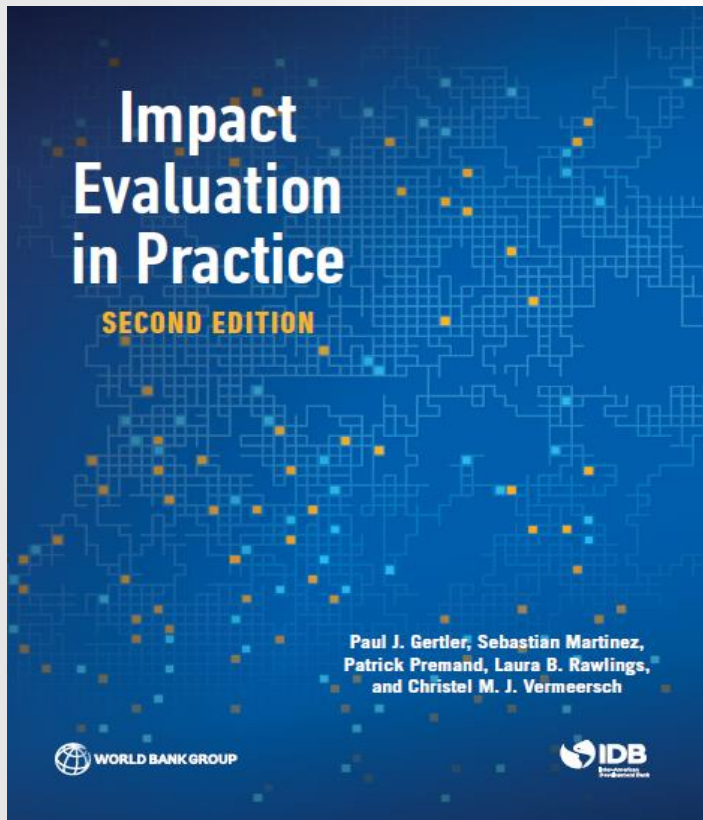




WORLD BANK GROUP
Social Protection & Labor

Impact Evaluation in Practice



Patrick Premand

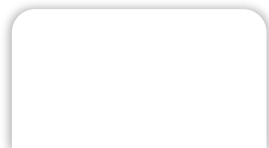
Bern, December 2016

Impact Evaluation in Practice: Our Perspective

- *Accessible introduction* to the topic of impact evaluation and its practice in development.
 - Overview of *quantitative impact evaluation methods*, informed by *program design*
 - *Practical, intuitive guidelines* for *designing and implementing* impact evaluations
-
- *Developed from dozens of training workshops, decades of experience*
 - *Since the first edition in 2011, one of the most downloaded World Bank publications*
 - *Used in training courses and universities globally*
 - *Available in English, French, Spanish and Portuguese*

What's New in the Second Edition?

- New techniques and perspectives on evaluating programs
- State-of-the-art implementation advice
- Expanded set of examples and cases
- New chapter on research ethics and open science
- New chapter on partnerships to conduct impact evaluation.
- Complementary on-line instructional material
- Up to date references on further resources in each chapter
- Updated glossary and key concepts



From M&E to impact evaluation

The main concepts of impact evaluation

Choosing the best design for your project

Randomization as an operational tool

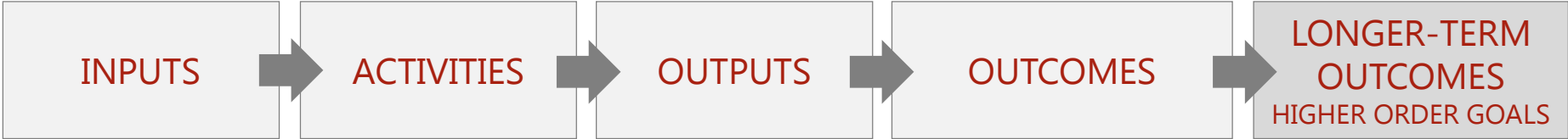
Results from Public Works IE in Cote d'Ivoire



WORLD BANK GROUP
Social Protection & Labor

Introduction to Impact Evaluation

The Results Chain in a Typical Program



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

- **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

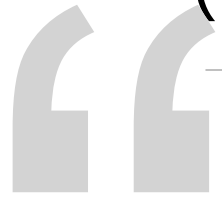
- Evaluate impact selectively, when project is:
 - Innovative
 - Replicable, scalable, or implemented at scale
 - Strategically relevant (e.g. large budget)
 - Evaluation will fill knowledge gap
 - Substantial policy impact
- Impact Evaluation can focus on selective innovations within projects
 - Beyond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

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**)



Public Works (THIMO) in Cote d'Ivoire

- Emergency Youth Employment and Skills Development Project set-up in 2012 following post-electoral crisis (US \$45 million)
 - Public Works Component
 - Skills Development Component (apprenticeships, internships, professional training, entrepreneurship training,...)
- **Public Works Program**
 - Covers 12,500 youths (18-30, 30% women) in 16 localities throughout the country
 - Daily wage rate of CFA 2,500 (~\$5) for 6 months
 - Complementary training:
 - 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 youths' employment and earnings during and after the program?

Design Question

Does the provision of complementary training (entrepreneurship training or job search training) improve labor-market outcomes after exit from the program?



WORLD BANK GROUP
Social Protection & Labor

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).

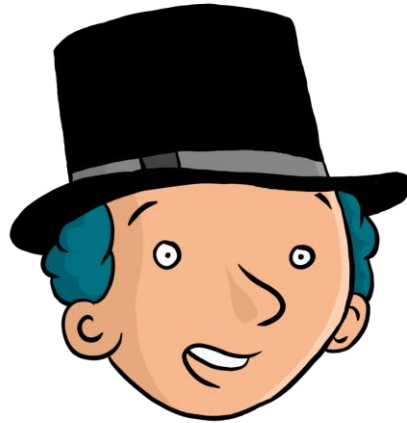
We call this the **Counterfactual**.

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 Bamba



additional money



(P)

on Bamba's consumption



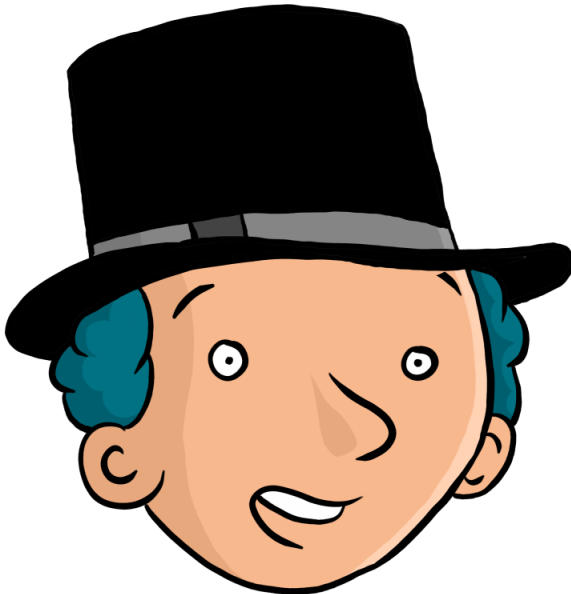
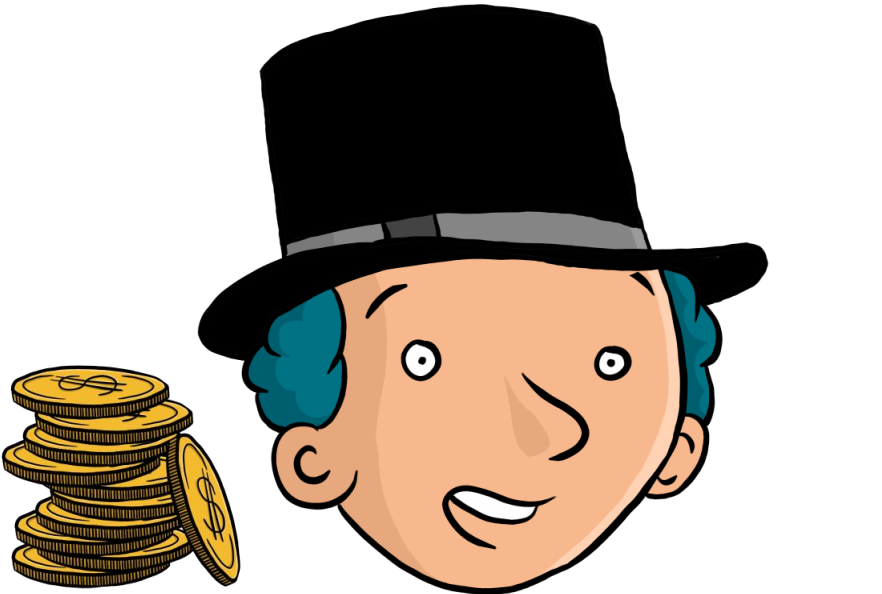
(Y)?



The Perfect Clone

Bamba

Bamba's Clone



6 candies

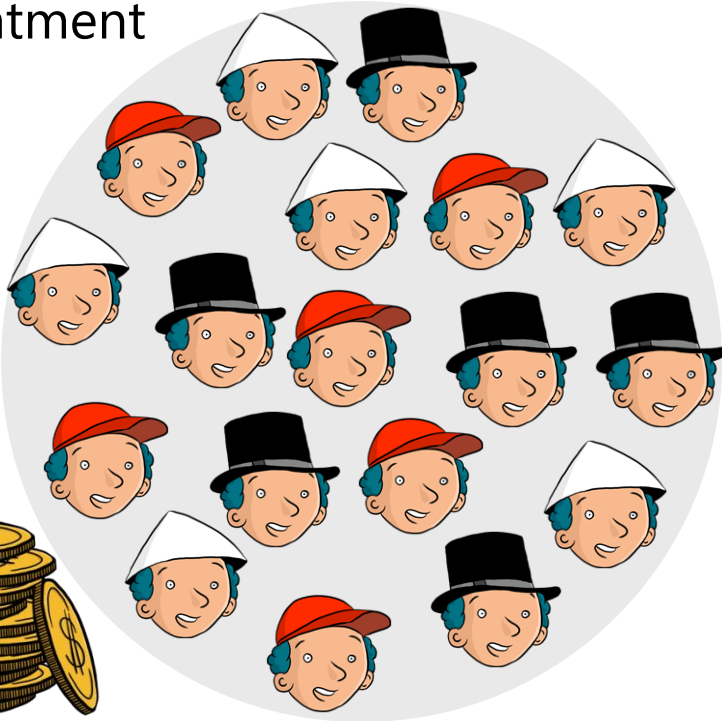


4 candies

$IMPACT = 6 - 4 = 2$ Candies

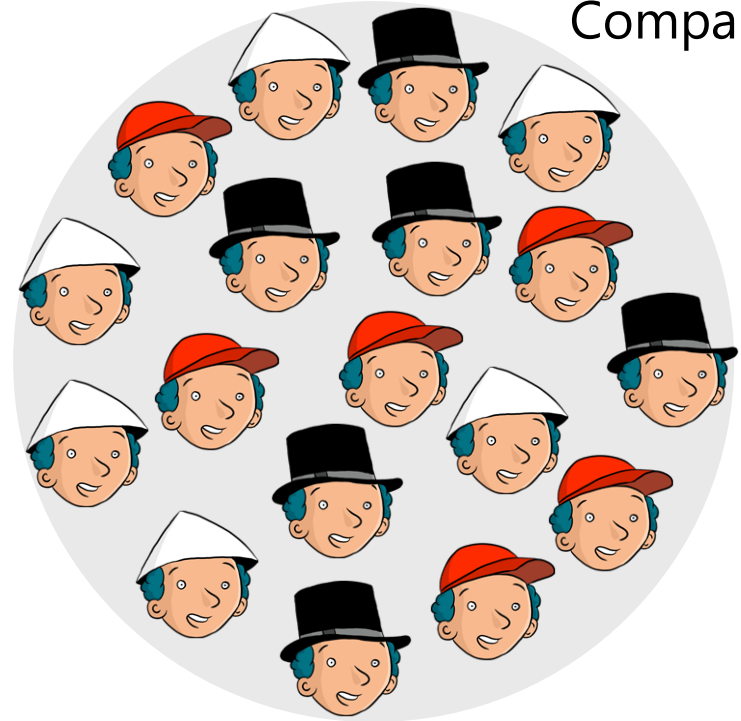
In reality, use statistics

Treatment



Average $Y=6$ candies

Comparison



Average $Y=4$ Candies

$IMPACT=6-4=2$ Candies



WORLD BANK GROUP
Social Protection & Labor

Choosing the best IE design for your project

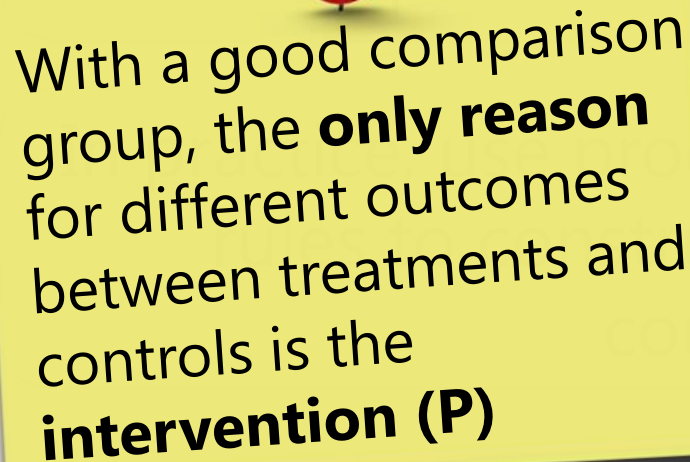
Finding good comparison groups

We want to find **clones** for the Bambas 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

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

Eligibility 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)

<i>Resources</i> →	<i>Excess demand</i>		<i>No Excess demand</i>	
<i>Eligibility</i> →	<i>Targeted</i>	<i>Universal</i>	<i>Targeted</i>	<i>Universal</i>
<i>Timing</i> ↓				
<i>Phased Roll-out</i>	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
<i>Immediate Roll-out</i>	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 Cote d'Ivoire

Resources →	<u>Excess demand</u>		No Excess demand	
Eligibility →	Targeted	<u>Universal</u>	Targeted	Universal
Timing ↓				
<u>Phased Roll-out</u>	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
Immediate Roll-out	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



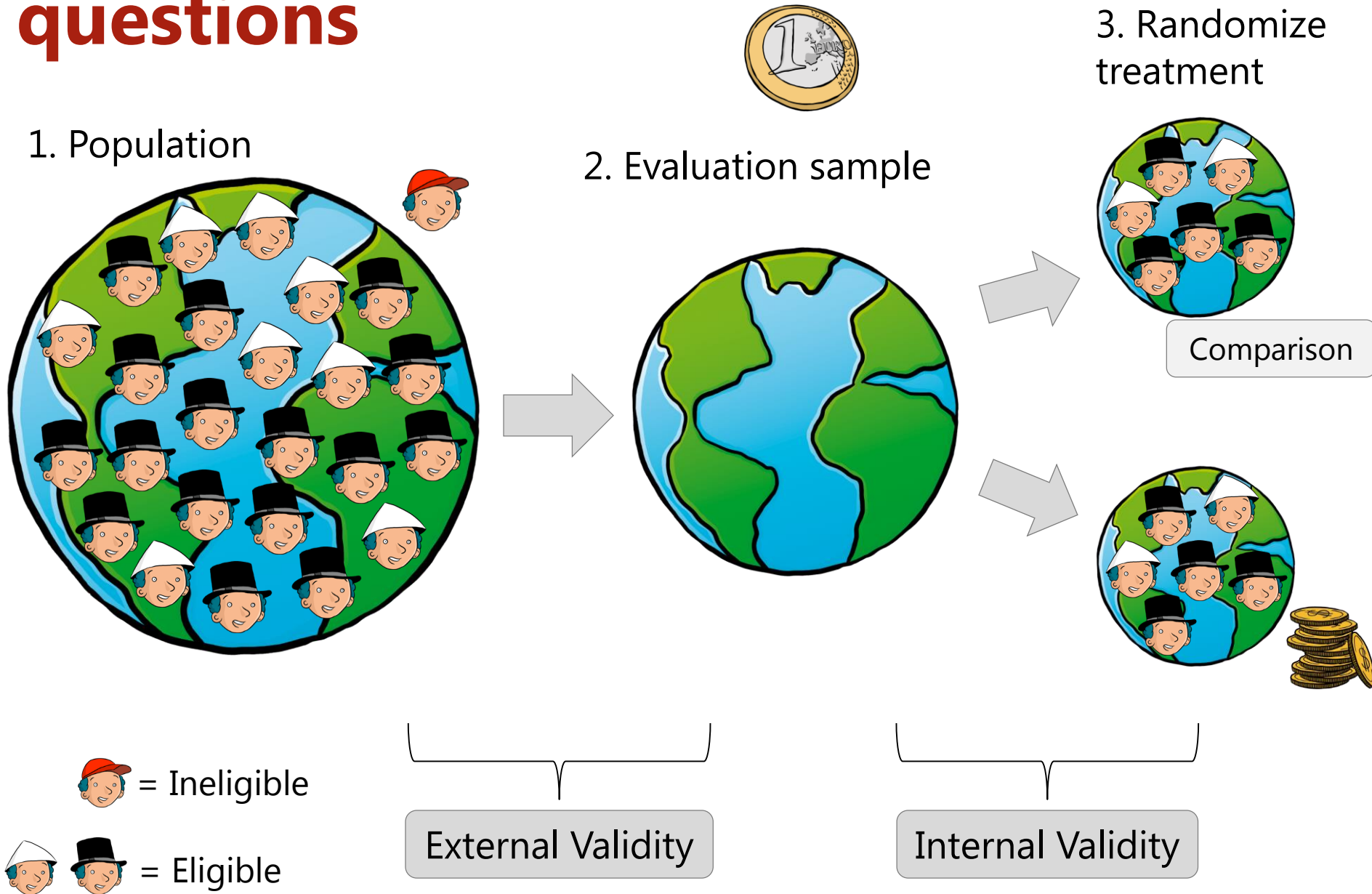
WORLD BANK GROUP
Social Protection & Labor

Randomization as an operational tool

Randomization is not only for the Impact Evaluation. In Cote d'Ivoire...

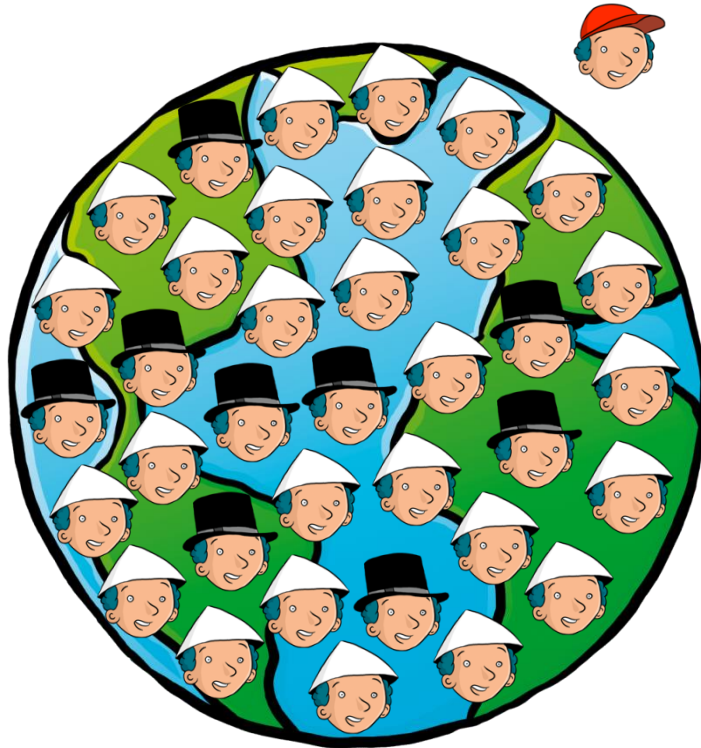
- Public works were initially introduced as an instrument to facilitate ex-combattants' reintegration
- Public lotteries were put in place to allocate the limited places available in the program among the eligible population
 - Transparent allocation mechanisms accepted by all
 - Minimized risks of tensions in the post-conflict context
 - Implemented separately in each locality, separately for men and women (stratification)
- The impact evaluation used existing lotteries.

Randomization to answer basic IE questions



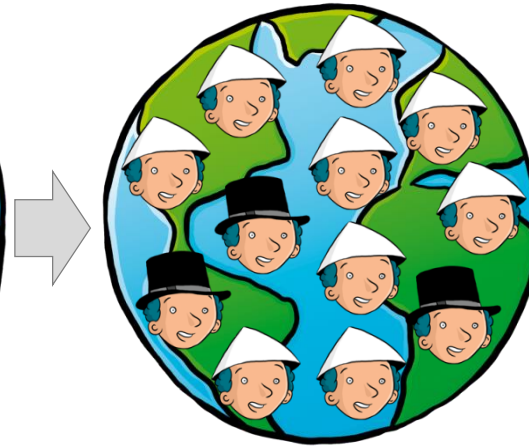
Randomized to alternative program modalities



1. Population



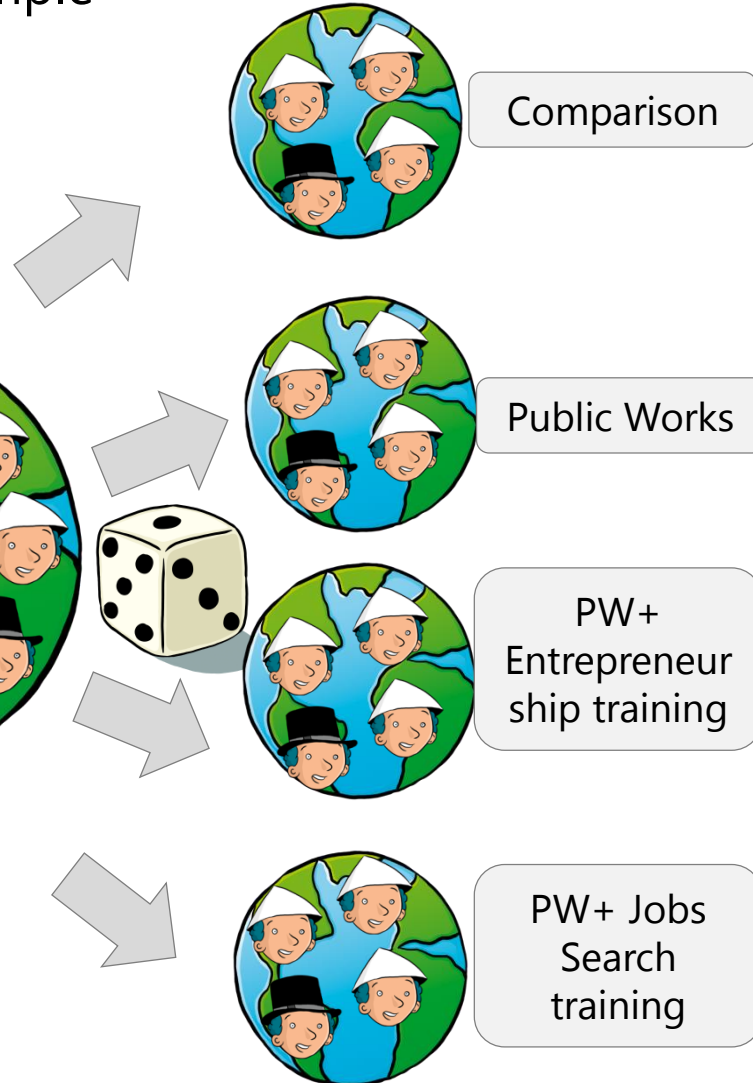
 = Not eligible

2. Evaluation Sample



  = Eligible

3. Randomize treatment.



Randomized Assignment



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

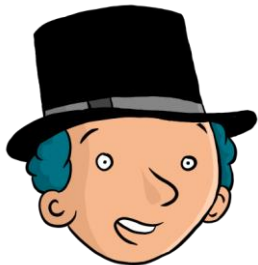
We have identified the perfect **clone**.

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

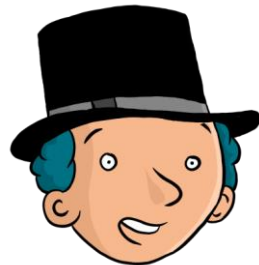
Most pilots and new programs fall into this category.

Consider evaluating relative effectiveness of alternative program design options.

Randomized beneficiary



Randomized comparison



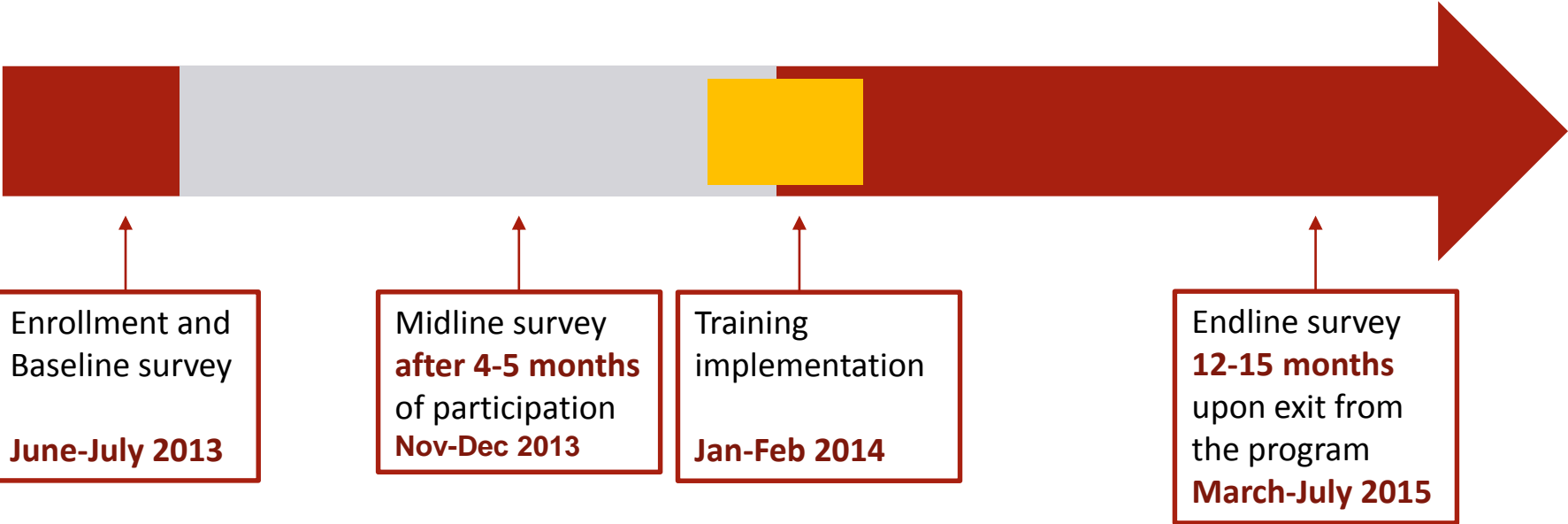


WORLD BANK GROUP
Social Protection & Labor

Results from Public Works Impact Evaluation in Cote d'Ivoire

Timeline of Cote d'Ivoire public works impact evaluation

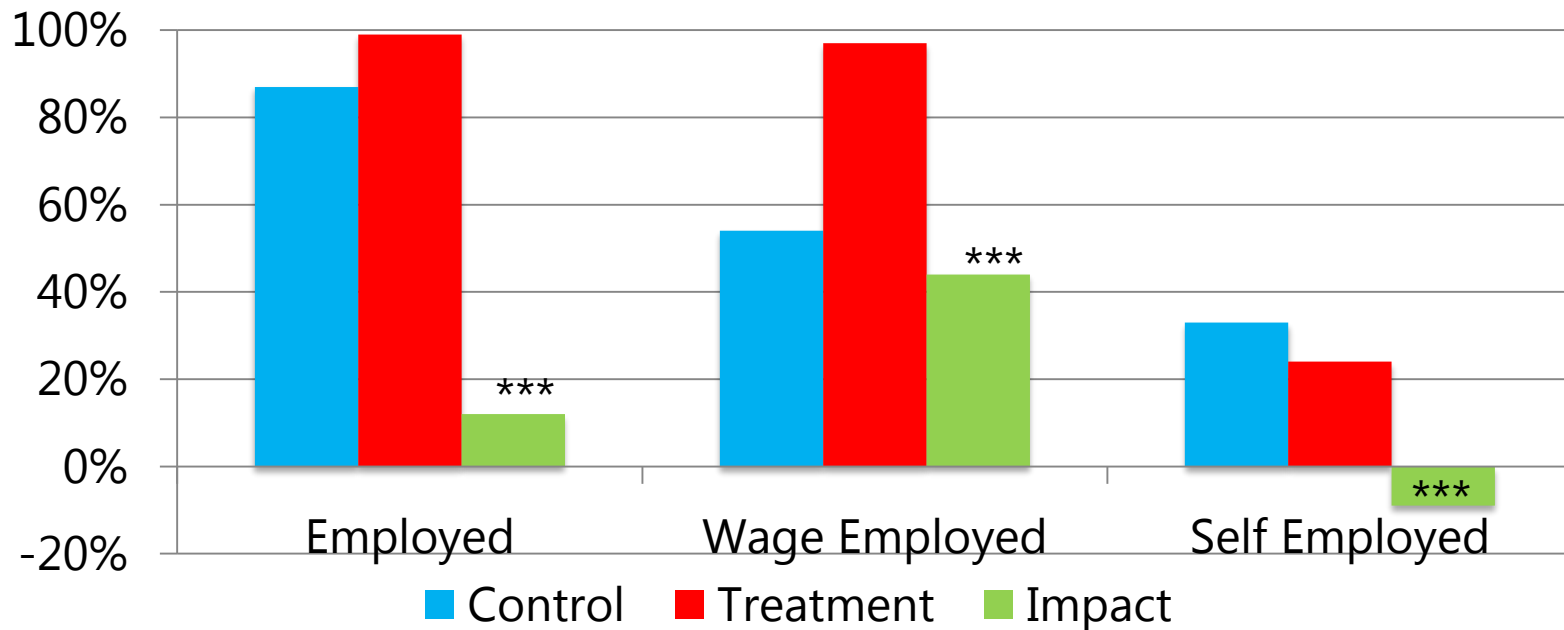
LIPW operates in waves, impact evaluation focuses on **wave implemented between July/August to February/March 2014**



**In the short-term...
(4-5 months after the start of the program)**

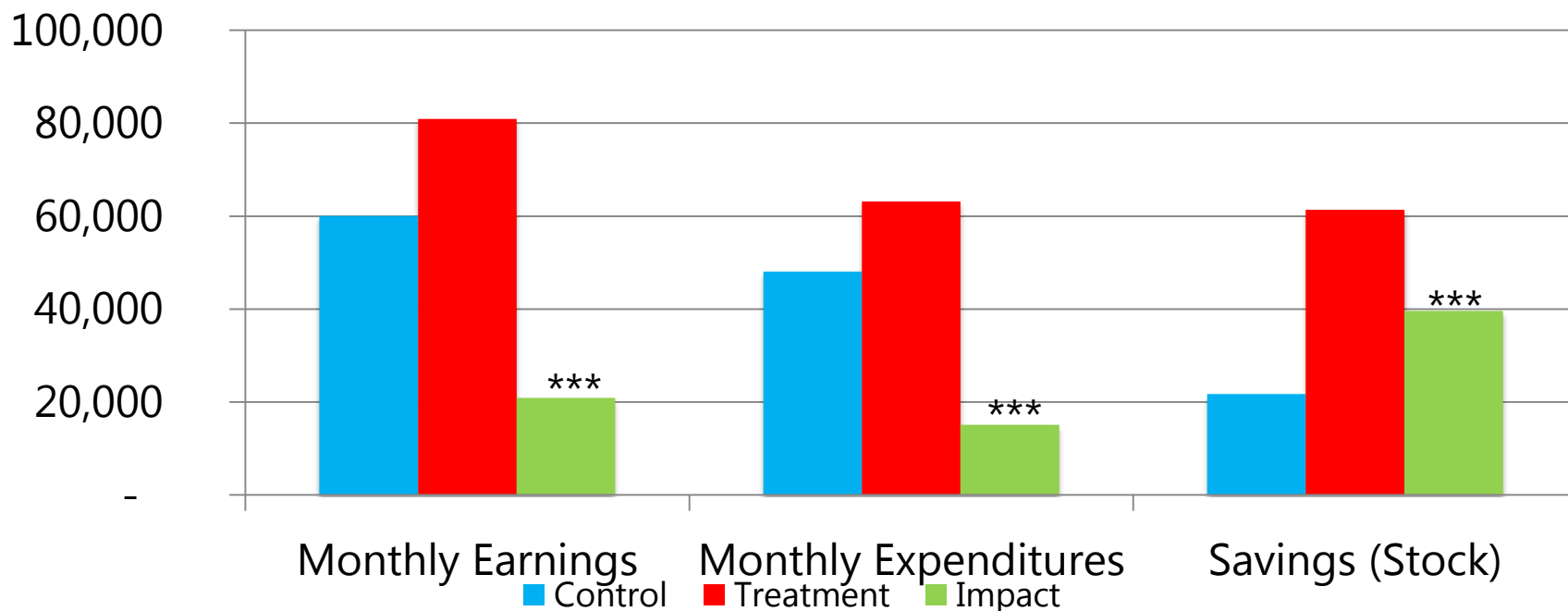


In the short-term, little increase in overall employment... but strong shift into wage employment



- 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)

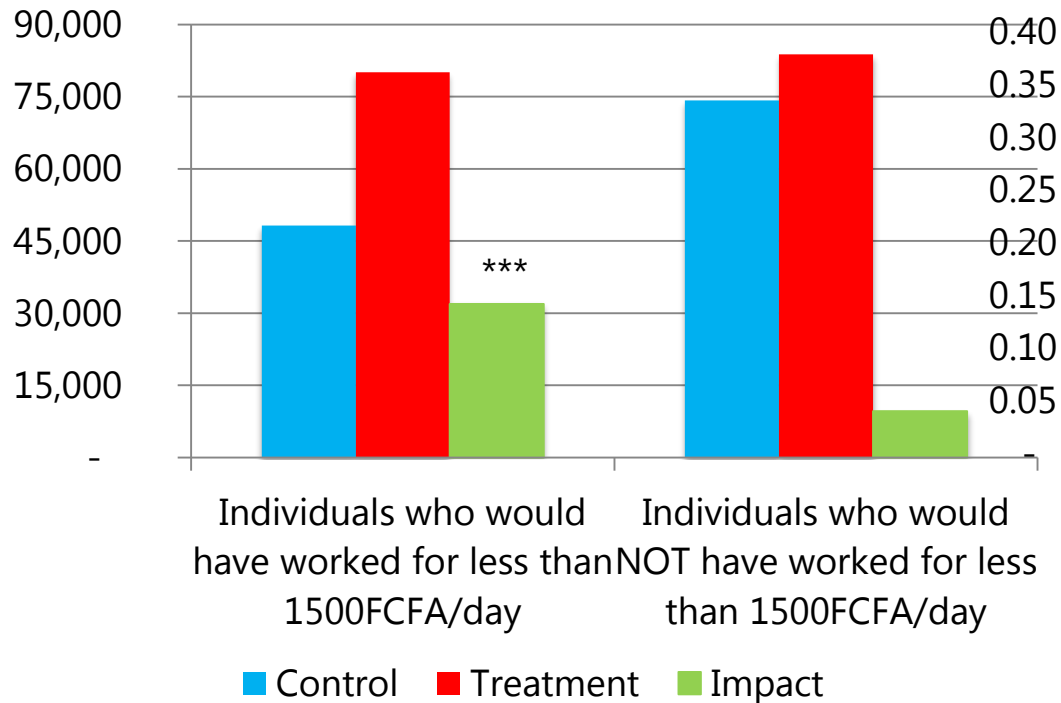
In the short-term, significant increase in earnings... but only by a relatively small share of transfer amount



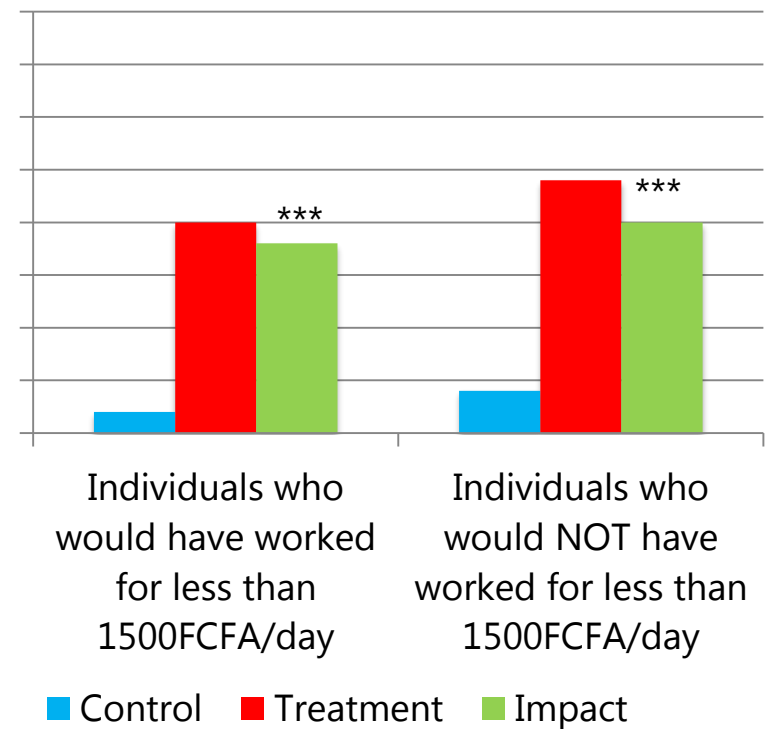
- Total monthly earnings increase from 60,000 FCFA to 81,000 FCFA (+21,000FCFA)
- **Earnings gain as a ratio of transfers $21,000/55,000 = 37.5\%$**
- Earnings gains contribute to **higher expenditures** (~+15,000/month) and **higher savings** (~+6,000/month)

Earnings increase much more for the vulnerable... but the less vulnerable benefit as much in terms of psychological well-being

Impacts on Earnings



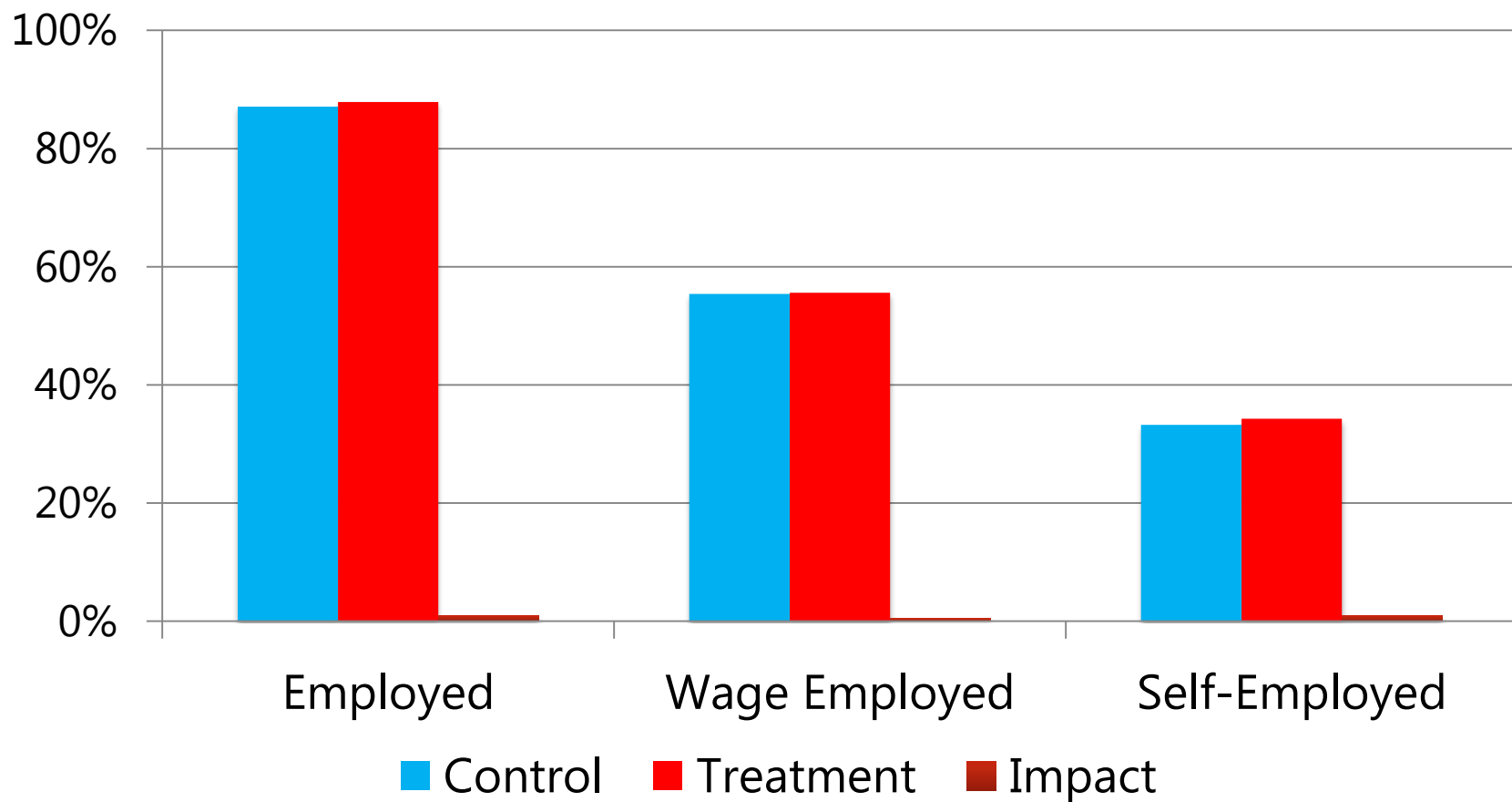
Impacts on psychological well-being



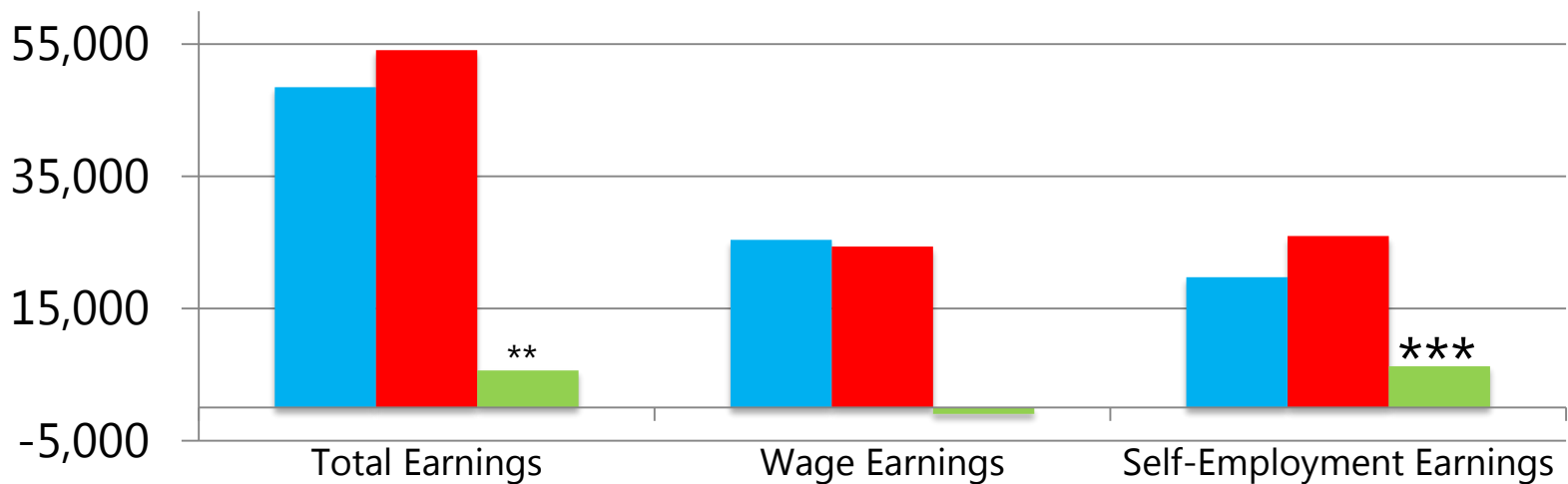
**In the medium-term...
(12-15 months after exit from the
program)**



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

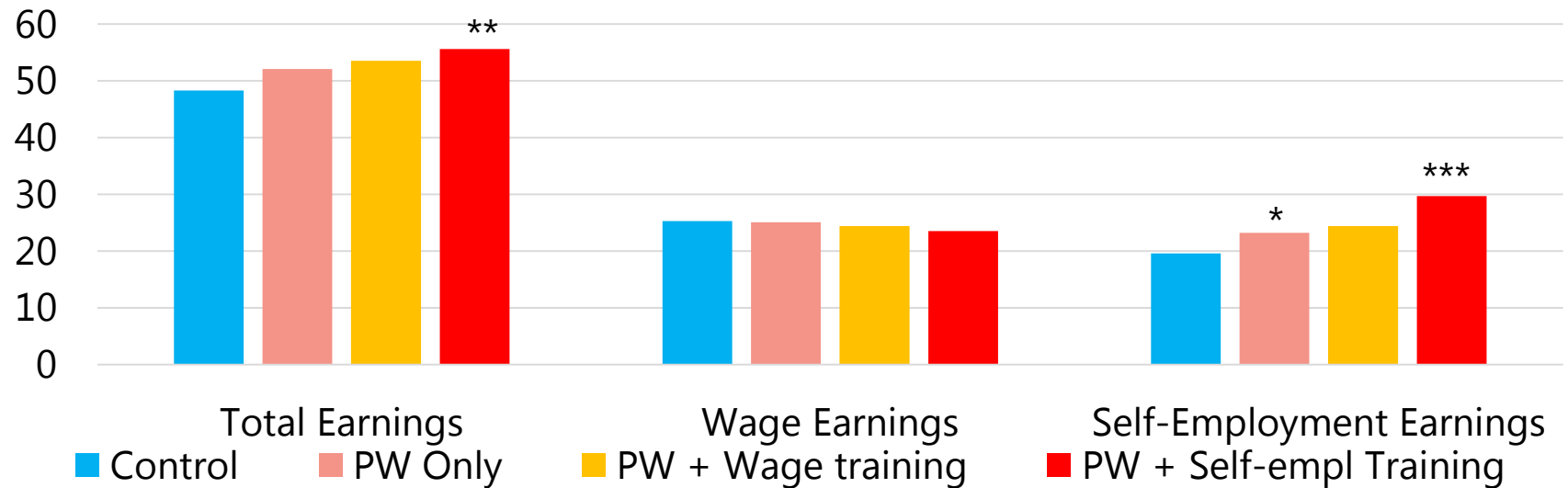
■ Control ■ Treatment ■ Impact

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 participants 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...

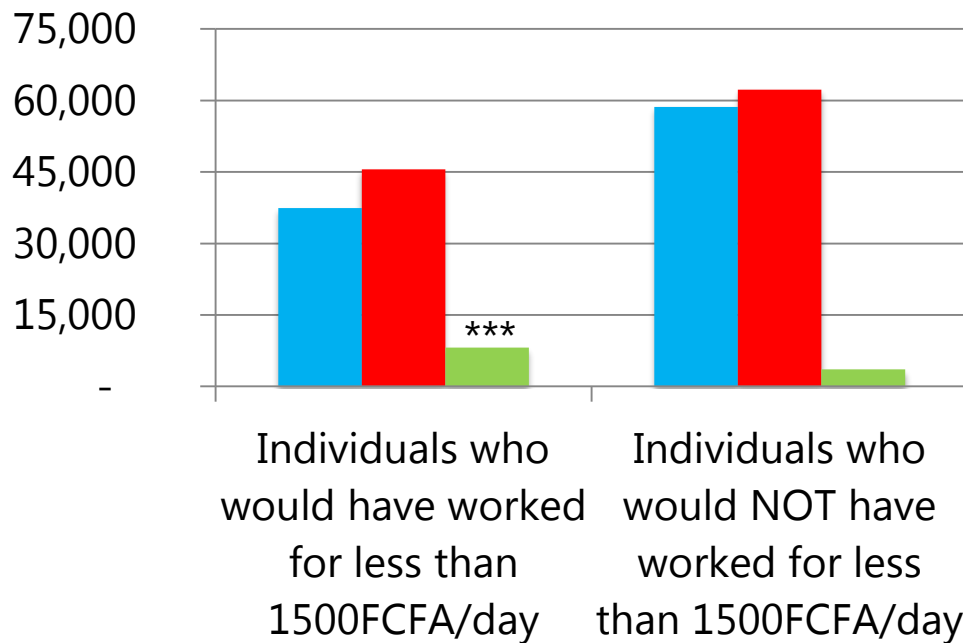


Unité : 000 FCFA

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

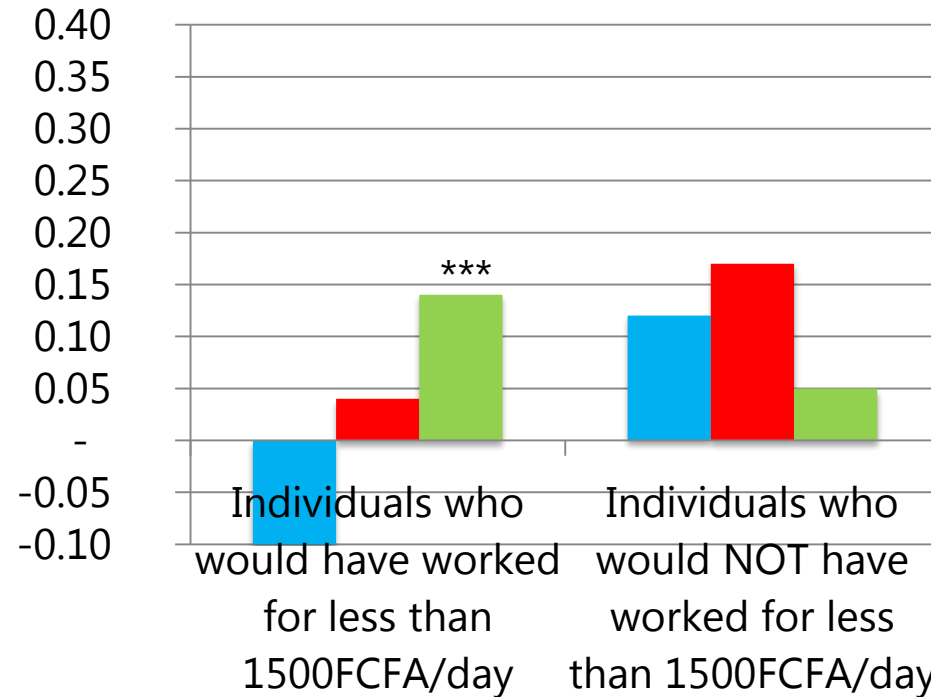
The most vulnerable who gain most in the short-term also benefit the most in the medium-term...

Impacts on earnings



■ Control ■ Treatment ■ Impact

Impacts on psychological well-being



■ Control ■ Treatment ■ Impact

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

Impact Evaluation as a collaboration in Cote d'Ivoire

- Impact evaluation implemented in **close collaboration** between policy and research teams:
 - Government of Cote d'Ivoire (in particular BCPE, Bureau de Coordination des Programmes d'Emploi)
 - World Bank
 - External researchers/academics
 - Funding from the PEJEDEC project as well as Trust Funds at the World Bank.
- Critical to align incentives and ensure visibility



WORLD BANK GROUP
Social Protection & Labor

Key Messages



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





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.





Policy and research teams need to work in **partnership** for the impact evaluation to succeed.



Thank you!



Reference also:
available in **Spanish**
and **Portuguese**

www.worldbank.org/ieinpractice