Learning from rigorous evaluation for more effective policy

Abhijit Banerjee

MIT and the Abdul Latif Jameel Poverty Action Lab (J-PAL)

A critical moment in history

- The challenge of managing the effects of climate change seem to be on us
  - Mostly a problem for the countries of the South
  - Estimates that under the business as usual scenario, life expectancy will drop by 10 years in India by 2100.
    - No change in the global North
- AI is poised to change the world we live in potentially for better and for worse
  - Immediate prospect of job loss in countries where the BPO sector is large
    - These are some of the rare « good » jobs
  - More long-term job loss everywhere

Recent crises threaten to push another

75 - 95 million people into extreme poverty

in 2022 (World Bank, 2022)

Urgent policy challenges

- How to get the most carbon-saving technologies adopted?

- How to identify the losers from climate change and AI if we want to compensate them?

- Is it income or work that we are trying to protect?

- How to design transfers for the most effective redistribution?

- How do we stimulate the creation of the most interesting new jobs?

More long-term questions

- How to improve the quality of the workforce (especially to work with AI?)?

- How to encourage the development of green/post-climate change technologies that are suited to various developing country contexts?
  - Lots of innovation in agriculture for example but very little of it suitable for developing countries

- How use AI to deliver better healthcare in remote locations?

- How to use the power of AI to improve financial access?

J-PAL is working on all of these, often in multiple locations

J-PAL uses randomised evaluations to learn how to allocate scarce resources

How do you choose between seemingly good program options to start with?

Pair knowledge of local conditions with rigorous evidence from around the world

How do you know whether the chosen program really works?

Understand how data and impact evaluations can help improve your effectiveness

And if it is effective, is it also scalable?

○ Cost effectiveness

○ Generalizability

Why the emphasis on randomised evaluations?

Well-designed randomised evaluations can test <u>specific hypotheses</u> and single out their causal impact

- They allow us to **rigorously test the individual elements of a policy**

- They provide rich insight into the ***mechanisms*** that lead to change

They often can surprise us

- Surprisingly little is known about what works & decisions are often taken **based on "intuition"**

- These intuitions are faulty, and often **based on ideology**
  - "We should never pay people to do what they should do on their own"

- And the general principles, even when correct, **are too general**
  - "Improving education is difficult without more resources"
    - What resources: teachers, books, charts?

What types of questions can we answer?

Does online tutoring support children's learning effectively?
Researchers designed and evaluated online tutoring for disadvantaged students in Italy while schools were closed during lockdown. This is now piloted in the Dominican Republic.

How does a universal basic income experiment help households respond to Covid-related income shocks?
We evaluated how a UBI experiment in rural Kenyan villages impacted economic and food security as well as health

How to reduce leakages in the delivery of social support programmes in Indonesia?
We worked with the Indonesian government to distribute ID cards with eligibility status and price information for eligible households.

How does an emissions trading scheme affect air quality and firms' profits?
Researchers partnered with Gujarat Pollution Control Board to evaluate the impact of the first ever emissions trading programme for particulate air pollution on air quality and compliance costs for industrial plants.
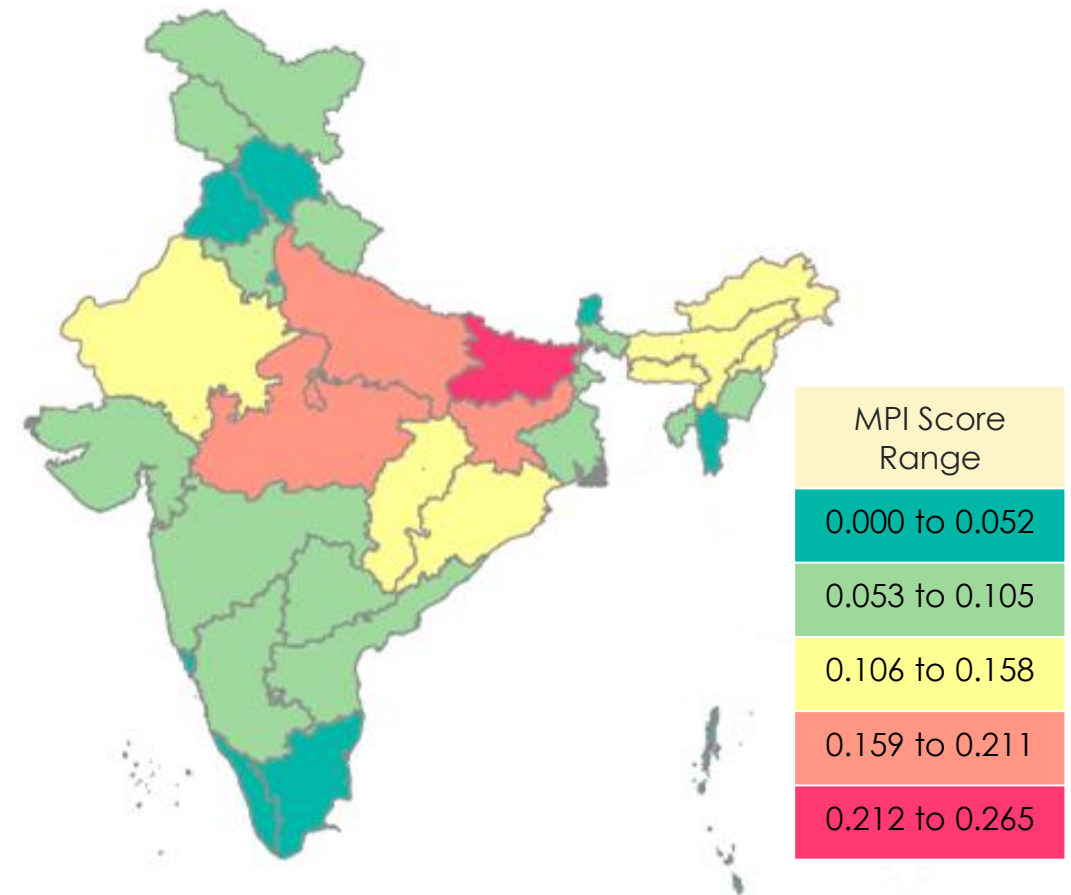
Ultra-poor households often face a variety of impediments to sustainably transition out of poverty

At around **25%** of the country's population, India has the largest number of people living in **multidimensional poverty** in the world

Poverty traps - are people poor because they are poor?

    ○   Example credit: Poor people cannot borrow enough to make necessary investments

Source: National Multidimensional Poverty Index:
https://www.niti.gov.in/sites/default/files/2021-11/National_MPI_India-11242021.pdf
Source: Socio Economic and Caste Census (2011)

Multidimensional Poverty Index across States and UTs

MPI Score Range

MPI Score Range
0.000 to 0.052
0.053 to 0.105
0.106 to 0.158
0.159 to 0.211
0.212 to 0.265

A big push might set off a virtuous cycle that takes them out of poverty

BRAC's Graduation programme

- Targeted to the poorest of the poor, who are often unable to reach mainstream support

- Package consisting of some assets, some training/handholding, some temporary cash assistance

- RCTs in 7 countries

Photo: BRAC

What drives poverty traps?

- Technology?

- Or psychology?

- The Graduation programme addresses both

  - There are assets but also hand-holding
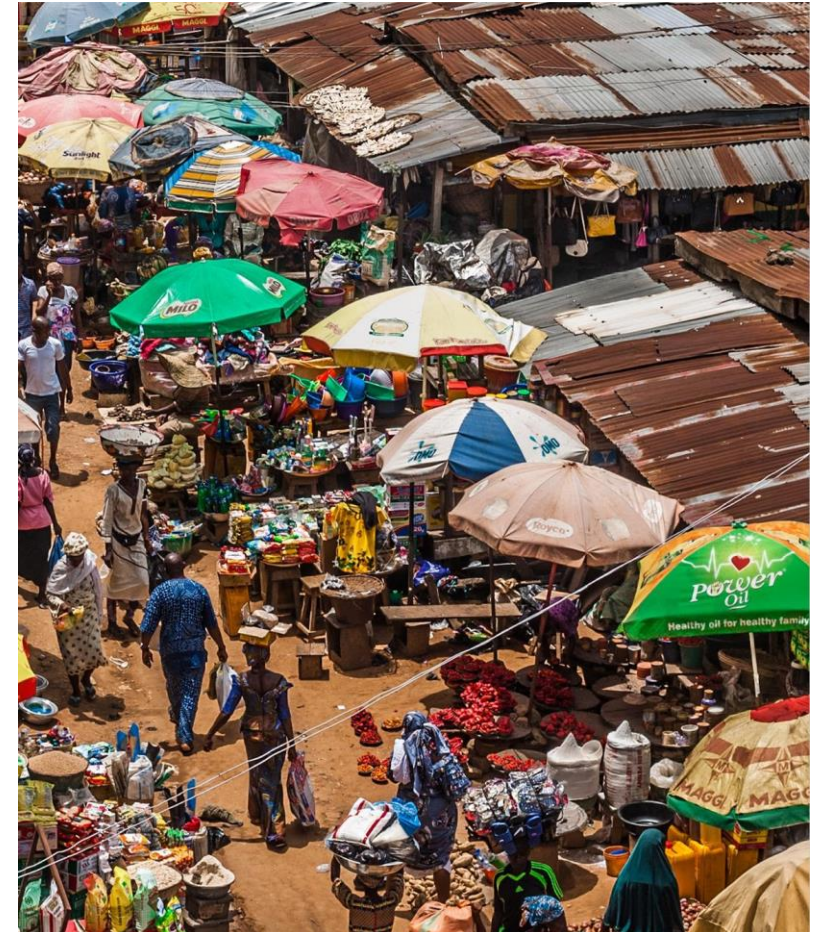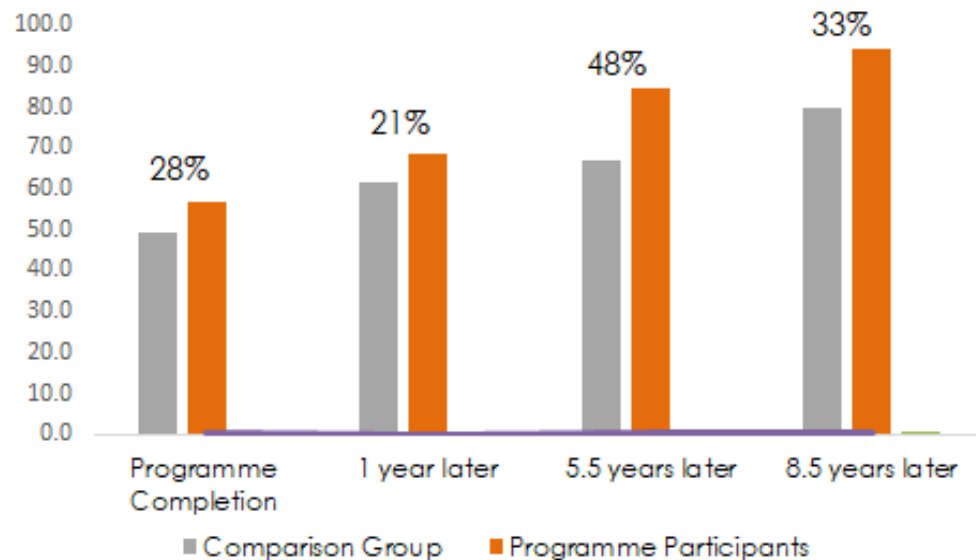
- Are both important?
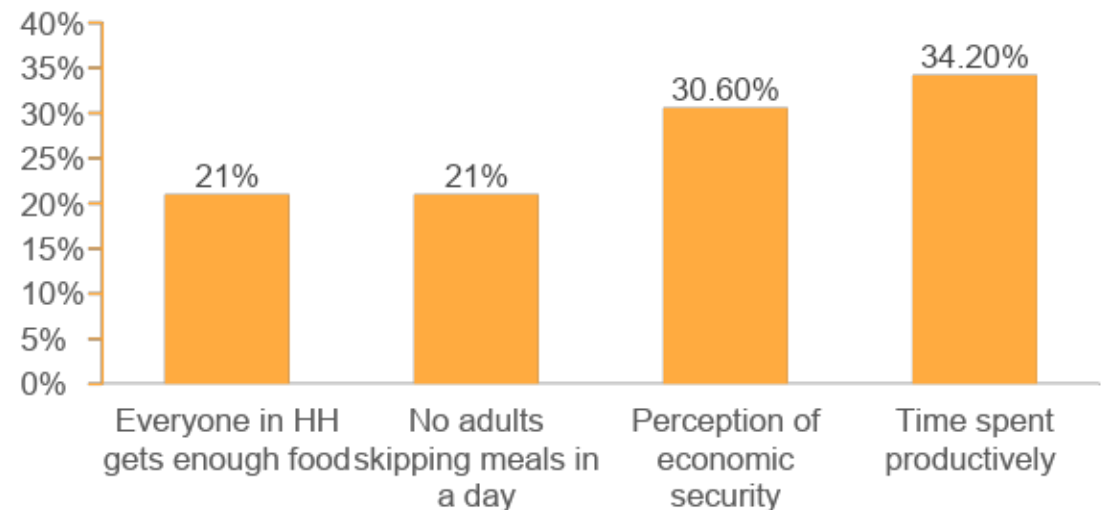

Photo: Toluade | Shutterstock.com

The Graduation approach was tested using a randomised evaluation in West Bengal, India

Measures of economic well-being **improve significantly and substantially faster** among treated households until year seven from the asset transfer, and the **resulting treatment control gap persists between years seven and ten.**
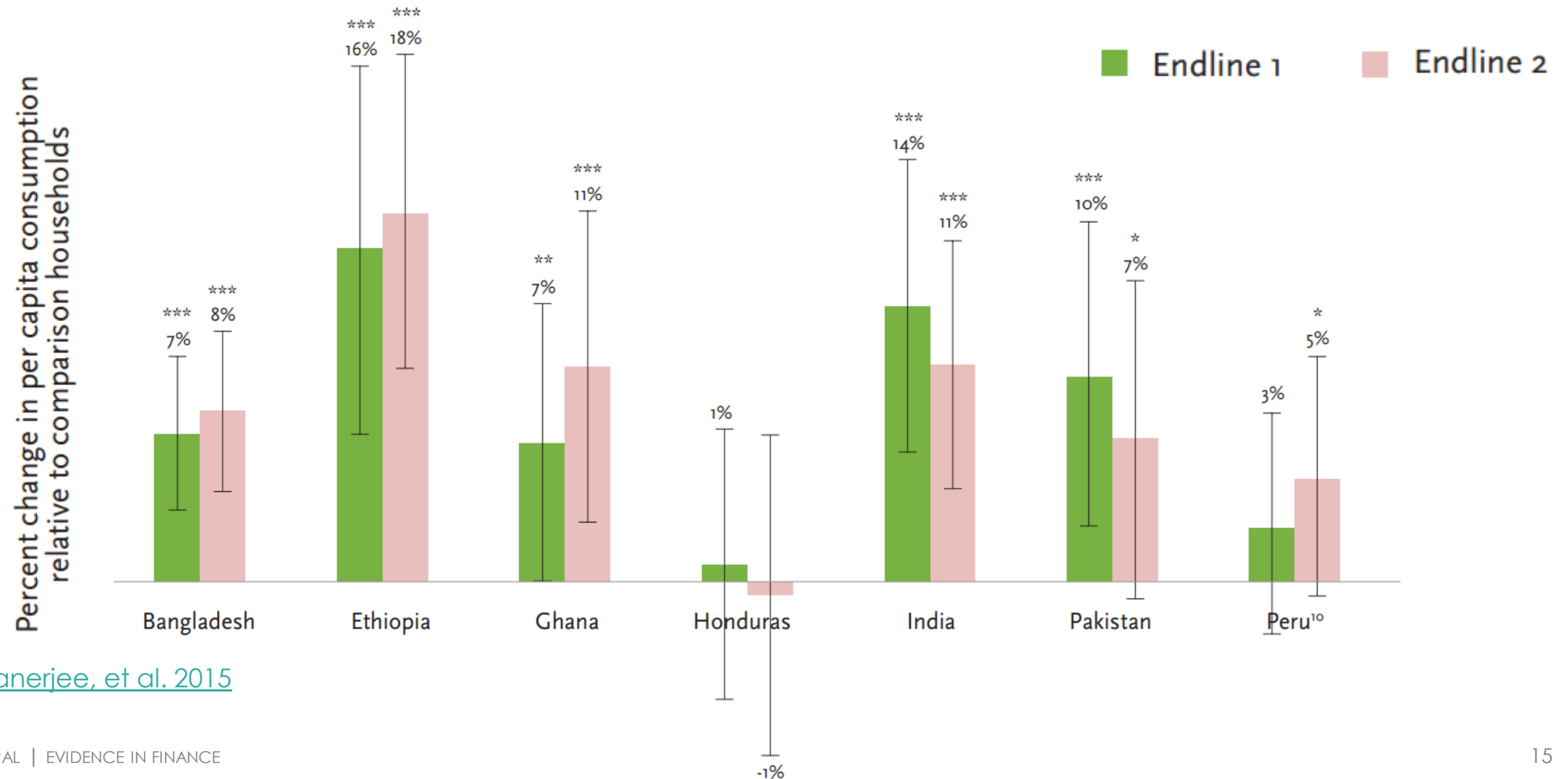
Consumption gains: compared with those who did not get the programme

Gains for beneficiaries, 8.5 years after graduating from the programme, as percentage increase against the comparison group

The Graduation programme caused broad and lasting economic impacts across countries

Banerjee, et al. 2015

What is driving this?

- ? The accumulation of RCT evidence suggests that feeling poor paralyses people

    – Consistent with evidence from Kaur, Mullainathan, Oh and Schilbach (2019)

- Which may explain why getting extra income does not make the poor lazy

- Indeed this exact question was motivated the first RCT that we are aware of: the Negative Income Tax experiment in the US

    – Heckman (1991) points to it as the ideal example of an RCT being used cumulatively unpack the elements of a theory

Scaling the model

BRAC scaled their approach to reach over 2 million households in Bangladesh, as well as an additional 1.1 million households in fourteen other countries

Plans to help a further 21 million people lift themselves out of poverty by 2026

New research focuses on how to most effectively optimize the model's components in different settings — including a model with a less-intensive coaching component and a version for urban areas

Summary on learning from poverty traps through randomised evaluations

- It is possible to durably increase earnings, at least for the poorest of the poor

- Support (but not quite proof) that poverty traps are real

- Very important for the design of social policy
  - Is a big push needed?

Randomised evaluations were the right method to answer these questions as they allow us to test specific hypothesis, measure causal impact, and understand mechanisms

Randomised evaluations uncover

where our theories were wrong

Economic theory is only a starting point in designing effective policies

Economic theory is **based on assumptions**, which come out of our need to limit the complexity of reality

➡️ Theory is always and necessarily **incomplete**

These assumptions are **about human behaviour** and the environment in which humans live and act

➡️ The assumptions are often **untested and unstable**

Economic models are very good at **alerting** us to various **possible mechanisms**
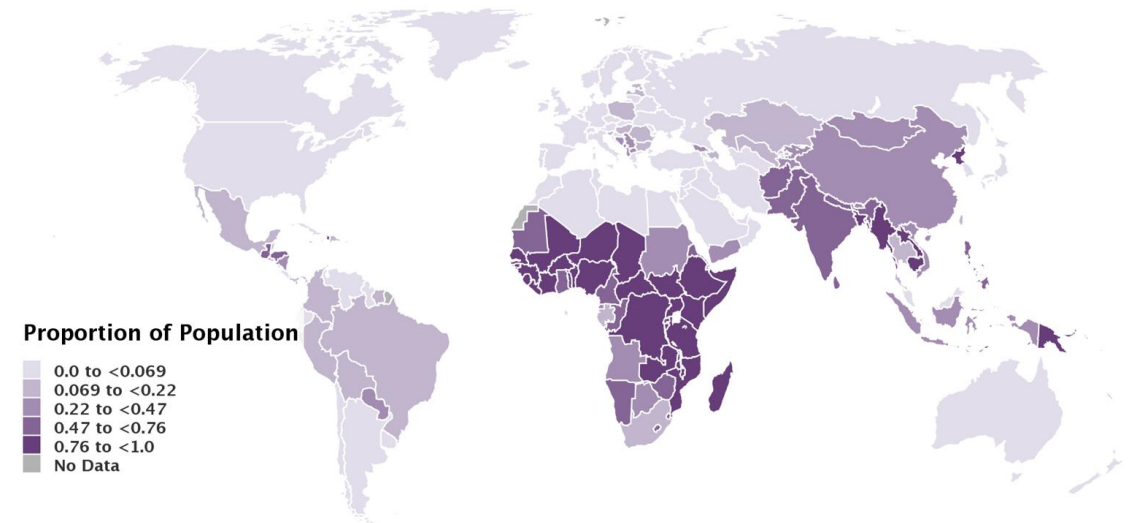
➡️ They do not tell us which of those **mechanisms actually work** and which don't

Indoor Air Pollution poses a serious threat to respiratory health around the world

- **Roughly one third of the world's households** burn biomass fuels, as well as coal, for energy.

- **Women and young children are disproportionately affected** by the resulting indoor air pollution

- **Possible solution:** cleaner cook stoves
  - cleaner fuels
  - chimneys to remove fumes

But is using these stoves really healthier?
Do they decrease fuel usage and GHG emissions?
Are people willing to use them?

Proportion of Population Using Solid Fuels in 2019

Proportion of Population

- 0.0 to <0.069
- 0.069 to <0.22
- 0.22 to <0.47
- 0.47 to <0.76
- 0.76 to <1.0
- No Data

Source: State of Global Air 2020

Study investigates the effect of improved cooking stoves on household behaviour and health

Randomisation
A public lottery randomly assigned the order in which stoves were constructed within 44 villages for 2,575 households across Orissa.

Installing "chulas"
The stoves were distributed in three waves with the first wave receiving stoves in 2006, the second in 2009, and the third at the end of the evaluation.

Tracking households
Households were followed for four years after initial stove offers, allowing for an examination of long-run use and impacts of stoves.

Duflo, Greenstone, Hanna, 2016

Results show that improved cooking stoves did not improve health outcomes because of low uptake

- Households did not use the stoves regularly or appropriately

- They did not make the necessary investments to maintain them properly

- Importance to **understand how new technologies fit people's lives**

  - Cooking stoves could not be moved outside which might explain why people rejected them

- Good example where out theory was wrong

Duflo, Greenstone, Hanna, 2016

Women using a cookstove ("chula") in India.

Photo: Thomas Chupein | J-PAL

Randomised evaluations ensure effective resource allocation

There may be different effective solutions - considering cost-effectiveness

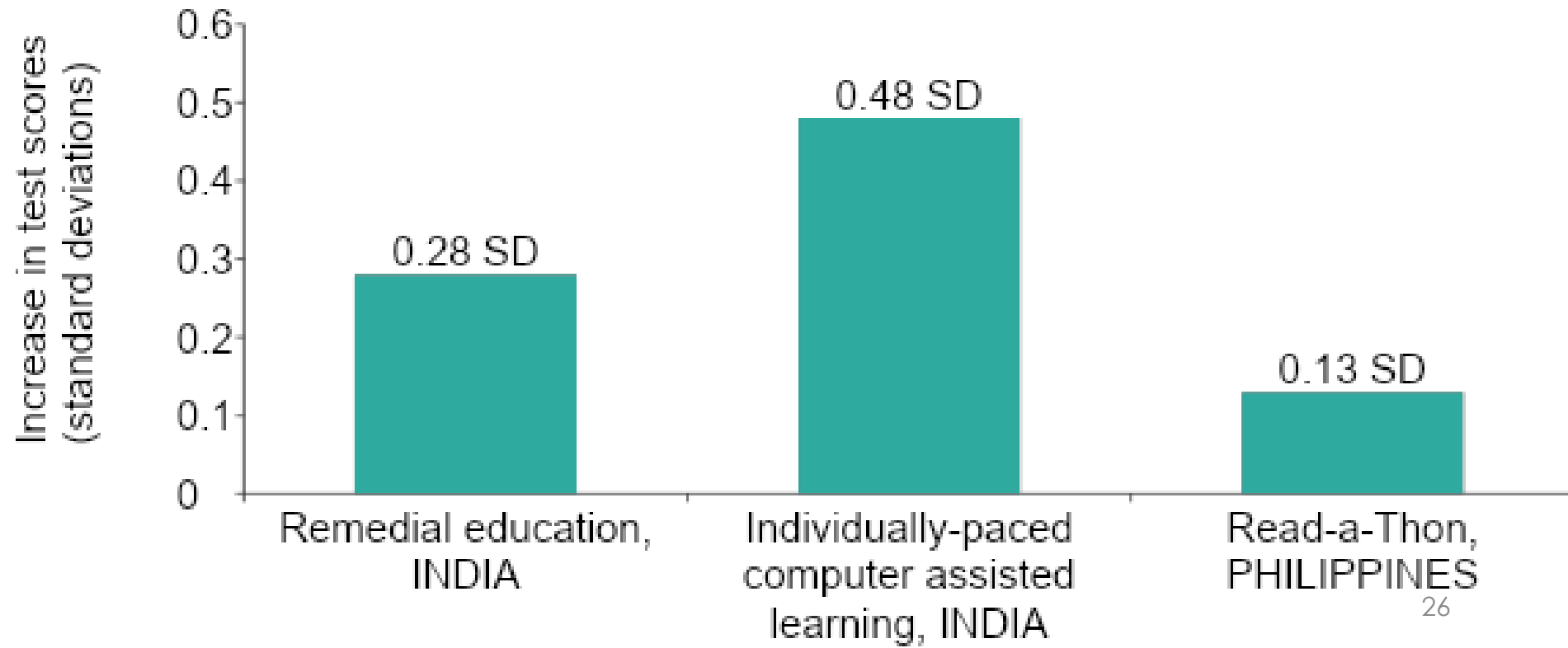**Consider three different promising approaches:**

- Supporting low-achieving students through a remedial education programme

- Providing additional inputs through computer-assisted learning

- Focusing on a specific skill through a read-a-thon programme to improve reading habits and skills
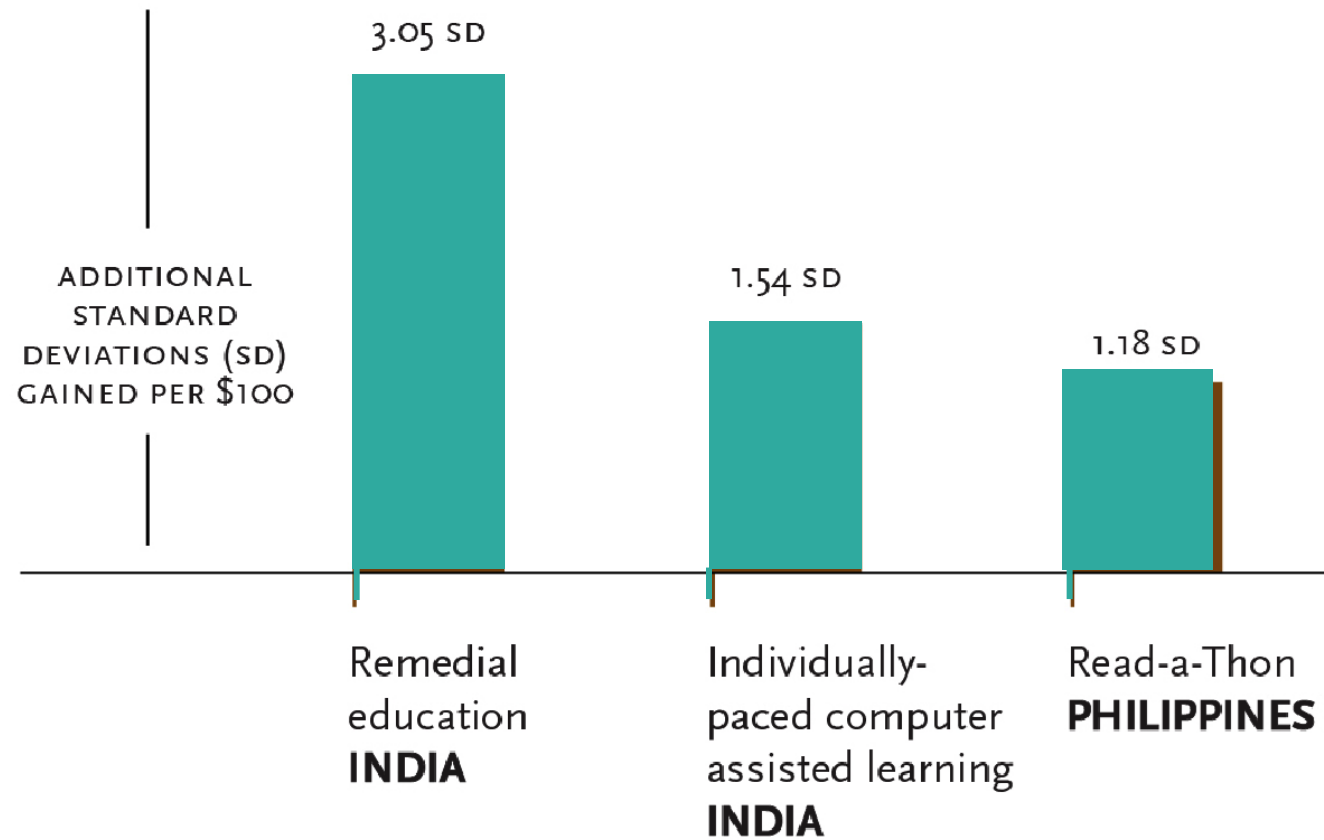

Photo: J-PAL

Comparing multiple education programmes

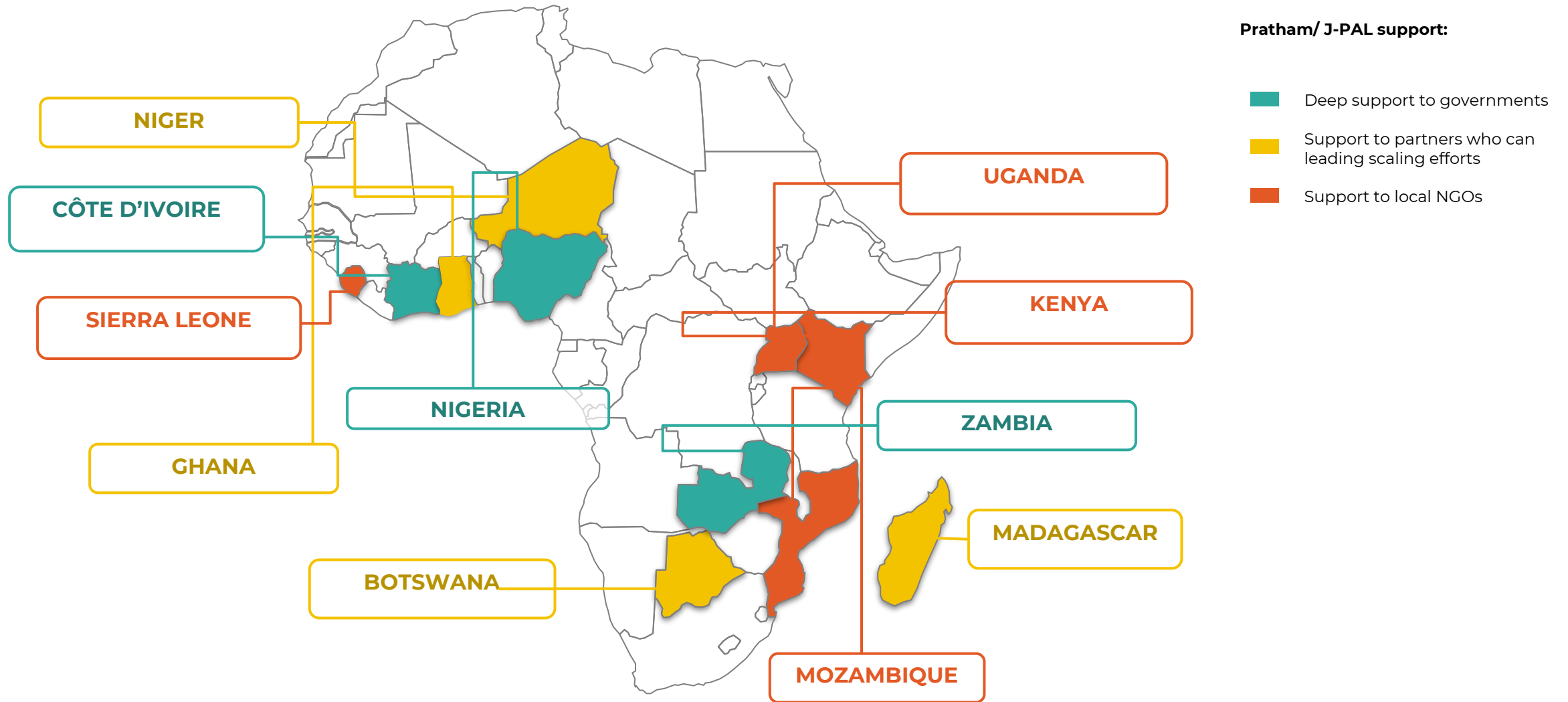Impact of education programs on student test scores

Incorporating costs

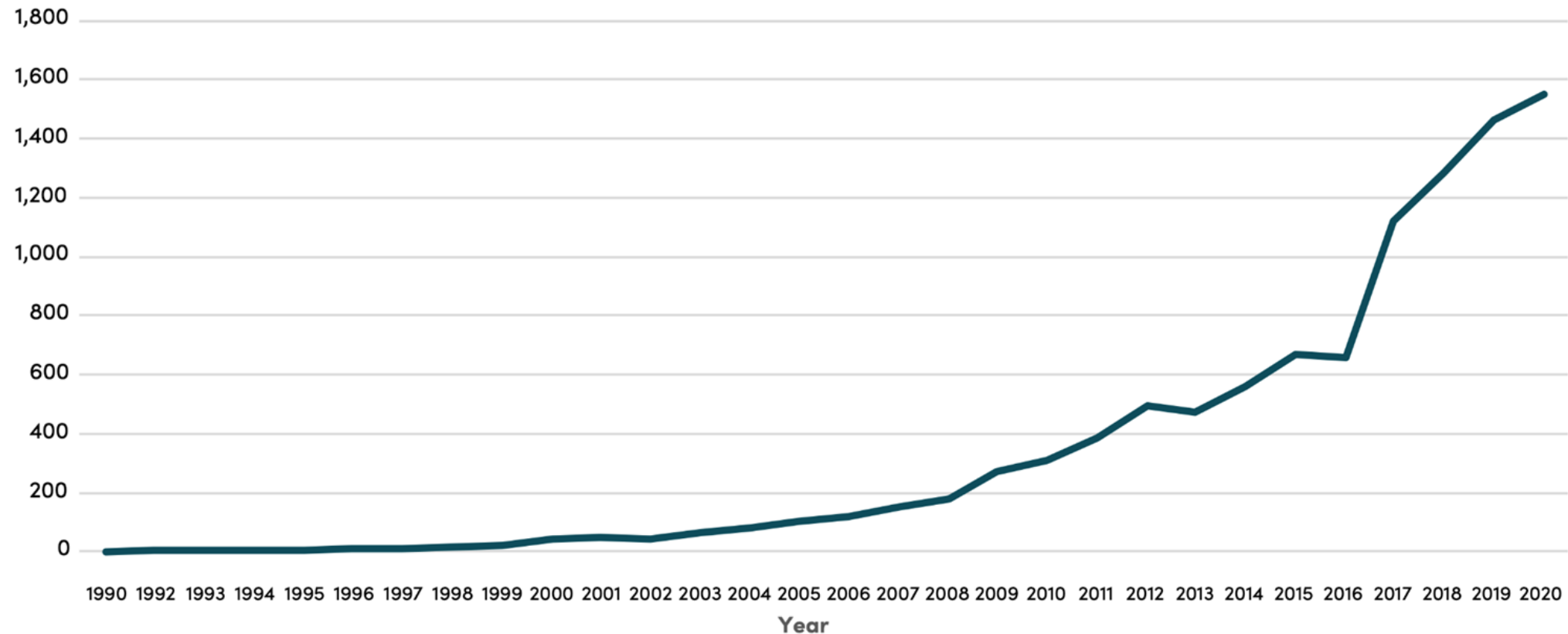### Cost-effectiveness of education programmes to improve student test scores

27

Scaling TaRL to 3 million children across Africa

Pratham/ J-PAL support:

- Deep support to governments
- Support to partners who can leading scaling efforts
- Support to local NGOs

NIGER

CÔTE D'IVOIRE

SIERRA LEONE

NIGERIA

GHANA

BOTSWANA

UGANDA

KENYA

ZAMBIA

MADAGASCAR

MOZAMBIQUE

Randomised evaluations can lead to transformative change

The last two decades have witnessed a large increase in the production of rigorous evaluations
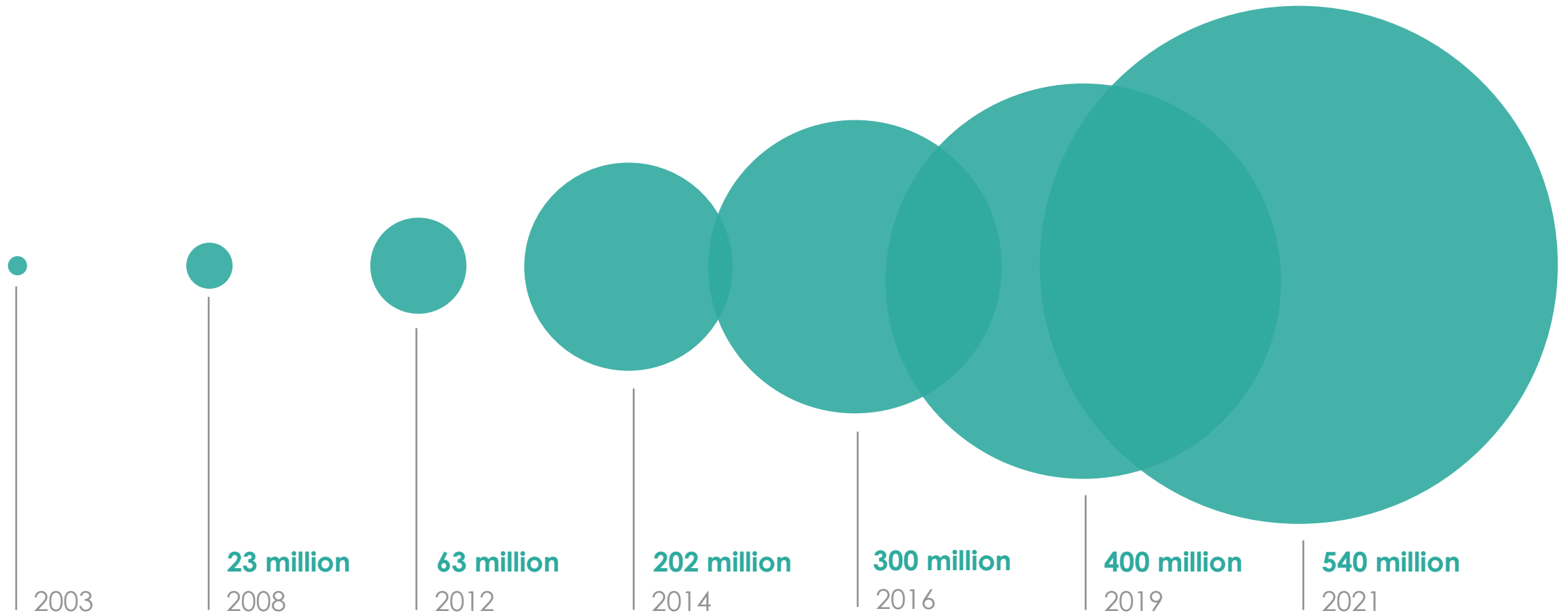

FIGURE 1 Number of published impact evaluations from 1990 to 2020

Source: Data from 3ie Development Evidence Portal (data as of March 2022), image from CGDev report "Breakthrough to Policy Use."

... and significant progress in the use of rigorous evaluations for public policy decisions
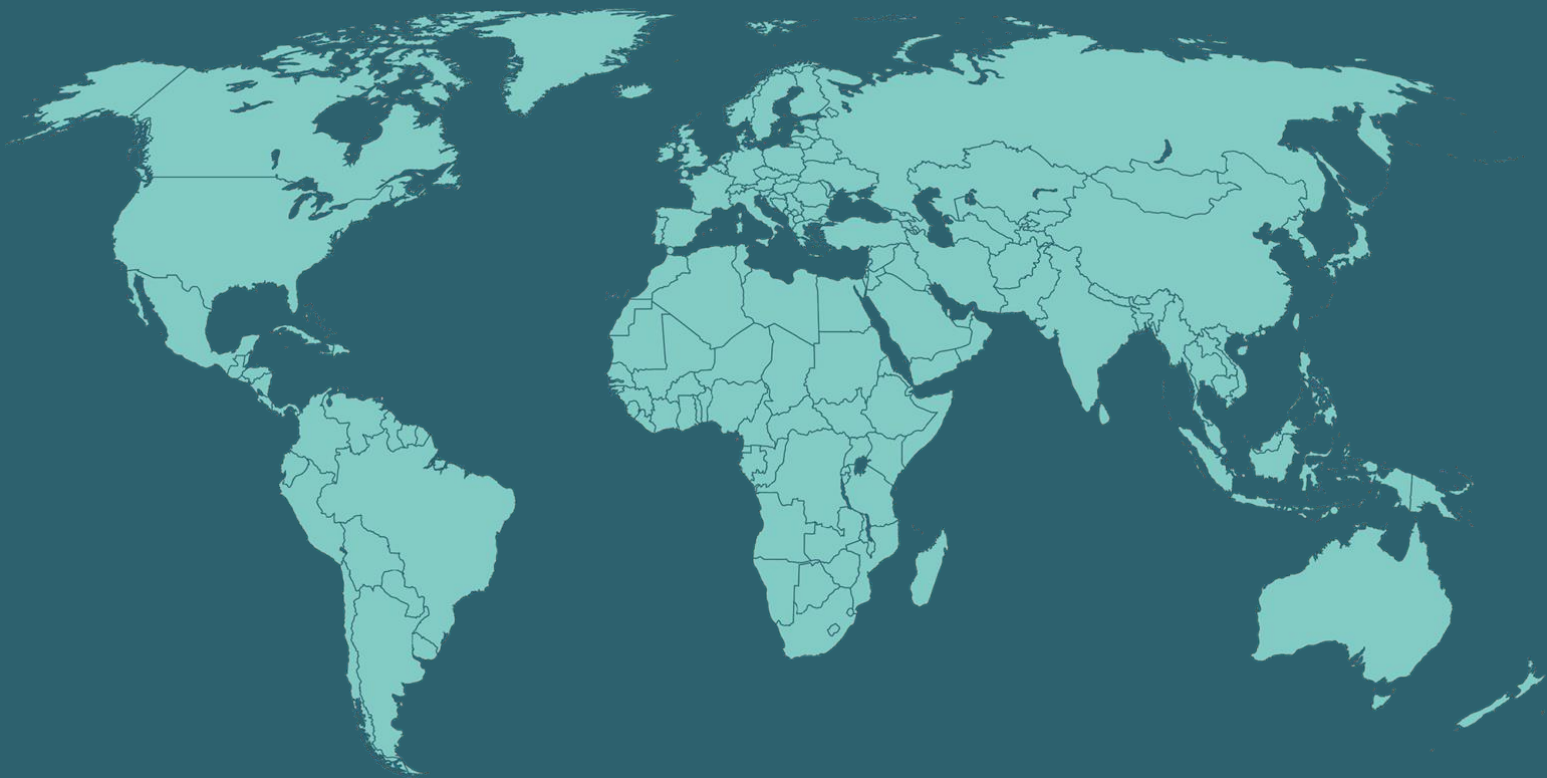
People reached by programs and policies informed by J-PAL evaluations

| | **23 million** | **63 million** | **202 million** | **300 million** | **400 million** | **540 million** |
| 2003 | 2008 | 2012 | 2014 | 2016 | 2019 | 2021 |

20 YEARS OF
EVIDENCE + IMPACT

600M

LIVES REACHED

Multilateral Financial Institutions play a criticl role in all of this

Consumers of evidence

- Effective lending requires knowing what works
- As an institutional lender MDBs have a stake in the success of the interventions that goes beyond financial considerations
  - Non-financial outcomes are often harder to measure and don't get reported on automatically?
- You provide advise and help to countries to choose effective policies
  - Participate in their design process
  - Bring ideas in from outside
  - Provide a check on the internal pressures
- You lead global conversations about the exigencies of the future

Producers of Evidence

- You have been at the forefront of the movement towards more reliable evidence in policy making
- By funding others to generate the evidence
  - The World Bank does it through its SIEF window for example
- By generating the evidence themselves
  - The Inter-American Development Bank organized some of the most influential early RCTs
    - On Progresa and other conditional cash transfers
  - The World Bank does it through its DIME window for example
- By using the research funds to strategically influence research in areas where there are policy needs but not much evidence

The NDB

- Can represent the South's perspectives on what the most urgent needs are

- Can show the way in evidence-based policy making

- Can shape the multi-lateral agenda by highlighting where the MDBs have often been lacking

Conclusion

Rigorous evaluation is a tool that can help us not only understand the success or failure of a programme…

Much as been achieved but especially as the new crises arrive, much more needs to be done

Its often said that this will be the century where knowledge capital will dominate. Lets make it happen, by generating the knowledge that really serves the world.

YEARS OF J-PAL
WORLDWIDE

J-PAL
ABDUL LATIF JAMEEL POVERTY ACTION LAB