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

Lecture Overview

- Attrition
- Spillovers
- Partial Compliance and Sample Selection Bias
- Intention to Treat & Treatment on Treated
- Choice of outcomes
- External validity
- Conclusion

Lecture Overview

- Attrition
- Spillovers
- Partial Compliance and Sample Selection Bias
- Intention to Treat & Treatment on Treated
- Choice of outcomes
- External validity
- Conclusion

Attrition

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

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 children in the treatment start going to school more
- First impact of your program: increased enrollment
- In addition, you want to measure the impact on child's growth
  - Second outcome of interest: 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**
	20	20	**22**	20
	25	25	**27**	25
	30	30	**32**	30
Ave.				
Difference			Difference	

	Before Treatment		After Treament	
	T	C	T	C
	20	20	**22**	20
	25	25	**27**	25
	30	30	**32**	30
Ave.	**25**	**25**	**27**	**25**
	Difference	0	Difference	2

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

Attrition Bias

- Devote resources to tracking participants in the experiment
  - Sample non respondent and devote additional resources
- If there is still attrition, check that it is not different in treatment and control. Is that enough?
- Good indication about validity of the main property of the RCT:
  - Compare outcomes of two populations that only differ because one of them receive the program
- **Internal validity**

Attrition Bias

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

Conclusion about attrition

It can be a serious issue

A threat on inernal validity: causal meaning of your parameter

A threat on external validity: even if it has a causal meaning it is not representative of the population

Need a special attention

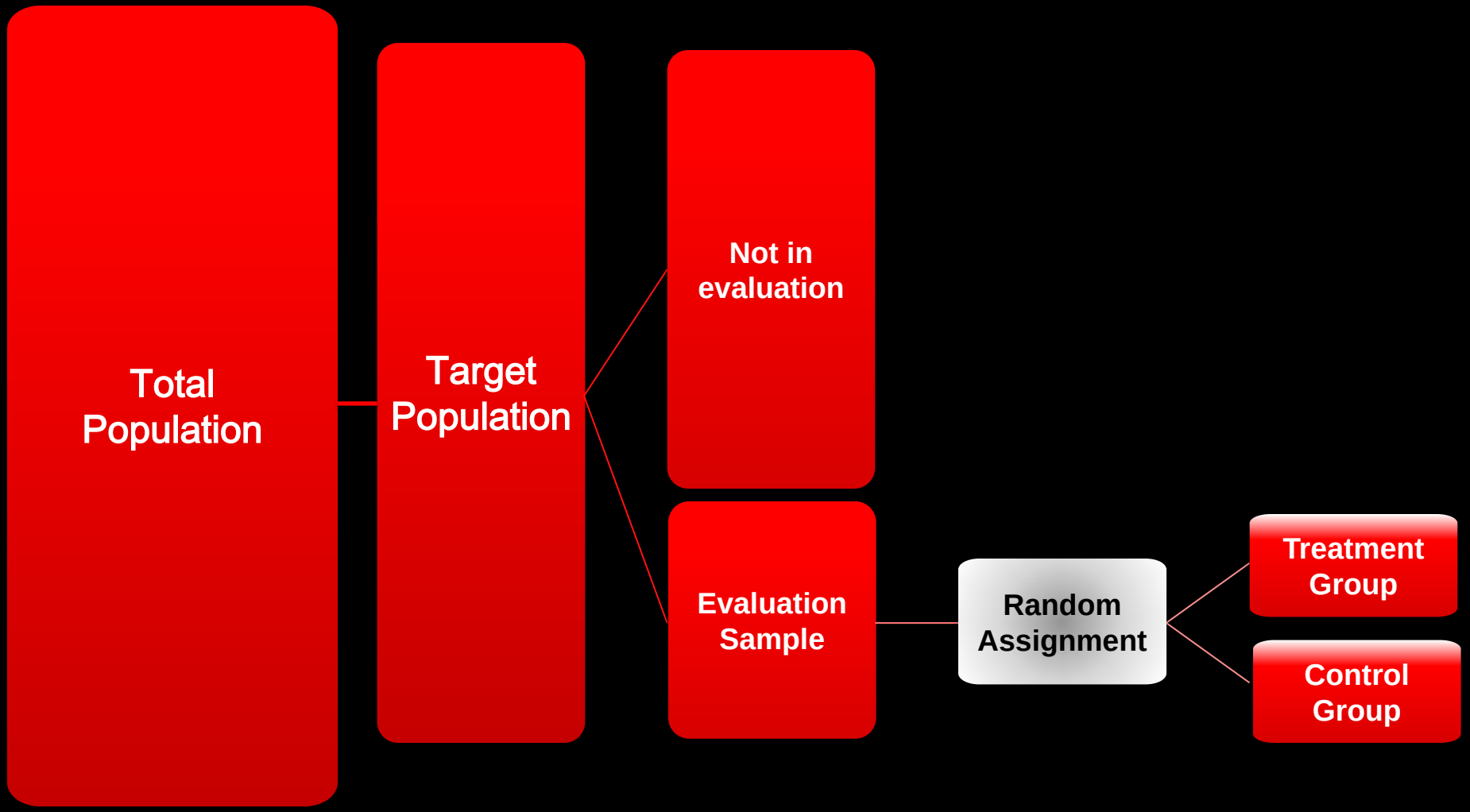
Not true that we cannot do anything

Requires a specific strategies and to secure funds for that

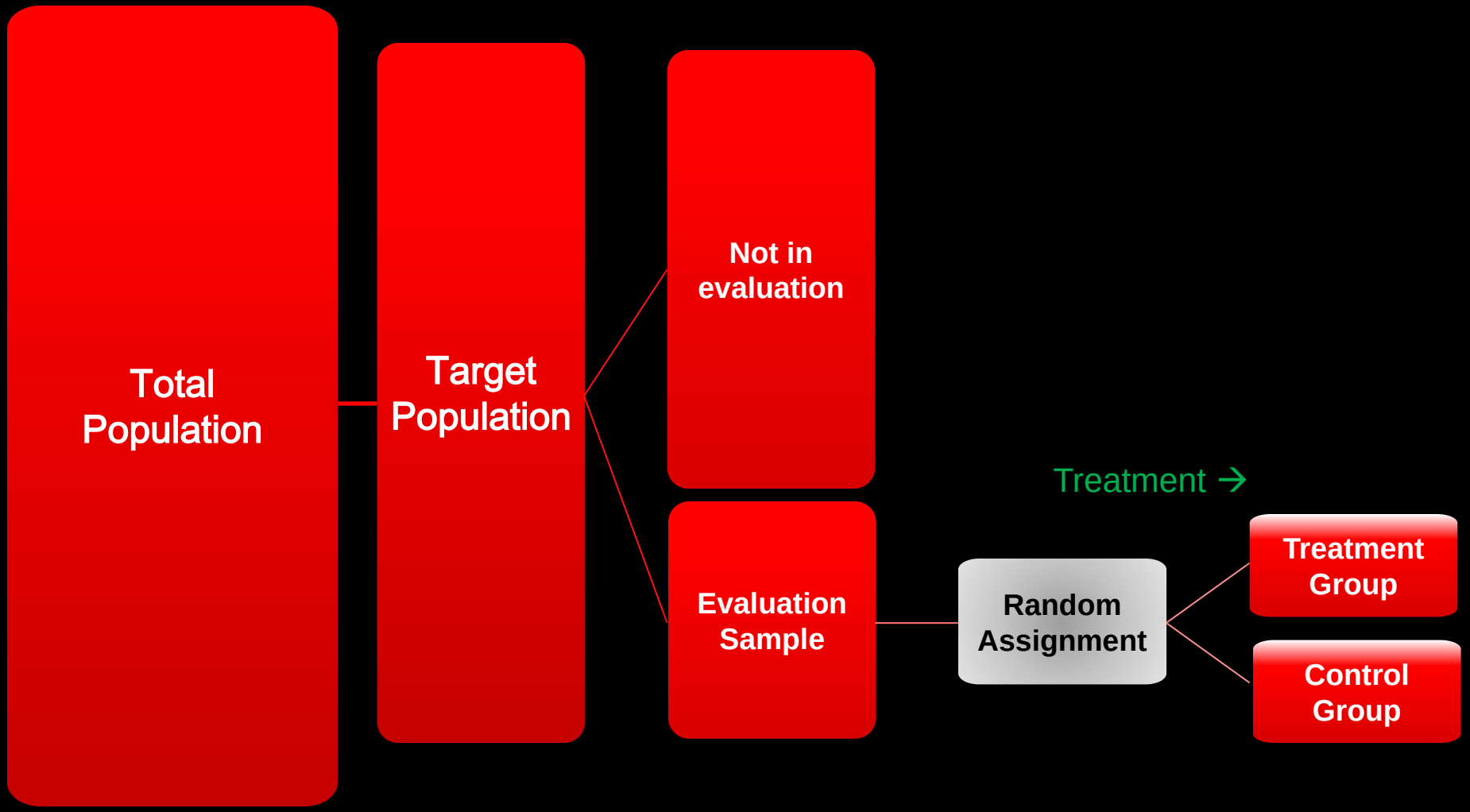
Lecture Overview

- Attrition
- Spillovers
- Partial Compliance and Sample Selection Bias
- Intention to Treat & Treatment on Treated
- Choice of outcomes
- External validity
- Conclusion

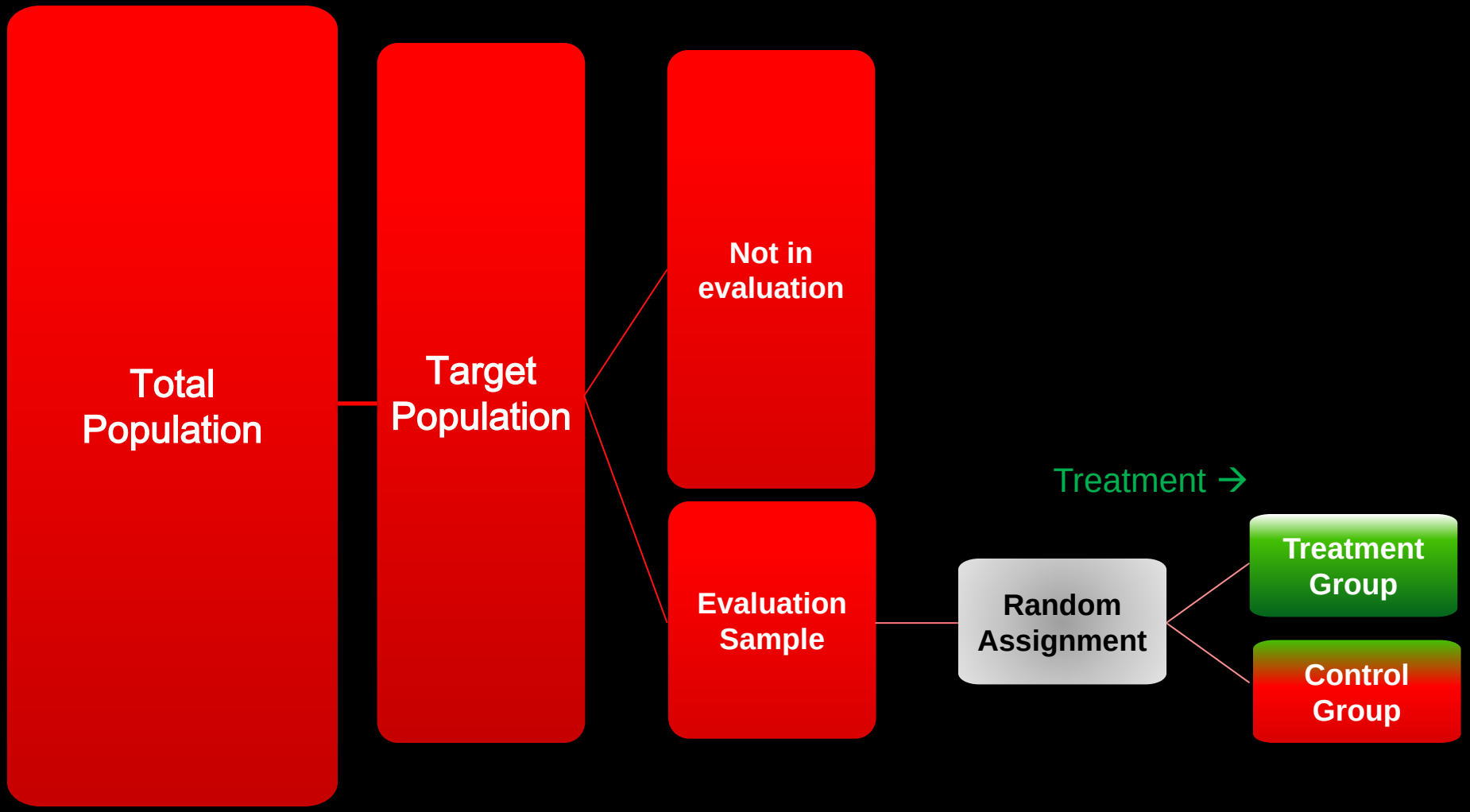
What else could go wrong?



Spillovers, contamination



Spillovers, contamination



Example: Vaccination for chicken pox

- Suppose you randomize chicken pox vaccinations *within* schools
  - Vaccinated youth do not get the disease
  - But, suppose that vaccination also prevents the transmission of disease, what problems does this create for evaluation?
  - Suppose these 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?

- Difficult to account for spillovers
  - Who knows whether pupil2 didn't get the chicken pox because 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: Price Information

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

Example: Price Information

- Actually can do both together!
- Randomly assign villages into one of two groups, A and B
- Group A Villages
  - SMS price information to randomly selected 50% of individuals with phones
  - Two random groups: Test A and Control A
- Group B Villages
  - No SMS price information
- Allows measuring the true effect of the program: Test A/B
- Also allows measuring the spillover effect: Control A/B

Lecture Overview

- Attrition
- Spillovers
- Partial Compliance and Sample Selection Bias
- Intention to Treat & Treatment on Treated
- Choice of outcomes
- External validity
- Conclusion

Conclusion about spillover

- It can also be a serious issue
- Only a threat on internal validity: causal meaning of your parameter
- Need a special attention
- But before running the experimet as the way to deal with it is in the design
- Minimum is to think about the type of spillovers and to choose the level of randomization accordingly

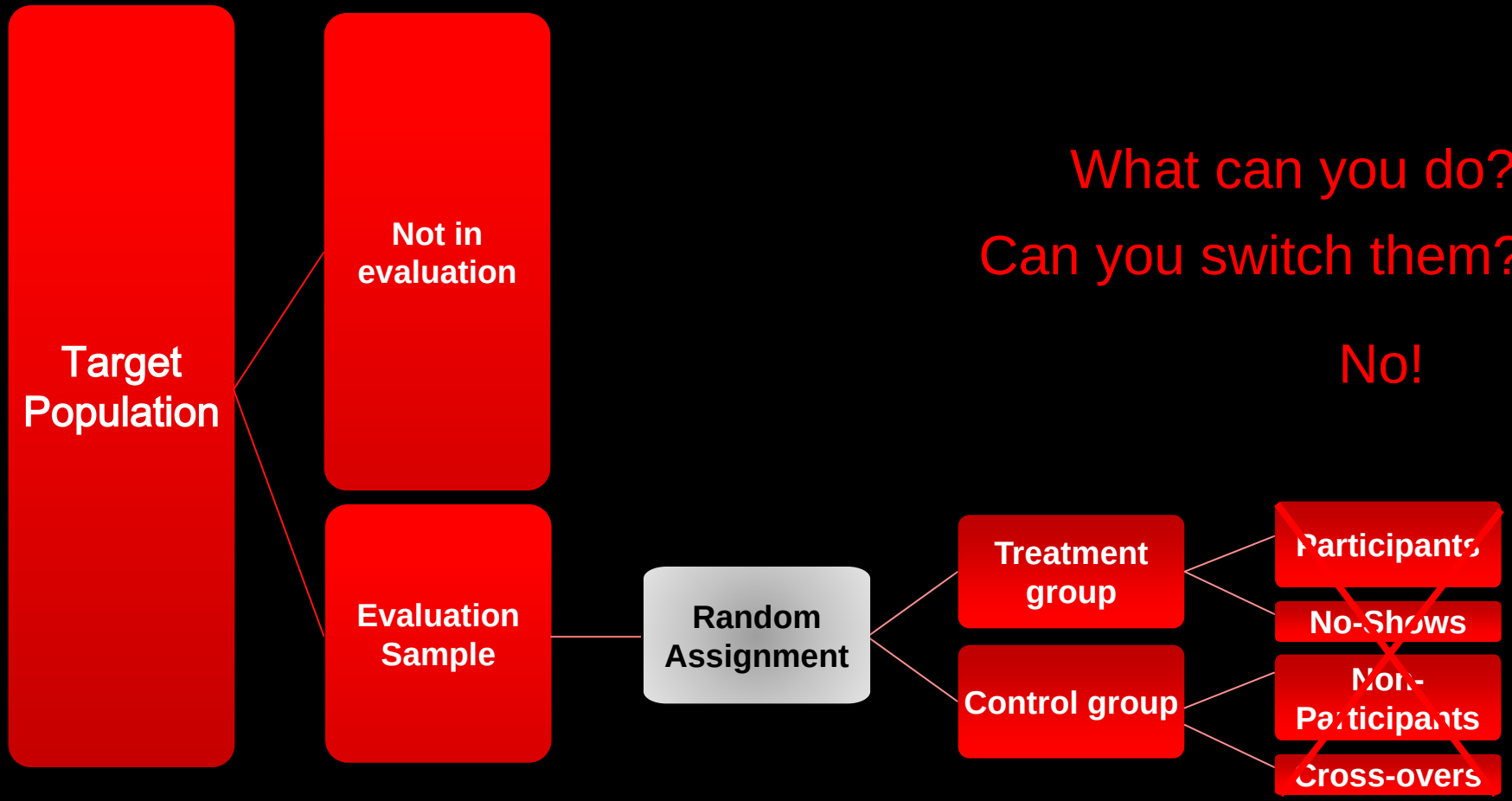
Sample selection bias

- 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

Sample selection bias

- Individuals assigned to comparison group could attempt to move into treatment group
  - School feeding program: parents could attempt to move their children from comparison school to treatment school
- Alternatively, individuals allocated to treatment group may not receive treatment
  - 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

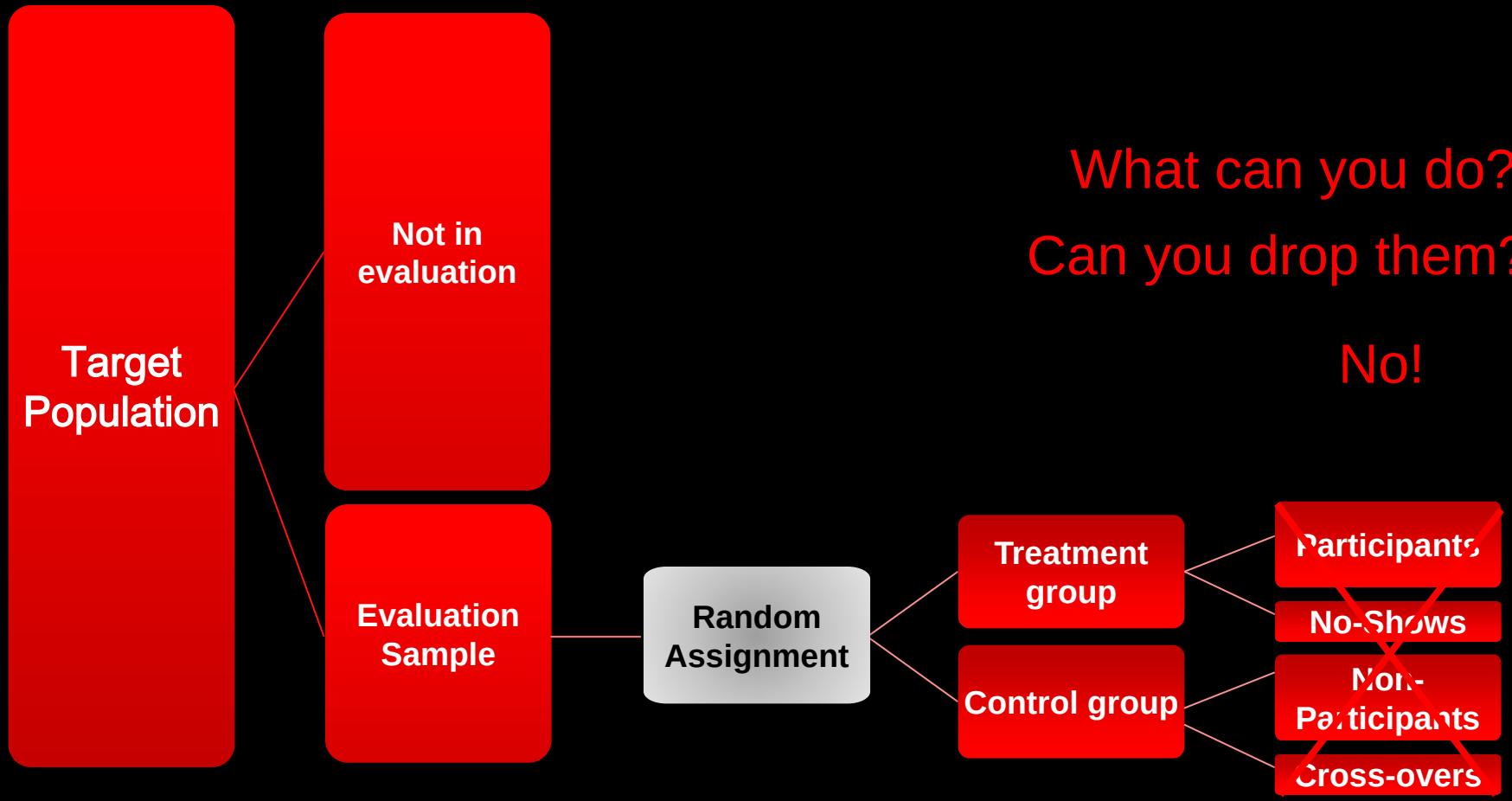


What can you do?

Can you switch them?

No!

Non compliers

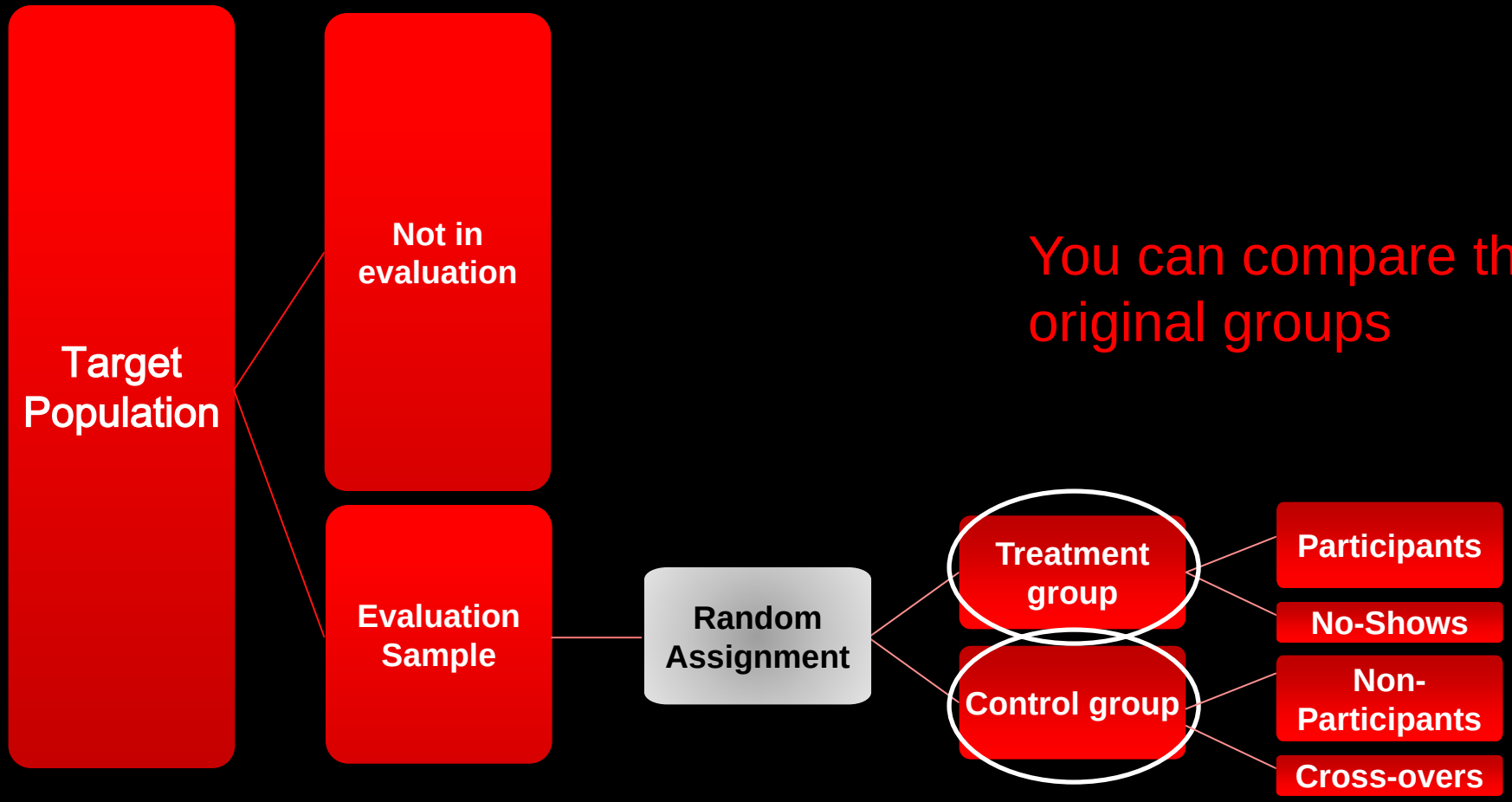


What can you do?

Can you drop them?

No!

Non compliers



You can compare the original groups

And so what?

- We will se in the next section that there is a robustness property of RCTs
- Even in this case of imperfect compliance it is possible to define parameters that have a causal meaning
- We will see that a it is possible to measure a causal impact of the program on Compliers

Lecture Overview

- Attrition
- Spillovers
- Partial Compliance and Sample Selection Bias
- Intention to Treat & Treatment on Treated
- Choice of outcomes
- External validity
- Conclusion

ITT and ToT

- Vaccination campaign in villages
- Some people in treatment villages not treated
  - 78% of people assigned to receive treatment received some treatment
- Same case as previously : there is **imperfect compliance**

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 2: treated – not infected 3: treated – infected	5: non-treated – infected 6: non-treated – not infected 7: non-treated – infected 8: non-treated – infected
4: non-treated – infected	

Intention to Treat - ITT

- Intention to Treat is a **key parameter**
- Simply the difference between
  - these **assigned** to the treatment
  - and these **assigned** to the control
- Ignore the decision to participate

Assigned to Treatment Group(AT): 50% infected

Assigned to Control Group(AC): 75% infected

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

Intention to Treat (ITT)

- What does "intention to treat" measure?

  "What happened to the average individual who is in a treated village in this population?"
- Is this difference a causal effect? Yes because we compare two identical populations
- But a causal effect of what?
  - Clearly not a measure of the vaccination
  - Actually a measure of the **global impact** of the intervention:
    - Information Campaign +Vaccination

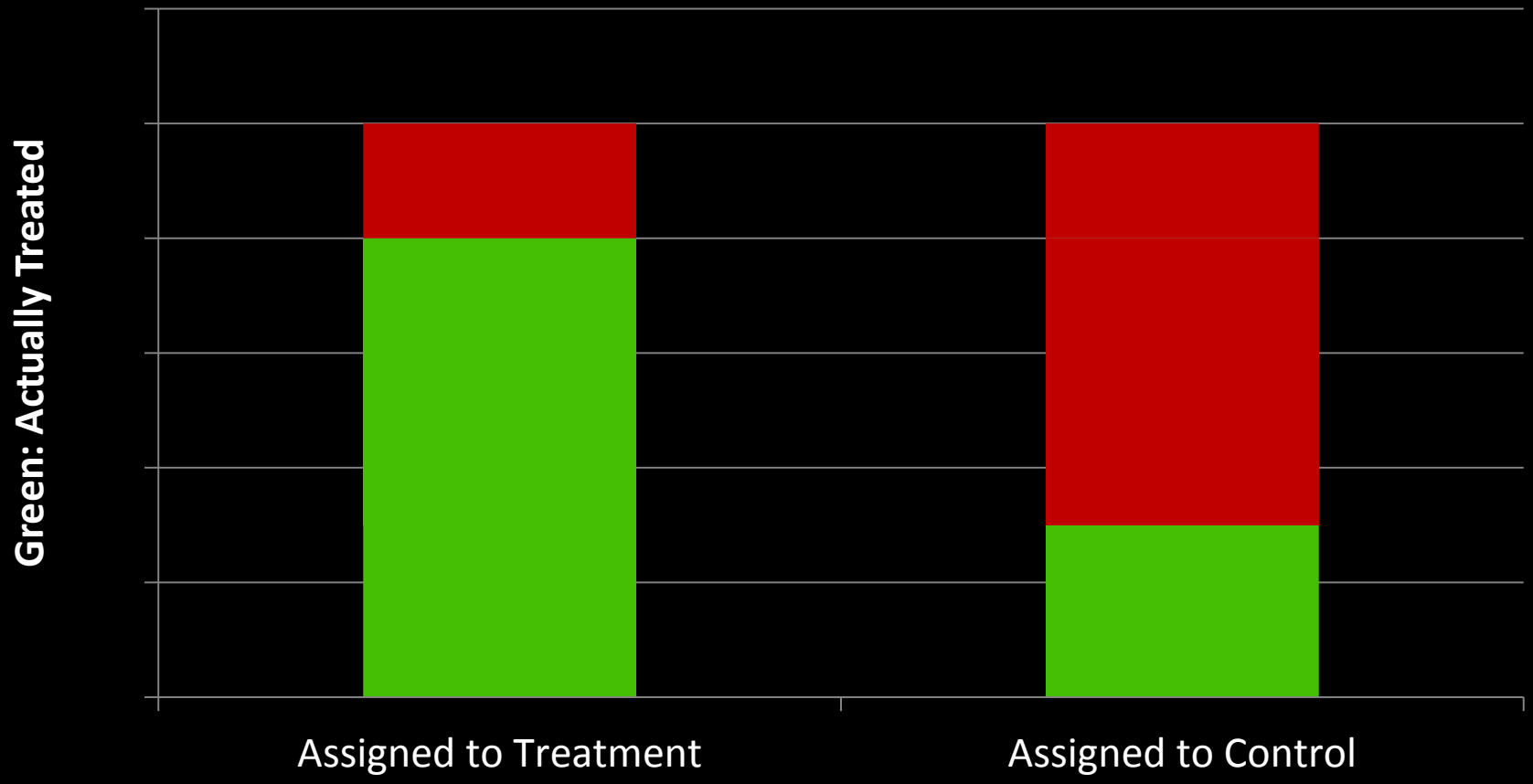
Intention to Treat (ITT)

- This global impact is the combination of two different impacts
- Impact on the decision to enter the program (get vaccinated)
  - Offering the vaccine for free in village has an impact on the fact that individuals get vaccinated
- Impact of the program itself
  - Does the vaccine prevent from getting the disease

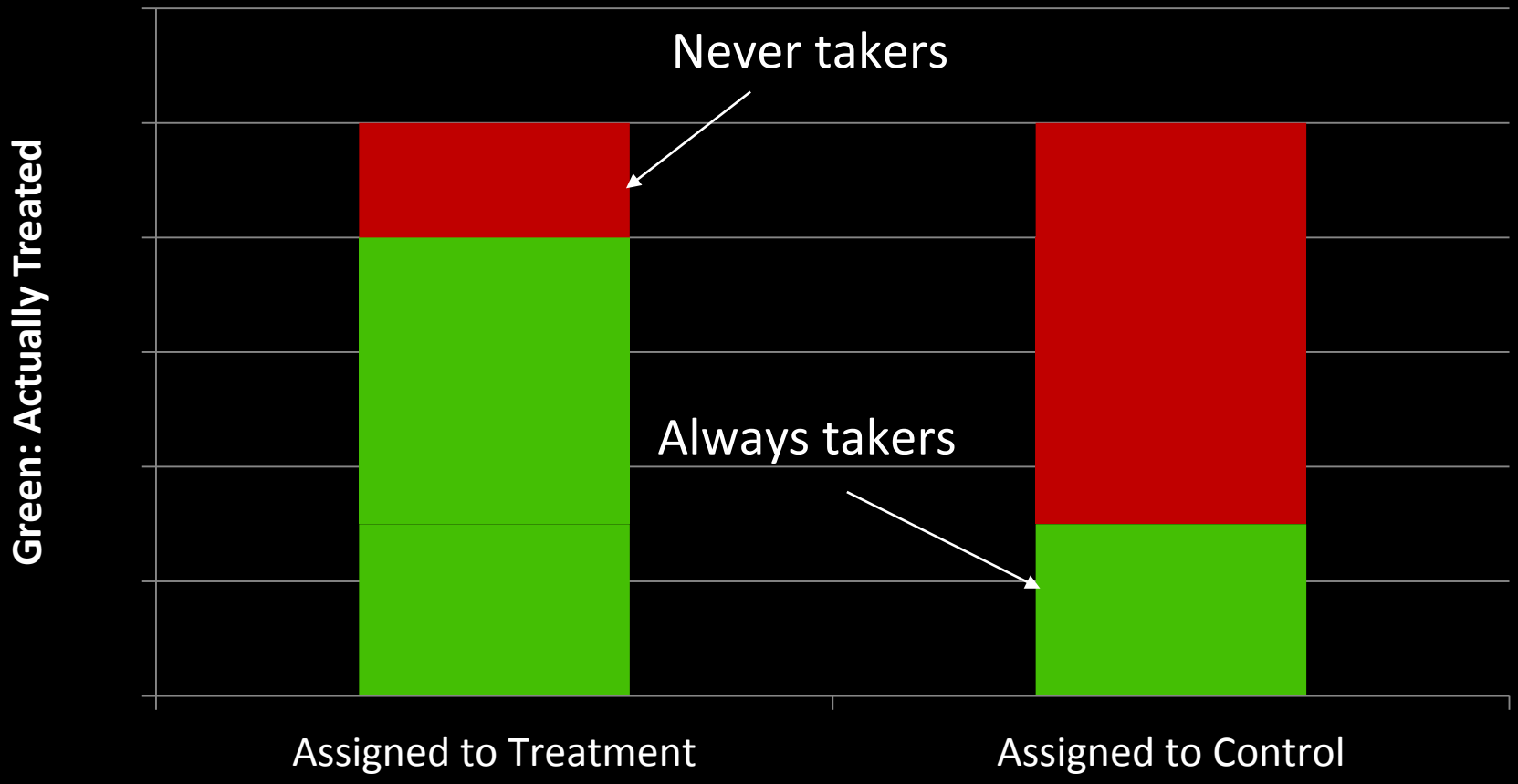
When is ITT useful?

- 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.

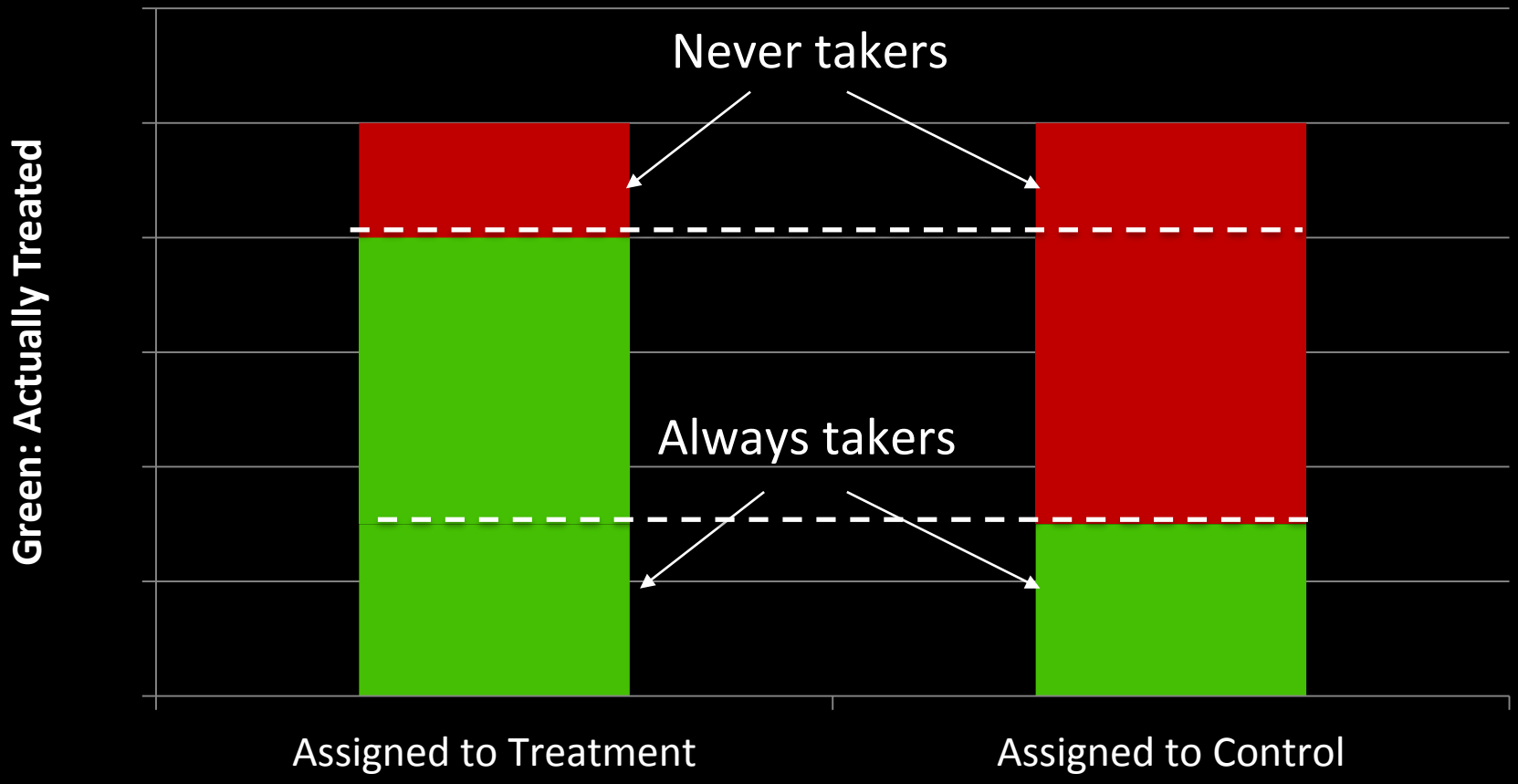
A more general setting



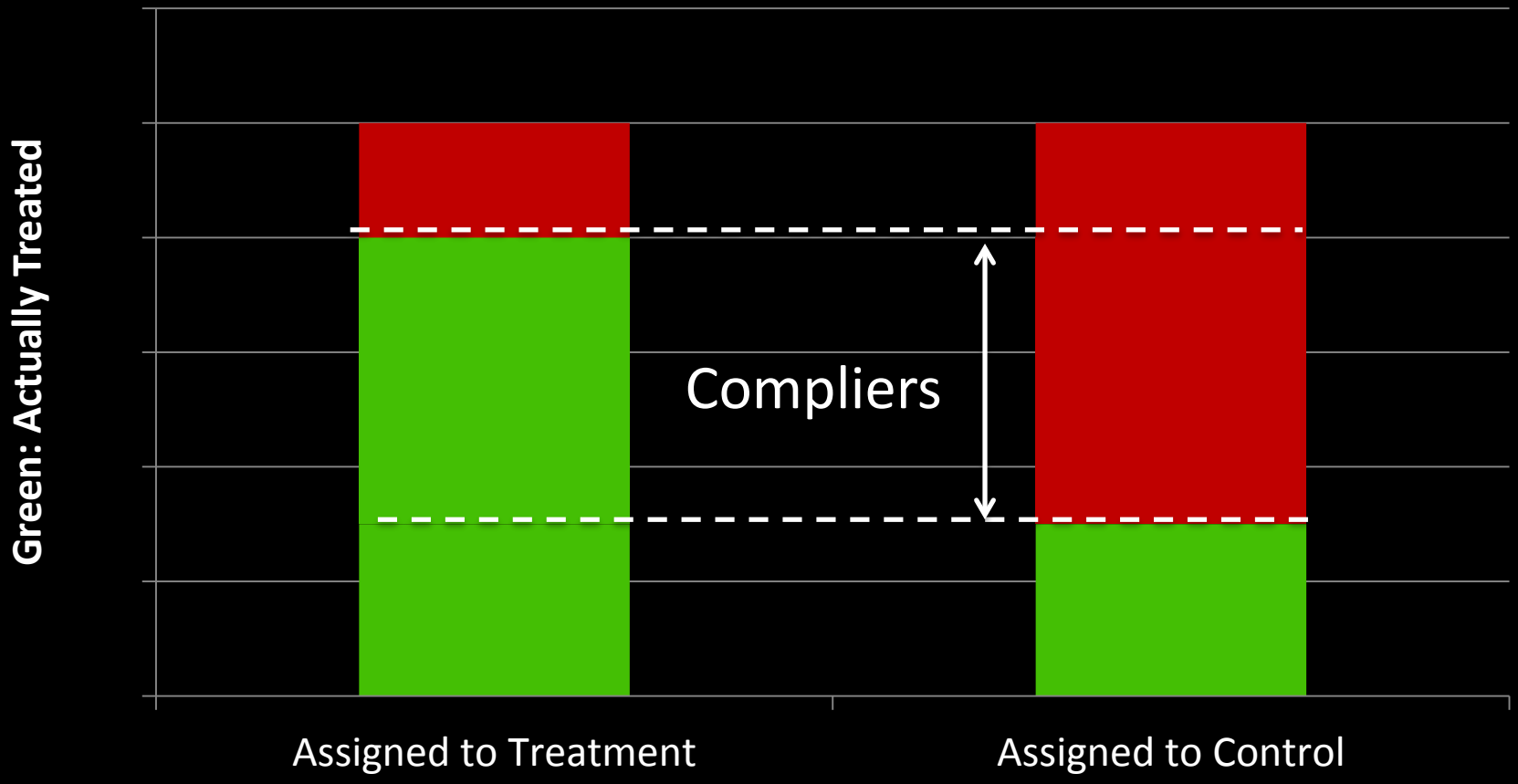
Specific populations



Specific populations



Specific populations



Intention To Treat

- This is as previously said the difference in the average of these assigned to treatment and to control
- It is a mixt of the decision on participation in the program due to the intervention
  - The compliance decision
- And the impact of the program

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

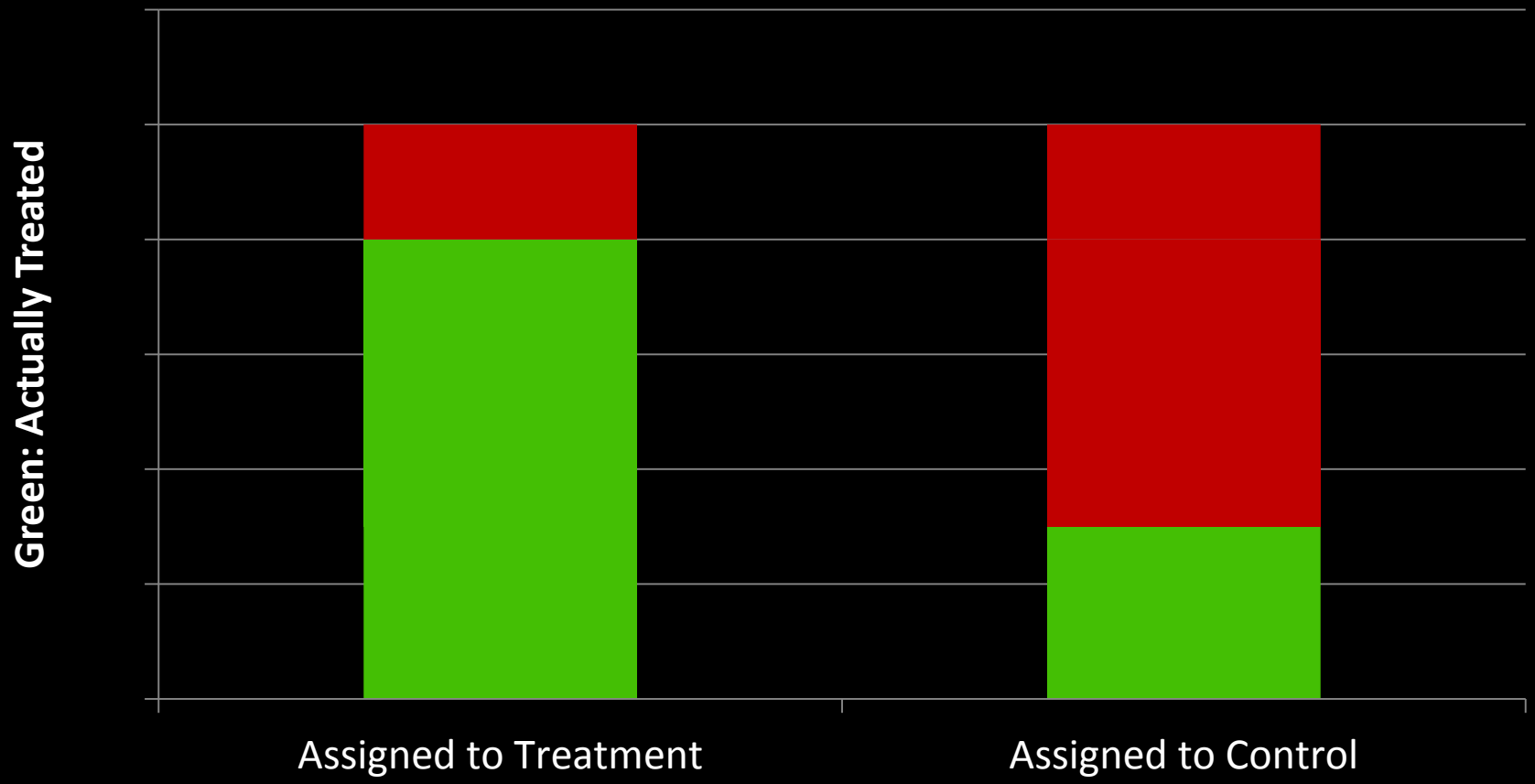
Impact on Compliers

- Actually possible to measure the impact on the compliers

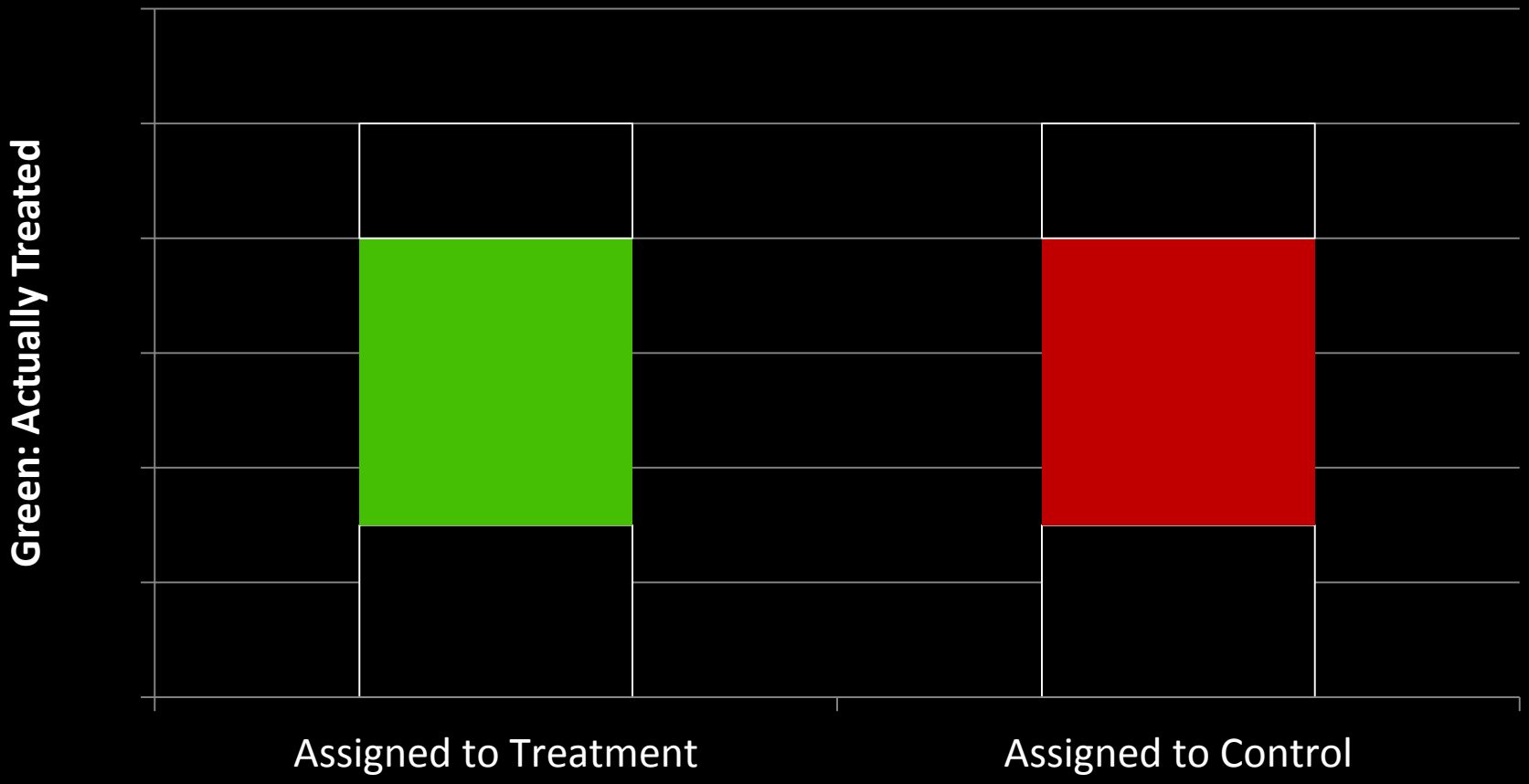
  Treatment On the Treated (TOT)

- Consider ITT
  - Never-takers and Always-takers are not affected
  - Never takers in the treatment group and the control group cancel out: impact on them is zero
  - Same for Always takers

Impact on compliers



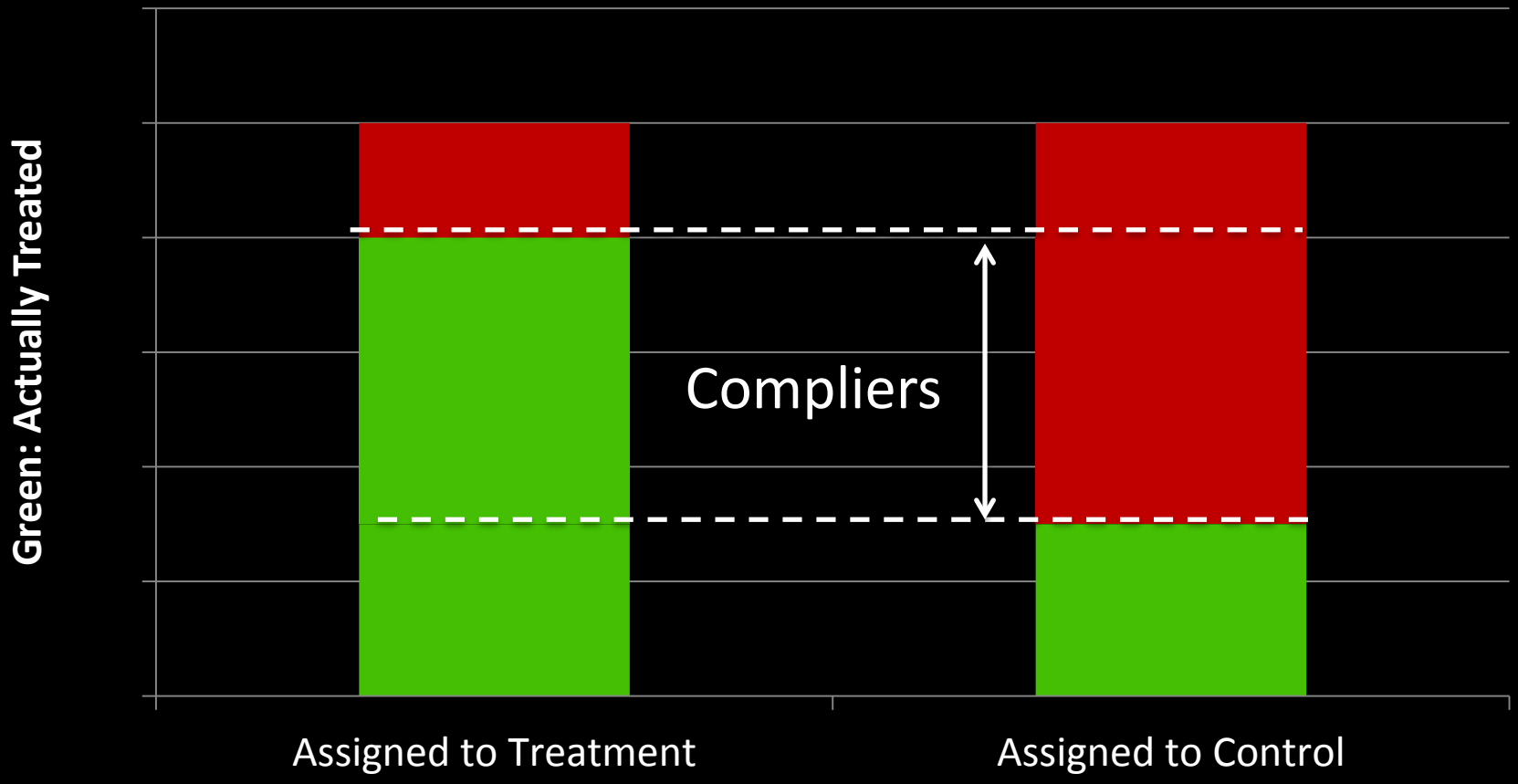
Impact on compliers



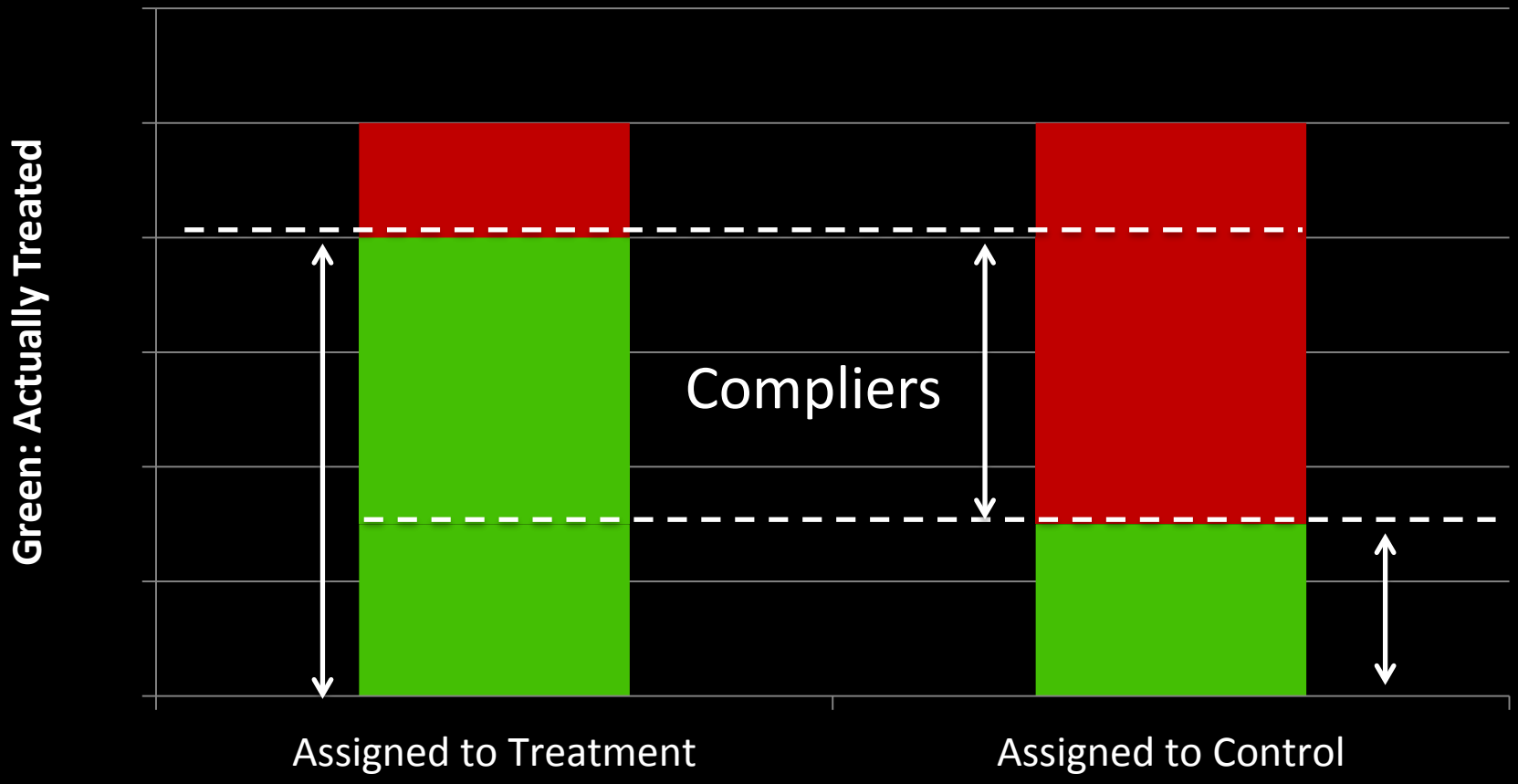
Impact on compliers

- ITT already almost the effect on the compliers
- They are the only one affected in the program
- The only thing to take into account is the size of the population of compliers
- How to measure it?

Size of the complier population



Difference in take-up



Treatment on the treated (TOT)

TOT=ITT/(Size of complier population)

Size of the Complier population

=

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)=	

A = Gain if Treated
B = Gain if not Treated

ToT Estimator: A-B

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

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) =	

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

A = Gain if Treated
B = Gain if not Treated

ToT Estimator: A-B

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

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

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

Meaning of parameters

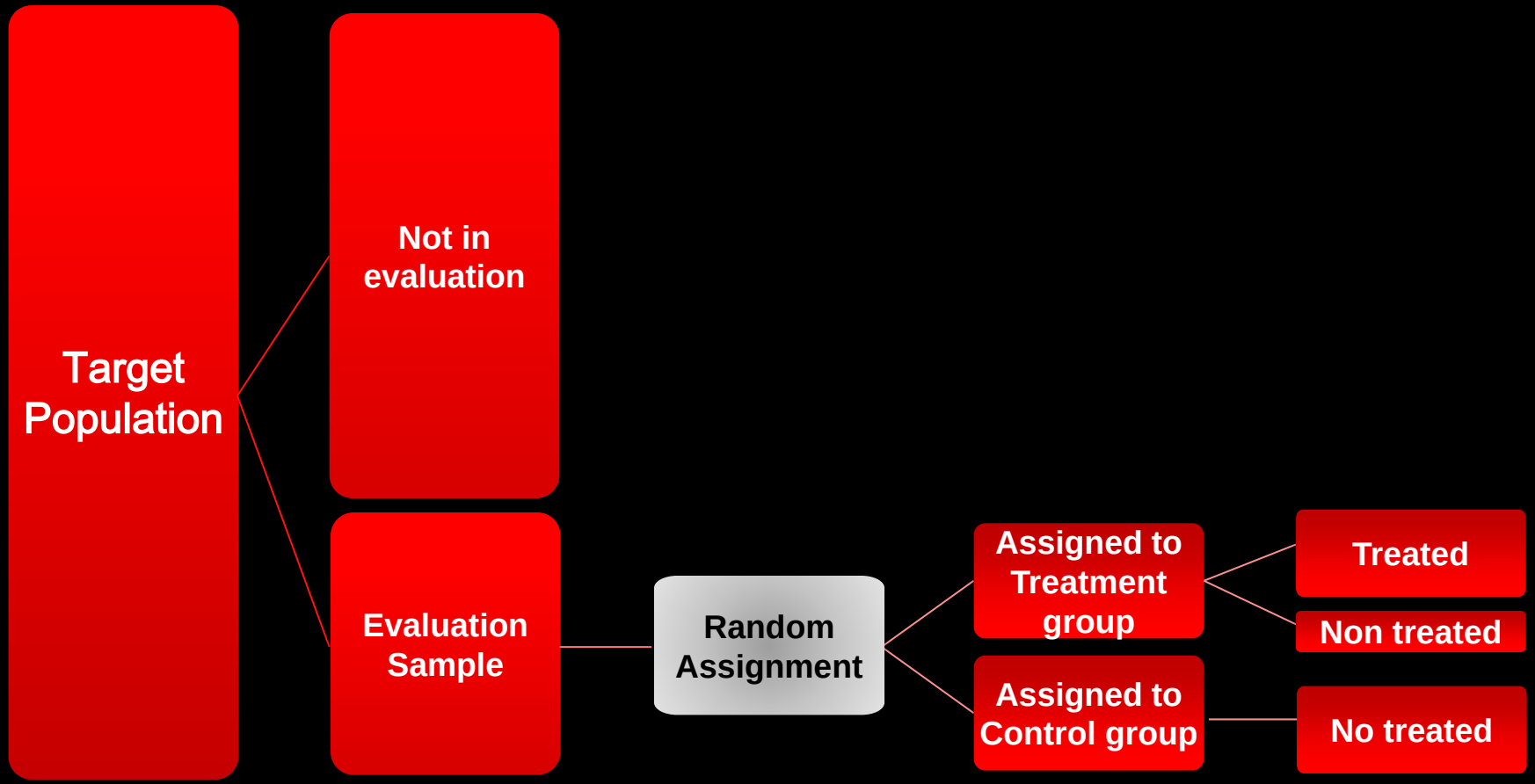
ITT: global impact of the intervention

A mixt of program participation and impact of the program

Difference in take-up: impact of the intervention on participation in the program

TOT: impact of program participation on these who participated thanks to the intervention

TOT not always appropriate…



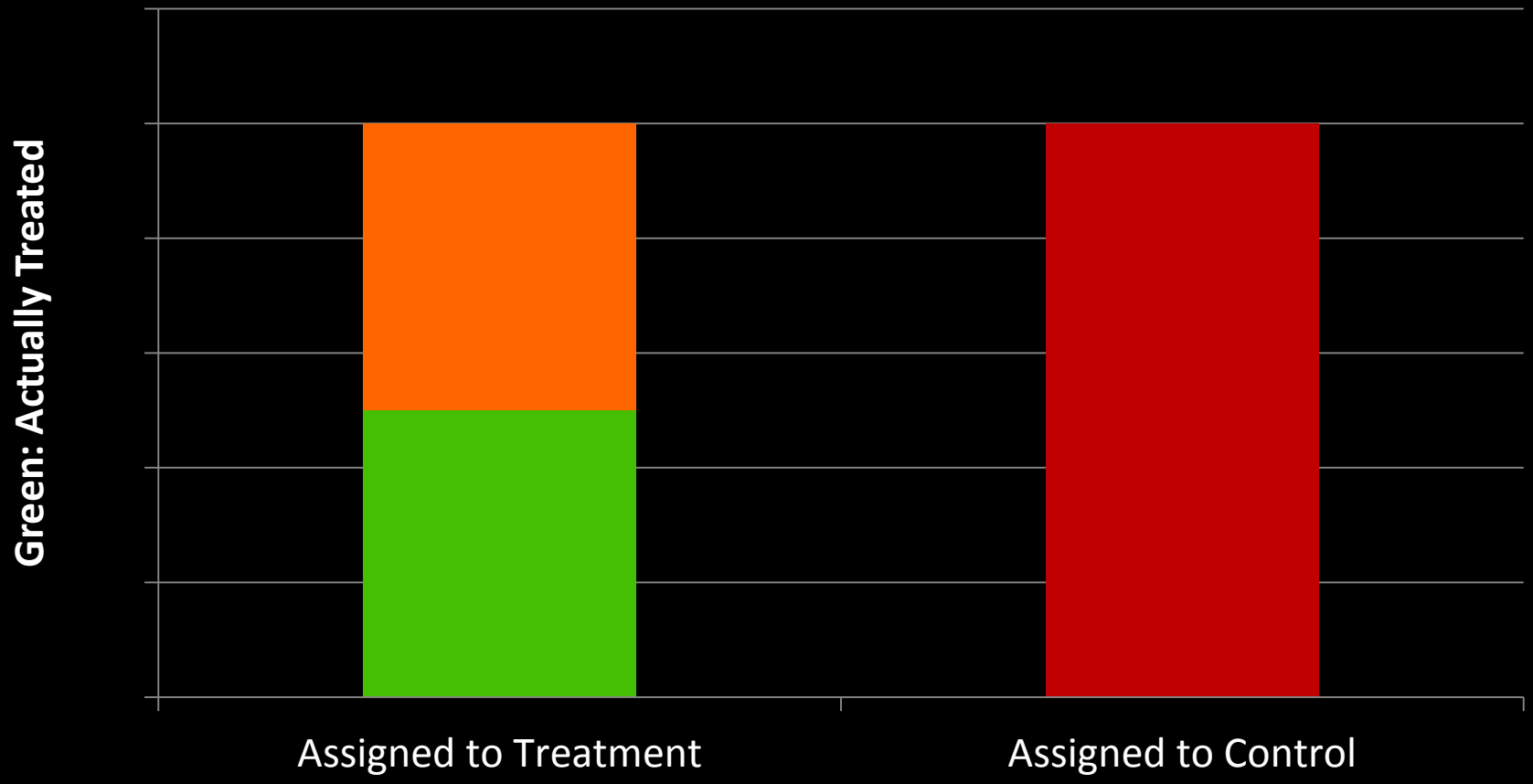
TOT not always appropriate…

- Actually a strong assumption behind TOT
- Never takers and Always takers are NOT affected by the intervention
- Might not be the case

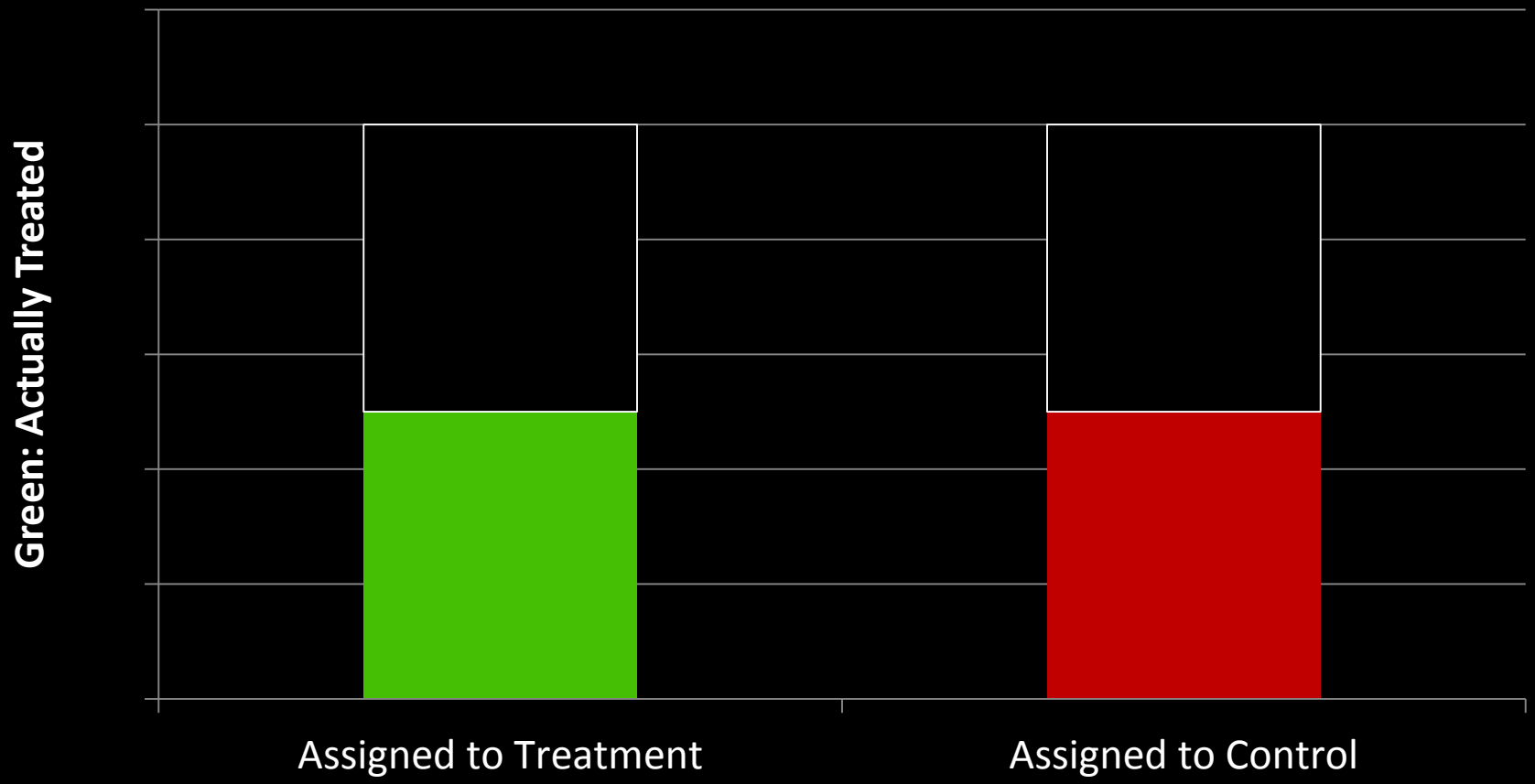
Example

- Intervention: send a letter to retired people in Paris warning of flu season, encourage them to get vaccines
- Suppose 50% in treatment, 0% in control get vaccines
- Suppose incidence of flu in treated group drops -5% relative to control group
- Is (-.05) / (.5 – 0) = -10% the correct estimate?
- What effect might letter alone have?
- Some retired people assigned to treatment might consider it is better not to get a vaccine but… to stay home
- They didn't get the treatment but they have been influenced by the letter

Never takers in the AT group impacted



Never takers do not cancel out



Lecture Overview

- Spillovers
- Partial Compliance and Sample Selection Bias
- Intention to Treat & Treatment on Treated
- Choice of outcomes
- External validity

Multiple outcomes

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

Covariates

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

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

Lecture Overview

- Spillovers
- Partial Compliance and Sample Selection Bias
- Intention to Treat & Treatment on Treated
- Choice of outcomes
- External validity
- Conclusion

Threat to external validity:

- Behavioral responses to evaluations
- 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

Lecture Overview

- Spillovers
- Partial Compliance and Sample Selection Bias
- Intention to Treat & Treatment on Treated
- Choice of outcomes
- External validity
- Conclusion

Conclusion

- There are threats to the internal and external validity of randomized evaluations…
  - …as are there for every other type of study
- It is possible however to anticipate some of them
  - With a specific design: spillover
- It is possible also to deal with some of them
  - Requires investments: tracking non-respondent
  - Also a robustness property: even if non compliance can measure impact on compliers