Evaluating Community-Based Programs

Robert J. LaLonde
The Harris School
The University of Chicago
November 2007
Introduction

- Outline challenges associated community-based programs designed to improve employment, health, education & other economic outcomes.
- Realistic expectations
- The evaluation problem
- The selection problem
- Evaluating outcomes for:
  - Individuals?
  - Communities?
Big Impacts Will Look Small

- An employment and training program raises annual earnings by $1,000 per year.
  - Combination of general skill, vocational skills, job search assistance.

- Participant (direct) costs: ~$3,000 - $5,000
  - Indirect costs (1): Training delays return to work?
  - Indirect costs (2): Trainees “displace” other job seekers?

- Is the $1,000 effect permanent?
  - “Rate” of return ~ 25%
  - Far better than one year of formal schooling!
Good Evaluations are Difficult, Time Consuming, & Expensive

- Evaluations are only cost-effective if they lead to some significant action or outcome.
  - Doing nothing could be a significant action!
  - High quality evaluations are sometimes supported by program opponents!

- Ask is it worthwhile to evaluate “this” program?
  - Concentrate scarce resources on a limited number of high quality evaluations
The Evaluation Problem

- The “evaluation problem” is …
  - a “**missing data**” problem.

- **What is missing?**
  - Data on participants’ “counterfactual” outcomes.

- **Use other data to “fill-in” or “estimate” participants’ counterfactual outcomes:**
  - A comparison community’s outcomes
  - A community’s pre-program outcomes
Community-Based Evaluations

- Empowerment/Enterprise zones
- TIFs (Tax Increment Financing)
- Community Policing
- Saturation designs
  - for youth employment
  - Reentry programs for prisoners
- Community organizations
- Political empowerment
Illustrate Four Evaluation Strategies

Program Communities: A

Non-Program Communities: B

Pre/post = A – A_{-1}         Timeline
Cross sectional = A - B
Panel Data = (A - B) - (A_{-1} - B_{-1})
Cohort = A – B_{-1}
Why Do Outcomes Differ Among Communities?

- Program Community A ➔ Outcome(A)
- Non-Program Community B ➔ Outcome(B)

**Outcome(A)**
- = Program Effect + Other community(A) variables.

**Outcome(B)**
- = Other community(B) variables.

**Outcome(A) – Outcome(B) = Program Effect +**
- [Other community(A) variables - Other community(B) variables].
It is Hard to Evaluate the Impact of One Community-Based Program

- Other variables in Community A =
  - (1) Variables we observe and measure
  - (2) Variable either we can not measure or observe.

- Can match Community A with a Community B that has the same values for the variables we can measure & observe.
  - E.g.: Percentage of households living in poverty

- Outcome(A) – Outcome(B) = Program Effect +
  - [Variables can not measure in community(A) - Variables can not measure in community(B)]

- The term in [ . ] is the “Matching Error.”
How do Program Evaluators Solve the Matching Error Problem?

- **Outcome(A) - Outcome(B) =**
  - Program Effect + Matching Error.

- **We need to have many program communities & non-program communities.**
  - Is the matching error on average = 0?
  - We can estimate the *average program effect*.
  - But *can not* estimate a program effect for any single program community.

- **Suppose we have 16 program communities ...**
  - Is it reasonable to assume the “matching error” averages out?
### Characteristics of PROGRESA Treatment, Control and Other Rural Mexican Communities

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Treatments</th>
<th>Controls</th>
<th>Other Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female Head?</td>
<td>8.3%</td>
<td>8.5%</td>
<td>13.4%</td>
</tr>
<tr>
<td>No Schooling</td>
<td>44.8%</td>
<td>46.0%</td>
<td>40.4%</td>
</tr>
<tr>
<td>Age of Head</td>
<td>42.2</td>
<td>42.6</td>
<td>47.2</td>
</tr>
<tr>
<td>No Bathroom</td>
<td>48.2%</td>
<td>48.9%</td>
<td>28.9%</td>
</tr>
<tr>
<td>Dirt Floor</td>
<td>72.9%</td>
<td>75.4%</td>
<td>20.3%</td>
</tr>
<tr>
<td>No Gas Stove</td>
<td>84.7%</td>
<td>83.4%</td>
<td>26.0%</td>
</tr>
</tbody>
</table>
The 1976 CETA Male Cohort

- Matching Participants & Non-Participants Prior Employment History.
  - Define: 0 = not employed; 1 = employed
  - Outcome: Employment Rates in 1977

<table>
<thead>
<tr>
<th>History</th>
<th>Non-participants</th>
<th>‘76 Trainees</th>
</tr>
</thead>
<tbody>
<tr>
<td>000000</td>
<td>.099</td>
<td>.674</td>
</tr>
<tr>
<td>111110</td>
<td>.538</td>
<td>.821</td>
</tr>
<tr>
<td>011111</td>
<td>.888</td>
<td>.863</td>
</tr>
<tr>
<td>101111</td>
<td>.866</td>
<td>.886</td>
</tr>
<tr>
<td>001111</td>
<td>.958</td>
<td>.830</td>
</tr>
<tr>
<td>111111</td>
<td>.918</td>
<td>.870</td>
</tr>
</tbody>
</table>
Matching Error also May Bias Pre/Post Comparisons

- Can the change in Community A’s post-program and pre-program outcomes estimate the program’s impact?
  - Outcome Next Year (A)
  - Outcome Last Year (A’)
  - Program Implemented This Year

- Outcome Next Year (A) - Outcome Last Year (A’) =
  - Program Effect + [Change in other variables in Community A]
  - The term [ . ] also is “matching error.”

- Is it possible for this “matching error” to “average out?”
The Matching Problem is a Selection Problem - 1

- Participants **choose** to participate in programs based on their own assessments of whether they will benefit from the program.

- Program operators **select** applicants that they believe will benefit from the program.
The Matching Problem is a Selection Problem - 2

- The same issues arise when evaluating community-based programs.
- Why is the program operating in Community A:
  - Strong community leaders?
  - Prior outcomes are extreme?
  - Community selected to receive program services or funding?
Because of participants’ and program operators’ decisions, …
- participants and non-participants are different.
- Expect their outcomes to differ even if program had no impact.

Challenge: How does the evaluation account for these decisions?
## Program and Non-Program Communities

<table>
<thead>
<tr>
<th>Program Communities</th>
<th>Non-Program Communities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not eligible &amp; does not participate</td>
<td>Household or establishment not eligible.</td>
</tr>
<tr>
<td>Eligible participant</td>
<td>Household or establishment eligible</td>
</tr>
<tr>
<td>Eligible non-participant</td>
<td></td>
</tr>
</tbody>
</table>

Household or establishment not eligible.
The 1964 MDTA Male Cohort

- Vocational classroom training for the permanently unemployed.

- Annual social security earnings of:

<table>
<thead>
<tr>
<th>Year</th>
<th>Participants</th>
<th>Non-participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962</td>
<td>$1,843</td>
<td></td>
</tr>
<tr>
<td>1963</td>
<td>$1,810</td>
<td></td>
</tr>
<tr>
<td>1964</td>
<td>$1,551</td>
<td>$3,750</td>
</tr>
<tr>
<td>1965</td>
<td>$2,923</td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>$3,750</td>
<td></td>
</tr>
</tbody>
</table>
The 1964 MDTA Male Cohort

- Vocational classroom training for the permanently unemployed.

- Annual social security earnings of:

<table>
<thead>
<tr>
<th>Year</th>
<th>Participants</th>
<th>Non-participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962</td>
<td>$1,843</td>
<td>$2,963</td>
</tr>
<tr>
<td>1963</td>
<td>$1,810</td>
<td>$3,108</td>
</tr>
<tr>
<td>1964</td>
<td>$1,551</td>
<td>$3,275</td>
</tr>
<tr>
<td>1965</td>
<td>$2,923</td>
<td>$3,458</td>
</tr>
<tr>
<td>1966</td>
<td>$3,750</td>
<td>$4,351</td>
</tr>
</tbody>
</table>

- 1963 earnings higher due to "selection"
- 1964 earnings higher due to "training year"
Conclusions

- Community-based training, health, education programs are **difficult** to evaluate.
- The “effect size” is very likely **small** compared to the “normal” variation in outcomes.
- Whether evaluating people or communities, must carefully identify **counterfactual** outcomes.
- Despite wishes of foundations, impact evaluations are **not cost-effective** for many programs.