

MIT OpenCourseWare
<http://ocw.mit.edu>

Abdul Latif Jameel Poverty Action Lab Executive Training: Evaluating Social Programs
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

How to randomize?

Abdul Latif Jameel
Poverty Action Lab

Outline

- I. Methods of randomization
- II. Gathering support for evaluation
- III. “Typical” plan

Good evaluation is not just “How did we do?”

Also “How should we do it?”

- Key component: Identify implementers’ questions
- Answer them



Make research a win-win for operations

Outline

I. Methods of randomization

- a. Basic Lottery
- b. Phase-in
- c. Rotation
- d. Encouragement

II. Gathering support for evaluation

III. “Typical” plan

Methods of randomization - overview

Design	Most useful when	Advantages	Disadvantages
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
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

Lottery: Start with clinical trial benchmark

- Take 1000 people and give half of them the drug
- Can we simply apply this approach to social programs?

What are the constraints when we randomize in social situations?

- Cannot interfere significantly with program operations
- Must be perceived as fair
- Must be politically feasible
- Must be ethical

Oversubscription or “Why resource constraints are an evaluator’s best friend”

- Many programs have limited resources
- Many more eligible recipients than resources will allow services for
- Quite common in practice:
 - Training for entrepreneurs or farmers
 - School vouchers



How do you allocate these resources?

Lotteries are simple, common and transparent

- Randomly chosen from applicant pool
- Participants know the “winners” and “losers”
- Simple lottery is useful when there is no a priori reason to discriminate
- Perceived as fair
- Transparent
- Often politically feasible

... and flexible

- When: Target the interested or the qualified?
- At what level: Individuals or groups?
 - Microcredit borrowers, groups or branches
 - Over communities: which villages should we enter?
 - Over schools (e.g., the balsakhi program)
- The extent: full or partial
 - Built in and “optimizing” operations (e.g., credit scoring in South Africa)

What if you have 500 applicants for 500 slots?

Consider non-standard lottery designs

- Could increase outreach activities
- Is this ethical?
- Consider carefully what you are trying to evaluate
 - Would you evaluate the program if it never grew beyond these original 500 applicants?

Sometimes screening matters

- Suppose there are 2000 applicants
- Screening of applications produces 500 “worthy” candidates
- There are 500 slots
- A simple lottery will not work



What are our options?

Consider the screening rules

- What are they screening for?
- Which elements are essential?
- Selection procedures may exist only to reduce eligible candidates in order to meet a capacity constraint

When screening matters: Partial Lottery

- Program officers can maintain discretion
 - “Worthy” individuals are selected
 - Others are allocated slots via lottery
- Example: Training program
 - 2000 candidates
 - 1000 meet key criteria: poor enough to qualify
 - 250 out of 1000 chosen based on particular attributes
 - NGO allocates remaining 250 slots by lottery
- Example: Expansion of consumer credit in South Africa
 - Top applicants were all selected
 - Below a certain cutoff, all were rejected
 - Between the two, credit was randomly assigned

Sometimes exclusion is not desirable

Consider other points of interaction

- Programs expand
- They also have initial stages



Think creatively about when to randomize

Phase-in: takes advantage of expansion

- Everyone gets program eventually
- Natural approach when expanding program faces resource constraints
 - “In five years, we will cover 500 schools”
- What determines which schools, branches, etc. will be covered in which year?
 - Some choices based on need, geography, etc.
 - Others largely arbitrary



Can randomize order of phase-in

Phase-in designs

Advantages

- Everyone gets something eventually
- Provides incentives to maintain contact

Concerns

- Can complicate estimating long-run effects
- Care required with phase-in windows
- Do expectations of change actions today?

Rotation design

- Groups get treatment in turns
- Group A gets treatment in first period
- Group B gets treatment in second period

Advantages

- Perceived as fairer; easier to get accepted

Concerns

- If people in Group B anticipate they'll receive the treatment the next period, they can have a different behavior in the first period
- Impossible to measure long-term impact since no control group after first period

When program benefits are necessary to ensure cooperation of control groups

- Phase-in may not provide enough benefit to late round participants
- Cooperation from control group may be critical



Consider within-group randomization

- E.g., balsakhi program
- All participants get some benefit
- Concern: increased likelihood of contamination

Encouragement design: What to do when you can't randomize access

- Sometimes it's practically or ethically impossible to randomize program access
 - Flu vaccines. Employee savings programs.
- But most programs have less than 100% take-up
- Randomize encouragement to receive treatment

What is “encouragement”?

- Something that makes some folks more likely to use program than others
- Not itself a “treatment”
 - Bad idea: Training as encouragement for credit
 - Good idea: Marketing efforts
- For whom are we estimating the treatment effect?
- Think about who responds to encouragement
 - Are they different from whole population

Multiple treatments

- Sometimes core question is deciding among different possible interventions
- You can randomize these programs
- Does this teach us about the benefit of any one intervention?
- Do you have a control group?
- Advantage: win-win for operations, can help answer questions for them, beyond simple “impact”!

Randomization in “the bubble”

- Sometimes a partner may not be willing to randomize among eligible people.
- Partner might be willing to randomize in “the bubble.”
- People “in the bubble” are people who are borderline in terms of eligibility
 - Just above the threshold → not eligible, but almost
- What treatment effect do we measure? What does it mean for external validity?

To summarize: Possible designs

- Simple lottery
- Partial lottery with screening
- Randomized phase-in
- Rotation
- Encouragement design
- Multiple treatments
- Randomization in the “bubble”
 - Note: These are not mutually exclusive.

Outline

- I. Methods of randomization
- II. Gathering support for evaluation**
 - a. Political: Gossip, Fairness and Ethics
 - b. Administrative: Cost and Timing
- III. “Typical” plan

Getting started

“But I already know the answer...

... and I don't want to risk learning that we do not have an impact.”

- **Listen**
- Consider others' perspectives & objectives
- Find ways to make research operationally useful
- Everyone should have interest in *successful* evaluation, not necessarily *favorable* evaluation

Address specific concerns

- Gossip and individual perceptions
- Fairness
- Ethics
- Cost
- Incentives
- Timing

Gossip: People will talk

- The control group will find out about the treatment
- You can take steps to minimize this (grouping, etc.) but it will happen
- This can be good: everyone wants word-of-mouth advertising
- But be mindful of impact
 - Set up experiment to measure this effect
- Often a key concern for partners

Fairness

- Rare is the case of unlimited resources
- When resources limited, “fairness” is in the framing:
 - Provide everyone with 50% of treatment
 - Provide treatment first-come, first-served: the first 50 have 100% probability of treatment; the rest 0%
- Randomization can help learn more about target participants

Ethics

*It's wrong to use people as guinea pigs
If it works, then it's wrong not to treat everyone*

Some answers

- Why are prescription drugs different?
- All limited initiatives in some way an 'experiment'
 - Here, we control the experiment to learn & improve
- Allocation of limited resources happens somehow
 - Why not learn while we are doing it?
 - Produce a public good for future generations.

Cost

- What is the cost of getting it wrong?
- Often cheaper than non-experimental evaluations
 - Which costs more for survey work?
 - Which has more downside risk by yielding imprecise or biased results?
 - Key question is not ex post versus randomized trial, but evaluation vs. no evaluation



Almost *certainly* cheaper when benefits included

Timing

- Almost all incremental work for a randomized evaluation happens up front
- Projects usually want to happen “yesterday”
- Begin thinking about evaluation concurrent with thinking about project

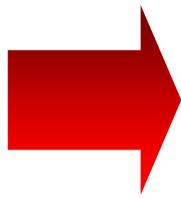
Outline

- I. Methods of randomization
- II. Gathering support for evaluation
- III. “Typical” plan

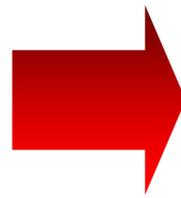
Randomization timeline



Plan



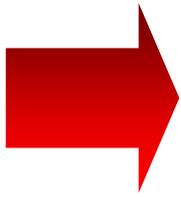
Pilot



Implement

1. Identify problem and proposed solution
2. Identify key players
3. Identify key operations questions to include in study
4. Design randomization strategy
5. Define data collection plan

Plan



Pilot



Implement

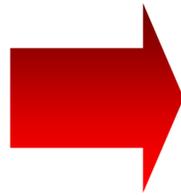
1. Identify problem and proposed solution

- Define the problem (which should lead to the key hypotheses)
- Define the intervention (sometimes already established)
- Learn key “hurdles” in design of operations (maybe opportunity to help answer them)

Plan



Pilot

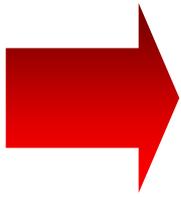


Implement

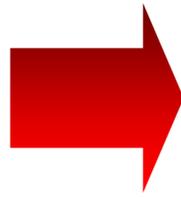
2. Identify key players

- Top management
- Field staff
- Donors

Plan



Pilot

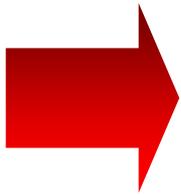


Implement

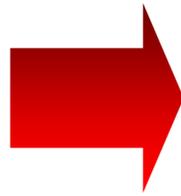
3. Identify key operations questions to include in study

- Find win-win opportunities for operations
- How to best market?
- How to sustain the program?
 - Pricing policy
 - Generating demand through spillovers
- Types or extent of training?

Plan



Pilot



Implement

4. Design randomization strategy

More on this later today

- Basic strategy
- Sample frame
- Unit of randomization
- Stratification

5. Define data collection plan

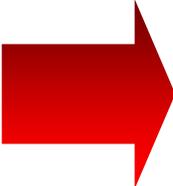
More on this tomorrow

Plan Pilot Implement

- Pilots vary in size & rigor
- Pilots & qualitative steps are *important*
- Something always goes wrong
- Better to find this out before study begins
- Often discover other interesting questions when you see the program working
- Complements lessons from other studies

Plan Pilot Implement

1. Identify “target” population
2. Collect baseline information (if applicable)
3. Randomize
4. Implement intervention to treatment group
5. Measure outcomes
6. Analyze & assess results

Plan  Pilot  Implement

1. Identify actual “target” population

More on this later today

2. Collect baseline data

More on this later



3. Randomize

More details later and in exercises

- Real-time randomization (e.g., credit scoring)
- All-at-once randomization
 - All villages known up front
- Waves
 - Want to learn from one wave to the next
 - Expansion plans not well defined throughout country. Selected districts in each wave

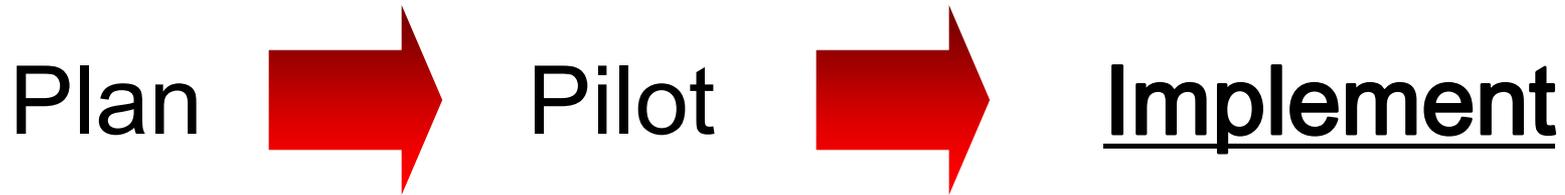
Plan Pilot Implement

4. Implement intervention to treatment groups

- Internal controls are critical
- Nothing worse than doing all this work, but not having control on the field!

5. Measure outcomes

- Common question: “How long should we wait?”
- No one-size-fits-all answer
- Operational considerations must be traded off
- Wait long enough for impacts to materialize, but not so long that the participants are hard to find.



6. Analyze & assess results

More on this later

- Must follow through on promises to partners
- A little goodwill goes a long way