ABDUL LATIF JAMEEL
Poverty Action Lab

TRANSLATING RESEARCH INTO ACTION

# Threats and Analysis

**Bruno Crépon**

**J-PAL**

# Course Overview

1. What is Evaluation?
2. Outcomes, Impact, and Indicators
3. Why Randomize and Common Critiques
4. How to Randomize
5. Sampling and Sample Size

7. Project from Start to Finish
8. Cost-Effectiveness Analysis and Scaling Up

# Lecture Overview

A. Attrition
B. Spillovers
C. Partial Compliance and Sample Selection Bias
D. Intention to Treat & Treatment on Treated
E. Choice of outcomes
F. External validity
G. Conclusion

# Lecture Overview

A. Attrition
B. Spillovers
C. Partial Compliance and Sample Selection Bias
D. Intention to Treat & Treatment on Treated
E. Choice of outcomes
F. External validity
G. Conclusion

# Attrition

A. Is it a problem if some of the people in the experiment vanish before you collect your data?
   A. It is a problem if the type of people who disappear is correlated with the treatment.
B. Why is it a problem?
   A. Loose the key property of RCT: two identical populations
C. Why should we expect this to happen?
   A. Treatment may change incentives to participate in the survey

# Attrition bias: an example

A. The problem you want to address:
    A. Some children don't come to school because they are too weak (undernourished)
B. You start a school feeding program and want to do an evaluation
    A. You have a treatment and a control group
C. Weak, stunted children start going to school more if they live next to a treatment school
D. First impact of your program: increased enrollment.
E. In addition, you want to measure the impact on child's growth
    A. Second outcome of interest: Weight of children
F. You go to all the schools (treatment and control) and measure everyone who is in school on a given day
G. Will the treatment-control difference in weight be over-stated or understated?

| | **Before Treatment** | |
|---|---|---|
| | **T** | **C** |
| | **20** | 20 |
| | **25** | 25 |
| | **30** | 30 |
| Ave. | | |
| Difference | | |

| | **After Treament** | |
|---|---|---|
| | **T** | **C** |
| | **22** | 20 |
| | **27** | 25 |
| | **32** | 30 |
| | | |
| Difference | | |

| | Before Treatment | |
|---|---|---|
| | **T** | **C** |
| | **20** | 20 |
| | **25** | 25 |
| | **30** | 30 |
| Ave. | **25** | **25** |
| | Difference | 0 |

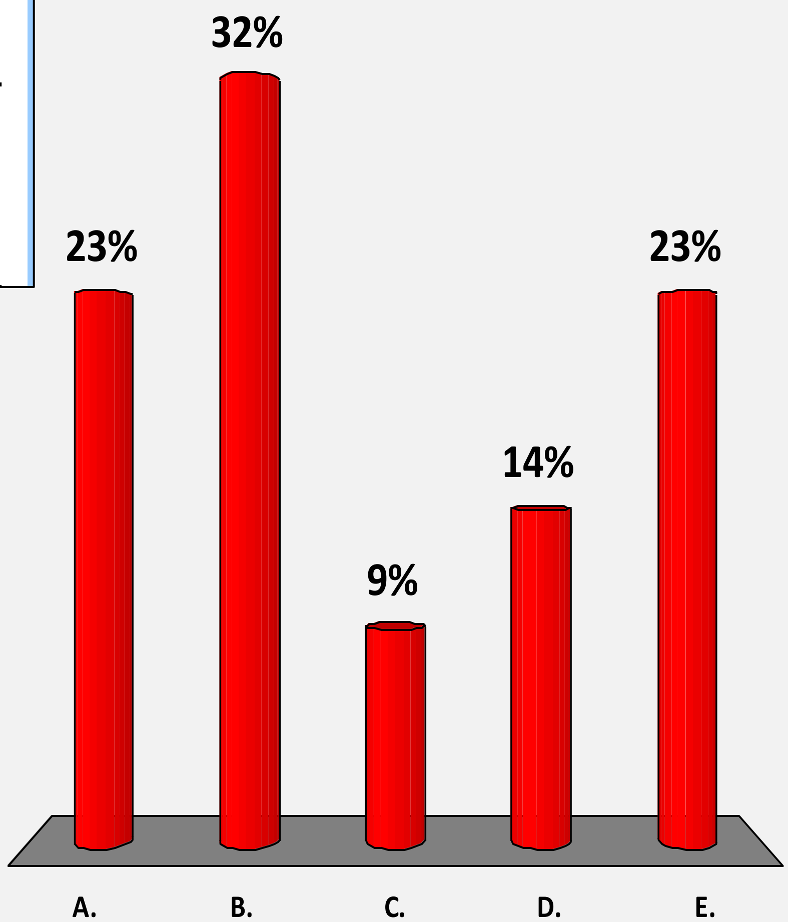
| | After Treament | |
|---|---|---|
| | **T** | **C** |
| | **22** | 20 |
| | **27** | 25 |
| | **32** | 30 |
| | **27** | **25** |
| | Difference | 2 |

# What if only children > 21 Kg come to school?

# What if only children > 21 Kg come to school?

| Before Treatment | | | After Treament | |
|---|---|---|---|---|
| T | C | | T | C |
| **20** | 20 | | **22** | 20 |
| **25** | 25 | | **27** | 25 |
| **30** | 30 | | **32** | 30 |

A. Will you underestimate the impact?
B. Will you overestimate the impact?
C. Neither
D. Ambiguous
E. Don't know

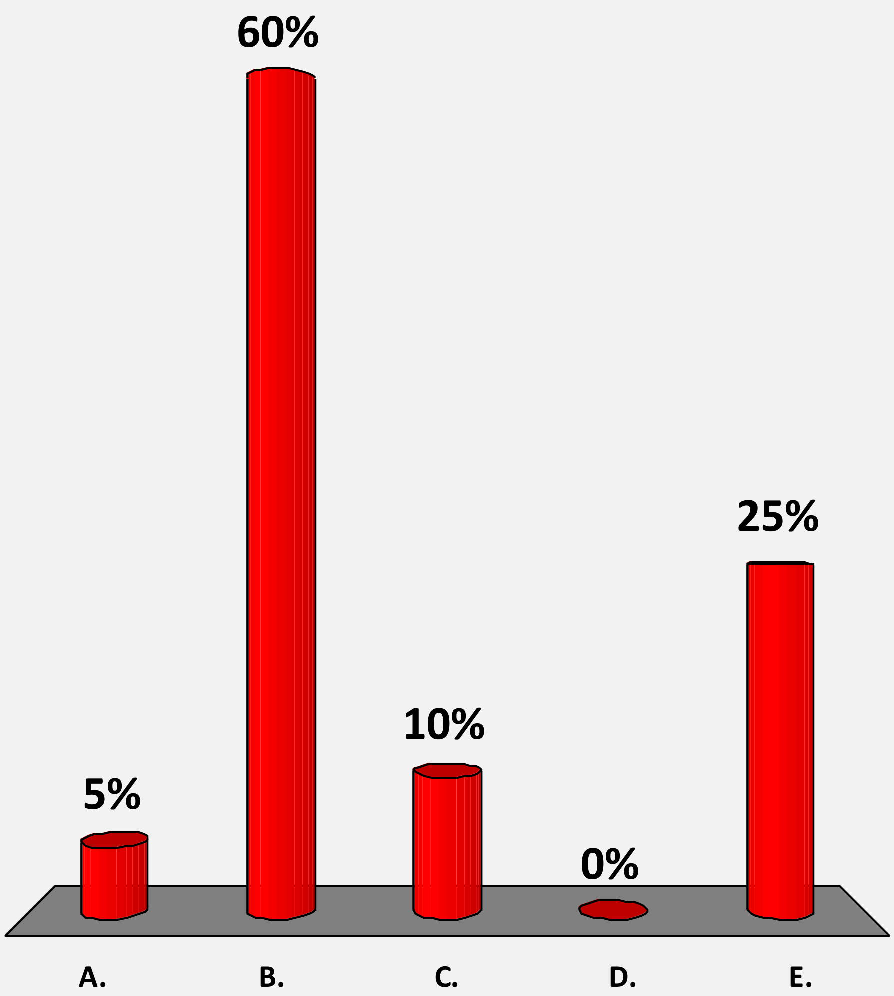


# What if only children > 21 Kg come to school absent the program?

| | **Before Treatment** | | **After Treament** | |
|---|---|---|---|---|
| | **T** | **C** | **T** | **C** |
| | [absent] | [absent] | **22** | [absent] |
| | **25** | 25 | **27** | 25 |
| | **30** | 30 | **32** | 30 |
| Ave. | **27,5** | **27,5** | **27** | **27,5** |
| | Difference | 0 | Difference | -0,5 |

# When is attrition not a problem?

A. When it is less than 25% of the original sample
B. When it happens in the same proportion in both groups
C. When it is correlated with treatment assignment
D. All of the above
E. None of the above



# Attrition Bias

A. Devote resources to tracking participants in the experiment
B. If there is still attrition, check that it is not different in treatment and control. Is that enough?
C. Good indication about validity of the first order property of the RCT:
   A. Compare outcomes of two populations that only differ because one of them receive the program
D. **Internal validity**

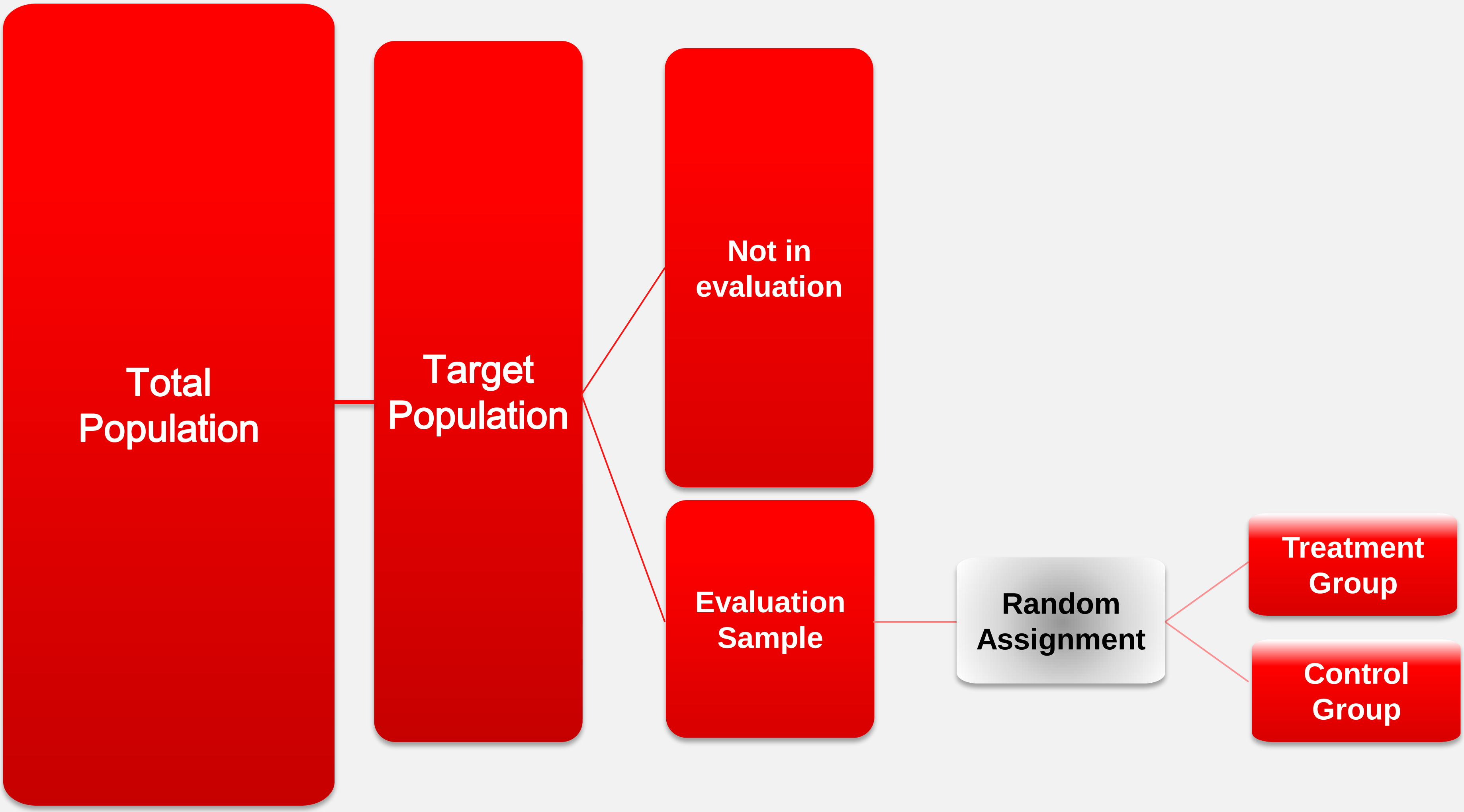
# Attrition Bias

A. If there is attrition but with the same response rate between test and control groups. Is this a problem?
B. It can
C. Assume only 50% of people in the test group and 50% in the control group answered the survey
D. The comparison you are doing is a relevant parameter of the impact but… **on the population of respondent**
E. But what about the population of non respondent
    A. You know nothing!
    B. Program impact can be very large on them,… or zero,… or negative!
F. **External validity** might be at risk

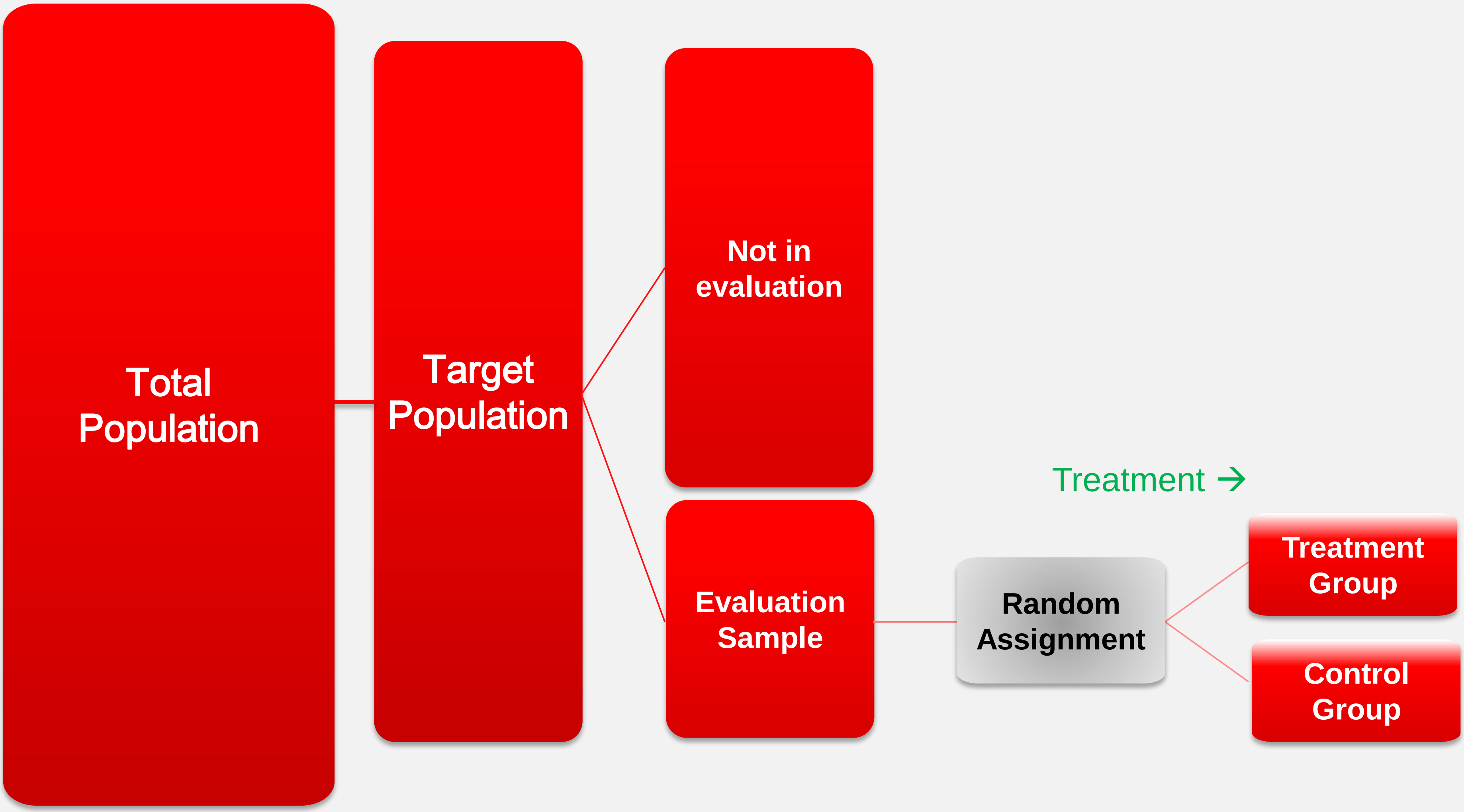
# Lecture Overview

A. Attrition
B. Spillovers
C. Partial Compliance and Sample Selection Bias
D. Intention to Treat & Treatment on Treated
E. Choice of outcomes
F. External validity
G. Conclusion

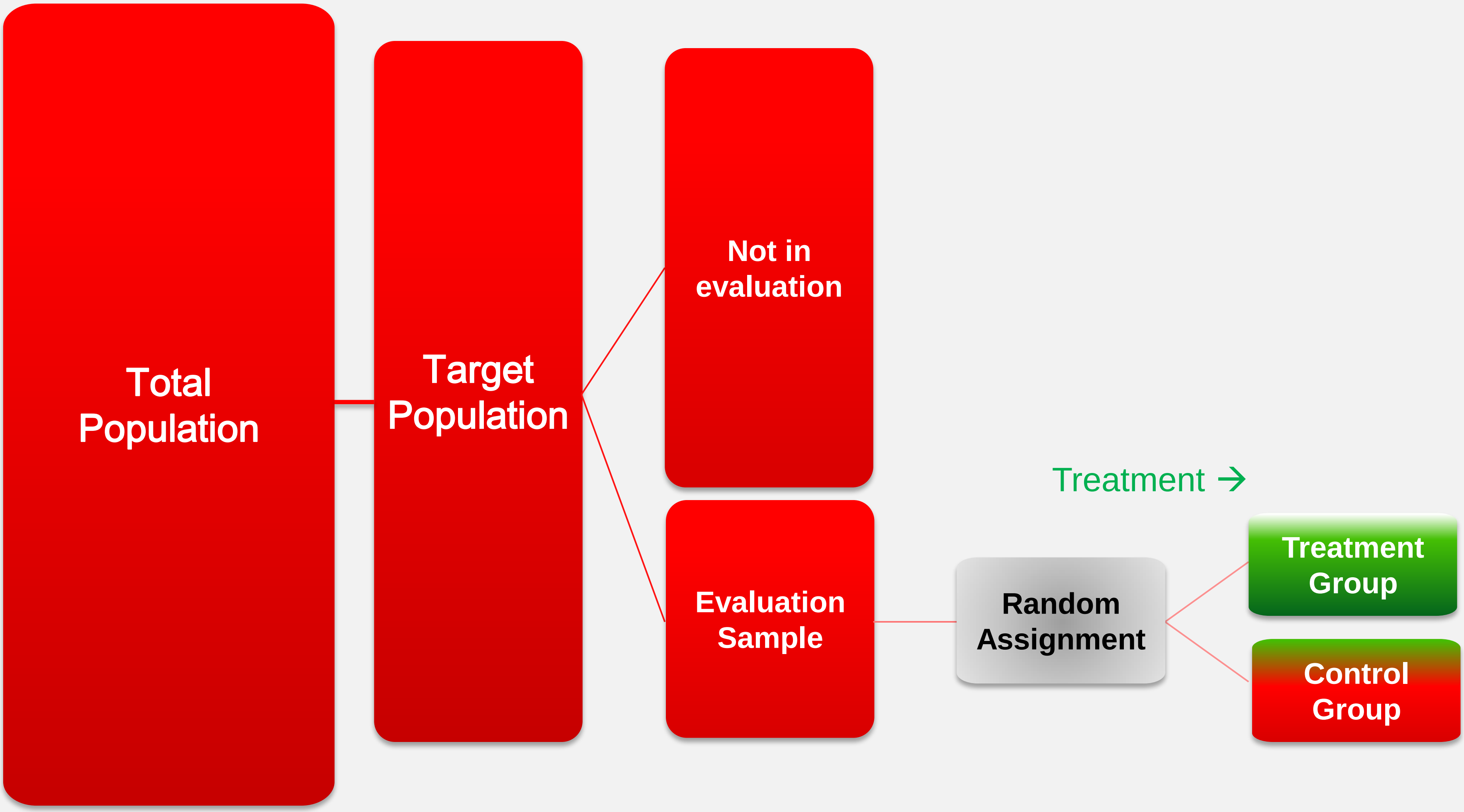
# What else could go wrong?



# Spillovers, contamination



# Spillovers, contamination



# Example: Vaccination for chicken pox

A. Suppose you randomize chicken pox vaccinations *within* schools
   A. Suppose that prevents the transmission of disease, what problems does this create for evaluation?
   B. Suppose externalities are local? How can we measure total impact?

# Externalities Within School

### Without Externalities

| School A | Treated? | Outcome |
|---|---|---|
| **Pupil 1** | Yes | no chicken pox |
| Pupil 2 | No | chicken pox |
| **Pupil 3** | Yes | no chicken pox |
| Pupil 4 | No | chicken pox |
| **Pupil 5** | Yes | no chicken pox |
| Pupil 6 | No | chicken pox |

Total in Treatment with chicken pox

Total in Control with chicken pox

**Treament Effect**

### With Externalities

Suppose, because prevalence is lower, some children are not re-infected with chicken pox

| School A | Treated? | Outcome |
|---|---|---|
| **Pupil 1** | Yes | no chicken pox |
| Pupil 2 | No | no chicken pox |
| **Pupil 3** | Yes | no chicken pox |
| Pupil 4 | No | chicken pox |
| **Pupil 5** | Yes | no chicken pox |
| Pupil 6 | No | chicken pox |

Total in Treatment with chicken pox

Total in Control with chicken pox

**Treatment Effect**

# Externalities Within School

## Without Externalities

| School A | Treated? | Outcome |
|---|---|---|
| **Pupil 1** | Yes | no chicken pox |
| Pupil 2 | No | chicken pox |
| **Pupil 3** | Yes | no chicken pox |
| Pupil 4 | No | chicken pox |
| **Pupil 5** | Yes | no chicken pox |
| Pupil 6 | No | chicken pox |

| | |
|---|---|
| Total in Treatment with chicken pox | 0% |
| Total in Control with chicken pox | 100% |
| **Treament Effect** | -100% |

## With Externalities

Suppose, because prevalence is lower, some children are not re-infected with chicken pox

| School A | Treated? | Outcome |
|---|---|---|
| **Pupil 1** | Yes | no chicken pox |
| Pupil 2 | No | no chicken pox |
| **Pupil 3** | Yes | no chicken pox |
| Pupil 4 | No | chicken pox |
| **Pupil 5** | Yes | no chicken pox |
| Pupil 6 | No | chicken pox |

| | |
|---|---|
| Total in Treatment with chicken pox | 0% |
| Total in Control with chicken pox | 67% |
| **Treatment Effect** | -67% |

# How to measure program impact in the presence of spillovers?

A. Design the unit of randomization so that it encompasses the spillovers

B. If we expect externalities that are all within school:

   A. Randomization at the level of the school allows for estimation of the overall effect

# Example: Price Information

A. Providing farmers with spot and futures price information by mobile phone
B. Should we expect spillovers?
C. Randomize: individual or village level?
D. Village level randomization
    A. Less statistical power
    B. "Purer control groups"
E. Individual level randomization
    A. More statistical power (if spillovers small)
    B. But spillovers might bias the measure of impact

# Example: Price Information

A. Actually can do both together!
B. Randomly assign villages into one of four groups, A, B and C
C. Group A Villages
    A. SMS price information to randomly selected 50% of individuals with phones
    B. Two random groups: Test A and Control A
D. Group B Villages
    A. No SMS price information
E. Allow to measure the true effect of the program: Test A/B
F. Allow also to measure the spillover effect: Control A/B

# Lecture Overview

A. Attrition
B. Spillovers
C. Partial Compliance and Sample Selection Bias
D. Intention to Treat & Treatment on Treated
E. Choice of outcomes
F. External validity
G. Conclusion

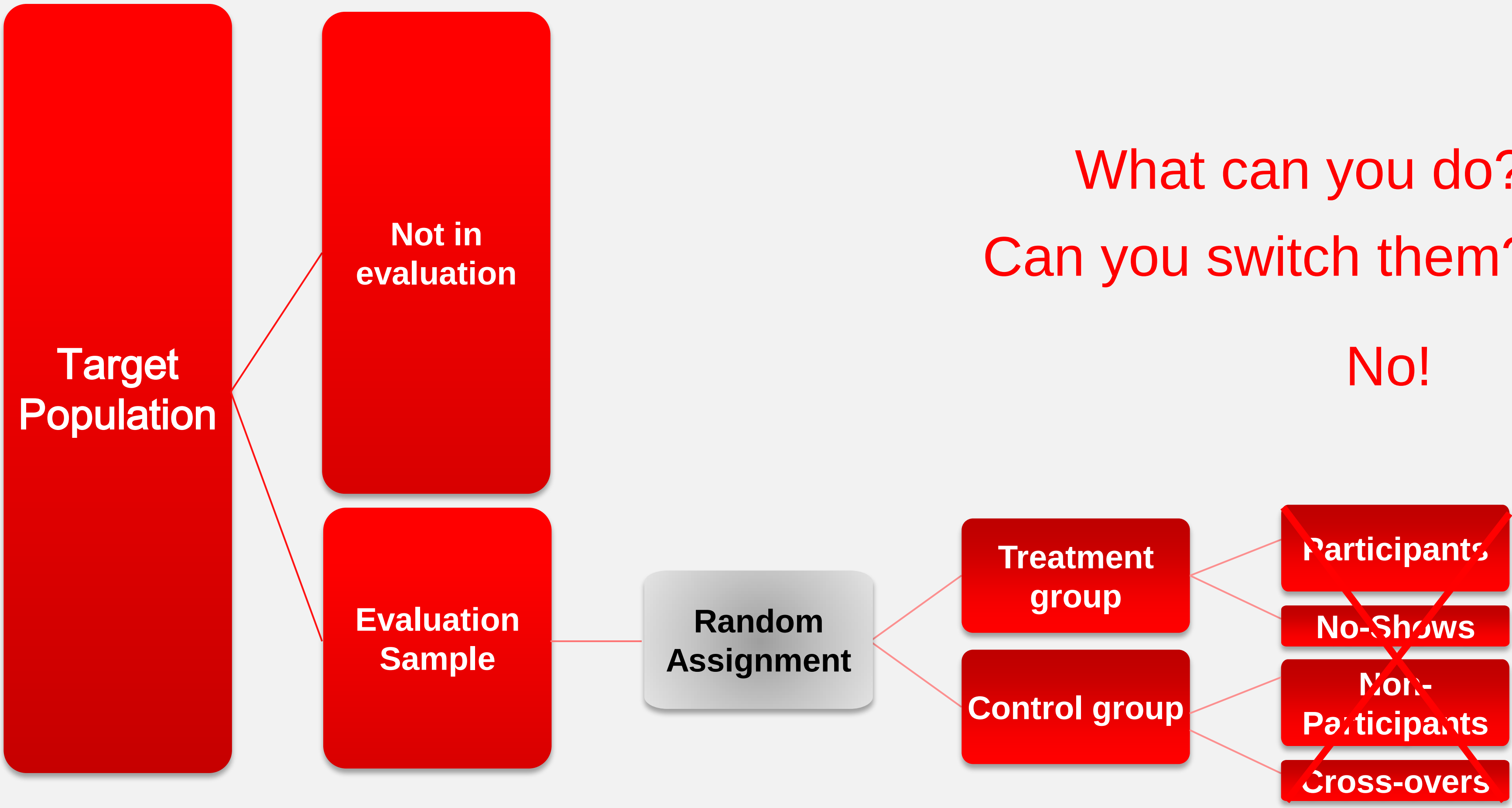
# Sample selection bias

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

# Sample selection bias

A. Individuals assigned to comparison group could attempt to move into treatment group
   A. School feeding program: parents could attempt to move their children from comparison school to treatment school
B. Alternatively, individuals allocated to treatment group may not receive treatment
   A. School feeding program: some students assigned to treatment schools bring and eat their own lunch anyway, or choose not to eat at all.

# Non compliers

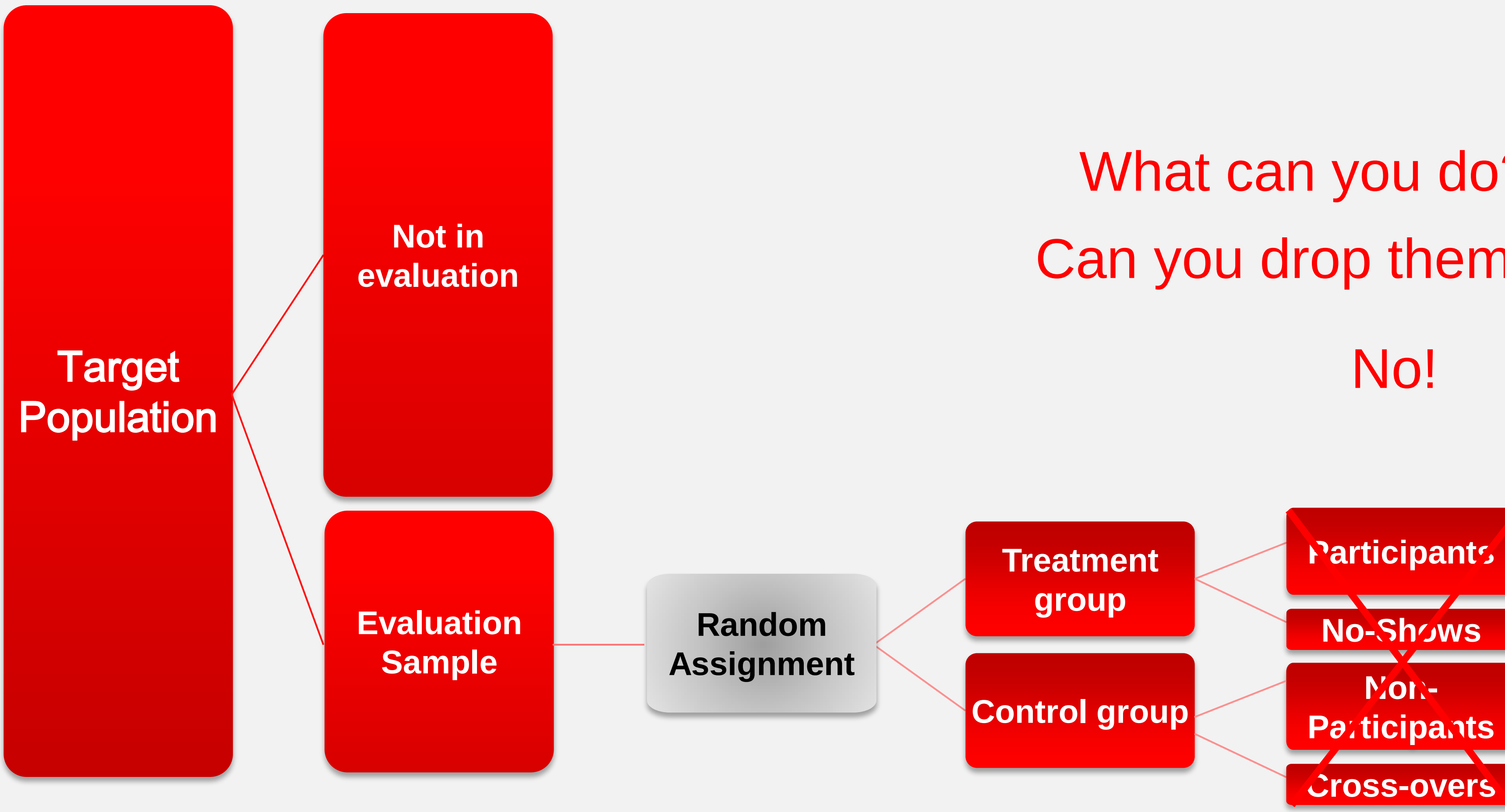



What can you do?

Can you switch them?

No!

# Non compliers

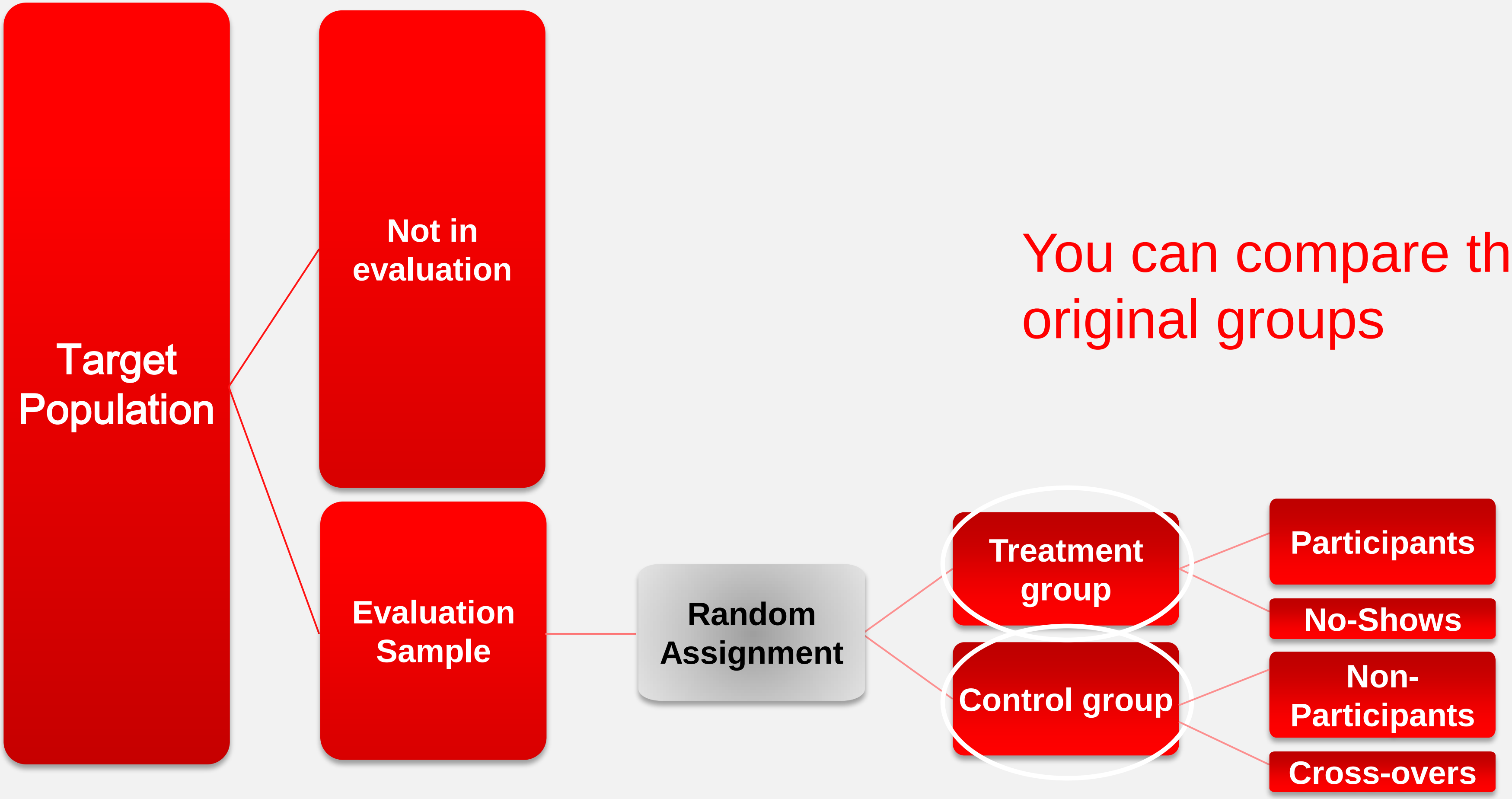



What can you do?

Can you drop them?

No!

# Non compliers




You can compare the original groups

# Lecture Overview

A. Attrition
B. Spillovers
C. Partial Compliance and Sample Selection Bias
D. Intention to Treat & Treatment on Treated
E. Choice of outcomes
F. External validity
G. Conclusion

# ITT and ToT

A. Vaccination campaign in villages

B. Some people in treatment villages not treated
   A. 78% of people assigned to receive treatment received some treatment

C. What do you do?
   A. Compare the beneficiaries and non-beneficiaries?
   B. Why not?

# Which groups can be compared ?

| **Assigned** to Treatment Group: Vaccination | **Assigned** to Control Group |
|---|---|
| TREATED | NON-TREATED |
| NON-TREATED | |

# What is the difference between the 2 random groups?

| Assigned to Treatment Group | Assigned to Control Group |
|---|---|
| 1: treated – not infected<br>2: treated – not infected<br>3: treated – infected | 5: non-treated – infected<br>6: non-treated – not infected<br>7: non-treated – infected<br>8: non-treated – infected |
| 4: non-treated – infected | |

# Intention to Treat - ITT

Assigned to Treatment Group(AT): 50% infected

Assigned to Control Group(AC): 75% infected

- Y(AT)= Average Outcome in AT Group
- Y(AC)= Average Outcome in AC Group

**ITT = Y(AT) - Y(AC)**

- ITT = 50% - 75% = -25 percentage points

# Intention to Treat (ITT)

A. What does "intention to treat" measure?

   *"What happened to the average child who is in a treated school in this population?"*

A. Is this difference a causal effect? Yes because we compare two identical populations

B. But a causal effect of what?

   A. Clearly not a measure of the vaccination

   B. Actually a measure of the global impact of the intervention

# When is ITT useful?

A. May relate more to actual programs
B. For example, we may not be interested in the medical effect of deworming treatment, but what would happen under an actual deworming program.
C. If students often miss school and therefore don't get the deworming medicine, the intention to treat estimate may actually be most relevant.

| 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 among Treated A=** | | | |

| 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 among Not-Treated B=** | | | |

School 1:
Avg. Change among Treated (A)
School 2:
Avg. Change among not-treated (B)

A-B

| 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 among Treated A=** | | | 3 |

| 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 among Not-Treated B=** | | | 0.9 |

| | | |
|---|---|---|
| School 1: Avg. Change among Treated | 3 | (A) |
| School 2: Avg. Change among not-treated | 0.9 | (B) |
| A-B | 2.1 | |

# From ITT to effect of Treatment On the Treated

A. What about the impact on those who received the treatment?

Treatment On the Treated (TOT)

A. Is it possible to measure this parameter?
    A. The answer is yes

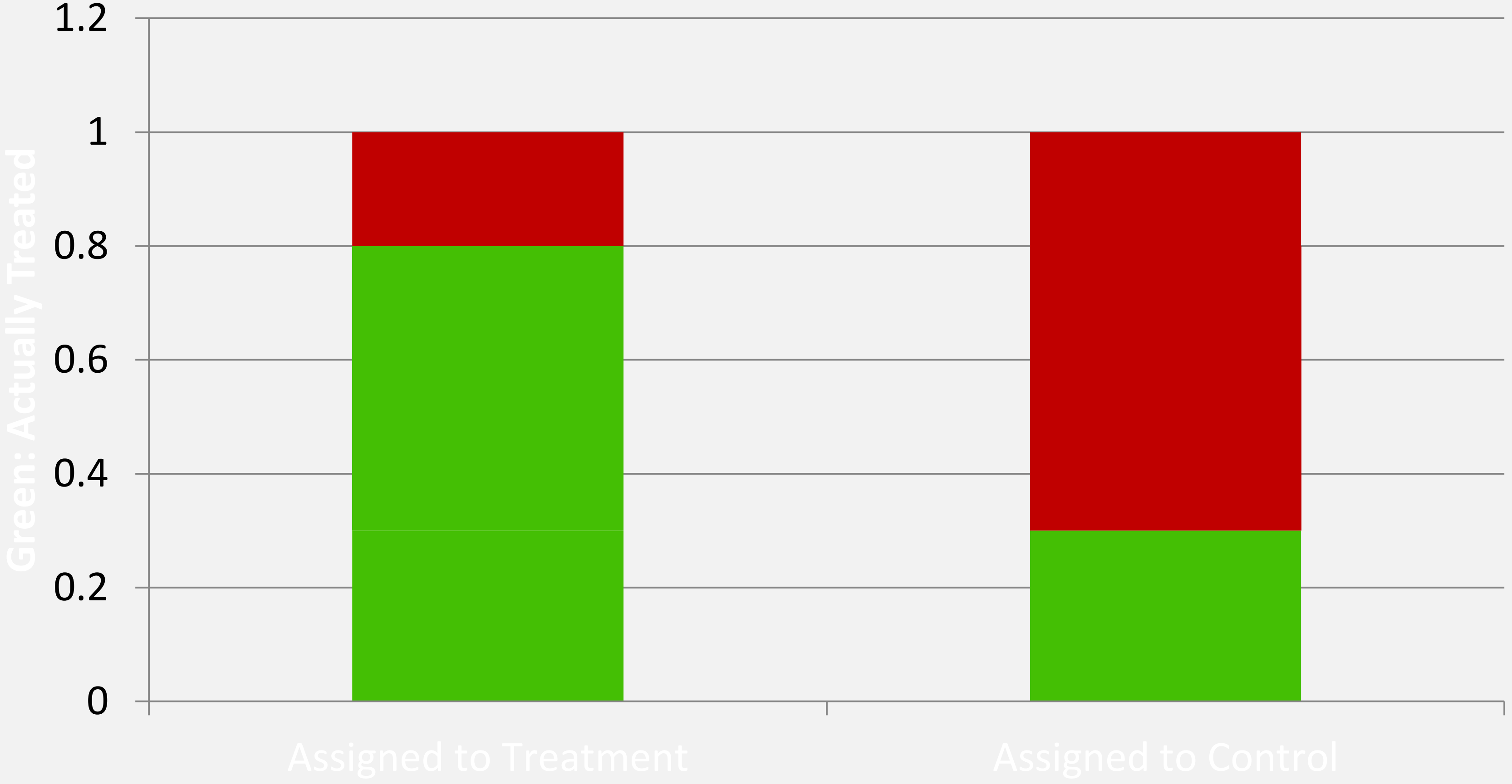
# From ITT to effect of Treatment On the Treated (TOT)

A. The point is that if there is such imperfect compliance, the comparison between those **assigned to treatment** and those **assigned to control** is smaller

B. But the difference in the probability of getting treated is also smaller

**C. The TOT parameter "corrects" the ITT, scaling it up by this "take-up" difference**

# Estimating ToT from ITT: Wald



# Interpreting ToT from ITT: Wald

# Estimating TOT

A. What values do we need?
B. Y(AT) the average value over the Assigned to Treatment group (AT)
C. Y(AC) the average value over the Assigned to Control group (AC)

A. Prob[T|AT] = Proportion of treated in AT group
B. Prob[T|AC] = Proportion of treated in AC group

C. These proportion are called **take-up** of the program

# Treatment on the treated (TOT)

A. Starting from a regression model

$$Y_i = a + B.T_i + e_i$$

A. Angrist and Pischke show

$$B = [E(Y_i | Z_i = 1) - E(Y_i | Z_i = 0)] / [P(T_i = 1 | Z_i = 1) - E(T_i = 1 | Z_i = 0)]$$

A. With Z=1 is assignement to treatment group

# Treatment on the treated (TOT)

$B=[E(Y_i|Z_i=1)-E(Y_i|Z_i=0)]/[P(T_i=1|Z_i=1)-E(T_i=1|Z_i=0)]$

A. Estimates will be

$[Y(\mathbf{AT})-Y(\mathbf{AC})]/[Prob[T|\mathbf{AT}] - Prob[T|\mathbf{AC}]]$

A. The **ratio** of the **ITT** estimates on the **difference in take-up**

# TOT estimate

| 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 Y(T)=** | |

| 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 Y(C) =** | |

A = Gain if Treated
B = Gain if not Treated

ToT Estimator: A-B

$$\textbf{A-B} = \frac{Y(T)-Y(C)}{Prob(Treated|T)-Prob(Treated|C)}$$

| | |
|---|---|
| Y(T) | |
| Y(C) | |
| Prob(Treated\|T) | |
| Prob(Treated\|C) | |

| | |
|---|---|
| Y(T)-Y(C) | |
| Prob(Treated\|T)-Prob(Treated\|C) | |

| | |
|---|---|
| **A-B** | |

# TOT 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 Y(T)=** | 3 |

| 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 Y(C) =** | 0.9 |

A = Gain if Treated
B = Gain if not Treated

ToT Estimator: A-B

$$\textbf{A-B} = \frac{Y(T)-Y(C)}{Prob(Treated|T)-Prob(Treated|C)}$$

| | |
|---|---|
| Y(T) | 3 |
| Y(C) | 0.9 |
| Prob(Treated\|T) | 60% |
| Prob(Treated\|C) | 20% |

| | |
|---|---|
| Y(T)-Y(C) | 2.1 |
| Prob(Treated\|T)-Prob(Treated\|C) | 40% |

| | |
|---|---|
| **A-B** | 5.25 |

# Generalizing the ToT Approach: Instrumental Variables

1. First stage regression

$$T=a_0+a_1Z+Xc+u$$

($a_1$ is the difference in take-up)

2. Get predicted value of treatment:

$$Pred(T|Z,X) = a_0+a_1Z+Xc$$

3. Perform the regression of Y on predicted treatment instead on treatment

$$Y=b_0+b_1Pred(T|Z,X)+Xd+v$$

# Requirements for Instrumental Variables

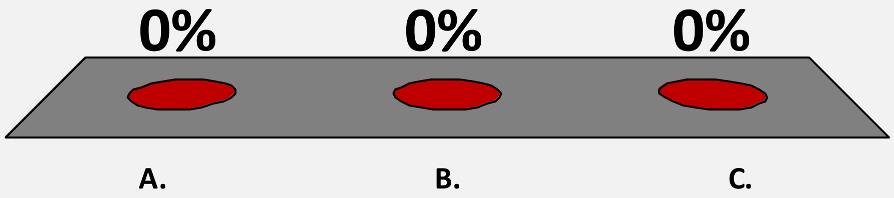
A. First stage
   A. Your experiment (or instrument) meaningfully affects probability of treatment
   B. Actually the experiment is "good" if there is a large effect of assignment to treatment on treatment participation (the difference in take-up)
B. Exclusion restriction
   A. Your experiment (or instrument) does not affect outcomes through another channel

# The ITT estimate will always be smaller (e.g., closer to zero) than the ToT estimate

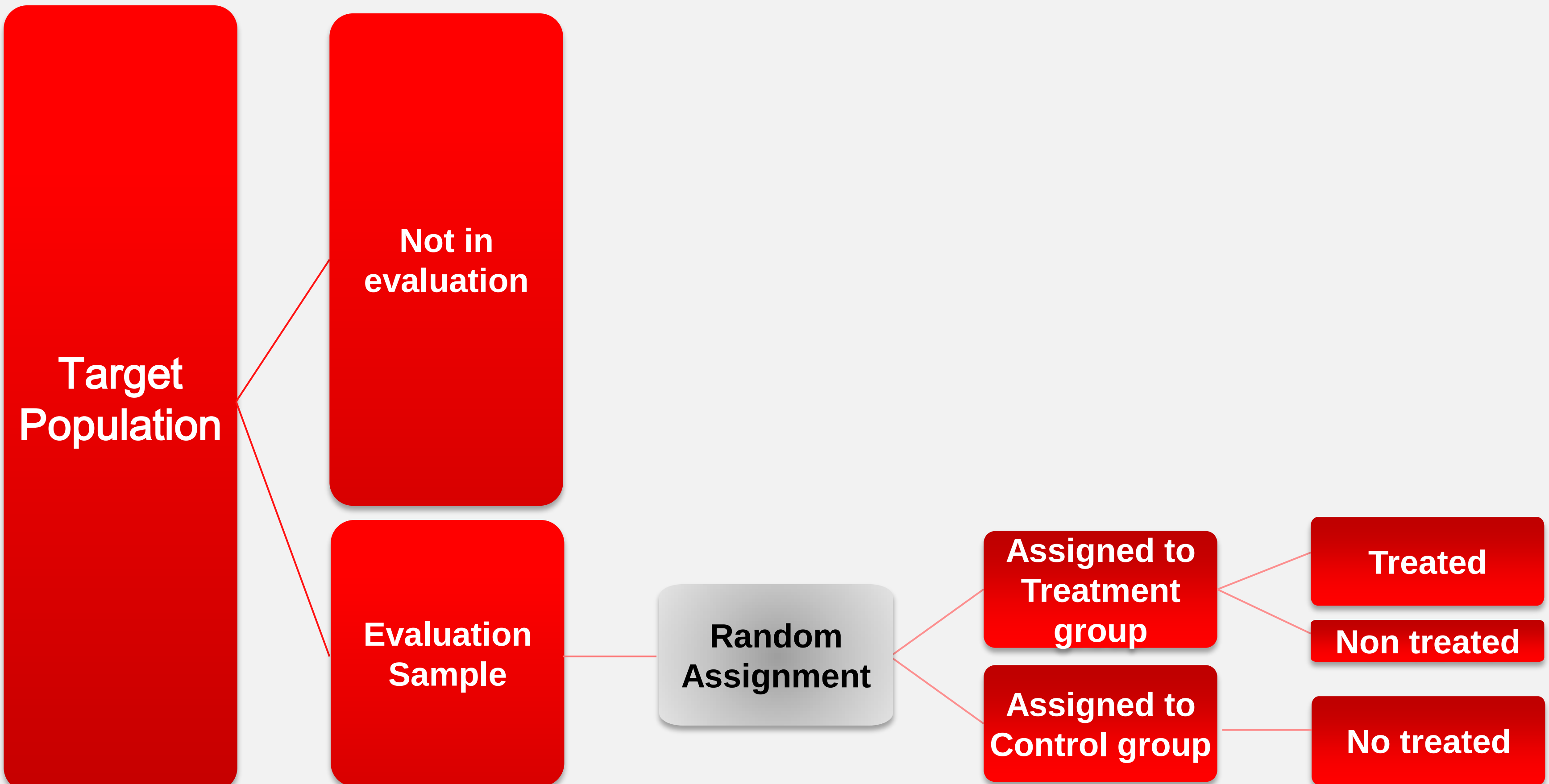
A. True

B. False

C. Don't Know



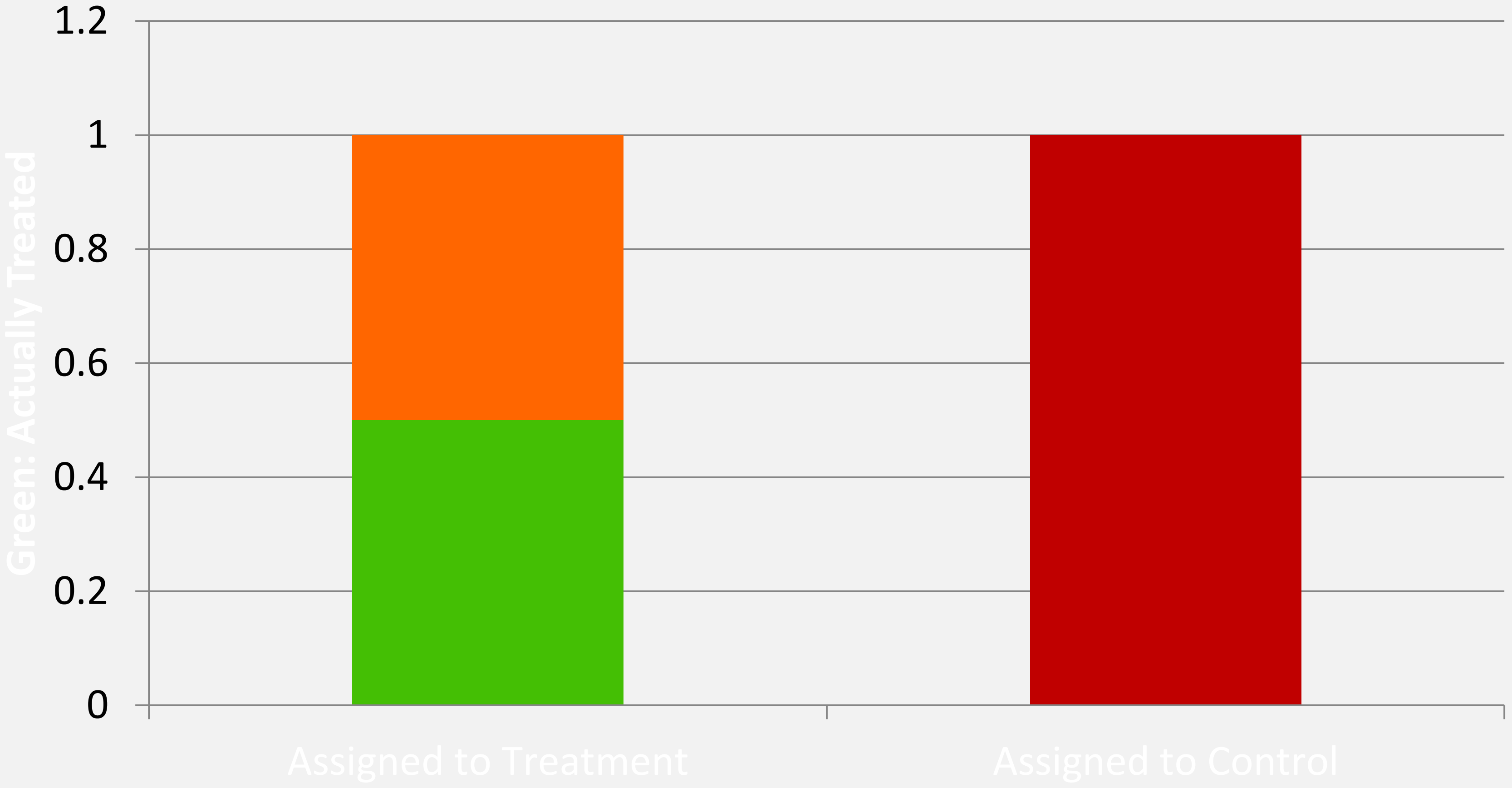
# TOT not always appropriate…



# TOT not always appropriate…

A. Example: send 50% of retired people in Paris a letter warning of flu season, encourage them to get vaccines
B. Suppose 50% in treatment, 0% in control get vaccines
C. Suppose incidence of flu in treated group drops 35% relative to control group
D. Is (.35) / (.5 – 0 ) = 70% the correct estimate?
E. What effect might letter alone have?
F. Some retired people in the assignment to treatment group might consider it is better not to get a vaccine but… to stay home
G. They didn't get the treatment but they have been influenced by the letter

# Non treated in the AT group impacted



# Non treated in AT group do not cancel out

# Lecture Overview

A. Spillovers
B. Partial Compliance and Sample Selection Bias
C. Intention to Treat & Treatment on Treated
D. Choice of outcomes
E. External validity

# Multiple outcomes

A. Can we look at various outcomes?
B. The more outcomes you look at, the higher the chance you find at least one significantly affected by the program
    A. Pre-specify outcomes of interest
    B. Report results on all measured outcomes, even null results
    C. Correct statistical tests (Bonferroni)

# Covariates

A. Why include covariates?
   A. May explain variation, improve statistical power
B. Why not include covariates?
   A. Appearances of "specification searching"
C. What to control for?
   A. If stratified randomization: add strata fixed effects
   B. Other covariates

Rule: Report both "raw" differences and regression-adjusted results

# Lecture Overview

A. Spillovers
B. Partial Compliance and Sample Selection Bias
C. Intention to Treat & Treatment on Treated
D. Choice of outcomes
E. External validity
F. Conclusion

# Threat to external validity:

A. Behavioral responses to evaluations

B. Generalizability of results

# Threat to external validity: Behavioral responses to evaluations

- One limitation of evaluations is that the evaluation itself may cause the treatment or comparison group to change its behavior
  - Treatment group behavior changes: Hawthorne effect
  - Comparison group behavior changes: John Henry effect
- Minimize salience of evaluation as much as possible
- Consider including controls who are measured at end-line only

# Generalizability of results

A. Depend on three factors:
   A. Program Implementation: can it be replicated at a large (national) scale?
   B. Study Sample: is it representative?
   C. Sensitivity of results: would a similar, but slightly different program, have same impact?

# Lecture Overview

A. Spillovers
B. Partial Compliance and Sample Selection Bias
C. Intention to Treat & Treatment on Treated
D. Choice of outcomes
E. External validity
F. Conclusion

# Conclusion

A. There are many threats to the internal and external validity of randomized evaluations…
B. …as are there for every other type of study
C. Randomized trials:
   A. Facilitate simple and transparent analysis
      A. Provide few "degrees of freedom" in data analysis (this is a good thing)
   B. Allow clear tests of validity of experiment

# Further resources

A. Using Randomization in Development Economics Research: A Toolkit (Duflo, Glennerster, Kremer)
B. Mostly Harmless Econometrics (Angrist and Pischke)
C. Identification and Estimation of Local Average Treatment Effects (Imbens and Angrist, Econometrica, 1994).