Abdul Latif Jameel Poverty Action Lab Executive Training: Evaluating Social Programs
Spring 2009

Thinking about measurement and outcomes

Abdul Latif Jameel
Poverty Action Lab

povertyactionlab.org

Outline

I. Example: Reservations for women

II. Possible outcomes

III. Key Hypotheses and Chain of Causality

IV. Data collection, management and entry

Today's objectives

- We now have a question and a design for the evaluation.

- How do we prepare the data collection?
  - What data should we collect?
  - What systems should we set up to ensure good quality data.
  - What sample size do we need to plan?

Outline

I.   **Example: Reservations for women**

II. Possible outcomes

III. Key Hypotheses and Chain of Causality

IV. Data collection, management and entry

The setting: Reservation in the panchayati raj

- What are the main goals of the Panchayati Raj?

- What are the main characteristics of the reservation policy?

Source: "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," by Raghabendra Chattopadhyay and Esther Duflo (2004a), Econometrica 72(5), 1409-1443.

The controversy about reservations

- Why were reservations deemed to be desirable in this context?

- Why did some people doubt that reservations would be effective?

Outline

I.   Example: Reservations for women

II.   **Possible outcomes**

III. Key Hypotheses and Chain of Causality

IV. Data collection, management and entry

The possible effects

- Let's start by drawing a list of everything we think reservations for women may affect.

The evaluation question

- Suppose budget was not an issue: should we go out, hire a large team of investigators, collect data on every single of these indicators in reserved and unreserved communities, and see what happens?
  - Pro:
  - Cons:

Interpreting multiple outcomes

- Suppose that you collect data on 20 different public goods
    - You find that for water wells, investment in the public good is significantly higher in Panchayats reserved for women.
    - You find that for irrigation, investment in the public good is significantly higher in Panchayats that are not reserved.
    - You find that for 18 goods, the outcomes are very similar, and not significantly different in both Panchayats
- What can you conclude?

Interpreting multiple outcomes

- Suppose that you collect data on 20 different public goods
    - You find that for water wells, investment in the public good is significantly higher in Panchayats reserved for women.
    - You find that for irrigation, investment in the public good is significantly higher in Panchayats that are not reserved.
    - You find that for 18 goods, the outcomes are very similar, and not significantly different in both Panchayats

- What can you conclude?

The hypotheses to test must be defined before the beginning of the experiment, or we have no good way to assess their validity

Outline

I.   Example: Reservations for women

II. Possible outcomes

III. **Key Hypotheses and Chain of Causality**

IV. Data collection, management and entry

Defining key hypotheses

- What might be examples of a few key hypotheses to test?

- Which variables, or combinations of variables, might you use to test these key hypotheses?

Drawing the chain of causality

- Answer more than <u>how</u> effective, but also <u>why</u>?

- We want to draw the link

Intervention ➡️ intermediary variables
➡️ final outcome

- What are the intermediate variables through which the intervention is affecting the final outcome of interest?

- How are they likely to be affected?

Defining and measuring Intermediate outcomes will enrich our understanding of the program, reinforce our conclusions, and make it easier to draw general lessons

Modeling the effects of reservations

- What are the possible chains of outcomes in the case of reservation?

- What are the critical steps needed to obtain the final results?

- What variables should we try to obtain at every step of the way to discriminate between various models?

One possible model
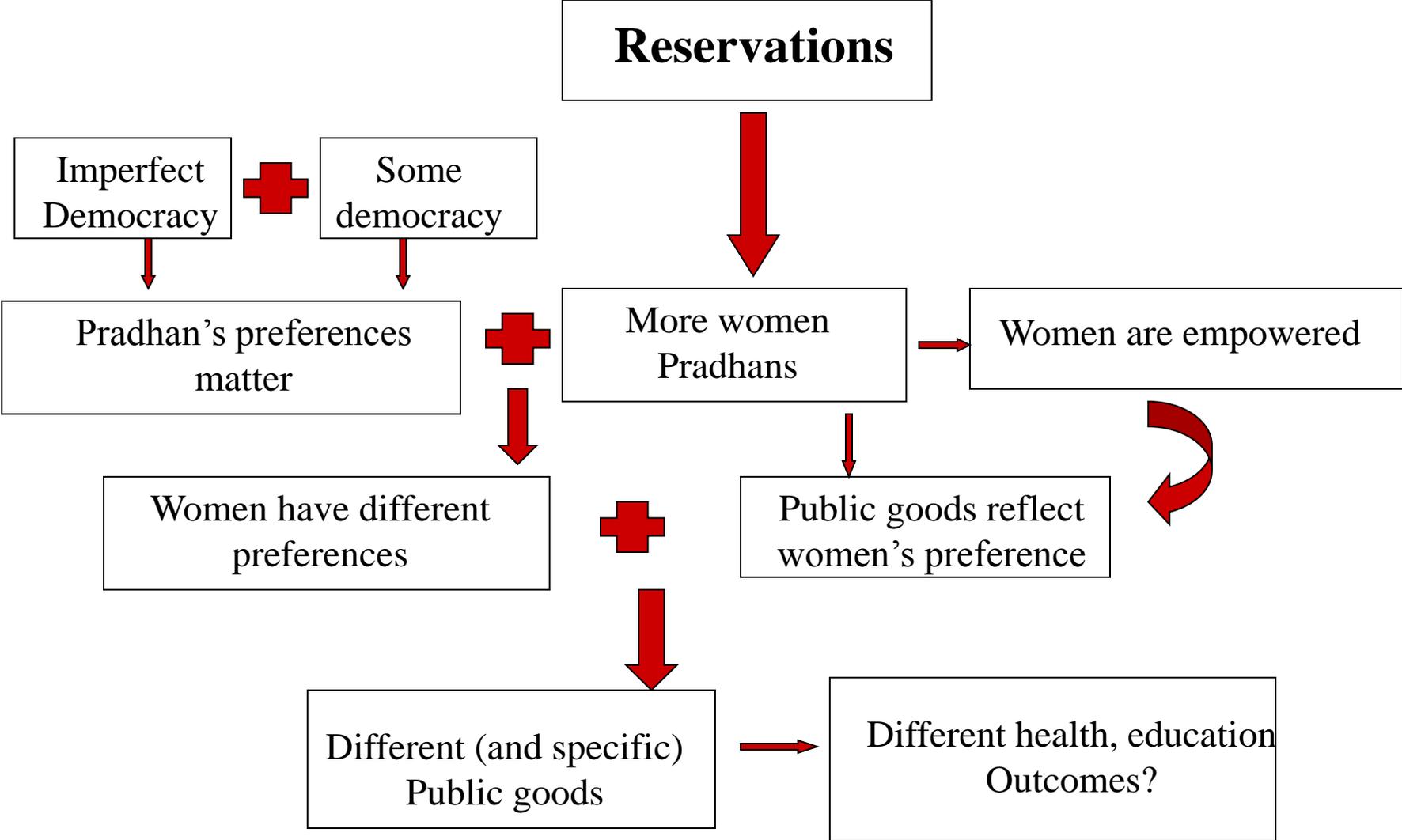
Outline

I. Example: Reservations for women

II. Possible outcomes

III. Key Hypotheses and Chain of Causality

IV. **Data collection, management and entry**

Getting ready to collect the data

- What variables do we need to test the entire model?

- What tools would allow us to collect what part of the data we need at each stage?
  - What is the best (richest) tool?
  - What is the cheapest tool?

Working within a budget

- Collecting data is expensive, and we may have to make choices.

- If we had to choose one variable in this chart to determine whether women made a difference, what would it be?

- You have computed the minimum budget required to detect the impact of water wells at the village level. What else can we find out with the same activities, without spending any extra money, and to what extent can we populate the flow chart?

Working within a budget

- With a larger budget, what would we want to do?

- Would data collection in several stages be advisable? What stages would you recommend?

The objective and chains in your proposals

- What is the main hypothesis in your proposal?

- What is the key variable of interest (on which you will declare success or failure)

- What are the intermediate outcomes that you are planning to collect data on?

Data collection and management

Data quality is at the heart of impact evaluation: with bad data, ALL previous efforts are lost.

Some important principles for data quality:

- Questionnaire must be thoroughly piloted to ensure that the questions make sense to people.

- Data entry format must be clear and should not leave room for interpretation by the enumerator: DK code, skip pattern.

- Training in these procedures is essential.

- There must be a weekly check of *all* the formats by a supervisor, and a re-check on a random basis by the research manager.

- There must be a re-survey of sample households on a random basis. **CHECK** that they produce the right results.

Data collection and management

- ## Other things to keep in mind:
  - – Your organization generates a lot of data: use it!
  - – However, we need to be careful to collect the SAME data in treatment and control group.
    - • Example: Seva Mandir NFE camera project: Cameras generate good data on quality. Yet, we need the same data on the control group as well!
  - – Attrition will bias the results and is very difficult to fix ex-post: put in place good procedures to contain attrition. Keep track of attrition as the survey progress.

Data entry

- Invest in a good data entry data base (you can get it programmed by a software firm on the basis of your survey).

    – It should guide the data entry (skip pattern etc…) and contain logical error checks.

- Enter sample data first to check that there are no logical error in the data base.

- Do double entry of all data: reconcile with the hard copies to detect mistakes.

- Re-enter some entries a third time (supervisor) and seek to achieve less than 0.5% error rate.

- After re-entry, there should be data cleaning: check that all the data is sensible (no 120 year old in school, no people with -5 chickens etc…)