How to evaluate the impact of poverty programs

Rachel Glennerster
Abdul Latif Jameel
Poverty Action Lab at MIT

povertyactionlab.org

Presentation to MIT DLab III Class – Thursday, April 19, 2007
Overview

• What is J-PAL?

• Why is evaluating what works so important?

• Why is measuring impact so hard and how do we solve these problems?

• What should I do?

• Some lessons from J-PAL studies
What is J-PAL?

• Established by 3 Professors of Economics at MIT, now a network of researchers throughout the country

• Goal is to fight poverty by ensuring that policy decisions are based on scientific evidence

• We do this by:
  – Running randomized evaluations of poverty programs
  – Encouraging others to rigorously evaluate the impact of their programs
  – Disseminating the results to decision makers

• Currently 50 ongoing programs in India, Kenya, Sierra Leone, Madagascar, Indonesia, Morocco, Peru, Philippines, and US
Why is rigorous evaluation important?

• There is surprisingly little hard evidence on what are the most effective ways to address the needs of the poor

• Evaluations can demonstrate that conventional wisdom needs to be rethought

• Important for maximizing the impact of limited resources

• By demonstrating impact and improving effectiveness, evaluations can increase support for social programs

• Evidence on what works has impact far beyond a given program
Different types of evaluation

(1) Process evaluation
   • Audit and monitoring
   • Did the intended policy actually happen?
     • How many people reached, books distributed etc

(2) Impact evaluation
   • What effect (if any) did the policy have?
     • How would individuals who did benefit from the program have fared in the absence of the program
     • How would those who did not benefit have fared if they had been exposed to the program
Why is measuring impact so hard?

• To know the impact of a program we must be able to answer the counterfactual:
  – How would an individual have fared without the program
  – But can’t observed same individual with and without the program

• Need an adequate comparison group
  – individuals who, except for the fact that they were not beneficiaries of the program, are similar to those who received the program

• Common approaches:
  – Before and after
  – Cross section
Randomized evaluation

- Determine treatment and control randomly
- By construction program beneficiaries are not more motivated, richer, more educated etc than non-beneficiaries
- Gives nice clean result everyone can understand—less fancy econometrics
- Need to plan the evaluation ex-ante, to ensure a reliable comparison group
- Randomized phase-in makes sure no-one looses out from evaluation
- Not all programs can be evaluated through randomized evaluations
Common evaluation strategies

• Pilot program—randomize who gets the pilot
• Gradual phase-in—randomize the order
• Program which is over subscribed—lottery
• Can sometimes evaluate national programs
  – look at impact of school fees by subsidizing in some areas
  – Promote take-up or knowledge of policy for some people or in some areas
What should I do?

• Doing a high quality impact study is hard and expensive
  – Not worth it when still in pilot phase

• Focus first on process, and take-up

• Understand the limits of nonrandomized approaches
  – Don’t invest in a bad evaluation

• Learn from other good impact evaluations
  – What are the issues to look out for
  – Lessons to improve my own program/innovation

• Sometimes can test limited pieces more easily
  – E.g. test alternative promotion strategies
  – Long run impacts (e.g. health impacts) much harder
Some of the lessons we have learnt

• Absence rates of public service providers are a major problem
  – 30-40 percent absence rates are common throughout developing world

• People respond to incentives
  – If people get paid whether or not they show up they don’t show up
  – If paid on basis of showing up, they show up
  – Have to be careful how you design an incentive

• Collective action is a major problem
  – Something that benefits community but benefits are diffuse are difficult
  – Eg, may be better to pay someone for maintenance rather than have committee in village to maintain

• Objective monitoring better than monitoring by people

• Major puzzles over take-up of new technologies
  – Procrastination and the benefit of small incentives
Example: Health in Rural India
Define the problem

• Seva Mandir wanted to expand into health programs in isolated rural, tribal communities with very low levels of literacy (5% female literacy)

• Major survey of health care and health status in Udaipur district
  – 1,000 households in 100 hamlets, 143 public facilities, 825 private doctors/faith healers

• Health status was very poor:
  – Very high levels of anemia (51% men, 56% women)
  – Low body mass index (93% men, 88% women below standard US cutoff)
  – Low immunization rates
  – High self reported disease, inability to work through fatigue etc

• Health delivery system was functioning very badly
  – Subcenters closed 56% of the time
  – 12% of these cases nurse could be found in the community
  – 45% totally absent (in line with national statistics for India and other developing countries)

• High use of private and faith healers
  – Visit private (often untrained) doctors twice as often as public facilities
  – Faith healers about as often as public facilities

• Large expenditures on health
  – Public meant to be free but cost as much as private visits
Design a solution

- Presented results to Seva Mandir, district health officials
- Jointly designed three programs to address issues which were potentially scalable/replicable, cost effective
- Decentralized fortification of flour
  - Flour normally fortified centrally, but this population grows, mills own
  - Education, taste test, community vote on fortification
  - Give fortified premix to local miller to add when milling (can opt out)
- Reduce absence among ANMs
  - Link pay to attendance—have to be there on specific days
  - Does regularity improve utilization and health?
- Immunization
  - Regular, predictable camps
  - Some get incentive
Evaluate impact

• Lottery held to determine which communities get which project

• Survey:
  – Utilization of facilities, e.g. attendance at immunization camps
  – Use of other providers: any reduction in visits to faith healers?
  – Impact on health
  – Economic and social impact (days worked, income, attendance at school etc)

• Analysis—compare means from the different groups (ITT)

• If we had not chosen communities at random how might our estimates of impact been biased?
  – E.g. if compare health in communities where nurse is regularly present with health in hamlets where nurse is usually not there
  – Note, some communities who were randomly selected to get fortified flour may vote not to take it up. Is this a problem? What do we do?
Why is this relevant to me?

• All of this takes a lot of work, knowledge of surveying, resources

• Important to understand the type of biases a simple before/after or simple observation may involve—seeing should not always be believing

• There are sometimes ways to introduce randomization in a cheap and easy way and measure impact through take-up (i.e. without collecting additional data)

• Service delivery is a major problem in poor communities and can be very complex to solve

• Growing evidence that small incentives can be highly effective in a way that is not straight economics
  – Immunization
  – HIV/AIDS test results