Compliance and the limits of RCTs

Matti Sarvimäki

Principles of Empirical Analysis, 2021
Lecture 7

Today's running example: Moving to Opportunity

- 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

Today's running example: Moving to Opportunity

- 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
  - experimental: offered housing vouchers to low-poverty neighborhoods

Today's running example: Moving to Opportunity

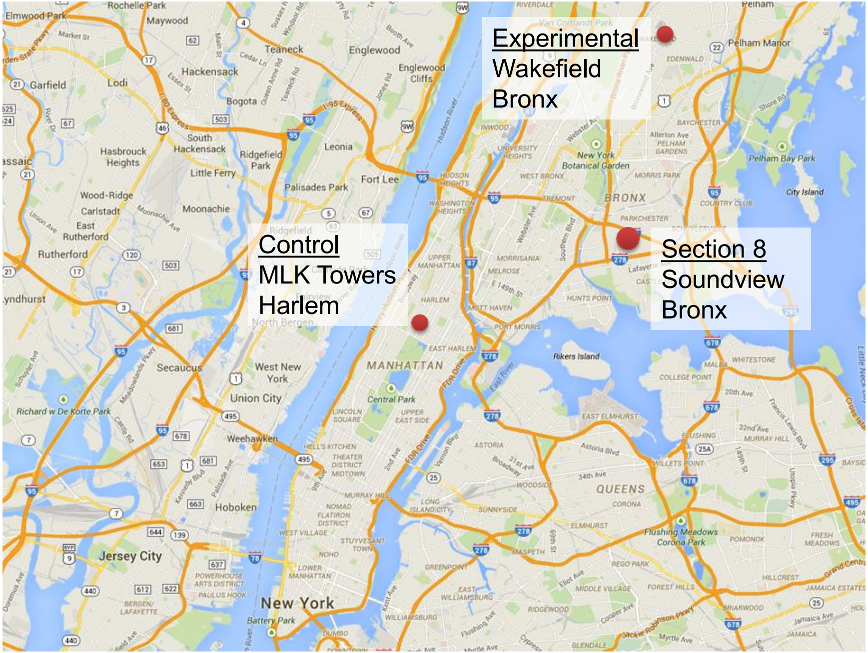
- One of the most famous social experiments of all time
  - target group: households with children living in high-poverty public housing projects (primarily minority, single mother families)
  - implemented in 1994-98 in Baltimore, Boston, Chicago, LA, New York
- Random assignment of 4,600 families into three groups:
  - control: not offered a voucher, stayed in public housing
  - section 8: offered conventional housing vouchers, no restrictions
  - experimental: offered housing vouchers to low-poverty neighborhoods
- Many families chose not to use the voucher they were offered
  - 48% of experimental group used voucher
  - 66% of Section 8 group used voucher

The MTO parts of these slides draw heavily from lecture 3 of Raj Chetty's excellent course Using Big Data to Solve Economic and Social Problems. I'm also borrowing quite a bit from Tuukka's (also excellent) Urban Economics course.

Today's learning objectives

- Today's question: How to deal with **imperfect compliance**?
  - some randomized into the treatment group do not get the treament
  - some randomized into the control group get the treament
- Key concepts
  1. compliers, always-takers and never-takers
  2. intention-to-treat (ITT)
  3. first-stage
  4. local average treatment effect (LATE)
  5. average treatment effect on the treated (ATT or TOT)
- We also briefly discuss the limits of RCTs

Common MTO Residential Locations in New York



Early research

MEAN EFFECT SIZES FOR SUMMARY MEASURES OF OUTCOMES[a]

	All Adults
	E – C (i)
Economic self-sufficiency	0.017 (0.031)
Absence of physical health problems	0.012 (0.024)
Absence of mental health problems	0.079* (0.030)
Absence of risky behavior	
Education	
Overall	0.036 (0.020)

[a] E – C denotes experimental – control

Robust standard errors adjusted for household clustering are in parentheses; * = p-value < 0.05.

Kling, Liebman, Katz (2007): Experimental Analysis of Neighborhood Effects.

Early research

Mean Effect Sizes for Summary Measures of Outcomes[a]

	All Adults	
	E – C (i)	S – C (ii)
Economic self-sufficiency	0.017 (0.031)	0.037 (0.033)
Absence of physical health problems	0.012 (0.024)	0.019 (0.026)
Absence of mental health problems	0.079* (0.030)	0.029 (0.033)
Absence of risky behavior		
Education		
Overall	0.036 (0.020)	0.028 (0.022)

[a] E – C denotes experimental – control; S – C denotes Section 8 – control. Estimates are the intent-to-treat mean effect sizes,

Robust standard errors adjusted for household clustering are in parentheses; * = p-value < 0.05.

Kling, Liebman, Katz (2007): Experimental Analysis of Neighborhood Effects.

Early research

MEAN EFFECT SIZES FOR SUMMARY MEASURES OF OUTCOMES[a]

	All Adults		All Youth		Female Youth		Male Youth		M – F Youth	
	E – C	S – C	E – C	S – C	E – C	S – C	E – C	S – C	E – C	S–C
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Economic	0.017	0.037								
self-sufficiency	(0.031)	(0.033)								
Absence of physical	0.012	0.019	−0.038	−0.020	0.025	0.077	−0.112*	−0.114	−0.138	−0.192*
health problems	(0.024)	(0.026)	(0.038)	(0.040)	(0.053)	(0.055)	(0.053)	(0.061)	(0.076)	(0.084)
Absence of mental	0.079*	0.029	0.102	0.138*	0.267*	0.192*	−0.052	0.054	−0.319*	−0.138
health problems	(0.030)	(0.033)	(0.053)	(0.056)	(0.062)	(0.067)	(0.080)	(0.092)	(0.101)	(0.113)
Absence of risky			−0.023	−0.039	0.142*	0.129*	−0.181*	−0.208*	−0.323*	−0.337*
behavior			(0.043)	(0.050)	(0.053)	(0.059)	(0.062)	(0.071)	(0.080)	(0.092)
Education			0.050	0.028	0.138*	0.056	−0.053	−0.001	−0.191*	−0.057
			(0.041)	(0.047)	(0.065)	(0.068)	(0.047)	(0.060)	(0.080)	(0.090)
Overall	0.036	0.028	0.018	0.018	0.136*	0.109*	−0.099*	−0.078*	−0.235*	−0.187*
	(0.020)	(0.022)	(0.025)	(0.026)	(0.034)	(0.034)	(0.031)	(0.037)	(0.047)	(0.051)

[a]E – C denotes experimental – control; S – C denotes Section 8 – control. Estimates are the intent-to-treat mean effect sizes, from Equation (1), fully interacted with gender in columns (v)–(x) as described in the text. The estimated equations all include site indicators and the baseline covariates listed in Appendix A with those in Table A1 included for adults and those in Tables A1 and A2 included for youth. M – F Youth is male – female difference. Adult economic self-sufficiency: + adult not employed and not on TANF + employed + 2001 earnings – on TANF – 2001 government income. Adult mental health: – distress index – depression symptoms – worrying + calmness + sleep. Adult physical health: – self-reported health fair/poor – asthma attack past year – obesity – hypertension – trouble carrying/climbing. Adult overall includes 15 measures in self-sufficiency, physical health, and mental health. Youth physical health: – self-reported health fair/poor – asthma attack past year – obesity – nonsports injury past year. Youth mental health: – distress index – depression symptoms – anxiety symptoms. Youth risky behavior: – marijuana past 30 days – smoking past 30 days – alcohol past 30 days – ever pregnant or gotten someone pregnant. Youth education: + graduated high school or still in school + in school or working + WJ-R broad reading score + WJ-R broad math score. Youth overall includes 15 measures in physical health, mental health, risky behavior, and education. Sample sizes in the E, S, and C groups are 1,453, 993, and 1,080 for adults and 749, 510, and 548 for youth ages 15–20 on 12/31/2001. Robust standard errors adjusted for household clustering are in parentheses; * = p-value < 0.05.

Kling, Liebman, Katz (2007): Experimental Analysis of Neighborhood Effects.

Early research

- Making sense of the previous table
  - outcomes: indices that aggregate information over multiple measures
    - for example, the index of economic self-sufficiency includes five measures of employment, earnings, and public assistance
  - each index has mean 0 and standard deviation 1

Early research

- Making sense of the previous table
  - outcomes: indices that aggregate information over multiple measures
    - for example, the index of economic self-sufficiency includes five measures of employment, earnings, and public assistance
  - each index has mean 0 and standard deviation 1
- Impacts of being offered an experimental voucher (4–7 years later)
  - no effects on adult economic self-sufficiency or physical health

Early research

- Making sense of the previous table
  - outcomes: indices that aggregate information over multiple measures
    - for example, the index of economic self-sufficiency includes five measures of employment, earnings, and public assistance
  - each index has mean 0 and standard deviation 1
- Impacts of being offered an experimental voucher (4–7 years later)
  - no effects on adult economic self-sufficiency or physical health
    - MDE for economic self-sufficiency: $2.8 \times .031 = .09$sd

Early research

- Making sense of the previous table
  - outcomes: indices that aggregate information over multiple measures
    - for example, the index of economic self-sufficiency includes five measures of employment, earnings, and public assistance
  - each index has mean 0 and standard deviation 1
- Impacts of being offered an experimental voucher (4–7 years later)
  - no effects on adult economic self-sufficiency or physical health
    - MDE for economic self-sufficiency: $2.8 \times .031 = .09$sd
  - improved mental health for adults
  - positive effect on teenage girls
  - negative effect on teenage boys

The most recent results

- Chetty, Hendren, Katz (2016) focus on those moving as children
  - group 1: younger than 13 (average 8.2) at assignment
  - group 2: 13-18 years old (average 15.1) at assignment

The most recent results

- Chetty, Hendren, Katz (2016) focus on those moving as children
  - group 1: younger than 13 (average 8.2) at assignment
  - group 2: 13-18 years old (average 15.1) at assignment
- MTO data linked to 1996–2012 federal income tax returns
  - 4,604 households and 15,892 individuals
    - primary focus on 8,603 children born in or before 1991
  - about 85% of children matched
    - match rates do not differ significantly across treatment groups
    - baseline covariates balanced across treatment groups in matched data

The most recent results

- Chetty, Hendren, Katz (2016) focus on those moving as children
  - group 1: younger than 13 (average 8.2) at assignment
  - group 2: 13-18 years old (average 15.1) at assignment
- MTO data linked to 1996–2012 federal income tax returns
  - 4,604 households and 15,892 individuals
    - primary focus on 8,603 children born in or before 1991
  - about 85% of children matched
    - match rates do not differ significantly across treatment groups
    - baseline covariates balanced across treatment groups in matched data
- Using administrative data (tax records) is quite new in the US
  - earlier work based typically on survey data
  - in the Nordic countries, we have a long tradition (and much better infrastructure) for using administrative data in research

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

	Housing voucher take-up (1)
Panel A. Children < age 13 at random assignment	
Exp. versus control	47.66***
	(1.653)
Sec. 8 versus control	65.80***
	(1.934)
Observations	5,044
Control group mean	0

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

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	
		ITT	TOT
	(1)	(2)	(3)
Panel A. Children < age 13 at random assignment			
Exp. versus control	47.66***	−17.05***	−35.96***
	(1.653)	(0.853)	(1.392)
Sec. 8 versus control	65.80***	−14.88***	−22.57***
	(1.934)	(0.802)	(1.024)
Observations	5,044	4,958	4,958
Control group mean	0	50.23	50.23

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	
		ITT	TOT
	(1)	(2)	(3)
Panel A. Children < age 13 at random assignment			
Exp. versus control	47.66***	−17.05***	−35.96***
	(1.653)	(0.853)	(1.392)
Sec. 8 versus control	65.80***	−14.88***	−22.57***
	(1.934)	(0.802)	(1.024)
Observations	5,044	4,958	4,958
Control group mean	0	50.23	50.23
Panel B. Children age 13–18 at random assignment			
Exp. versus control	40.15***	−14.00***	−34.70***
	(2.157)	(1.136)	(2.231)
Sec. 8 versus control	55.04***	−12.21***	−22.03***
	(2.537)	(1.078)	(1.738)
Observations	2,358	2,302	2,302
Control group mean	0	49.14	49.14

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

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

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

Compliance

- Often only part of the treatment group actually gets the treament
  - e.g. only 48% of those randomized into the experimental group in MTO chose to use the voucher (column 1 of the previous slide)
  - similarly, 66% of the section 8 group used the voucher

Compliance

- Often only part of the treatment group actually gets the treament
  - e.g. only 48% of those randomized into the experimental group in MTO chose to use the voucher (column 1 of the previous slide)
  - similarly, 66% of the section 8 group used the voucher
- This is often referred to as **compliance**
  - "compliers": randomized into treatment group and get the treatment
  - "never-takers": in treatment group, but don't take the treatment
  - "always-takers": in control group, but get the treatment nevertheless
    - the challenge is symmetrical for both types of "noncompliance"

Compliance

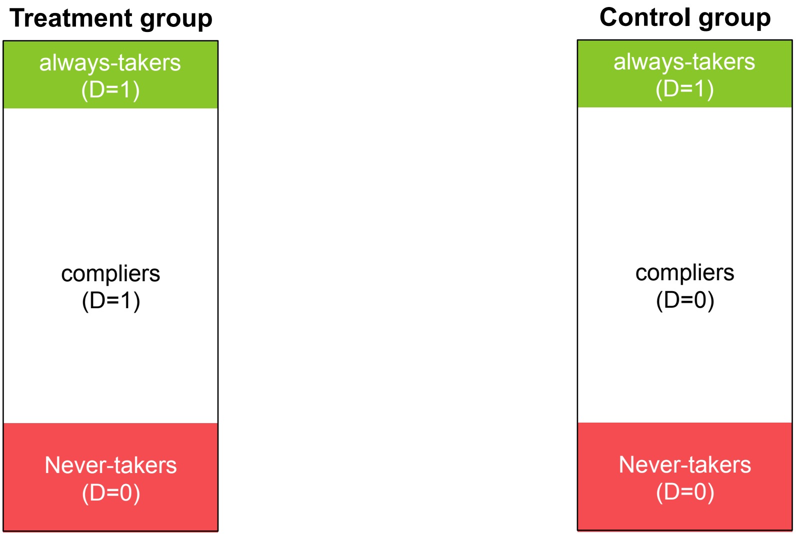
- Often only part of the treatment group actually gets the treament
  - e.g. only 48% of those randomized into the experimental group in MTO chose to use the voucher (column 1 of the previous slide)
  - similarly, 66% of the section 8 group used the voucher
- This is often referred to as **compliance**
  - "compliers": randomized into treatment group and get the treatment
  - "never-takers": in treatment group, but don't take the treatment
  - "always-takers": in control group, but get the treatment nevertheless
    - the challenge is symmetrical for both types of "noncompliance"
- Compliance *choice* is potentially affected by potential outcomes
  - e.g. those expecting to benefit the least becoming never-takers

Compliance

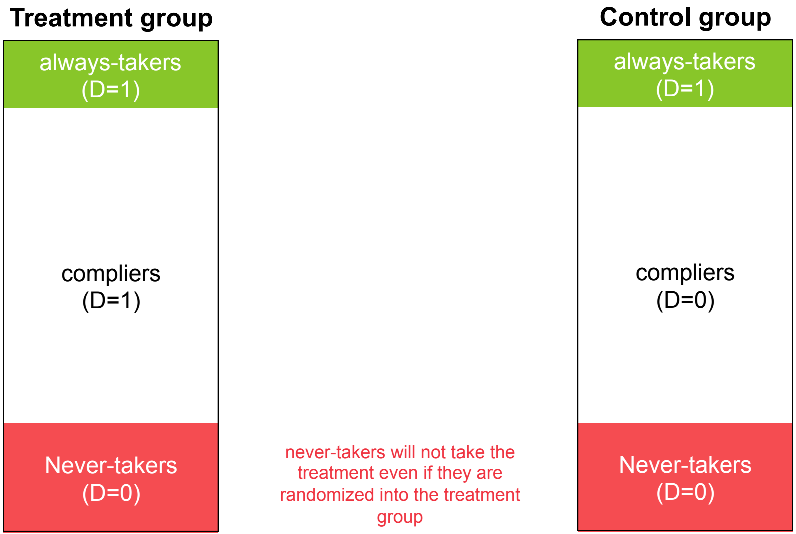
- Often only part of the treatment group actually gets the treament
  - e.g. only 48% of those randomized into the experimental group in MTO chose to use the voucher (column 1 of the previous slide)
  - similarly, 66% of the section 8 group used the voucher
- This is often referred to as **compliance**
  - "compliers": randomized into treatment group and get the treatment
  - "never-takers": in treatment group, but don't take the treatment
  - "always-takers": in control group, but get the treatment nevertheless
    - the challenge is symmetrical for both types of "noncompliance"
- Compliance *choice* is potentially affected by potential outcomes
  - e.g. those expecting to benefit the least becoming never-takers

  → comparing those who actually gets the treatment to the entire control group is not a valid comparison

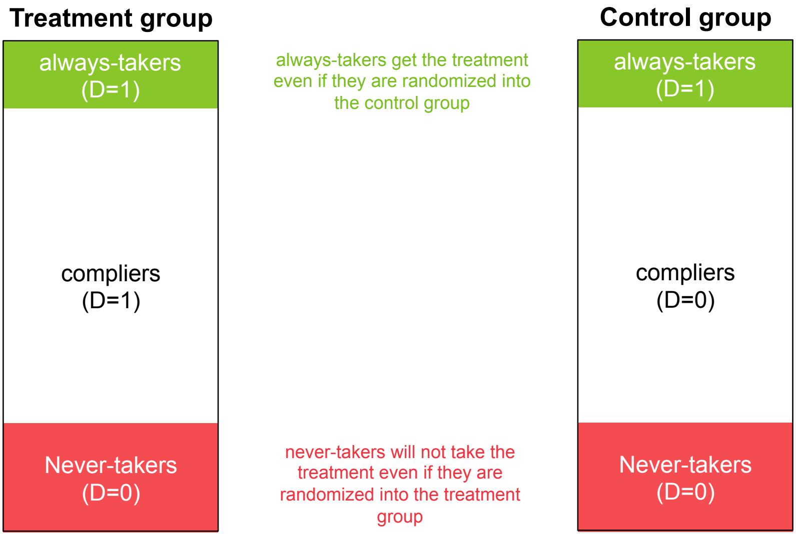
Compliers, always-takers and never-takers



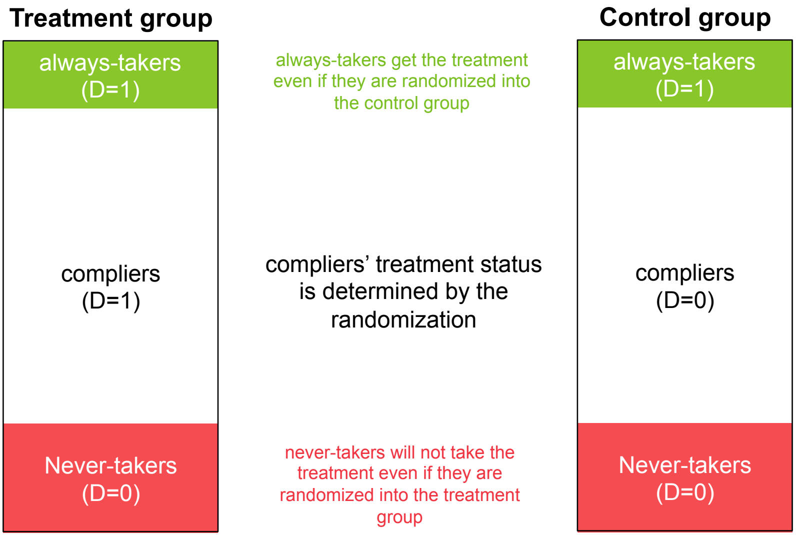
Compliers, always-takers and never-takers



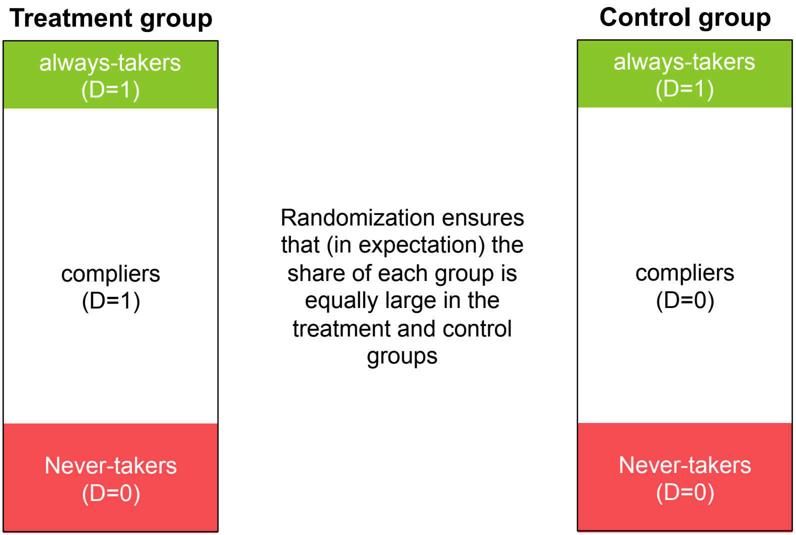
Compliers, always-takers and never-takers



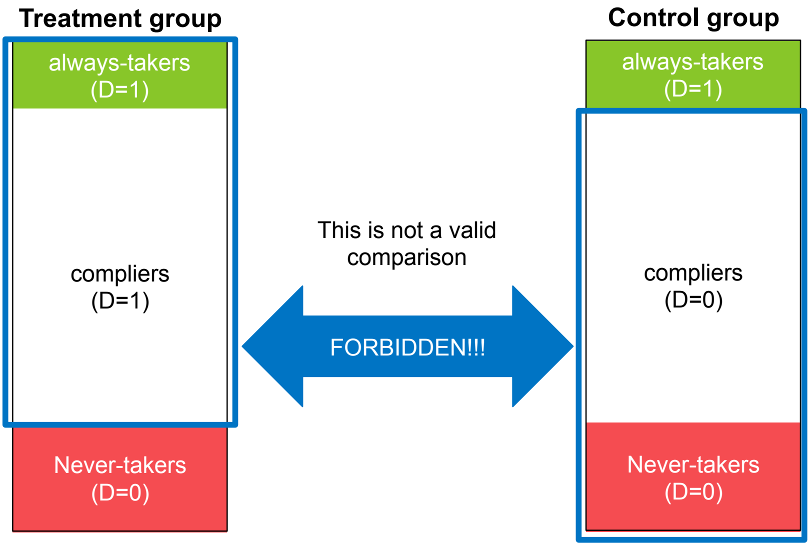
Compliers, always-takers and never-takers



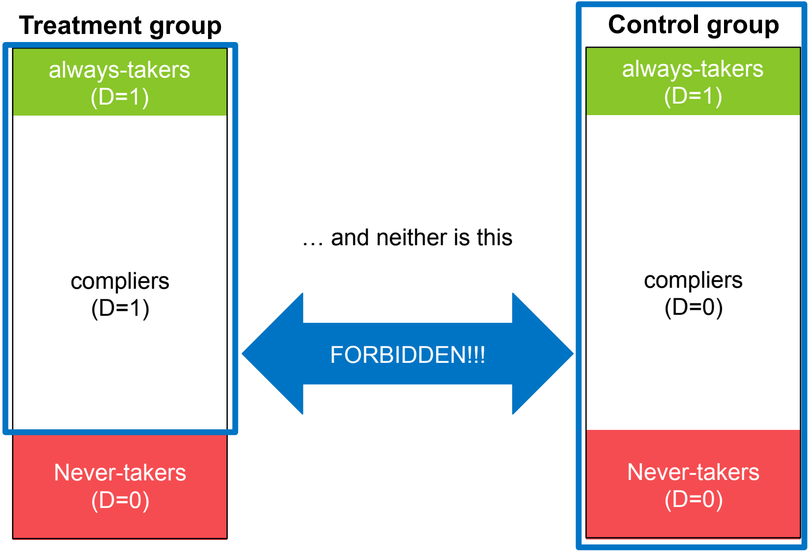
Compliers, always-takers and never-takers



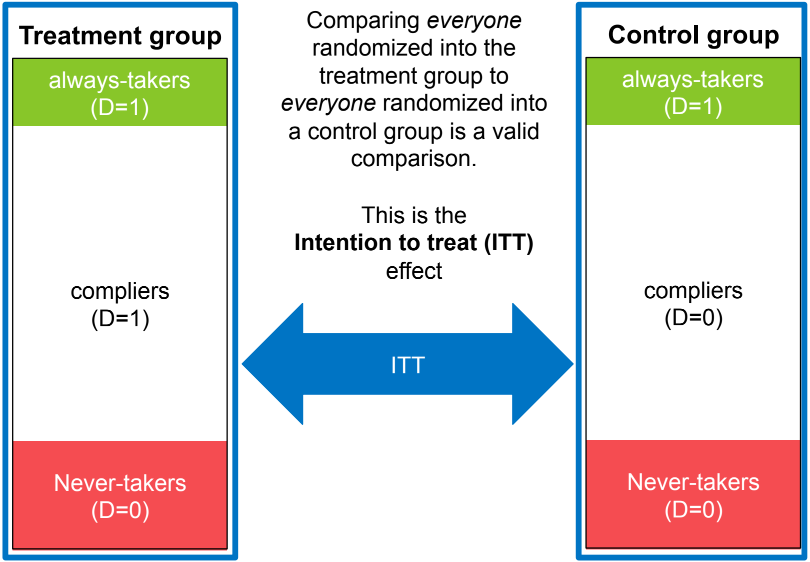
Invalid comparisons



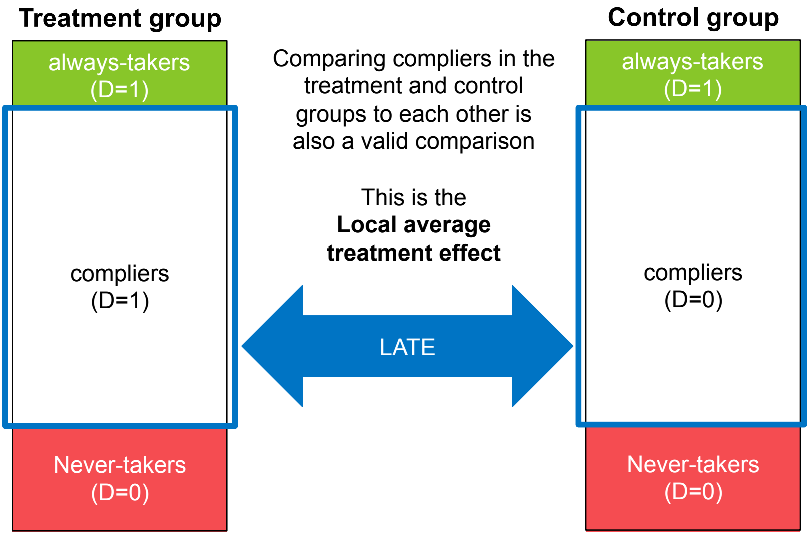
Invalid comparisons



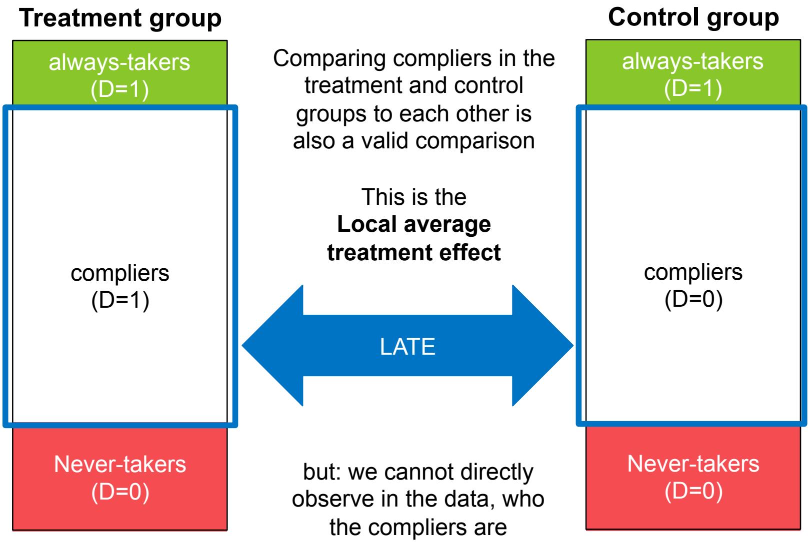
Intention to treat (ITT)



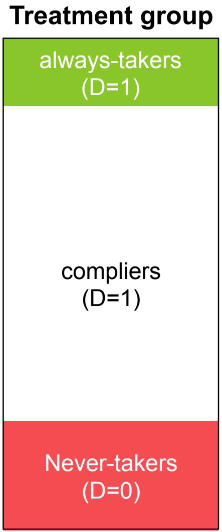
Local average treatment effect (LATE)



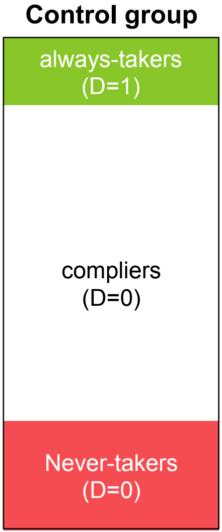
Local average treatment effect (LATE)



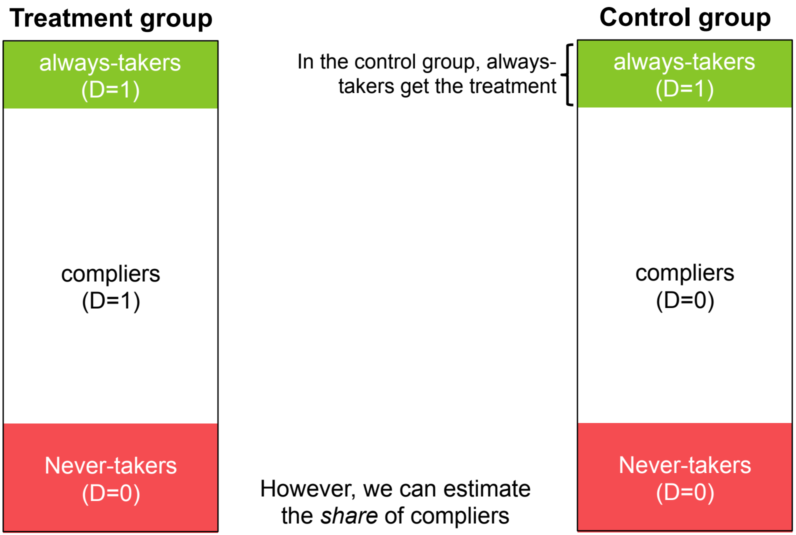
Share of compliers (first-stage)



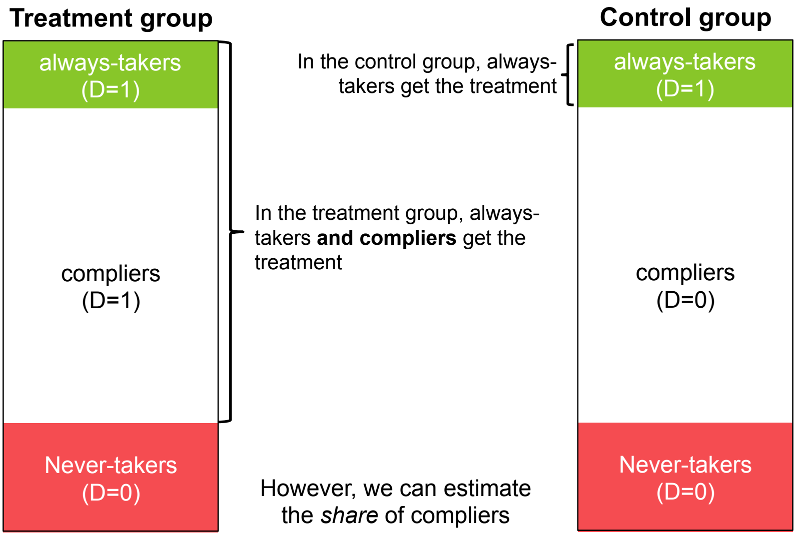
However, we can estimate the *share* of compliers



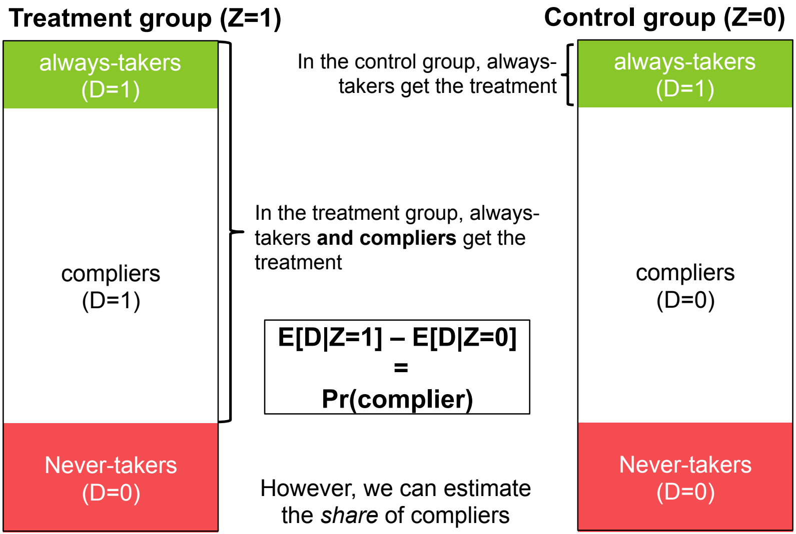
Share of compliers (first-stage)



Share of compliers (first-stage)

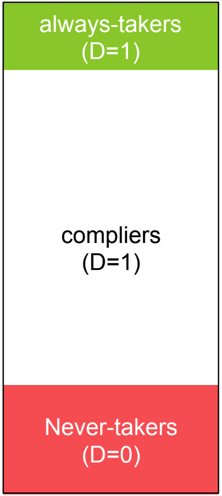


Share of compliers (first-stage)



Interpretting the difference in treatment/control average

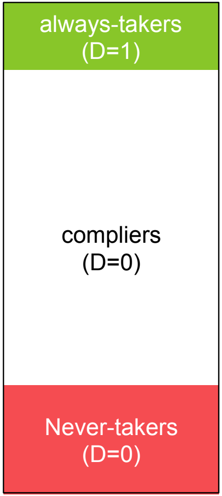
Treatment group (Z=1)



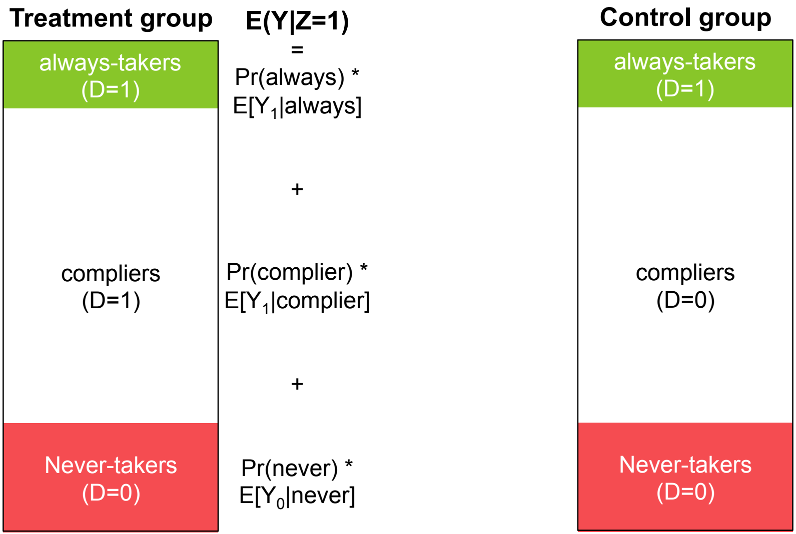
Let's denote the expected outcome of the treatment group as E[Y|Z=1], where Z denotes randomization status.

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.

Control group (Z=0)



Interpretting the difference in treatment/control average



Interpretting the difference in treatment/control average

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)

Interpretting the difference in treatment/control average

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)

Interpretting the difference in treatment/control average

Treatment group

always-takers (D=1)

compliers (D=1)

Never-takers (D=0)

$$E(Y|Z=1) - E(Y|Z=0)$$

$$=$$

$$Pr(complier) * E[Y_1|complier] - Pr(complier) * E[Y_0|complier]$$

<=>

$$\boxed{\begin{array}{c} E(Y|Z=1) - E(Y|Z=0) \\ = \\ Pr(complier)* \\ E[Y_1-Y_0|complier] \end{array}}$$

Control group

always-takers (D=1)

compliers (D=0)

Never-takers (D=0)

ITT, LATE and TOT

- We just derived the **Wald estimator**

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

  - Y is the outcome
  - Z is a 0/1 indicator for being randomized into the treatment group
  - D is a 0/1 indicator for actually receiving the treatment

ITT, LATE and TOT

- We just derived the **Wald estimator**

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

  - Y is the outcome
  - Z is a 0/1 indicator for being randomized into the treatment group
  - D is a 0/1 indicator 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)
  - $\beta_{LATE} = \mathbb{E}[Y_1 - Y_0|\text{complier}]$ is the local average treatment effect
    - the impact of receiving the treatment for the *compliers*
    - may differ from the impact on never-takers and always-takers

ITT, LATE and TOT

- We just derived the **Wald estimator**

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

  - Y is the outcome
  - Z is a 0/1 indicator for being randomized into the treatment group
  - D is a 0/1 indicator 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)
  - $\beta_{LATE} = \mathbb{E}[Y_1 - Y_0|\text{complier}]$ is the local average treatment effect
    - 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
  - you'll see more in later courses (no need to get this now)

ITT, LATE and TOT

- Sometimes ITT is the most relevant estimate
  - in the context of the MTO, it is the impact of *offering* 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)

ITT, LATE and TOT

- Sometimes ITT is the most relevant estimate
  - in the context of the MTO, it is the impact of *offering* 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 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")

ITT, LATE and TOT

- Sometimes ITT is the most relevant estimate
  - in the context of the MTO, it is the impact of *offering* 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 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 never-takers LATE = TOT (average treatment effect on the treated)
    - this is the case in MOT, thus the "TOT"s in the tables

ITT, LATE and TOT

- Sometimes ITT is the most relevant estimate
  - in the context of the MTO, it is the impact of *offering* 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 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 never-takers LATE = TOT (average treatment effect on the treated)
    - this is the case in MOT, thus the "TOT"s in the tables
- We now have everything we need to understand the MTO results

Table 3—Impacts of MTO on Children's Income in Adulthood

	Individual earnings 2008–2012 ($)	
	ITT (2)	TOT (4)
Panel A. Children < age 13 at random assignment		
Exp. versus control	1,624.0**	3,476.8**
	(662.4)	(1,418.2)
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 ($) 2008–2012	Individual earnings 2008–2012 ($)		
	ITT	ITT	ITT w/ controls	TOT
	(1)	(2)	(3)	(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 ($) 2008–2012	Individual earnings 2008–2012 ($)		
	ITT	ITT	ITT w/ controls	TOT
	(1)	(2)	(3)	(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 ($) 2008–2012 ITT (1)	Individual earnings 2008–2012 ($) ITT (2)	Individual earnings 2008–2012 ($) ITT w/ controls (3)	Individual earnings 2008–2012 ($) TOT (4)	Individual earnings ($) Age 26 ITT (5)	Individual earnings ($) 2012 ITT (6)	Employed (%) 2008–2012 ITT (7)	Hhold. inc. ($) 2008–2012 ITT (8)	Inc. growth ($) 2008–2012 ITT (9)
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

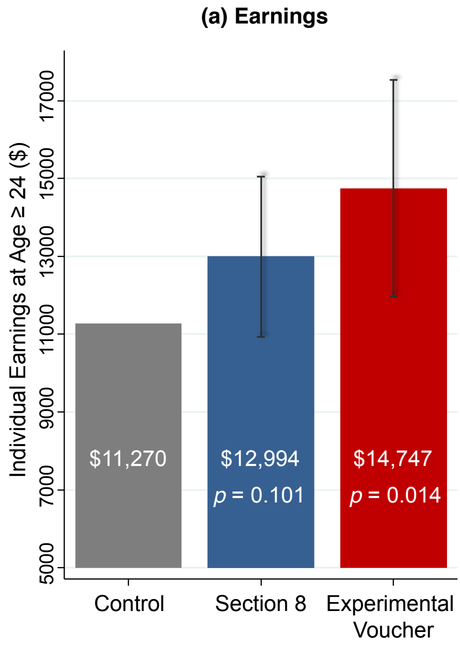
Chetty, Hendren, Katz (2016): The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment

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

	W-2 earnings ($) 2008–2012 ITT (1)	Individual earnings 2008–2012 ($) ITT (2)	Individual earnings 2008–2012 ($) ITT w/ controls (3)	Individual earnings 2008–2012 ($) TOT (4)	Individual earnings ($) Age 26 ITT (5)	Individual earnings ($) 2012 ITT (6)	Employed (%) 2008–2012 ITT (7)	Hhold. inc. ($) 2008–2012 ITT (8)	Inc. growth ($) 2008–2012 ITT (9)
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

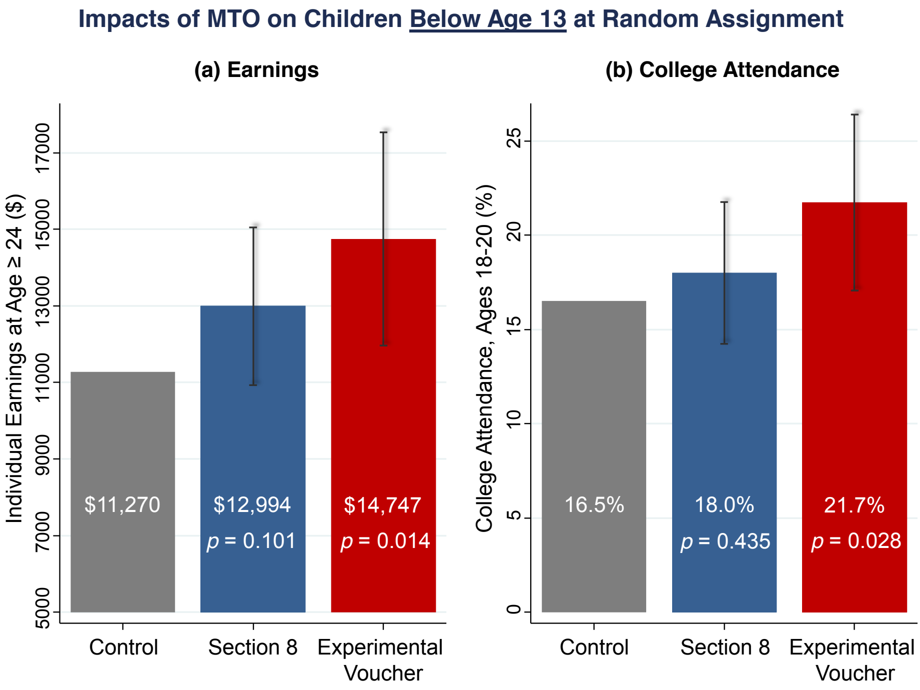
Chetty, Hendren, Katz (2016): The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment

Impacts of MTO on Children Below Age 13 at Random Assignment

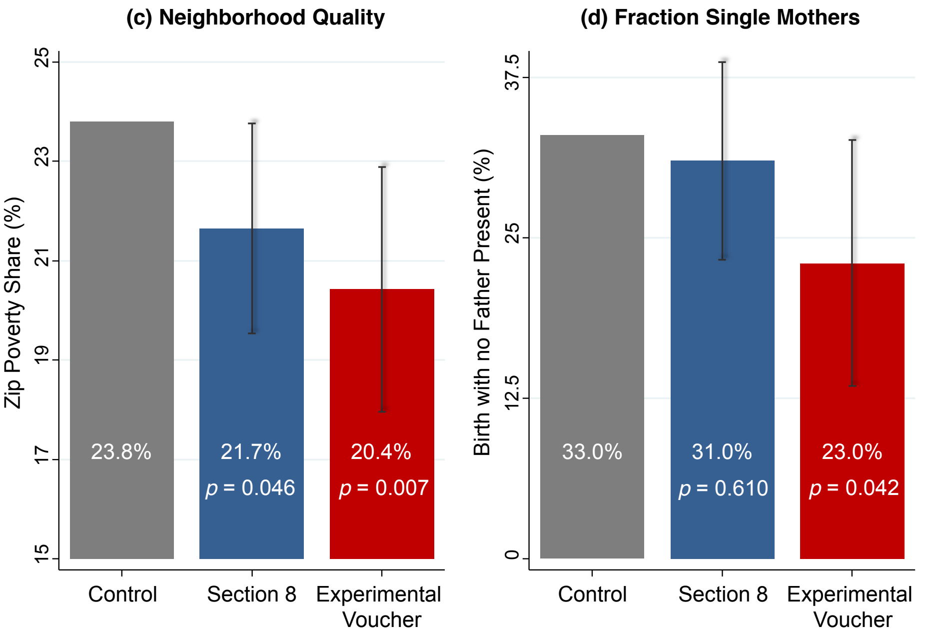


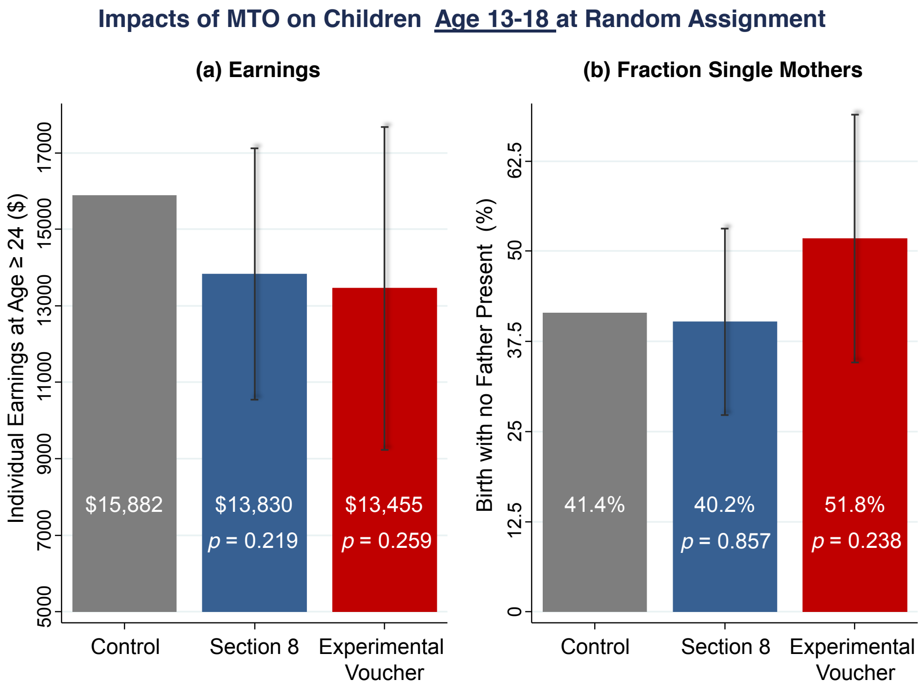
95% confidence intervals = [ToT-1.96*SE, ToT+1.96*SE] = [$698, $6,255] (for the experimental group)

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
(a) Earnings
Individual Earnings at Age ≥ 24 ($)
5000
7000
9000
11000
13000
15000
17000
$11,270
$12,994
p = 0.101
$14,747
p = 0.014
Control
Section 8
Experimental Voucher
(b) College Attendance
College Attendance, Ages 18-20 (%)
0
5
10
15
20
25
16.5%
18.0%
p = 0.435
21.7%
p = 0.028
Control
Section 8
Experimental Voucher

Impacts of MTO on Children Below Age 13 at Random Assignment




Impacts of MTO on Children Age 13-18 at Random Assignment
(a) Earnings
Individual Earnings at Age ≥ 24 ($)
5000
7000
9000
11000
13000
15000
17000
$15,882
$13,830
p = 0.219
$13,455
p = 0.259
Control
Section 8
Experimental Voucher
(b) Fraction Single Mothers
Birth with no Father Present (%)
0
12.5
25
37.5
50
62.5
41.4%
40.2%
p = 0.857
51.8%
p = 0.238
Control
Section 8
Experimental Voucher

Take-aways from the MTO experiment

- Strong evidence on the existence of neighborhood effects
  - might seem obvious, but hard evidence on them is scarce

Take-aways from the MTO experiment

- Strong evidence on the existence of neighborhood effects
  - might seem obvious, but hard evidence on them is scarce
- Putting the effects into a 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?
    - you'll discuss these points in more depth with Tuukka

Take-aways from the MTO experiment

- Strong evidence on the existence of neighborhood effects
  - might seem obvious, but hard evidence on them is scarce
- Putting the effects into a 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?
    - you'll discuss these points in more depth with Tuukka
- 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

Limits of RCTs

- RCTs are often the best way the evaluate the impact of "treatments"
  - simple and transparent $\rightarrow$ everyone can understand the results
  - requires less (untestable) assumptions than the alternative approaches

Limits of RCTs

- RCTs are often the best way the evaluate the impact of "treatments"
  - simple and transparent $\rightarrow$ everyone can understand the results
  - requires less (untestable) assumptions than the alternative approaches
- So, why don't we always run an RCT?
  - ethical and practical limitations
  - fundamental limitations

Ethical and practical limitations of RCTs

- Experiments should not knowingly harm anyone
  - but we still need to understand the effect of potentially harmful things

Ethical and practical limitations of RCTs

- Experiments should not knowingly harm anyone
  - but we still need to understand the effect of potentially harmful things
- Meaningful experiments are sometimes very expensive
  - on the other hand, policy and business mistakes can also be very costly
    $\rightarrow$ even large investments in experimentation can be justified

Ethical and practical limitations of RCTs

- Experiments should not knowingly harm anyone
  - but we still need to understand the effect of potentially harmful things
- Meaningful experiments are sometimes very expensive
  - on the other hand, policy and business mistakes can also be very costly
    $\rightarrow$ even large investments in experimentation can be justified
- The relevant time horizon may be very long
  - sometimes many decades!

Ethical and practical limitations of RCTs

- Treatments are often a bundle of many things
  - then we don't learn the importance of individual components
  - on the other hand, we can overcome this limitation by running several experiments where we vary one component at a time
- Attrition
  - often hard to follow participants over long periods of time
  - less of a problem with administrative data
  - this is equally an issue also for other approaches than RCTs

Ethical and practical limitations of RCTs

- Treatments are often a bundle of many things
  - then we don't learn the importance of individual components
  - on the other hand, we can overcome this limitation by running several experiments where we vary one component at a time
- Attrition
  - often hard to follow participants over long periods of time
  - less of a problem with administrative data
  - this is equally an issue also for other approaches than RCTs
- Hawthorne and John Henry Effects
  - the evaluation itself may push people to change their behavior
  - likely less of a problem with administrative data and long follow-up periods (subjects not reminded about being evaluated)

Fundamental limitations of RCTs

- Spillovers / general equilibrium (GE) effects
  - the treatment also affects the control group $\rightarrow$ cannot use the control group to infer what would have happened without the treatment
  - the GE effects may be the main value of some treatments
    - RCTs never capture economy-wide GE effects
    - some examples of measuring more limited spillovers with RCTs exist

Fundamental limitations of RCTs

- Spillovers / general equilibrium (GE) effects
  - the treatment also affects the control group $\rightarrow$ cannot use the control group to infer what would have happened without the treatment
  - the GE effects may be the main value of some treatments
    - RCTs never capture economy-wide GE effects
    - some examples of measuring more limited spillovers with RCTs exist
- Scarcity of potential observations
  - some treatments affect entire countries or even the whole world
  - we'll never have experimental designs for these treatments

Summary

- Methodological take-away: 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
- RCTs are a powerful tool, but they also have important limitations
  - alternative 1: quasi-experimental approaches,
  - alternative 2: "structural" methods