



THE WORLD BANK



Operational Issues in Impact Evaluation

**Laura Rawlings
World Bank**

**Paul Gertler
UC Berkeley**

Operational Issues

- I. Political economy of results
- II. Planning an impact evaluation
- III. Policy context
- IV. Designing the evaluation
 - Retrospective versus prospective designs
 - Making the design compatible with operations
 - Ethical issues
- V. Importance of monitoring
- VI. Overall messages



I. Political Economy of Results



□ Cultural shift

- From retrospective evaluation
 - **Look back and judge**
- To prospective evaluation
 - **Decide what need to learn**
 - **Experiment with alternatives**
 - **Measure and inform**
 - **Adopt better alternatives overtime**

□ Change in incentives

- Rewards for changing programs
- Rewards for generating knowledge
- Separating job performance from knowledge generation



- Focus on results and use of prospective evaluations facilitates:
 - Tailoring policy questions
 - Precise unbiased estimates
 - Using resources wisely:
 - Better methods
 - Cheaper data
 - Timely feedback and program changes
 - Improve results on the ground



II. Planning an Impact Evaluation



- Evaluate impact when project is:
 - Innovative
 - Replicable/scalable
 - Strategically relevant for reducing poverty
 - Evaluation will fill knowledge gap
 - Substantial policy impact
- Ingredients for success:
 - Political and financial support
 - Advance planning
 - Understanding context
 - Stakeholder involvement



Planning an Impact Evaluation



Address the following **as early as possible**:

- ❑ *Objectives* – What are the key questions that need to be answered? Fit these into a RESULTS framework.
- ❑ *Methodology* – What type of assessment is needed? Is an impact evaluation needed? Relate the evaluation design to program operation rules.
- ❑ *Timeframe* - Depends largely on depth and breadth of evaluation, design and availability of existing data. Sufficient time must have elapsed for results to be produced for an impact evaluation.



- ❑ *Cost of Evaluation* – varies widely for impact evaluation: average between 0.25% and 2.0% of project costs. Data collection often highest expense. Power calculations needed.
- ❑ *Financing* – often under financed; can be financed by project, other government resources, grants or combinations. Explore donor resources given public good value of results.
- ❑ *Capacity* - coordination and execution is challenging. Often good balance between using international technical assistance and local capacity.
- ❑ *Policy Relevance* - Ensure full engagement of policymakers, program implementers and the use of evaluation results to inform key decisions.



III. The Policy Context

- Address policy-relevant questions:
 - What policy questions need to be answered?
 - What outcomes answer those questions?
 - What indicators measures outcomes?
 - How much of a change in the outcomes would determine success?
 - Who cares? What decisions will results inform?

- Example:
 - Scale up pilot?
 - Criteria: Need at least a $X\%$ average increase in beneficiary **outcome over a given period**



IV. Designing the Evaluation

- What is the evaluation's identification strategy?
 - Depends on the implementation of the program
 - Control groups are generated from the operational design of the program, not as a separate strategy

- Retrospective vs. Prospective

Retrospective Designs

- Retrospective Analysis is necessary when we have to work with a **pre-assigned program** (expanding an existing program) and **existing data** (baseline?)

- Examples:
 - Randomization: Auditorias de corrupción (Brazil)
 - Regression Discontinuity: Bono Sol (Bolivia)
 - Difference in Differences: AGES (México)
 - Instrumental variables: Piso firme (México)

Prospective Designs

□ Prospective Analysis

- The evaluation is designed in parallel with the assignment of the program
- Baseline data can be gathered

□ Example:

- Progresas/Oportunidades (México)

Prospective Designs

Use opportunities to generate good controls

- The majority of programs cannot assign benefits to all the entire eligible population
 - Budget limitations:
 - Eligible beneficiaries that receive benefits are potential treatments
 - Eligible beneficiaries that do not receive benefits are potential controls
 - Logistical limitations:
 - Those that go first are potential treatments
 - Those that go later are potential controls
- Not all eligible receive the program
 - Randomized Promotion



The evaluation method depends on rules of operation



Program Implementation		Targeted	Universal
In Stages	Without cut-off	Randomization	Randomized Rollout
	With cut-off	RD/DiD Match/DiD	RD/DiD Match/DiD
Immediately	Without cut-off	Randomized Promotion	Randomized Promotion
	With cut-off	RD/DiD Match/DiD	Randomized Promotion



Who gets the program? When?

- Who gets the program? Eligibility criteria
 - Are benefits targeted?
 - How are they targeted?
 - Can we rank eligible's priority?
 - Are measures good enough for fine rankings?
 - Can data on non-eligible be a control group?

- When? Roll out
 - Equal chance to go first, second, third?
 - Can use later entries as control group for early beneficiaries



Ethical Considerations

- ❑ Do not delay benefits: Rollout based on budget/administrative constraints
- ❑ Equity: equally deserving beneficiaries deserve an equal chance of going first
- ❑ Transparent & accountable & equitable method
 - Give everyone eligible an equal chance
 - If rank based on some criteria, then criteria should be quantitative and public



IV. The Importance of Monitoring



- Projects/programs regularly collect data for management purposes
- Typical content
 - Lists of beneficiaries
 - Distribution of benefits
 - Expenditures
 - Outputs
- Cost data needed for cost-benefit analysis



- Monitoring information is needed for impact evaluation
 - Verify who is beneficiary
 - When started
 - What benefits were actually delivered

Necessary condition for program to have an impact:

- benefits need to get to targeted beneficiaries



V. Overall Messages

- Ensure that impact evaluation is **useful for key policy decisions**
 - Validating program design
 - Adjusting program structure
 - Communicating to finance ministry & civil society
- **Engage policymakers and operational staff** in evaluation design to ensure its utility
- **Plan ahead-** prospective designs are better



Overall Messages

- Respect the **rules of program operation** and use them to generate a robust counterfactual
 - What would have happened to beneficiaries if had not received the program
 - Need to know all reasons why beneficiaries got program & others did not
- **Monitoring systems** & administrative data critical to good impact evaluation and effective program management
- Incorporate principles of **transparency, accountability and equity**

