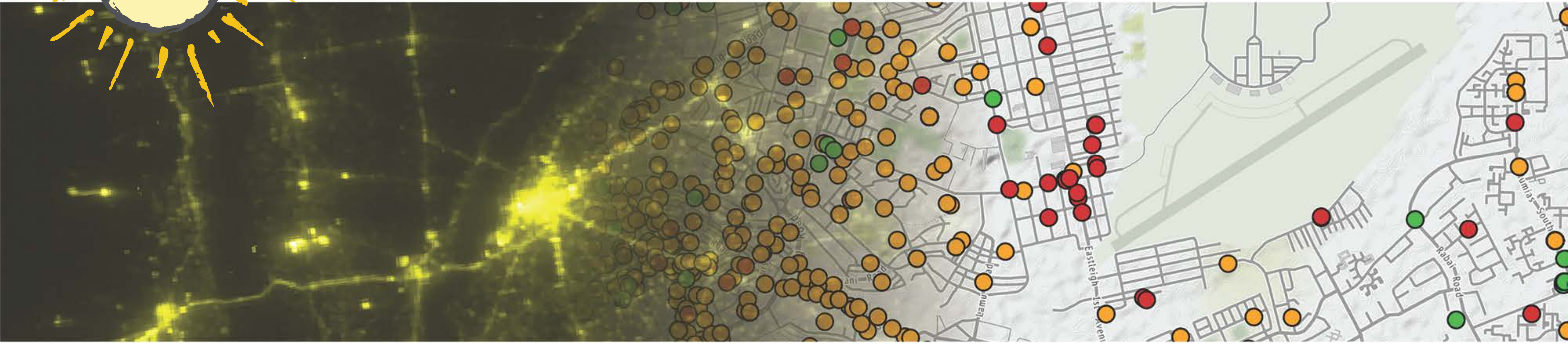
IE CONNECT FOR IMPACT

Transforming the Growth Potential of Transport Investments

Experimental Methods

Maria Jones

4/12/19



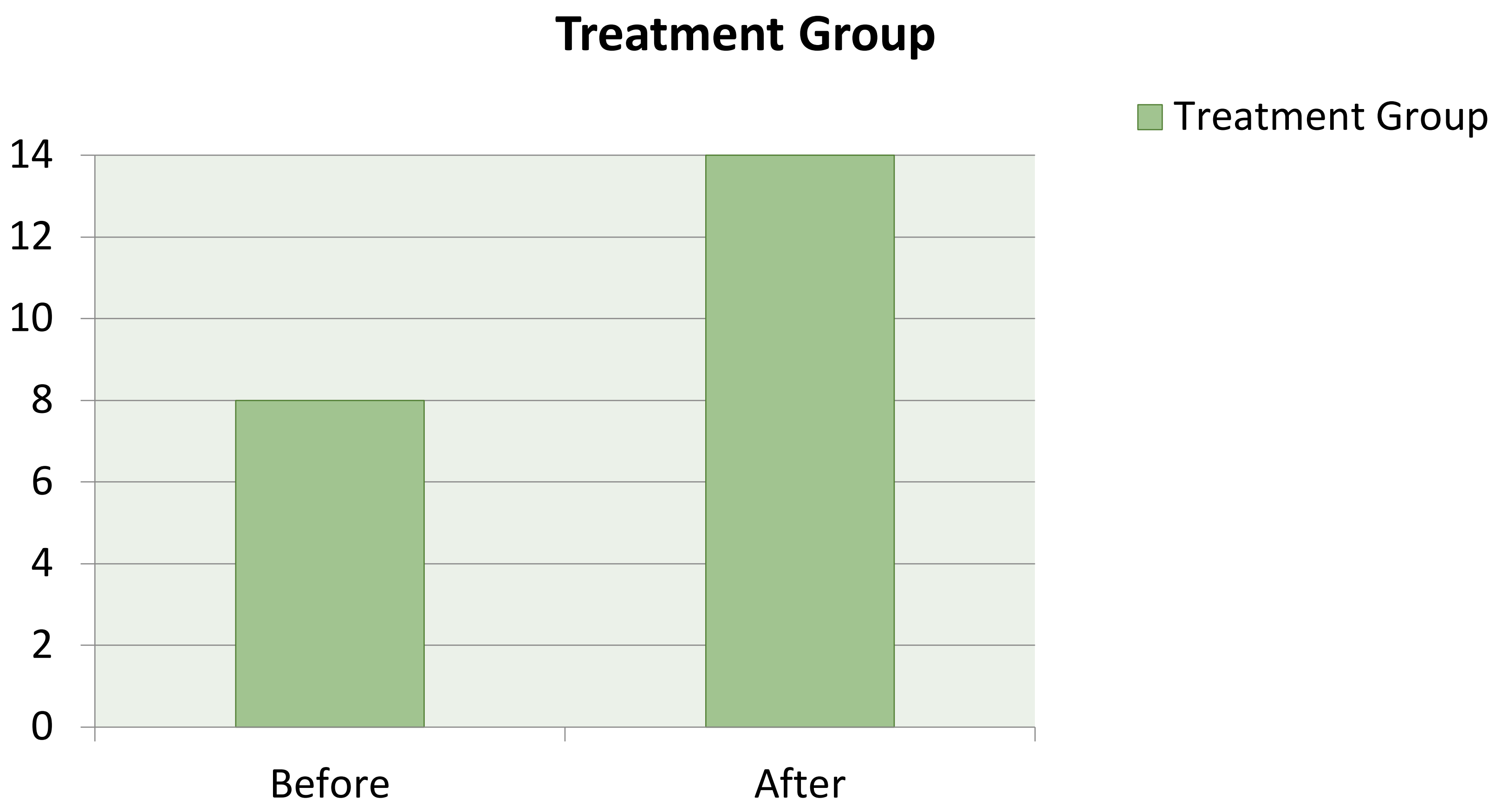
i2i DIME TRANSFORM DEVELOPMENT
WORLD BANK GROUP Transport
SUSTAINABLE mobility FOR ALL
UKaid from the British people

“for their experimental approach to alleviating global poverty”

A preview

- Impact Evaluation measures **causal impact**
- To find out what would have happened without the project, or with an alternative project, we must compare outcomes to a control group
- **Randomly** selecting treatment and control groups is the "gold standard" for impact evaluation
- Opportunities for randomization abound, even in large infrastructure projects!

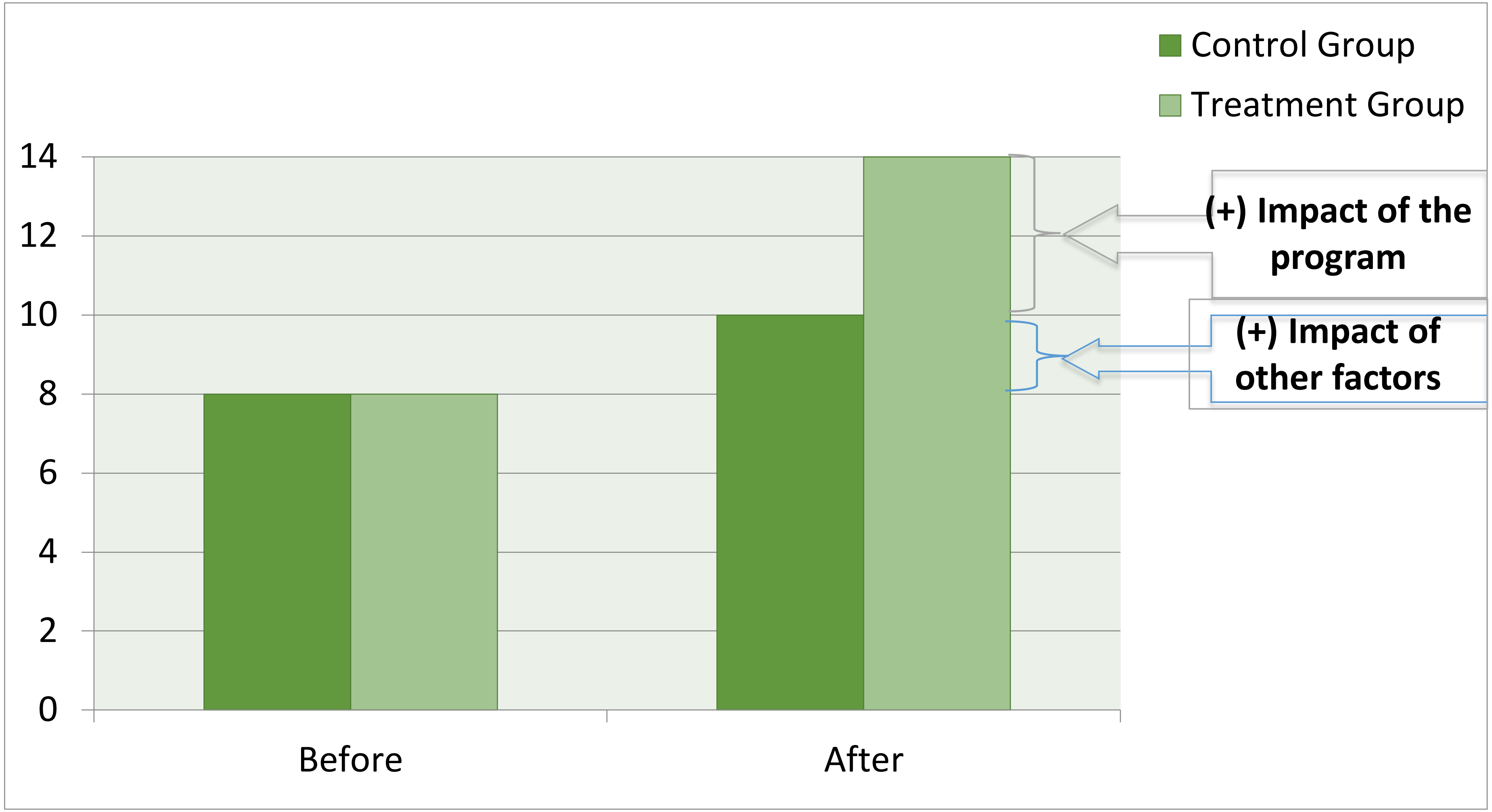
Is this the impact of a project?



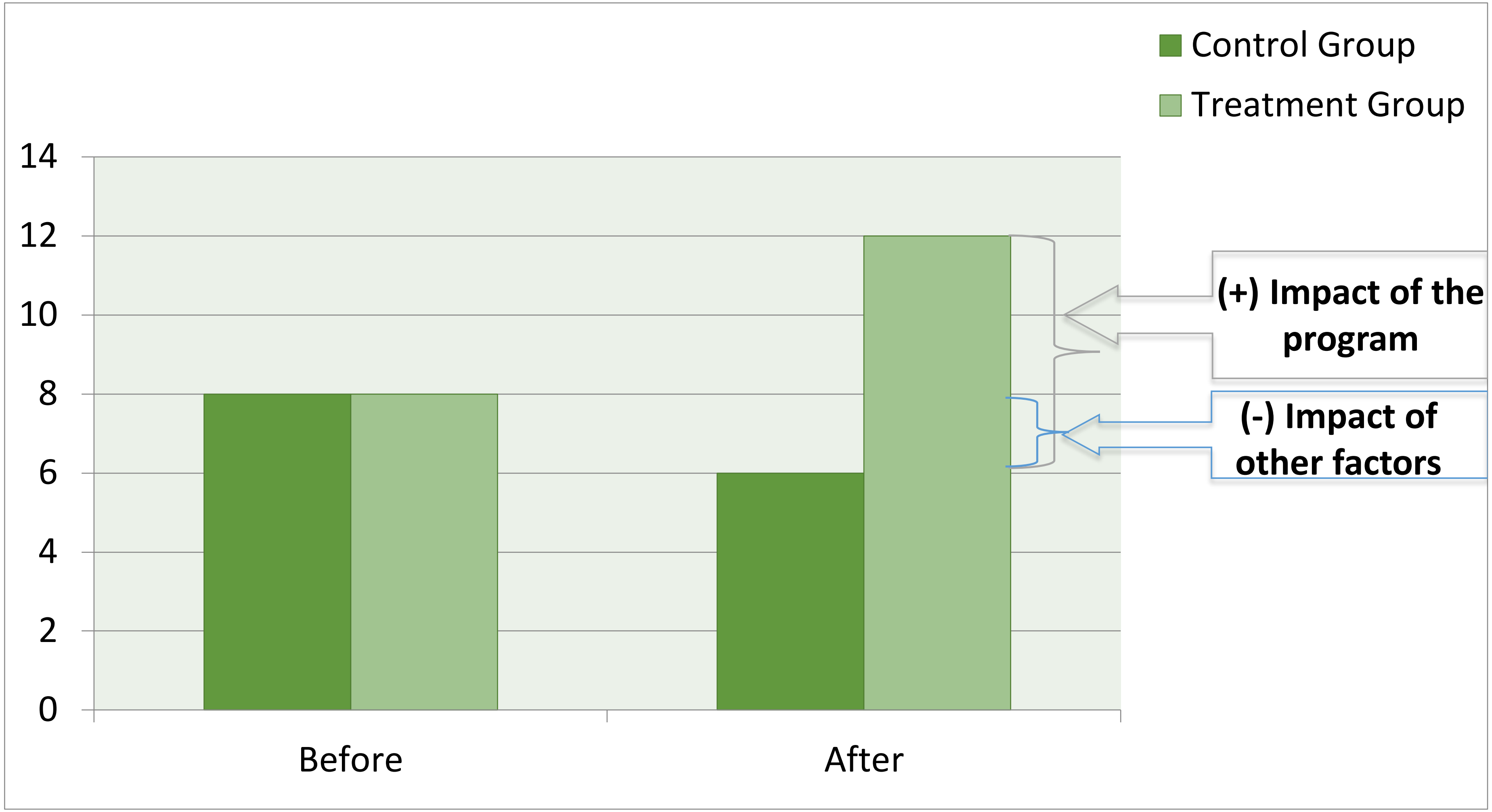
Counterfactual analysis

Ideal	Compare same individual with & without project	This is impossible!
Instead	Identify a control group that is similar in all ways except project participation	Measure treatment group outcomes and compare to control group to evaluate project impact

The importance of a control group



The importance of a control group



What makes a good control?

- Treatment and control need to be as similar as possible

- Can we compare people who received the project to anyone who didn’t receive the project?
  - What was the reason that some people received it and others didn’t?

- Selection bias is important to understand
  - Projects started at specific times and places for particular reasons
  - Participants may self-select into programs

Randomization is gold standard

- Randomly assign potential beneficiaries to treatment or control group
- By design, treatment and control have the same characteristics (observed and unobserved), on average
  - With large enough sample, all characteristics average out
- Only difference is treatment, so impact estimates are unbiased

Case study

- Republic of Atlantis wants to upgrade rural roads to reduce transport costs and promote local economic development
- IE Question: Does road upgrading lead to an increase in the households' per capita consumption?
- Is randomization possible? What could it look like?

Case study: possibility 1

- Rural Revitalization Program
  - 2000 villages invited to express interest
  - 1021 village expressed interest in first phase

- Compare: 1021 villages to 979 that did not express interest

- Will this give us good (unbiased) impact estimates?

Case study: possibility 1

- **No!**

- Are the 1021 villages that expressed interest in the project early likely similar to those that did not?

- Selection bias - differences are likely both observable and *non-observable*

- Can only control for observable differences

Treatment and control group balance

	TREATMENT	COMPARISON	DIFFERENCE
Number of users	44.26	31.83	12.43*
Pop. density	111.90	109.46	2.44*
Local market [1= Yes]	0.86	0.85	0.01
Number of children per HH	4.83	5.27	-0.44*
Diversification (%)	25.90	25.33	0.57
Sample size	1021	979	

- Village characteristics are significantly different at baseline
- Treatment and control are not "balanced"

Case study: possibility 2

- Rural Revitalization Program
  - 4000 villages identified as having high priority roads
  - Funding sufficient to upgrade roads for 2000 villages now

- Conduct a lottery to select 2000 for upgrading now. The remaining 2000 will be upgraded in the future

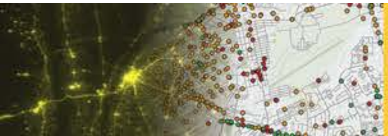
- Will this give us good (unbiased) impact estimates?

Case study: possibility 2

- **Yes!**

- Only reason for assignment to one of the groups is random chance
  → we are confident treatment and control villages are similar in all observable and non-observable characteristics

- Only difference between groups is that one group receives upgraded road, the other does not
  → Comparison of treatment and control group gives a causal impact



Treatment and control group balance

	TREATMENT	COMPARISON	DIFFERENCE
Number of users	37.85	38.27	0.42
Pop. density	10.20	10.89	-0.69
Local market [1= Yes]	0.34	0.37	-0.03
Number of children per HH	5.05	5.03	0.02
Diversification (%)	25.51	25.67	-0.16
Sample size	2,000	2,000	

- Village characteristics are "balanced" at baseline
  - difference between control and treatment is not statistically significantly different from zero before the intervention starts
- This happens because of randomization as long as the sample is "large enough"

Can I randomize??

- Randomizing infrastructure placement might be not feasible
  - One (or a few) big intervention – corridor, port
  - Road networks, zoning laws
- But, possible to randomize complementary interventions
  - Department of Social and Economic Affairs interested in promoting income diversification for rural women working in agriculture
  - Can transport subsidy help women take advantage of the improved roads to market their produce?

Case study: possibility 3

- IE Question: does a transport subsidy help women market their produce?

- In phase 1 villages, DSEA randomly selects 900 women in villages of the phase 1 program. The 900 women are randomly dived into 2 groups: 450 women receive a bike, 450 women do not.

- Will this give us good (unbiased) impact estimates?

- Yes!

Treatment and control group balance

	TREATMENT	COMPARISON	DIFFERENCE
Number of users	38.14	38.08	0.06
Pop. density	10.52	10.70	-0.18
Local market [1= Yes]	0.34	0.37	-0.03
Number of children per HH	5.02	4.90	0.12
Diversification (%)	24.64	25.20	-0.56
Sample size	450	450	

- Village characteristics are "balanced" at baseline
  - difference between control and treatment is not statistically significantly different from zero before the intervention starts
- This happens because of randomization as long as the sample is "large enough"

Multiple Treatments

- Not sure how much to treat? Test different treatment intensity
  - Randomly assign participants to different levels of treatment (e.g. different levels of subsidy)

- No evidence on which alternative is best? Test variations in treatment
  - Randomly assign subjects to different interventions
  - Compare one to another
  - Assess complementarities

Multiple Treatments

2x2 (factorial design) to assess complementarities and overall effect

		Intervention 2	
		Control	Treatment
Intervention 1	Control	*x*	Road upgrade
	Treatment	Transport subsidy	Both programs

Opportunities to randomize

- Budget constraints → prevent full coverage (at one point in time)
  - Random assignment (lottery) is fair and transparent
- Interest in complementary interventions to maximize impact of infrastructure investments
  - Randomize interventions among project participants
- No evidence on which alternative is best
  - Randomize allocation of treatments with equal ex ante chances of success

Unit of Randomization

- Choose according to type of program
  - Individual/Household
  - Street/Neighborhood
  - Block/Village/Community
  - Ward/District/Region

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

- Keep in mind
  - Need "sufficiently large" number of units to detect minimum desired impact
  - Spillovers/contamination
  - Operational and survey costs

Takeaways

- Impact Evaluation measures **causal impact**
- To find out what would have happened without the project, or with an alternative project, we must compare outcomes to a control group
- **Randomly** selecting treatment and control groups is the “gold standard” for impact evaluation
- Opportunities for randomization abound, even in large infrastructure projects!

Thank you!

Questions?