Abdul Latif Jameel Poverty Action Lab Executive Training: Evaluating Social Programs
Spring 2009

Managing threats
to evaluation and data analysis

Abdul Latif Jameel
Poverty Action Lab

A good evaluation strategy is necessary…. but not sufficient

- Even if you have managed to set up a proper randomized trial, there can still be problems with impact measurement and analysis

- Today we will study what these problems can be and how to avoid them

Outline

I. Attrition

II. Externalities

III. Partial Compliance and Sample Selection Bias

Outline

I. **Attrition**
“Is it a problem if some of the people in the experiment vanish before you collect your data?”

II. Externalities

III. Partial Compliance and Sample Selection Bias

Attrition bias

- It is a problem if the type of people who disappear is correlated with the treatment.
- Why?
- This is called *attrition bias*.
- Why should we expect this to happen?

Attrition bias: an example

- The problem you want to address:
  - Some children don’t come to school because they are too weak (undernourished)
- You start a school feeding program and want to do an evaluation
  - You have a treatment and a control group
- Weak, stunted children start going to school more if they live next to a treatment school
- Want to know program impacts:
  - increased enrollment.
  - weight of children
- You go to all the schools (treatment and control) and measure everyone who is in school on a given day
- Will the treatment-control difference in weight be over-stated or understated?

	Before Treatment		After Treament	
	T	**C**	**T**	**C**
	30	30	**32**	30
	35	35	**37**	35
	40	40	**42**	40
Ave.				
	Difference		Difference	

What if only children > 30 kg come to school?

	Before Treatment		After Treament	
	T	**C**	**T**	**C**
	[absent]	[absent]	**32**	[absent]
	35	35	**37**	35
	40	40	**42**	40
Ave.				
	Difference		Difference	

Attrition bias: another example

- Suppose the treatment is a harder math course.
- Those who cannot handle it drop out of school.
- You give the same math test in treatment and control schools
- You only have data on those who have not dropped out.
- What is the direction of the bias?

Attrition bias

- What source of attrition bias did we worry about in the de-worming program with regards to testing?

- How did we deal with it?

- Always the same answer: make sure that no one drops out from your original treatment and control groups.
-
- Pick a sample of those who will be tested before the treatment and follow them (no matter where they go!)

Attrition bias

- If there is still attrition, check that it is not different in treatment and control. Is that enough?

- Also check that it is not correlated with observables.

- Try to bound the extent of the bias
  - Suppose everyone who dropped out from the treatment got the lowest score that anyone got
  - OR suppose everyone who dropped out of control got the highest score that anyone go.
  - Why does this help?

Outline

I. Attrition

II. Externalities

"What if the treatment affects those not directly treated?"

- **Spillovers**

III. Partial Compliance and Sample Selection Bias

Example: Deworming

Previous studies randomize deworming treatment <u>within</u> schools

- Suppose that deworming prevents the transmission of disease, what problems does this create for evaluation?
- Suppose externalities are local? How can we measure total impact?

Externalities Within School

Without Externalities

School A	Treated?	Outcome
Pupil 1	Yes	no worms
Pupil 2	No	worms
Pupil 3	Yes	no worms
Pupil 4	No	worms
Pupil 5	Yes	no worms
Pupil 6	No	worms

Total in Treatment with Worms	
Total in Control with Worms	
Treament Effect	

With Externalities

Suppose, because prevalance is lower, some children are not re-infected with worms

School A	Treated?	Outcome
Pupil 1	Yes	no worms
Pupil 2	No	no worms
Pupil 3	Yes	no worms
Pupil 4	No	worms
Pupil 5	Yes	no worms
Pupil 6	No	worms

Total in Treatment with Worms	
Total in Control with Worms	
Treatment Effect	

How to measure program impact in the presence of spillovers?

- Design the unit of randomization so that it encompasses the spillovers

- If we expect externalities that are all *within* school:
  - Randomization at the level of the school allows for estimation of the overall effect
    - Example: Deworming of all the children in a school

Measuring total impact in the presence of spillovers

GROUP	TREATMENT SCHOOLS	CONTROL SCHOOLS	PROGAM EFFECT
% of children with a moderate or heavy infection	27%	52%	-25%***
% of children who were sick the week before the survey	41%	45%	-4%**
% of children who are anemic	2%	4%	-2%*

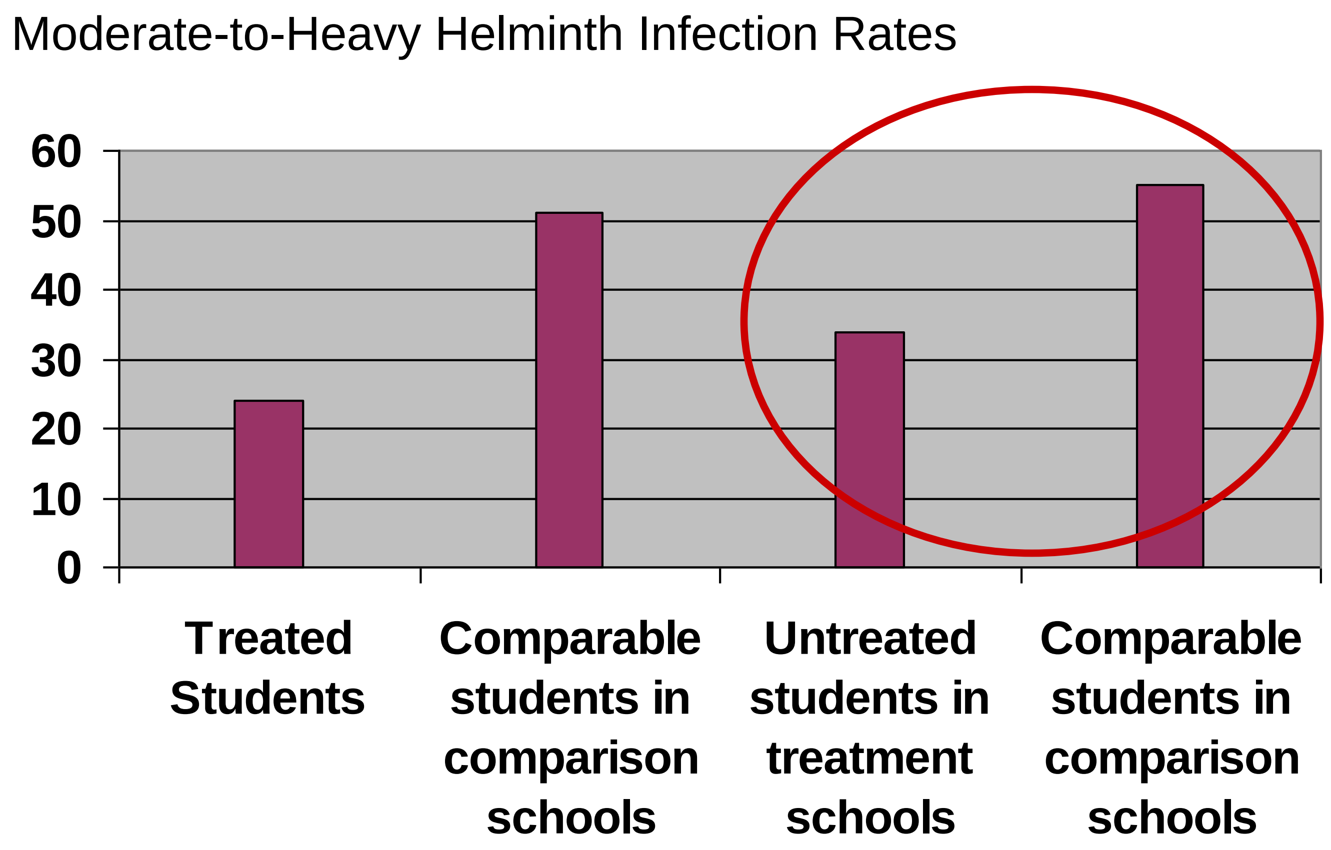
Within-school health externalities

- What if we wanted to measure the spillovers?

- Deworming study
  - Because girls above 12 could not be treated in the treatment schools, we can compare girls above 12 in treatment schools to girls above 12 in comparison schools.

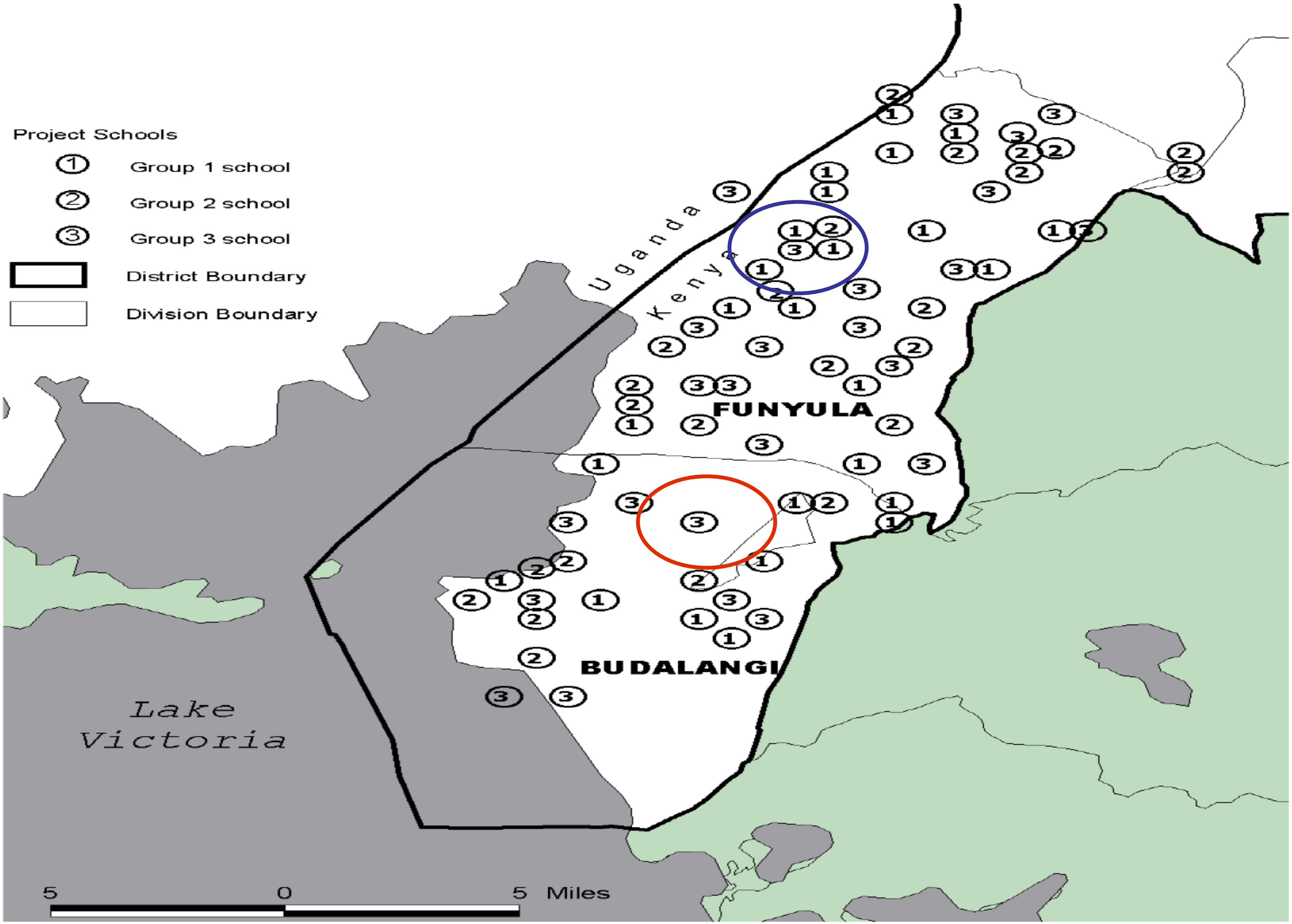
- More generally: need to randomize treatment within the unit so as to be able to learn about spillovers.

Deworming treatment effects on health: spillovers



Measuring externalities

- If we expect externalities across schools or outside of schools
  - Need to use *random* variation in density of treatment



Project Schools
Group 1 school
Group 2 school
Group 3 school
District Boundary
Division Boundary
Uganda
Kenya
FUNYULA
BUDALANGI
Lake Victoria
5
0
5 Miles

Cross-school externalities

- Can you look at the number of treatment schools nearby?
- What is potential problem?
- What do we need to control for?

Results: controlling for density

- Infection rates 26 percentage points lower per 1000 pupils in treatment schools within 3 km

- Infection rates 14 percentage points lower per 1000 pupils in treatment schools between 3-6 km away

Estimating the overall effect

- Comparison schools:
  - 1.5 percentage point increase in school participation
  - 3 pupils in control schools for every treated child

- Treatment schools:
  - 7 percentage point increase in school participation for all children
  - 1 untreated child for every 2 treated children

- Overall effect of treating one child:
  - (.015 * 3) + (.07 * .5) + (.07 * 1) = .15 years
  - For each child treated, school participation increased by .15 years.

Outline

I. Attrition

II. Externalities

III. Partial Compliance and Sample Selection Bias

“Does randomization always guarantee that there is no sample selection bias?”

Sample selection bias

- *Sample selection bias* could arise if factors other than random assignment influence program allocation
  - Even if intended allocation of program was random, the actual allocation may not be
  - *Why?*

Sample selection bias

- Individuals assigned to comparison group could attempt to move into treatment group
  - *De-worming program*: parents could attempt to move their children from comparison school to treatment school
- Alternatively, individuals allocated to treatment group may not receive treatment
  - *De-worming program*: some students assigned to treatment in treatment schools did not receive medical treatment

Sample selection bias

- Some students in treatment schools not treated
  - 1998: 78% of pupils assigned to receive treatment received at least some treatment
  - 1999: around 72%
  - Absence from school the major cause of non-compliance
- Some students in comparison schools treated
  - 5% received treatment outside of program
- *What do you do?*

Sample selection bias

- Use the original assignment
  - If a child ended up in a treatment school but was from the control, she should be assigned to control when calculating the effect.

- This gives us the Intention to Treat estimate (ITT).

Intention to treat (ITT)

- What does “intention to treat” measure?

  “What happened to the average child who is in a treated school in this population?”

- Is this the right number to look for?

- Remember: In the deworming case many children in treatment schools were not treated and some children in comparison schools were.

ITT Estimator

School 1	Intention to Treat ?	Treated?	Observed Change in weight
Pupil 1	yes	yes	4
Pupil 2	yes	yes	4
Pupil 3	yes	yes	4
Pupil 4	yes	no	0
Pupil 5	yes	yes	4
Pupil 6	yes	no	2
Pupil 7	yes	no	0
Pupil 8	yes	yes	6
Pupil 9	yes	yes	6
Pupil 10	yes	no	0
		Avg. Change	

School 2			
Pupil 1	no	no	2
Pupil 2	no	no	1
Pupil 3	no	yes	3
Pupil 4	no	no	0
Pupil 5	no	no	0
Pupil 6	no	yes	3
Pupil 7	no	no	0
Pupil 8	no	no	0
Pupil 9	no	no	0
Pupil 10	no	no	0
		Avg. Change	

Avg. Change School 1	
Avg. Change School 2	
==>ITT Effect is	

When is ITT useful?

- May relate more to actual programs

- For example, we may not be interested in the medical effect of deworming treatment, but what would happen under an actual deworming program.

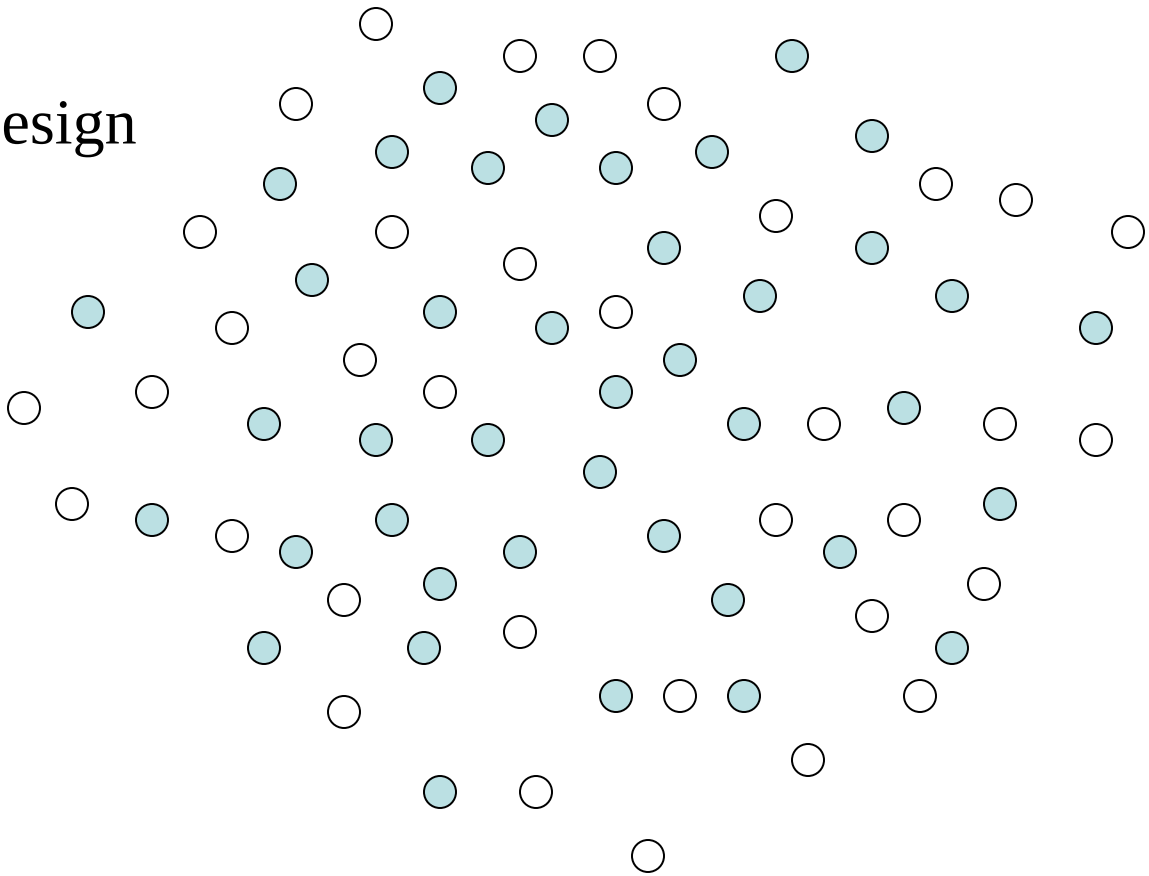
- If students often miss school and therefore don't get the deworming medicine, the intention to treat estimate may actually be most relevant.

ITT: Another example

- You are starting a malaria prevention program
- You sample 40 villages for the pilot study
  - 20 villages randomly assigned to receive the treatment in the first year
  - Remaining 20 villages will be the “comparison” during the pilot and will receive the treatment later if it works.
- Some of the villages that are “comparison” are unhappy. Their leaders talk to your program manager and repeatedly ask him to treat their village.
- The program manager cannot resist the pressure; in the end he cannot fully respect the initial design:
  - He implements the program in only 15 of the 20 villages you selected
  - And also in 2 villages that were in the “comparison”, and in 3 villages out of your sample
- What do you do to measure the impact of your program?

ITT: Another example

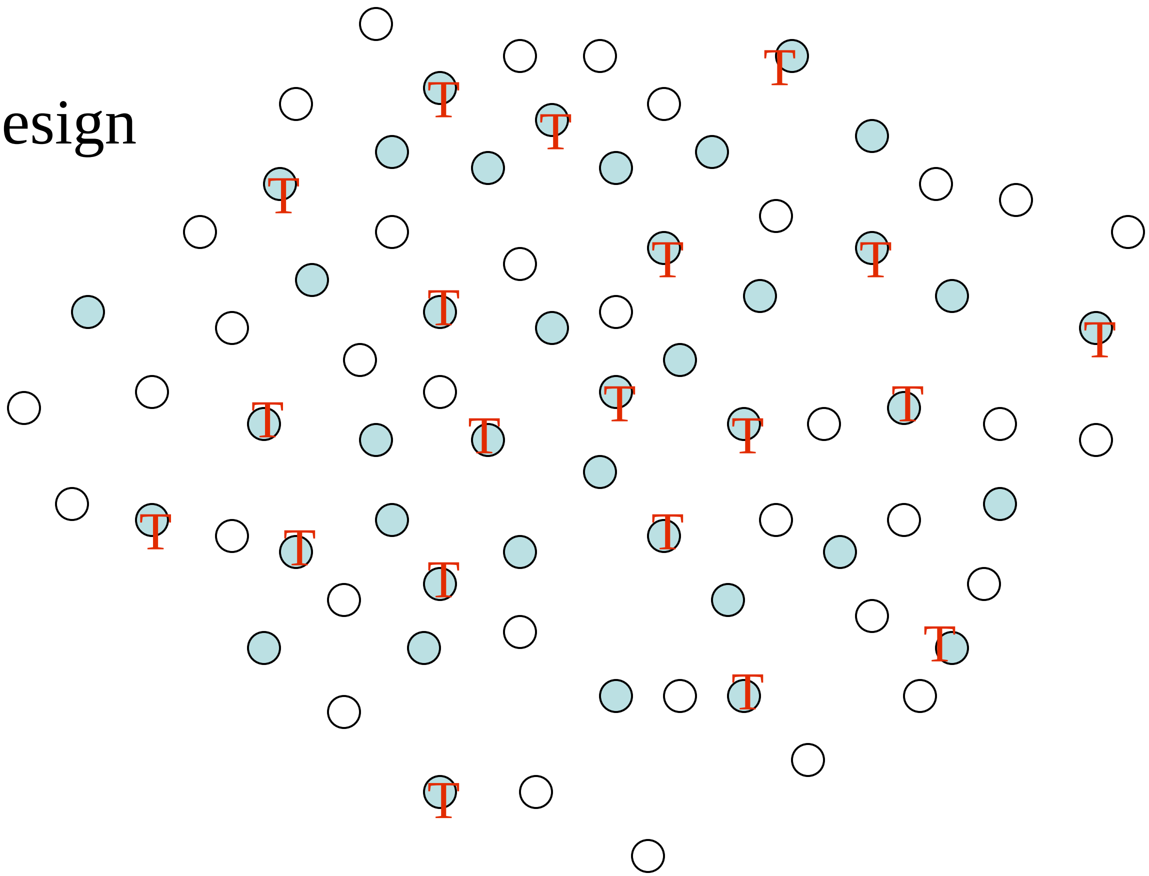
Initial Design



- Your sample
- Other villages

ITT: Another example

Initial Design



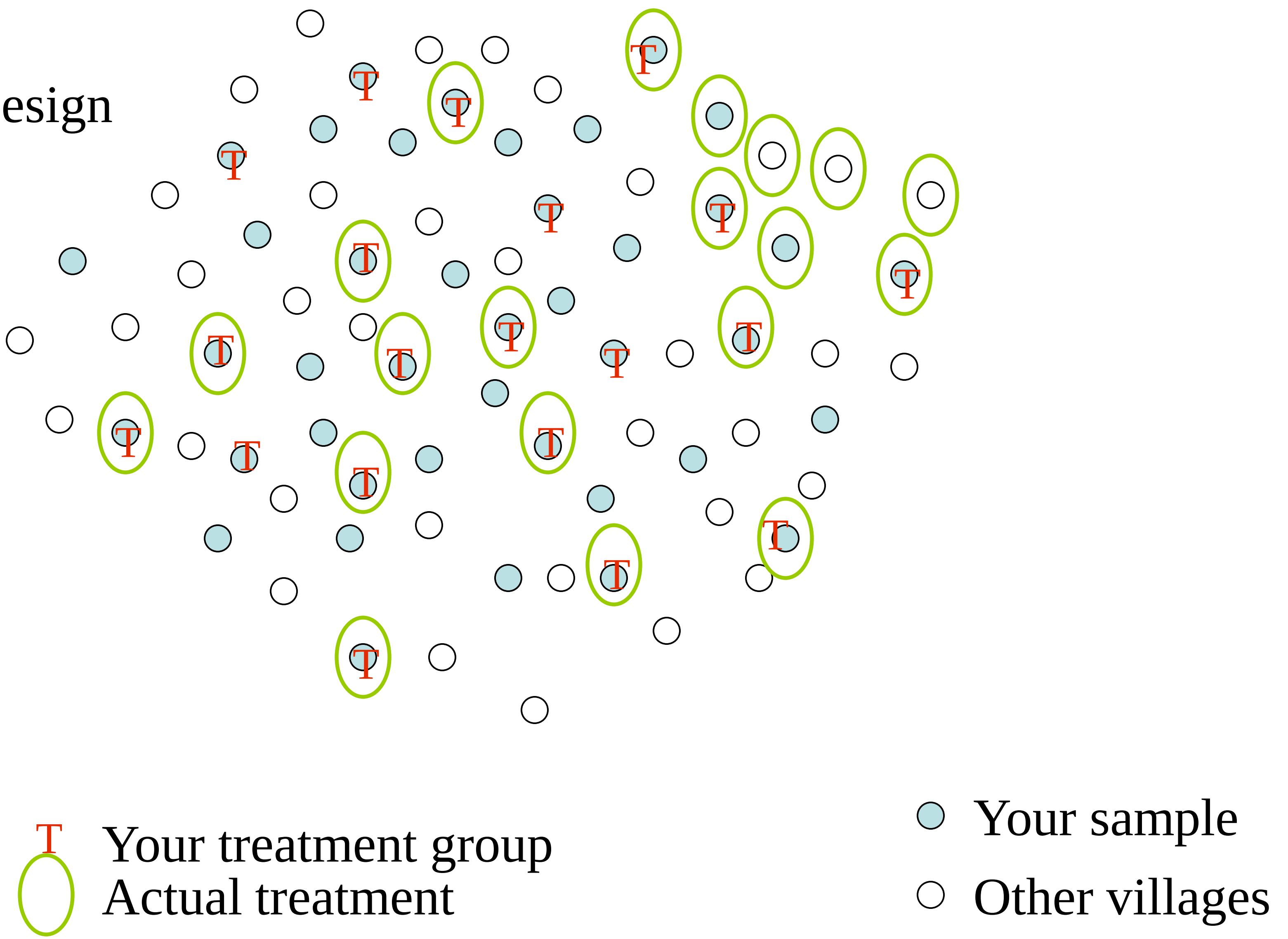
T Your treatment group

Your sample

Other villages

ITT: Another example

Initial Design



How do you measure your impact?

- You cannot compare the villages with the villages.

- Why?
  - Because the villages that should not have been treated but were treated are not randomly selected
  - They are villages with particularly motivated / vocal leaders. They are likely to have better outcomes in any case

How do you measure your impact?

- You cannot compare the villages with the villages.
- Why?
  - Because the villages that should have been treated but were NOT treated are not randomly selected
  - They are villages with particularly uncaring leaders (they didn’t do anything when they were dropped from the treatment list). They are likely to have worse outcomes in any case

How do you measure your impact?

- Respect the initial assignment!
- You should compare the initial 20 treatment villages with the initial 20 comparison villages
  - Even if some of the treatment villages were not treated
  - Even if some of the comparison villages were treated
  - Ignore the villages outside of your initial sample
- That is, compare [T + T] with [+]
- This is the ITT estimator: “Intention to Treat”