Answering Causal Questions Using Non-experimental Data

Roadmap

- Defining causal effects
- Estimating causal effects with randomized experiments
- Estimating causal effects with observational studies
  - Stratification/regression
  - Propensity score methods
  - Instrumental variables
  - Difference in differences
  - Fixed effects
  - Regression discontinuity
  - Sensitivity analysis

Today's roadmap

1. Observational studies
2. Cautionary Tale
3. Types of Quasi-experiments
4. Assumptions/Estimands
5. Methods
    - Subclassification
    - Regression

Polio and the Salk Vaccine

- Polio first appeared in the US in 1916
- In 1954 it had become enough of a public health issue that the Public Health Service was ready to perform a study to investigate the effectiveness of a vaccine invented by Salk
- The evidence from observational studies did not appear to provide evidence that the vaccine was terribly effective, particularly because it was assumed that they were biased in favor of making the vaccine look more effective than it truly was.
- Why was it assumed that the bias worked in that direction?

Observational Design

- Children can only be vaccinated with parents' permission. So NFIP thought to use those gave permission as treatment group and those did not as controls.
- Children of higher-income parents more likely to have permission and more likely to contract polio.
- The difference in response may be *confounded* by having higher income families in the treatment group and lower income families in the control group .
- Given this confounding, what direction to you expect the bias to go (negative or positive)?

Polio and the Salk Vaccine

- Polio first appeared in the US in 1916
- In 1954 it had become enough of a public health issue that the Public Health Service was ready to perform a study to investigate the effectiveness of a vaccine invented by Salk
- The evidence from observational studies did not appear to provide evidence that the vaccine was terribly effective, particularly because it was assumed that they were biased in favor of making the vaccine look more effective than it truly was.
- However, some school districts agreed to perform randomized controlled trials and the results from these experiments suggest that the vaccine was very effective
- Many lives were saved

Observational Studies

Observational Studies

- We will use this term to refer to any type of study where the treatment/causing variable has not been randomly assigned
- They are not, in general, well suited for answering causal questions, though they can be planned or analyzed in ways that minimize potential problems

Observational studies: flawed for causal inference

- The key feature in observational studies is that observations "self-select" into treatment groups
- This means that, in addition to receiving different "treatments" (welfare programs, pollution levels in the town they live in, child care services, drugs, job training) the people in each group are likely to be different in other important ways (age, income levels, racial composition, "motivation", health)
- If these characteristics predict outcomes as well they are called *confounding covariates*
- So if we see differences in outcomes across treatment groups, we don't know if this is caused by the treatment or is instead related to the confounding covariates.
- The bias caused by this self-selection is often referred to as *selection bias*

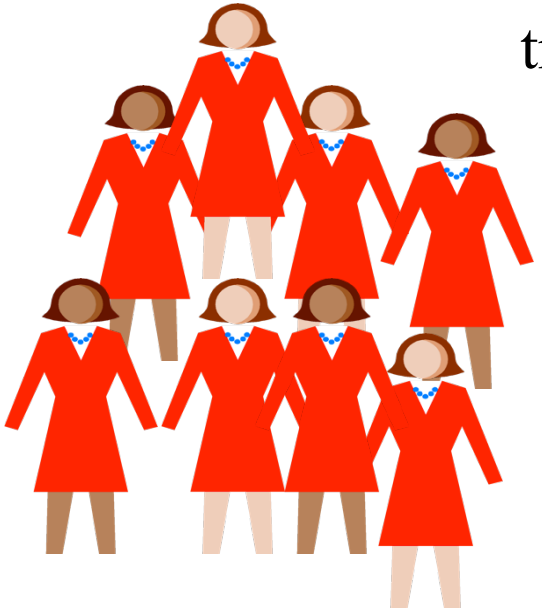
Randomized controlled experiments that can start to look like observational studies

- Randomized experiments:
  - "Broken" randomized experiments (broken through missing data, noncompliance, failure to randomize faithfully) (NYC school choice example, Tennessee Star)
  - Randomized encouragement designs (randomized encouragement for flu shots)
- Natural experiments (lottery example, Vietnam draft)

Self-selection into treatment groups

Receives placebo

Receives treatment



Definitions

The following definitions will be used

- Treatment, *Z*: A manipulable variable we want to know the effect of (pollution, a vaccine, welfare work requirements, tax increase, child maltreatment)
- Outcome, *Y*: The variable we hypothesize might be affected by the treatment
- Covariates or confounders, *X*: Background variables that we think may influence the treatment that people participate in as well as their outcome

Hypothetical Example: observational study

Person	Treat	Educ.	Age	Y0	Y1	Y
1	1	1	26	10	14	14
2	1	1	21	8	12	12
3	1	1	30	12	16	16
4	1	1	19	8	12	12
5	1	0	25	6	10	10
6	1	0	22	4	8	8
7	0	0	21	4	8	4
8	0	0	26	6	10	6
9	0	0	28	8	12	8
10	0	0	20	4	8	4
11	0	1	26	10	14	10
12	0	1	21	8	12	8
13	0	0	16	2	6	2
14	0	0	15	1	5	1

$$\bar{e}_{Z=1} - \bar{e}_{Z=0} = .67 - .25 = .42$$

$$\bar{a}_{Z=1} - \bar{a}_{Z=0} = 23.83 - 21.63 = 2.2$$

$$\bar{Y}_{Z=1} - \bar{Y}_{Z=0} = 12.0 - 5.375 = 6.625$$

Observational studies: causal effects

- Suppose we are interested in the ATE (Average Treatment Effect),

$$E[Y(1)-Y(0)] = E[Y(1)] - E[Y(0)]?$$

- Well in an observational study
  - $E[Y(1)] \neq E[Y(1)|Z=1]$
  - $E[Y(0)] \neq E[Y(0)|Z=0]$
- So we *cannot* unbiasedly estimate
  - $E[Y(1)]$ with $\overline{Y}_{Z=1}$ or
  - $E[Y(0)]$ with $\overline{Y}_{Z=0}$
- So the problem remains

Observational studies: causal effects

- What about the average causal *effect of the treatment on the treated*,

  $E[Y(1)-Y(0) \mid Z=1] = E[Y(1) \mid Z=1] - E[Y(0) \mid Z=1]$?
- Well in an observational study
  - $E[Y(0) \mid Z=1] \neq E[Y(0) \mid Z=0]$
- So we *cannot* unbiasedly estimate
  - $E[Y(0) \mid Z=1]$ with $\overline{Y}_{Z=0}$
- So the problem remains

Hypothetical Example

Person	Treat	Educ.	Age	Y0	Y1	Y
1	1	1	26	10	14	14
2	1	1	21	8	12	12
3	1	1	30	12	16	16
4	1	1	19	8	12	12
5	1	0	25	6	10	10
6	1	0	22	4	8	8
7	0	0	21	4	8	4
8	0	0	26	6	10	6
9	0	0	28	8	12	8
10	0	0	20	4	8	4
11	0	1	26	10	14	10
12	0	1	21	8	12	8
13	0	0	16	2	6	2
14	0	0	15	1	5	1

These data were sample from a population with the following properties:

$$E[Y(0)] = 6.5;\ \ E[Y(0) \mid Z = 0] = 5.375$$

$$E[Y(1)] = 10.5;\ \ E[Y(1) \mid Z = 1] = 12$$

Solutions to selection bias problem?

- How can we fix this?
- What if we knew that education level was the only confounding covariate?

Hypothetical Example (τ_c=4)

Person	Treat	Educ.	Age	Y0	Y1	Y
1	1	1	26	10	14	14
2	1	1	21	8	12	12
3	1	1	30	12	16	16
4	1	1	19	8	12	12
5	1	0	25	6	10	10
6	1	0	22	4	8	8
7	0	0	21	4	8	4
8	0	0	26	6	10	6
9	0	0	28	8	12	8
10	0	0	20	4	8	4
11	0	1	26	10	14	10
12	0	1	21	8	12	8
13	0	0	16	2	6	2
14	0	0	15	1	5	1

$$\bar{e}_{Z=1} - \bar{e}_{Z=0} = .67 - .25 = .42$$

$$\bar{a}_{Z=1} - \bar{a}_{Z=0} = 23.83 - 21.63 = 2.2$$

$$\bar{Y}_{Z=1} - \bar{Y}_{Z=0} = 12.0 - 5.375 = 6.625$$

If, however, we stratify on education…

Person	Treat	Educ.	Age	Y0	Y1	Y
1	1	1	26	10	14	14
2	1	1	21	8	12	12
3	1	1	30	12	16	16
4	1	1	19	8	12	12
5	1	0	25	6	10	10
6	1	0	22	4	8	8
7	0	0	21	4	8	4
8	0	0	26	6	10	6
9	0	0	28	8	12	8
10	0	0	20	4	8	4
11	0	1	26	10	14	10
12	0	1	21	8	12	8
13	0	0	16	2	6	2
14	0	0	15	1	5	1

$$educ = 1 : \bar{Y}_{Z=1} - \bar{Y}_{Z=0} = 13.5 - 9 = 4.5$$

$$educ = 0 : \bar{Y}_{Z=1} - \bar{Y}_{Z=0} = 9 - 4.17 = 4.83$$

$$\hat{\tau} = (6/14) * 4.5 + (8/14) * 4.83 = 4.69$$

Now for effect of treatment on treated

Person	Treat	Educ.	Age	Y0	Y1	Y
1	1	1	26	10	14	14
2	1	1	21	8	12	12
3	1	1	30	12	16	16
4	1	1	19	8	12	12
5	1	0	25	6	10	10
6	1	0	22	4	8	8
7	0	0	21	4	8	4
8	0	0	26	6	10	6
9	0	0	28	8	12	8
10	0	0	20	4	8	4
11	0	1	26	10	14	10
12	0	1	21	8	12	8
13	0	0	16	2	6	2
14	0	0	15	1	5	1

$$educ = 1 : \bar{Y}_{Z=1} - \bar{Y}_{Z=0} = 13.5 - 9 = 4.5$$

$$educ = 0 : \bar{Y}_{Z=1} - \bar{Y}_{Z=0} = 9 - 4.17 = 4.83$$

$$\hat{\tau}_{tot} = (2/3) * 4.5 + (1/3) * 4.83 = 4.61$$

Observational studies: stratifying on 1 covariate

- Suppose we want to estimate the average causal effect, E[Y(1)-Y(0)] = E[Y(1)] - E[Y(0)]
- Even though we have an observational study, if we stratify on X (e.g. education) and if X is the only covariate
  - E[Y(1) | X=x] = E[Y(1) | Z=1, X=x] = E[Y | Z=1, X=x]
  - E[Y(0) | X=x] = E[Y(0) | Z=0, X=x] = E[Y | Z=0, X=x]
- So we can unbiasedly estimate
  - E[Y(1) | X=x] with $\bar{Y}_{Z=1,X=x}$ and
  - E[Y(0) | X=x] with $\bar{Y}_{Z=0,X=x}$
- We do this for each level of education and then weight up these individual effects to get the appropriate overall average

Generalizing to adjustments to more covariates

- In our case stratifying on education level removed much of the bias
- Controlling for age (by matching, regression, etc) should remove the rest of the bias (because in this example there were only 2 confounders)
- What is the best way to control for *many* confounding covariates, however?
- The good news is the *theory* extends effortlessly to many covariates (just replace single variable X in the previous slide with a vector **X** of confounding covariates)
- Controlling for many covariates *in practice* however is not always an easy task

Strategies for causal inference with observational data assuming all confounding covariates are observed

- Subclassification/Stratification: Make comparisons of the outcome variable only within like groups of people (stratify by important variables)
- Regression analysis can sometimes be used to control for confounding covariates but only if certain strict assumptions hold (we'll explore soon)
- More fancy adjustments/modeling techniques exist (as we shall discuss)

Formalization of the assumptions

Important: in order for covariance adjustment to result in causal inference we need to make an **untestable** assumption that X is the only confounding covariate, formally,

$$Y(0), Y(1) \perp Z \mid X$$

referred to as "**ignorability** of the assignment mechanism" (Rubin, 1978) or "selection on observables" (Heckman) or "all confounders measured" or "no hidden bias" or "no omitted variables"

- Intuitively this says that two people who are similar on all their X values are equally likely to have received the treatment

!!!Pop Quiz!!!

- Does the set-up for observational studies seem similar to any of the randomized experiments we discussed?
  - Which design is this most similar to? randomized block design
  - How these designs are different? randomization
  - Would we estimate the ATE differently? Why or why not?

Observational studies: stratifying on covariates

Note that the justification for unbiased estimation when you control for covariates is the same as the justification for unbiased estimation in a randomized block experiment with just one exception:

It is important to remember however that in the observational study you also have to assume you have controlled for all potential confounding covariates (ignorability)!

Using linear regression for causal inference

Linear regression: Interpretation of coefficients

Interpret β_1 in each of the following regressions.

1) $Y = \beta_0 + \beta_1 H + \varepsilon$

H is a binary indicator indicating high school was completed

2)...

3)...

4)...

Linear regression: Interpretation of coefficients

Interpret β_1 in each of the following regressions.

1) $Y = \beta_0 + \beta_1 H + \varepsilon$

H is a binary indicator indicating high school was completed

2) $Y = \beta_0 + \beta_1 E_1 + \beta_2 E_2 + \beta_3 E_3 + \varepsilon$

E_1-E_3 indicate the highest level of education completed

(high school, college, grad school)

3) ....

4) ....

Linear regression: Interpretation of coefficients

Interpret β_1 in each of the following regressions.

1) $Y = \beta_0 + \beta_1 H + \varepsilon$

H is a binary indicator indicating high school was completed

2) $Y = \beta_0 + \beta_1 E_1 + \beta_2 E_2 + \beta_3 E_3 + \varepsilon$

E_1-E_3 indicate the highest level of education completed

(high school, college, grad school)

3) $Y = \beta_0 + \beta_1 R_1 + \beta_2 R_2 + \beta_2 R + ... + \beta_{12} R_{12} + \varepsilon$

R_1-R_{12} indicate years of school completed past 8th grade

(high school, college, grad school) in a sample where everyone completed between 0 and 12 years of school past 8th grade

4) ....

Linear regression: Interpretation of coefficients

Interpret β_1 in each of the following regressions.

1) $Y = \beta_0 + \beta_1 H + \varepsilon$

H is a binary indicator indicating high school was completed

2) $Y = \beta_0 + \beta_1 E_1 + \beta_2 E_2 + \beta_3 E_3 + \varepsilon$

E_1-E_3 indicate the highest level of education completed

(high school, college, grad school)

3) $Y = \beta_0 + \beta_1 R_1 + \beta_2 R_2 + \beta_2 R + ... + \beta_{12} R_{12} + \varepsilon$

R_1-R_{12} indicate years of school completed past 8th grade

(high school, college, grad school) in a sample where everyone completed between 0 and 12 years of school past 8th grade

4) $Y = \beta_0 + \beta_1 D + \varepsilon$

D indicates years of schooling completed past 8th grade

Now let’s think about causal inference

Controlling for confounders using regression

- Recall that our goal is to estimate things like E[Y(1) | X] and E[Y(0) | X]
- Given our data limitations (i.e. "the fundamental problem of causal inference" we actually model E[Y | X]
- Here X can represent many covariates.
- Technically this is the definition of a regression. Simply put, a regression is just a way of describing the relationship between the means of different subpopulations defined by X.
- Linear regression says these means are linearly related. Linear regression is probably the most popular form of covariance adjustment

Linear regression and causal inference

- What if we knew that ignorability was satisfied and that the following was true about the world

$$E[Y(1) \mid X] = \alpha + \beta X + \tau$$

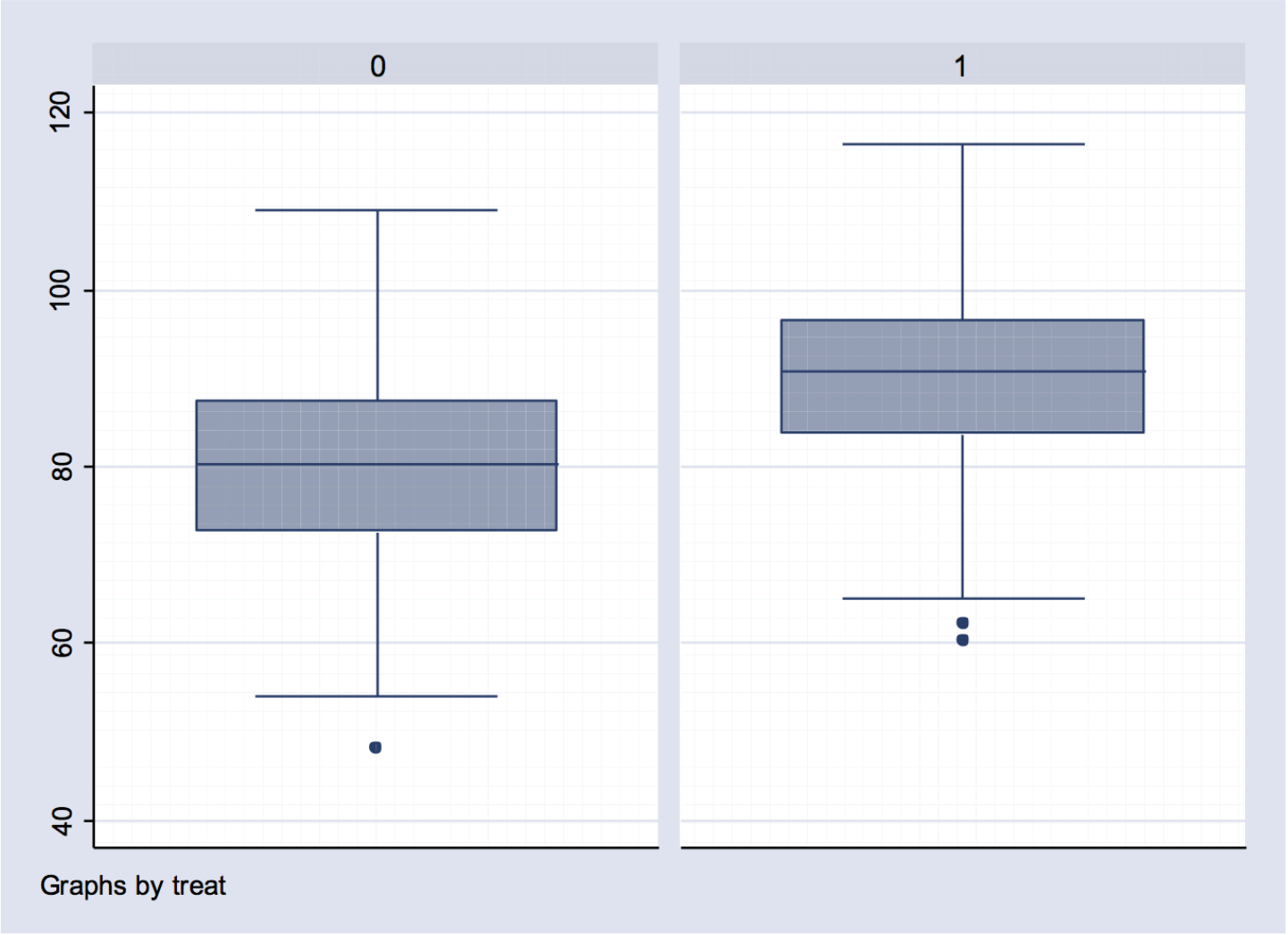
$$E[Y(0) \mid X] = \alpha + \beta X$$

$$\tau = E[Y(1) \mid X] - E[Y(0) \mid X]$$

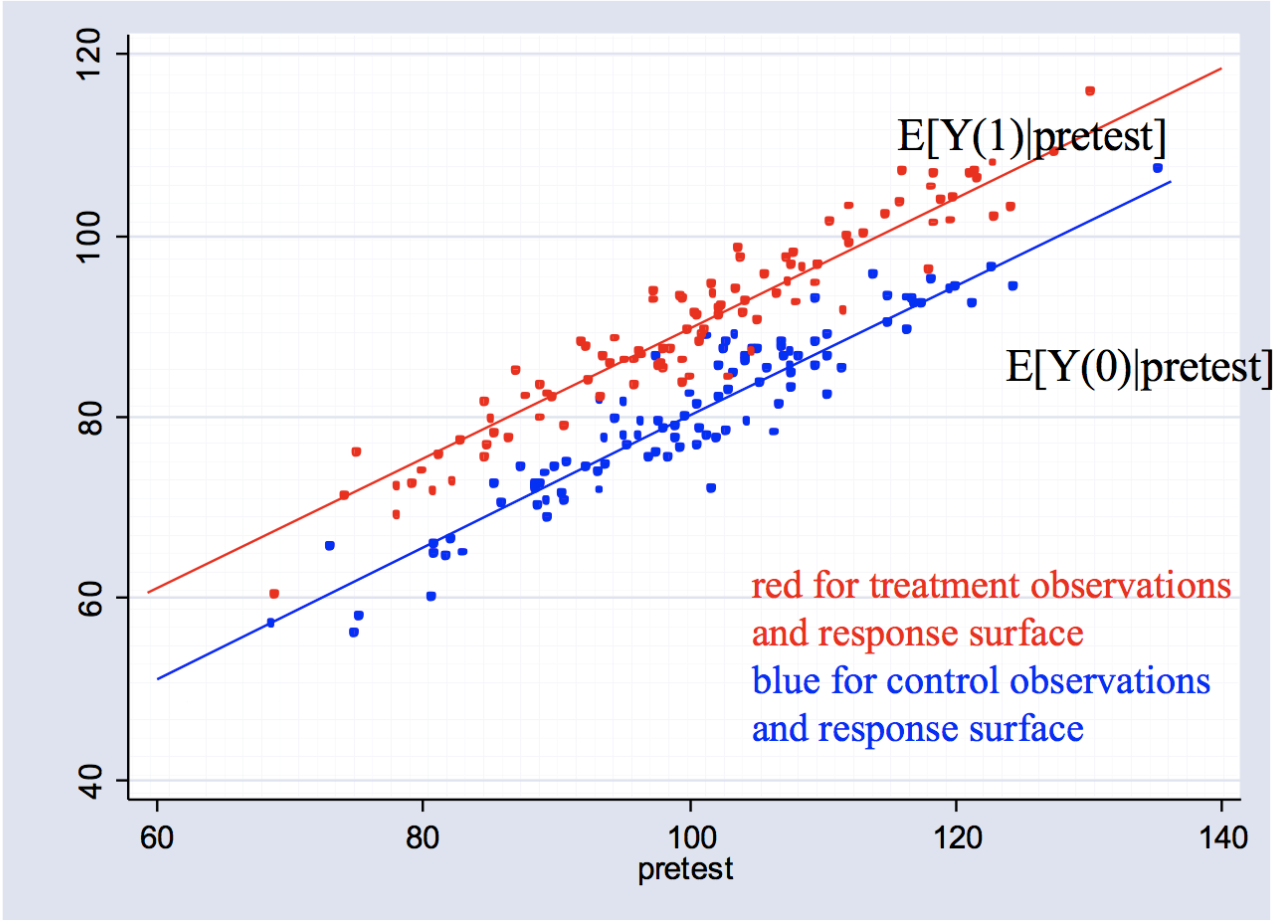
- Could we use linear regression to estimate causal effects? Yes

Average treatment effect is 10

Randomized experiment

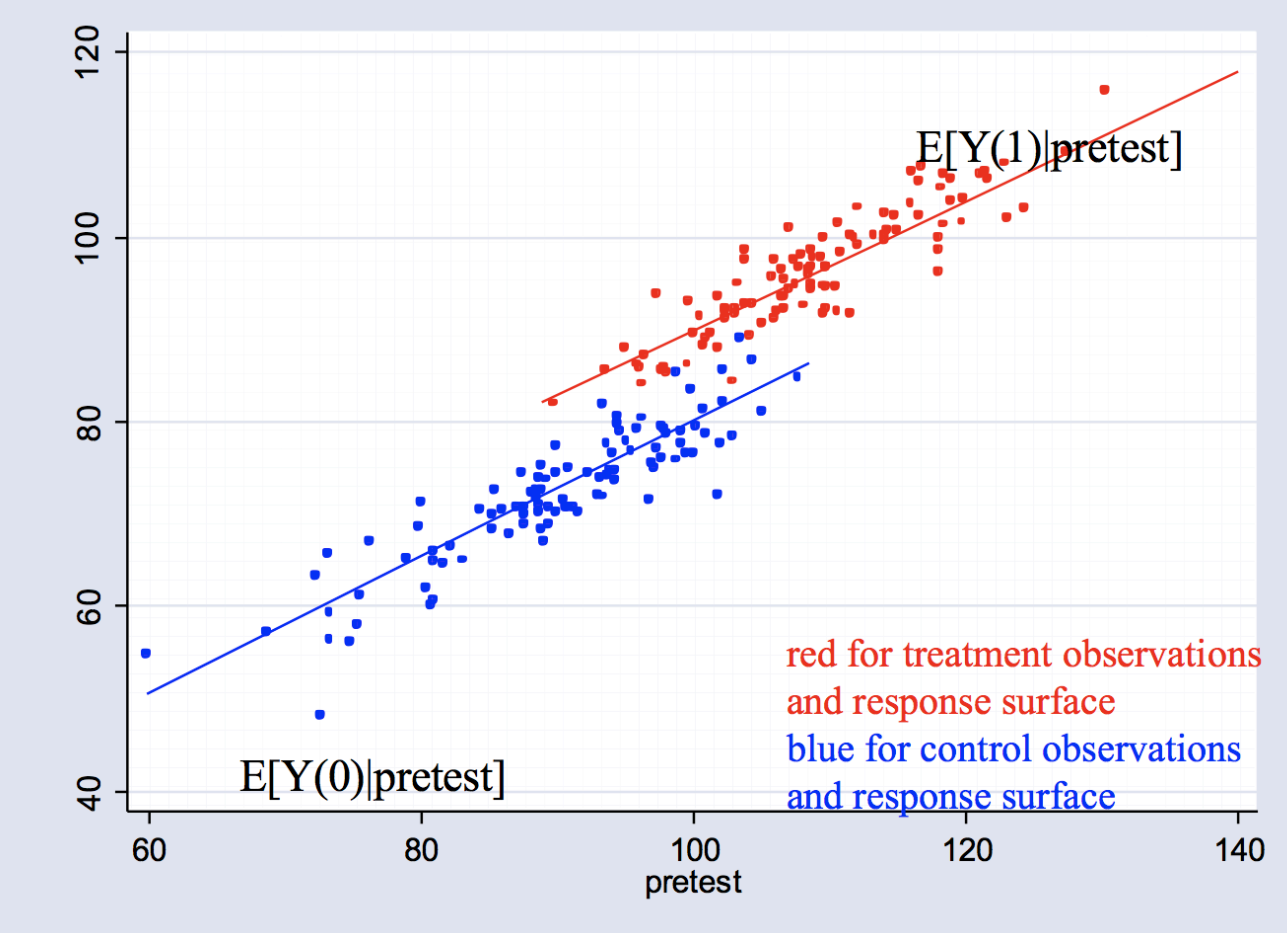


Randomized experiment: regression modeling for more precision



Average treatment effect still 10

Observational studies: one confounding covariate, linearity holds



Average treatment effect still 10

Linear regression pitfalls

- Linear regression has certain pitfalls for causal inference, however, such as reliance on linear specifications and extrapolations over portions of the covariate space where there are no data
- Importantly, there are no "red flags" produced by standard regression software to alert you to when you're making a bad decision

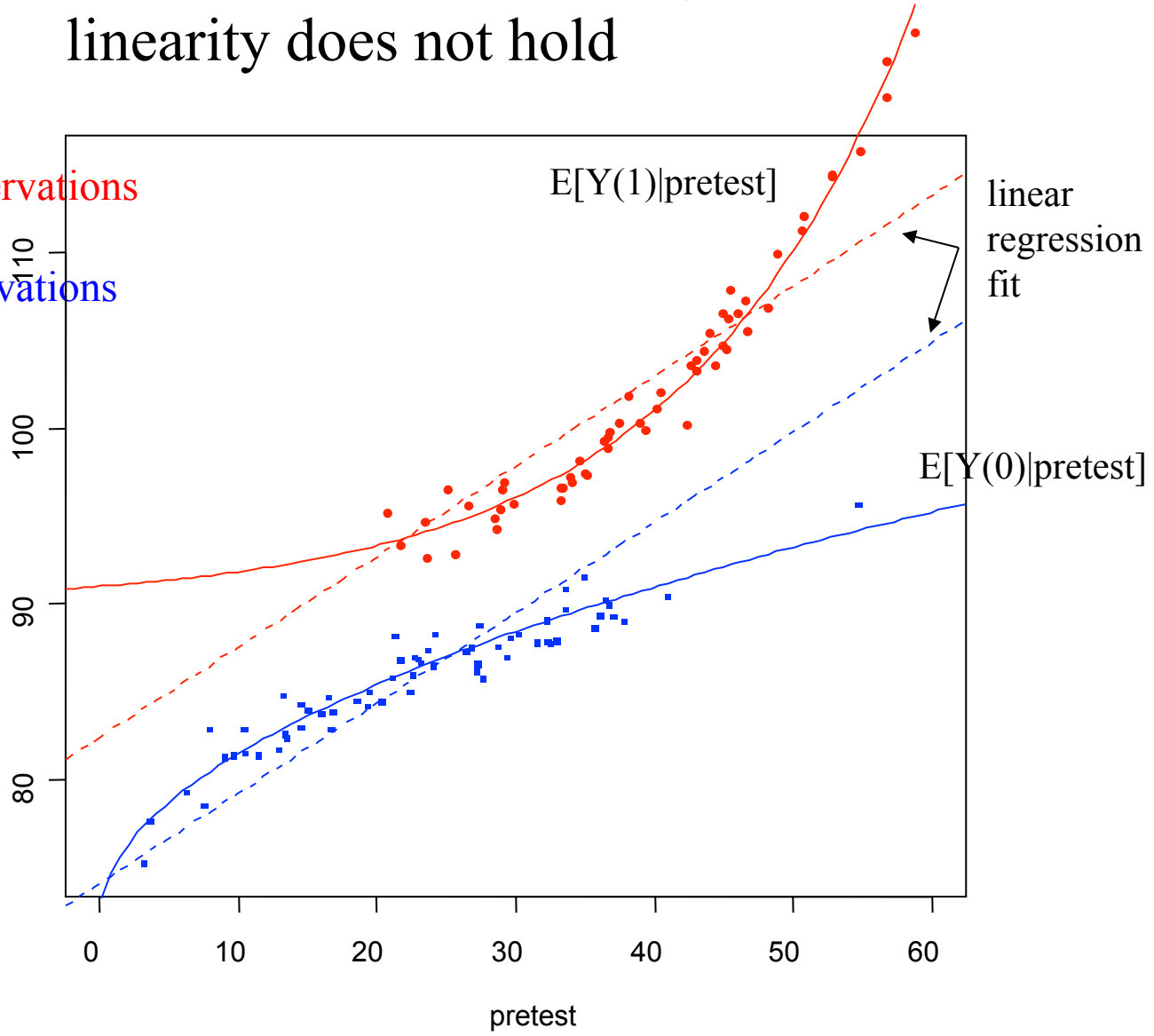
Observational studies: one confounding covariate, linearity does not hold

red for treatment observations and response surface

blue for control observations and response surface

Linearity is not an appropriate model here

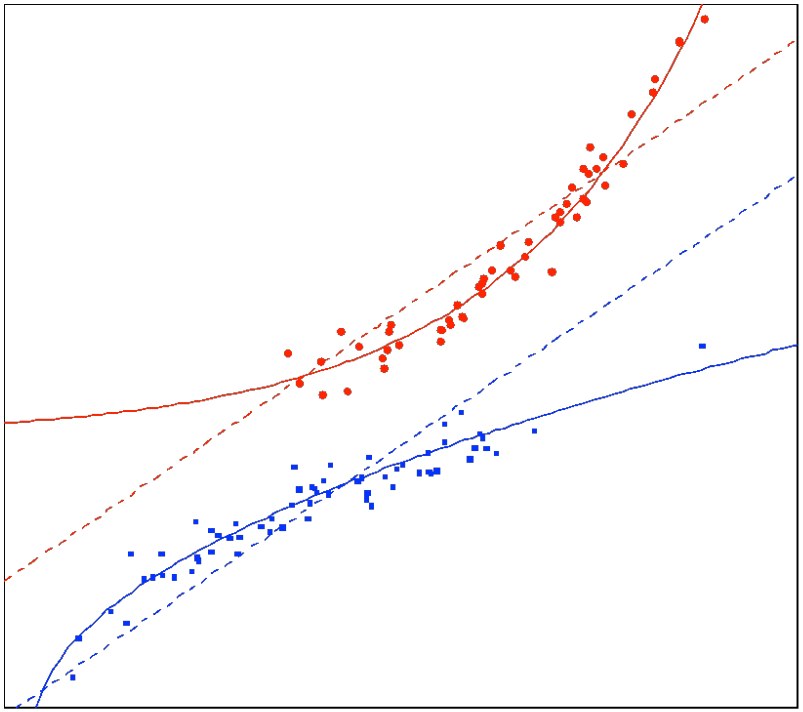
Lack of overlap in pretest exacerbates the problem by forcing model extrapolation

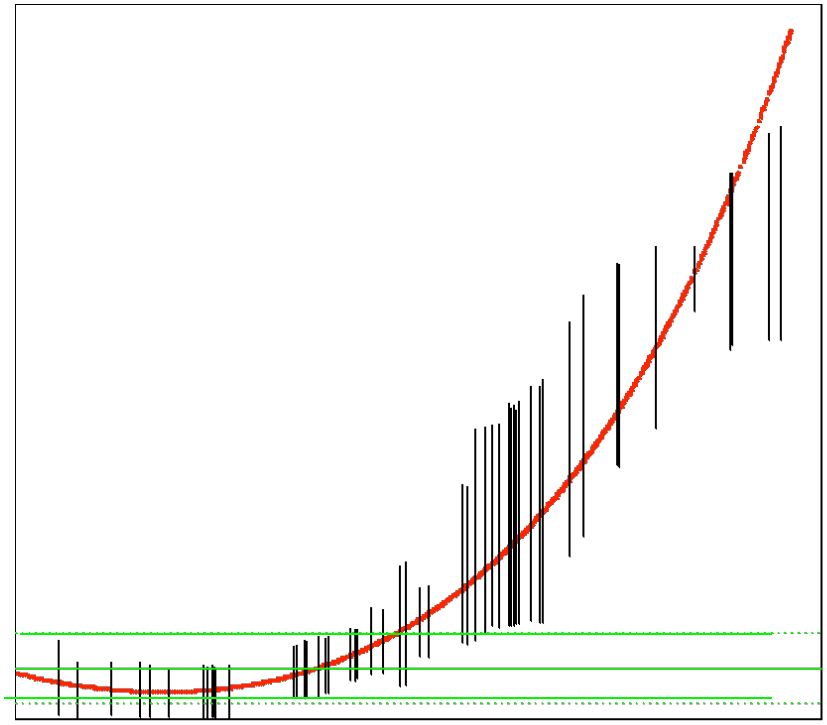


How deal with non-linearity, imbalance, extrapolation?

1) It can be helpful to restrict to areas of overlap and match, subclassify or weight in order to achieve balance and avoid undue model extrapolation – in other words, focus is on the assignment mechanism, p(Z|X), just like a randomized experiment does
2) Growing in popularity are methods that model *both* the assignment mechanism and response surface (we'll talk about propensity score matching with additional covariance adjustment). Some versions of these strategies have the formal classification of "doubly robust" methods

Linear regression missing the mark

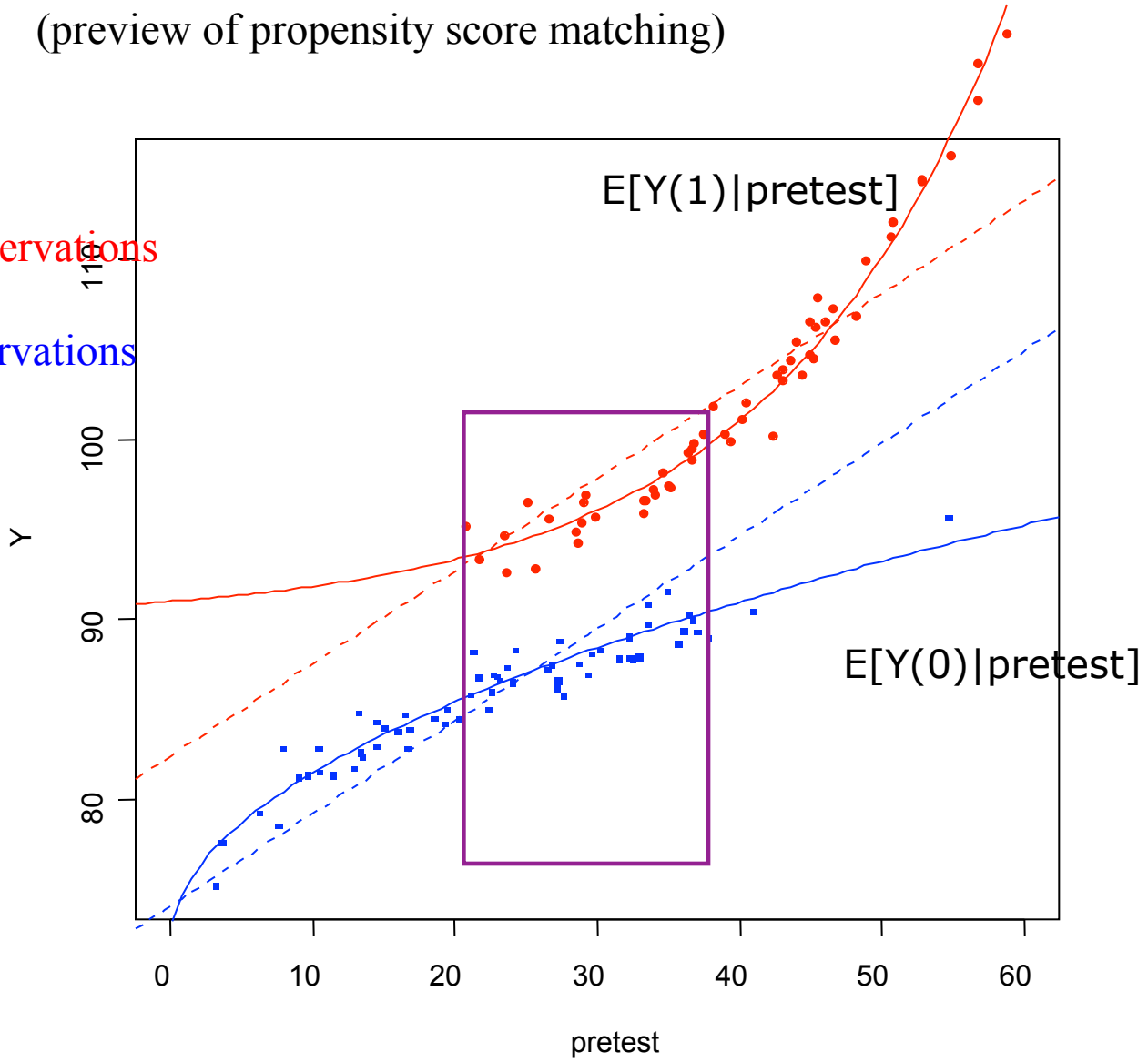




This plot adds regression estimate of treatment effect and uncertainty bounds in green

(preview of propensity score matching)

red for treatment observations
and response surface
blue for control observations
and response surface



Propensity score matching would constrain inference to areas of common support (overlap) – purple box

Linear regression: the estimand

- Suppose that ignorability holds and the following model is true for our data such that every has the same treatment effect,

$$E[Y(1) \mid X] = \alpha + \beta X + \tau$$
$$E[Y(0) \mid X] = \alpha + \beta X$$
$$\tau = E[Y(1) \mid X] - E[Y(0) \mid X]$$

- Then regression estimates this constant treatment effect

Linear regression: estimand

- What happens if effects aren't actually constant across individuals?

$$E[Y_i(1) \mid X_i = x] = \alpha + \beta X_i + \tau_x$$

$$E[Y_i(0) \mid X_i = x] = \alpha + \beta X_i$$

$$\tau_x = E[Y_i(1) \mid X_i = x] - E[Y_i(0) \mid X_i = x]$$

- Then regression estimates a variance-weighted estimate of the individual-level treatment effects, where the weights take the form

$$\frac{[\Pr(Z_i=1 \mid X_i=x)(1-\Pr(Z_i=1 \mid X_i=x))]\Pr(X_i=x)}{\Sigma[\Pr(Z_i=1 \mid X_i=x)(1-\Pr(Z_i=1 \mid X_i=x))]\Pr(X_i=x)}$$

Linear regression (pitfall rebuttal?)

- Linear regression has the property that even if it is not the correct model, it will still represent the best linear approximation to the correct model
- This is why linear regression actually performs better than one might think at times
- This doesn't help so much once we start extrapolating beyond the areas of common support

Linear regression parametric assumptions

- Final caveats
  - Whenever you fit a model you have to be aware of the assumptions involved
  - (Arguably) *the most important assumption for linear regression is that of the linear functional form* – this is the assumption we have highlighted thus far.
  - Other model violations (e.g. heteroskedasticity and lack of independence) can also cause problems depending on severity.
  - Normality of the predictors is not an assumption. Neither is normality of the outcome variable. The normality assumption has to do with the *residuals*. This is hard to check without a large sample size. With a large sample size the problems caused by deviations from this assumption will often go away (Central Limit Theorem).

Using linear regression for causal inference: assumptions

Structural assumptions

- Ignorability
- (overlap)
- SUTVA

Parametric assumptions

- The parametric form of the model is correct
- (other regression assumptions hold)

What happens if ignorability is not satisfied?

Sometimes this is referred to as “omitted variable bias”

Algebra fun!
Linear regression and "Omitted Variable Bias"

The argument typically presented is something like…

- Suppose we know that the "correct" specification for Y is
  $Y_i = \alpha + \beta X_i + \tau Z_i + \varepsilon_I$
- What happens if we ignore X when fitting the model? We fit
  $Y_i = \alpha' + \tau' Z_i + \varepsilon_I'$
- It's helpful to define a third regression
  $X_i = \gamma + \pi Z_i + \omega_I$
- If we substitute this representation of x into the original equation and rearrange terms we get
  $Y_i = \alpha + \beta(\gamma + \pi Z_i + \omega_I) + \tau Z_i + \varepsilon_I$
  $= \alpha + \beta\gamma + (\beta\pi + \tau) Z_i + \varepsilon_I + \beta\omega_I$
- Equating the coefficients on Z across the two models we get
  $\tau' = \beta\pi + \tau$
- What does that say about when bias occurs?

Summary of assumptions

Structural

- Ignorability
- (Overlap)
- SUTVA Stable Unit Treatment Value Assumption

Parametric (depends on the method)

- Stratification: non-parametric but doesn't accommodate many covariates
- Linear Regression: makes fairly strong assumptions about the relationship between the outcome and the covariates

Other considerations with observational studies

Observational Studies: other considerations

Other aspects of observational studies differentiate them from randomized experiments.

- What does it mean to label something as a treatment variable?
- What is the counterfactual?
- What is the temporal ordering of the variables?
- Can several variables be considered to be treatment variables?

Treatment variables

There are studies that examine the effect of height on outcomes such as earnings? What does it mean to think of height as a "treatment"?

How about breastfeeding?

How about parents' marital status in a study examining children's outcomes?

Think: What **intervention** does this treatment map to?

Implicit counterfactual states

Besides defining the treatment it is important to carefully define the control condition.

- In a study examining the effect of a new welfare program, what is the counterfactual condition?
- How about examining the effect of Head Start?

Temporal ordering of the variables

- Causality (in this framework at least) implies that the treatment, or "causing variable" occurs prior to the outcome.
- We also know that any potential confounders that we adjust for must reflect pre-treatment phenomena
- This type of temporal ordering will be hard to achieve with cross-sectional data
- Panel data (repeated cross-sectional) will give us a better chance of ensuring that this holds

Multiple treatments

- It is not uncommon in social science to see articles that interpret many coefficients in a regression model causally. What's wrong with this?
- Unless none of these variables have effects on the others then some variables are in fact *outcomes* of the others! It is not appropriate to control for post-treatment outcomes (it will give the wrong answer) -- we will discuss this more at the end of the semester, and touch on it a little when we talk about instrumental variables methods

Thought experiment

- If you are having trouble thinking about how to design or interpret an observational study, then perform the following thought experiment. Figure out the randomized experiment that the observational study maps to and/or the randomized study you would like to have done to answer the causal question you are interested in.

Consider the following causal questions that observational studies have attempted to answer. Can you imagine a randomized experiment that could be used to answer each of these questions?
If not , how could you use observational study to answer causal questions? What confounders would you pick up ?

1. The effect of neighborhood on residents' health
2. The effect of marital status on children's behavioral outcomes
3. The effect of HIV status on intimate partner violence
4. The effect of race on earnings
5. The effect of number of children on parents' labor supply
6. The effect of poverty on mental health
7. The effect of virginity pledges on future sexual activity
8. The effect of the gender of the candidate on election results

So, unless we can perform a randomized control trial, we can't reconstruct the counterfactual and identify causal relationships?

Not necessarily!

We'll spend the rest of the semester looking at alternatives.

To motivate the next method that we'll talk about. Let's look at a study that addressed the same question using both a randomized experiment and an observational study

Constructed observational studies

- How can we ever judge whether non-experimental methods can work in "real-life" empirical settings? When do we ever know if our estimate is a good approximation of the "truth"?
- *Constructed observational studies* combine treatment groups from a randomized experiment to comparison groups from
  - General survey data (e.g. LaLonde,1983) to test the efficacy of a variety of econometric techniques. (Dehejia & Wahba redid using p-score with favorable results)
  - Control groups in different sites or time periods but within the same experiment (Friedlander & Robins, Dynarski et al., Bloom et al.)
  - Survey data from program non-participants (Heckman et al.)

Constructed observational study: first calculate experimental benchmarks

Receives control

Receives treatment

Random assignment into treatment groups

$\overline{Y}_0$

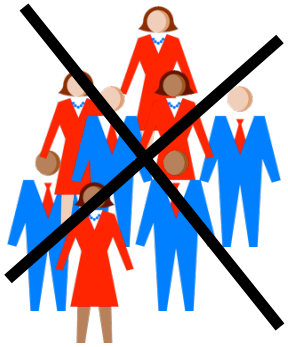
$\overline{Y}_1$

$$\hat{\tau} = \overline{Y}_1 - \overline{Y}_0$$

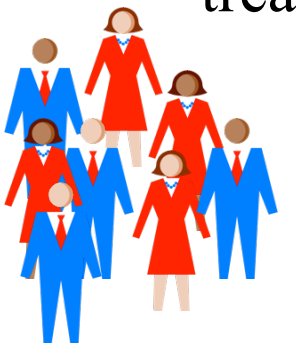
Constructed observational study:
then create new comparison group

Random assignment into
treatment groups

Receives
control



Receives
treatment



Use comparison between experimental treatment group and observational comparison group instead to estimate T.

Our constructed observational study

- Uses the (randomized experiment) evaluation of the Infant Health and Development Program (IHDP)
- As our constructed comparison group we use data from the National Longitudinal Survey of Youth (NLSY) child sample (n=4511)

For further reading, see Chapter 5 by Hill, J., Reiter, J., Zanutto, E. "A comparison of experimental and observational data analyses."

Research question

- What was the effect of the IHDP intervention on child test scores at age 3?
- Recall that in the intervention the treatment effect on PPVT (Peabody Picture Vocabulary Test) scores was 6.4 points (s.e=1.2) -- this will act as our benchmark

Analyses performed: methods

- Regression: linear regression of outcome on treatment indicator on covariates
- P-score Direct: difference in means across propensity-score-matched groups
- P-score Regression: linear regression of outcome on treatment indicator on covariates across propensity-score-matched groups

Pre-treatment Data (variables that could be collected across both datasets)

- Mother variables: age, ethnicity, marital status, education level, worked pre-birth, prenatal care
- Child variables: birth weight, days in hospital, age in 1990, weeks preterm, first born, sex
- Geographic variables: unemployment rate, state

Balance in constructed observational study

- In the constructed observational study there were big differences in terms of:
  - Mother's: race, marital status, education
  - Child's: birth weight, days in hospital, weeks preterm
  - Family's state of residence

Analyses performed: confounding covariates included

- DE: just individual-level demographics
- DE+U: individual-level demographics plus county-level unemployment rates
- DE+U+S: individual-level demographics plus county-level unemployment rates plus state indicators included in models
- DE+U+X: individual-level demographics plus county-level unemployment rates plus propensity score analyses exact matched within states

  [In last two analyses control observations from other states are excluded because otherwise there is an aliasing problem between the "left out" state for IHDP sample and all states not represented in IHDP sample but in NLSY sample.]

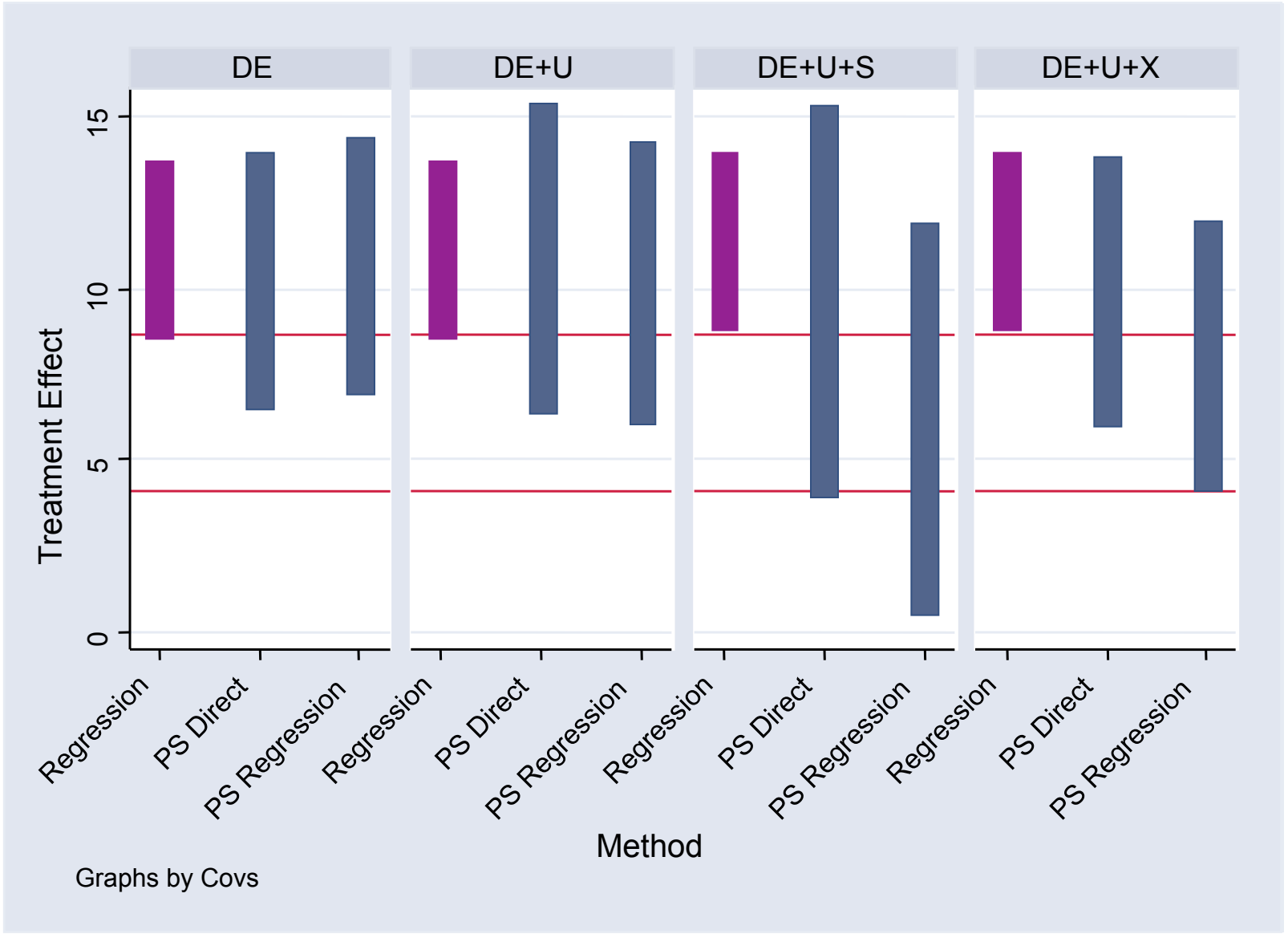
Results of analyses

Method	DE	DE+U	DE+U+S	DE+U+X
Regression	11.16 (1.3)	11.16 (1.3)	11.39 (1.3)	11.39 (1.3)
P-score Direct	10.26 (1.9)	10.88 (2.3)	9.64 (2.9)	9.94 (2.0)
P-score Regression	10.67 (1.9)	10.16 (2.1)	6.20 (2.9)	8.04 (2.0)

Method controls for:				
Demographics	X	X	X	X
Unemployment		X	X	X
States			X	X
Exact state match				X

Table 1: Point estimates and standard errors (in parentheses) of treatment effects. The treatment effect for the experiment equals 6.39 with a standard error of 1.17.

Results


DE
DE+U
DE+U+S
DE+U+X
Treatment Effect
0
5
10
15
Regression
PS Direct
PS Regression
Method
Graphs by Covs

Propensity score matching

This example provides an “existence theorem” for propensity score matching.

In other words it provides an example where a propensity score approach can recover a reasonable estimate.

Next week we will start learning about propensity score approaches to causal inference!

Conclusions

- Randomized experiments are the gold standard for causal inference (though not without flaws)
- Without randomized experiments we need to another way to control for differences between those who receive the treatment and those who do not
- Simple methods such as regression may not be sufficient!
- More complicated methods (e.g. propensity score matching) may also not be sufficient, but may allow us to make slightly less heroic assumptions
- Crucial to be clear about the assumptions were are making with each analysis strategy and to try to minimize the strength of these assumptions