Introduction

In the evolving fight against world poverty, the need for reliable scientific evaluation of development programs has become paramount. Resources are not infinite and the leaders of international aid and non-government organizations require a methodological standard to guide the decision to allocate funds across the broad spectrum of anti-poverty strategies. More than ever before, program advocators must provide concrete evidence of success or failure to ensure expansion of their programs. Researchers have a responsibility to use sound methods in evaluating new approaches. Only if we trust in our assessments can we learn from the success and failure of past experiences in combating poverty; moreover, only with reliable results to build upon can we ensure our future policies will move us forward in the struggle against poverty.

What makes a good evaluation? What makes a bad one? These are some of the questions that confront many practitioners. MIT's Poverty Action Lab has developed an executive education program to educate practitioners on the ideas and statistics behind successful evaluation. The course provides participants with the necessary knowledge and training to scientifically test the accountability of development programs. A wide variety of examples will be discussed, from fields as diverse as education, health, decentralization and infrastructure projects. Participants will walk away with a new understanding of the broad issues related to randomized evaluation as well as the specific requirements necessary to implement their own successful ones. This should help everyone, be they heads of NGOs or managers in donor organizations, to become better consumers and producers of evaluations.

Program Objectives

The program is designed for people from a variety of backgrounds---managers and researchers from international development organizations, managers in non-Governmental organizations from around the world as well as trained economists looking to retool. It will provide a thorough understanding of randomized evaluation and a pragmatic step-by-step training in conducting one's own evaluation.

Specifically, the following key concepts will be covered:

- The common pitfalls of evaluation
- Why randomized evaluations circumvent these pitfalls
- The key components to a good randomized evaluation design
- The guidelines to successful implementation of a randomized evaluation
- The significance and proper techniques of statistical analysis
- The interpretation and marketing of results

The program will achieve these goals through a diverse set of integrated teaching methods. Expert researchers will provide both theoretical and example-based classes complemented by workshops where participants can apply key concepts to real world examples. By examining both successful and problematic evaluations, participants will better understand the significance of various specific details of randomized evaluations. Furthermore, the program will offer extensive opportunities to apply these ideas ensuring that participants will leave with the knowledge, experience, and confidence necessary to conduct their own randomized evaluations.

Teaching Team

Joshua Angrist
Ph.D., Princeton University
Professor, Labor Economics

Abhijit Banerjee
Ph.D., Harvard University
Professor, Applied Theory and Development Economics, MIT

Esther Duflo
Ph.D., MIT
Castle Krob Career Development Associate Professor of Economics, MIT

Dean Karlan
Ph.D., MIT
Assistant Professor of Economics and International Affairs
Economics Department and Woodrow Wilson School of Princeton University

Michael Kremer
Ph.D., Harvard University
Gates Professor for Developing Societies, Harvard

Sendhil Mullainathan
Ph.D., Harvard
Mark Hyman Jr. Career Development Associate Professor of Economics, MIT

Curriculum

I. Introduction to effective evaluations

A program evaluation attempts to answer the basic question: How would an individual have fared in the absence of the program? This unit examines the various methods of program evaluation including both non-randomized evaluations (such as propensity score matching, difference-in-difference estimation, and regression discontinuity design) and randomized evaluations, retrospective evaluations, and impact evaluations. First, the arguments for the superiority of randomized evaluations are laid out. The focus will then shift to introducing two real-world examples: PROGRESA in Mexico and the district primary education program (DPEP) in India, respectively a success and failure at doing a proper evaluation. These two evaluations will serve as key sources of examples throughout the duration of the course.

a. Why evaluations matter
b. Different evaluation types
c. What to look for in an evaluation
d. Why randomized evaluations are the gold standard
e. Two examples: PROGRESA in Mexico and DPEP in India

II. Randomized Evaluation Design

This unit focuses on the design stage of randomized evaluations. In this planning stage, researchers must make several key decisions. This unit will develop a systematic checklist to guide the randomization process. Lessons will also focus on inventive and elegant ways to add randomization into a policy.

a. Choosing the level of intervention: Individual, School, Country, or other
b. Specifying the program to be studied
c. Choosing objectives
d. Choosing the population and sample size
e. Choosing what variables to survey
f. Randomizing in the real world

III. Implementation

Implementing a randomized evaluation can be a bewildering experience. This unit established guidelines for survey and instrument design as well as good practices in data collection and data entry. Field conditions often differ from expectations, and researchers must plan ahead to standardize data to the greatest extent possible. Participants will examine case studies to raise their awareness of problems with

program implementation, and this practice will better enable project managers to adapt to varied field conditions.

a. Baseline data collection: survey & instrument design
b. Following up on program implementation
c. Ensuring the quality and consistency of regular data collection and entry
d. Common pitfalls and their solutions

IV. Analyzing the Data

Evaluation results are only useful if the resulting data are properly analyzed. This unit introduces and reviews the relevant statistical knowledge for analyzing data from a randomized evaluation. Using both theoretical and data-driven workshops, participants will reinforce their data analysis skills. Lessons will cover both general statistical techniques and trouble shooting solutions specific to problems encountered in randomized evaluation experiments.

a. Reviewing statistics
b. Determining if an estimate is spurious or significant
c. Interpreting the size of the estimates
d. Understanding how to estimate results for different subgroups
e. Troubleshooting problems
    i. Hawthorne and John Henry effects
    ii. Differential attrition rates
    iii. Spillover effects

V. Using Results

This final unit covers both interpretation and presentation of results from a randomized evaluation. Once data has been analyzed, a researcher faces several key decisions: What is the bottom line for the program? Can it be expanded? Should it be expanded? What other evaluations tests should be done? These answers are not always explicit, yet this final integration of results will often singularly direct the program's future. Project managers must offer a convincing argument of both their results and their methodology to ensure policymakers fully appreciate the concrete evidence provided by randomized evaluations. Only with a clear presentation can the results of a randomized evaluation contribute as a global public good to the overall anti-poverty research network.

a. Conducting Cost/Benefit Analysis
b. Scaling up the results
c. Deciding whether to conduct additional experiments
d. Forecasting the impact of the program to broader samples
e. Presenting results effectively
f. Political Economy of Presenting randomized trials