Principles of Empirical Analysis

Lecture 11: Instrumental Variables (IV)

Spring 2024
Miri Stryjan

Observational alternatives to experiments

1. **Selection on observables: treatment and control groups differ from each other only w.r.t. observable characteristics**
   - We saw an example in Lecture 7
2. **Selection on unobservables: treatment and control groups differ from each other in unobservable characteristics**
   - Treatment and controls are observed before and after treatment – difference-in-differences (DID)
   - Selection mechanism is known – regression discontinuity designs (RDD)
   - Exogenous variable induces variation in treatment – instrumental variables (IV)

Outline

- **Prelude - *Compliance*** in experiments
- **Basic idea of IV designs**
  - Setup, motivation and assumptions
- **IV assumptions and components**
  - Relevance, exogeneity, exclusion restriction
  - First stage, Reduced form and IV estimate
- **Testing the IV assumptions**
- **the limitations of IV**
  - Conditions not fulfilled
  - LATE – the Local Average Treatment Effect limits interpretation
- **Applications**

Outline - compliance

Our first topic: How to deal with imperfect compliance?

First, what is "imperfect compliance"?

some randomized into the treatment group do <u>not</u> get treated

some randomized into the control group still get the treated

Key concepts

- compliers, always-takers and never-takers
- intention-to-treat (ITT)
- first-stage
- local average treatment effect (LATE)
- average treatment effect on the treated (ATT or TOT)

Recall Moving to Opportunity (Lecture 6)?

One of the most famous social experiments of all time

- target group: households with children living in high-poverty public housing projects (primarily minority, single mother families)
- implemented in 1994-98 in Baltimore, Boston, Chicago, LA, New York

Random assignment of 4,600 families into three groups:

- **control**: not offered a voucher, stayed in public housing
- **section 8**: offered conventional housing vouchers, no restrictions on where they would move to.
- **experimental**: offered housing vouchers to **low-poverty neighborhoods.**

Impacts of being offered an experimental voucher (4–7 years later)

- no effects on adult economic self-sufficiency or physical health
- improved mental health for adults
- positive effect on teenage girls but negative effect on teenage boys

Most recent results

You discussed the earlier results from MTO with Prottoy in Lecture 6.

Today we will look at the most recent results, where the authors followed up on the individuals who moved as children, to track their labor market outcomes as adults:

Chetty, R., Hendren, N. and Katz, L.F., 2016. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. American Economic Review, 106(4), pp.855-902.

One feature of this study setup becomes important: **Take-up of was imperfect**: many households allocated to treatment and offered the vouchers did not actually move to another neighborhood.

Take-up in MTO

Table 2—First-Stage Impacts of MTO on Voucher Take-Up and Neighborhood Poverty Rates (*Percentage Points*)

	Housing voucher take-up	Poverty rate in tract one year post- RA		Mean poverty rate in tract post-RA to age 18		Mean poverty rate in zip post-RA to age 18	
		ITT	TOT	ITT	TOT	ITT	TOT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Children < age 13 at random assignment							
Exp. versus control	47.66***	−17.05***	−35.96***	−10.27***	−21.56***	−5.84***	−12.23***
	(1.653)	(0.853)	(1.392)	(0.650)	(1.118)	(0.425)	(0.752)
Sec. 8 versus control	65.80***	−14.88***	−22.57***	−7.97***	−12.06***	−3.43***	−5.17***
	(1.934)	(0.802)	(1.024)	(0.615)	(0.872)	(0.423)	(0.622)
Observations	5,044	4,958	4,958	5,035	5,035	5,035	5,035
Control group mean	0	50.23	50.23	41.17	41.17	31.81	31.81
Panel B. Children age 13–18 at random assignment							
Exp. versus control	40.15***	−14.00***	−34.70***	−10.04***	−24.66***	−5.51***	−13.52***
	(2.157)	(1.136)	(2.231)	(0.948)	(1.967)	(0.541)	(1.113)
Sec. 8 versus control	55.04***	−12.21***	−22.03***	−8.60***	−15.40***	−3.95***	−7.07***
	(2.537)	(1.078)	(1.738)	(0.920)	(1.530)	(0.528)	(0.921)
Observations	2,358	2,302	2,302	2,293	2,293	2,292	2,292
Control group mean	0	49.14	49.14	47.90	47.90	35.17	35.17

Notes: Columns 1, 2, 4, and 6 report ITT estimates from OLS regressions (weighted to adjust for differences in sampling probabilities across sites and over time) of an outcome on indicators for being assigned to the experimental voucher group and the Section 8 voucher group as well as randomization site indicators. Columns 3, 5, and 7 report

Take-up in MTO

of the children below age 13 at time of randomization..

- only 48% of those randomized into the experimental group in MTO actually used the voucher.
- similarly, 66% of those allocated to the section 8 group used the voucher.

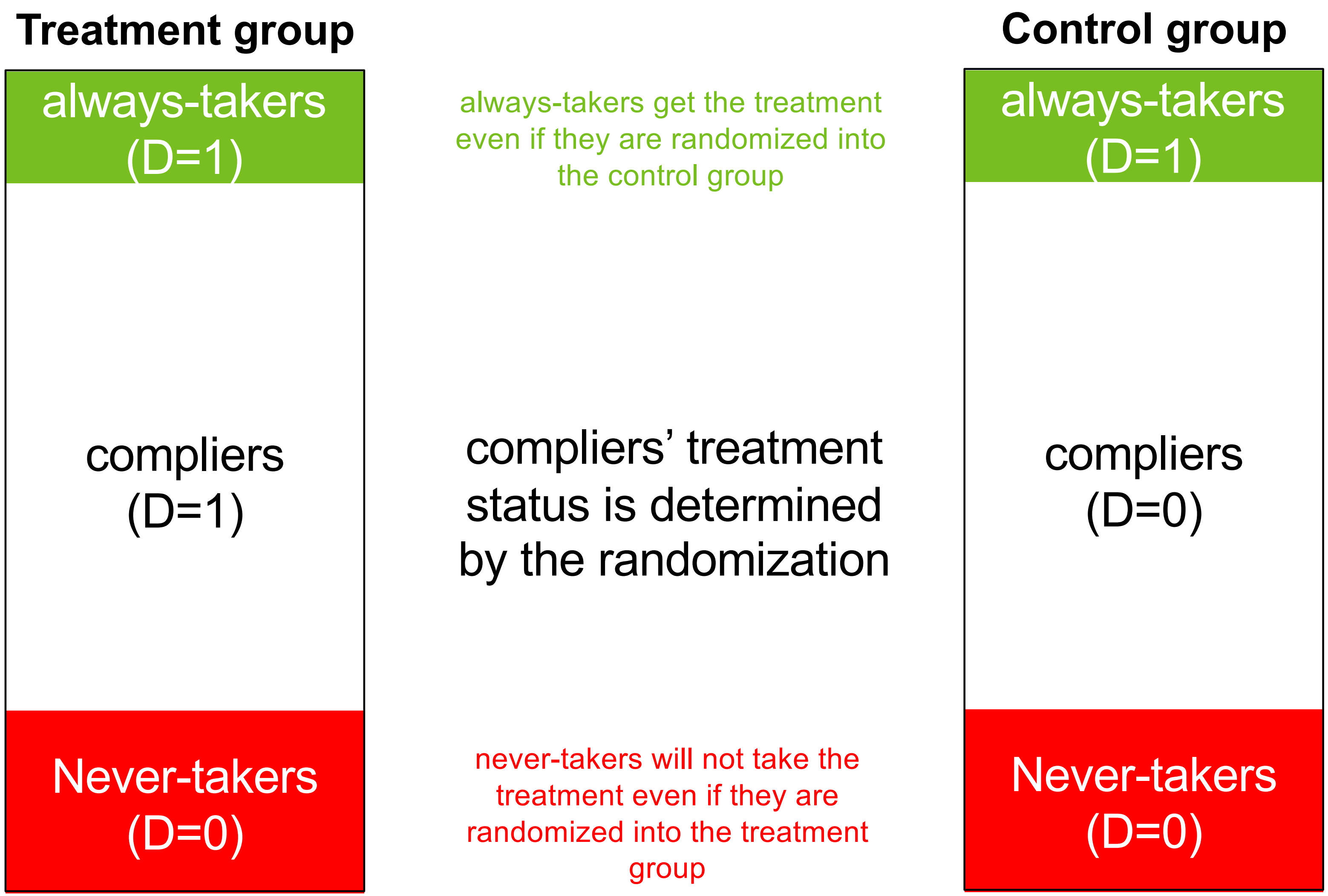
Compliance choice is potentially affected by potential outcomes

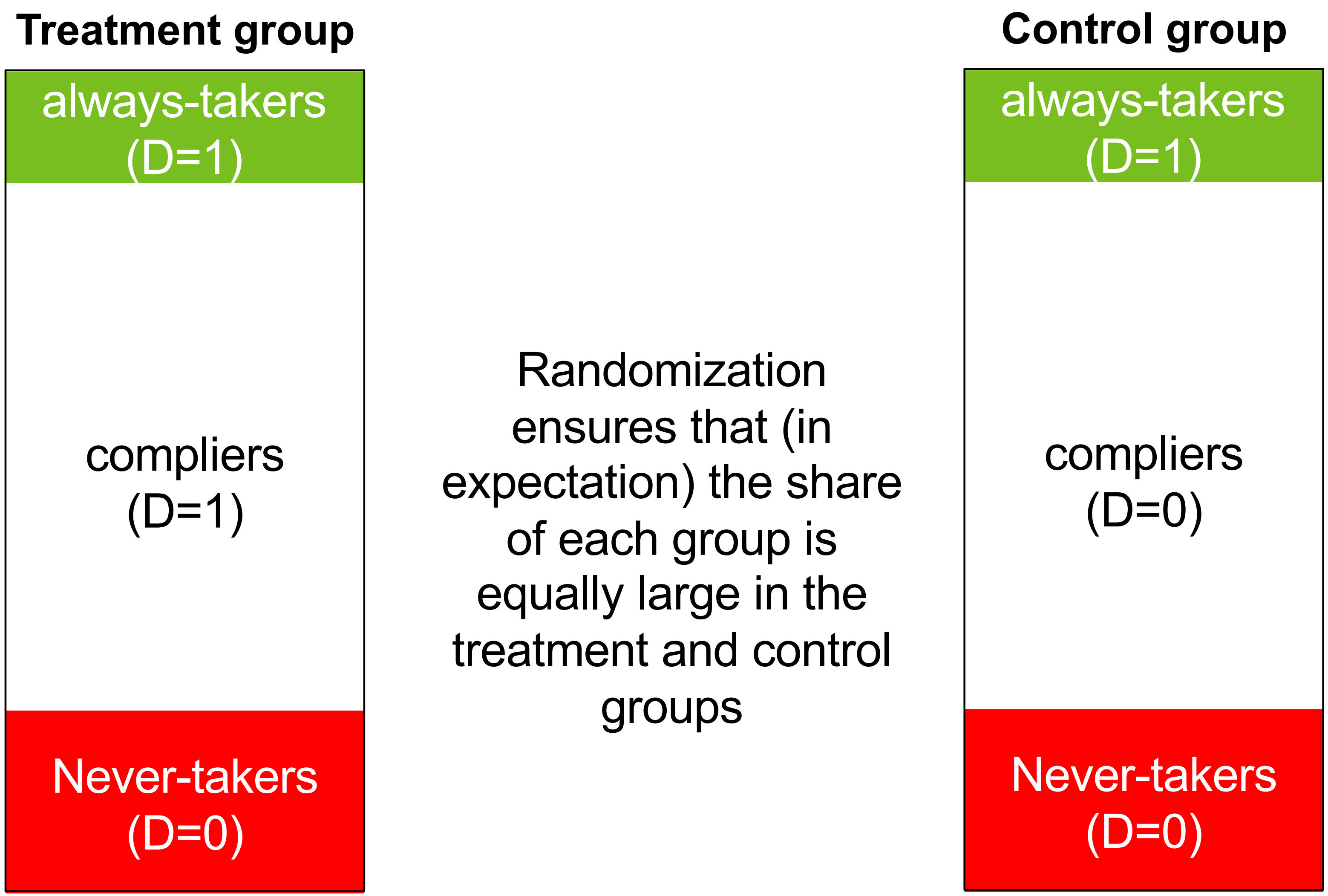
- e.g. those expecting to benefit the least become “never-takers”.

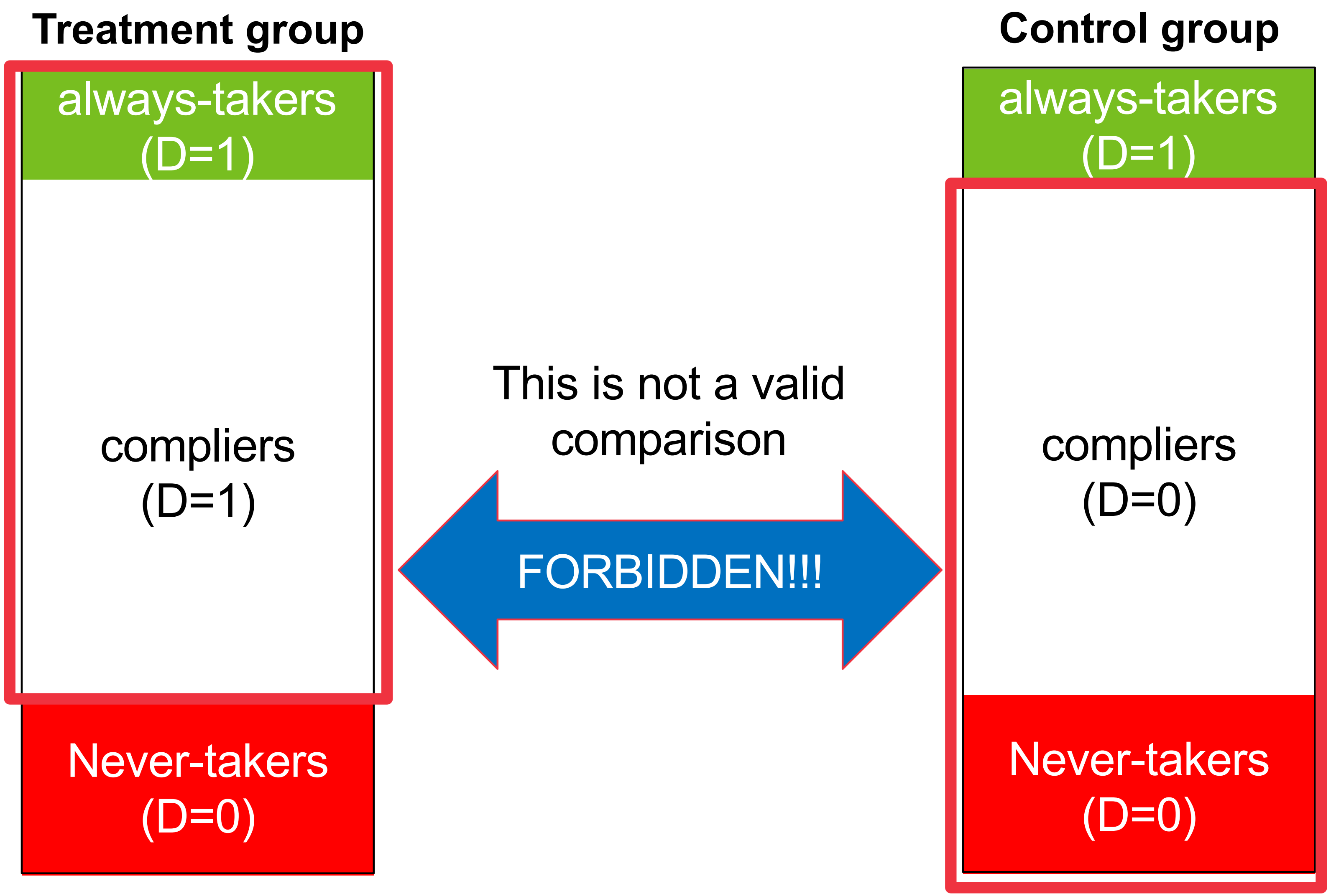
In that case comparing those who actually get the treatment to the entire control group is not a valid comparison!

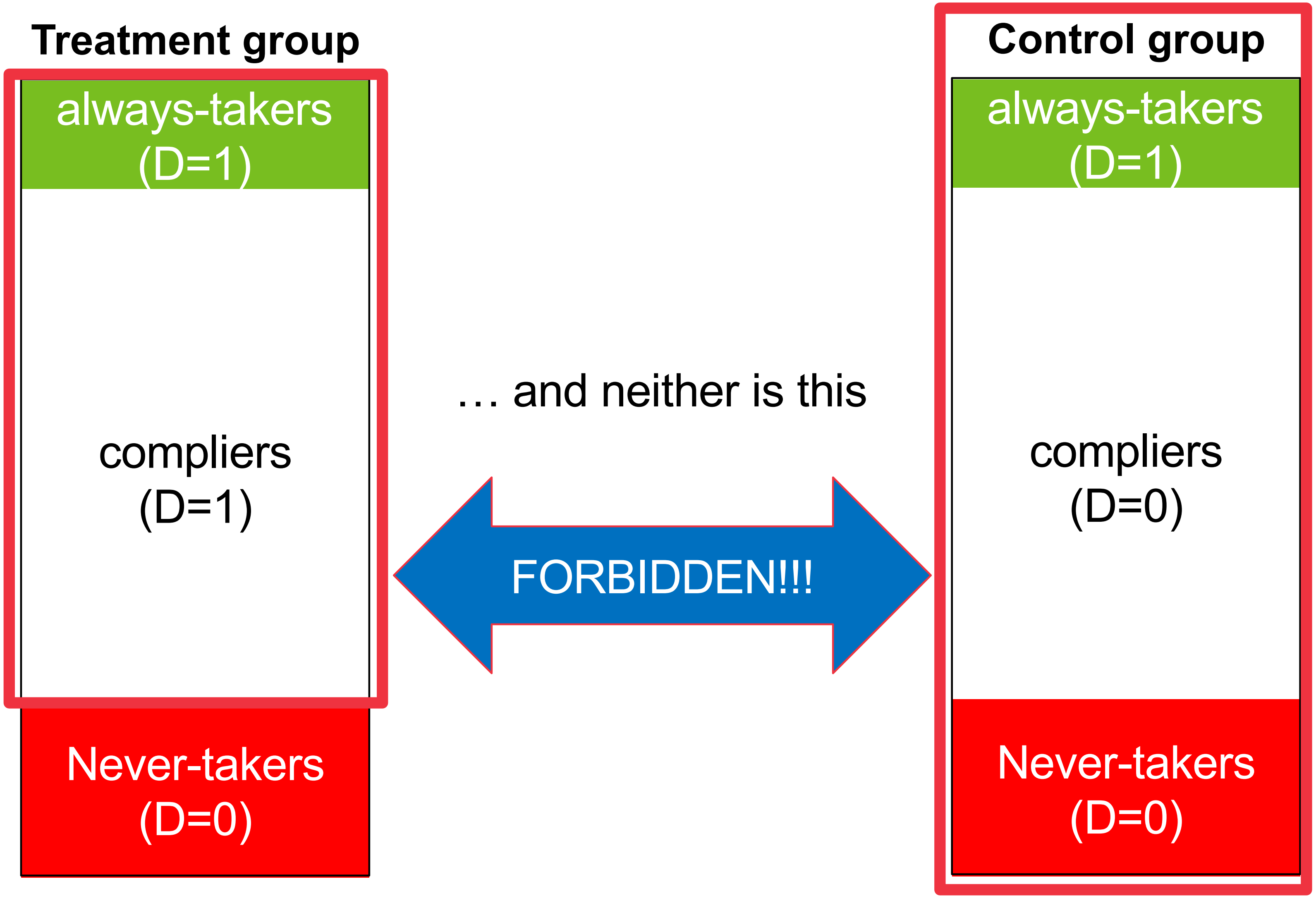
	Housing voucher take-up (1)
Panel A. Children < age 13 at random assignment	
Exp. versus control	$\boxed{47.66^{***}}$
	$\boxed{(1.653)}$
Sec. 8 versus control	$\boxed{65.80^{***}}$
	$\boxed{(1.934)}$
Observations	5,044
Control group mean	0
Panel B. Children age 13–18 at random assignment	
Exp. versus control	40.15***
	(2.157)
Sec. 8 versus control	55.04***
	(2.537)
Observations	2,358
Control group mean	0

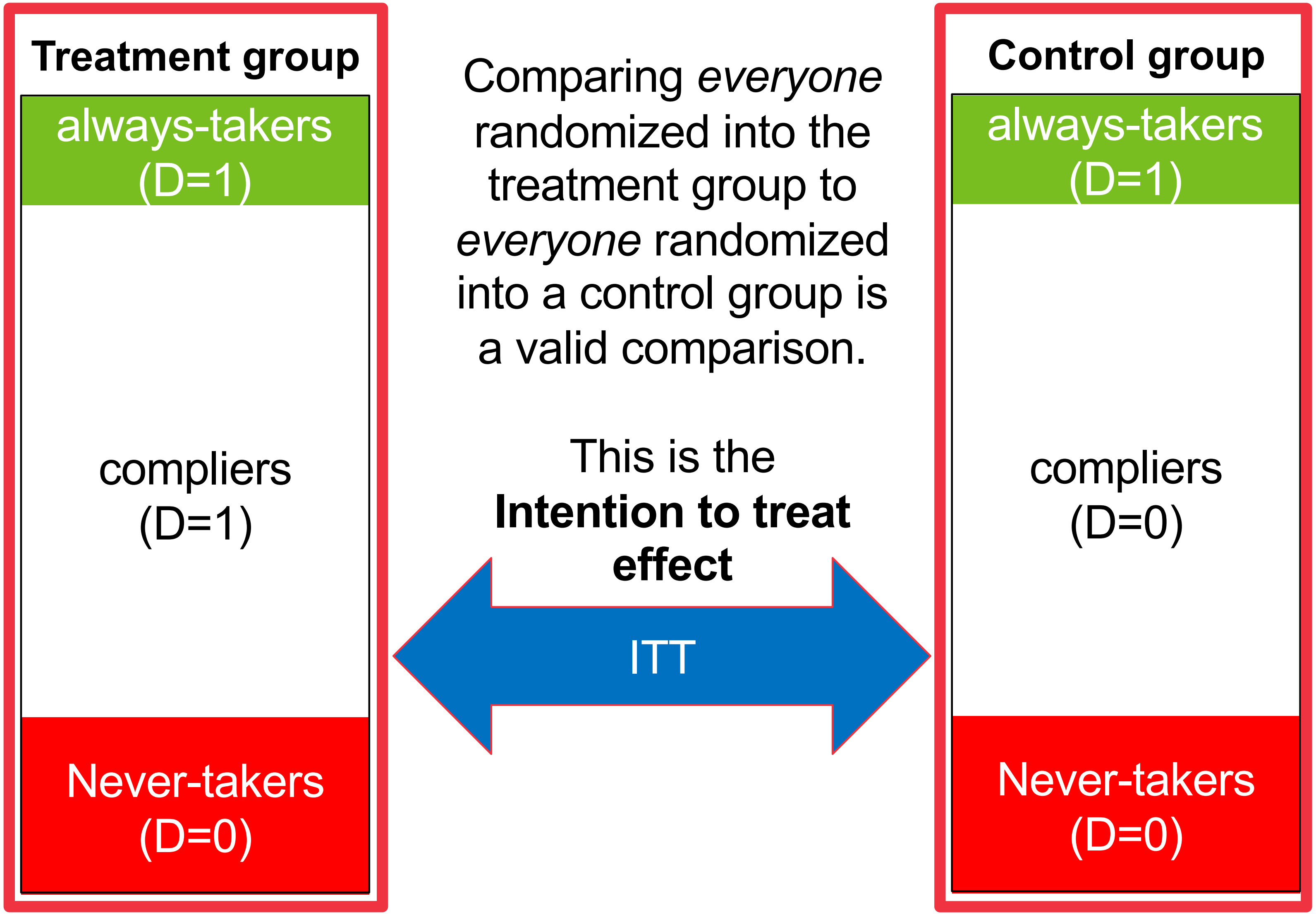
Aalto University
School of Business

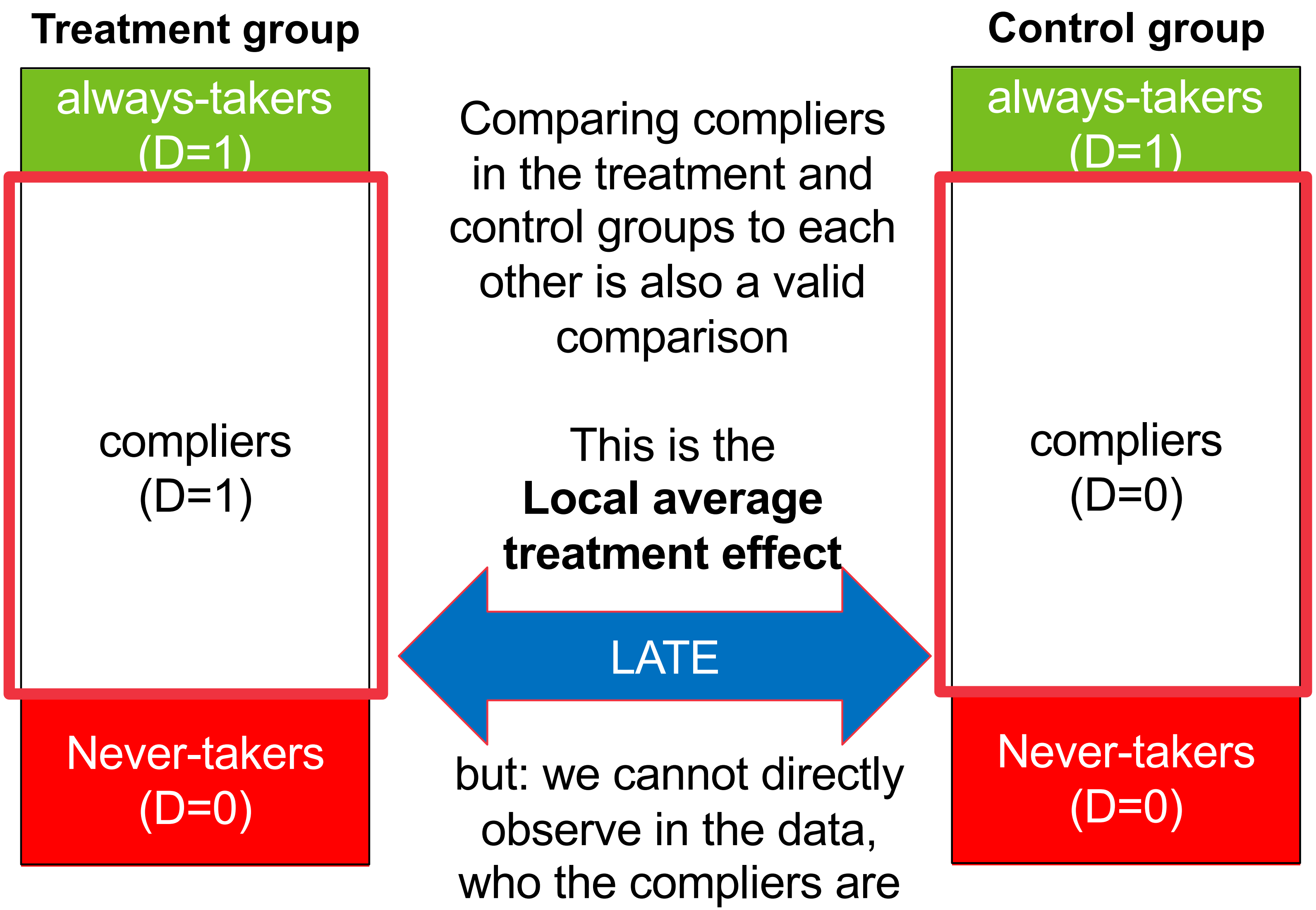

Treatment group
always-takers (D=1)
compliers (D=1)
Never-takers (D=0)
always-takers get the treatment even if they are randomized into the control group
compliers' treatment status is determined by the randomization
never-takers will not take the treatment even if they are randomized into the treatment group
Control group
always-takers (D=1)
compliers (D=0)
Never-takers (D=0)

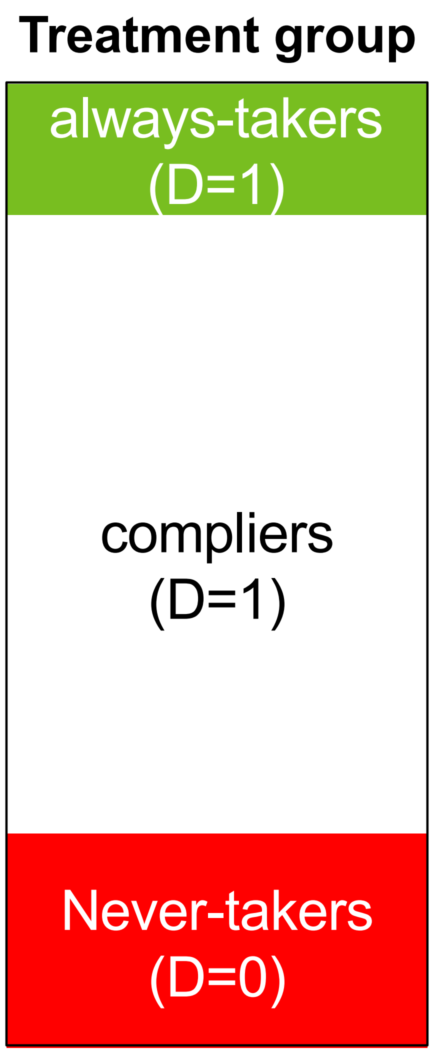

Treatment group
always-takers (D=1)
compliers (D=1)
Never-takers (D=0)
Randomization ensures that (in expectation) the share of each group is equally large in the treatment and control groups
Control group
always-takers (D=1)
compliers (D=0)
Never-takers (D=0)

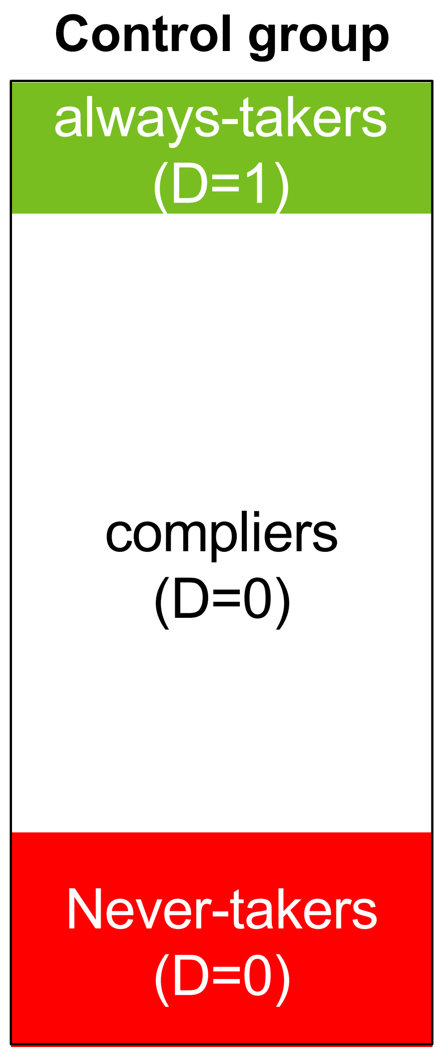

Treatment group
always-takers
(D=1)
compliers
(D=1)
Never-takers
(D=0)
This is not a valid
comparison
FORBIDDEN!!!
Control group
always-takers
(D=1)
compliers
(D=0)
Never-takers
(D=0)


Treatment group
always-takers
(D=1)
compliers
(D=1)
Never-takers
(D=0)
… and neither is this
FORBIDDEN!!!
Control group
always-takers
(D=1)
compliers
(D=0)
Never-takers
(D=0)

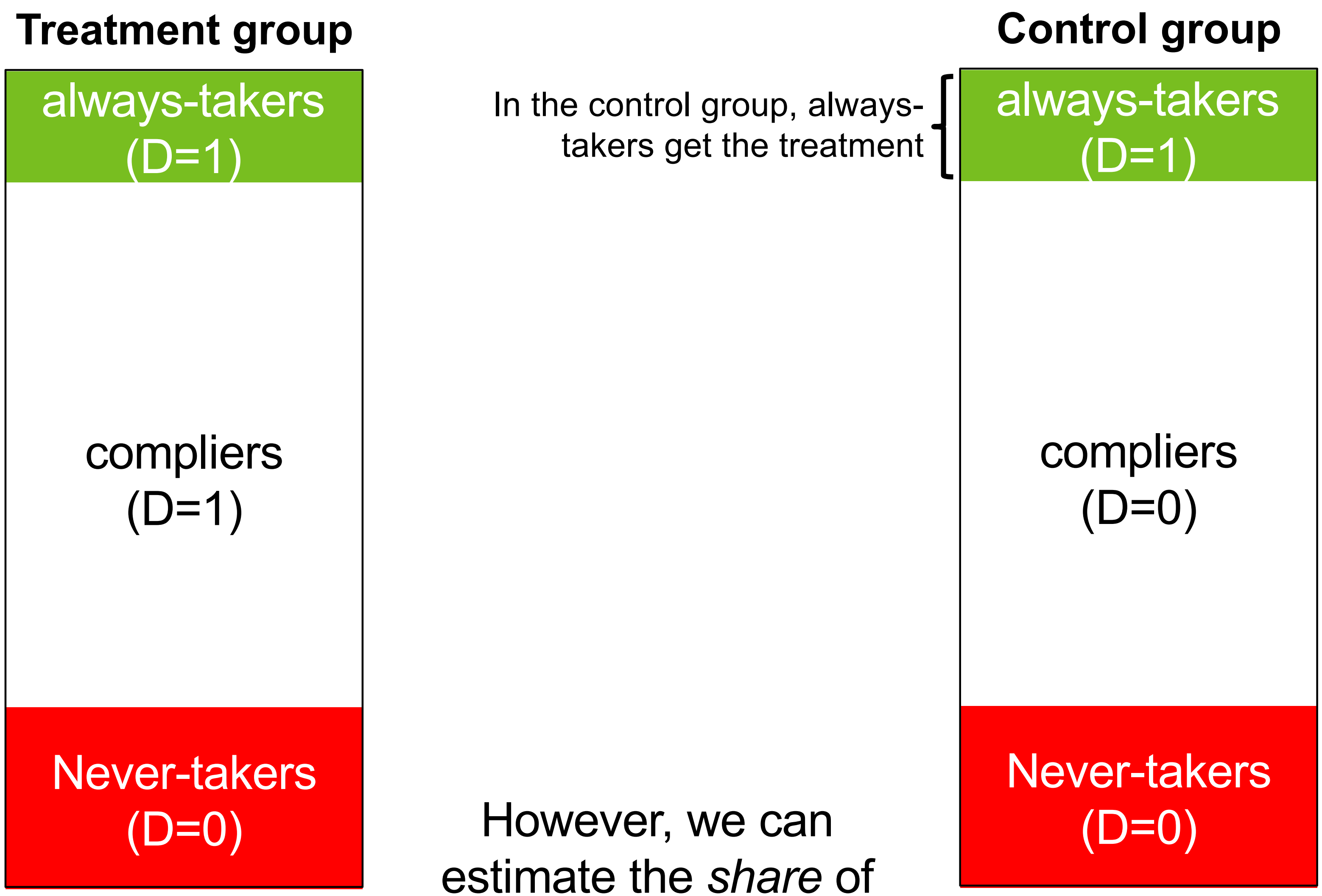

Treatment group
always-takers
(D=1)
compliers
(D=1)
Never-takers
(D=0)
Comparing *everyone* randomized into the treatment group to *everyone* randomized into a control group is a valid comparison.
This is the **Intention to treat effect**
ITT
Control group
always-takers
(D=1)
compliers
(D=0)
Never-takers
(D=0)


Treatment group
always-takers
(D=1)
compliers
(D=1)
Never-takers
(D=0)
Control group
always-takers
(D=1)
compliers
(D=0)
Never-takers
(D=0)
Comparing compliers in the treatment and control groups to each other is also a valid comparison
This is the **Local average treatment effect**
LATE
but: we cannot directly observe in the data, who the compliers are

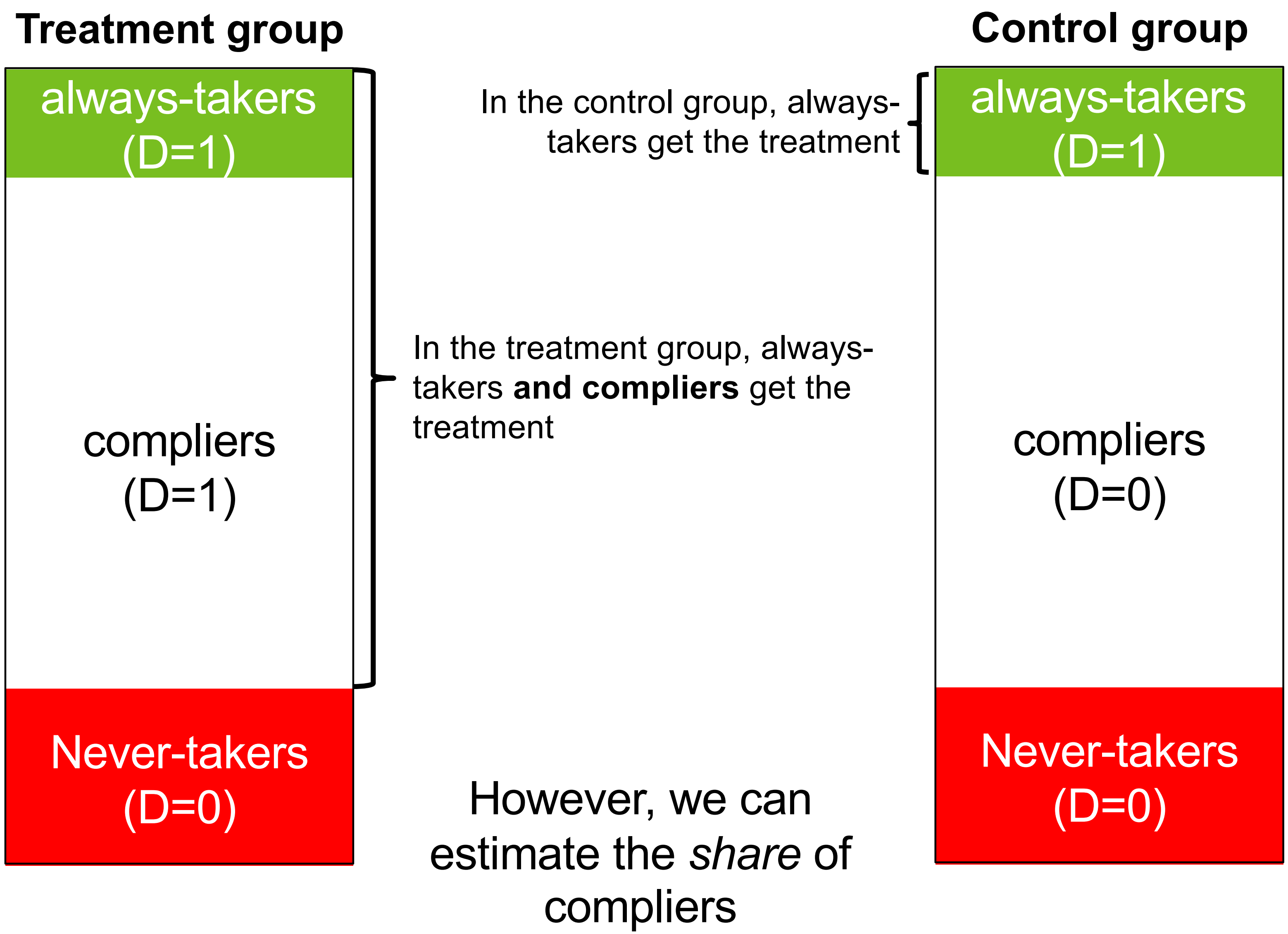




However, we can estimate the *share* of compliers



However, we can estimate the *share* of compliers


Treatment group
always-takers (D=1)
compliers (D=1)
Never-takers (D=0)
Control group
always-takers (D=1)
compliers (D=0)
Never-takers (D=0)
In the control group, always-takers get the treatment
In the treatment group, always-takers **and compliers** get the treatment
However, we can estimate the *share* of compliers

Treatment group (Z=1)

always-takers (D=1)

compliers (D=1)

Never-takers (D=0)

In the control group, always-takers get the treatment

In the treatment group, always-takers **and compliers** get the treatment

$$\boxed{\begin{array}{c}\mathbf{E[D|Z=1] - E[D|Z=0]} \\ \mathbf{=} \\ \mathbf{Pr(complier)}\end{array}}$$

However, we can estimate the *share* of compliers

Control group (Z=0)

always-takers (D=1)

compliers (D=0)

Never-takers (D=0)

Treatment group (Z=1)

always-takers
(D=1)

compliers
(D=1)

Never-takers
(D=0)

Control group (Z=0)

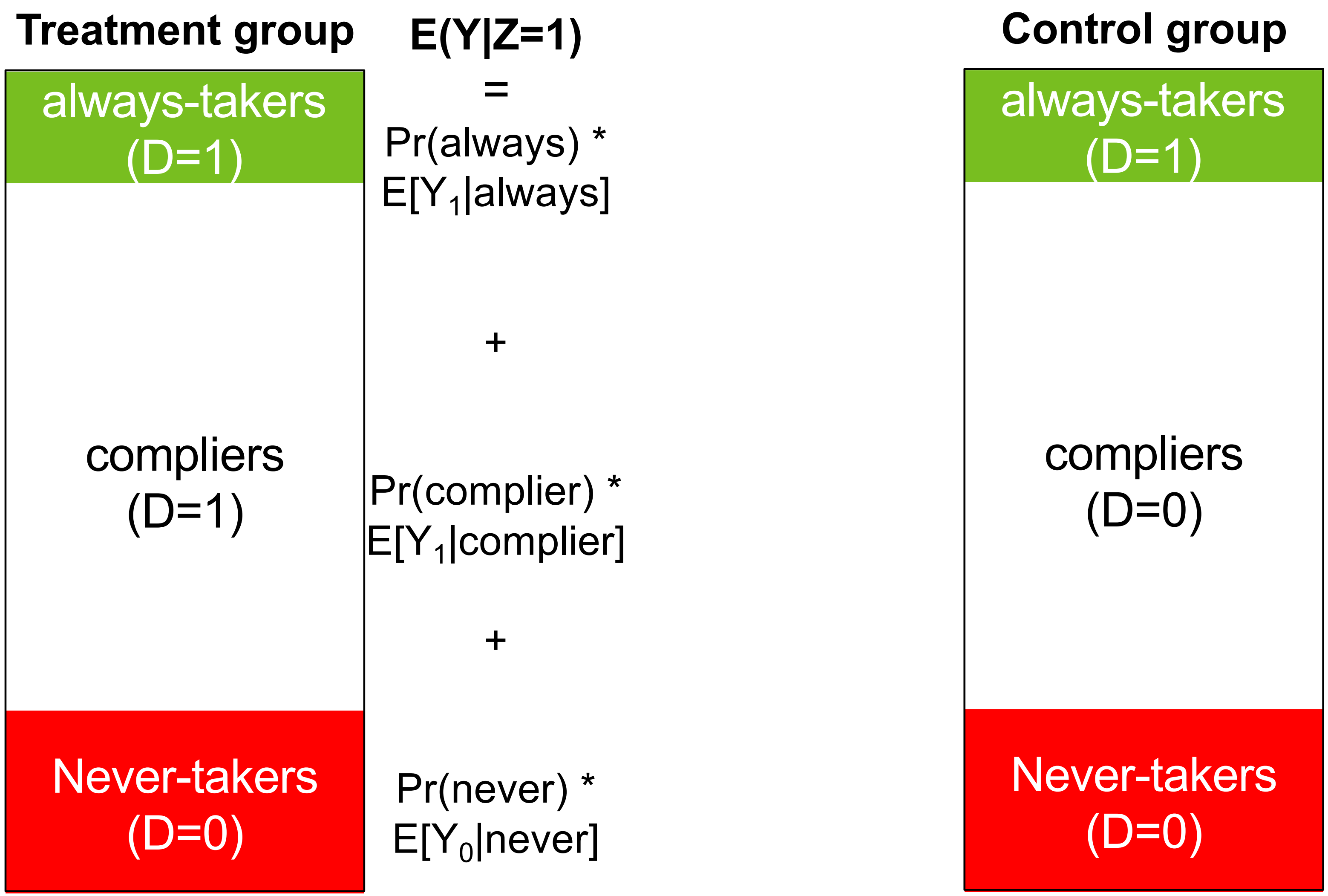
always-takers
(D=1)

compliers
(D=0)

Never-takers
(D=0)

Let’s denote the expected outcome of the treatment group as E[Y|Z=1], where Z denotes randomization status.

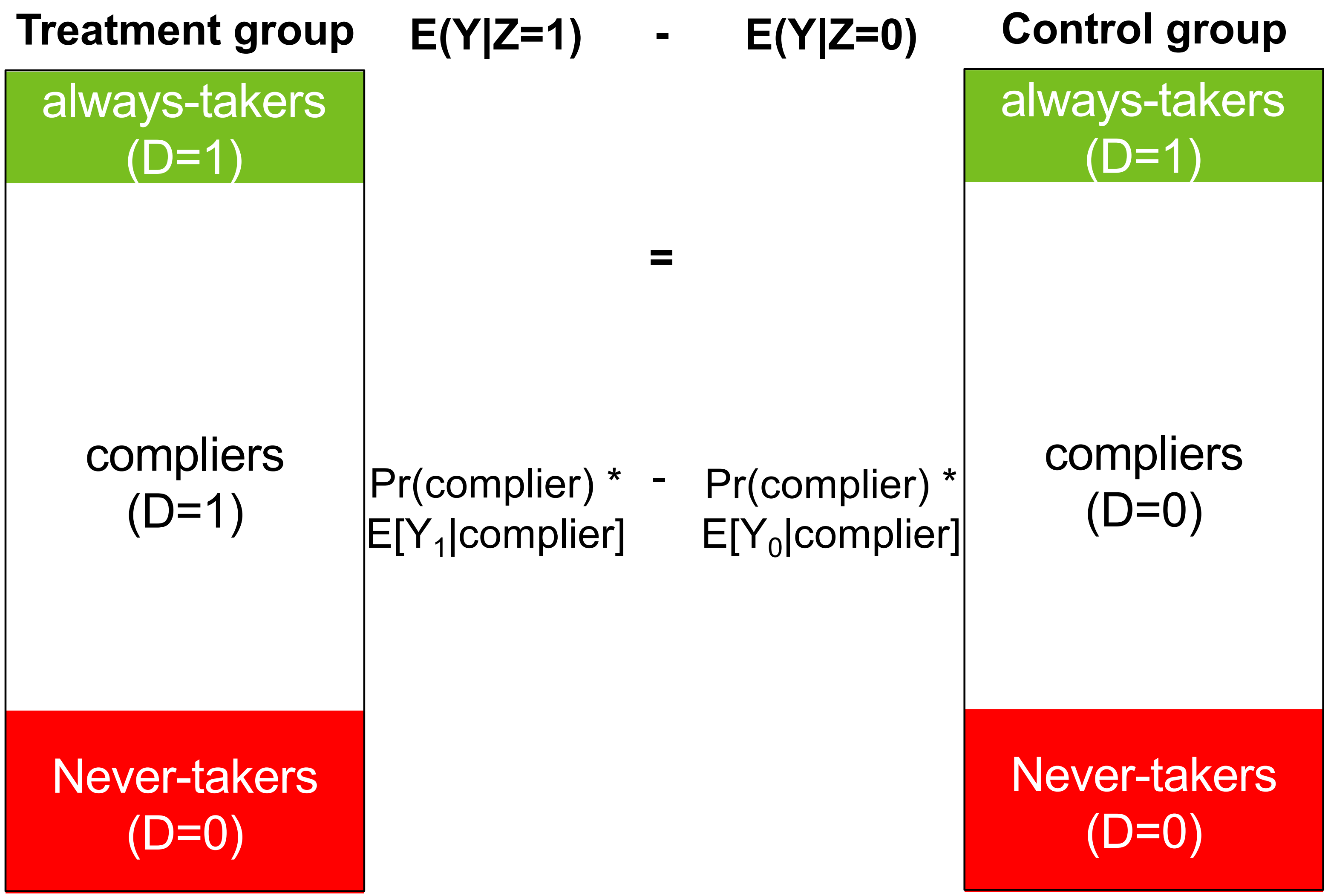
This is just the weighted average of the expectations among the always takers, compliers and never-takers in the treatment group, where the weights correspond to the shares of each group.

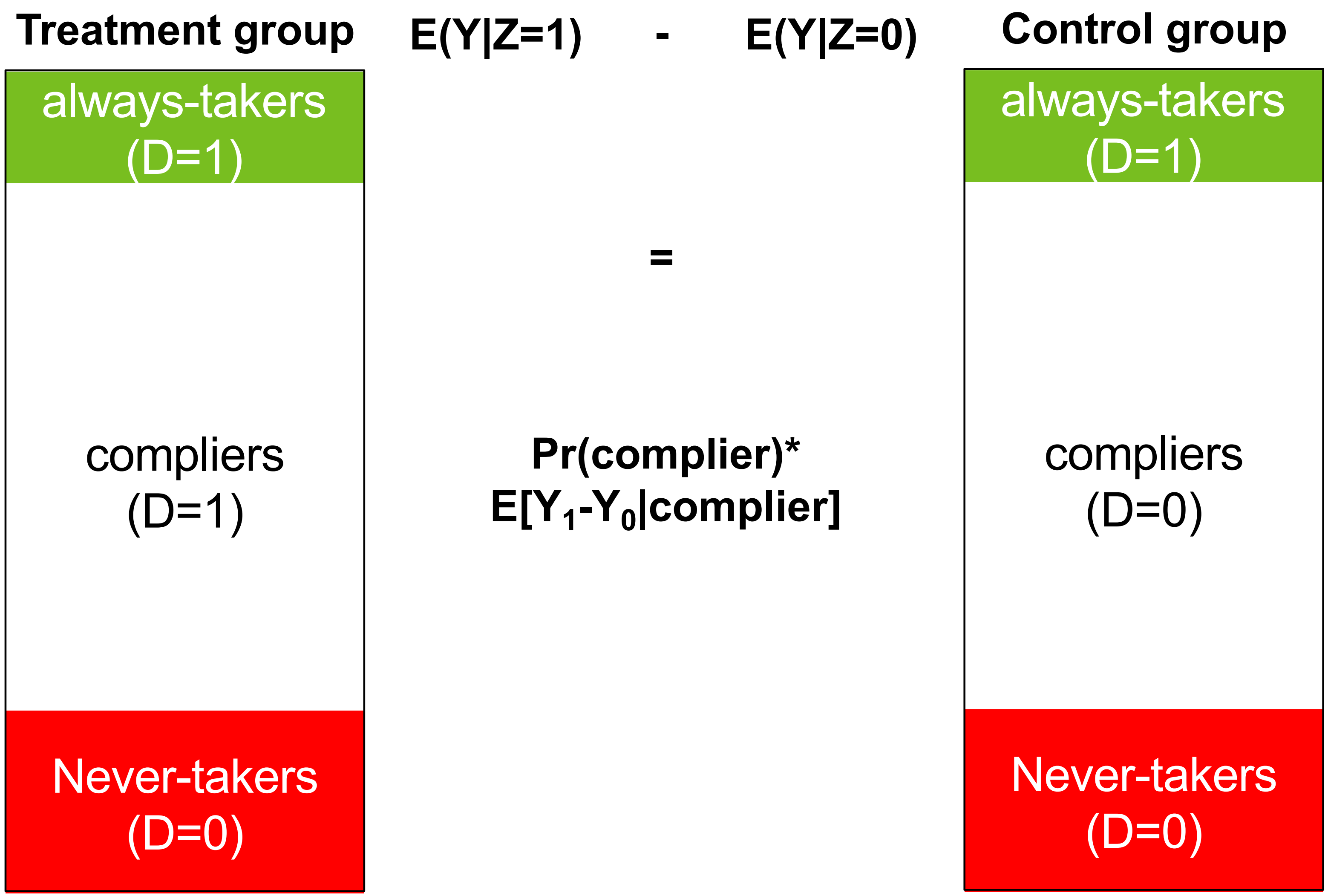


Treatment group
E(Y|Z=1)
=
Pr(always) *
E[Y1|always]
+
Pr(complier) *
E[Y1|complier]
+
Pr(never) *
E[Y0|never]
always-takers
(D=1)
compliers
(D=1)
Never-takers
(D=0)
Control group
always-takers
(D=1)
compliers
(D=0)
Never-takers
(D=0)

Treatment group	**$E(Y \mid Z=1)$**	**$E(Y \mid Z=0)$**	**Control group**
always-takers (D=1)	= Pr(always) * $E[Y_1 \mid \text{always}]$	= Pr(always) * $E[Y_1 \mid \text{always}]$	always-takers (D=1)
	+	+	
compliers (D=1)	Pr(complier) * $E[Y_1 \mid \text{complier}]$	Pr(complier) * $E[Y_0 \mid \text{complier}]$	compliers (D=0)
	+	+	
Never-takers (D=0)	Pr(never) * $E[Y_0 \mid \text{never}]$	Pr(never) * $E[Y_0 \mid \text{never}]$	Never-takers (D=0)

Treatment group	$E(Y\|Z=1)$		$E(Y\|Z=0)$	Control group
always-takers (D=1)	Pr(always) * $E[Y_1\|\text{always}]$	=	Pr(always) * $E[Y_1\|\text{always}]$	always-takers (D=1)
compliers (D=1)	Pr(complier) * $E[Y_1\|\text{complier}]$	≠	Pr(complier) * $E[Y_0\|\text{complier}]$	compliers (D=0)
Never-takers (D=0)	Pr(never) * $E[Y_0\|\text{never}]$	=	Pr(never) * $E[Y_0\|\text{never}]$	Never-takers (D=0)


Treatment group
always-takers (D=1)
compliers (D=1)
Never-takers (D=0)
$E(Y|Z=1) - E(Y|Z=0)$
=
$Pr(complier) * E[Y_1|complier] - Pr(complier) * E[Y_0|complier]$
Control group
always-takers (D=1)
compliers (D=0)
Never-takers (D=0)


Treatment group
E(Y|Z=1)
-
E(Y|Z=0)
Control group
always-takers (D=1)
always-takers (D=1)
=
compliers (D=1)
Pr(complier)*
E[Y_1-Y_0|complier]
compliers (D=0)
Never-takers (D=0)
Never-takers (D=0)

Treatment group

always-takers
(D=1)

compliers
(D=1)

Never-takers
(D=0)

Solve for $\mathbf{E[Y_1\text{-}Y_0|complier]}$:

$$\frac{\mathbf{Pr(complier)^* \; E[Y_1\text{-}Y_0|complier]}}{\mathbf{Pr(complier)}}$$

$$\mathbf{=}$$

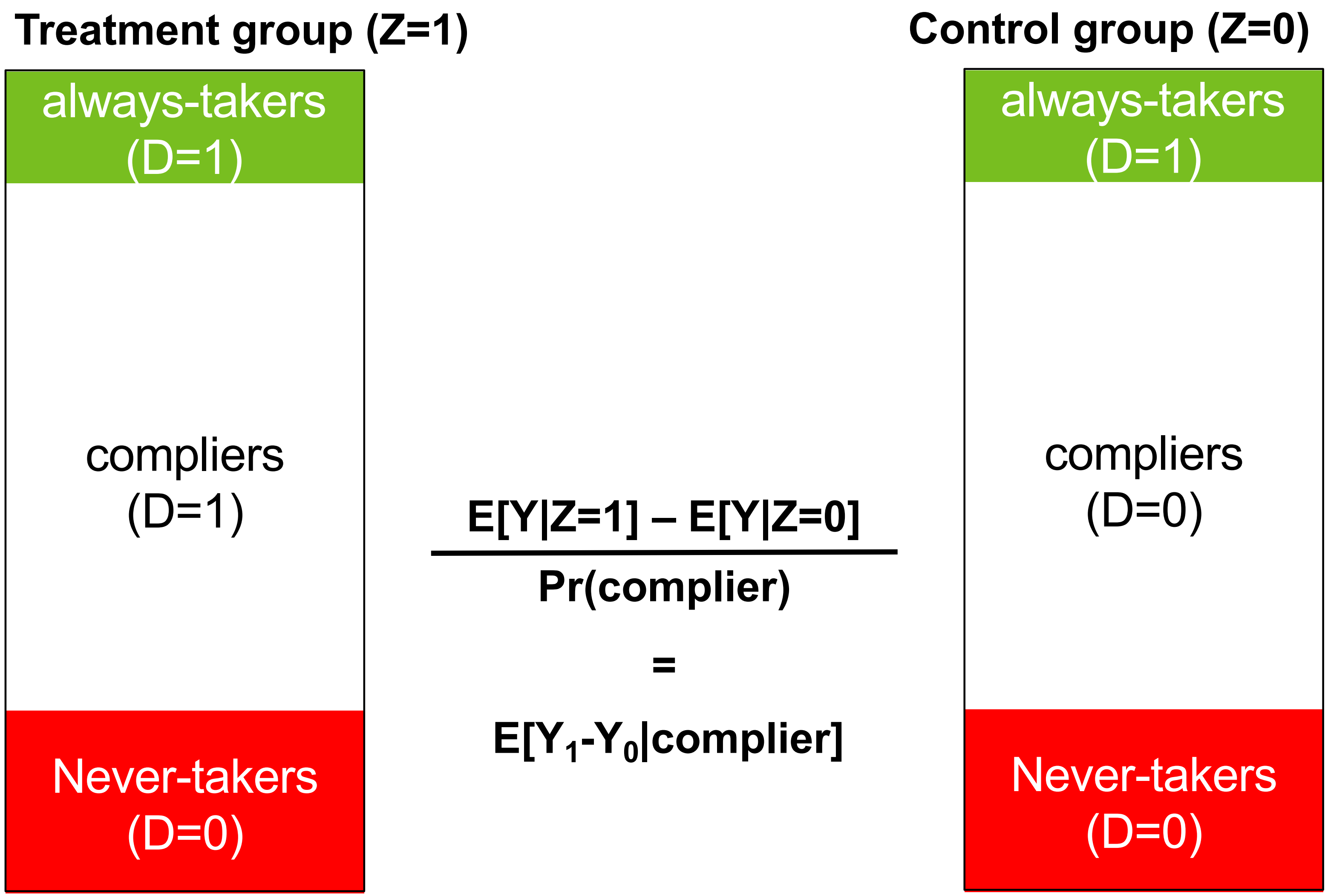
$$\mathbf{E[Y_1\text{-}Y_0|complier]}$$

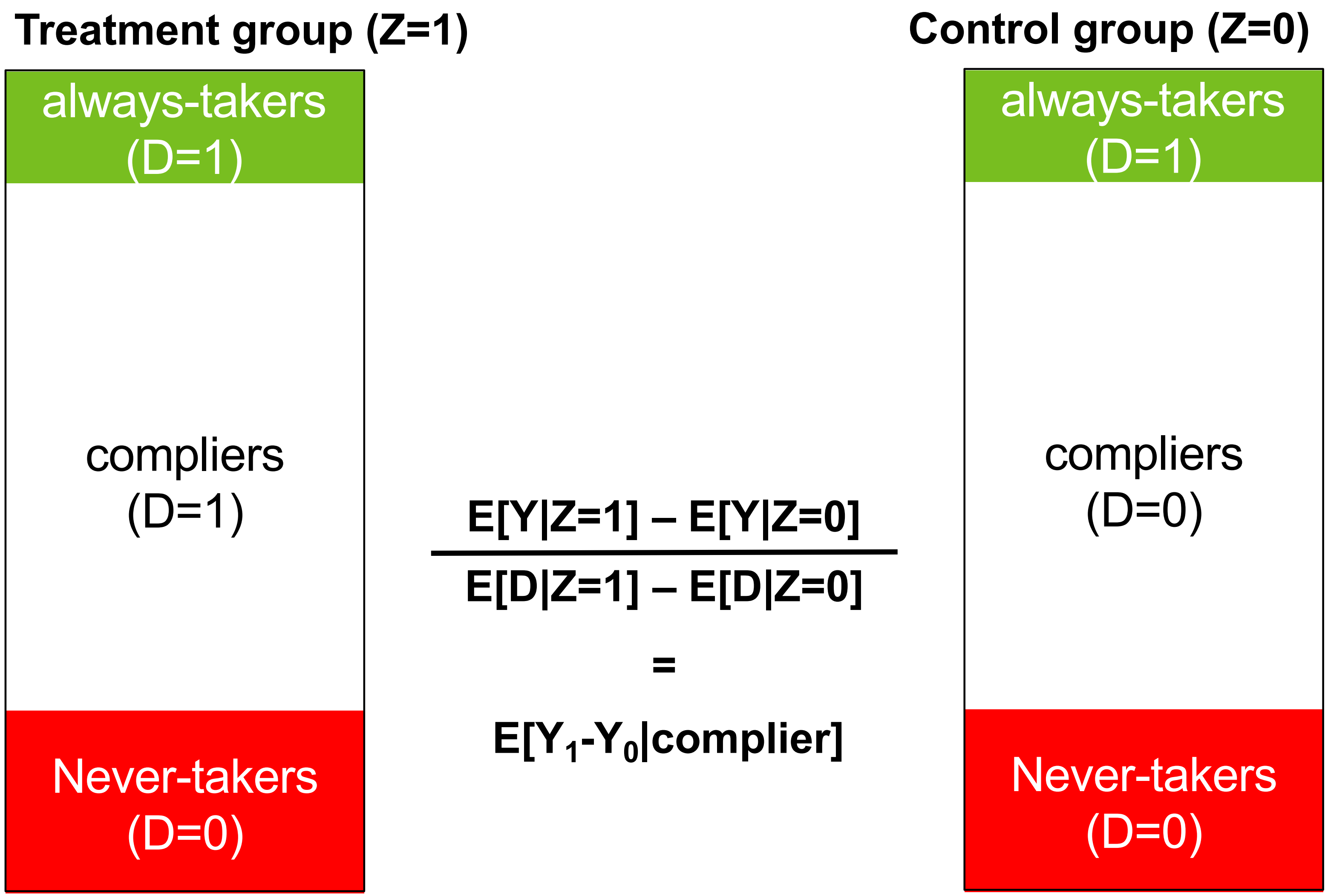
Control group

always-takers
(D=1)

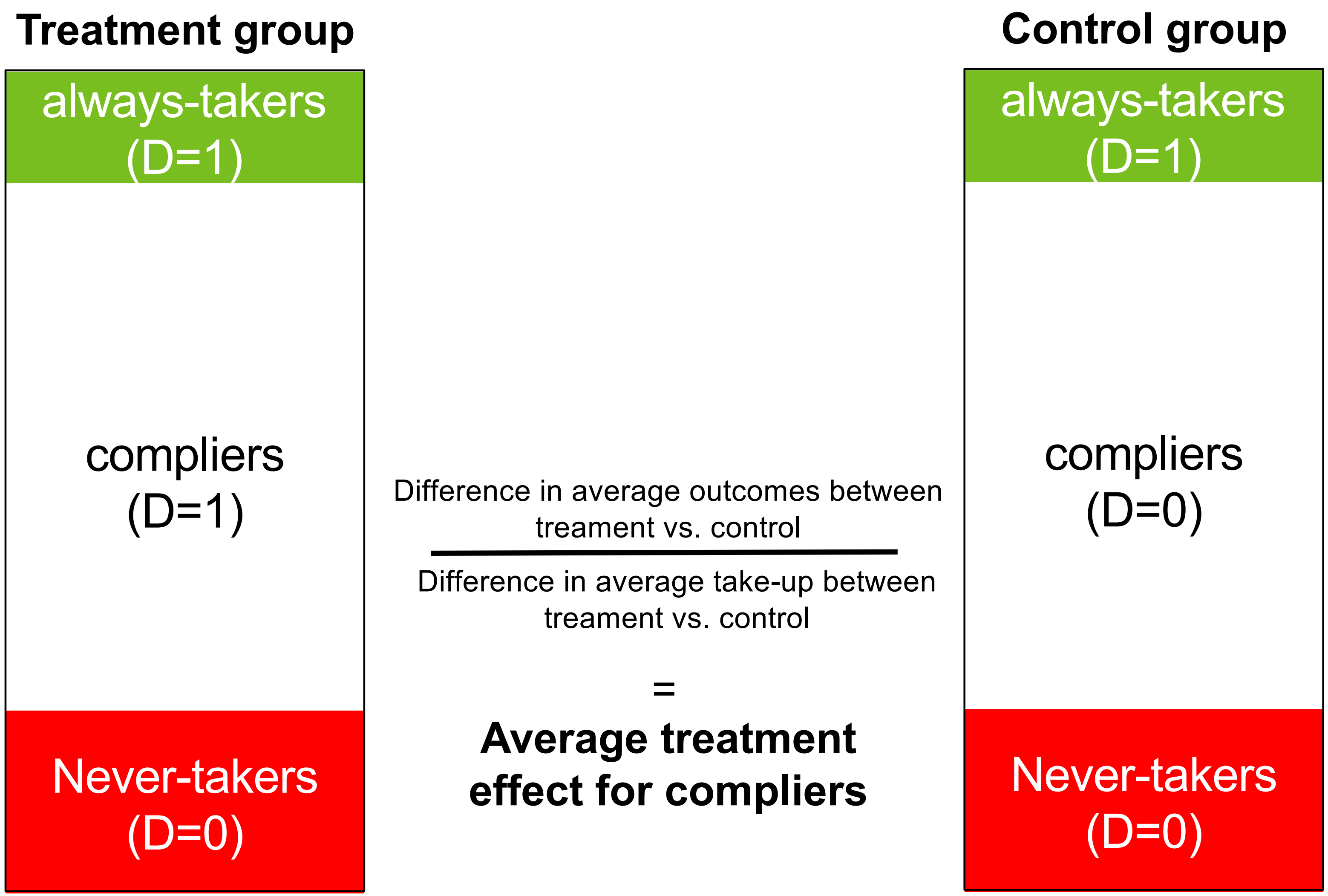
compliers
(D=0)

Never-takers
(D=0)


Treatment group (Z=1)
always-takers (D=1)
compliers (D=1)
Never-takers (D=0)
Control group (Z=0)
always-takers (D=1)
compliers (D=0)
Never-takers (D=0)
$\frac{E[Y|Z=1] - E[Y|Z=0]}{Pr(complier)} = E[Y_1 - Y_0|complier]$



$$\frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]} = E[Y_1 - Y_0|\text{complier}]$$


Treatment group
always-takers (D=1)
compliers (D=1)
Never-takers (D=0)
Control group
always-takers (D=1)
compliers (D=0)
Never-takers (D=0)
Difference in average outcomes between treament vs. control
Difference in average take-up between treament vs. control
=
Average treatment effect for compliers

Wald estimator

- **We just derived the *Wald estimator*!**

$$\boldsymbol{\beta_{LATE}} = \frac{\boldsymbol{E[Y|Z=1] - E[Y|Z=0]}}{\boldsymbol{E[D|Z=1] - E[D|Z=0]}}$$

ITT

Share compliers

- ***Y*** is the **outcome**
- **Z** is an indicator (or dummy variable) for being randomized or **allocated into the treatment group**
- **D** is an indicator (or dummy variable) for actually **receiving the treatment**

Components of the Wald estimator

- the numerator is the intention to treat effect (ITT)
- the denominator is the share of compliers (first-stage)

$\beta_{LATE} = E[Y_1 - Y_0|complier]$ is the *local average treatment effect*

- *Importantly, the impact of receiving the treatment for the compliers may differ from the impact on never-takers and always-takers*

This is one version of the **instrumental variables (IV) estimators** which we will discuss this week.

ITT, LATE and ATT

Sometimes ITT is the most relevant estimate

- in the context of the MTO, it is the impact of being offered housing vouchers
- this is arguably the most relevant effect given that offering vouchers is likely to be the relevant policy (rather than forcing everyone to move)

Sometimes LATE is more relevant

- in Moving to Opportunity (MTO), it is the impact of living in better neighborhoods
- potentially informative for policy discussion on whether we should invest in improving existing neighborhoods (”place-making policies”)

LATE informs us only about the impact on compliers

- usefulness depends on how representative the compliers are
- when there are no always-takers LATE = ATT (average treatment effect on the treated)
- this is the case in MTO, thus the tables report ”TOT” (same as ATT)*

*ATT (Average Treatmente Effect on the Treated is also referred to as ToT or TOT (Treatment on the Treated)

Back to MTO

- We now have everything we need to understand the newest MTO results!

TABLE 3—IMPACTS OF MTO ON CHILDREN'S INCOME IN ADULTHOOD

	W-2 earnings ($) 2008–2012 ITT (1)	Individual earnings 2008–2012 ($) ITT (2)	ITT w/ controls (3)	TOT (4)
Panel A. Children < age 13 at random assignment				
Exp. versus control		1,624.0**		3,476.8**
		(662.4)		(1,418.2)
Sec. 8 versus control		1,109.3		1,723.2
		(676.1)		(1051.5)
Observations		8,420		8,420
Control group mean		11,270.3		11,270.3

TOT
=
ITT / First stage
=
$1,624/.467
=
$3,476.8

Chetty, Hendren, Katz (2016): The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment

TABLE 3—IMPACTS OF MTO ON CHILDREN'S INCOME IN ADULTHOOD

	W-2 earnings ($) 2008–2012 ITT (1)	Individual earnings 2008–2012 ($) ITT (2)	ITT w/ controls (3)	TOT (4)
Panel A. Children < age 13 at random assignment				
Exp. versus control	1,339.8**	1,624.0**	1,298.9**	3,476.8**
	(671.3)	(662.4)	(636.9)	(1,418.2)
Sec. 8 versus control	687.4	1,109.3	908.6	1,723.2
	(698.7)	(676.1)	(655.8)	(1051.5)
Observations	8,420	8,420	8,420	8,420
Control group mean	9,548.6	11,270.3	11,270.3	11,270.3

TOT
=
ITT / First stage
=
$1,624/.467
=
$3,476.8

Chetty, Hendren, Katz (2016): The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment

TABLE 3—IMPACTS OF MTO ON CHILDREN'S INCOME IN ADULTHOOD

	W-2 earnings ($) 2008–2012 ITT (1)	Individual earnings 2008–2012 ($) ITT (2)	ITT w/ controls (3)	TOT (4)
Panel A. Children < age 13 at random assignment				
Exp. versus control	1,339.8**	1,624.0**	1,298.9**	3,476.8**
	(671.3)	(662.4)	(636.9)	(1,418.2)
Sec. 8 versus control	687.4	1,109.3	908.6	1,723.2
	(698.7)	(676.1)	(655.8)	(1051.5)
Observations	8,420	8,420	8,420	8,420
Control group mean	9,548.6	11,270.3	11,270.3	11,270.3
Panel B. Children age 13–18 at random assignment				
Exp. versus control	−761.2	−966.9	−879.5	−2,426.7
	(870.6)	(854.3)	(817.3)	(2,154.4)
Sec. 8 versus control	−1,048.9	−1,132.8	−1,136.9	−2,051.1
	(932.5)	(922.3)	(866.6)	(1,673.7)
Observations	11,623	11,623	11,623	11,623
Control group mean	13,897.1	15,881.5	15,881.5	15,881.5

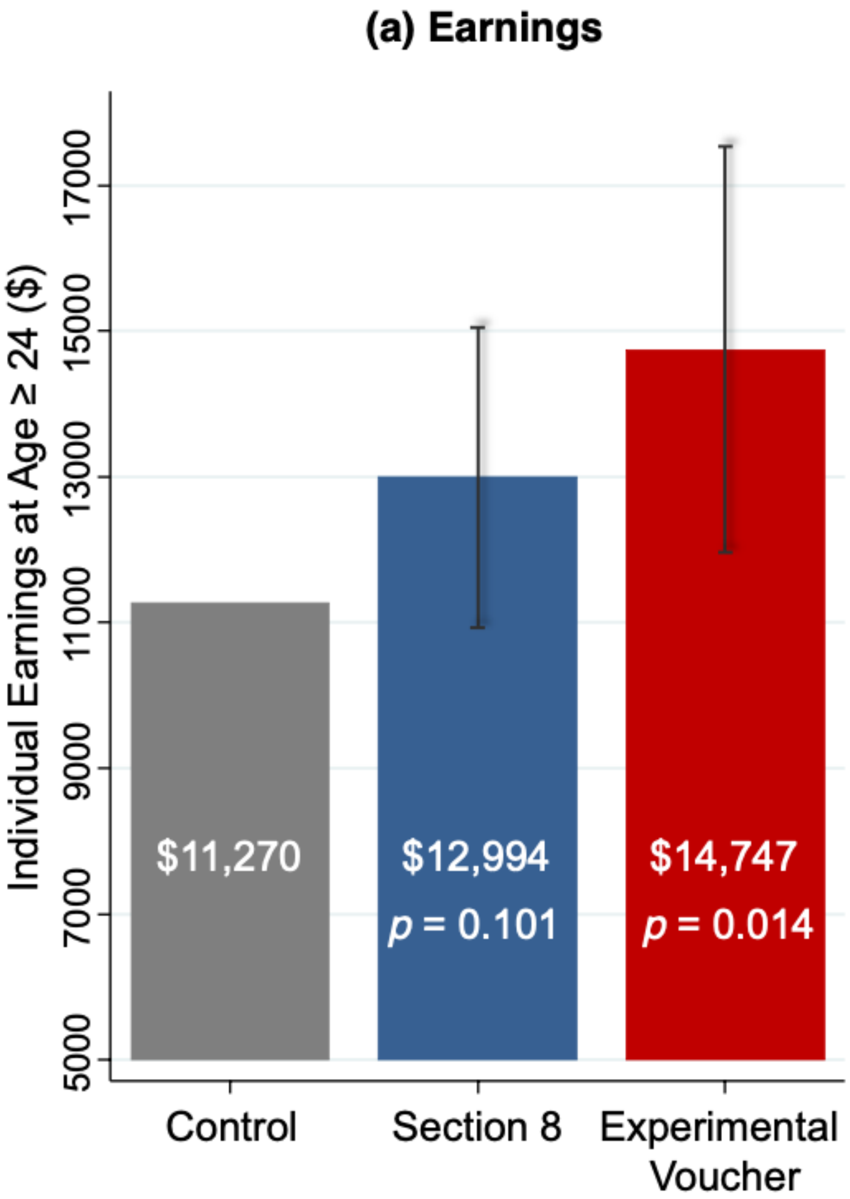
Chetty, Hendren, Katz (2016): The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment

TABLE 3—IMPACTS OF MTO ON CHILDREN'S INCOME IN ADULTHOOD

	W-2 earnings ($) 2008–2012 ITT (1)	Individual earnings 2008–2012 ($)			Individual earnings ($)		Employed (%) 2008–2012 ITT (7)	Hhold. inc. ($) 2008–2012 ITT (8)	Inc. growth ($) 2008–2012 ITT (9)
		ITT (2)	ITT w/ controls (3)	TOT (4)	Age 26 ITT (5)	2012 ITT (6)			
Panel A. Children < age 13 at random assignment									
Exp. versus control	1,339.8**	1,624.0**	1,298.9**	3,476.8**	1,751.4*	1,443.8**	1.824	2,231.1***	1,309.4**
	(671.3)	(662.4)	(636.9)	(1,418.2)	(917.4)	(665.8)	(2.083)	(771.3)	(518.5)
Sec. 8 versus control	687.4	1,109.3	908.6	1,723.2	551.5	1,157.7*	1.352	1,452.4**	800.2
	(698.7)	(676.1)	(655.8)	(1051.5)	(888.1)	(690.1)	(2.294)	(735.5)	(517.0)
Observations	8,420	8,420	8,420	8,420	1,625	2,922	8,420	8,420	8,420
Control group mean	9,548.6	11,270.3	11,270.3	11,270.3	11,398.3	11,302.9	61.8	12,702.4	4,002.2
Panel B. Children age 13–18 at random assignment									
Exp. versus control	−761.2	−966.9	−879.5	−2,426.7	−539.0	−969.2	−2.173	−1,519.8	−693.6
	(870.6)	(854.3)	(817.3)	(2,154.4)	(795.4)	(1,122.2)	(2.140)	(11,02.2)	(571.6)
Sec. 8 versus control	−1,048.9	−1,132.8	−1,136.9	−2,051.1	−15.11	−869.0	−1.329	−936.7	−885.3
	(932.5)	(922.3)	(866.6)	(1,673.7)	(845.9)	(1213.3)	(2.275)	(11,85.9)	(625.2)
Observations	11,623	11,623	11,623	11,623	2,331	2,331	11,623	11,623	11,623
Control group mean	13,897.1	15,881.5	15,881.5	15,881.5	13,968.9	16,602.0	63.6	19,169.1	4,128.1

Chetty, Hendren, Katz (2016): The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment

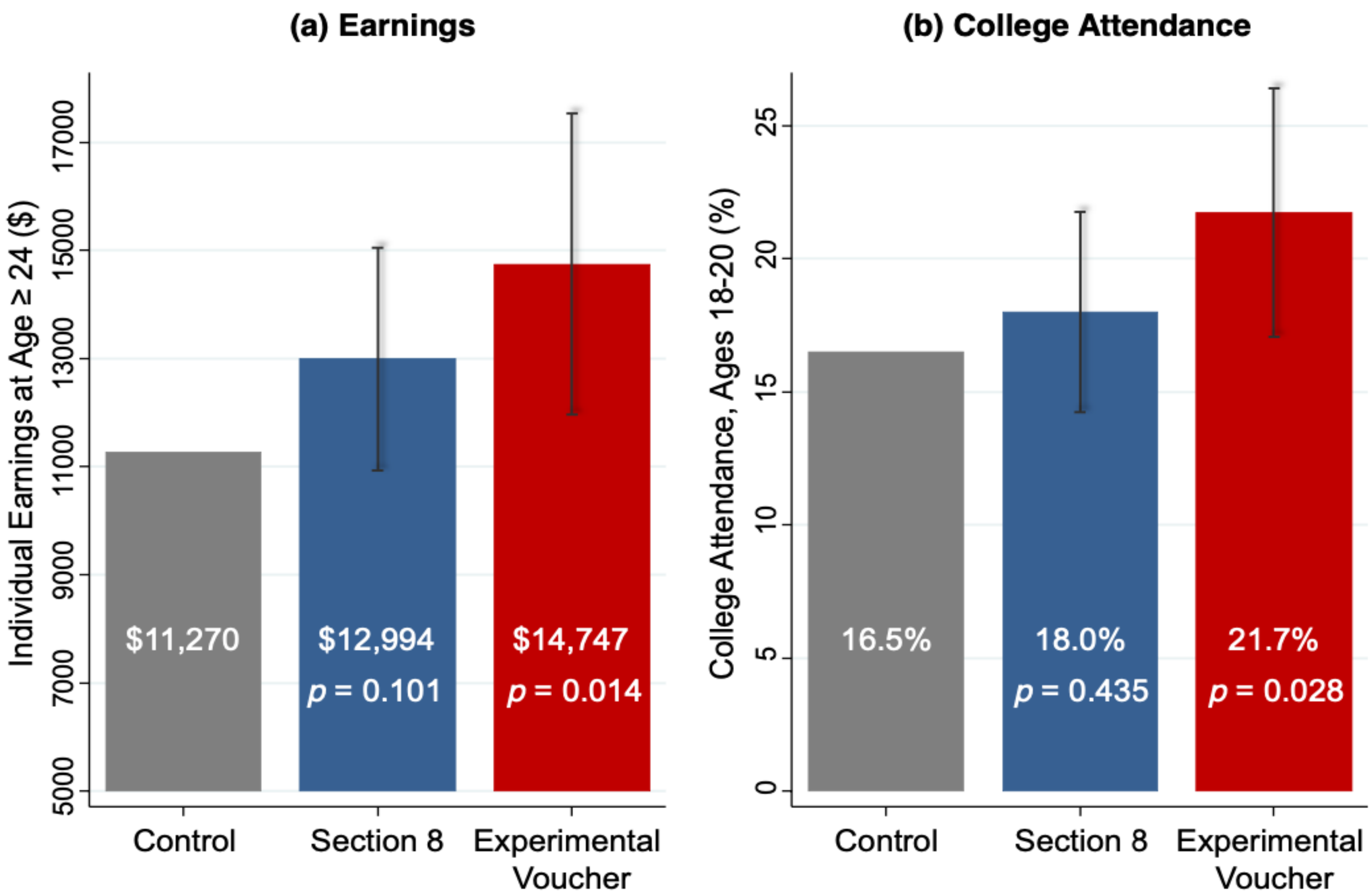
Impacts of MTO on Children Below Age 13 at Random Assignment



95% confidence intervals = [ToT-1.96*SE, ToT+1.96*SE] = [$11,967.4, $17,526,8 (for the experimental group)

Average income for the experimental group = baseline + ToT = $11,270 + $3,477 = $14,747

Impacts of MTO on Children Below Age 13 at Random Assignment



Take-aways from MTO follow up

Strong evidence on the existence of neighborhood effects

- might seem obvious, but hard evidence on them is scarce.

Putting the effects into context

- the average income of the participants remains far below average (even though it is much higher in comparison to the control group)
- **external validity**: would the effects be similar also in other contexts?
- Chyn 2018 (discussed in Lecture 7) supports the findings and arguably has higher external validity.

Methodological lesson: how to deal with partial compliance

- manipulation of the likelihood of being treated can take us a long way
- but: important to think about who the compliers are

Instrumental Variable approach

- **First published use 1928 in Appendix of a book by Philip G. Wright, "The Tariff on Animal and Vegetable Oils".**
  - In the example, the author (who may have been Philip G. Wright's son Sewall) shows that if one can find one variable that shifts demand but not supply, that variable can be used to determine the slope of the demand curve (Stock & Trebbi, 2003).
- **Became a popular tool among applied economists for isolating causality in the past two decades.**
- **Strong internal validity, but sometimes low external validity due to focus on a particular group**
  - LATE, compliers, who are we interested in for policy?

Instrumental Variable (IV) general idea

Task: answering a causal question: does T affect Y?

Recall: for many (most) causal questions we may think of in social sciences, selection problems and OVB will usually render this task complicated.

Why? There are probably factors that we cannot measure that are correlated with both the independent and the dependent variables.

- **Idea of an "Instrument": find something exogenous that we *can* measure, and that**

  - affects only T (whose effect on Y we want to estimate)
  - cannot affect the outcome Y directly

Instrumental Variable (IV) example

Let's exemplify with a causal question, for example: "does the amount or quality of schooling one gets affect ones' wages in the future?"

T: schooling   Y: later wages

A naïve comparison between individuals with fewer vs. more years of schooling would be affected by selection problems/OVB: individuals who complete high school are different ex ante that those who do not.

For example: they have higher "**ability**" than those who not. Ability affects both their completion of high school and how well they later do in the labor market. It may also be that those who think they have **higher returns** to schooling will invest more in it.

Instrumental Variable (IV) example

Let's exemplify with a causal question, for example: "does the amount or quality of schooling one gets affect ones' wages in the future?"

- Idea of an "Instrument": find something that is as good as random and affects only amount of schooling but cannot affect wages directly.
- Example from literature: quarter of birth (Angrist & Krueger 1991) Because one is allowed to drop out of school after age 16: Everyone born in 2000 started school in the same year, but those born in January 2000 are allowed to drop out earlier – in Jan 2016 than those born in December 2000 who can only drop out legally in Dec 2016. Therefore those born early in year on average get less schooling.

Example: School admission lotteries

KIPP program evaluation in the US

Setting:

- Two types of schools: public schools and charter schools. KIPP is a type of charter school with a special curriculum.
- Large gaps between white students and black & hispanic students.
- KIPP schools seemed to achieve better results for black and hispanic students than other schools – is this a causal effect of the KIPP program or something else?
- **T**= attending a KIPP school; **Y**= Grades
- The study focuses on the first KIPP school, Lynn in New England. Since 2005, number of applicants> number of student spots, and school started **to allocate spots by lottery**.

KIPP program evaluation in the US

Application forms Submitted to school

Some pre-screening e.g. too old applicants are screened out, those w siblings in Lynn KIPP are screened in.

Lottery

Winners

Losers

Not enrolled

Enrolled 73%

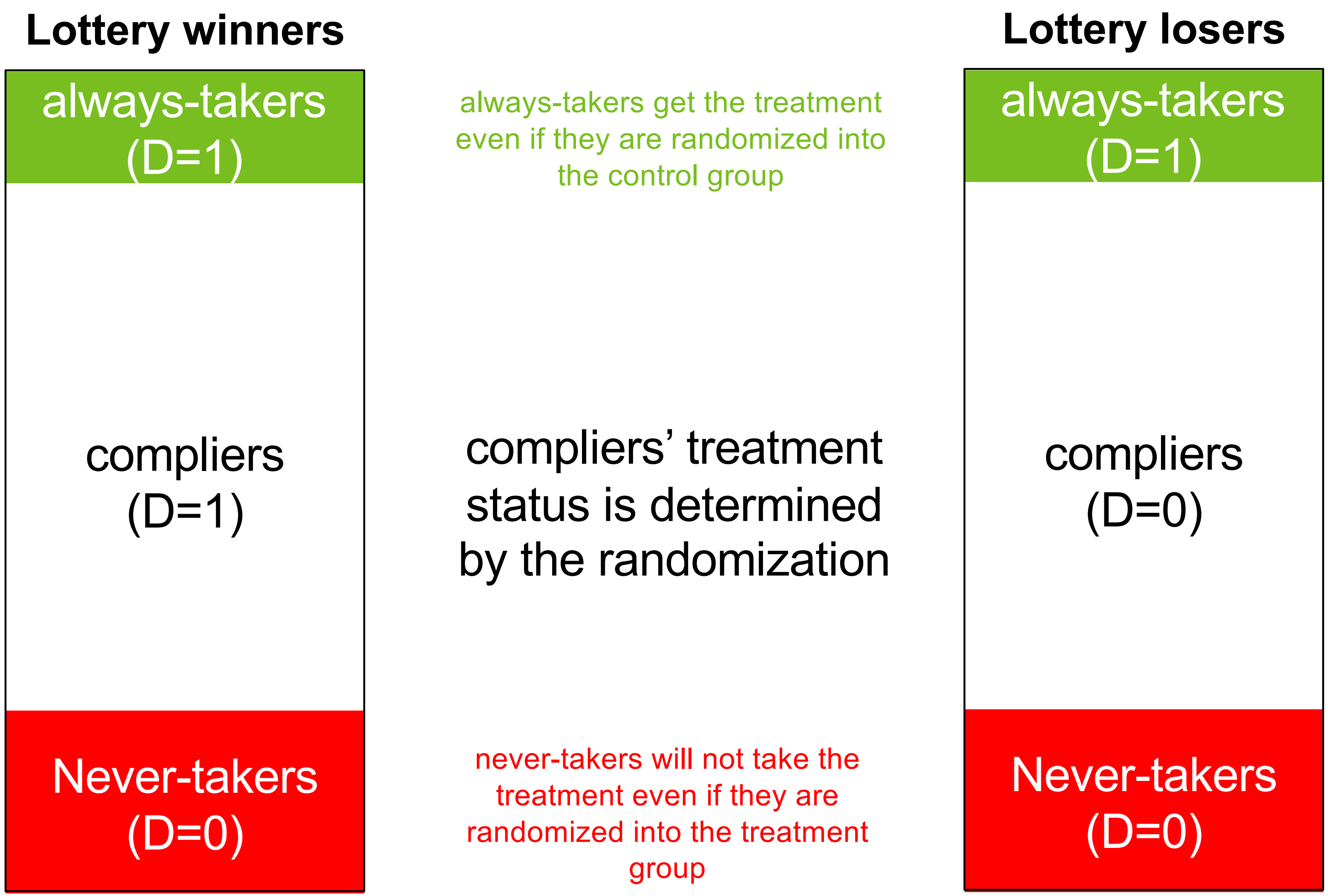
Not enrolled

Enrolled 3.5%

Who should we compare?

All those attending to all those not attending (after the lottery)?

Lottery winners to lottery losers?


Lottery winners
always-takers
(D=1)
compliers
(D=1)
Never-takers
(D=0)
always-takers get the treatment even if they are randomized into the control group
compliers' treatment status is determined by the randomization
never-takers will not take the treatment even if they are randomized into the treatment group
Lottery losers
always-takers
(D=1)
compliers
(D=0)
Never-takers
(D=0)

Who should we compare?

All those attending to all those not attending (after the lottery)?

- No: This would be problematic as there is *selection* into the enrolled group among lottery winners and lottery losers. Not all winners decide to enrol, some losers still manage to enrol.

Lottery winners to lottery losers?

- No: This would allow us to measure the causal effect of winning the lottery (ITT) but not the effect of actually enrolling and attending the charter school.

Using an Instrumental Variable

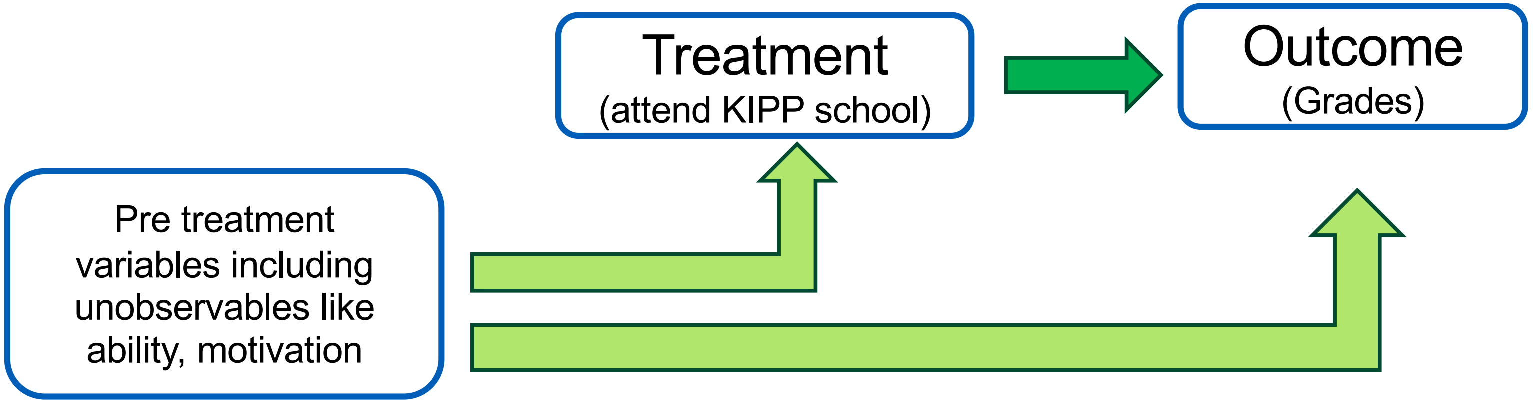
An instrumental Variable allows the researcher to identify causal effects if she can find an exogenous variable that leads to variation in the variable of interest (T). Generally:

- we worry that there is selection on omitted variables into ”treatment”
- Using an Instrument we can treat *part of the variation in treatment status* as random or exogenous.

If we use **the lottery** as an **instrument for attending** the KIPP school, this would be equivalent to measuring the causal affect of **attending the school because of winning the lottery.**

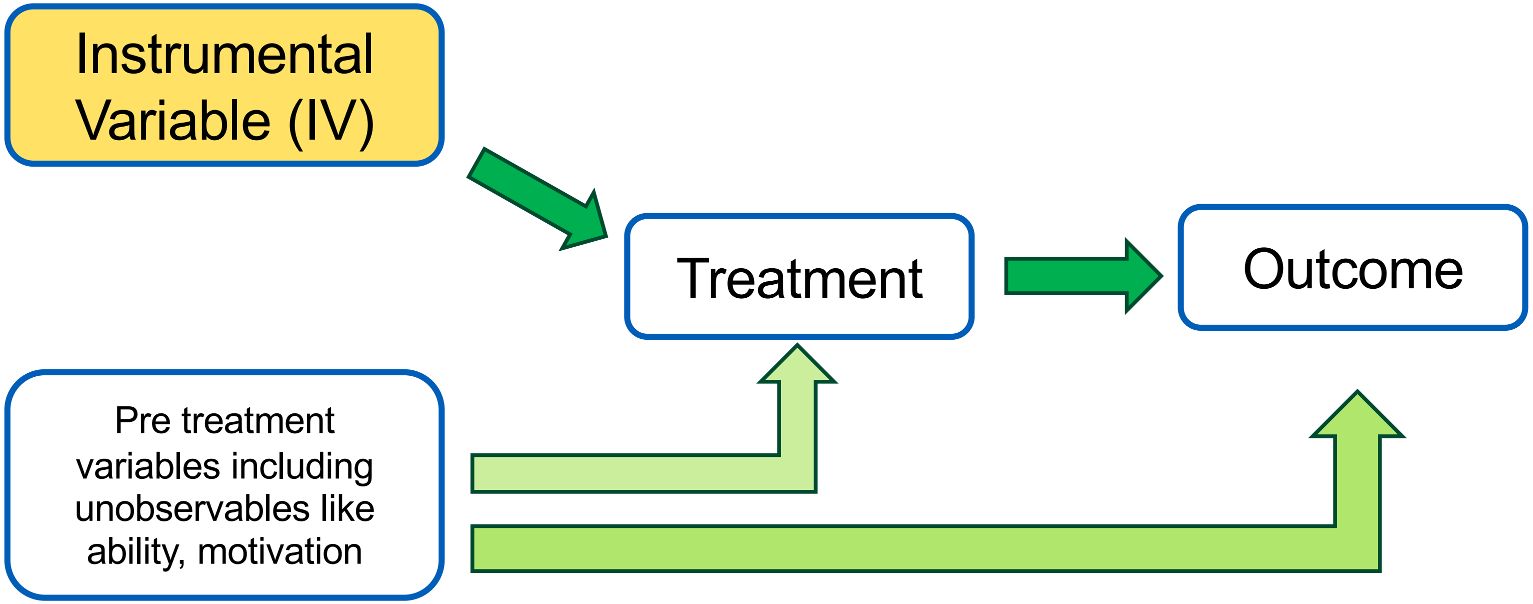
Instrumental variables idea

Starting point: We want to estimate the effect of a Treatment on an outcome. But we suspect that there are variables that affect both the treatment status (here: whether a person goes to a KIPP school) and the outcome, “pre-treatment variables” (confounders).



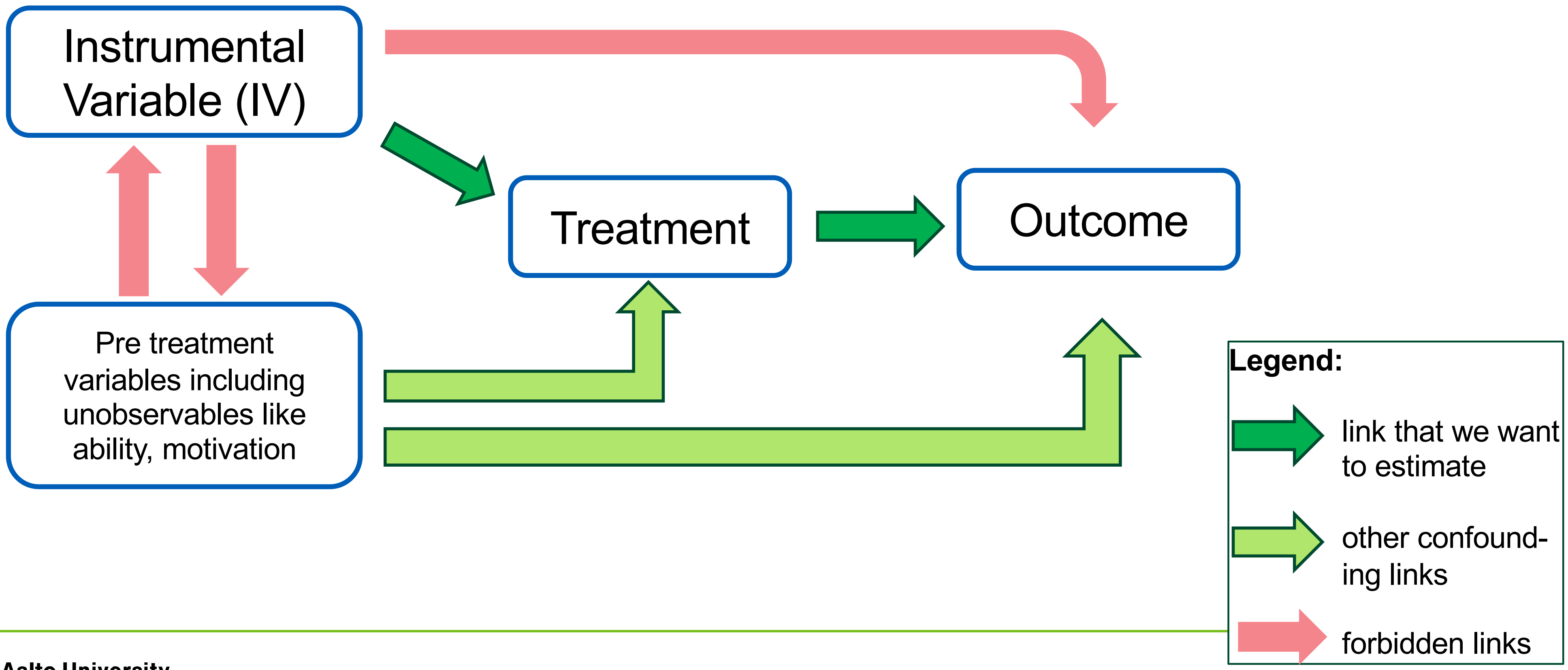
Instrumental variables idea

Possible solution: We find some variable that is random or almost random, that affects the treatment and introduces exogenous ("random") variation in the treatment. We call such a variable an "instrumental variable" or IV.



Instrumental variables challenges

Threats: The IV should not affect the outcome directly, and should not be systematically correlated with the same pre-treatment variables that we believe affect the treatment and the outcome (as said above, it should be as good as random).



Conditions on a valid instrument

There are three conditions that need to be satisfied for a variable to work as a good or "valid" instrument:

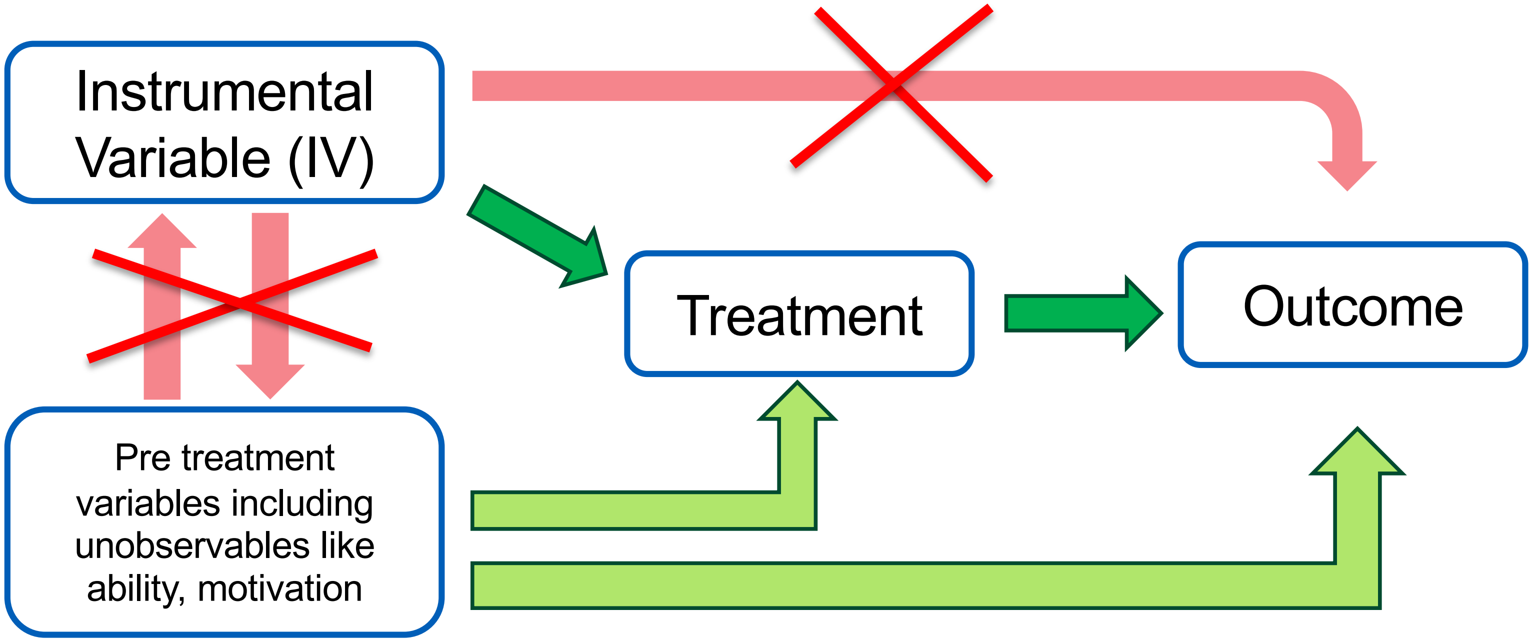
1. **Relevance condition: the instrument should be correlated with the variable of interest!** *The instrument has a causal effect on the variable whose effects we're trying to capture.*
2. **Exogeneity** *The instrument is randomly assigned or "as good as randomly assigned," in the sense of being unrelated to the omitted variables we'd like to control for if we could (e.g. family background, ability or motivation).*
3. **Exclusion restriction** *the chosen instrument should only affects outcomes through the variable of interest, or the "treatment variable".*

Relevance

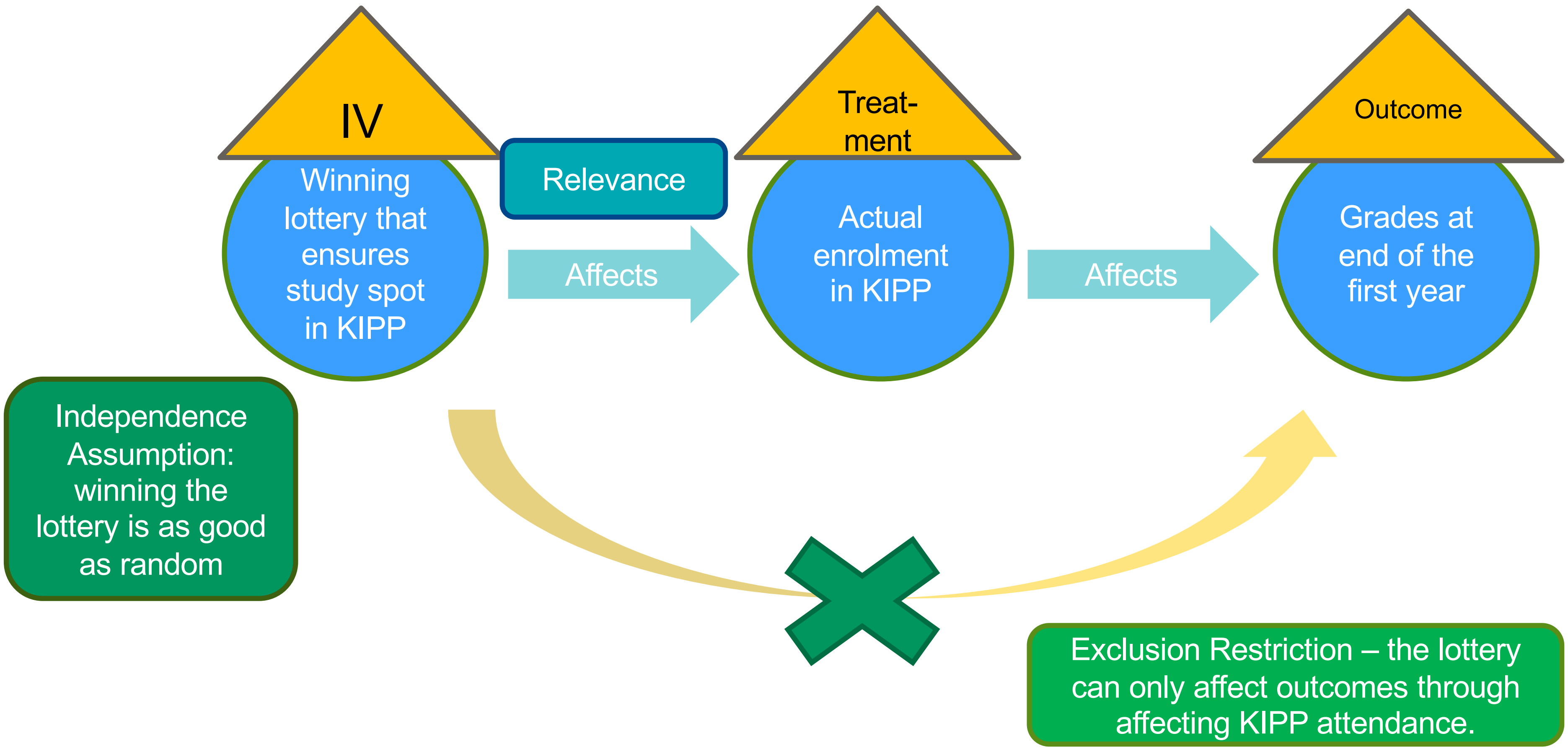
Exogeneity condition

Exclusion Restriction

Instrumental variables conditions added



Assumptions in the KIPP program evaluation



Using an Instrumental Variable

In the example, we would really want to measure the effect of attending a KIPP school on grades. But **being in a KIPP school is an endogenous outcome of unobservable characteristics.**

If all conditions (relevance, independence assumption and exclusion restriction) are satisfied, then:

Effect of winning lottery on likelihood to attend × Effect of attending on grades = Effect of winning lottery on grades

Using an Instrumental Variable

$$\overbrace{\boxed{\text{Effect of winning lottery on likelihood to attend}}}^{\text{First stage}} \times \boxed{\text{Effect of attending on grades}} = \overbrace{\boxed{\text{Effect of winning lottery on grades}}}^{\text{Reduced form}}$$

If we rearrange this expression we can isolate the effect of enrolling on grades:

$$\boxed{\text{Effect of attending on grades}} = \frac{\overbrace{\boxed{\text{Effect of winning lottery on grades}}}^{\text{Reduced form}}}{\underbrace{\boxed{\text{Effect of winning lottery on likelihood to attend}}}_{\text{First stage}}}$$

Using an Instrumental Variable

In the KIPP example (fig 3.2 in Mastering Metrics):

- Standardized grades among lottery winners are -0.003, while grades among lottery losers are -0.358.
- The likelihood to attend among lottery winners is 78.7% while among lottery losers it is 4.6%.

$$\boxed{\text{Effect of attending on grades}} = \frac{\boxed{\text{Effect of winning lottery on grades}}}{\boxed{\text{Effect of winning lottery on likelihood to attend}}} = \frac{-0.003 - (-0.358)}{0.787 - (0.046)} = \mathbf{0.486}$$

compare slide 29 (Wald estimator)!

Testing if the IV conditions are fulfilled

Conditions on a valid instrument

1. **Relevance : the instrument should be correlated with the variable of interest!** *The instrument has a causal effect on the variable whose effects we're trying to capture*
2. **Exogeneity** *The instrument is randomly assigned or "as good as randomly assigned," in the sense of being unrelated to the omitted variables we might like to control for (e.g. family background, ability or motivation).*
3. **Exclusion restriction** *the chosen instrument should only affects outcomes through the variable of interest, or the "treatment variable".*

Relevance condition

Exogeneity condition

Exclusion Restriction

Testing IV-conditions: Relevance

1. **Relevance: the instrument should be (strongly) correlated with the variable of interest!**

Relevance condition

How can this be tested? Ideas?

We can check the correlation between the instrument (winning lottery) and the var. of interest (being admitted). Often we do this by running a regression.

This regression is referred to as the “First stage”. A strong correlation i.e. a large and statistically significant coefficient= a strong first stage.

Testing IV-conditions: Exogeneity or Independence

2. Exogeneity *The instrument is randomly assigned or "as good as randomly assigned". It is not correlated with pre-treatment characteristics that may affect the potential outcome (this is the endogeneity problem that we try to solve by using the IV strategy!).*

In the case of a lottery, this randomness is almost guaranteed. For other instruments, this may be less evident.

Exogeneity condition

Can exogeneity be tested?

Not 100%, but the researcher can show that the instrument is not correlated with *observable* pre treatment variables that may affect the potential outcome. Show that the characteristics of winners and losers are similar. There can still be *unobservable* differences.

Testing IV-Conditions: Exclusion restriction

3. Exclusion restriction *the chosen instrument should only affects outcomes through the variable of interest, or the "treatment variable".*

Can this be tested? *No, but the researcher needs to provide convincing arguments in favour of the exclusion restriction. Usually this is done by suggesting different "threats" to the exclusion restriction – i.e. different alternative ways that the IV may affect the outcome, and showing with data that these do not seem to be problematic.*

Exclusion Restriction

Using an Instrumental Variable

We often face endogeneity problems (there is selection on omitted variables into ”treatment”)

An Instrument is an exogenous variable that leads to variation in the endogenous Treatment (variable of interest). Generally:

- Using an Instrument we can treat *part of the variation in treatment status* as random or exogenous.
- What part?

LATE (Local Average Treatment Effect)

The IV strategy allows us to estimate the treatment effect for a specific group of individuals or “units”:

Those whose treatment status changes because of the instrument!

This is equivalent to the **“compliers” discussed earlier!**

LATE in the KIPP study

The effect of attendance on grades is estimated for kids who:

1. Have families who selected into signing up for the lottery
2. Attended the KIPP *because* they won the lottery (i.e.“compliers” with respect to this instrument.)

Probably, other groups of kids would have had a different effect from attending these schools.

- Kids who did not sign up for the lottery, and kids who did not attend despite winning the lottery (never takers)
- Kids who attended even if they lost the lottery (always takers)

If treatment effects are heterogenous (different) between different groups of kids, then the effect that we estimate for the compliers is not the general ATE, but the ”Local Average Treatment Effect”, LATE.

LATE in Card (1995)

In another famous IV paper on returns to education, David Card (1995) wants to estimate the effects of higher education on earnings. Since there is selection into college, Card suggests an IV that leads to variation in whether an individual attends college:

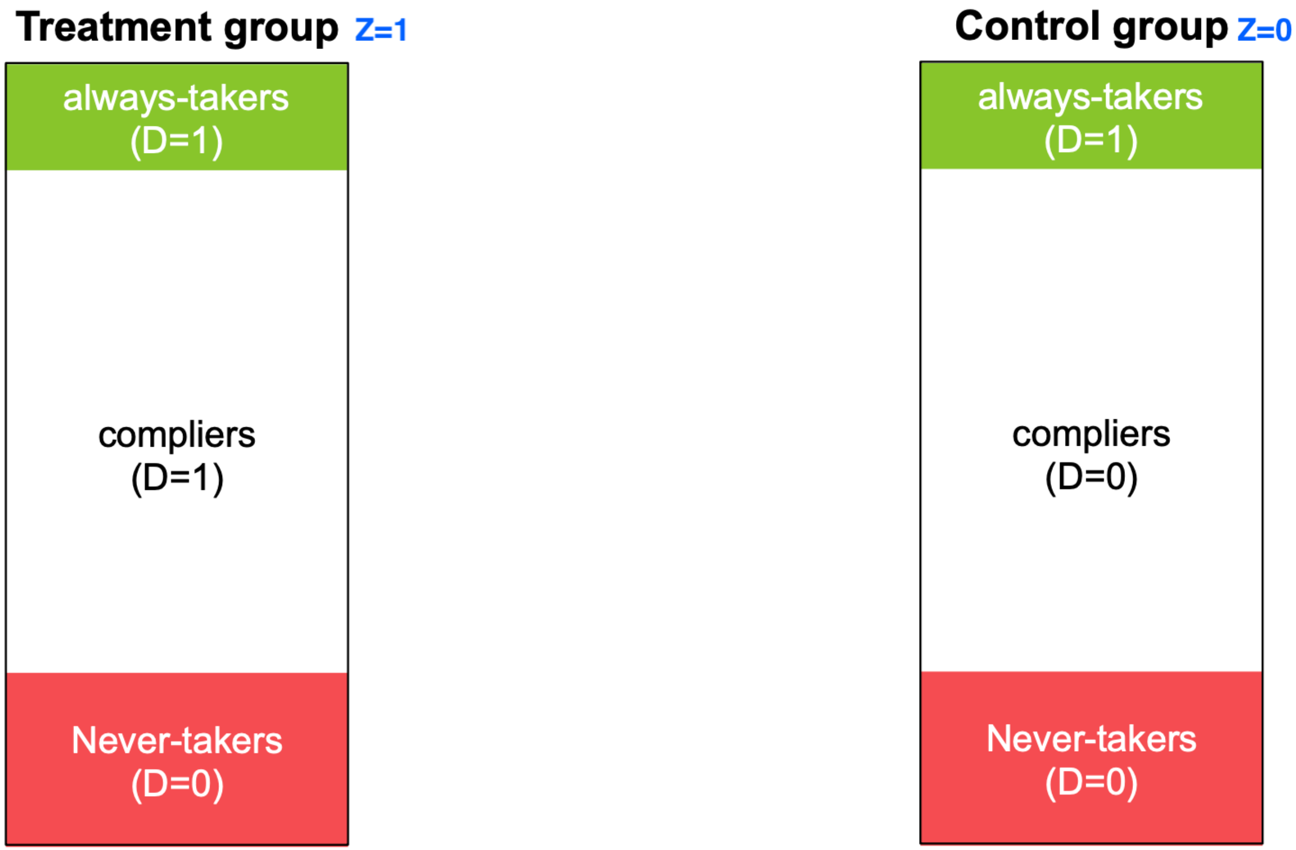
The IV is: ”living in the same county as a college”

Apparently, there is a positive correlation between living near a college and attending college.

The Instrument allows us to estimate the effect of schooling on those who attend college only if they live in a county with a college, and do not attend college otherwise.

LATE and compliance

Think of the instrument as the “treatment assignment”, Z.



In Card 1995: always takers are students who go to college regardless if they live near one.

Compliers are those who go only if they live in a county with a college.

Variation in the IV “college in county” allows us to estimate the effect on the compliers.

LATE and drawbacks of IV

If treatment effects are different for compliers than for non-compliers:

‘IV only identifies the “LATE” and that may or may not be a policy relevant variable. It’s value ultimately depends on how closely the compliers’ average treatment effect resembles that of the other subpopulations’. (Cunningham)

Important to think about to what kind of individuals’ results can be generalized.

→ What possible school reforms could the study of living close to a college teach us more about?

This depends on why would living in the same county as a college make someone more likely to go to college? Lower costs – can remain at home while studying.

If the instrument mainly affects attendence through lowering college costs, perhaps lowering tuition fees or paying housing allowance would attract similar students?

More on IV

For more explanation on this and other examples, see

- https://www.youtube.com/watch?v=eoJUPd6104Q

  (video for pre-lecture assignment from Marginal revolution University (Master Joshway/Josh Angrist))
- Mastering Metrics chapter on Instrumental Variables
- (Cunningham Mixtape chapter on Instrumental Variables – this is too technical for your level but has some nice discussion and examples)

Other Examples 1:
Rainfall as an instrument

Miguel, Satyanath and Sergenti, 2004 (JPE)

Economic Shocks and Civil Conflict: An Instrumental Variables Approach

Edward Miguel

University of California, Berkeley and National Bureau of Economic Research

Shanker Satyanath and Ernest Sergenti

New York University

Rainfall

- **T= Economic shocks Y= civil conflict**
- **Geographic focus on Sub-Saharan Africa**
- **Why is it (generally) hard to estimate the effect of economic shocks/economic conditions on conflict?**
  - Endogeneity concerns, ovb, reverse causality…
- **using deviations in rainfall as IV.**
  - Idea: when it rains **much more or much less than usual**, this affects agriculture (in Sub Saharan Africa, most agriculture is rainfed) and as agriculture is the biggest sector, this affects economic conditions of many people.

Conditions on a valid instrument

1. **Relevance : the instrument should be correlated with the variable of interest!** *The instrument has a causal effect on the variable whose effects we're trying to capture*
2. **Exogeneity** *The instrument is randomly assigned or "as good as randomly assigned," in the sense of being unrelated to the omitted variables we might like to control for (e.g. family background, ability or motivation).*
3. **Exclusion restriction** *the chosen instrument should only affects outcomes through the variable of interest, or the "treatment variable".*

Relevance condition

Exogeneity condition

Exclusion Restriction

What is required in order to fulfill the three conditions/assumptions of IV?

T= economic shock, Y= Conflict, Z= Rainfall shocks

What is required for:

Relevance:

Exogeneity:

Exclusion restriction:

What is required in order to fulfill the three conditions/assumptions of IV?

T= economic shock, Y= Conflict, Z= Rainfall shocks

What is required for:

Relevance: The instrument Z and the treatment T are strongly correlated (with the expected sign).

Exogeneity: The instrument Z is “as good as random”. Whether it rains less than usual in a given place is as good as random.

Exclusion restriction: rainfall shocks only affect conflict and unrest through their effect on economic conditions/shocks.

In your opinion – are these conditions fulfilled?

Rainfall as an IV and the exclusion restriction

Madestam, A., Shoag, D., Veuger, S. and Yanagizawa-Drott, D., 2013. Do political protests matter? evidence from the tea party movement. *The Quarterly Journal of Economics*, *128*(4), pp.1633-1685.

Rainfall during days of political rallies across the US affected local turnout, which then affected how big the Tea party movement became in that city. *If this would later lead to conflict, rainfall would have affected conflict through the participation in political rallies.*

Rogall, T., 2021. Mobilizing the masses for genocide. American economic review, 111(1), pp.41-72.

Rainfall in combination with bad quality roads affected what villages the militia chose to target during the Rwandan genocide. *Rainfall affected conflict through transportation costs.*

Moreno-Medina, J., 2022. Sinning in the Rain: Weather Shocks, Church Attendance, and Crime. Review of Economics and Statistics, pp.1-16.

Rainfall on Sundays leads to decrease in church attendance and to an increase in crime in the US. *Increased crime could potentially increase conflict locally.*

Aalto University
School of Business

Other example of IV

The effect of family size on parents' labor supply

Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size

By JOSHUA D. ANGRIST AND WILLIAM N. EVANS*

Research on the labor-supply consequences of childbearing is complicated by the endogeneity of fertility. This study uses parental preferences for a mixed sibling-sex composition to construct instrumental variables (IV) estimates of the effect of childbearing on labor supply. IV estimates for women are significant but smaller than ordinary least-squares estimates. The IV are also smaller for more educated women and show no impact of family size on husbands' labor supply. A comparison of estimates using sibling-sex composition and twins instruments implies that the impact of a third child disappears when the child reaches age 13. (JEL J13, J22)

Treatment: family size, Outcome: how much the parents work. But family size is an endogenous choice of the parents. This study uses the idea that parents often want to have both boys and girls. The authors use the sex of the first 2 children as an instrument for whether parents have a third child.

The effect of Sweden - US mass emigration (1860-) on Swedish political development in the following decades

< PREVIOUS ARTICLE NEXT ARTICLE >

Details | Figures | References | Cited by

Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States

Mounir Karadja and Erik Prawitz

PDF | PDF PLUS | Abstract | Full Text | Supplemental Material

Abstract

We study the political effects of mass emigration to the United States in the nineteenth century using data from Sweden. To instrument for total emigration over several decades, we exploit severe local frost shocks that sparked an initial wave of emigration, interacted with within-country travel costs. Our estimates show that emigration substantially increased the local demand for political change, as measured by labor movement membership, strike participation, and voting. Emigration also led to de facto political change, increasing welfare expenditures as well as the likelihood of adopting more inclusive political institutions.

Journal of Political Economy
Volume 127, Number 4
August 2019

Treatment: emigration to the US **Outcome**: political participation and political change in Sweden. But it is not random who migrates or from which areas people migrate. Therefore the authors use frost shocks in 1860's together with the distance from emigration ports as an IV for initial emigration.

The effect of slave trades on economic outcomes in Africa

THE LONG-TERM EFFECTS OF AFRICA'S SLAVE TRADES*

NATHAN NUNN

Can part of Africa's current underdevelopment be explained by its slave trades? To explore this question, I use data from shipping records and historical documents reporting slave ethnicities to construct estimates of the number of slaves exported from each country during Africa's slave trades. I find a robust negative relationship between the number of slaves exported from a country and current economic performance. To better understand if the relationship is causal, I examine the historical evidence on selection into the slave trades and use instrumental variables. Together the evidence suggests that the slave trades had an adverse effect on economic development.

Treatment: **country being exposed to the slave trade**. Outcome: **country GDP in 2000**. But there may be omitted variables that affected both these variables (e.g. historic economic performance). To get around this Nunn uses an IV based on the distance from countries in Africa to slave destinations.

The effect of (colonial) institutions on current GDP in former colonies

The Colonial Origins of Comparative Development: An Empirical Investigation

By Daron Acemoglu, Simon Johnson, and James A. Robinson*

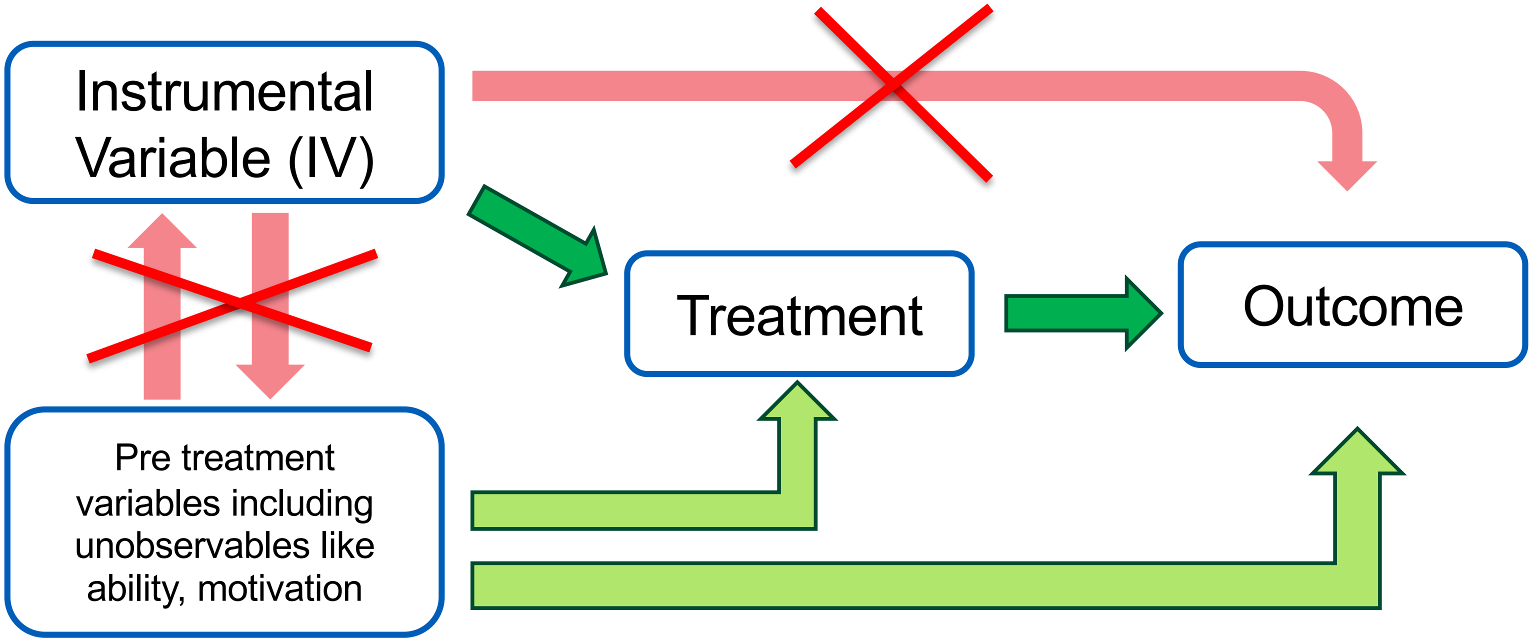
We exploit differences in European mortality rates to estimate the effect of institutions on economic performance. Europeans adopted very different colonization policies in different colonies, with different associated institutions. In places where Europeans faced high mortality rates, they could not settle and were more likely to set up extractive institutions. These institutions persisted to the present. Exploiting differences in European mortality rates as an instrument for current institutions, we estimate large effects of institutions on income per capita. Once the effect of institutions is controlled for, countries in Africa or those closer to the equator do not have lower incomes. (*JEL* O11, P16, P51)

Treatment: **country having stronger institutions**, Outcome: **country GDP in 1995**.
But there could be *reverse causality*. Therefore the authors restrict their analysis to countries that are former colonies, and use *settler mortality rates* as an IV for institutional quality (idea: settler mortality affected historic institutions set up during colonial times, and those affect the institutions present today. More specifically, where many settlers died, colonizers set up extractive institutions, and where settlers did not die, they settled down and set up more “long term” institutions that help economic growth.)

Aalto University
School of Business

Conclusion and main takeaways

Instrumental variables recap



Instrumental variables recap

In words:

1. We face a situation with potential selection problem/OVB: pre treatment characteristics are likely to impact both treatmemt status and outcome.
2. We find a variable, the “IV” that is correlated with the outcome of interest.
3. The IV must not affect the outcome directly. *[Exclusion restriction – an assumption that must be justified by the researcher]*
4. Assume the IV is as good as randomly assigned *[Exogeneity – an assumption that can be partly confirmed with balance tests]*
5. The IV must be strongly correlated with the treatment (the variable of interest) *[Relevance – can be tested by the First stage]*

   -> If 4 holds, the First stage reveals the *causal* effect if the IV on the treatment status: since the IV is “randomly assigned”, it is not correlated with any other confounders, except the treatmemt.

The components of IV

The reduced form gives the effect of the instrument, Z on Y. But we are interested in the effect of T on Y. Because Z is exogenous/as good as random/uncorrelated with pre treat variables, the only reason Z affects the outcome is because if affects T (schooling).
The first stage= how the instrument affect actually getting treatment. **This is = the share of compliers**.

If we divide the RF by the FS, we rescale the effect of the instrument on Y by the effect of the instrument on T, that produces a causal effect that is measured in treatment units.

$$\boxed{\text{Effect of attending on grades}} = \frac{\overbrace{\boxed{\text{Effect of winning lottery on grades}}}^{\text{Reduced form}}}{\underbrace{\boxed{\text{Effect of winning lottery on likelihood to attend}}}_{\text{First stage}}}$$

Conclusion: key points about IV

- Instrumental Variable approach can be useful for solving endogeneity problems in observational data.
- The IV approach rests on 3 conditions, one is assumed, one can be partially tested (but not for unobservable variables) and one can be tested.
- problematic aspects

  1. The three assumptions are not always plausibly fulfilled

  2. What we estimate with IV is the Local average treatment effect (LATE). It is important to consider who are the compliers: who are the people whose treatment can be expected to be exogenously changed when the instrument is switched on?