Impact Evaluation as a tool for decision-making

Ariel Fiszbein
Chief Economist
Human Development Network
World Bank

Human Development Network

Middle East and North Africa Region

Spanish Impact Evaluation Fund

Introduction

- Main Objective: Policy and program decisions based on evidence
- Devil's Advocate: There are diverse types of evidence (simpler and cheaper!)
- Argument: When I have doubts concerning the effects of an intervention, and those doubts have strong practical relevance, the impact evaluation is generally justified

The evaluation problem

- Impact: difference between the outcome with and without the program. We cannot observe simultaneously.
- Simply comparing outcomes for people with and without the program can be deceptive if program participation depends on attributes of individuals that also influence outcomes
- Simply comparing outcomes before/after program can be misleading if other things happened during the period which may explain changes in outcomes

Participants vs. non-participants

- Does job training increase employment and earnings?
- Compare employment & earning of those who sign up to those who did not
- Who signs up?
  - Those who are most likely to benefit, i.e. those with more ability
  - Would have higher earnings than non-participants without job training
- Cannot separate effect of ‘ability’ and training

Before and After

- Effect of School scholarship program on enrollment
  - Financial assistance to poor students
  - Compare enrollment before and after
  - Find fall in enrollment
  - Did the program fail?
  - Before is normal economy, but after is recession
  - Could not separate (identify) effect of financial assistance program from effect of the recession

Solving the evaluation problem

- Counterfactual: what would have happened without the program
- Need to estimate counterfactual
  - i.e. find a control or comparison group
- Counterfactual Criteria
  - Treated & counterfactual groups have identical characteristics on average,
  - Only reason for the difference in outcomes is due to the intervention

Impact Evaluation design options

- Randomized Experiments
- Quasi-experiments/non-experimental
  - Regression Discontinuity (RD)
  - Difference in difference – panel data
  - Other (using Instrumental Variables, matching, etc)
- In all cases, these will involve knowing the rule for assigning treatment

Two paths to Control Groups

- Retrospective (very hard):
  - Try to evaluate after program implemented
  - Statistically model how governments & individuals made allocation choices
  - Cannot alter treatment or control group
- Prospective:
  - Can introduce some reasons for participation that are uncorrelated with outcomes (randomization)
  - Can exploit selection rules (e.g. RD)
  - Easier and more robust

Three uses for decision-making

1. How much better off are the beneficiaries as a result of the program? Informs decisions about continuity, size and budget of the program

2. Is the program cost-effective? Informs about alternatives to achieve an objective.

3. How would results change if the program design were altered? Informs program design decisions.

Impact evaluation can influence continuity and size of programs

- **D.A.R.E. (U.S.A.)**
  - Created in 1983 as part of the effort to reduce drug use
  - Police educated students in the 5th and 6th grades
  - Multiple experimental evaluations in different parts of the country
  - Lost federal and local funding as a result of evaluations that revealed no significant impact (GAO, 2003, West and O'Neil, 2004, and Weiss, 2003).

Impact evaluation can influence continuity and size of programs

- **PROGRESA-Oportunidades (Mexico)**
  - The program began in 1998 and was phased-in over 2 years, accompanied by an experimental evaluation
  - Evaluation results were released shortly after elections in 2000, findings revealed improved education and health among participants
  - Decision to continue the program, include more families, and extend the transfers to all high school students was influenced by the rigorous independent research effort and international confirmation of the evaluation's results
  - The program now reaches over 5 million households

...hopefully by using cost-benefit analysis

- **Job Corps (EE.UU.)**
  - Training for young adults 16-24
  - Experimental design beginning in 1994-1995 and continuing over the course of 4 years
  - In the fourth year, earnings of the beneficiaries were 12% higher than the control group and the arrest rate was 16% lower
  - Cost-benefit analysis supported the program's funding (Mathematica, 2001)

...hopefully by using cost-benefit analysis

- Familias en Acción (Colombia)
  - A conditional cash transfer program that seeks to improve child health and increase school enrollment
  - On average, child weight increased by 420 grams, height-for-age increased by 0.25 SD, and secondary school enrollment increased by 5-8 percentage points
  - Using the impact evaluation results, the net present value of increased earnings as a result of better health and increased schooling were calculated
  - The value of the benefits (US$259 million) are 60% greater than the costs (US$163 million)

Three uses for decision-making

1. How much better off are the beneficiaries as a result of the program? Informs decisions about continuity, size and budget of the program

2. **Is the program cost-effective? Informs about alternatives to achieve an objective.**

3. How would results change if the program design were altered? Informs program design decisions.

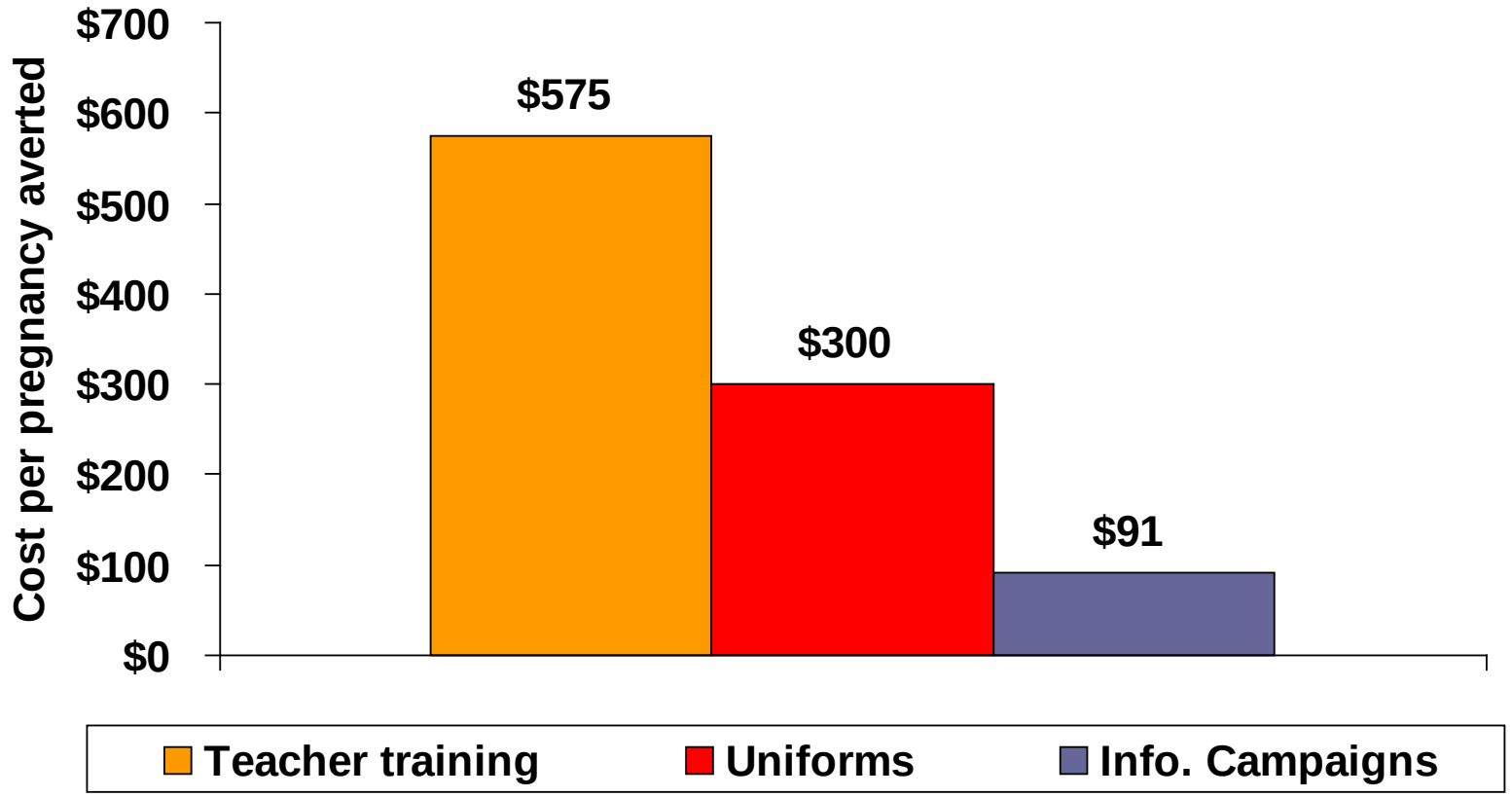
Informs about alternatives to achieve an objective (cost effectiveness)

- **Teacher incentives (India)**
  - Improve learning outcomes by:
    - Monetary incentives to groups of teachers and individual teachers conditioned on student performance
    - More inputs to schools (extra teacher, block grant for school supply needs)
  - Incentives increased test scores by 0.15 SD
  - Inputs increased test scores by 0.09 SD (Muralidharan and Sundararaman, 2006)

Informs about alternatives to achieve an objective (cost effectiveness)

- **HIV/AIDS prevention among girls in Kenya:**
  - Four methods were tested
    1. training teachers in the Kenyan government's HIV/AIDS-education curriculum
    2. organized debate and essay contest on the role of condoms in protecting teens against HIV/AIDS
    3. reduced cost of education through the supply of school uniforms
    4. information campaign for Kenyan teenagers to spread the awareness of high HIV prevalence among adult men
  - Informing girls about the high level of HIV among adult males was the most cost-effective method (Duflo, Dupas, Kremer, and Sinei, 2006).

Prevention of HIV/AIDS



Source: Duflo, Dupas, Kremer, and Sinei (2006)

Three uses for decision-making

1. How much better off are the beneficiaries as a result of the program? Informs decisions about continuity, size and budget of the program

2. Is the program cost-effective? Informs about alternatives to achieve an objective.

3. **How would results change if the program design were altered? Informs program design decisions.**

Informs program design decisions

- **RAND Health insurance experiment (U.S.A.)**
  - 5809 people in 6 sites throughout the country were randomly assigned to insurance plans that either had no cost sharing, or 25, 50, or 95% coinsurance with a maximum family out-of-pocket payment of $1000 in current dollars
  - Rates of hospitalization was equivalent for the groups
  - The study found that cost sharing reduces spending
  - Health Savings Accounts, established by the federal government in 2003, applied high-deductibles to insurance options

Informs program design decisions

- **Conditional cash transfers for school attendance (Bogotá, Colombia)**
  - Pilot program involving 10,000 students
  - Experiments with different payment methods
    1. Transfer of $15/month conditional on school attendance
    2. Transfer of $10/month, plus $50 at the end of the school year
    3. Transfer of $10/month, plus $240 at the end of secondary school
  - Allows for different timing and size of payments to be tested

If it is so obvious, why aren't there more impact evaluations?

- Ignorance has political advantages (Pritchett, 2002)
- Technical capacity is limited (and there is a strong tradition among other evaluation methods)
- Benefits are not clearly appropriated to those who bear the costs: Evaluations as a public good

There is an array of possible responses

- Legal mandates (e.g. Social Development Law in Mexico)
- Institutionalization within the state (e.g. SINERGIA in Colombia)
- Facilitating independent evaluation: methods, information, case studies (www.policyhub.gov.uk)

There is an array of possible responses

- International coordination
  - Exchange of information from evaluation results (individual and meta-analysis, guides, etc.)
    - Campbell Collaboration
      - Aims to prepare systematic reviews of social and educational policies and practices
      - (www.campbellcollaboration.org)
  - Multi-country projects for impact evaluation
    - DIME initiative
      - Coordinates impact evaluation in critical, policy-relevant areas (i.e. CCTs, education, slum upgrading)
      - Systematic review of development interventions
    - 3IE

Road map: The next 5 days

- Today: The Context
  - Why do results matter?
  - Linking monitoring with evaluation
  - Importance of evidence for policy
- Today, Monday, Tuesday: The Tools
  - Cost-benefit and cost effectiveness
  - Identification strategies
  - Data collection
  - Operational issues
- Wednesday, Thursday: The Experience
  - Group work on evaluation design and presentations