces resource constraints
III. Phase-in design

Round 1
- Treatment: 1/3
- Control: 2/3

Round 2
- Treatment: 2/3
- Control: 1/3

Round 3
- Treatment: 3/3
- Control: 0

randomised evaluation ends
Advantages

• Everyone gets something eventually
• Provides incentives to maintain contact

Concerns

• Can complicate estimating long-run effects
• Do expectations of change actions today?
III. Phase-in design

[Graph showing outcome and time with intervention point labeled as 'Intervention'. Counterfactual is marked on the graph.]
How can we randomize on who gets vaccinated?

• Sometimes it’s practically or ethically impossible to randomise program access
• Randomise encouragement to receive treatment
• Encouragement = Something that makes some folks more likely to use program than others

→ Where have we seen this in the course?
IV. Encouragement design

- Encourage
- Do not encourage
- participated
- did not participate
- Complying
- Not complying

**compare encouraged to not encouraged**
These must be correlated
**do not compare participants to non-participants**
IV. Encouragement design

• Example: in the control group 15% get vaccinated, in the treatment group 65%

• Comparing groups assigned to T and C: Intention to Treat (ITT)
  \[ \Delta \text{Outcome} = 20 - 10 \]

• Treatment on the Treated (TOT)
  \[ \Delta \text{Prob} = 0.65 - 0.15 = 50\% = \frac{10}{50\%} = 20 \]
V. Rotation design

Round 1
Treatment: 1/2
Control: 1/2

Round 2
Treatment from Round 1 → Control
Control from Round 1 → Treatment
V. Rotation design

• Groups get treatment in turns
• Advantages?
• Concerns?
To summarize: Possible designs

- Simple lottery
- Randomisation in the “bubble”
- Randomised phase-in
- Encouragement design
- Rotation Design

→ The optimal randomisation mechanism depends on the nature of the intervention and the environment of the study
Factors to consider when deciding on study design

(Political) Interference
- Do partners agree to randomize?
- Do we run the risk of interference?
  - Transparency / Public lottery

Resources
- Does research project have enough funding?
- Is the available sample large enough?
  - Reduce number of treatment arms

Spillovers:
- Is there a risk of spillovers? At what level?
  - Cluster Randomization
Factors to consider when deciding on study design

Logistics
• Can sample be reached in our timeline?
• Can the program be implemented at this scale?

Monitor intervention
• Can we realistically monitor if the program is administered?
  – Ex. Iron Fortification
• Can we check adherence to random assignment?
  – Ex. Community Health Workers

Fairness / Ethics
• Do we cause any damage through our study
  – For study participants?
  – For the work of our partner organization?
Outline

• Intervention: Variations on simple treatment-control
• Unit of randomisation
• Method of randomisation
• Stratification
Example: Exec Ed Course – Stratification

There are less government groups and more NGOs in the treatment groups
→ Is this imbalance a problem?

→ What can we do?
What is it:
- dividing the sample into different subgroups
- Run a separate random assignment (lottery) in each subgroup

Example:
- we have 10 NGO (red) and 10 Government (blue) participants in our evaluation sample
- We randomly divide them into control and treatment groups of 10 students
Stratification

- **What is it:**
  - dividing the sample into different subgroups
  - Run a separate random assignment (lottery) in each subgroup
- **Example:**
  - we have 10 female (red) and 10 male (blue) students in our evaluation sample
  - We randomly divide them into control and treatment groups of 10 students

→ Stratification forced our control and treatment groups to be balanced along student gender
When to stratify

• Stratify on variables (or index variables you create) that could have important impact on outcome variable (bit of a guess)

• We can stratify along multiple variables, e.g. gender **and** perception on returns to education

<table>
<thead>
<tr>
<th></th>
<th>Lower percept.</th>
<th>Higher percept.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

→ Limitations?
• Stratify on subgroups that you are particularly interested in (where may think impact of program may be different)

• Stratification can increase the precision of our estimate

• Stratification more important when we have a small sample size
Questions?
# Methods of randomization—recap

<table>
<thead>
<tr>
<th>Design</th>
<th>Most useful when</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| **Basic lottery** | Program oversubscribed  
OK for some to get nothing | Familiar  
Easy to understand  
Easy to implement  
Can be implemented in public | Control group may not cooperate  
Differential attrition |
| **Phase in**  | Expanding over time  
Everyone must receive treatment eventually | Easy to understand  
Constraint easy to explain  
Control comply as expect to benefit later | Anticipation of treatment may impact short run behavior  
Difficult to measure long term impact |
| **Rotation**  | Everyone must get something at some point, not enough resources a year for all | More data points than phase in | Difficult to measure long term impact |
| **Encouragement** | Program has to be open to all comers  
When take up in general is low but can be impacted with incentive easily. | Can randomize at individual level even when program isn't | Measures impact of those who respond to the incentive  
Need big enough enducement to get change in take up  
Encouragement may have direct effect |