IV II &
RDD II

Session 12

PMAP 8521: Program evaluation
Andrew Young School of Policy Studies

Plan for today

Treatment effects and compliance

Randomized promotion

Fuzzy regression discontinuity

Treatment effects and compliance

Potential outcomes

$$\delta = (Y|\, P = 1) - (Y|\, P = 0)$$

δ (delta) = causal effect

P = Program

Y = Outcome

$$\delta = Y_1 - Y_0$$

Fundamental problem of causal inference

$$\delta_i = Y_i^1 - Y_i^0 \quad \text{in real life is} \quad \delta_i = Y_i^1 - ???$$

Individual-level effects are impossible to observe!

Average treatment effect

Difference between average/expected value when program is on vs. expected value when program is off

$$\text{ATE} = E(Y_1 - Y_0) = E(Y_1) - E(Y_0)$$

Can be found for a whole population, on average

$$\delta = (\bar{Y} \mid P = 1) - (\bar{Y} \mid P = 0)$$

Every individual has a treatment/causal effect

ATE = average of all unit-level causal effects

ATE = Average effect for the whole population

Other versions of causal effects

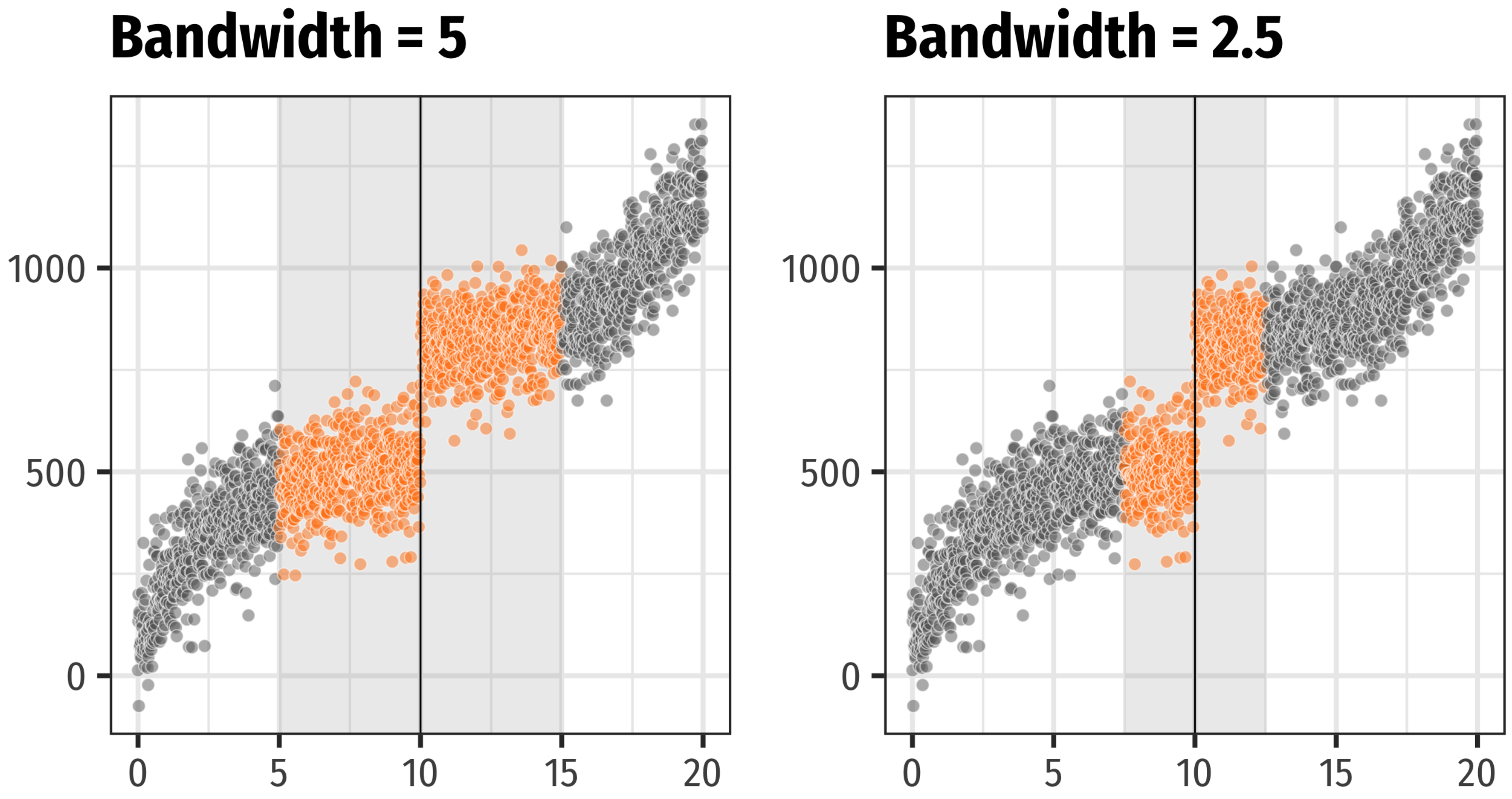
Average treatment on the treated

ATT/TOT

Conditional average treatment effect

CATE

Local effects



LATE

Local average treatment effect (LATE) = weighted ATE

Narrower effect; only applies to some of the population

You can't make population-level claims with LATE

(But that can be okay!)

LATE

In RDD, LATE = people in the bandwidth

In RCTs and IVs, LATE = **compliers**

Compliance

Complier

Treatment
follows assignment

Always taker

Gets treatment
regardless of assignment

Never taker

Rejects treatment
regardless of assignment

Defier

Does the opposite
of assignment

Ignoring defiers

We can generally assume that defiers don't exist

In drug trials this makes sense; you can't get access to medicine without being in treatment group

In development it can make sense; in a bed net RCT, a defier assigned to treatment would have to tear down all existing bed nets out of spite

Ignoring defiers

Monotonicity assumption

Assignment to treatment only has an effect in one direction

Assignment to treatment can only increase—not decrease—your actual chance of treatment

More causal effects

Intent to treat (ITT)

Effect of assignment (not actual treatment!)

Assigned to treatment

Compliers + always takers — Y

Never takers — N

Assigned to control

Y — Always takers

N — Compliers + never takers

More causal effects

Complier Average Causal Effect (CACE)

LATE for the compliers

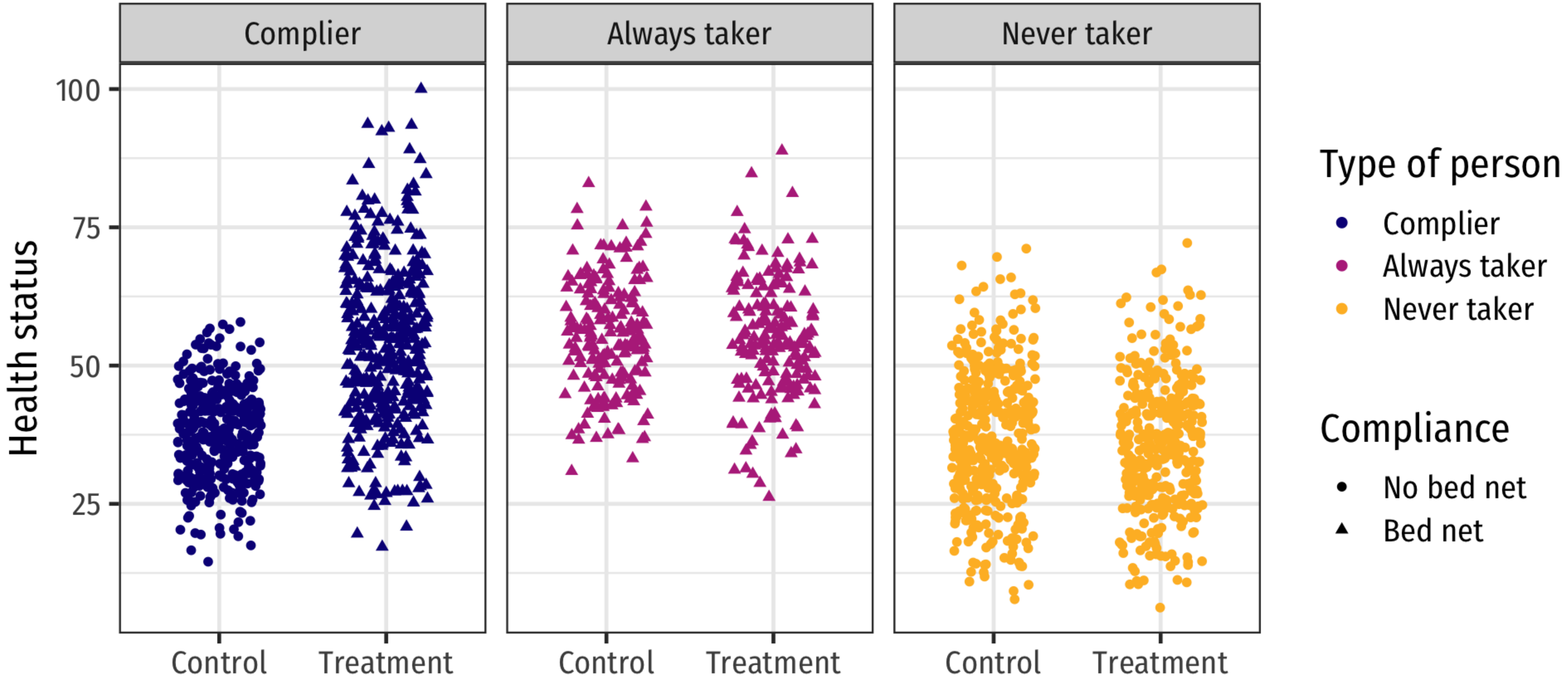
Assigned to treatment			Assigned to control	
Compliers + always takers	Y		Y	Always takers
Never takers	N		N	Compliers + never takers

Hypothetical bed net program

An NGO distributes mosquito bed nets to help improve health by reducing malaria infection rate

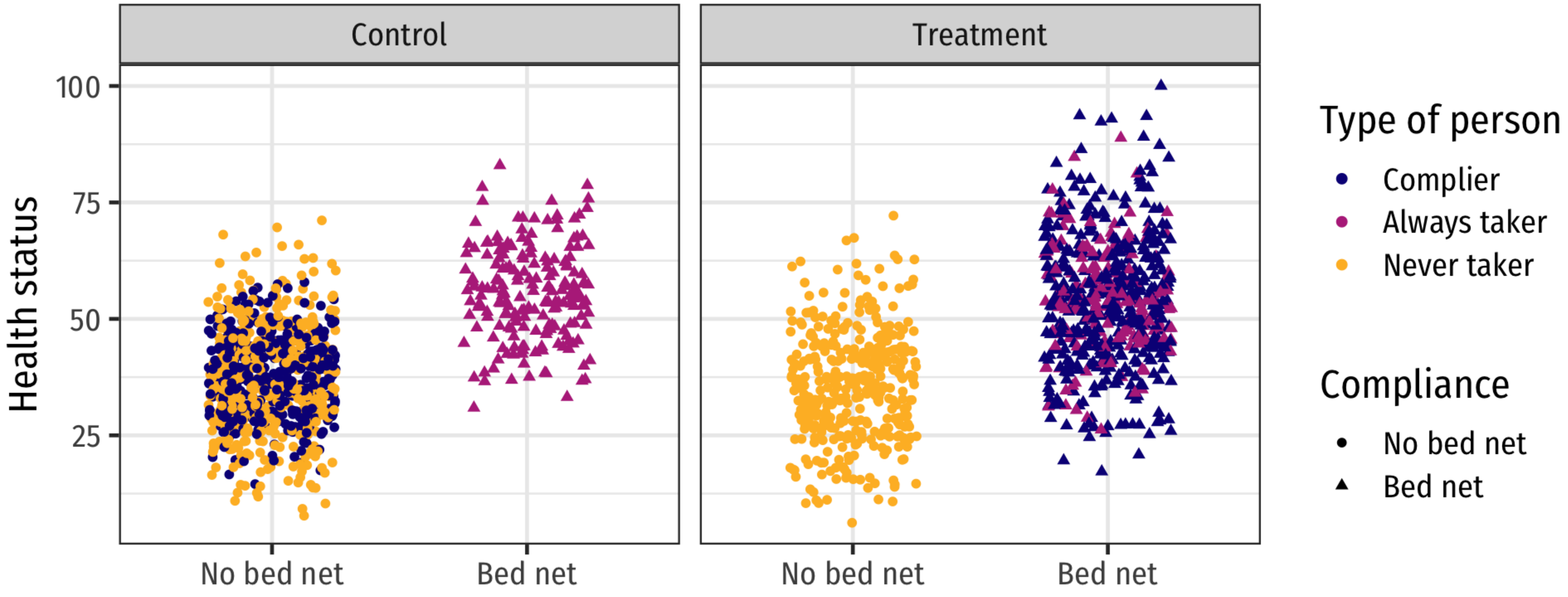
We can read everyone's minds and we know if people are always takers, never takers, or compliers

Mind reading



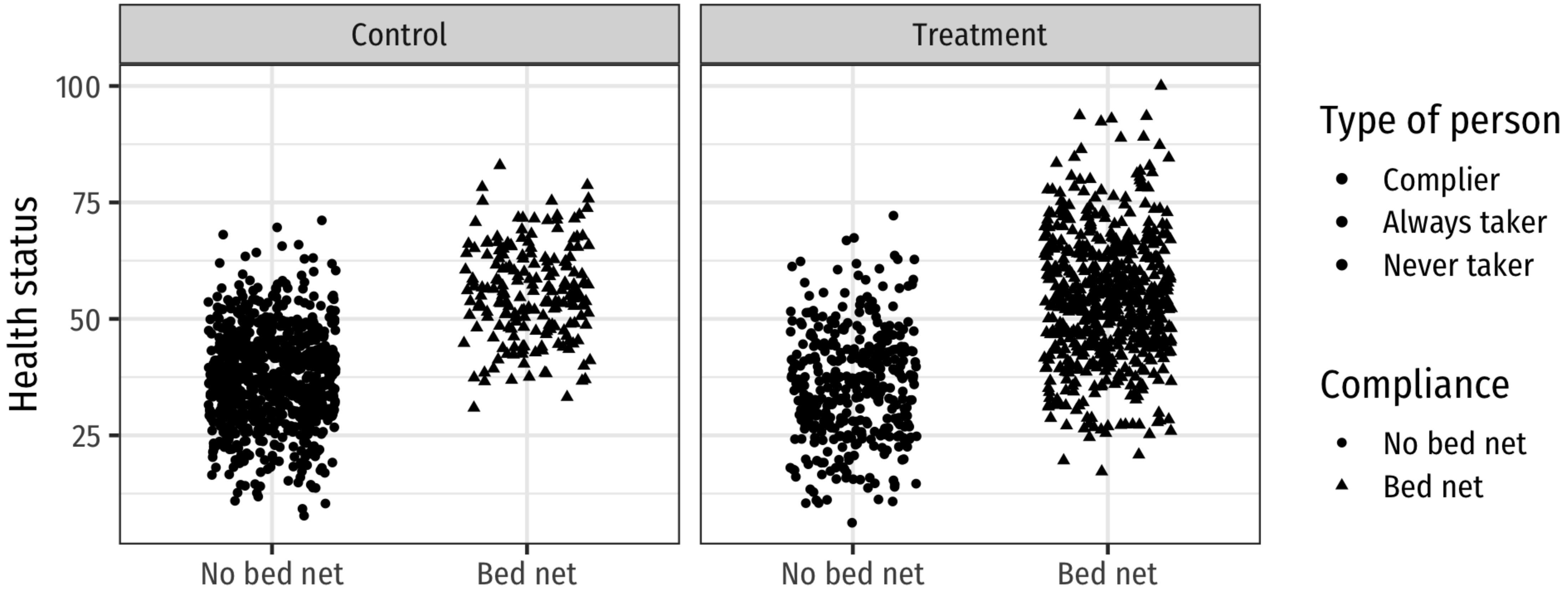
Actual data

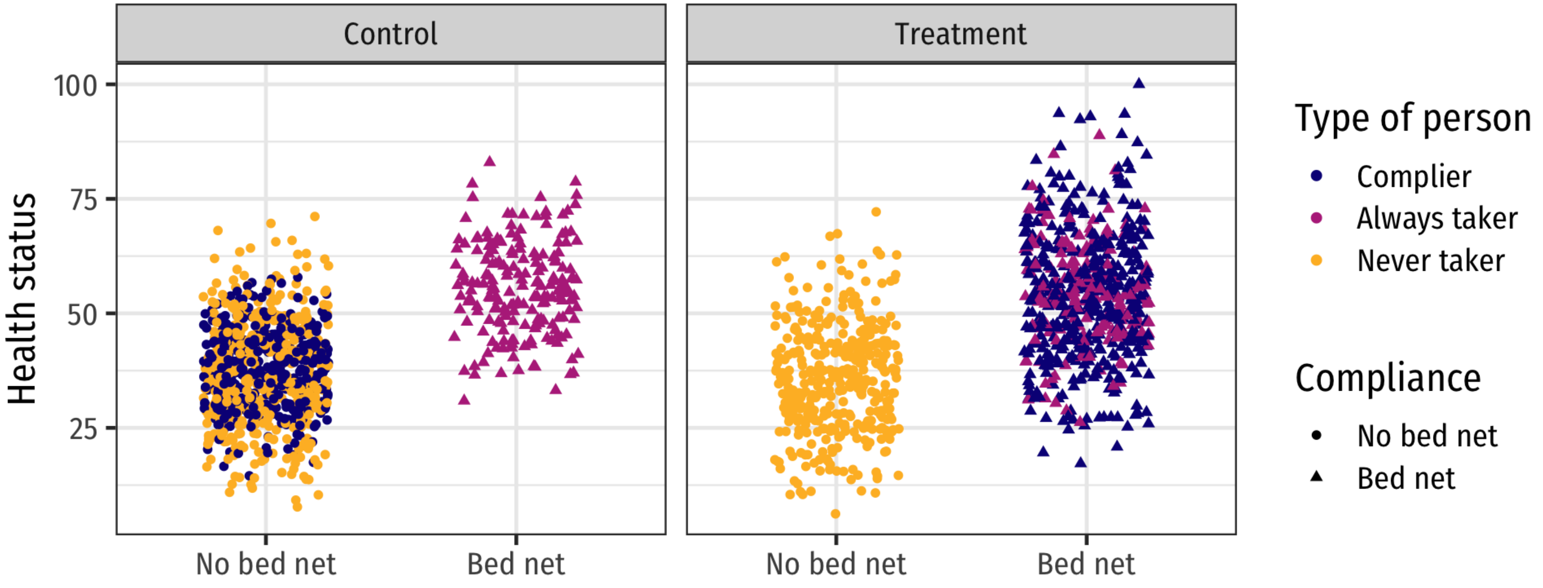
But we can't read minds! This is what we actually see:



Actual data

(Actually *this* is what we see)





Assigned to control

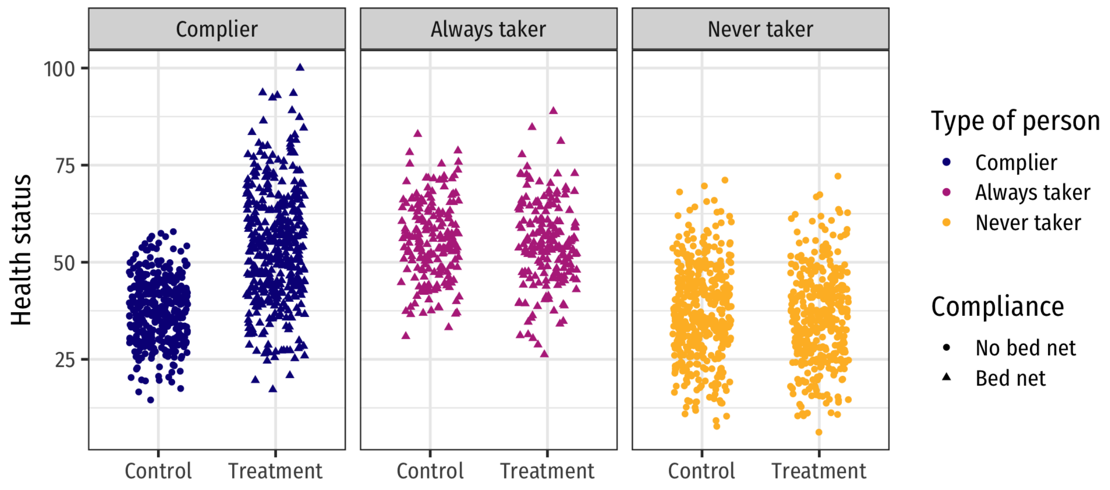
π_A — Y

$\pi_C + \pi_N$ — N

Assigned to treatment

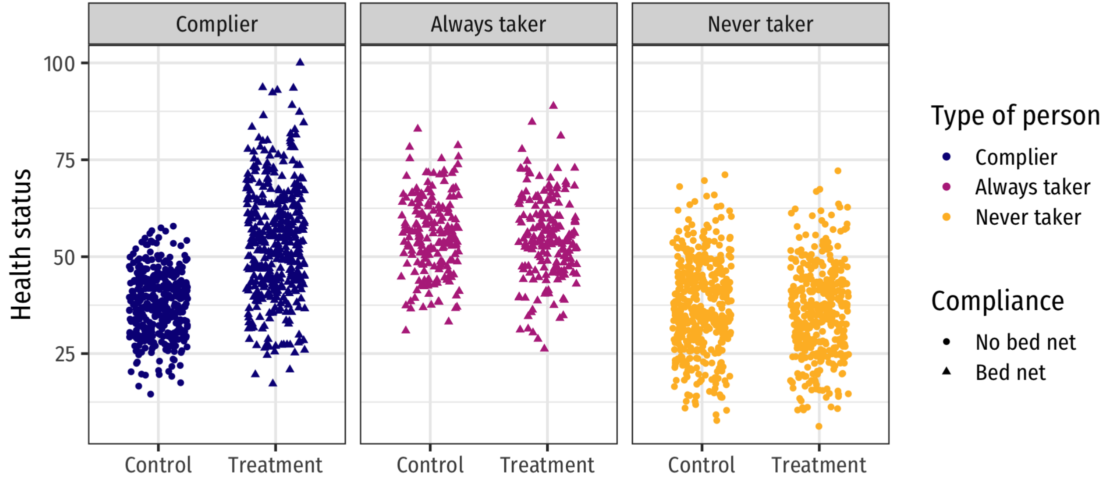
Y — $\pi_C + \pi_A$

N — π_N



$$\begin{aligned}\text{ITT} = \; & \pi_{\text{compliers}} \times (\text{T} - \text{C})_{\text{compliers}} + \\ & \pi_{\text{always takers}} \times (\text{T} - \text{C})_{\text{always takers}} + \\ & \pi_{\text{never takers}} \times (\text{T} - \text{C})_{\text{never takers}}\end{aligned}$$

$$\text{ITT} = \; \pi_{\text{C}}\text{CACE} + \pi_{\text{A}}\text{ATACE} + \pi_{\text{N}}\text{NTACE}$$



$$\text{ITT} = \pi_{\text{C}}\text{CACE} + \pi_{\text{A}}\text{ATACE} + \pi_{\text{N}}\text{NTACE}$$

Treatment received is same regardless of assignment!
Being assigned to treatment doesn't influence ATs and NTs

$$\text{ITT} = \pi_{\text{C}}\text{CACE} + \pi_{\text{A}} \times 0 + \pi_{\text{N}} \times 0$$

$$\begin{aligned}\text{ITT} &= \pi_{\text{C}}\text{CACE} + \pi_{\text{A}}\text{ATACE} + \pi_{\text{N}}\text{NTACE} \\ &= \pi_{\text{C}}\text{CACE} + \pi_{\text{A}} \times 0 + \pi_{\text{N}} \times 0 \\ \text{ITT} &= \pi_{\text{C}}\text{CACE}\end{aligned}$$

$$\text{CACE} = \frac{\text{ITT}}{\pi_{\text{C}}}$$

ITT and π_C are both findable!

Finding the ITT

ITT = effect of assignment to treatment on outcome

$$\text{ITT} = (\bar{y} \mid \text{Treatment}) - (\bar{y} \mid \text{Control})$$

```
bed_nets %>%
  group_by(treatment) %>%
  summarize(avg = mean(health))
```

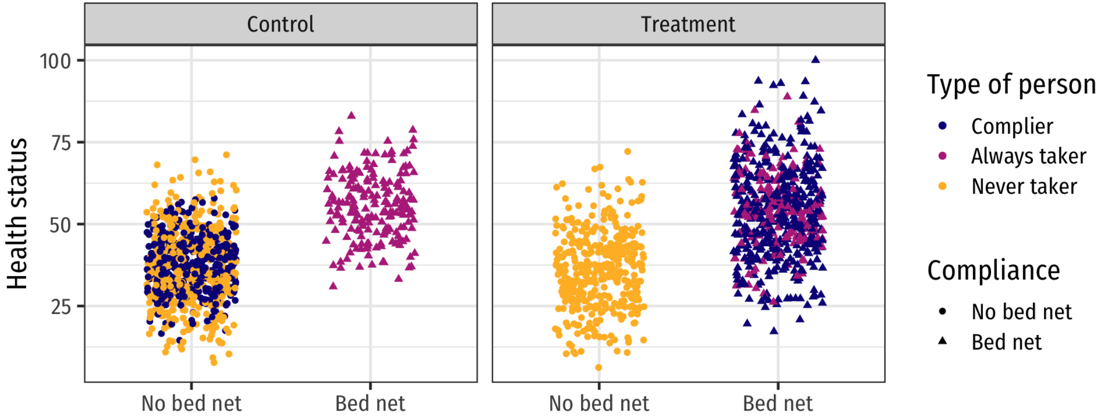
```
## # A tibble: 2 x 2
##   treatment   avg
##   <chr>      <dbl>
## 1 Control     40.9
## 2 Treatment   46.9
```

```
itt_model <- lm(health ~ treatment,
                data = bed_nets)
tidy(itt_model)
```

```
## # A tibble: 2 x 2
##   term               estimate
##   <chr>                 <dbl>
## 1 (Intercept)           40.9
## 2 treatmentTreatment     5.99
```

Finding the π_C

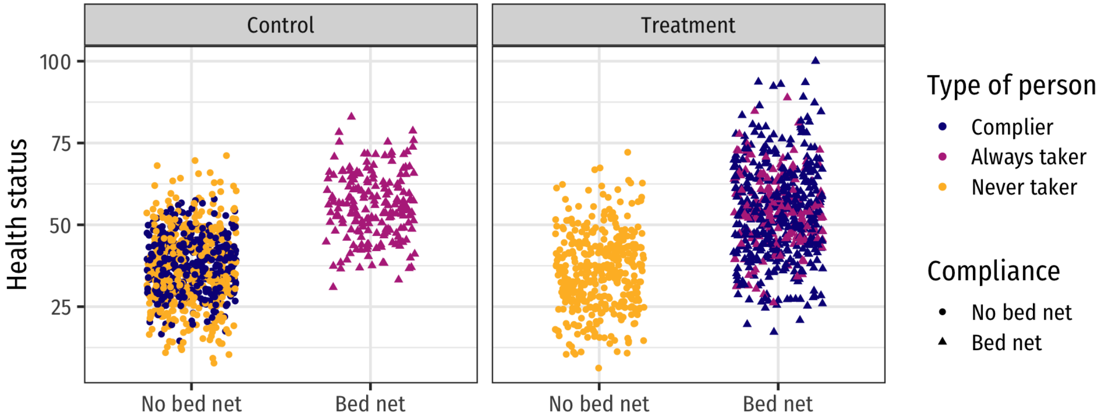
People in treatment group who complied are a combination of Always Takers and Compliers



$$\pi_A + \pi_C = \% \text{ yes in treatment; or}$$
$$\pi_C = \% \text{ yes in treatment} - \pi_A$$

Can we know π_A?

$$\pi_C = \% \text{ yes in treatment} - \pi_A$$



We can assume that the proportion of Always Takers is the same across treatment and control

We know how many people were in control but still used nets—that's π_A!

Isolating π_C

$$
\begin{aligned}
\pi_{\text{C}} &= \text{\% yes in treatment} - \pi_{\text{A}} \\
&= \text{\% yes in treatment} - \text{\% yes in control}
\end{aligned}
$$

```r
bed_nets %>%
  group_by(treatment, bed_net) %>%
  summarize(n = n()) %>%
  mutate(prop = n / sum(n))
```

```
## # A tibble: 4 x 4
## # Groups:   treatment [2]
##   treatment bed_net        n  prop
##   <chr>     <fct>      <int> <dbl>
## 1 Control   No bed net   808 0.805
## 2 Control   Bed net      196 0.195
## 3 Treatment No bed net   388 0.390
## 4 Treatment Bed net      608 0.610
```

```r
# pi_c = prop yes in treatment -
#          prop yes in control
pi_c <- 0.6104418 - 0.1952191
pi_c
```

```
## [1] 0.4152227
```

41.5% compliers!

Finding the CACE, finally!

$$\text{CACE} = \frac{\text{ITT}}{\pi_\text{C}}$$

```r
ITT <- tidy(itt_model) %>%
  filter(term == "treatmentTreatment") %>%
  pull(estimate)
ITT
```

```
## [1] 5.987992
```

```r
pi_c
```

```
## [1] 0.4152227
```

```r
CACE <- ITT / pi_c
CACE
```

```
## [1] 14.42116
```

Bed nets *cause* 14.4 point increase in health for compliers

$$\text{CACE} = \frac{\text{ITT}}{\pi_{\text{C}}}$$

$$\text{ITT} = (\bar{y} \mid \text{Treatment}) - (\bar{y} \mid \text{Control})$$

$$\begin{aligned} \pi_{\text{C}} \;=\; & \,\%\text{ yes in treatment} - \\ & \,\%\text{ yes in control} \end{aligned}$$

A faster way with 2SLS

LATE for the compliers

If you use assignment to treatment as an instrument, you can find the causal effect for just compliers

Instrumental variables in general give you the CACE

CACE with 2SLS

```
model_2sls <- iv_robust(health ~ bed_net | treatment,
                        data = bed_nets)
tidy(model_2sls)
```

```
##               term estimate std.error statistic      p.value
## 1      (Intercept) 38.12285 0.5150818  74.01320 0.000000e+00
## 2 bed_netBed net 14.42116 1.2538198  11.50178 1.086989e-29
```

Same 14.421 effect!

Promotion as an instrument

Universal programs

What if you have a program that anyone can opt in to?

ACA, voting, employer retirement matching

You can't just look at outcomes of participants vs. non-participants!

Selection bias!

You can't randomly assign people to it either

Ethics!

Randomized promotion

What if you *encourage* some people to participate?

What if the encouragement is randomized?

Valid treatment/control groups?

Not really...

Randomized promotion

...but also, kind of!

Encouragement/promotion = an instrument!

Not something weird? Does that work!?

Relevant?

$Z \rightarrow X \quad Cor(Z, X) \neq 0$

Promotion causes people to use the program. Yep.

Exclusive?

$Z \rightarrow X \rightarrow Y \quad Z \not\rightarrow Y \quad Cor(Z, Y \mid X) = 0$

Promotion causes outcome *only through* program? Yep.

Exogenous?

$U \not\rightarrow Z \quad Cor(Z, U) = 0$

Unobserved things that influence outcome don't also influence promotion? Yep.

Program compliance

Always Takers

People who will always enroll in program

Never Takers

People who will never enroll in program

Compliers / Enrollers-if-Promoted

People who will enroll in the program if encouraged to

LATE for compliers

id	outcome	program	promotion
1	45	TRUE	TRUE
2	55	TRUE	FALSE
3	52	FALSE	FALSE
4	39	FALSE	TRUE

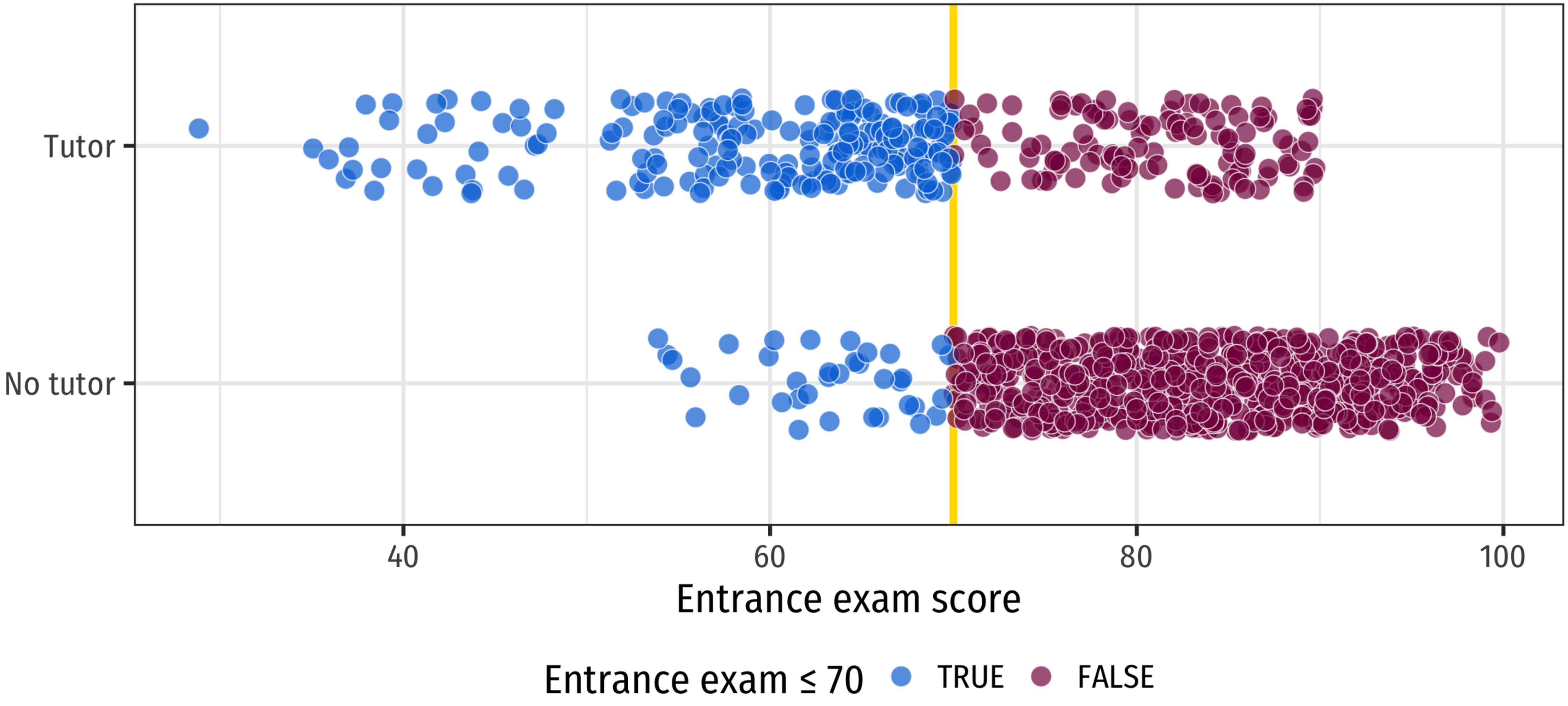
```
iv_robust(outcome ~ program | promotion)
```

This will show the LATE for promoted-ees!

Says nothing about the effect of the program on Always Takers or Never Takers

Fuzzy RDD

Fuzzy discontinuities



Fuzzy discontinuities

Fuzzy discontinuities imply noncompliance

Address noncompliance with instrumental variables

What do we use as instrument?

Instrument = above/below cutoff

i.e. what they were supposed to do

(This is just like the CACE we just did!)

Not something weird? Does that work!?

Relevant?

$Z \longrightarrow X \quad \mathrm{Cor}(Z, X) \neq 0$

Cutoff causes program? Yep.

Exclusive?

$Z \longrightarrow X \longrightarrow Y \quad Z \not\longrightarrow Y \quad \mathrm{Cor}(Z, Y \mid X) = 0$

Cutoff causes outcome *only through* program? Yep.

Exogenous?

$U \not\longrightarrow Z \quad \mathrm{Cor}(Z, U) = 0$

**Unobserved things that influence outcome don't also influence cutoff?
It's an arbitrary cutoff, so sure.**

Doubly local LATE

Effect is only for (1) compliers (2) near the cutoff

Be specific when talking about effects; definitely don't make population-level claims

Parametric fuzzy RD

Step 1: Center running variable + make threshold variable

```
tutoring_centered <- tutoring %>%
  mutate(entrance_centered = entrance_exam - 70,
         below_cutoff = entrance_exam <= 70)
head(tutoring_centered, 6)
```

```
## # A tibble: 6 x 6
##      id entrance_exam tutoring exit_exam entrance_centered below_cutoff
##   <int>         <dbl> <lgl>        <dbl>             <dbl> <lgl>
## 1     1          92.4 FALSE         78.1            22.4   FALSE
## 2     2          72.8 FALSE         58.2             2.77  FALSE
## 3     3          53.7 TRUE          62.0           -16.3   TRUE
## 4     4          98.3 FALSE         67.5            28.3   FALSE
## 5     5          69.7 TRUE          54.1            -0.288 TRUE
## 6     6          68.1 TRUE          60.1            -1.93  TRUE
```

Parametric fuzzy RD

Step 2: Use cutoff as instrument in 2SLS model

```
# Bandwidth ± 10
fuzzy1 <- iv_robust(
  exit_exam ~ entrance_centered + tutoring | entrance_centered + below_cutoff,
  data = filter(tutoring_centered, entrance_centered >= -10 & entrance_centered <= 10)
)

tidy(fuzzy1)
```

```
##                term   estimate  std.error statistic       p.value
## 1       (Intercept) 60.1413558 1.01765573 59.097939 9.746624e-200
## 2 entrance_centered  0.4366281 0.09929619  4.397229  1.407213e-05
## 3      tutoringTRUE  9.7410444 1.91184891  5.095091  5.384163e-07
```

Nonparametric fuzzy RD

Use the `fuzzy` argument in `rdrobust()`

Important! Specify actual treatment status, *not* the instrument of above/below the cutoff

```
rdrobust(y = tutoring$exit_exam, x = tutoring$entrance_exam,
         c = 70, fuzzy = tutoring$tutoring) %>%
  summary()
```

```
## =============================================================================
##         Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
## =============================================================================
##   Conventional     9.683     1.893     5.116     0.000     [5.973 , 13.393]
##         Robust         -         -     4.258     0.000     [5.210 , 14.095]
## =============================================================================
```