



---

# Impact Evaluation for Safety Nets

---

## Methods and Applications

**Patrick Premand**  
**(Africa HD)**

**Safety Nets Course, December 2013**

# Outline

From M&E to Impact Evaluation

The Main Concepts of Impact Evaluation

Choosing the best IE design for your project

Randomization as an operational tool

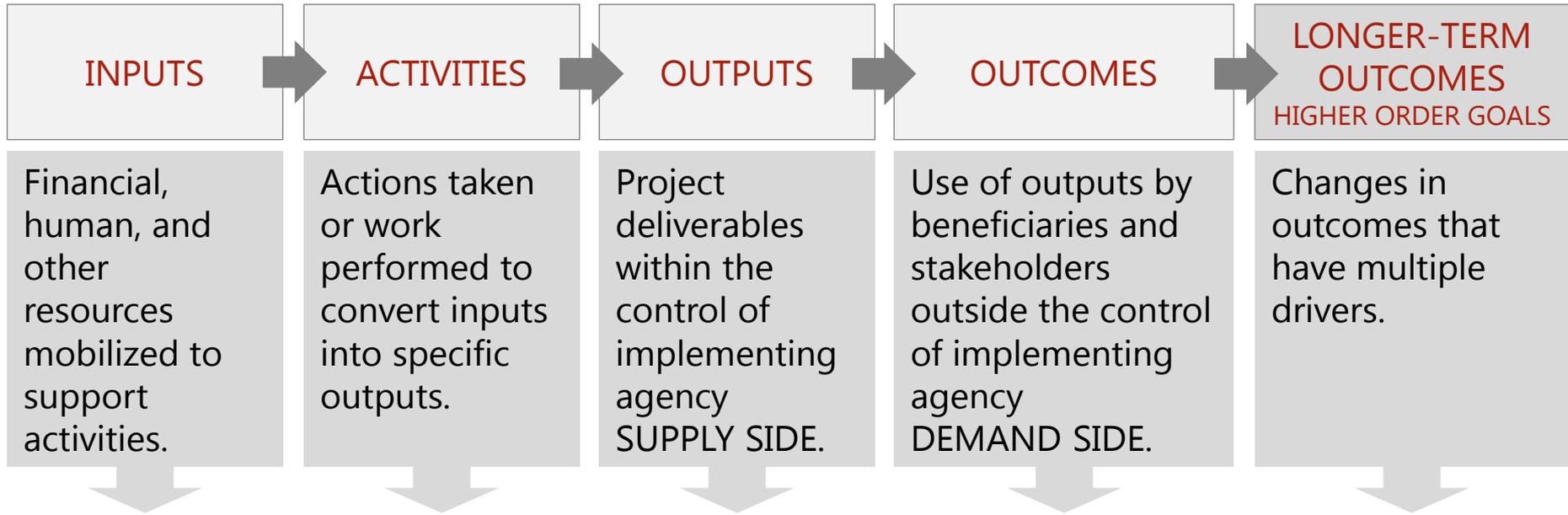
Illustration of IE results

Operational aspects



# From M&E to Impact Evaluation

# The Results Chain in a Typical Program



- Budget
- Staffing

- Training
- Studies
- Construction

- Training plan completed
- Cash transfer delivered
- Road constructed
- School built

- New practices adopted
- Use of the road
- School attendance up
- Health service use up

- Poverty reduced
- Income inequality reduced
- Labor productivity increased

Focus of traditional M&E

Focus of Impact Evaluation

Results-based management

# Monitoring vs. Evaluation

	<b>Monitoring</b>	<b>Evaluation</b>
Frequency	Regular, Continuous	Periodic
Coverage	All programs	Selected program, aspects
Data	Universal	Sample based
Depth of Information	Tracks implementation, looks at WHAT	Tailored, often to performance and impact/ WHY
Cost	Cost spread out	Can be high
Utility	Continuous program improvement, management	Major program decisions

# Evaluations



A systematic, objective assessment of an on-going or completed project, program, or policy, its design, implementation and/or results, asking

- **Descriptive Questions** to seek to determine what is taking place and describe aspect of a process.
- **Normative Questions** to compare what is taking place to what should be taking place. (**PROCESS EVALUATION**)
- **Cause-and-Effect Questions** to examine outcomes and assess what difference the intervention makes in outcomes (**IMPACT EVALUATION**)

# Impact Evaluation is not for every project

- Evaluate impact selectively, when project is:
  - Innovative
  - Replicable, scalable, or implemented at scale
  - Strategically relevant (e.g. large budget)
  - Evaluation will fill knowledge gap
  - Substantial policy impact
- Impact Evaluation can focus on selective innovations within projects
  - Beyond 'does my program work'?
  - Towards 'which design is more effective?'

# Public Works (THIMO) in Cote d'Ivoire

- Emergency Youth Employment and Skills Development Project (US \$45 million)
  - Public Works Component and Skills Development Component (apprenticeships, internships, professional training, entrepreneurship training,...)
- Public Works
  - 12,500 youths (18-30) by 2015
  - Daily wage rate of CFA 2,500 for 6 months
- Graduation elements in public works:
  - Entrepreneurship training to help youth enter into self-employment
  - Sensitization on wage employment opportunities to help youth transition into wage jobs
  - Also: payment into bank accounts, basic life skills training

# Key questions for the impact evaluation of Public Works in Cote d'Ivoire?

## Basic Question

What is the impact of participation in the public works program on employment and earnings of youths and their households?

## Design Question (1)

Does the provision of basic entrepreneurship training facilitate the creation of household enterprises after graduation?

## Design Question (2)

Does the provision of sensitization on wage employment opportunities facilitate insertion into wage jobs after graduation?

# The Niger Safety Nets Project

- Niger safety nets project (US \$70 million)
  - Building block for a national social protection system
  - Cash transfer and cash for work components
- [Video](#) and [Story](#)
- Large-Scale cash transfer program:
  - 80,000 households by 2017
  - Unconditional cash transfer, US\$20/month for 24 months (PMT targeting)
- Accompanying measures as “soft conditionalities”
  - 18-month parenting training (community assemblies, meetings, household visits) covering holistic early childhood development
  - Aims to trigger ‘behavioral changes’
  - Similar approach in several new projects in Africa

# Key questions for the impact evaluation of cash transfers in Niger?

## Basic Question

Does the cash transfers program reduce poverty, improve food security and improve children's nutrition?

## Design Question

Do the accompanying measures (parenting training) generate higher impacts on children's nutrition and development?



# **The Main Concepts of Impact Evaluation**

# Impact Evaluation needs to be distinguished from other “evaluations”



The objective of impact evaluation is to estimate the **causal** effect or **impact** of a program on outcomes of interest.

---

# The Objective



**Estimate the causal effect (impact) of intervention (P) on outcome (Y).**

*(P) = Program or Treatment*

*(Y) = Outcome Indicator, Measure of Success*

---

**Example:** What is the effect of a cash transfer program (P) on Household Consumption (Y)?

# Solution

Estimate what **would** have happened to outcomes ( $Y$ ) in the absence of the program ( $P$ ).

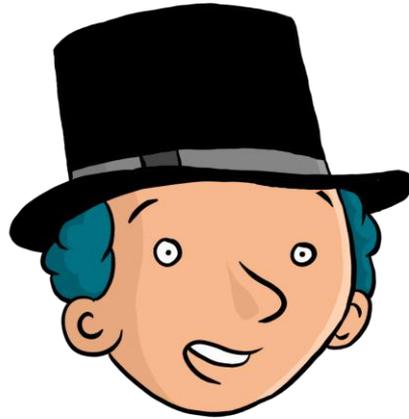
We call this the **Counterfactual**.

The key to a good impact evaluation is a valid estimate of the **counterfactual!**

Counterfactual is estimated based on **comparison groups**

# Communicating complex concepts in 3 slides... Example: What is the Impact of...

giving Ali



additional money



(P)

on Ali's consumption

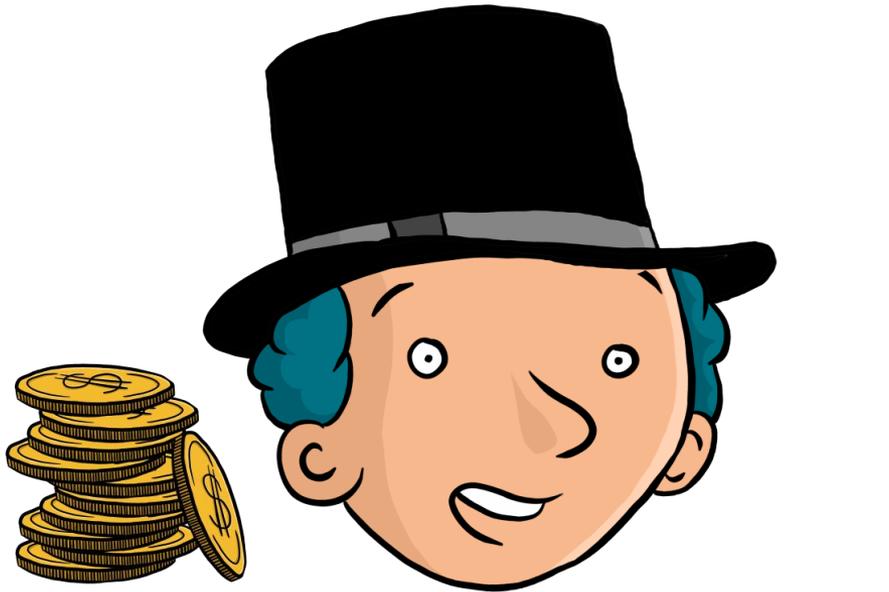


(Y)?

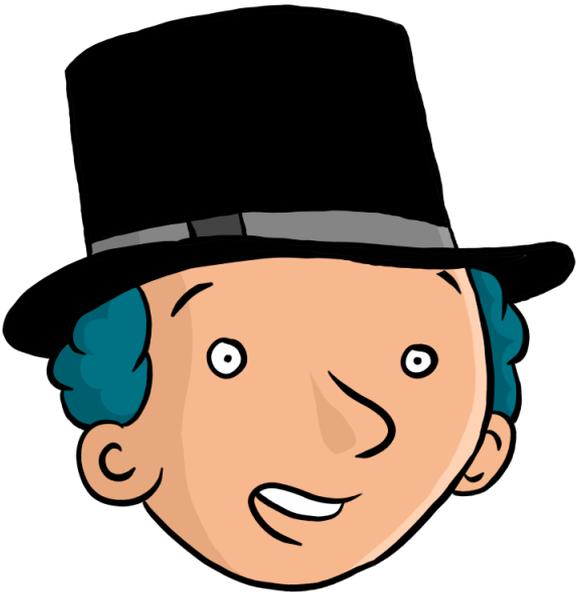
# The Perfect Clone

Ali

Ali's Clone



6 candies

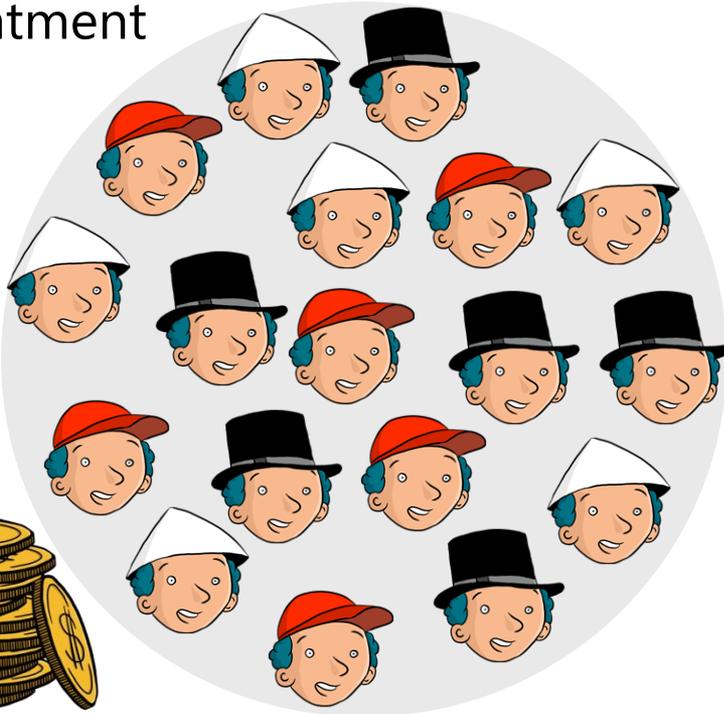


4 candies

$IMPACT = 6 - 4 = 2$  Candies

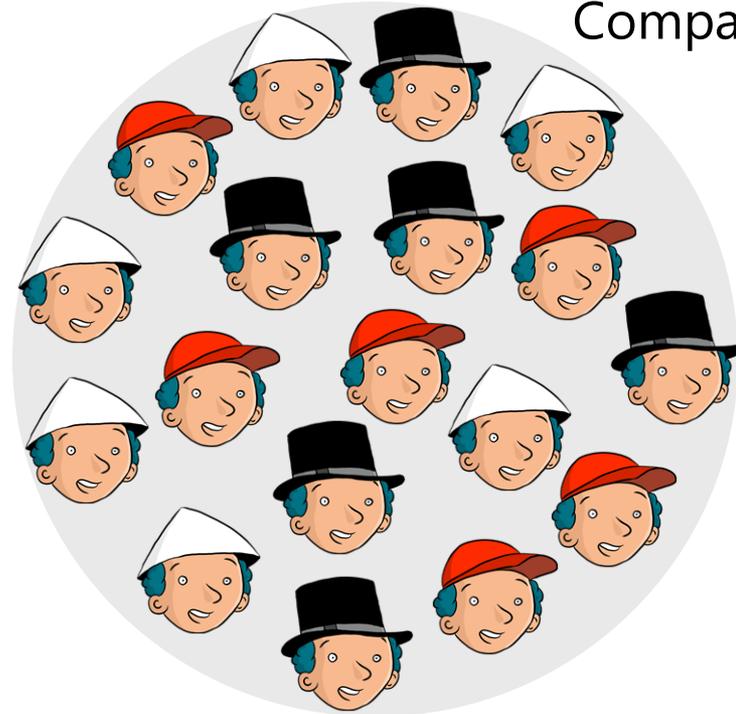
# In reality, use statistics

Treatment



Average  $Y=6$  candies

Comparison



Average  $Y=4$  Candies

$IMPACT=6-4=2$  Candies



# **Choosing the best IE design for your project**

# Finding good comparison groups

We want to find **clones** for the Alis in our programs.

The treatment and comparison groups should

- have identical characteristics

benefiting from the intervention.

With a good comparison group, the **only reason** for different outcomes between treatments and controls is the **intervention (P)**

program eligibility & assignment  
to get valid estimates of the  
counterfactuals

# Two false counterfactuals to avoid



Before  
vs After

**Compare:** Same individuals  
**Before and After** they  
receive **P**.

**Problem:** Other things may  
have happened over time.

Enrolled  
vs Not Enrolled

**Compare:** Group of  
individuals **Enrolled** in a  
program with group that  
**chooses not to enroll**.

**Problem:** Selection Bias.  
We don't know why they  
are not enrolled.

**Both counterfactuals lead  
to biased estimates of the  
counterfactual and the  
impact.**

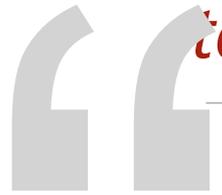
# The conversation needs to start early



**Retrospective Evaluation** is necessary when we have to work with a program that has already been roll-out and existing data.

*Rarely feasible: baseline data? Information on targeting?*

---



In **Prospective Evaluation**, the evaluation is designed in parallel with the program (and targeting decisions).

*The way to go: ensure baseline data is collected, and comparison group exists.*

# Where do good Comparison Groups come from?

- The rules of program operation determine the evaluation strategy.
- We can almost always find a valid comparison group if:
  - the operational rules for selecting beneficiaries are equitable, transparent and accountable;
  - the evaluation is designed prospectively.
- Evaluation design and program design go hand-in-hand.

# 5 methods in IE Toolbox

- 5 methods in IE toolbox take different approaches to generate comparison groups and estimate the counterfactual:

1 Randomized Assignment

2 Randomized Promotion

3 Regression Discontinuity Design

**RDD**

4 Difference-in-Differences

**DD**

5 Matching

# Choosing an IE design for your program

- Design IE prospectively to generate good comparison groups and collect baseline data
- 3 operational questions to determine which method is appropriate for a given program

**Money:** Does the program have sufficient resources to achieve scale and reach full coverage of all eligible beneficiaries?

**Targeting Rules:** Who is eligible for program benefits? Is the program targeted based on an eligibility cut-off or is it available to everyone?

**Timing:** How are potential beneficiaries enrolled in the program – all at once or in phases over time?

# Choosing your IE method(s)

Money →	<i>Excess demand</i>		<i>No Excess demand</i>	
Targeting →	<i>Targeted</i>	<i>Universal</i>	<i>Targeted</i>	<i>Universal</i>
Timing ↓				
<b><i>Phased Roll-out</i></b>	1 Randomized assignment 4 RDD	1 Randomized assignment 2 Randomized promotion 3 DD with 5 Matching	1 Randomized Assignment 4 RDD	1 Randomized assignment to phases 2 Randomized Promotion to early take-up 3 DD with 5 matching
<b><i>Immediate Roll-out</i></b>	1 Randomized Assignment 4 RDD	1 Randomized Assignment 2 Randomized Promotion 3 DD with 5 Matching	4 RDD	If less than full Take-up: 2 Randomized Promotion 3 DD with 5 Matching

# Choosing the IE method in Niger

Money →	<u>Excess demand</u>		<b>No Excess demand</b>	
Targeting →	<u>Targeted</u>	<b>Universal</b>	<b>Targeted</b>	<b>Universal</b>
Timing ↓				
<b><u>Phased Roll-out</u></b>	1 Randomized assignment 4 RDD	1 Randomized assignment 2 Randomized promotion 3 DD with 5 Matching	1 Randomized Assignment 4 RDD	1 Randomized assignment to phases 2 Randomized Promotion to early take-up 3 DD with 5 matching
<b><u>Immediate Roll-out</u></b>	1 Randomized Assignment 4 RDD	1 Randomized Assignment 2 Randomized Promotion 3 DD with 5 Matching	4 RDD	If less than full Take-up: 2 Randomized Promotion 3 DD with 5 Matching



# **Randomization as an operational tool**

# Randomization is not only for the Impact Evaluation. In Niger...

- Geographical targeting only works with higher administrative units:
  - Regions, departments, communes...
- How to choose between long list of criteria to select villages within units?
- Public lottery was deemed the most transparent, and least controversial approach.
- Project decided to keep using the randomization approach for purely operational reasons, including when not needed for the IE.

# Randomization can help with transparency

**Project staff in Niger:** “Now political authorities cannot interfere with the village selection. All the village chiefs were present and signed that they agreed with the procedure before we did the selection. No one can complain to us and try to change the result.”



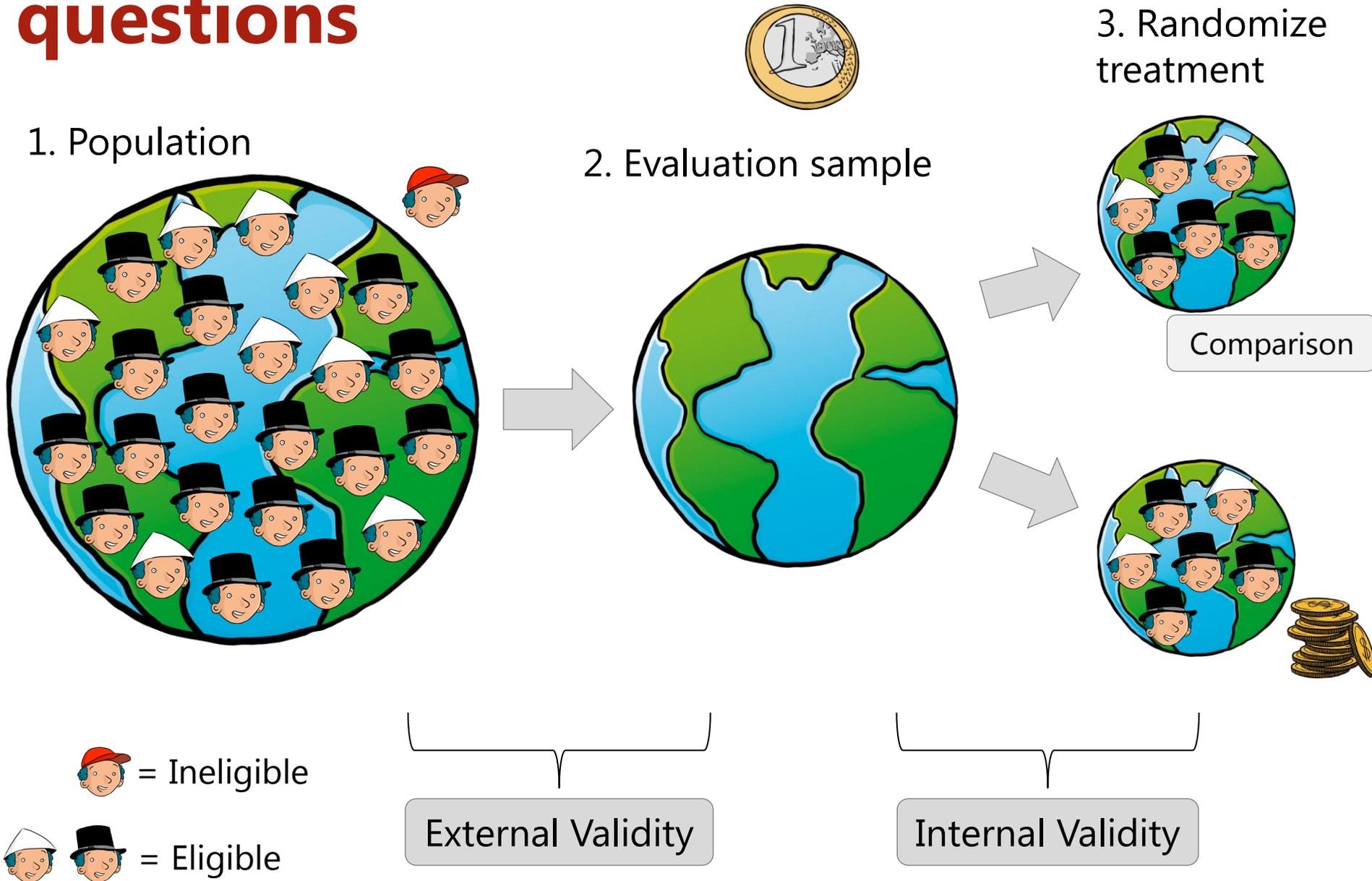
**Beneficiaries in Nicaragua:** “At least this time we know why we were not chosen for the program. Usually decisions are made and we don’t know why our village cannot participate.”



# Randomization

- Randomization can be used only in certain contexts.
- BUT excess demand happens for most programs
  - Even after applying all existing targeting criteria, not everyone can be served
  - Randomization is fair, transparent and ethical way to assign benefits to equally deserving populations
  - Provides equal chance of participation among equally deserving units.
- Randomization is the Gold Standard
  - Most robust method
  - But also the most simple, and the cheapest.
- Multiple ways to perform randomization

# Randomization to answer basic IE questions

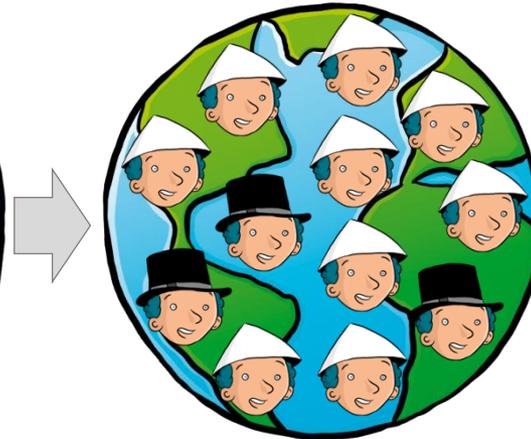


# Randomized to answer IE “design” questions

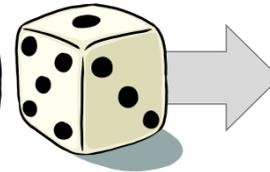
1. Population



2. Evaluation Sample



3. Randomize treatment.



Comparison



 = Not eligible

  = Eligible

# Randomized Assignment



In **Randomized Assignment**, large enough samples, produces 2 statistically equivalent groups.

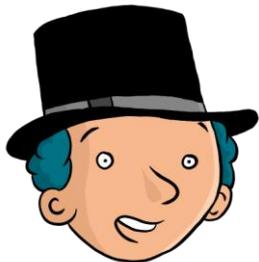
We have identified the perfect **clone**.

Feasible for prospective evaluations with over-subscription/excess demand.

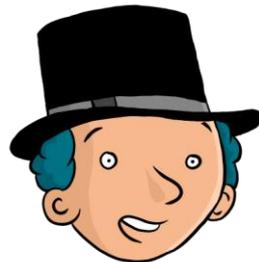
Most pilots and new programs fall into this category.

Consider evaluating relative effectiveness of alternative program design options.

Randomized beneficiary



Randomized comparison





# Illustration of IE results

# Productive safety net pilot in Nicaragua

- Key question:
  - Can safety nets provides reduce poverty and protect households from shocks in the short-term
  - While also help household invest and manage risk in the medium term?
- Combine a basic CCT with 2 productive interventions
- Where: 6 municipalities in rural Nicaragua
  - High poverty , dependent on agriculture
- 3 groups of households:

**GROUP 1**

**Basic CCT**

**GROUP 2**

**Basic CCT +  
Productive investment grant**

**GROUP 3**

**Basic CCT +  
Vocational training**

# Impact Evaluation Design

1. **Public Lottery within selected municipalities,** public lottery to randomly select
  - 50 Control communities
  - 56 Treatment communities
2. **Within each treatment communities,** public lottery to assign households to 3 packages



1000

**Basic CCT**

1000

**Basic CCT +  
Productive investment grant**

1000

**Basic CCT +  
Vocational training**

## CCT + Business Grant



## CCT + Training

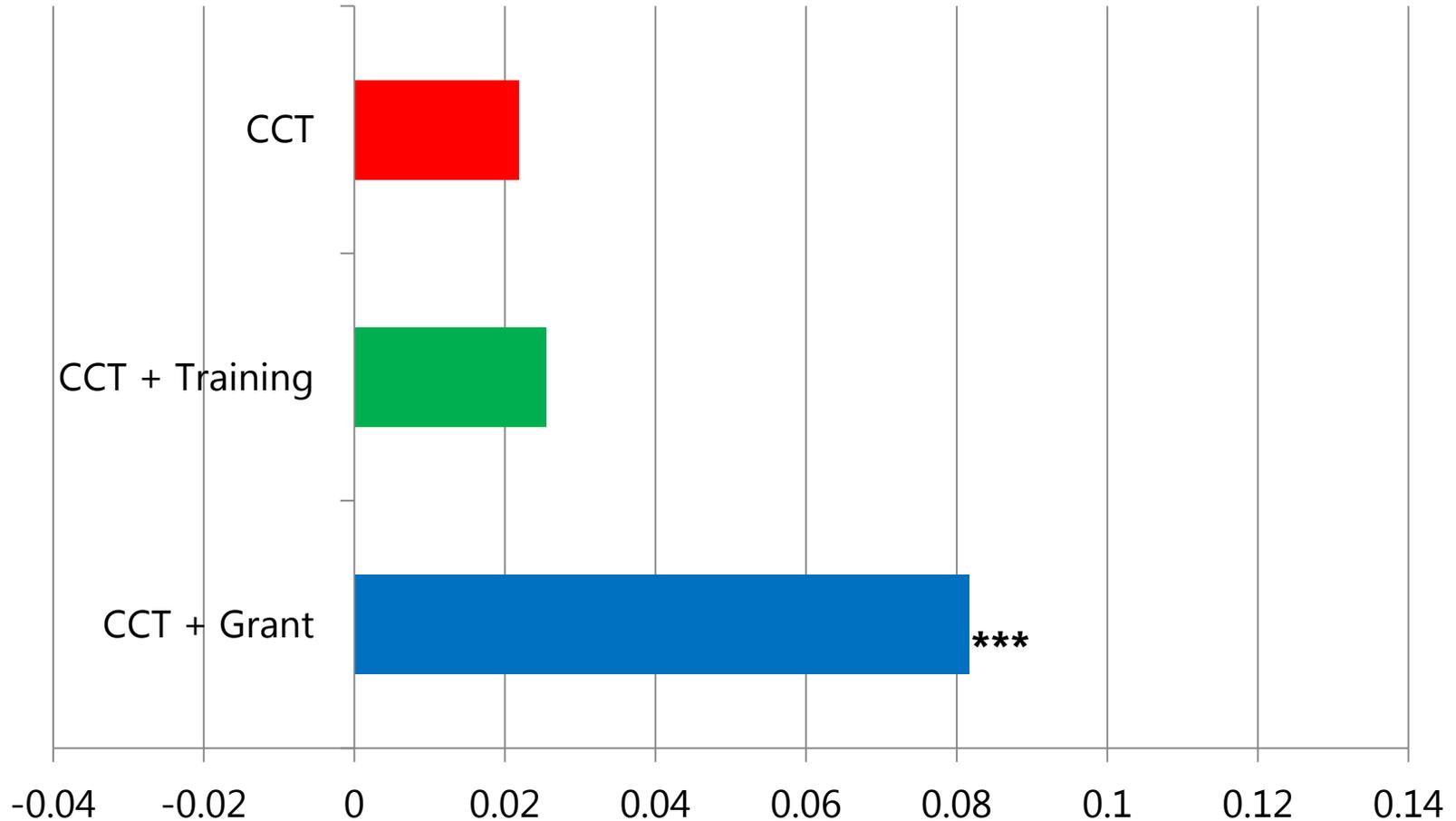


- The program lasted for one year
- What were the results **two years** after the end of the intervention?
- (There was a range of short-term impacts, more info:

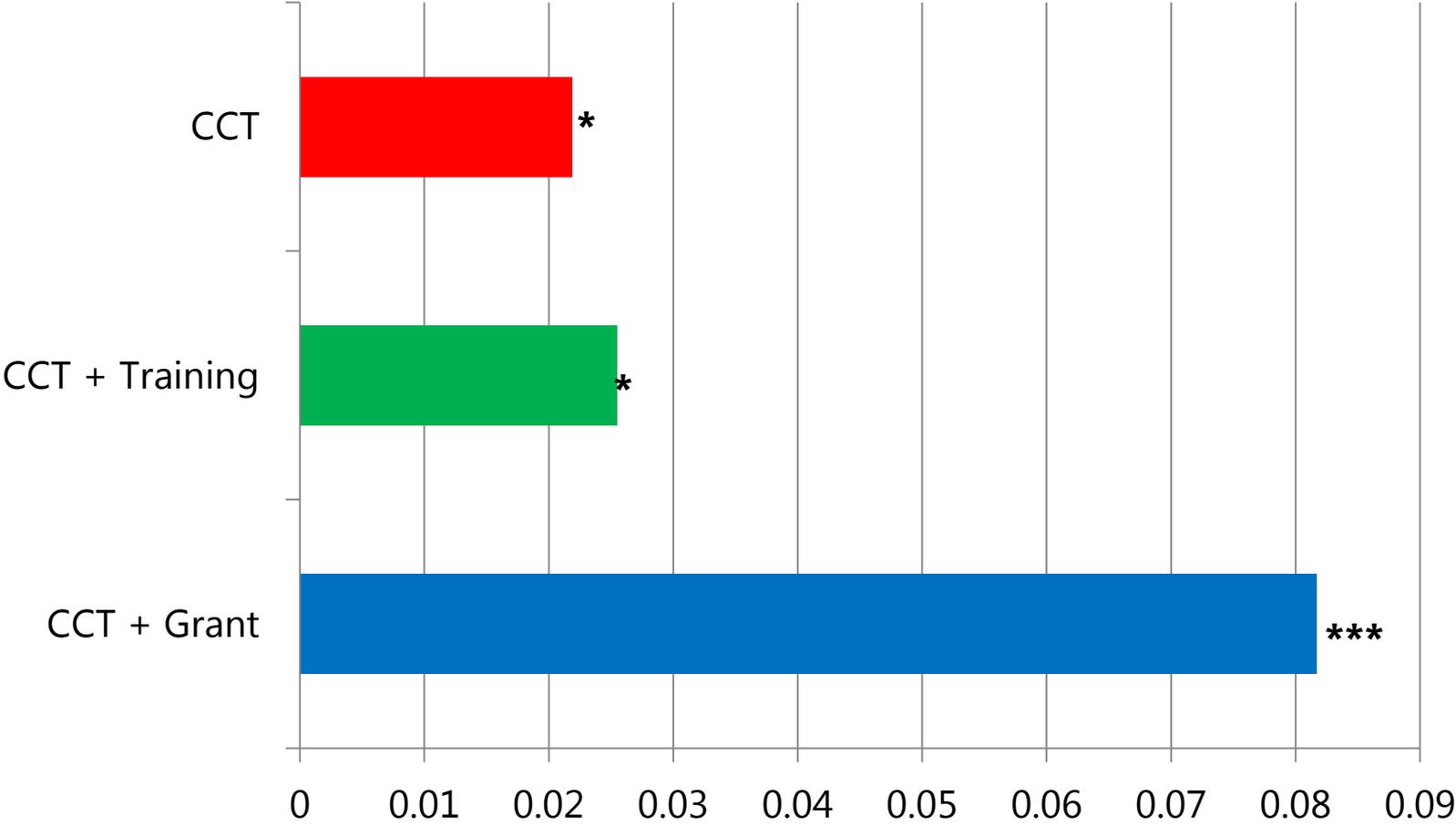


[www.worldbank.org/atencionacrisisevaluacion](http://www.worldbank.org/atencionacrisisevaluacion)

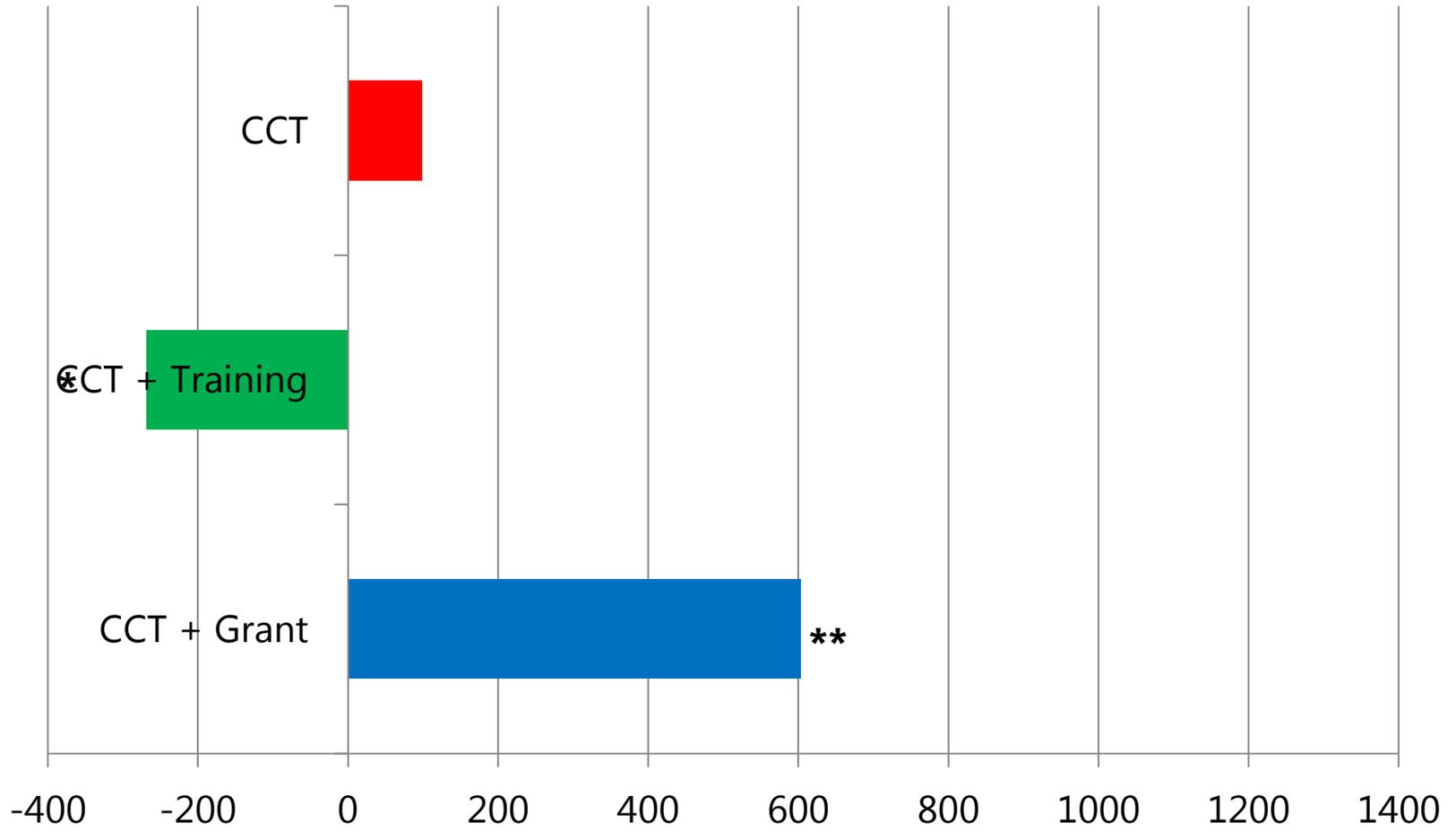
# The CCT + grant had a lasting impact on welfare two years after the end of the program...



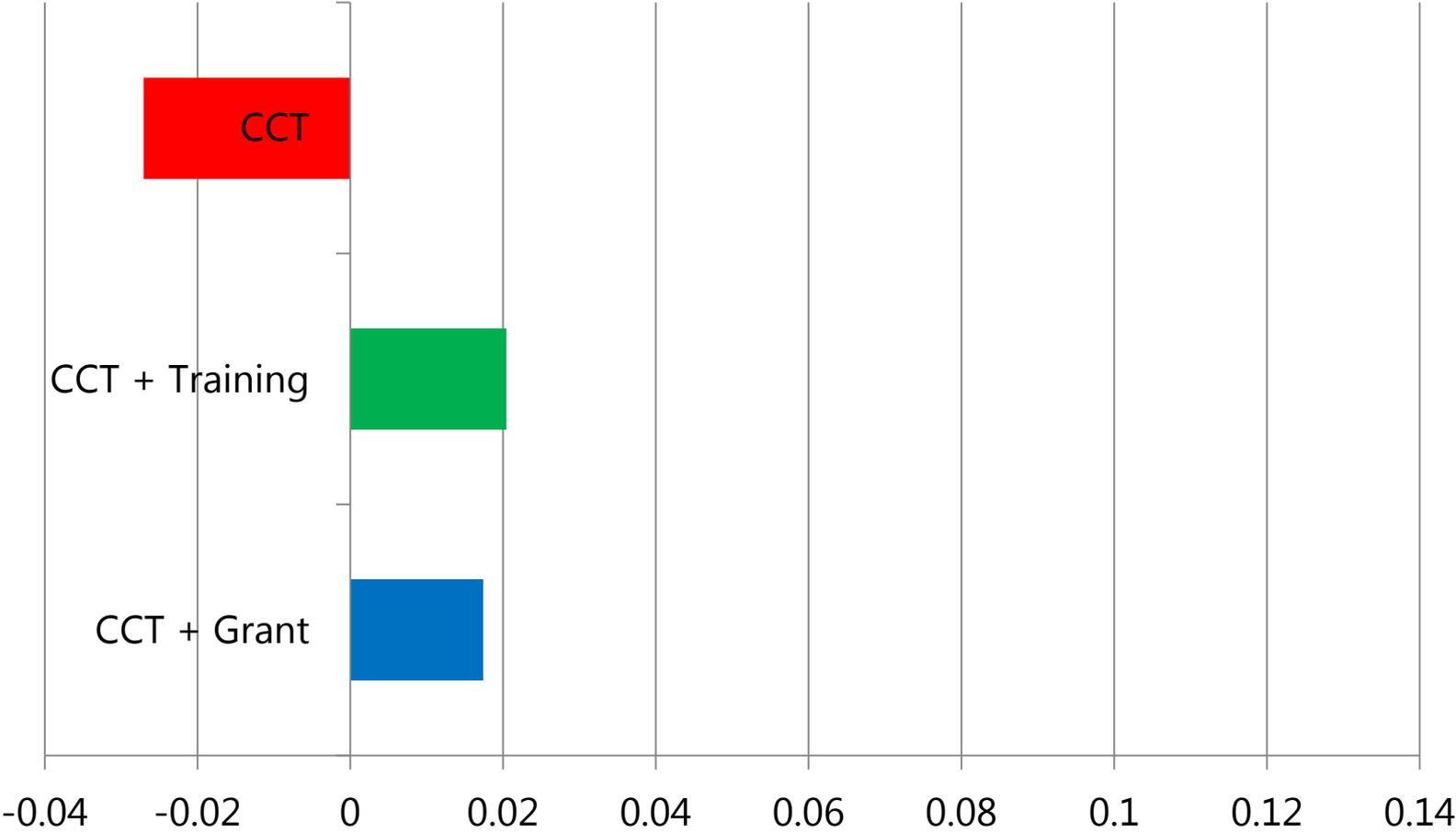
# The CCT + grant had the largest impact on entry into non-agricultural self-employment



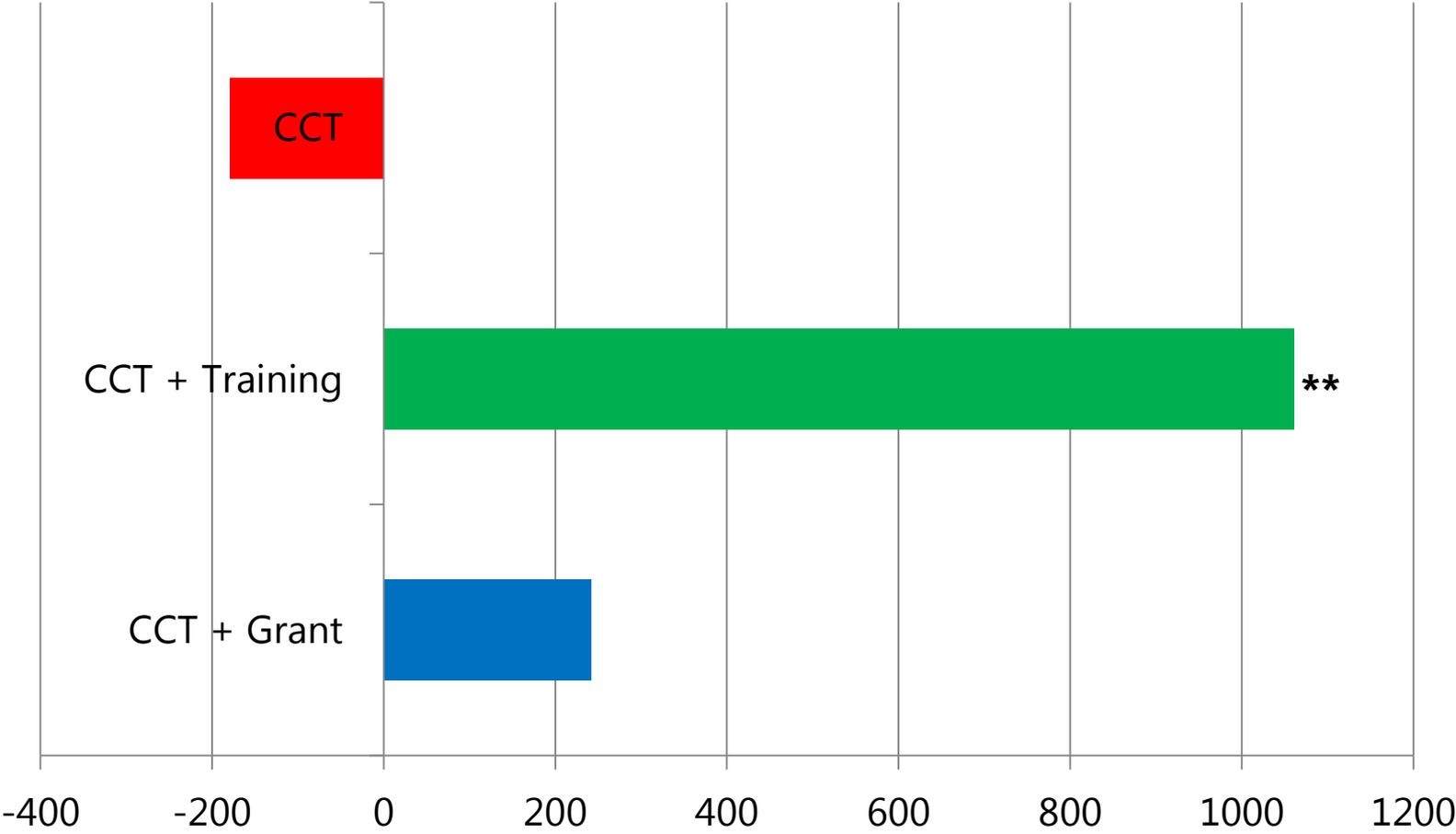
# ... and the largest impact on profits in non-agricultural businesses



# There was no significant impact on entry into non-agricultural wage employment



# But the CCT + training increased wages for those in private sector jobs





# Operational aspects

# What to Evaluate?



**Efficacy Studies** are carried out in a specific setting to test a “model” implemented in best-possible way.

(e.g. **Pilots** for proof of concept)

---



**Effectiveness Studies**, provide evidence from interventions taking place under normal circumstances

(e.g. **Scalable National Programs**)

# Who does the Impact Evaluation?

- Critical to start discussing IE early:
  - Clarify role of different types of “evaluations”
  - Large potential value-added of IE... but it is an investment
- Essential to design the evaluation with operational team
  - Framing of evaluation question
  - Program design and IE design go together.
  - Implementing of IE requires close coordination with project implementation
- IE best as seen of collaboration between implementers and evaluators
  - Quality/Validity of design is what makes results legitimate
  - Consider which components to outsource

# Design and implementation of IE in Niger?

- Broad outline of Impact Evaluation in project design document
- After effectiveness, 6-months preparation phase (typically longer)
  - Multiple workshops with local stakeholders to finalize design
  - Participation of project team in a regional impact evaluation workshop
- Randomization performed by project (through geographical targeting process)
- Baseline data collection contracted to National Statistical Agency (June 2012)
  - Oversight by WB and project teams.
- Next steps:
  - Focus on ensuring MIS colleges comprehensive data on project implementation
  - Qualitative (process) evaluation to gather information on quality of implementation, as well as beneficiaries' perceptions
  - Follow-up data collection (December 2014)

# Benefits of Impact Evaluation

- IE is the only way to know if a program is effective.
  - Clear value-added, but needs to be used selectively.
- Doing IEs can improve project implementation
  - The process of doing IE can change the way we work (“Science of Delivery”)
  - Learning through the process in Niger:
    - Learning quality control in baseline helped with PMT survey
    - Randomization helped with transparency
    - Solid baseline helps for profiling of beneficiaries
    - Allows analysis of targeting efficiency
    - Complementary qualitative evaluation helps improve project implementation

# Impact Evaluation, MIS, and Process evaluation are complementary

- Monitoring and Process Evaluation data is key for the evaluation
  - MIS: Lists of beneficiaries, distribution of benefits
  - Targeting data
  - Process evaluation
- Common issues:
  - For the IE design, can you do targeting (e.g. PMT) in comparison group?
  - How to set up unique identifiers across all sources of data
- Data on project implementation is essential to interpret IE result. Does the project work/not work:
  - because of the **intervention model**,
  - or because of **implementation of the intervention model**?

# Remember



The objective of impact evaluation is to estimate the **causal** effect or **impact** of a program on outcomes of interest.

---

# Remember



To estimate impact, we need to estimate the **counterfactual**.

- what would have happened in the absence of the program and
- use comparison or control groups.

# Remember



We have a **toolbox** with **5 methods** to identify good comparison groups.

---

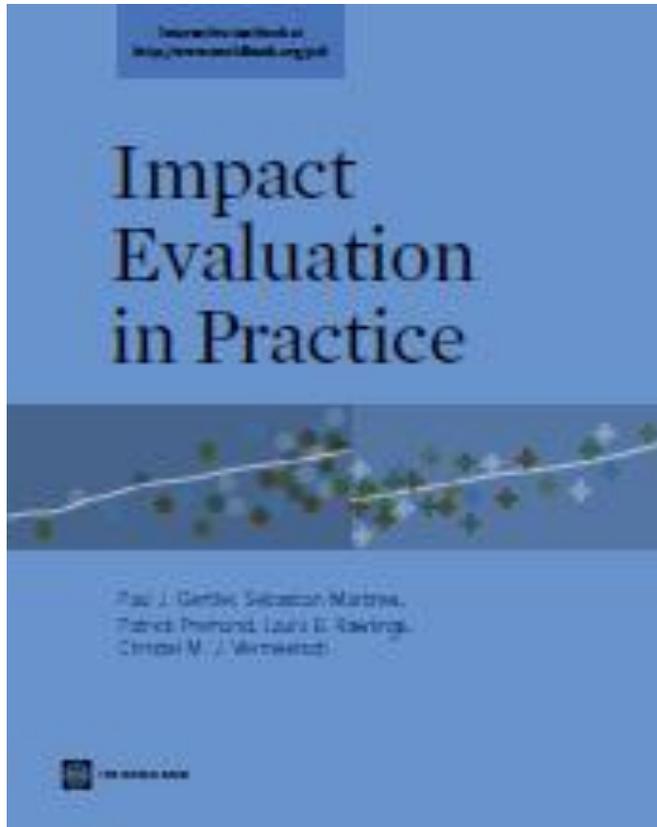
# Remember



**Choose the best evaluation method** that is feasible in the program's operational context.

---

# Thank you!



Reference also:  
available in **Spanish**  
**French** and  
**Portuguese** (soon)

[www.worldbank.org/ieinpractice](http://www.worldbank.org/ieinpractice)

