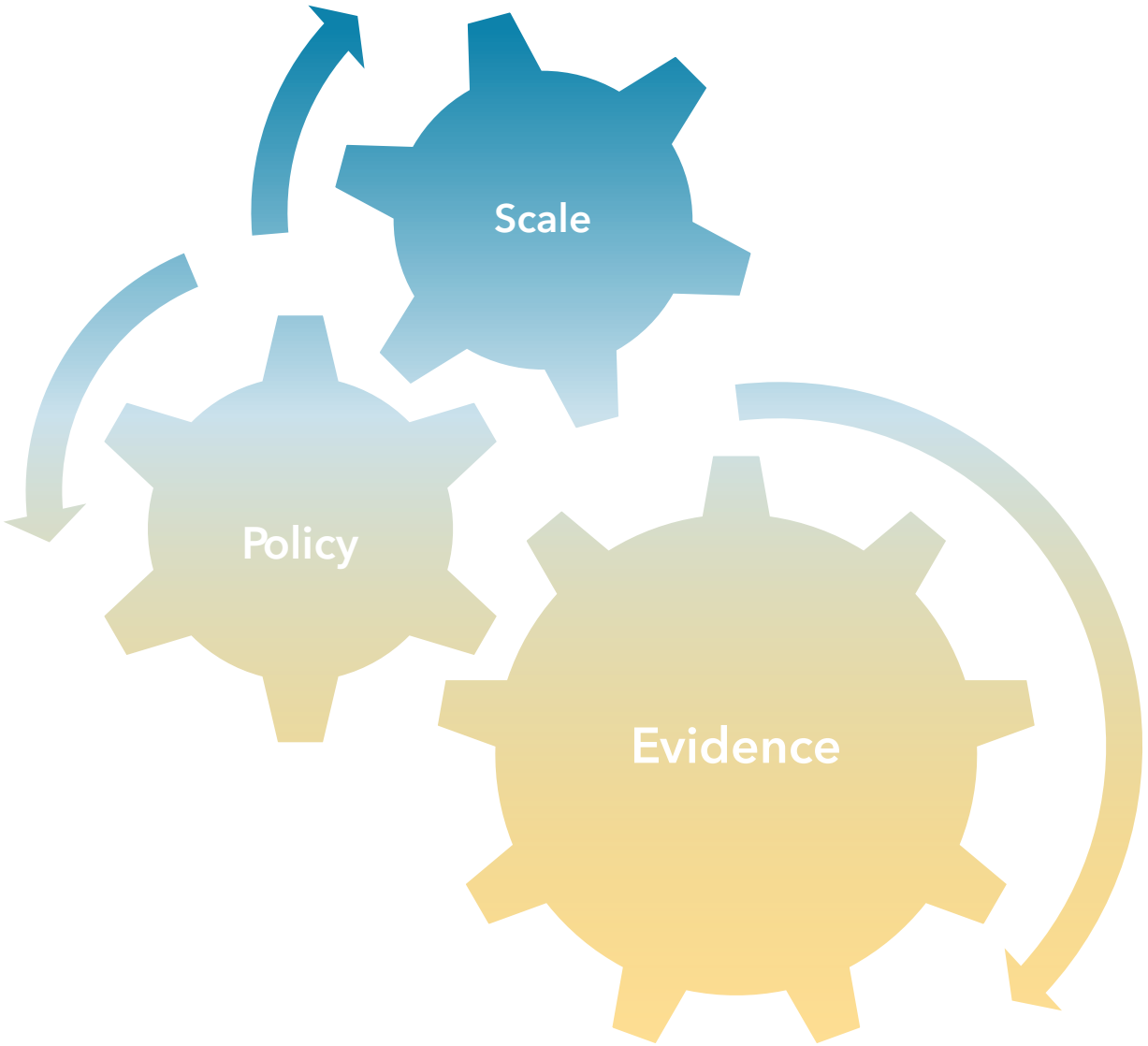
PEI Impact Evaluation Workshop

Moving Economic Inclusion to scale

PARTNERSHIP FOR ECONOMIC INCLUSION

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

i2i DIME
TRANSFORM DEVELOPMENT



Experimental Impact Evaluation Methods

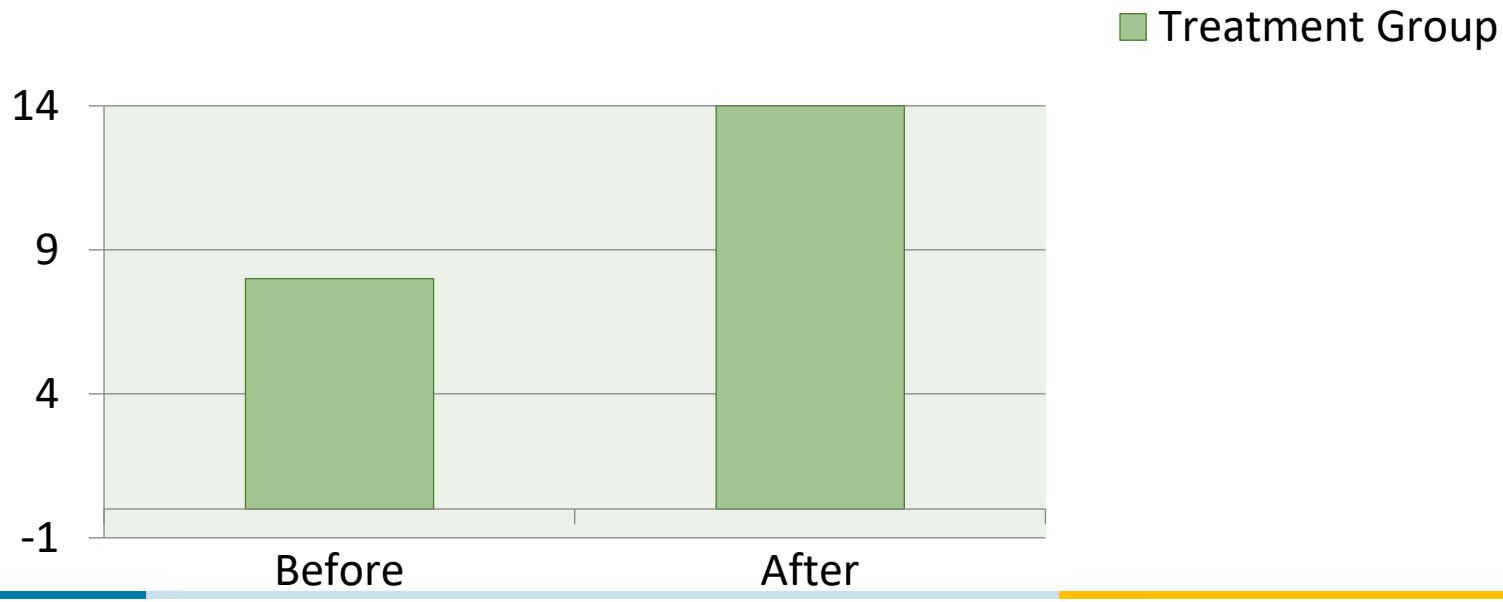
PARTNERSHIP FOR ECONOMIC INCLUSION

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

i2i DIME
TRANSFORM DEVELOPMENT

Measuring Causal Impacts

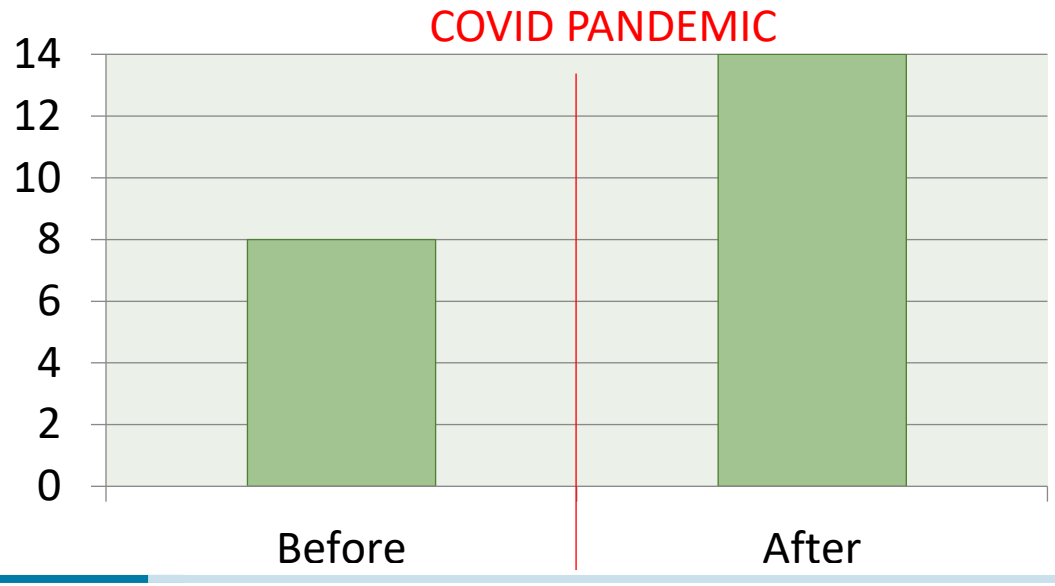
- We want to measure the **causal impact** of an intervention
- Can we do a **comparison over time**?



Measuring Causal Impacts

- We want to measure the **causal impact** of an intervention
- Can we do a **comparison over time**?

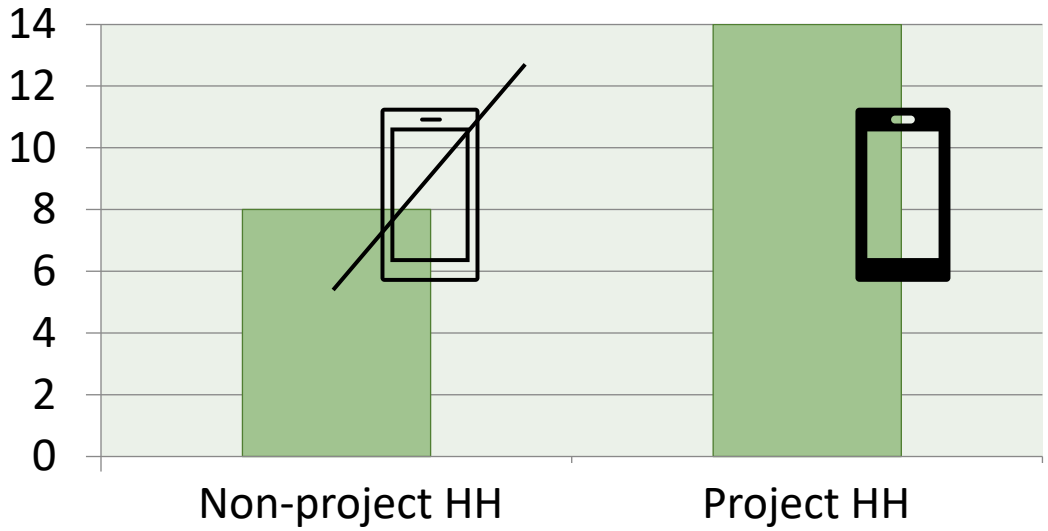
No!
Other things may have happened in the meantime



Measuring Causal Impacts

- We want to measure the **causal impact** of an intervention
- Can we do a **comparison across households**?

No!
Suppose the project gave mobile money transfers. We cannot compare those who have vs do not have mobile money!

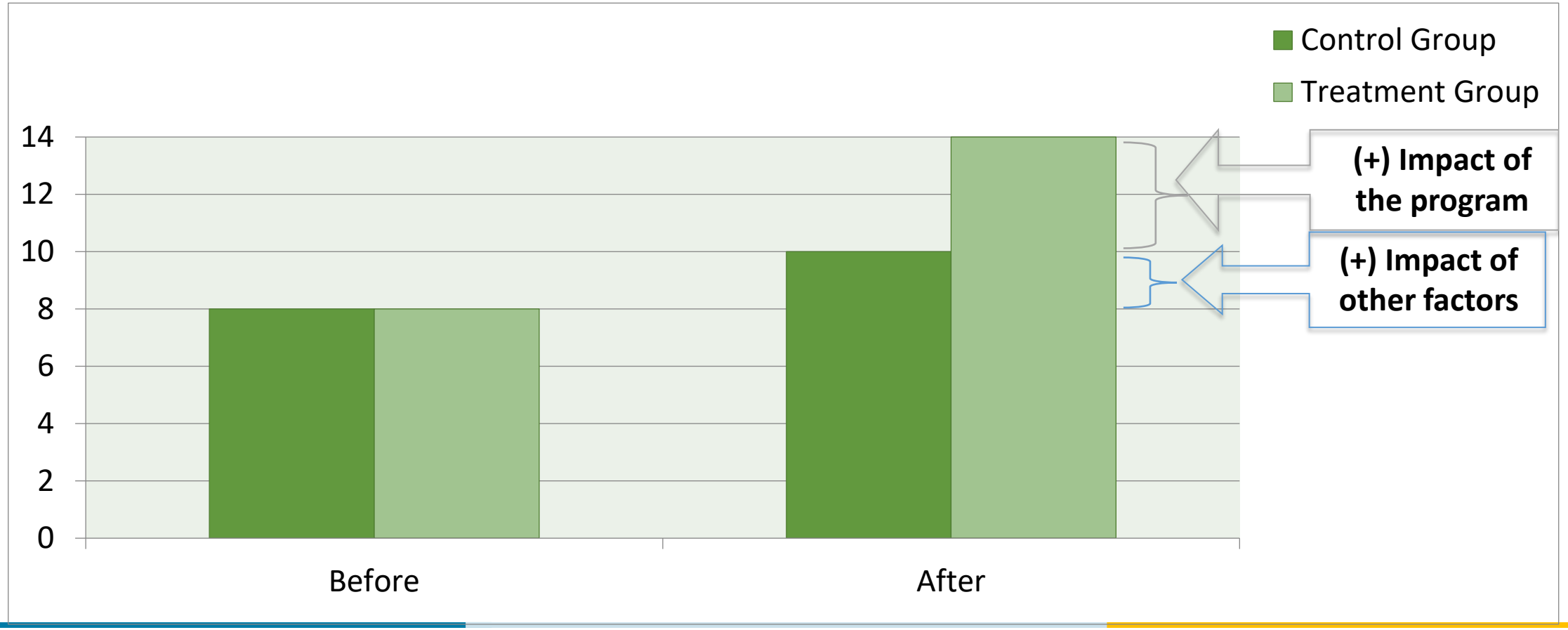


Measuring Causal Impacts

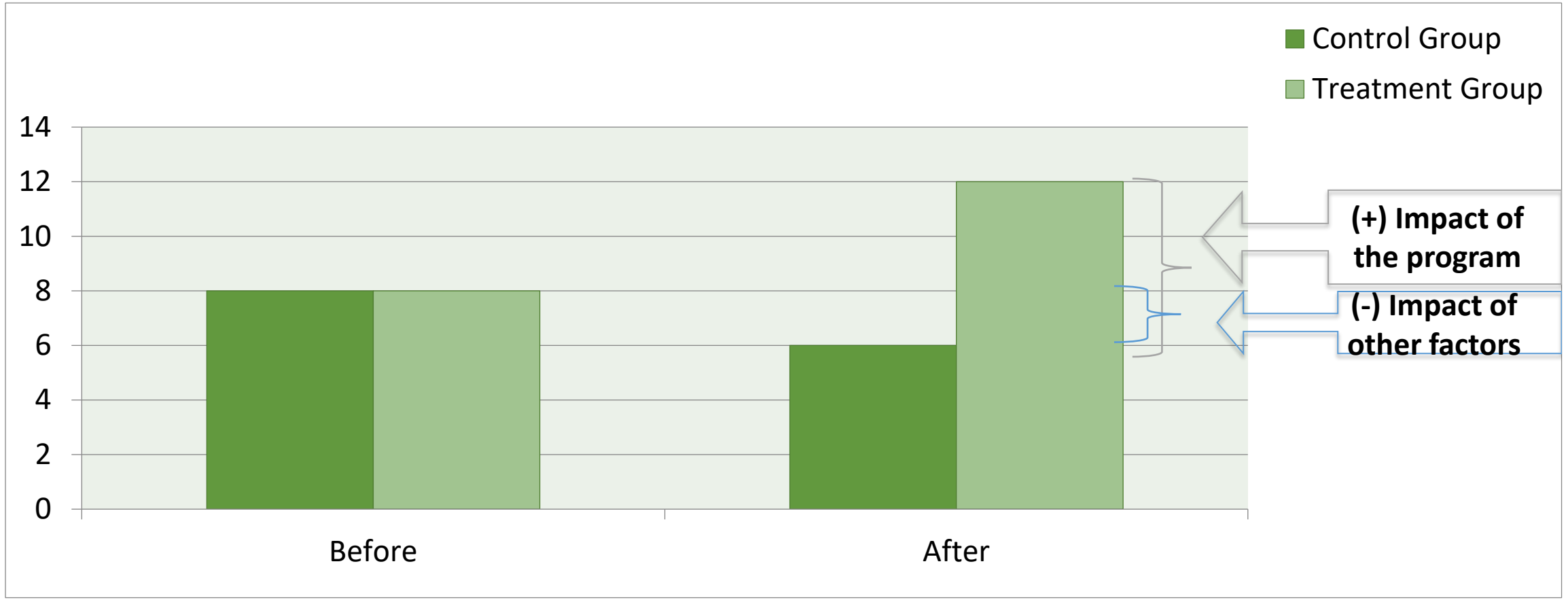
- We want to measure the **causal impact** of an intervention
- To assess causal impacts, ideally want to compare:
  - What happens to people **receiving** the program
  - What would have happened to **same** people **in absence** of the program

- Impossible! What is the best alternative? A credible control group
  - Measure treatment group outcomes, control group outcomes
  - Compare treatment to control group: **difference** is project impact

Measuring Causal Impacts



Measuring Causal Impacts



What Makes a Credible Control Group?

- Treatment and Control group need to be as similar as possible
- Recap: cannot compare same people over time
  - Other intervening factors
- Cannot compare people who received project to people who did not
  - Why did project leaders choose to carry our project there and then?
  - Why did people choose to be part of the project?
  - → Selection bias can threaten our results

- But then, how to build a control group?

Randomization is the Gold Standard

- Before starting the program:
  - assign potential beneficiaries to treatment or control group…
  - …in a randomized fashion: a randomization algorithm

- With a **large enough** sample, the two groups will be identical
  - In terms of observable characteristics: age, gender, occupation
  - In terms of unobservable characteristics: motivation, entrepreneurship, ability

- Only difference between two groups is treatment
  - → impact estimates are unbiased

Case Study: TUP in Bangladesh (Bandiera et al 2017)

- Want to study impacts of a graduation program in 1,309 villages
- Can I compare
  - the 6,732 **ultra-poor** households who are eligible
  - To the 6,743 **near-poor** who are not eligible?
- **NO!** The two groups are not comparable (see assets, savings)

	(1) Ultra-poor	(2) Near-poor	(3) Middle class	(4) Upper class
Consumption and assets				
Household is below the $1.25 a day poverty line	0.530	0.493	0.373	0.121
Consumption expenditure (per adult equivalent)	627.8	645.1	759.5	1,234.2
Household assets [$]	36.5	68.1	279.9	1,663.4
Household savings [$]	7.9	22.1	84.5	481.9
Household receives loans	0.191	0.393	0.498	0.433
Household gives loans	0.012	0.018	0.030	0.067
Business assets (excl. livestock and land) [$]	22.9	54.4	286.1	1,569.8
Livestock				
Household owns cows	0.055	0.154	0.469	0.733
Household owns goats	0.092	0.142	0.300	0.425
Value of cows [$]	33.8	120.2	633.8	1,559.1
Value of goats [$]	7.97	12.8	39.8	71.3
Household rents cows for rearing	0.070	0.148	0.118	0.030
Household rents goats for rearing	0.111	0.157	0.102	0.021

Case Study: TUP in Bangladesh (Bandiera et al 2017)

- Want to study impacts of a graduation program in 1,309 villages
- Can I randomize villages in T and C and compare
  - **Ultra-poor** households in T villages
  - To **ultra-poor** households in C villages?
- **YES!** The two groups are comparable

	(1) Treated Villages	(2) Control Villages	(3) t-test [Treatment=Control]	(4) Normalized Differences
B. Poverty, Expenditures and Financial Wealth				
Below the $1.25 a day poverty line [yes=1]	.556	.584	.524	-.040
	(.400)	(.398)		
Consumption expenditure, per adult equivalent	629	613	.501	.047
	(246)	(236)		
Value of household assets	36	37	.829	-.011
	(48)	(63)		
Household savings	6.2	9.2	.071	-.059
	(28)	(43)		
Household receives loans	.20	.18	.441	-.044
	(.40)	(.38)		
Household gives loans	.011	.014	.356	-.022
	(.10)	(.12)		
Average standardized difference (p-value)			*.849*	
C. Productive Assets				
Cows value	36	30	.575	.023
	(176)	(166)		
Goats value	6.5	8.5	.261	-.050
	(25)	(31)		
Household rents in land [yes=1]	.058	.061	.875	-.007
	(.235)	(.239)		
Household owns land [yes=1]	.068	.062	.738	.017
	(.252)	(.241)		
Value of land owned	175	238	.390	-.027
	(997)	(2190)		
Value of other business assets	23	23	.991	-.0004
	(79)	(101)		

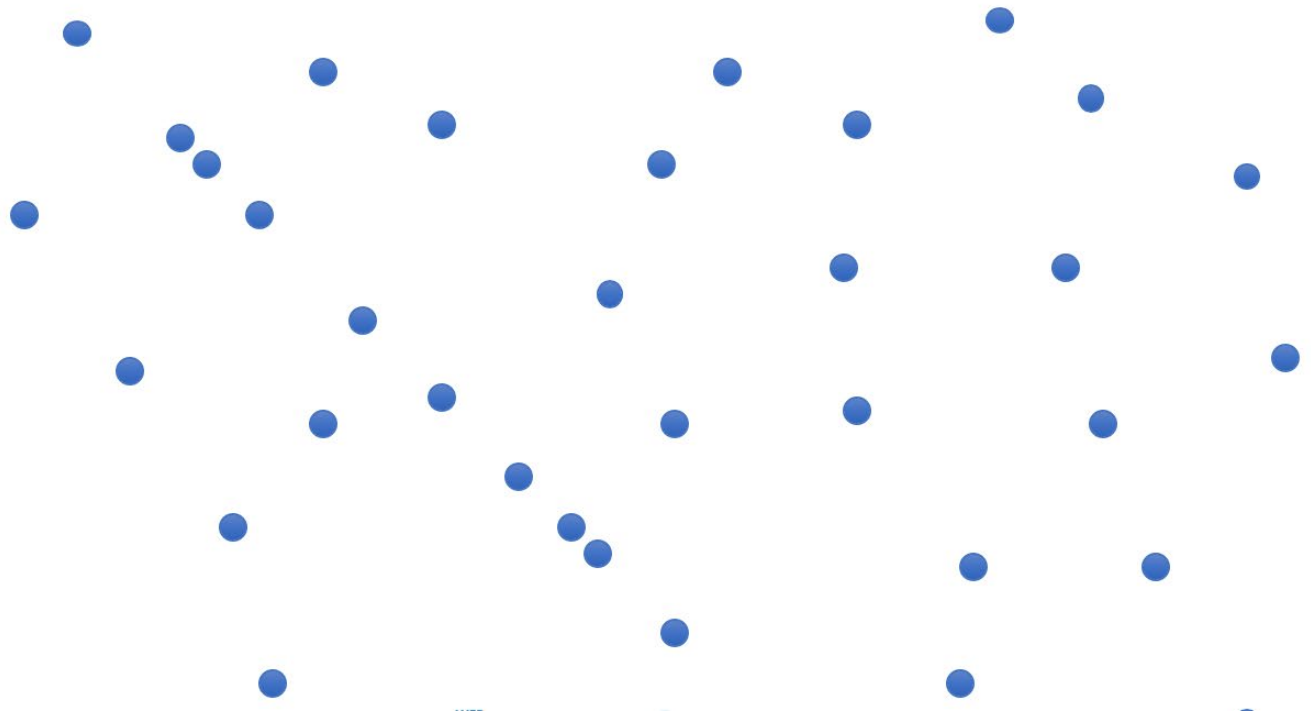
Key Steps for RCTs

1. Decide (recruit) the universe of individuals that would be eligible
2. Decide unit of randomization (individual, household, village, group…)
3. Randomize into treatment(s) and control
4. Test that randomization worked by checking balance on observables
5. Conduct intervention with treatment group(s)

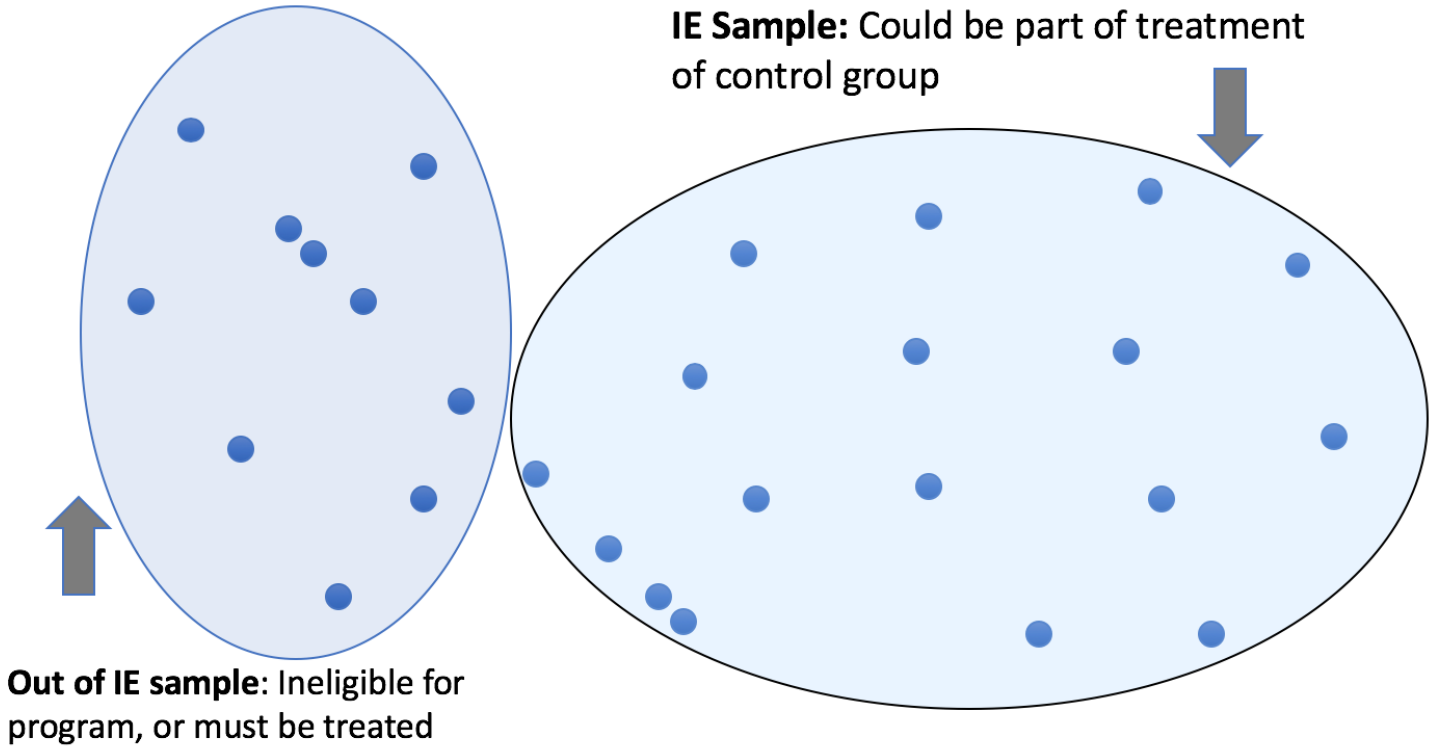
Key Steps for RCTs

1. **Decide (recruit) the universe of individuals that would be eligible**
2. Decide unit of randomization (individual, household, village, group…)
3. Randomize into treatment(s) and control
4. Test that randomization worked by checking balance on observables
5. Conduct intervention with treatment group(s)

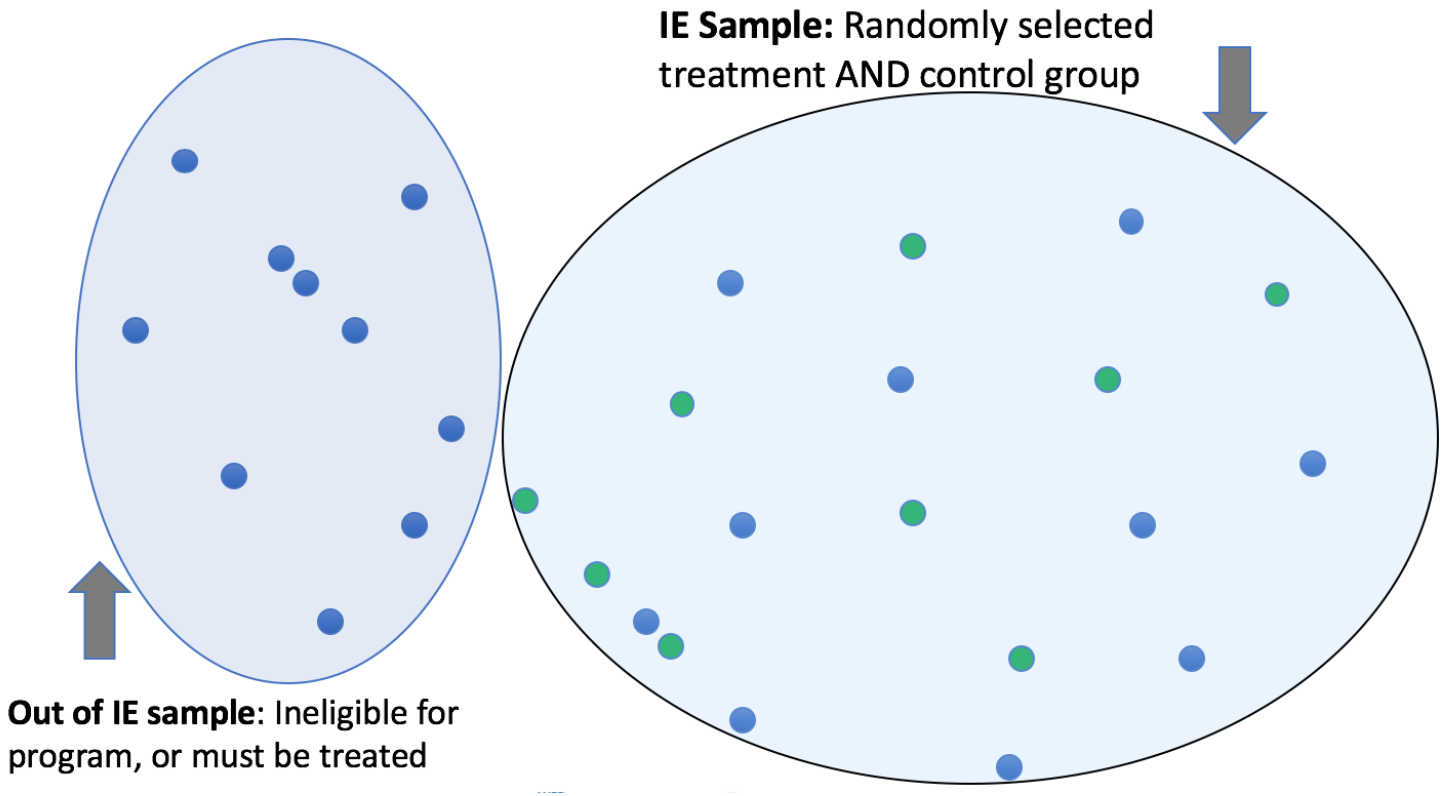
Step 1 – Recruit eligible population



Step 1 – Recruit eligible population



Step 1 – Recruit eligible population



Key Steps for RCTs

1. Decide (recruit) the universe of individuals that would be eligible
2. **Decide unit of randomization (individual, household, village…)**
3. Randomize into treatment(s) and control
4. Test that randomization worked by checking balance on observables
5. Conduct intervention with treatment group(s)

Step 2 - Unit of Randomization

- Choose according to how program is administered and feasibility
  - Individual or Household
  - Street or Neighborhood
  - Block or Village or Community
  - Ward or District or Region

Most Common for EI programs

As a rule of thumb, randomize at the smallest viable unit of implementation.

- Be mindful of spillovers/contamination
  - Outcomes of controls can be affected by treatment
  - Set unit of randomization so treatments and controls are separate
  - Measure spill-overs

Key Steps for RCTs

1. Decide (recruit) the universe of individuals that would be eligible
2. Decide unit of randomization (individual, household, village, group…)
3. **Randomize into treatment(s) and control**
4. Test that randomization worked by checking balance on observables
5. Conduct intervention with treatment group(s)

Step 3 - Randomization Opportunities

1. Budget or Capacity Constraints:
   cannot implement program everywhere at the same time
   - Randomization is fair & transparent
   - Structured targeting prevents nepotism

2. Interest in complementary interventions to maximize impacts
   - Randomize complementary interventions among participants

3. No evidence on which alternative is best
   - Randomize interventions with equal ex-ante chance of success among participants

Step 3 – Randomization Opportunities

- There are opportunities for randomization in almost every project
- In case of economic inclusion interventions, can randomize:
    - Targeting mechanisms to identify vulnerable populations
    - Size, timing and duration of transfer
    - Conditionality required to receive the transfer
    - Complimentary interventions e.g trainings, investment support
- RCTs are doable in high-stakes contexts
    - Teacher incentive schemes
    - Audits of government officials
    - Country-wide cash transfers

Step 3 - Multiple Treatment Experiments

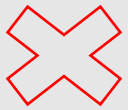
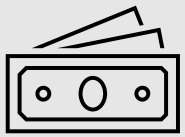
1. How much treatment should we provide?
   - Randomly assign different **treatment intensity** to different treatment groups
   - Example: High Cash vs Low Cash;

2. Which program component should we provide?
   - Randomly assign **different treatments** to different groups
   - Example: Cash vs. In-kind; Training vs. Input Subsidy

3. Are components necessary for impacts?
   - Randomly assign more features to some groups
   - Example: cash + training vs cash only → comparison tells us if training matters

Step 3 - Multiple Treatment Experiments: Example

- Point 3 in previous slide: a very frequent dilemma for policy makers
- Factorial designs isolate 1) contribution of each component; 2) overall effect

	Control	Treatment 2 - Training
Control		
Treatment 1 - Cash		

Key Steps for RCTs

1. Decide (recruit) the universe of individuals that would be eligible
2. Decide unit of randomization (individual, household, village, group…)
3. Randomize into treatment(s) and control
4. **Test that randomization worked by checking balance on observables**
5. Conduct intervention with treatment group(s)

Step 4 – Test Randomization Worked

- Using baseline (pre-intervention) data
  - compare observable characteristics of T and C
  - verify they are similar

Sample: Ultra-Poor Women and their Households

	(1) Treated Villages	(2) Control Villages	(3) t-test [Treatment=Control]	(4) Normalized Differences
A. Labor Market Outcomes				
Hours devoted to livestock rearing (cows/goats)	115	129	.584	-.036
	(258)	(275)		
Earnings from livestock rearing	7.85	8.90	.654	-.013
	(53.2)	(60.4)		
Hours devoted to agricultural labor	269	237	.740	.042
	(537)	(539)		
Hourly wage in agricultural labor	.330	.360	.431	-.195
	(.103)	(.114)		
Hours devoted to domestic maid	325	479	.013	-.152
	(651)	(774)		
Hourly wage in maid services	.256	.261	.823	-.028
	(.107)	(.113)		
Earnings from casual labor	164	191	.340	-.085
	(218)	(239)		
Total earnings	241	289	.172	-.117
	(275)	(300)		
Total days worked in the past year	247	259	.327	-.060
	(141)	(130)		
Average standardized difference (p-value)			.207	
B. Poverty, Expenditures and Financial Wealth				
Below the $1.25 a day poverty line [yes=1]	.556	.584	.524	-.040
	(.400)	(.398)		
Consumption expenditure, per adult equivalent	629	613	.501	.047
	(246)	(236)		
Value of household assets	36	37	.829	-.011
	(48)	(63)		
Household savings	6.2	9.2	.071	-.059
	(28)	(43)		
Household receives loans	.20	.18	.441	-.044
	(.40)	(.38)		
Household gives loans	.011	.014	.356	-.022
	(.10)	(.12)		
Average standardized difference (p-value)			*.849*	
C. Productive Assets				
Cows value	36	30	.575	.023
	(176)	(166)		
Goats value	6.5	8.5	.261	-.050
	(25)	(31)		
Household rents in land [yes=1]	.058	.061	.875	-.007
	(.235)	(.239)		
Household owns land [yes=1]	.068	.062	.738	.017
	(.252)	(.241)		
Value of land owned	175	238	.390	-.027
	(997)	(2190)		
Value of other business assets	23	23	.991	-.0004
	(79)	(101)		

Key Steps for RCTs

1. Decide (recruit) the universe of individuals that would be eligible
2. Decide unit of randomization (individual, household, village, group…)
3. Randomize into treatment(s) and control
4. Test that randomization worked by checking balance on observables
5. **Conduct intervention with treatment group(s)**

Step 5 – Conduct Intervention

- Local staff should be prepared with implementation plans
  - Have list of villages, households, etc to be treated
  - Have timeline on when they should be treated

- Implementation should follow the research design closely

Example of RCT

- A large NGO wants to undertake a cash transfer program for targeted communities

- It has identified 100 villages where
  - Project can be implemented
  - The resident population is vulnerable to potential economic/climate shocks

- The NGO wants to create a rigorous counterfactual
  - To document impacts
  - To improve delivery of intervention

Example of RCT

- Possible constraint 1: NGO has **funds** to reach 50 of 100 eligible villages

- Solution: ?

Example of RCT

- Possible constraint 1: NGO has **funds** to reach 50 of 100 eligible villages

- Solution: RCT where 50 villages are assigned with randomization algorithm to receive program

Example of RCT

- Possible constraint 2: NGO has **staff** to reach only 50 of 100 eligible villages in first year

- Solution: ?

Example of RCT

- Possible constraint 2: NGO has **staff** to reach only 50 of 100 eligible villages in first year

- Solution: RCT where 50 villages are selected at random to receive program in first year, remaining 50 in following years
  - Called a randomized phase-in

- Drawback: can only measure one-year impact. When remaining 50 villages are phased-in, they cease to be a control group.

Example of RCT

- Possible constraint 3: NGO wants to support gender equality and women's empowerment with program, but does not know how to do it

- Solution: ?

Example of RCT

- Possible constraint 3: NGO wants to support gender equality and women's empowerment with program, but does not know how to do it

- Solution: include a business training favoring female-oriented businesses and/or who the cash recipient is
  - If best program is unknown ex-ante, randomization can provide evidence on what works best
  - Opportunity for learning in early phase of project

Caveats to Keep in Mind

1. Sample Size
    - Need a sufficiently large number of units to detect minimum impacts
2. Spillovers/Contamination
    - Outcomes of controls can be affected by treatment
    - Set unit of randomization so treatments and controls are separate
    - Measure spill-overs
3. Operational and Survey Costs
4. General Equilibrium Effects
    - Programs can have impacts on whole economy (e.g. prices)
    - Caution when interpreting results or embed in design (e.g. vary treatment intensity)
5. RCTs do not guarantee external validity
    - Set up similar RCTs in different countries

Takeaways

- Impact evaluations measure causal impacts
- We compare outcomes of a treatment and a control group
  - To find out what would have happened in counterfactual world

- The gold standard for impact evaluations is to **randomize assignment** of potential beneficiaries to treatment and control group

- Many opportunities to randomize: programs, components, intensity

Thank you!

Benedetta Lerva
blerva@worldbank.org

PEI FUNDING PARTNERS

PARTNERSHIP FOR ECONOMIC INCLUSION

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

brac

Co—Impact

german cooperation
DEUTSCHE ZUSAMMENARBEIT

Implemented by

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Irish Aid
Rialtas na hÉireann
Government of Ireland