



THE WORLD BANK



Measuring Results and Impact Evaluation: From Promises into Evidence

Paul Gertler
University of California, Berkeley

How do we turn this teacher...



05.01.2006 12:40

...into this teacher?



05.01.2006 12:00

Answer Three Questions:

1. Why is evaluation valuable?
2. What makes a good impact evaluation?
3. How to implement an impact evaluation?

Why Evaluate?

- Need evidence on what works
 - Limited budget & bad policies could hurt
- Improve program/policy implementation
 - Design: eligibility, benefits
 - Operations: efficiency & targeting
- Information key to sustainability
 - Budget negotiations
 - Informing beliefs and the press
 - Results agenda & Aid effectiveness

Allocate limited resources?

- Benefit-Cost analysis
 - allows comparison of choices
 - indicates highest return investment

- Benefit:
 - change in outcome indicators
 - measured through impact evaluation

- Cost:
 - additional cost of providing benefit
 - Economic versus Accounting costs

Impact Evaluation Answers

- ❑ What was the effect of the program on outcomes?
- ❑ How much better off are the beneficiaries because of the program/policy?
- ❑ How would outcomes change if changed program design?
- ❑ Is the program cost-effective?
- ❑ Traditional M&E cannot answer these

Impact Evaluation Answers...

- ❑ What is effect of scholarships on school attendance & performance (test scores)?
- ❑ Does contracting out primary health care lead to an increase in access?
- ❑ Does replacing dirt floors with cement reduce parasites & improve child health?
- ❑ Do improved roads increase access to labor markets & raise income

Types of Impact Evaluation

□ Efficacy:

- Proof of Concept
- Pilot under ideal conditions

□ Effectiveness:

- Normal circumstances & capabilities
- Impact will be lower
- Impact at higher scale will be different
- Costs will be different as there are economies of scale from fixed costs

Using impact evaluation to....

- ❑ Scale up pilot-interventions/programs
- ❑ Kill programs
- ❑ Adjust program benefits
- ❑ Inform (i.e. Finance & Press)
- ❑ e.g. PROGRESA in Mexico
 - Transition across presidential terms
 - Expansion to 5 million households
 - Change in benefits
 - Battle with the press
 - Educate the world (Brazil versus Mexico case)

Answer Three Questions:

1. Why is evaluation valuable?
2. What makes a good impact evaluation?
3. How to implement an impact evaluation?

How to assess impact

- e.g. How much does an education program improve test scores (learning)?
- What is beneficiary's test score with program compared to without program?
- Formally, program impact is:

$$\alpha = (Y \mid P=1) - (Y \mid P=0)$$

- Compare same individual with & without programs at same point in time

Solving the evaluation problem



- ❑ Counterfactual: what would have happened without the program
- ❑ Estimated impact is difference between treated observation and counterfactual
- ❑ Never observe same individual with and without program at same point in time
- ❑ Need to estimate counterfactual
- ❑ Counterfactual is key to impact evaluation

Counterfactual Criteria...

- Treated & counterfactual
 - have identical characteristics,
 - except for benefiting from the intervention
- No other reason for differences in outcomes of treated and counterfactual
- Only reason for the difference in outcomes is due to the intervention

2 “Counterfeit” Counterfactuals

1. Before and after:
 - Same individual before the treatment
2. Those not enrolled
 - Those who choose not to enroll in program
 - Those who were not offered the program

□ Problem:

Cannot completely know why the treated are treated and the others not

1. Before and After Examples

- Agricultural assistance program
 - Financial assistance to purchase inputs
 - Compare rice yields before and after
 - Before is normal rainfall, but after is drought
 - Find fall in rice yield
 - Did the program fail?
 - Could not separate (identify) effect of financial assistance program from effect of rainfall
- School scholarship program on enrollment

1. Before and After

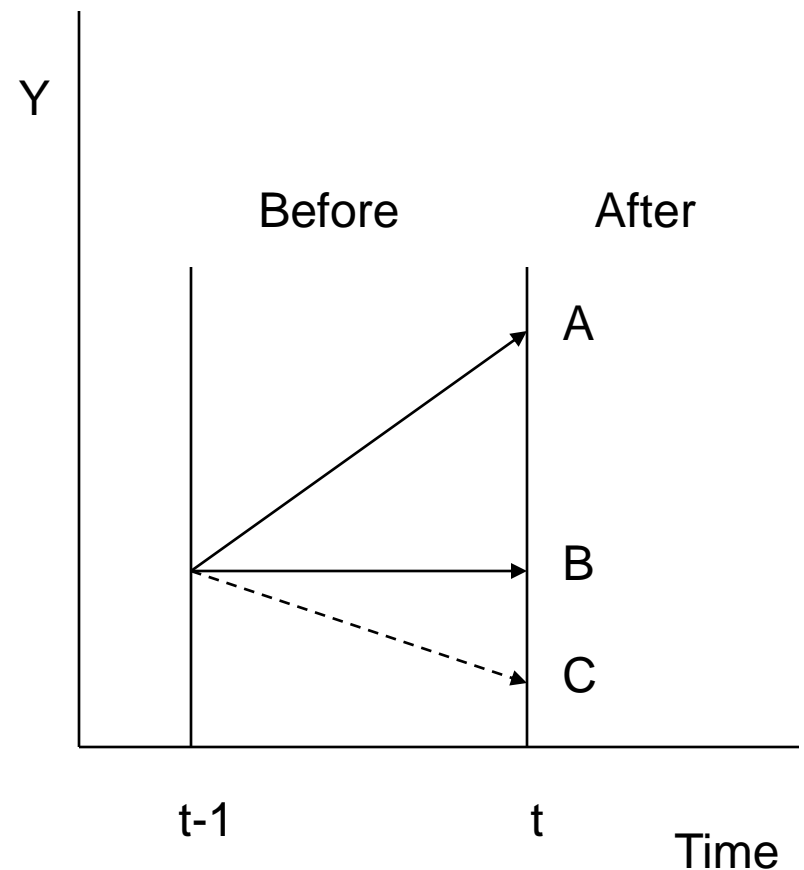
- Compare Y before and after intervention

$$a_i = (Y_{it} | P=1) - (Y_{i,t-1} | P=0)$$

- Estimate of counterfactual

$$(Y_{i,t-1} | P=0) = (Y_{i,t} | P=0)$$

- Does not control for time varying factors



2. Non-Participants....

- Compare non-participants to participants
- Counterfactual: non-participant outcomes

- Impact estimate:

$$a_i = (Y_{it} | P=1) - (Y_{j,t} | P=0) ,$$

- Assumption:

$$(Y_{j,t} | P=0) = (Y_{i,t} | P=0)$$

- Problem: why did they not participate?

2. Non-participants Example 1

- Job training program offered
- 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
- Poor estimate of counterfactual

2. Non-participants Example 2

- Health insurance offered
- Compare health care utilization of those who got insurance to those who did not
 - Who buys insurance: those that expect large medical expenditures
 - Who does not: those who are healthy
- With no insurance: Those that did not buy have lower medical costs than that did
- Poor estimate of counterfactual

What's wrong?

- Selection bias: People choose to participate for specific reasons
- Many times reasons are related to the outcome of interest
 - Job Training: ability and earning
 - Health Insurance: health status and medical expenditures
- Cannot separately identify impact of the program from these other factors/reasons

Program placement example

- ❑ Gov't offers family planning program to villages with high fertility
- ❑ Compare fertility in villages offered program to fertility in other villages
- ❑ Program targeted based on fertility, so
 - Treatments have high fertility
 - Counterfactuals have low fertility
- ❑ Estimated program impact confounded with targeting criteria

Need to know...

- ❑ Know all reasons why someone gets the program and others not
- ❑ reasons why individuals are in the treatment versus control group
- ❑ If reasons correlated w/ outcome
 - cannot identify/separate program impact from
 - other explanations of differences in outcomes

Possible Solutions...

- ❑ Need to guarantee comparability of treatment and control groups
- ❑ ONLY remaining difference is intervention
- ❑ In this seminar we will consider
 - Experimental design/randomization
 - Quasi-experiments
 - ❑ Regression Discontinuity
 - ❑ Double differences
 - Instrumental Variables

These solutions all involve...

- Knowing how the data are generated
- Randomization
 - Give all equal chance of being in control or treatment groups
 - Guarantees that all factors/characteristics will be on average equal btw groups
 - Only difference is the intervention
- If not, need transparent & observable criteria for who is offered program

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
 - Identification strategies
 - Sample size and power
 - Operational issues
- Wednesday, Thursday: The Experience
 - Group work on evaluation design and presentations

THANK YOU