



Massachusetts  
Institute of  
Technology

# The power of Evidence (and randomization)

Esther Duflo

Lecture 2-14/73



**Finding out what works: power of social experiments**

# Why so much focus on impact evaluation in this class?

- In this class, you will see many example of “randomized controlled trials”
- Today we will discuss what this is, and why they are important

Let's be concrete. What potential policies would help increase participation in school?

## Evaluating these policies

- Suppose that you are given full freedom to pick the best one (or maybe combinations of the best ones) to scale up in the entire country. You have some time, (let's say 3 or 4 years) to come up with the best plan, and money to try things out
- What questions do you need to answer about each of these policies to know whether to recommend them or not?
- For example if we chose the example of providing free school meals to poor kids. What kinds of questions might you want to ask to describe how such a program is going?

# Evaluating school meals: the questions

- Are the school meals served regularly?
- Is there wastage?
- Do kids eat them?
- Are the kids better nourished?
- Are kids more likely to come to school now?
- Are the poor kids the ones who are really getting the meals?
- Do the kids learn more in school?

# Organizing these questions

---

Needs  
Assessment

Process  
Evaluation

Impact  
Evaluation

# Needs Evaluation

---

- Who is the targeted population?
  - All children? The poor ones?
  - Why do we need to answer this question?
- What's the nature of the problem being solved?
  - How will school meals solve it
  - Why do we need to answer this question?
- How does the service fit the environment?
  - Do teachers feel comfortable cooking?

## Process Evaluation

- Are the services being delivered?
  - Money is being spent
  - School meals are delivered, children are eating them
- Are there ways of improving cost effectiveness?
  - Substituting expensive inputs with less costly alternatives, substituting costly inputs with labor, delivery methods
  - Are children spending all day at school eating instead of studying?
- Are the services reaching the right population?
  - Schools with large absence problem
- Are the clients satisfied with service?
  - Teachers', students' response to meals

# Impact Evaluation

- Key question: Did school meals cause students to attend school more?
- Auxiliary questions:
  - What was the effect on enrollment?
  - What was the effect on attendance?
  - What was the effect on learning?
  - Did some types of people benefit more than others?
    - Students who were doing worse, poorer students, etc.

# Why impact evaluation?

- Surprisingly little hard evidence on what works, and evidence is often not based on data analysis, more on general impression.
- Central issue in the debate on aid
  - Do we know that anything is working?
  - How do we identify what works?
  - Pick what really works
- Key ingredient to “Effective altruism” (and good policy making)
- It is also how we learn what is important in people’s lives and how they behave..



GiveWell

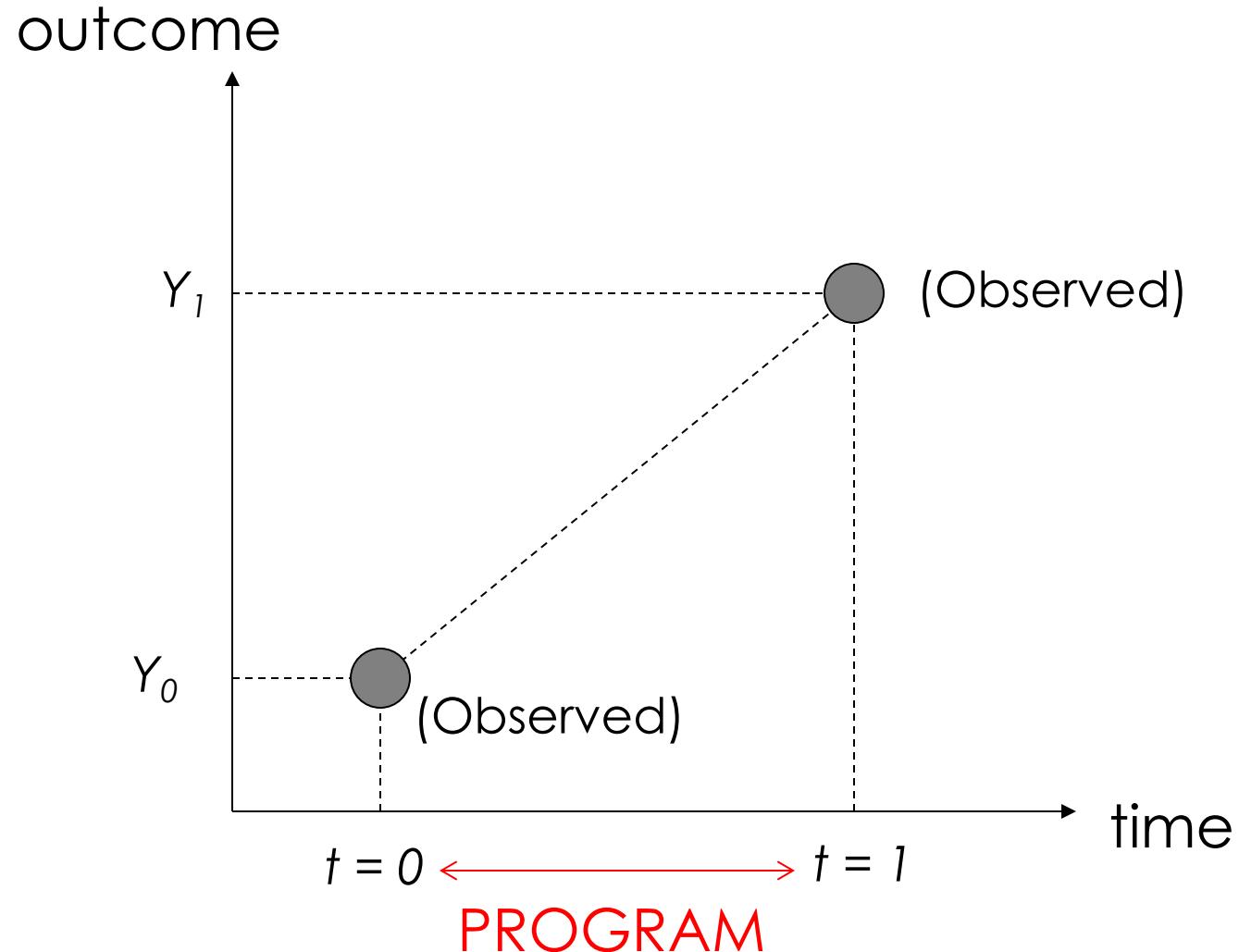
We search for the charities that save or improve lives the most per dollar.

Our goal is to produce the world's top research on where to give. Free, for everyone. We recommend a small number of charities that do an incredible amount of good.

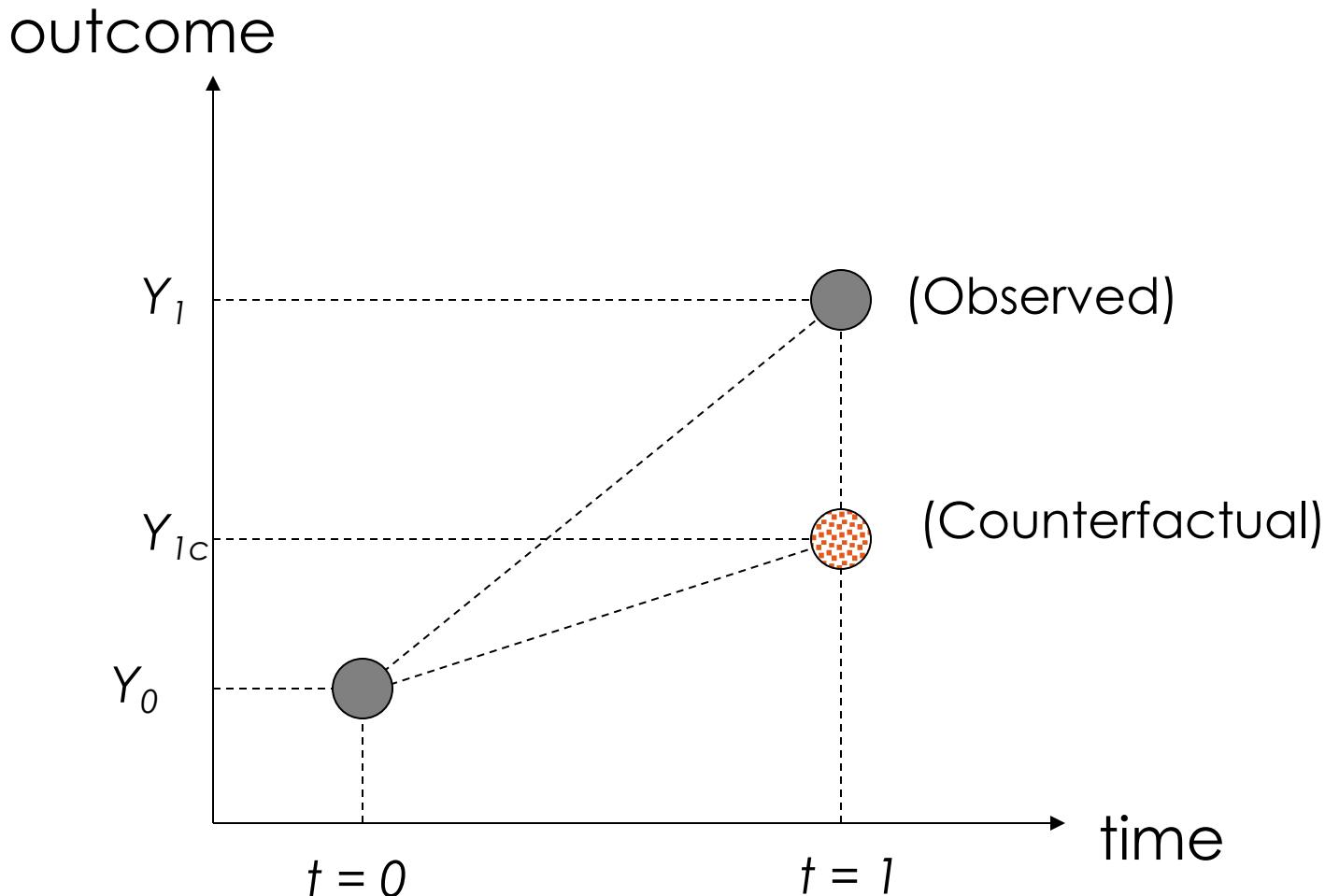
# Why is impact evaluation difficult?

- When we answer a process question, we need to describe **what happened**.
  - This can be done from reading documents, interviewing people etc.
- To determine the **impact** of the program we need knowledge of **counterfactuals**, that is, what would have happened in the absence of the program?
- Problem: The true counterfactual is **not observable**
  - The fundamental problem of impact evaluation is thus a problem of **missing data**
  - We don't know what would have happened in the absence of the program (the counterfactual)
- The key goal of all program/impact evaluation methods is to **construct** or “mimic” the counterfactual **as best as possible**.

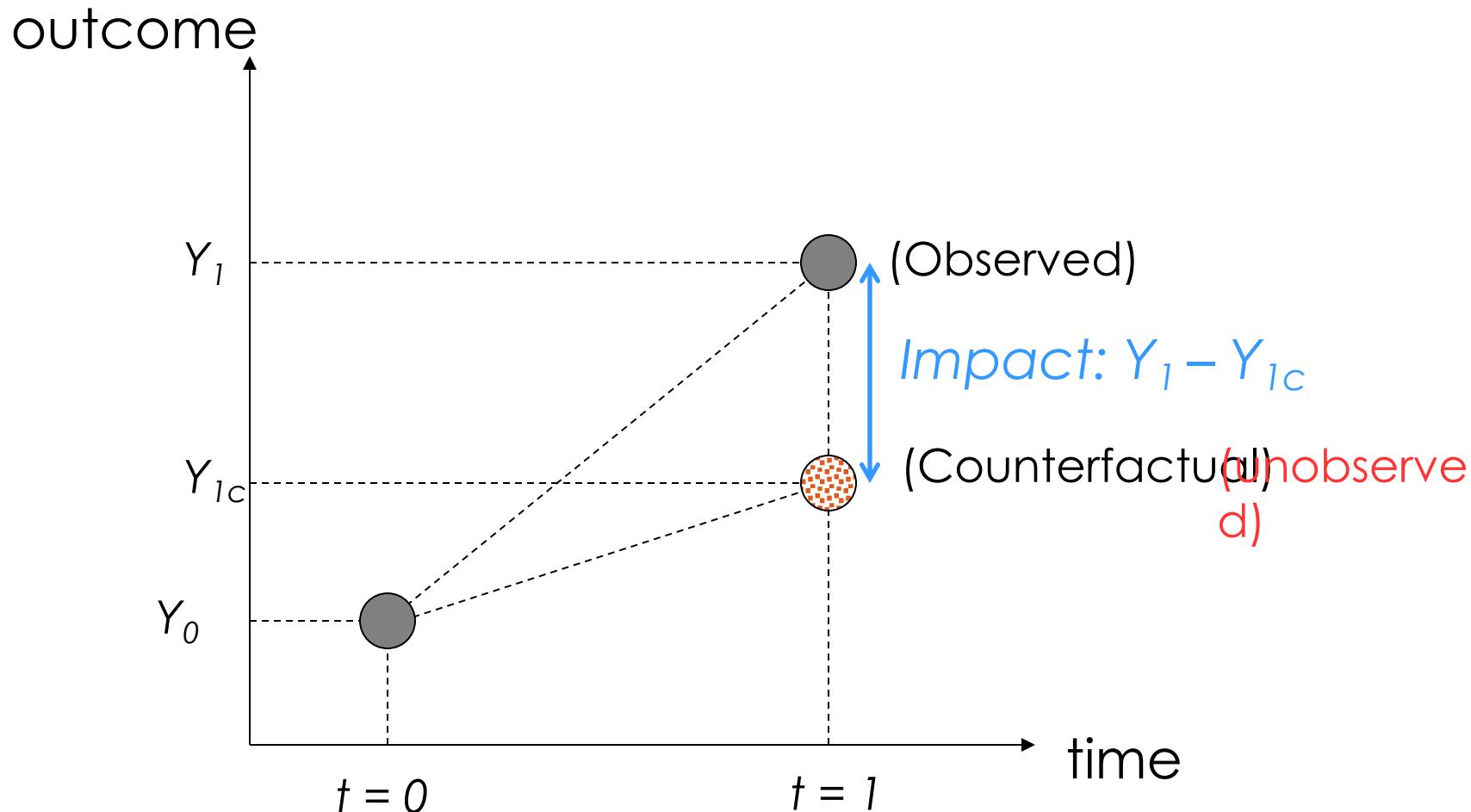
We observe an outcome ...



We need to identify the counterfactual:  
what would have happened in the absence of the  
program



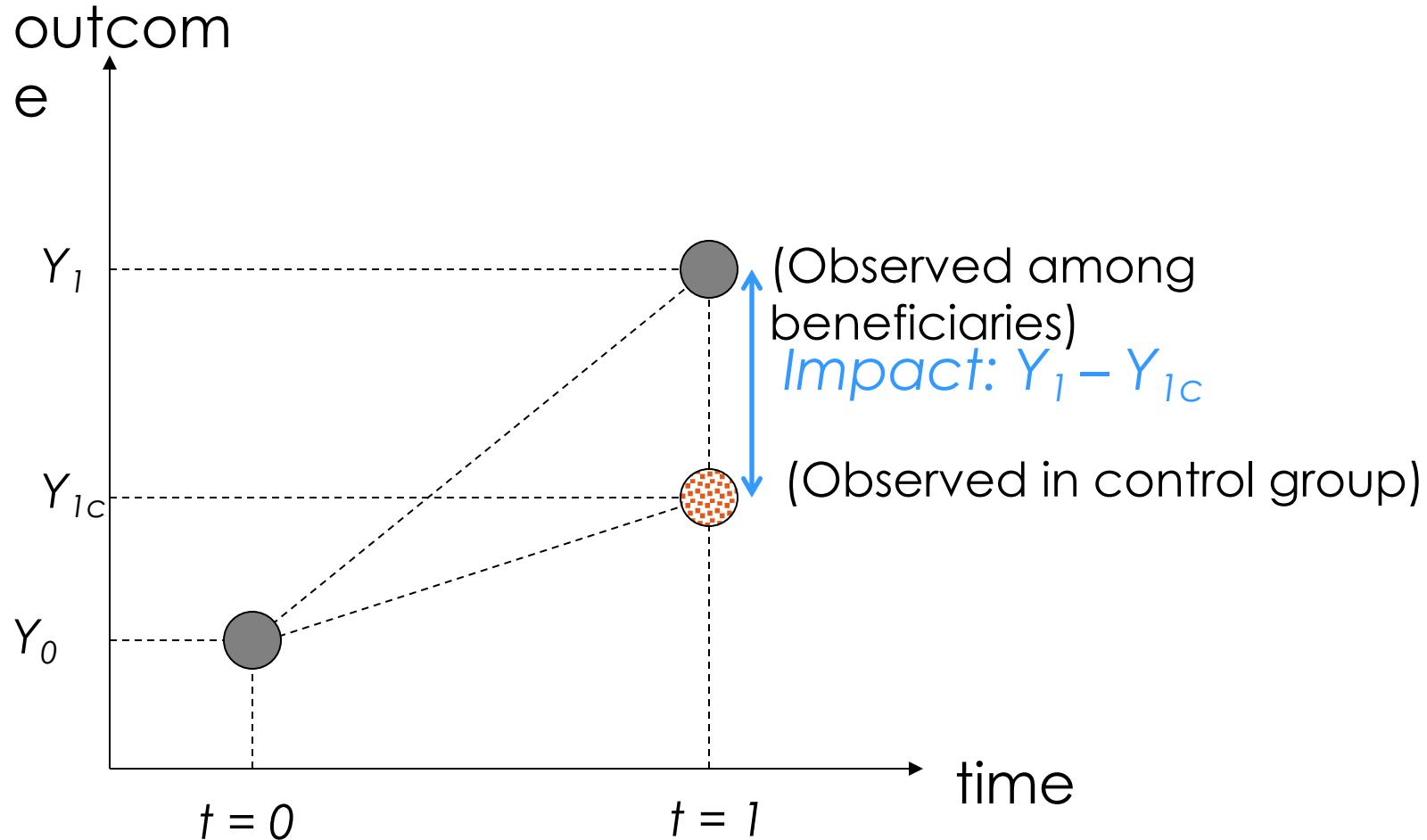
We need to identify the counterfactual:  
what would have happened in the absence of the  
program



# But the problem is...

- We will never have a child both with and without a school meal at the same time ...
- So the counterfactual is not observed
- Solution:
  - Use non-participants as point of comparison  
= “Control” Group
  - E.g.: use kids who did not

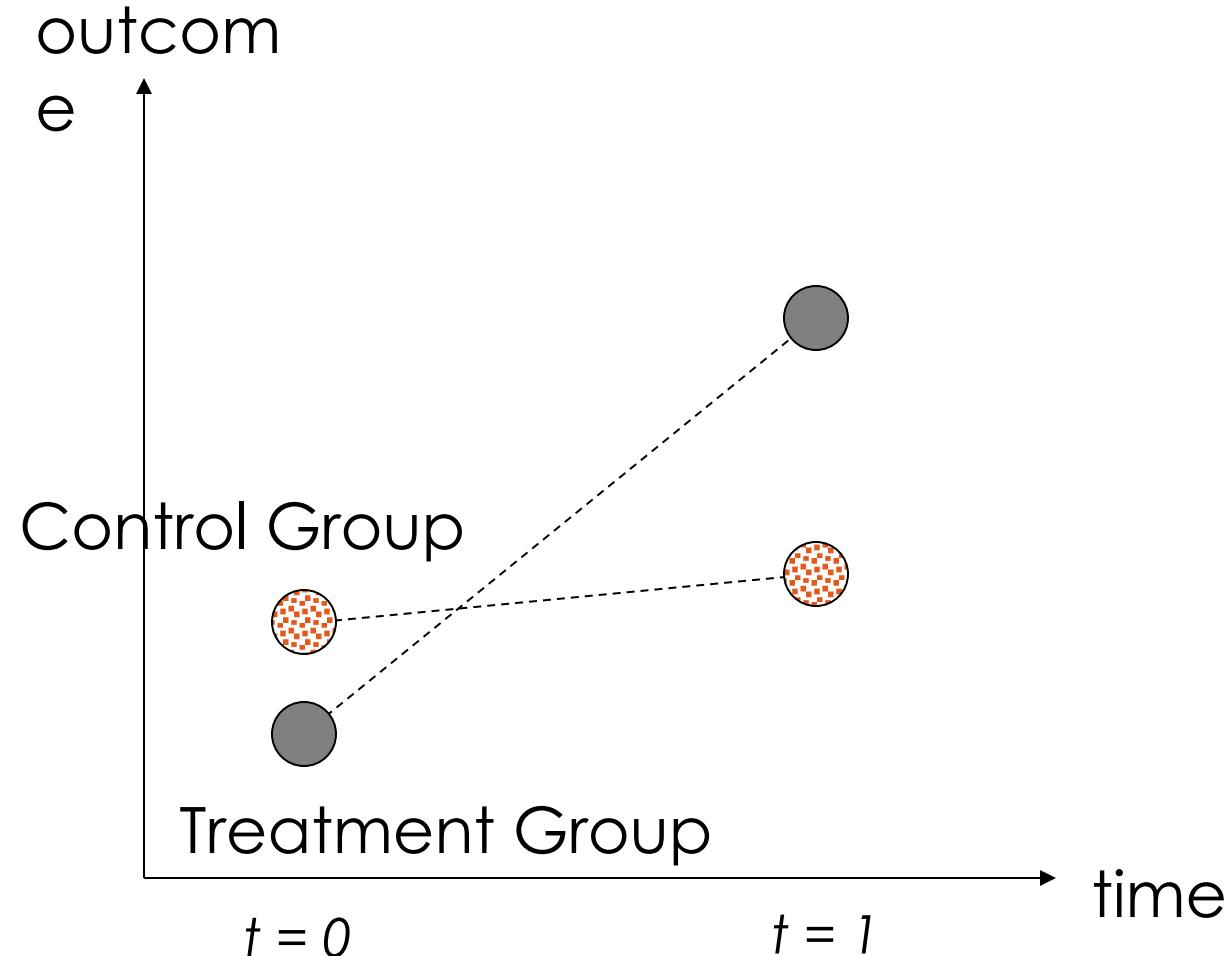
# Simple Difference



## But still...a few problems

- If there are differences in background characteristics between the group of participants and the non-participants
    - E.g., if only kids who are very poor are offered a school lunch
- This will bias the comparison ...
- This biased is called “selection bias”

# Selection Bias



## How to get rid of all possible selection biases?

- Random assignment of program to treatment and control group
- This creates a comparison group that is not systematically different from the participants
- *i.e.*, one that is not subject to any selection bias
- Why?

# Why does Random assignment work

- Because of the law of large numbers...
  - Take 200 villages and randomly split them into two groups of 100
  - The average participation
  - Note: not true if you have only 10 villages to split into 2 groups
- Suppose 50% of a group of individuals are randomly ‘treated’ to a program (without regard to their characteristics).
  - If successfully randomized, individuals assigned to the treatment and control groups differ only in their exposure to the treatment.
  - Implies that the distribution of both observable and unobservable characteristics in the treatment and control groups are statistically identical.
- Any difference between treatment and control can be attributed to the treatment

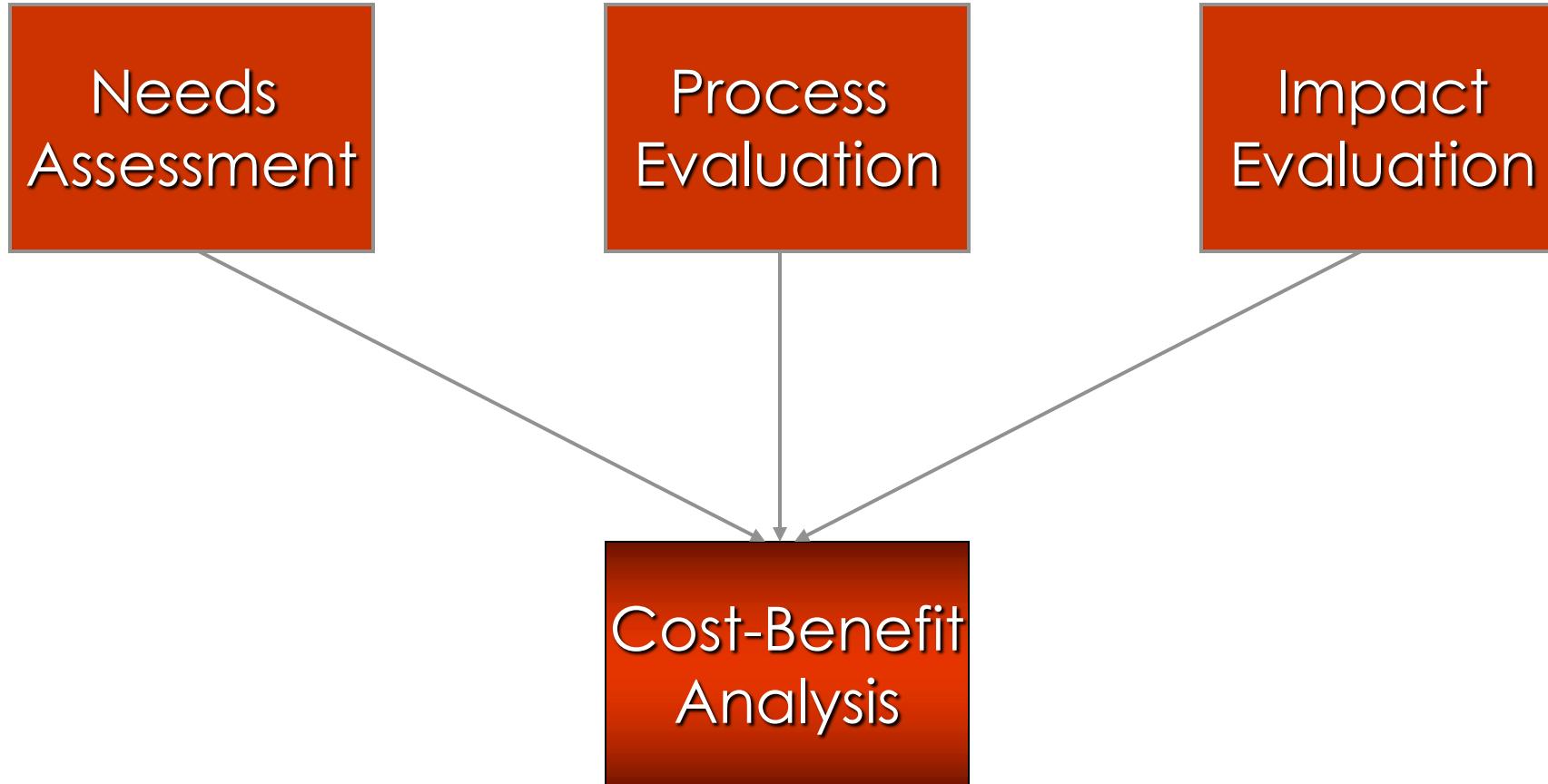
# Participation in education

- Reducing the cost of education:
  - Conditional Cash Transfers: PROGRESA in Mexico
    - 3.4% increase in enrollment on average. Larger impact at the secondary school levels.
  - School Uniforms in Kenya
    - School Uniforms distributed to 10,000 students in grade 6, and then 7 in 163 randomly selected schools
    - Drop out fell from 18% to 12% for girls, 13% to 9% for boys
- School meals
  - Evaluation for Pre-schools in Kenya: participation was 30% higher in schools where free breakfast was given

# Participation in education

- School health
  - Deworming in Kenya: 0.15 years of extra education (25% increase in presence)
  - Replicated in India (pre-school).
- Incentives for Students
  - Girls scholarship program based on good performance on tests scores in Kenya
- Informing parents about the returns to education
  - Madagascar: increase participation

# Cost Benefit Analysis



# Evaluation and cost-benefit analysis

---

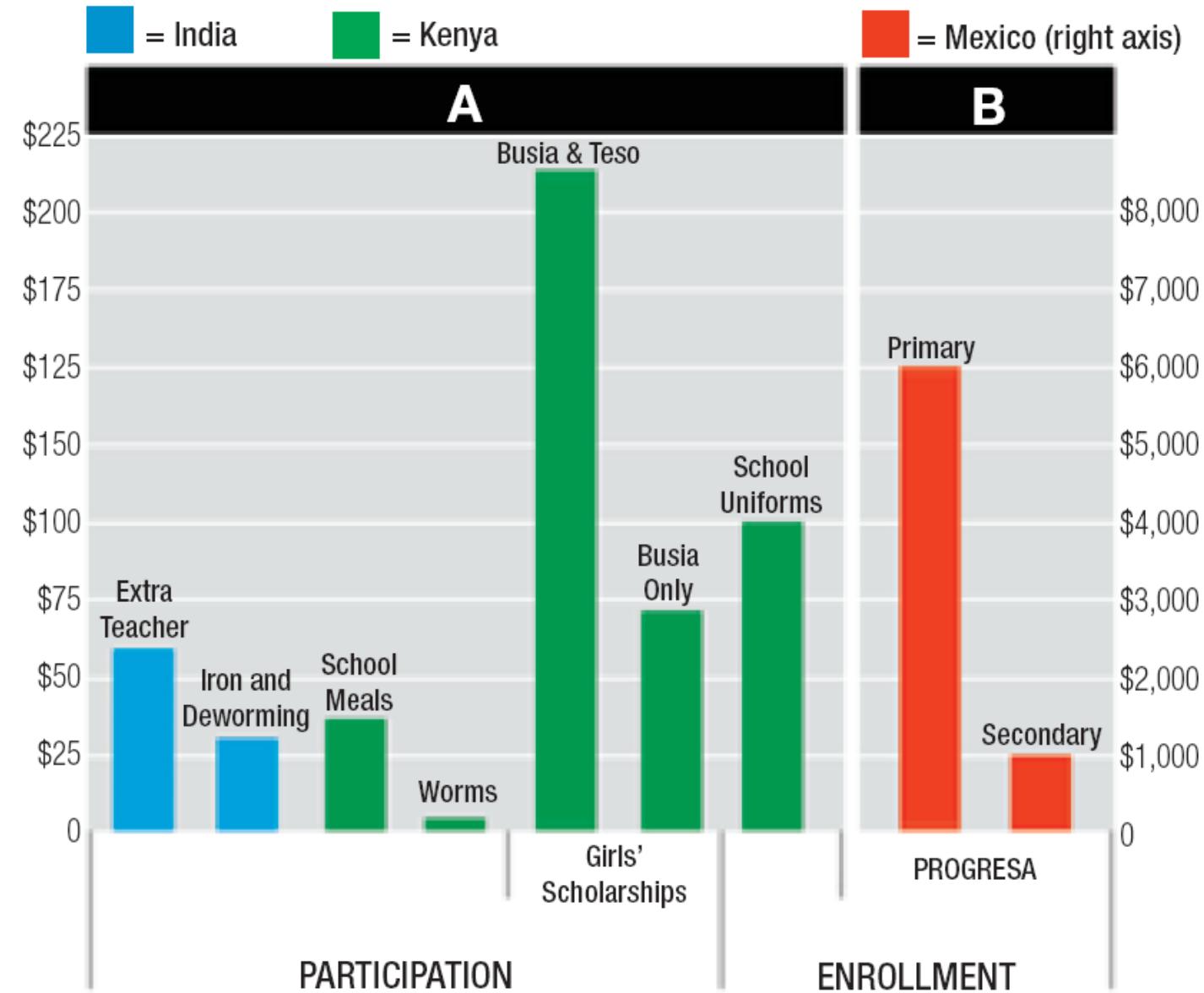
- Needs assessment gives you the metric for defining the cost/benefit ratio
- Process evaluation gives you the costs of all the inputs
- Impact evaluation gives you the quantified benefits
- Identifying alternatives allows for comparative cost benefit

# Cost benefit analysis

---

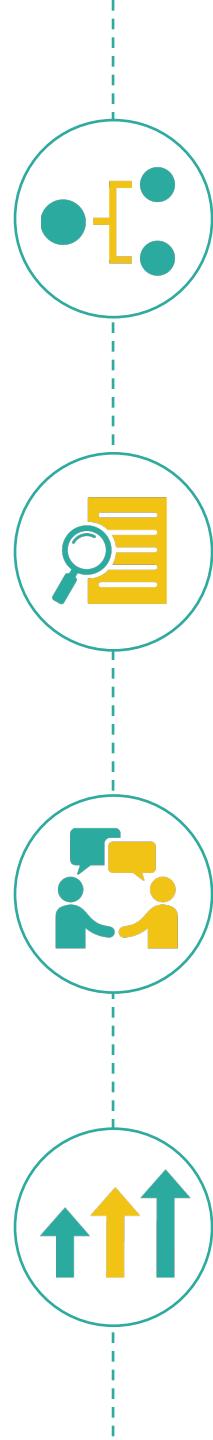
- Use the cost of the program to calculate how much it would have cost you to do this program for X children (e.g. 1000).
- Then use the program impact to calculate how many extra year of education you got for this 1000 children, thanks to the program.
- Deworming example:
  - Cost per child: 0.5 dollars per year
  - Increase in year of education: 0.15 years
- Deworming cost 3 dollars per **extra year of education induced**. This is different from the usual price per program that the so called “rating agencies” give you for NGOs, because now the price is put in perspective with the benefits.

## Cost Per Extra Year of Education Induced



# **From Evidence to Policy: The long journey**

## The strawman



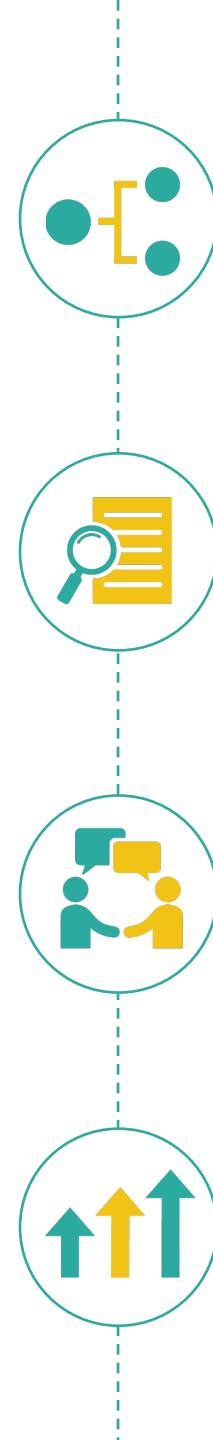
**Run a small, well controlled experiment**

**Get the results**

**Prepare a shiny policy brief and peddle to policy makers**

**Get full scale adoption**

# The strawman subject to all sorts of critics



## Run a small, well controlled experiment

"Gold plated experiments"—samples are too small

## Get the results

Results only valid in one place, might not replicate elsewhere;  
might not even be internally valid (imperfect take up, spillovers  
on non beneficiaries)

## Prepare a shiny policy brief and peddle to policy makers

May not fit with the policy makers interest at the time

## Get full scale adoption

Results will be quite different if adopted at scale: equilibrium  
effects, Political economy effects

**But really, it is not the way policy influence works**

# The Miracle of Microcredit?



# Microcredit in the 2000s

## The Seattle Times

Opinion

### Microfinancing changes lives around the world – measurably

Originally published April 7, 2010 at 10:21 am | Updated April 7, 2010 at 12:31 pm

Microfinancing can help poor people around the world with small loans that can change their lives, writes guest columnist Brigit Helms. The anecdotes are plentiful but a new study shows the benefit.

By [Brigit Helms](#)

#### Share story

 Share

 Email

 Tweet



### The Nobel Peace Prize 2006

Muhammad Yunus, Grameen Bank

Today, Grameen Bank gives loans to nearly 7.0 million poor people, 97 percent of whom are women in 73,000 villages in Bangladesh.

Grameen Bank gives collateral-free income generating, housing, student and micro-enterprise loans to poor families and offers a host of attractive savings, pension funds and insurance products for its members. Since it introduced them in 1984, housing loans have been used to construct 640,000 houses. The legal ownership of these houses belongs to the women themselves. We focused on women because we found giving loans to women always brought more benefits to the family.

In a cumulative way the bank has given out loans totaling about US \$6.0 billion. The repayment rate is 99%. Grameen Bank routinely makes profit. Financially, it is self-reliant and has not taken donor money since 1995. Deposits and own resources of Grameen Bank today amount to 143 percent of all outstanding loans. According to Grameen Bank's internal survey, 58 percent of our borrowers have crossed the poverty line.

# Then the tone shifted...

## Trapped in micro debt

Peace Prize winner Yunus' famous microcredit bank Grameen Bank takes 30 percent interest and brings many poor women into a debt spiral.



Focal point documentary "[Trapped in micro-debt](#)" shows a very different side of microcredit than Muhammad Yunus and his peace-prize-winning bank Grameen Bank enjoy in public.

### debt Spiral

Surveys Danish journalist Tom Heinemann has done for NRK Brennpunkt through several trips to Bangladesh and talks with a number of international experts around the world, shows that Grameen Bank leads many poor women into a devastating debt spiral.

The women pay about 30 percent interest on the loans, which they must already start paying back after a week. In the documentary, the poor tell about hard collection methods from Grameen Bank, which received NOK 400 million in assistance from Norway.

The Food First website features a prominent red header with the organization's name in a large, white, cursive font. Below the header is a navigation bar with links to "Press Room" and "Blog". A secondary navigation bar below the main one includes "ABOUT US", "BOOK RELEASES", "ISSUE AREAS", and "PUBL". The main content area contains a large, framed photograph of three people: an older woman in a green sari, a young girl in a green sari, and a young boy in a green shirt. Below the photo is the title "The Limits of Microcredit: A Bangladesh Case" and the author's name "Jason Cons and Kasia Paprocki | 12.01.2008". A quote is presented in a large, stylized font, flanked by double quotes. The quote discusses the failure of microcredit to achieve poverty alleviation, financial independence, and gender equality in rural Bangladesh.

**Food First**

Press Room   Blog

ABOUT US | BOOK RELEASES | ISSUE AREAS | PUBL

The image shows three individuals: an elderly woman in a green patterned sari, a young girl in a green sari, and a young boy in a green shirt. They appear to be in a rural, possibly impoverished environment.

### The Limits of Microcredit: A Bangladesh Case

Jason Cons and Kasia Paprocki | 12.01.2008

In rural Bangladesh, microcredit is not achieving its core goals of poverty alleviation, financial independence, and gender equality.

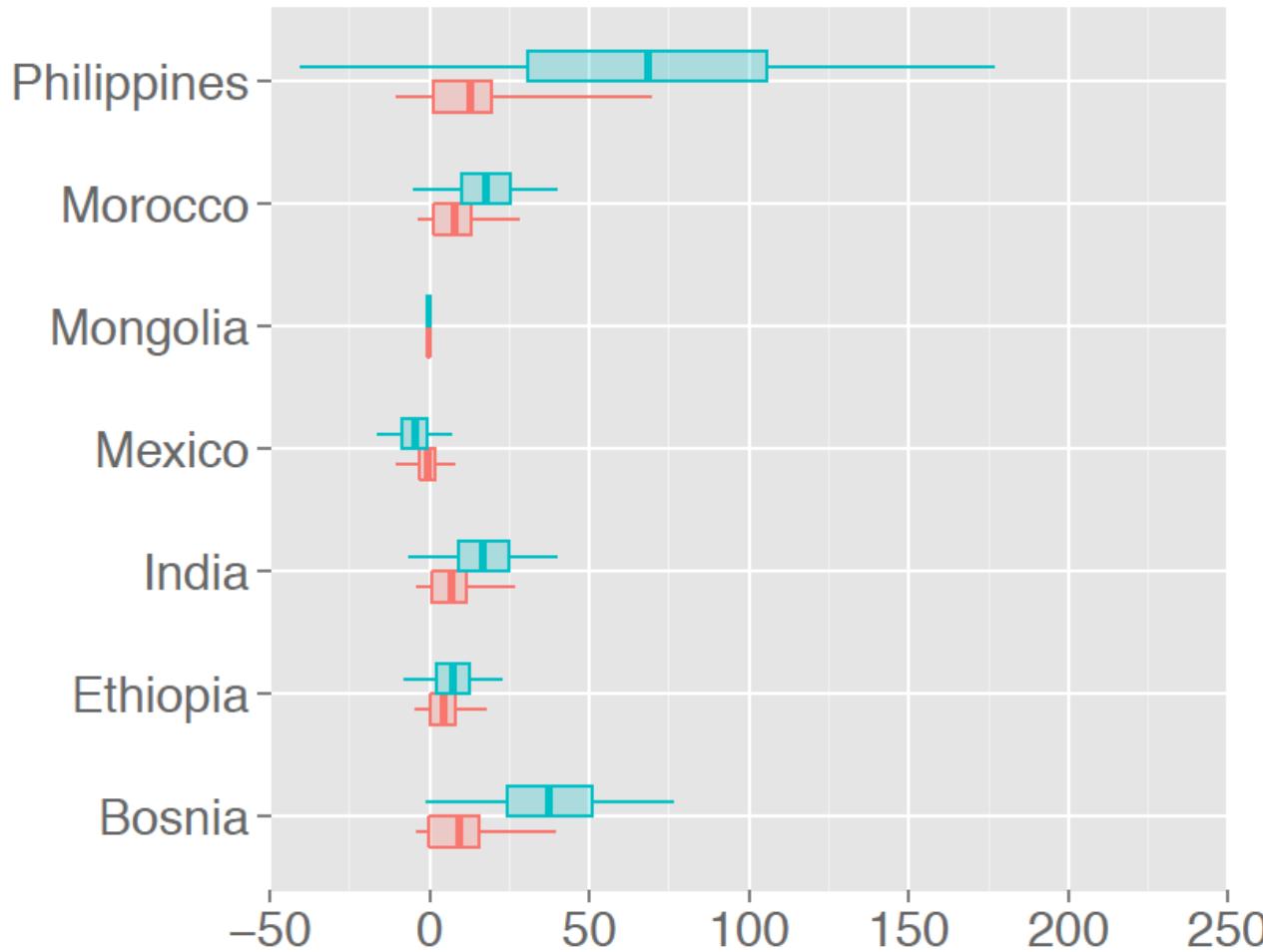
# The evaluations

- The first evaluations conducted where in India and in the Philippines.
  - Philippines: rather good
  - India: rather weak
- India is a very unique context: hotbed of microfinance
- To get a useful answer, we had to wait...
- Until seven studies came out together, all in different contexts and carried out by different teams
- But published together with the same outcomes.



Photo: iStock.com

# Bayesian Hierarchical Modelling of all the MF results: Profits



BHM Posterior  
OLS

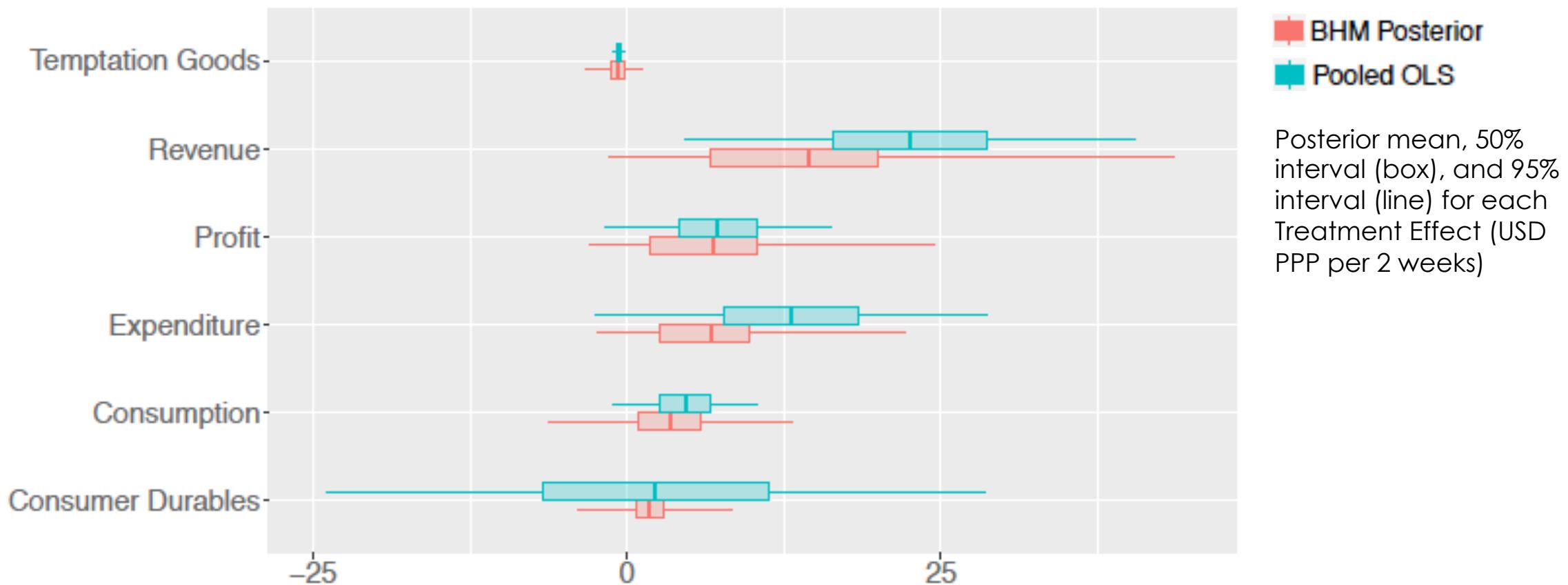
Posterior mean, 50%  
interval (box), and 95%  
interval (line) for each  
Treatment Effect (USD  
PPP per 2 weeks)



**Rachel Meager**  
LSE

# Meta-analysis: Average estimated effect and range, 6 countries

## Posterior distribution of average effect



# Changing the debate



© THE WALL STREET JOURNAL.

TUESDAY, MAY 22, 2012

Home      Budget 2012      Student Enterprises      The Small Picture

Views RSS

## The Microfinance Mess

While the current attacks on microfinance are unwarranted, overblown claims about it also need to be tempered

Economist.com FINANCE & ECONOMICS  
**ECONOMICS FOCUS**

Economics focus

### A partial marvel

Jul 16th 2009  
From The Economist print edition

Microcredit may not work wonders but it does help the entrepreneurial poor



Small is smart  
Maitreesh Ghatak  
Posted online: 2009-08-24 21:49:32+05:30

Microfinance is big these days. It may be called 'micro' but there is nothing small-time about it. It serves more than worldwide through various microfinance programmes (MFPs). While its true origins are lost in the mists of time, Yunus and the Grameen Bank of Bangladesh are often cited as the microfinance institution (MFI) of the modern era.

HOME / GLOBE / IDEAS

### Small change

Billions of dollars and a Nobel Prize later, it looks like 'microlending' doesn't actually do much to fight poverty

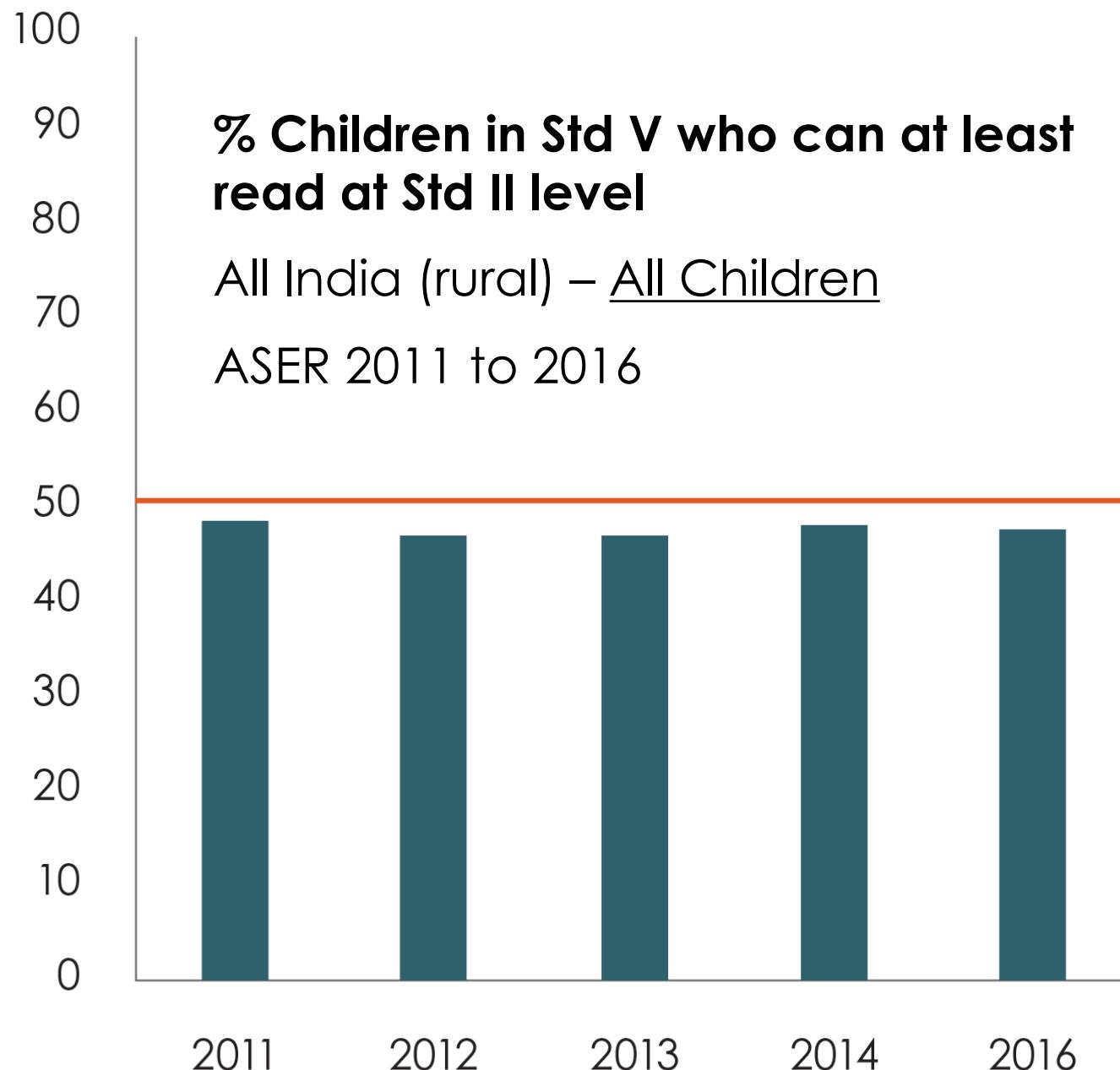
The Boston Globe

# Changing microfinance

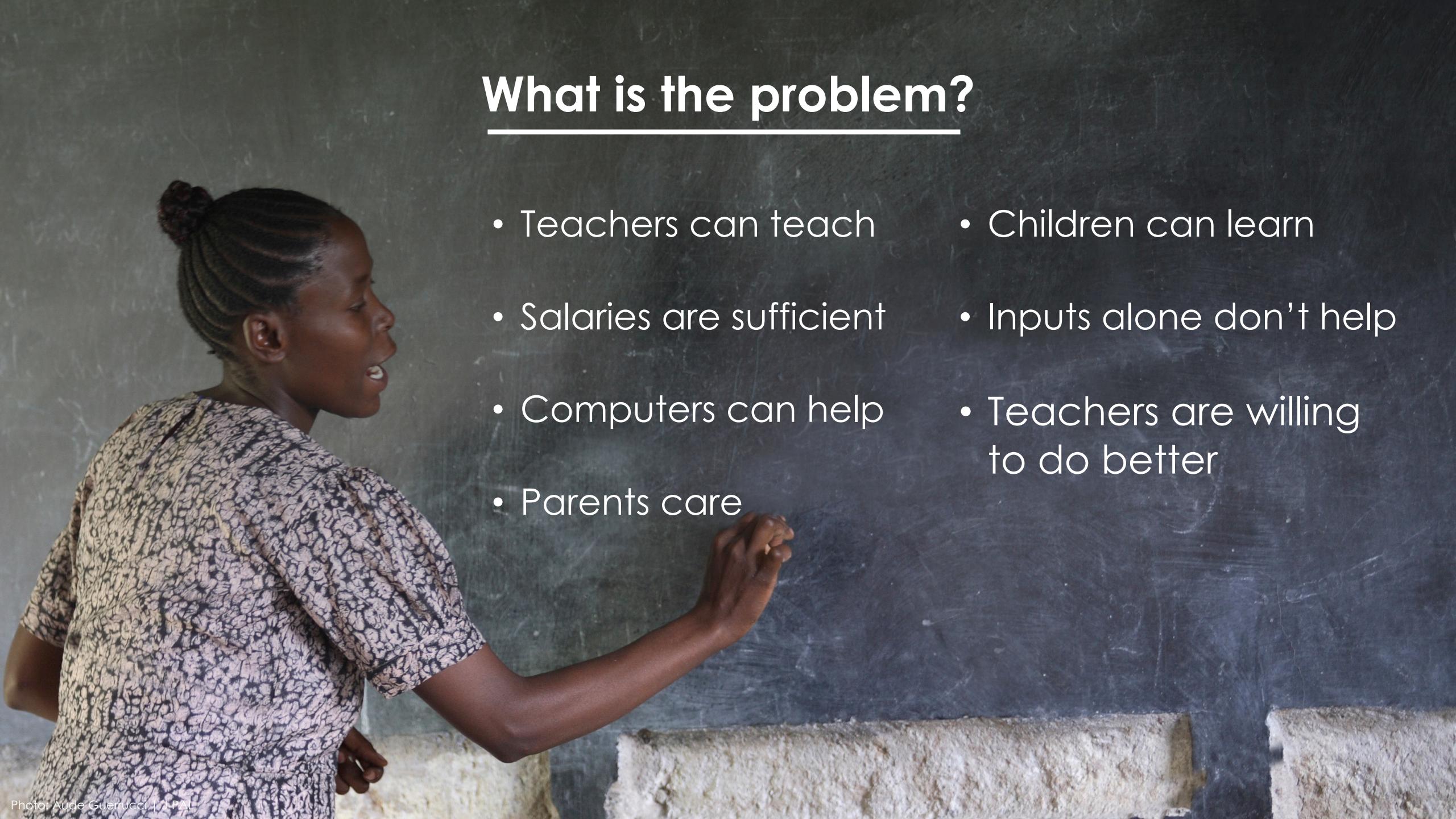
- One sized approach (small term loan) does not work for all.
- Many experiments since then have focused on the terms of lending.
- Focus on all the financial services the poor need, and the right fit for each
  - Transaction
  - Savings
  - Insurance
  - Ultra poor programs
- Focus on heterogeneity: Creating methods to identify real entrepreneurs who will benefit from a microcredit.
- Interplay between theory and experiments is what makes economics useful to policy (and policy useful to economics)

# Teaching at the right level





# What is the problem?

- 
- A young girl with dark skin and her hair in braids is shown from the side, facing right. She is wearing a light-colored dress with a dark, intricate pattern. Her right arm is extended towards a large, dark, textured chalkboard, holding a piece of chalk and appearing to write or draw on it. The chalkboard has some faint white marks on its surface.
- Teachers can teach
  - Children can learn
  - Salaries are sufficient
  - Inputs alone don't help
  - Computers can help
  - Teachers are willing to do better
  - Parents care

## TaRL Support Pieces



## In a TaRL classroom

### **STEP 1**

#### ASSESSMENT

Test children on the basics using simple tools

### **STEP 2**

#### GROUPING

Create homogeneous learning level groups

### **STEP 3**

#### FOUNDATIONAL SKILLS

Focus on basic skills for a period of the day or year



**LEARNING OUTCOMES IMPROVE**

Children are reassessed and moved through the levels as they progress.



Dr. Rukmini Banerji



Dr. Mdhav Chavan

# For 15+ years of experimentation

**2001-2003**

"Balsakhi" program; Pratham community volunteer "pull out" remedial program in urban schools



**2000**

**2005-2006**

Village volunteers conducted community classes for rural primary school children



**2008**

In-school one month gov't teacher-led summer camp with support by rural village volunteers



**2005**

**2008-2010**  
In-school gov't teacher-led learning improvement program & support by Pratham volunteers (rural)

**2010**

**2010-2013**

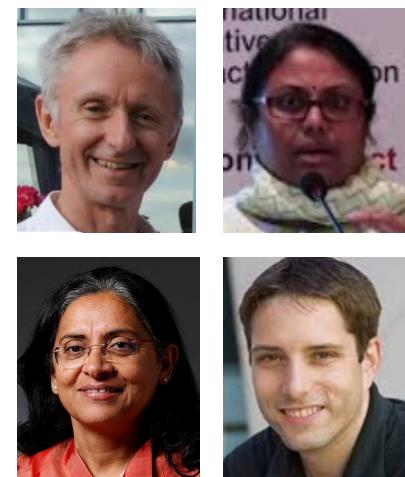
Ghana trials of teacher-led vs. tutor-led in school and out of school



**2013-2014**

"Learning Camps" in gov't primary schools; led by Pratham teams supported by village volunteers

**2015**



# Teaching at the right level today



# Improving programs that run at scale



Helping government address  
the “plumbing problems”

# Evaluating programs at scale: Targeted information to improve rice distribution



# Raskin Social Assistance ID Cards

- Raskin is Indonesia's US\$2.2 billion rice subsidy program for poorest families
- Problem:
  - Poor receive just 30% of the intended subsidy
  - Pay 25% more for Raskin



Photo: J-PAL



## KARTU PERLINDUNGAN SOSIAL ( KPS )

**Kartu Perlindungan Sosial (KPS)  
Hanya Untuk Yang Miskin!**  
*Selengkapnya mengenai KPS [...]*

# Experimental design

The experiment varied 4 aspects of the cards

- Information on the cards
- Who gets the card
- Common knowledge through posters
- Create impression of accountability



Researchers:



Abhijit Banerjee



Rema Hanna



Jordan Kyle



Benjamin A. Olken



Sudarno Sumarto

# Results

Banerjee, Hanna, Kyle, Olken, Sumarto

- Poor families get 26% increase in subsidy
- Driven by reduction in leakage
- Cost Effective: \$1 for ~\$8 increase in subsidy



Photo: Ben Olken | J-PAL

# Scale-up

- Government rolled out “social protection” cards in 2013 to 15.5 million poor families, reaching 66 million people
- Continued partnership to improve service delivery, with planned evaluations of a new reformed social benefit scheme to be implemented through electronic vouchers



Photo: Ritwik Sarkar | J-PAL

# Reforming the auditing of firms in Gujarat



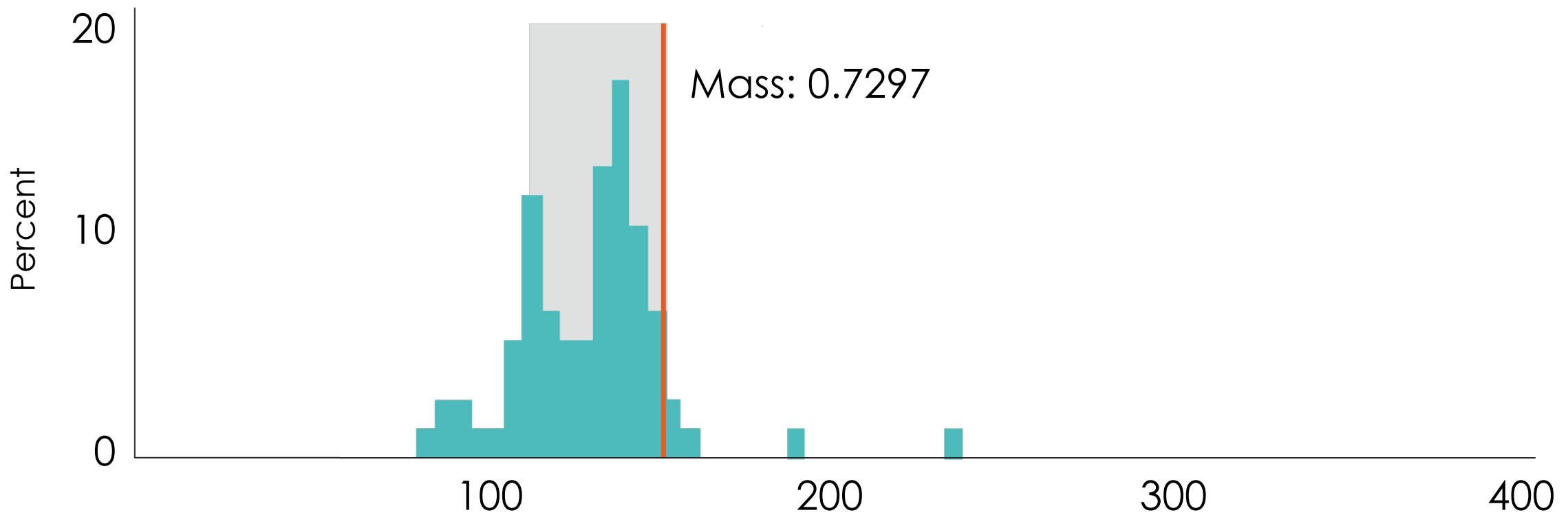
# Third party audit



Photos: Vipin Awatramani | J-PAL

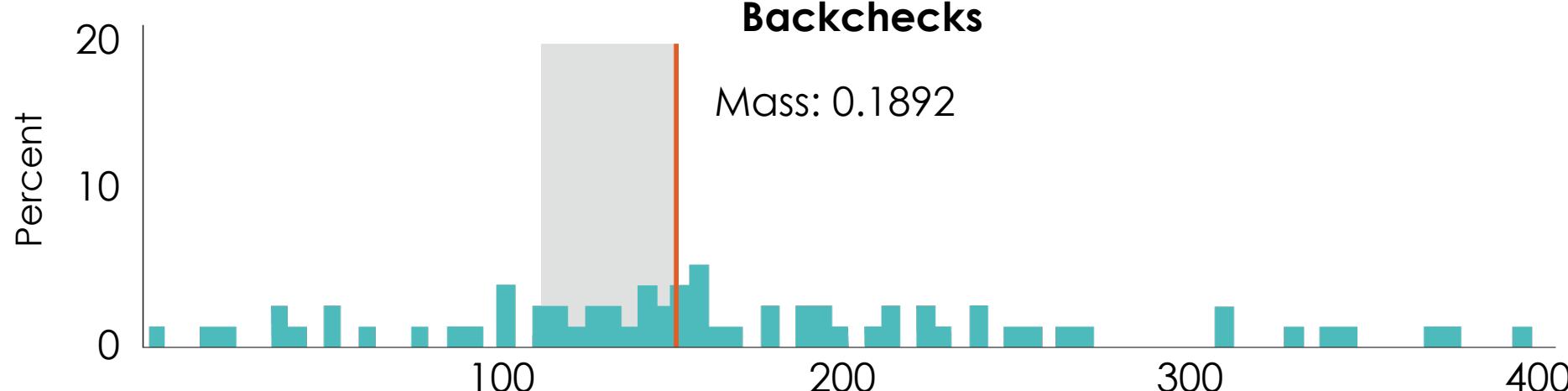
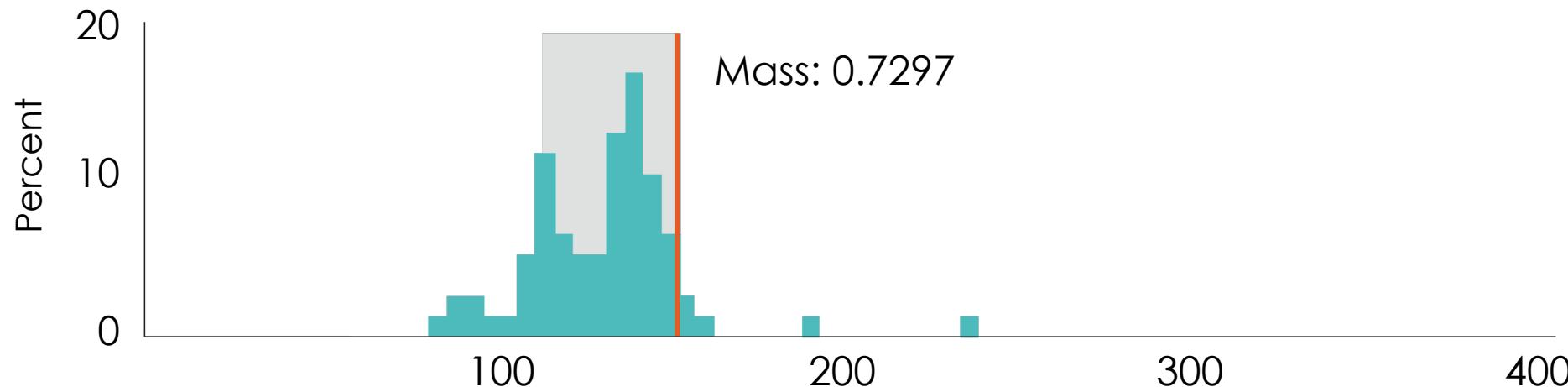
## What audits say...

### Suspended particulate matter, mg/Nm<sup>3</sup> | A. Control, Midline



# Vs reality

Suspended particulate matter, mg/Nm<sup>3</sup> | A. Control, Midline



# The reform we proposed

Duflo, Greenstone, Pande, Ryan

1. Random assignment of auditors and fixed payments from a central pool
2. Back check auditors for accuracy
3. Payment (or continuation with the scheme) based on accuracy
  - Ideas based on basic economics, and a solid understanding of the institutions

## Researchers:



Esther Duflo



Michael Greenstone



Rohini Pande

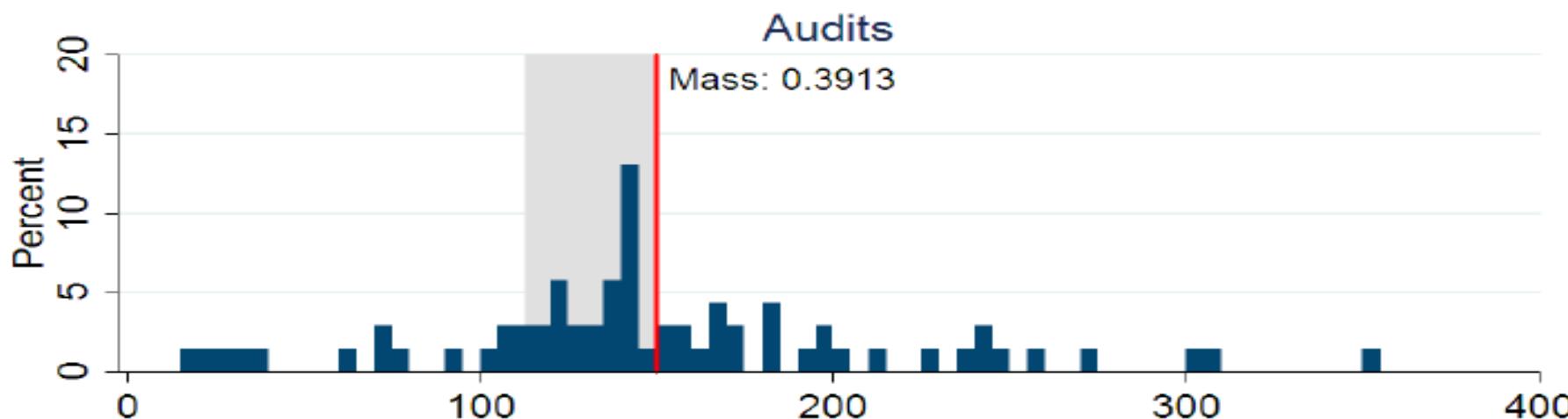


Nicholas Ryan

# Impact of the reform

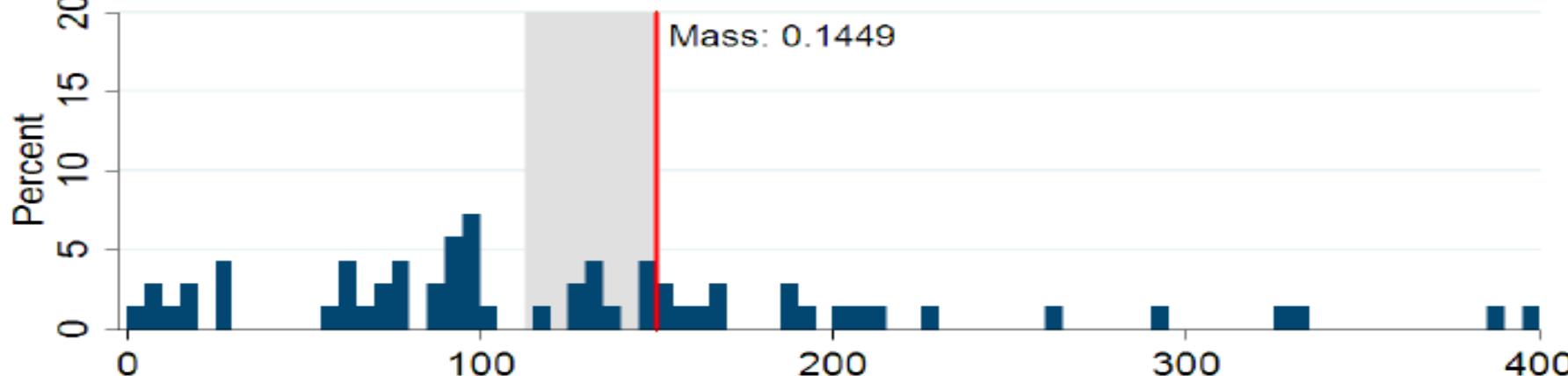
Suspended particulate matter, mg/Nm<sup>3</sup>

B. Treatment, midline



Backchecks

Mass: 0.1449



# Fostering a culture of learning inside Governments

- Many governments have launched either long run partnership with J-PAL or their own “learning units” (e.g. Minedulab in Peru, Tamil Nadu research partnership).
- World Bank Supports hundreds of RCT and training with various governments
- Many of the governments we meet want us to run an RCT, rather than listen to any evidence we might bring.
- May be one day we can make ourselves irrelevant...