# Impact Evaluation for Safety Nets

## Patrick Premand, DEC/DIME and SPJ GP

## Safety Net Core Course, November 2019

# Outline

Impact Evaluation Questions

The Main Concepts of Impact Evaluation

Choosing the best IE design for your project
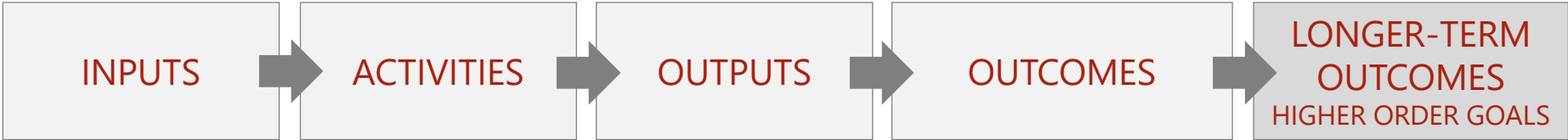
Cases (including on IE and Delivery systems)

Summary of Key Messages

# The questions impact evaluations can help answer

# The Results Chain in a Typical Program

| INPUTS | → | ACTIVITIES | → | OUTPUTS | → | OUTCOMES | → | LONGER-TERM OUTCOMES HIGHER ORDER GOALS |
|---|---|---|---|---|---|---|---|---|

| Financial, human, and other resources mobilized to support activities. | Actions taken or work performed to convert inputs into specific outputs. | Project deliverables within the control of implementing agency SUPPLY SIDE. | Use of outputs by beneficiaries and stakeholders outside the control of implementing agency DEMAND SIDE. | Changes in outcomes that have multiple drivers. |
|---|---|---|---|---|

| Focus of traditional M&E | Focus of Impact Evaluation |
|---|---|

**Results-based management**

# Evaluations

"

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

o Descriptive Questions to seek to determine what is taking place and describe aspect of a process.

o Normative Questions to compare what is taking place to what should be taking place. (PROCESS EVALUATION)

o Cause-and-Effect Questions to examine outcomes and assess what difference the intervention makes in outcomes (IMPACT EVALUATION)
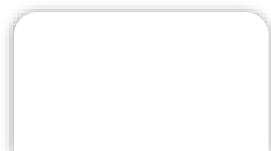
# Impact Evaluation is not for every project

- 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?'

# Using Impact Evaluation to inform investment decision...

**Objective**: Improve quality of jobs and earnings in a sustainable way

**Proposed program**: Give grants of $400 to youths, in a single transfer, without conditions

**Will this program be effective?**

Will it improve employment opportunities and earnings for youths?

Will youths invest in micro-enterprises or use them for consumption or other leisure expenditures?

# What would you decide?

**Would you decide to invest in this employment program providing cash grants to youths?**

1. You decide TO INVEST
2. You decide NOT TO INVEST

An [impact evaluation in Uganda](#) has shown that this program was cost-effective

# Using Impact Evaluation to optimize program benefits ...

**Objective**: Improve girls' access to secondary education

**Proposed program**: Scholarship for girls

**What is the optimal amount of the scholarship?**
Yearly scholarship of $45 or yearly scholarship of $60?

# What would you decide?

**How would you invest the budget for scholarships?**

1. Give 225,000 girls a transfer of $60?
2. Give 300,000 girls a transfer of $45?

An [Impact Evaluation in Cambodia](#) has shown that scholarships of $45 were just as effective as scholarships of $60.

# Using Impact Evaluation to chose delivery modality…

**Objectif**: Accelerate the transition of unemployed individuals back to work

**Proposed program**: Job counselling program

**What is the most effective delivery mechanism?**

- **Public standard** (monthly, cost €120)

- **Public intensive counselling** (weekly 6 months, € 657)

- **Private intensive counselling** (weekly 6 months, cost €900)

# What would you decide?

Which program modality would you choose?

1. **Public standard** (monthly, cost €120)

2. **Public intensive counselling** (weekly 6 months, € 657)

3. **Private intensive counselling** (weekly 6 months, cost €900)

An [Impact Evaluation in France](#) found that both intensive models were more effective than the standard public model, but that intensive public delivery was more cost-effective.

# Using impact evaluation to optimize program design

**Objectif**: Improve employment and earnings for youths in urban area in Cote d'Ivoire

**Proposed program** (PEJEDEC): Labor-intensive public works (LIPW), with monthly salary at minimu wage ($60) for 6 months, 30 hours of work per week.

**Proposed innovation/graduation component**

- Entrepreneurship training to help entry into self-employment
- Job search training to help youth transition into wage jobs
- (Also: payment into bank accounts, basic life skills training)

# Impact Evaluation Questions: Public Works in Cote d'Ivoire

## Basic Question

What is the impact of the public works program on youth employment and earnings?

## Design Question (1)

Does the entrepreneurship training facilitate more productive self-employment after exit from the program?

## Design Question (2)

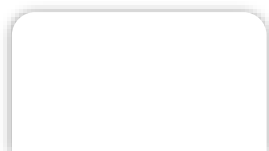Does the job search training facilitate insertion into wage jobs after exit from the program?

Results: here (French) or here (English)

# 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.
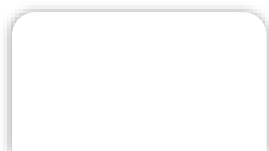
# The Objective

**"**

**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

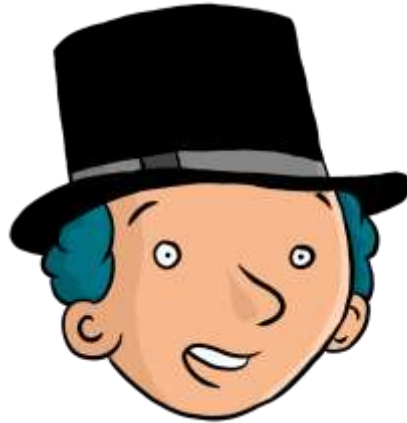Estimate what ***would*** have happened to outcomes (*Y)* in the absence of the program (*P)*.
We call this the **Counterfactual.**

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

Counterfactual is estimated based on **comparison groups**

# Example: **What is the Impact of...**
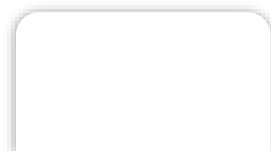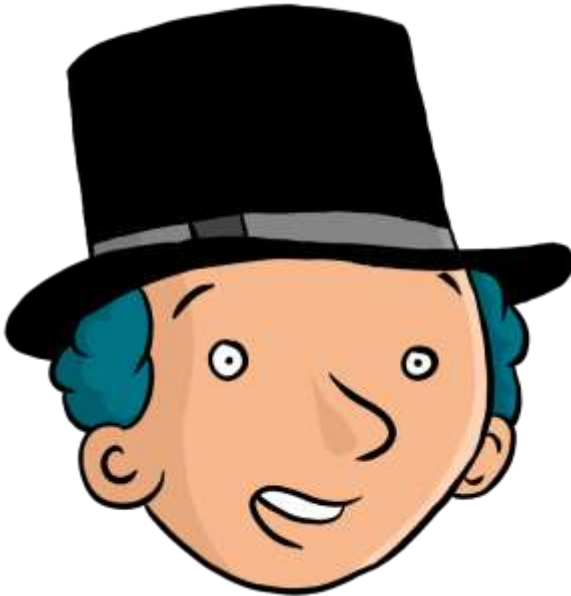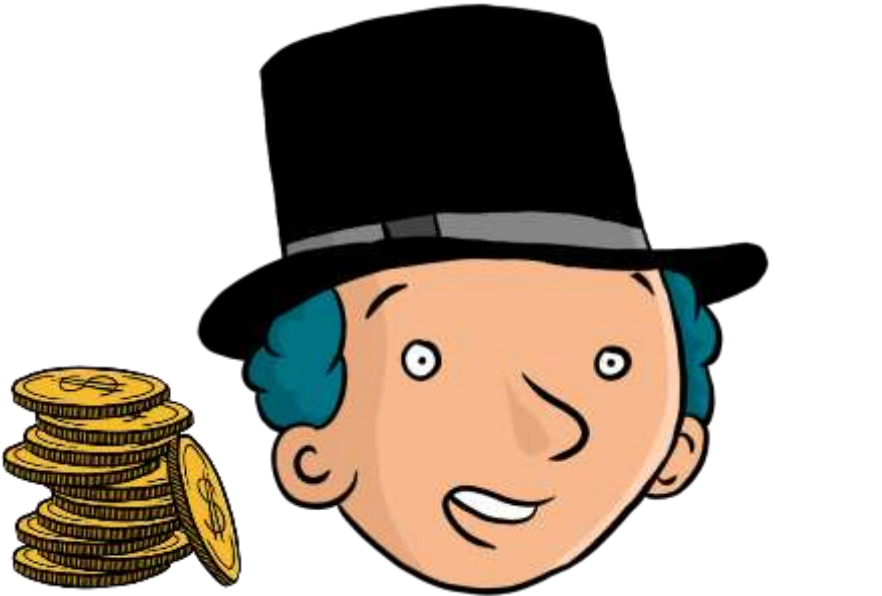
giving Ugo

additional money (P)

on Ugo's consumption (Y)?
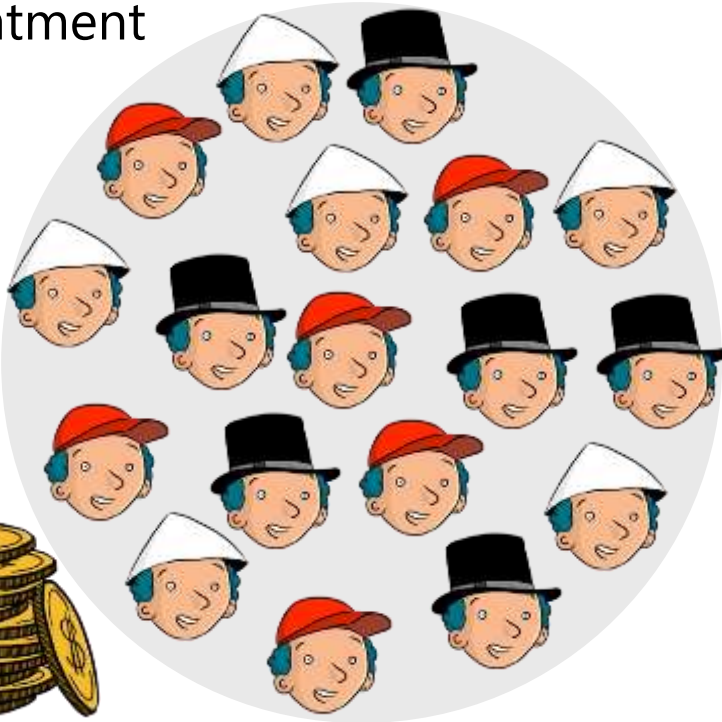
# The Perfect *Clone*

Ugo

Ugo's Clone

6 candies

4 candies

IMPACT=6-4=2 Candies

# In reality, use statistics

Treatment

Comparison

Average Y=6 candies
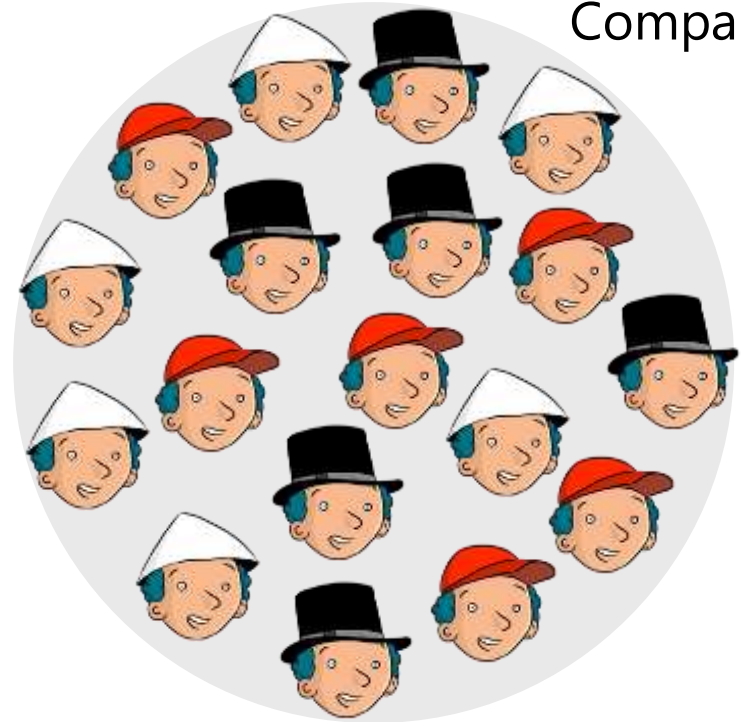
Average Y=4 Candies

IMPACT=6-4=2 Candies

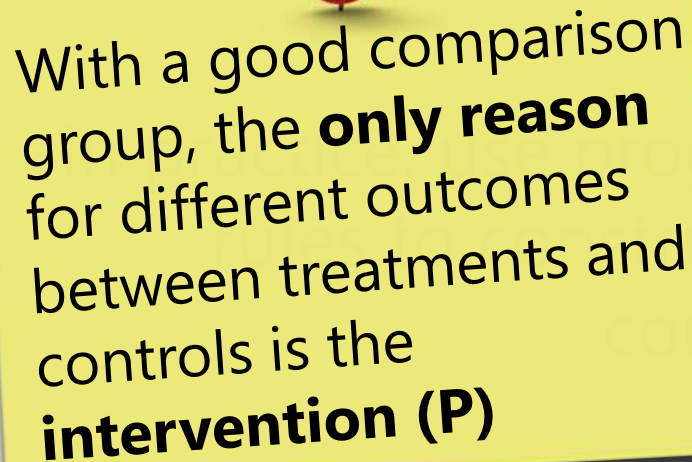# Choosing the best IE design for your project

# Finding good comparison groups

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

The treatment and comparison groups should
- have identical characteristics
- except for benefiting from the intervention.

...gram eligibility & assignment ...ct valid estimates of the counterfactuals

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

# 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?*

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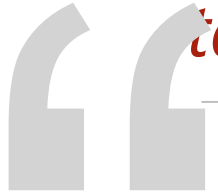
In **Prospective Evaluation**, the evaluation is designed in parallel with the program (and targeting decisions).

*The way to go: ensure baseline data is collected, and comparison group exists.*
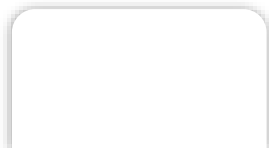
# Where do good Comparison Groups come from?

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

# Choosing the IE method in Niger

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

# Randomization as an operational tool

# Randomization is not only an evaluation tool. In Côte d'Ivoire…

- Public works were introduced as an instrument right after a period of conflict.

- How to chose between the large number of youth interested to participate?

- Public lottery was deemed the most transparent, and least controversial approach.

- Project decided to use randomization/lottery for purely operational reasons.

- The impact evaluation came later and just built on this.
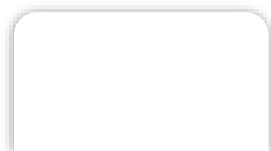
# Randomization can help with transparency



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. Noone 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."
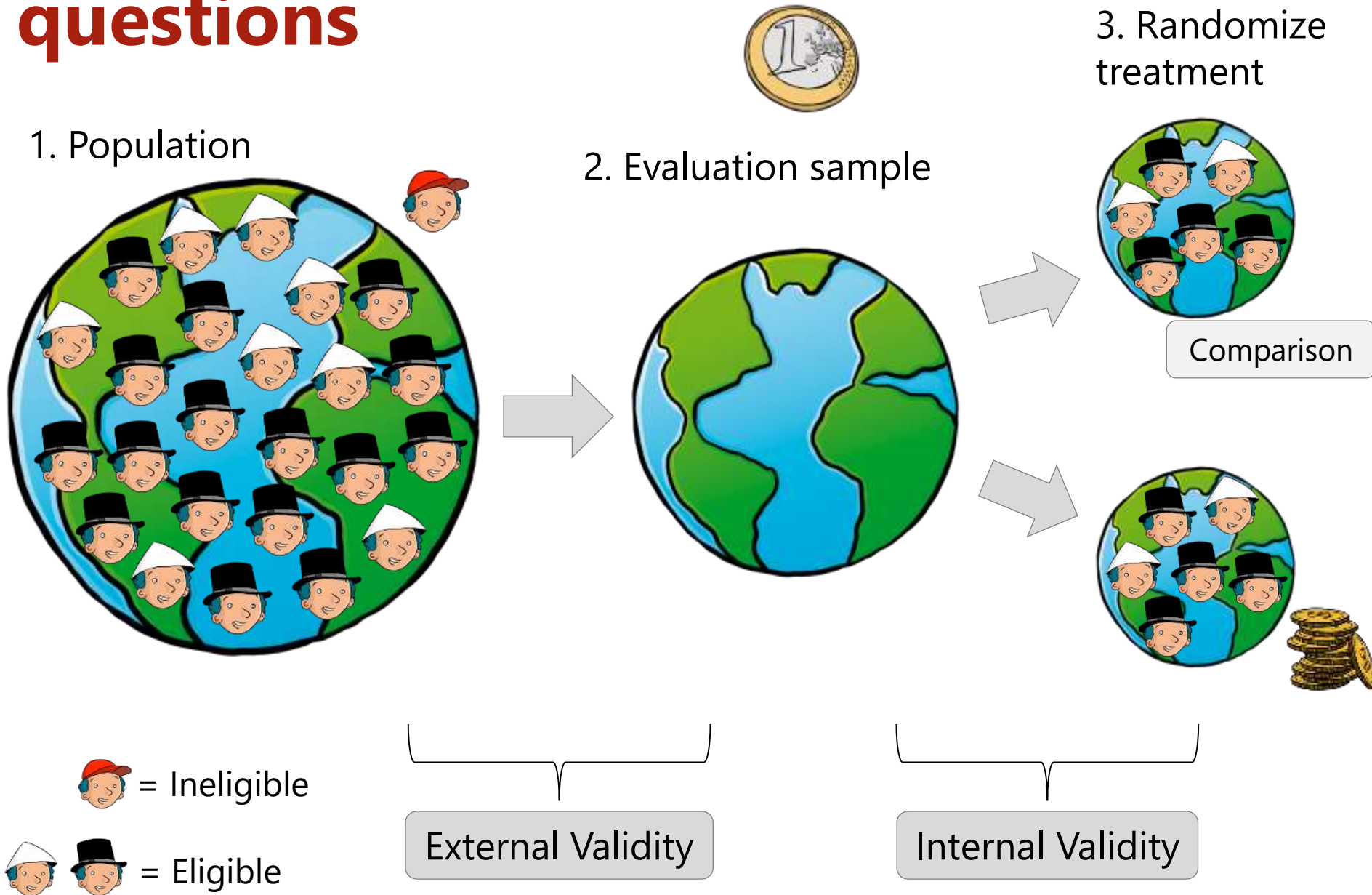
# Randomization

- 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 <u>equally deserving</u> populations
  - Provides <u>equal chance</u> of participation among <u>equally deserving</u> units.
- Randomization is often the preferred method for impact evaluation
  - Most robust method
  - But also the most simple, and the cheapest.
- Multiple ways to perform randomization
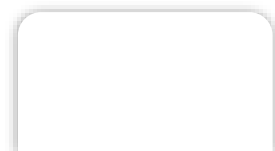
# Randomization to answer basic IE questions

1. Population

2. Evaluation sample

3. Randomize treatment

Comparison

= Ineligible

= Eligible

External Validity
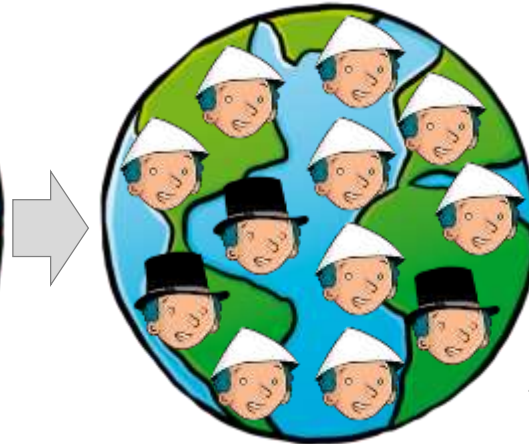
Internal Validity

# Randomized to alternative program modalities



1. Population
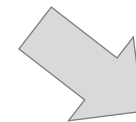
2. Evaluation Sample

3. Randomize treatment.

Comparison

Public Works

PW+ Entrepreneurship training

PW+ Jobs Search training

= Not eligible

= Eligible

# Randomized Assignment

**!**

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

We have identified the perfect **clone**.

Randomized beneficiary

Randomized comparison



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.

# Cote d'Ivoire case results

# In the short-term…
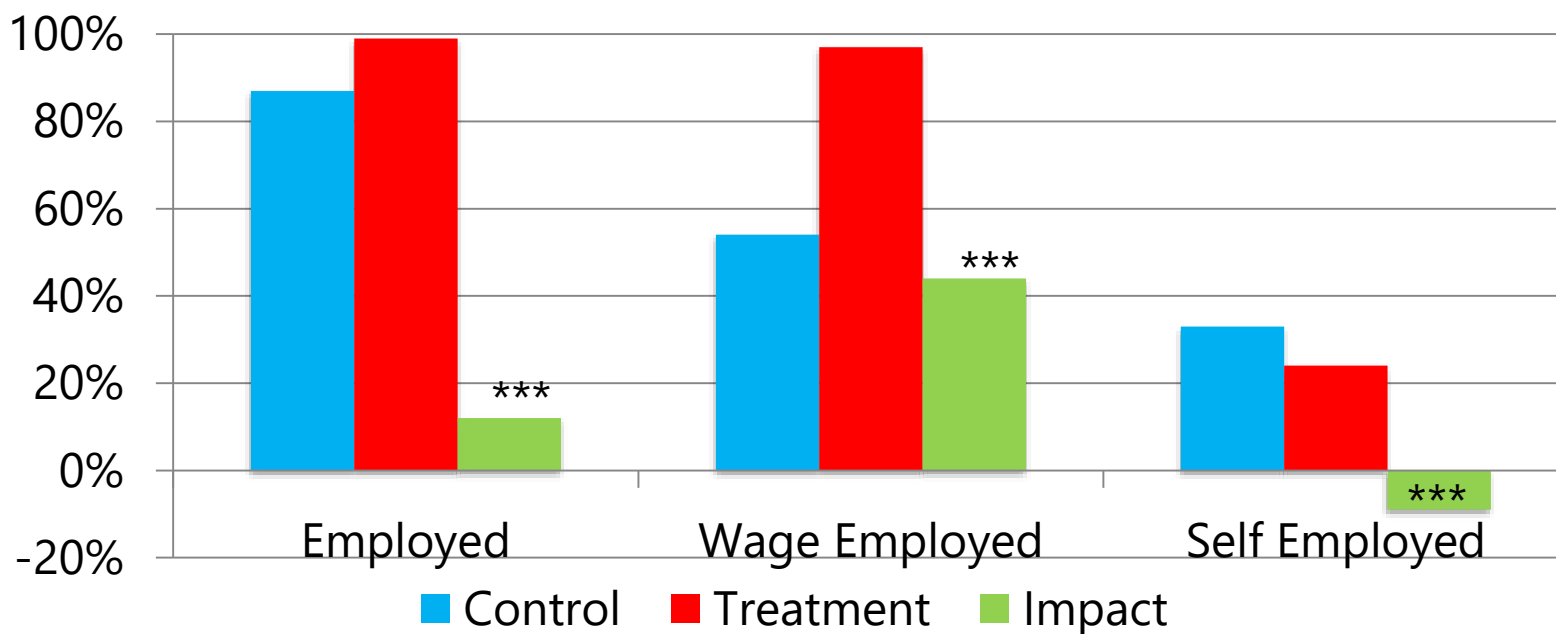# (4-5 months after the start of the program)

# In the short-term, little increase in overall employment... but strong shift into wage employment



- Small impact on overall employment (from 86% to 98%, +12 pp)
  - Only very small reduction of unemployment or inactivity
- Strong impact on **wage employment (from 53% à 97%, +44 pp)**
- But small decrease in self-employment and other types of employment (-9 pp)

# In the short-term, significant increase in earnings… but only by a relatively small share of transfer amount
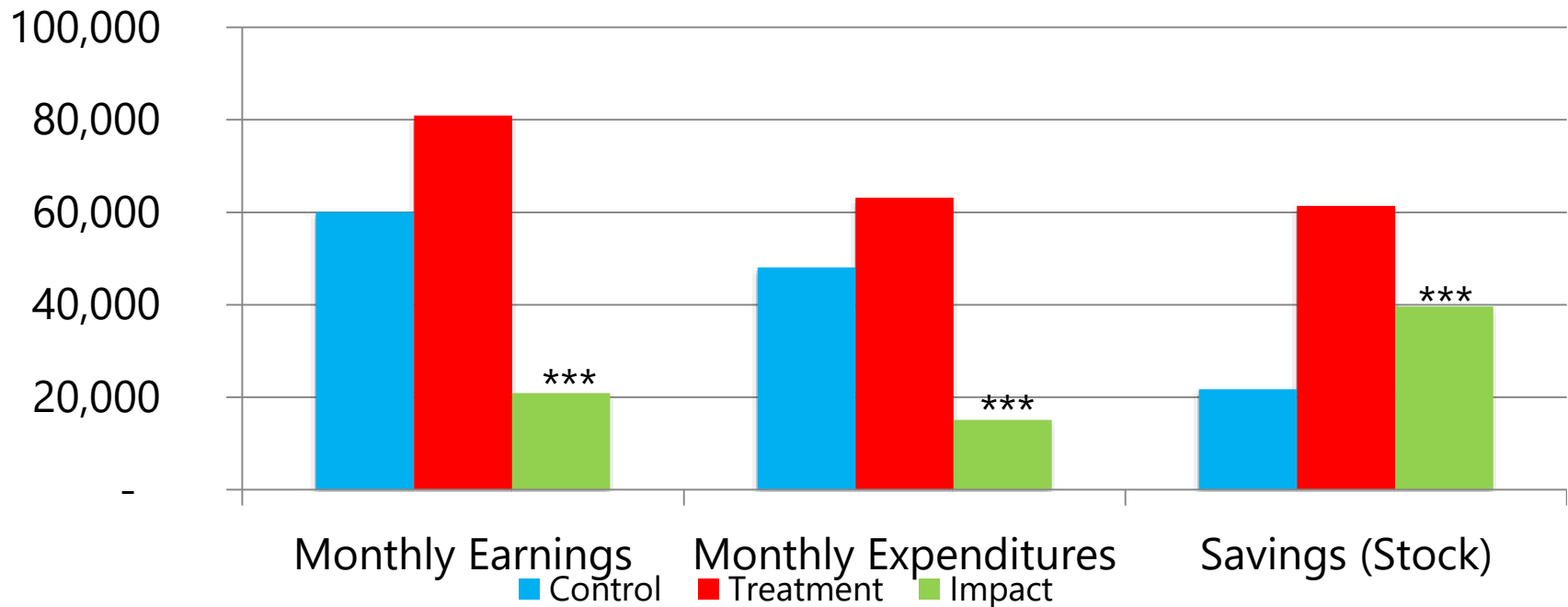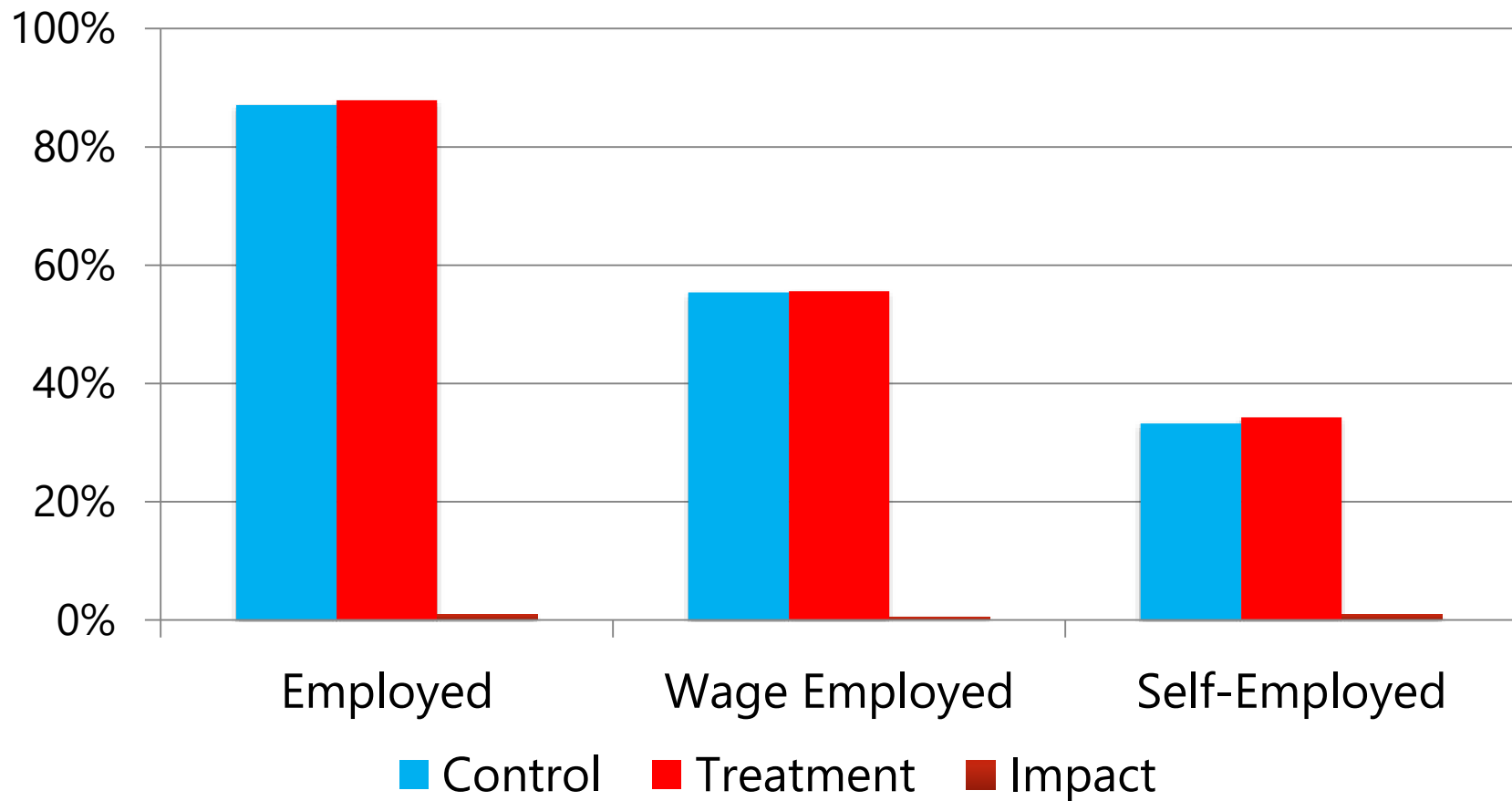


- Total monthly earnings increase from 60,000 FCFA to 81,000 FCFA (+21,000FCFA)
- **Earnings gain as a ratio of transfers 21,000/55,000 = 37.5%**
- Earnings gains contribute to **higher expenditures** (~+15,000/month) and **higher savings** (~+6,000/month)

# In the medium-term…
# (12-15 months after exit from the program)

# In the medium-term… there are no impacts on the level or composition of employment

# But medium-term impacts on earnings are observed



Unit : FCFA

Significant increase in earnings: 5,600 FCFA, or 12% increase
Significant and substantial increase in self employment earnings: 6,200 FCFA, or 32% increase
Former participates are self-employed in slightly larger-scale activities

# Increases in earnings come mostly from the group who participated in public works and self-employment training…
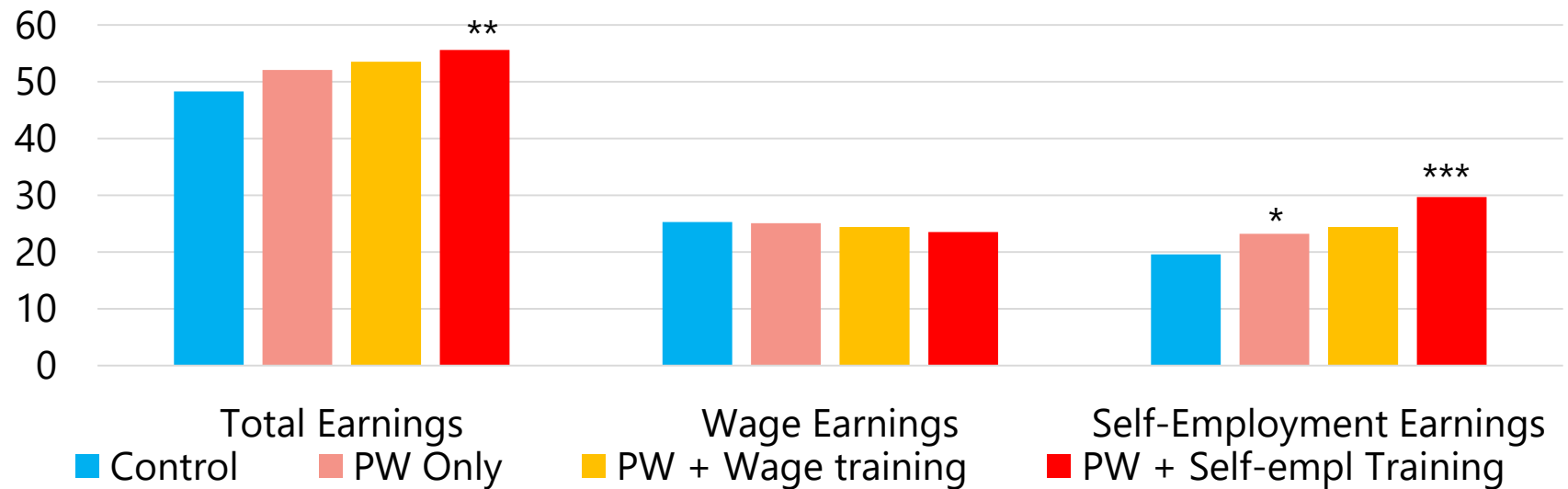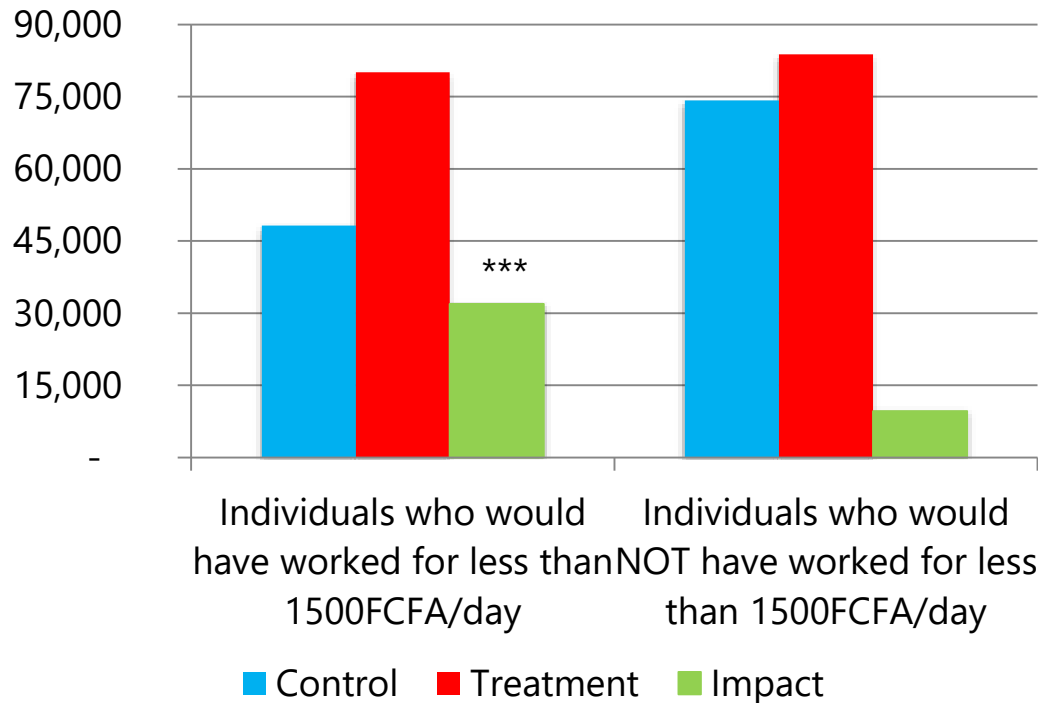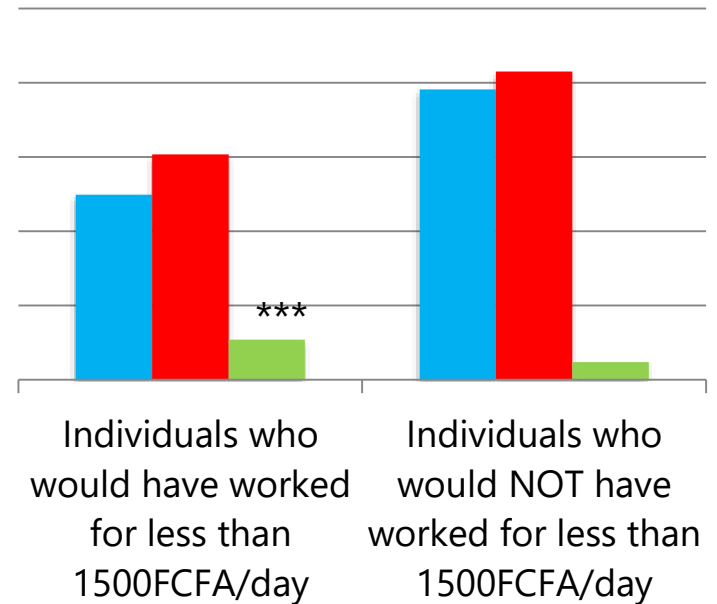


Unité : 000 FCFA

Caveat: differences in earnings between different treatment arms is not significant. Cannot formally say that one type of training is more effective than another

# Earnings increase much more for the vulnerable… improving targeting would make the program much more cost-effective



**Short-term Impacts on Earnings**

**Medium-term Impacts on Earnings**

Short-term chart:
- Individuals who would have worked for less than 1500FCFA/day
- Individuals who would NOT have worked for less than 1500FCFA/day
- *** 

Medium-term chart:
- Individuals who would have worked for less than 1500FCFA/day
- Individuals who would NOT have worked for less than 1500FCFA/day
- ***

Legend: ■ Control ■ Treatment ■ Impact

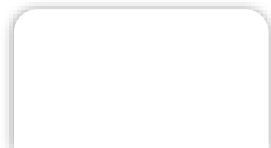# Using Impact Evaluation tools to assess SP delivery system

# Using impact evaluation tools to assess delivery systems

Improving enrollment process

Increasing take up

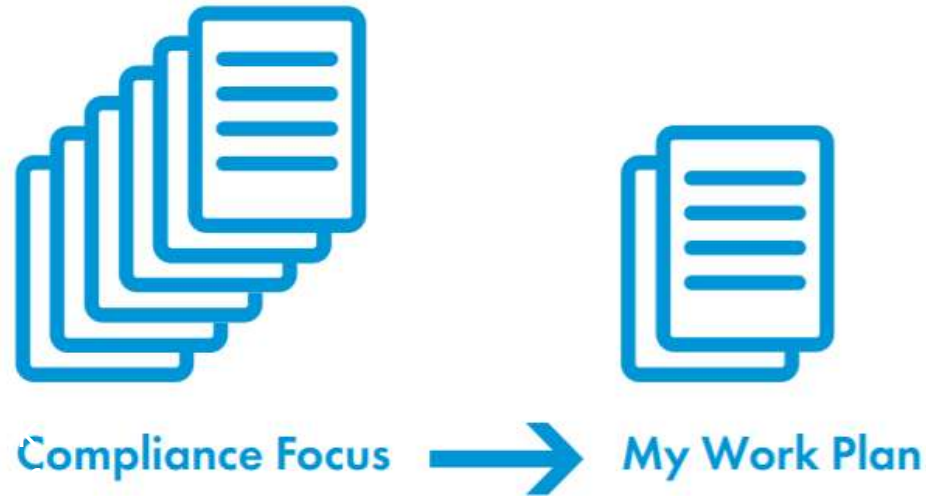Selecting service provider

Monitoring

# Improving enrollment and follow-up process

How to get job seekers into work faster? (reduce duration of unemployment benefits)

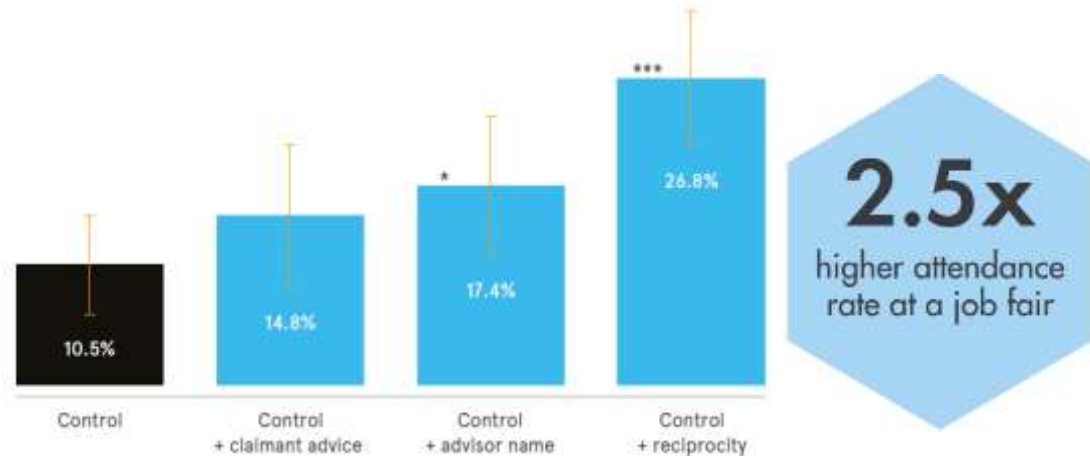UK JobCentre Plus tested shifting from from a 'compliance-focused' approach to an 'outcome-focused' approach



Compliance Focus ➡ My Work Plan

3% of jobseeker find a job faster

# Increasing participation in recruitment events

| Group | Message |
|-------|---------|
| Control | A simple fact-based message |
| Treatment 1 | The control message personalised with the job seeker's name |
| Treatment 2 | The control message personalised with the job seeker's name and the job advisor's name |
| Treatment 3 | The control message personalised with the job seeker's name, the job advisor's name, and a 'good luck' message. |

Using text messages to increase attendance



| | | | | |
|---|---|---|---|---|
| 10.5% | 14.8% | 17.4% * | 26.8% *** | **2.5x** higher attendance rate at a job fair |
| Control | Control + claimant advice | Control + advisor name | Control + reciprocity | |

# Improving monitoring for school feeding program

How to ensure that children are receiving adequate quantities of food from contractors for national school feeding program in Colombia?



**5 RCTs to iteratively test different messages**

• Send **personalized messages** –**27%** higher response rate.

• Ask **closed questions** (rather than open questions), e.g. "Did your child eat X?" – **163%** higher response rate.

• Name **three food items** the child should have eaten (rather than one item), e.g. "Did your child eat rice, broccoli and beans?" – **59%** higher response rate.

• Send messages about the **parent's own child's food** (rather than thousands of similar children) – **32%** higher response rate.

• Send messages in the **morning** (rather than afternoon) – **45%** higher response rate.

# Doing Impact Evaluation can also improve project implementation

- The process of doing [impact evaluation can help delivery projects](#) ("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
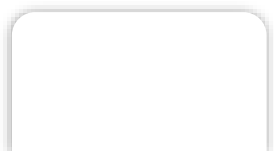  - Complementary qualitative evaluation helps improve project implementation
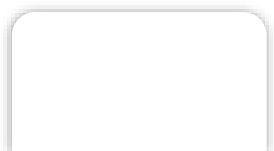
# Key Messages

"

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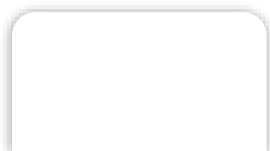
"

Impact evaluation can test whether a program works or not, but also **how to optimize program design and delivery**
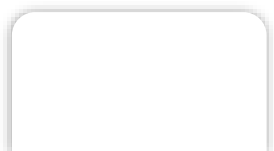
"

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

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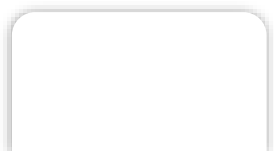
- ○ use comparison or control groups.

"

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

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

# Thank you!



Impact Evaluation in Practice
SECOND EDITION

Paul J. Gertler, Sebastian Martinez,
Patrick Premand, Laura B. Rawlings,
and Christel M. J. Vermeersch

WORLD BANK GROUP    IDB

Reference also: available in Spanish French and Portuguese

**www.worldbank.org/ieinpractice**