

# STA 235H - Observational Studies

Fall 2022

McCombs School of Business, UT Austin

# What happens when the ignorability assumption doesn't hold?

- Imagine **we can't randomize**

**Observational Study**

Can we compare two groups to get a causal effect?

# What happens if the ignorability assumption doesn't hold?

- Now, let's assume  $(Y(0), Y(1)) \not\perp\!\!\!\perp Z$

$$\tau = E[Y_i(1)] - E[Y_i(0)] \neq E[Y_i|Z = 1] - E[Y_i|Z = 0]$$

**Correlation does not imply causation**

# What happens if the ignorability assumption doesn't hold?

- Now, let's assume  $(Y(0), Y(1)) \not\perp\!\!\!\perp Z$

$$\begin{aligned}\tau &= E[Y_i(1) - Y_i(0)] = \\ &= E[Y_i(1) - Y_i(0)|Z = 1]Pr(Z = 1) + E[Y_i(1) - Y_i(0)|Z = 0](1 - Pr(Z = 1))\end{aligned}$$

# What happens if the ignorability assumption doesn't hold?

- Now, let's assume  $(Y(0), Y(1)) \not\perp\!\!\!\perp Z$

$$\begin{aligned}\tau &= E[Y_i(1) - Y_i(0)] = \\ &= \underbrace{E[Y_i(1) - Y_i(0) | Z = 1]}_{\text{ATT}} Pr(Z = 1) + \overbrace{E[Y_i(1) - Y_i(0) | Z = 0]}^{\text{ATC}} (1 - Pr(Z = 1))\end{aligned}$$

- Weighted average of the ATT and ATC.

# What happens if the ignorability assumption doesn't hold?

- After some simple math, you can get to:

$$\tau = E[Y_i(1) - Y_i(0)] = ATE$$

$$ATE = E[Y_i|Z = 1] - E[Y_i|Z = 0]$$

$$= (E[Y_i(0)|Z = 1] - E[Y_i(0)|Z = 0])$$

$$= (1 - Pr(Z = 1))(ATT - ATC)$$

Check out Scott Cunningham's "Causal Inference: The Mixtape" (Ch. 4.1.3) for the decomposition

# What happens if the ignorability assumption doesn't hold?

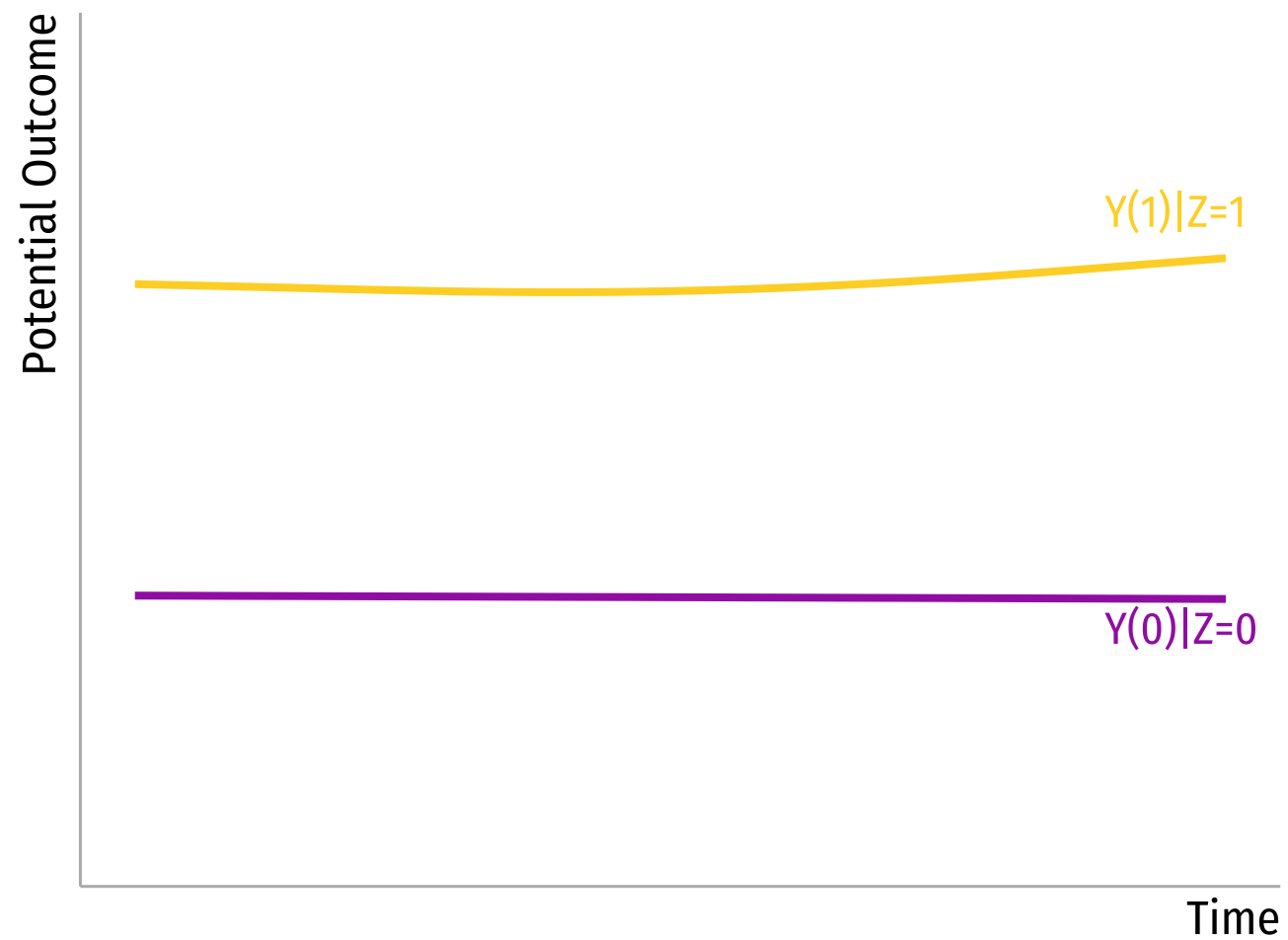
- After some simple math, you can get to:

$$\begin{aligned}\tau &= E[Y_i(1) - Y_i(0)] = ATE \\ ATE &= \underbrace{E[Y_i|Z = 1] - E[Y_i|Z = 0]}_{\text{Obs diff in means}} \\ &\quad - \underbrace{(E[Y_i(0)|Z = 1] - E[Y_i(0)|Z = 0])}_{\text{Selection bias}} \\ &\quad - \underbrace{(1 - Pr(Z = 1))(ATT - ATC)}_{\text{Heterogeneous treat. effect bias}}\end{aligned}$$

- **Selection Bias:** Difference between groups if they both were under control (e.g. baseline differences).
- **Heterogeneous Treatment Effect Bias:** Difference in returns to treatment for the two groups (weighted by the control population).

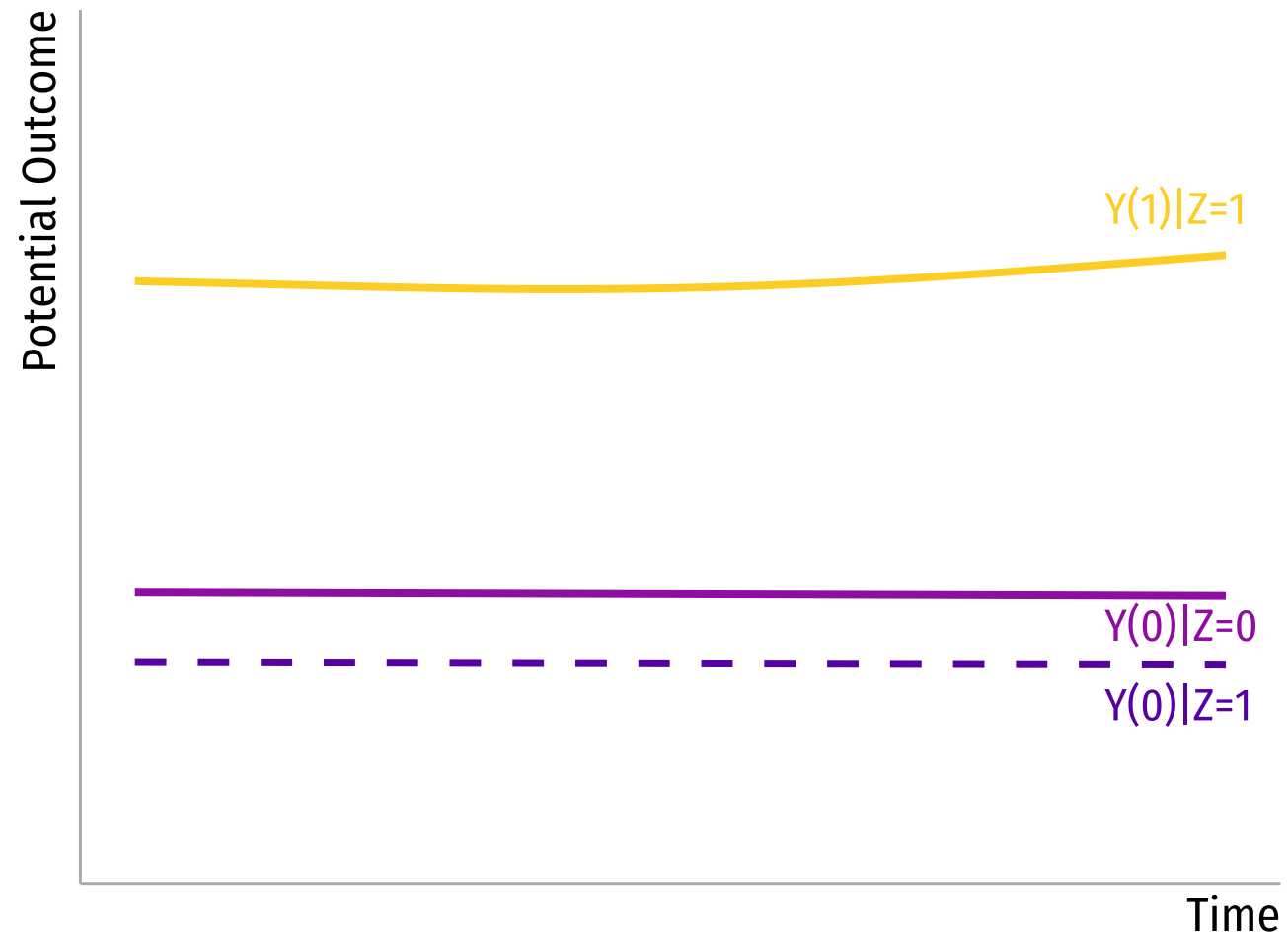
Check out Scott Cunningham's "Causal Inference: The Mixtape" (Ch. 4.1.3) for the decomposition

# How would bias look like?

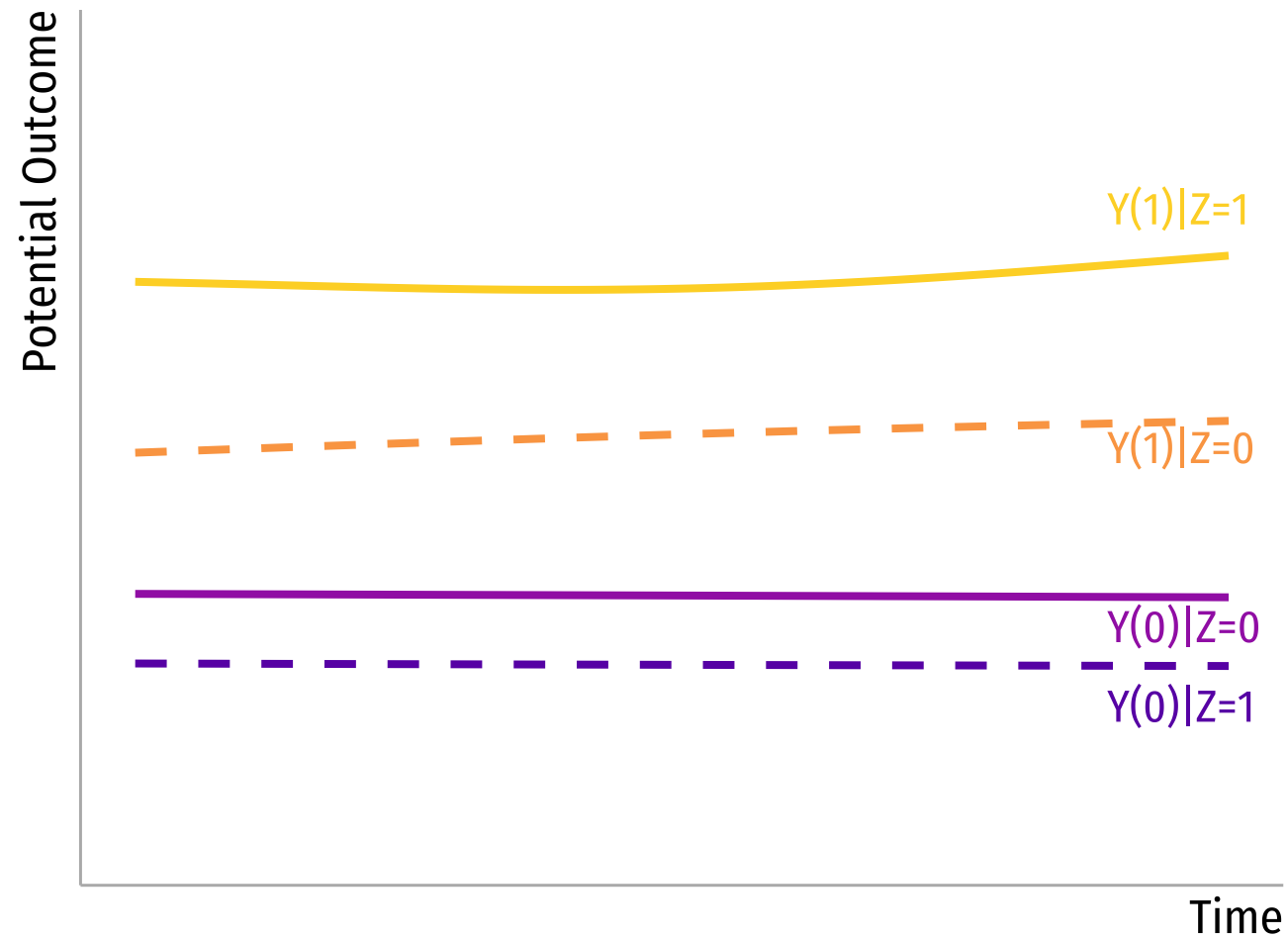




# How would bias look like?



# How would bias look like?



# Example: Effect of types of advertising on sales

You want to know whether is more convenient to **e-mail** or **physically mail** potential customers to increase your subscribers.



**freshdirect**  
The freshest groceries. Delivered.

Hey Bellport, get everything you need to grill, thrill, and chill, delivered all summer long!

**\$25 OFF\***  
YOUR NEXT ORDER OF \$99+  
USE CODE: BEACH10

Fresh-picked produce | Farm-fresh dairy | Custom-cut meats  
Sustainable seafood | Your favorite grocery brands | Hundreds of weekly deals  
Plus FreshDirect Wines & Spirits

\*This offer for \$25 OFF is good on your next residential order delivered when promo code BEACH10 is entered at checkout. May not be combined with any other promotion code. Valid only for your order totaling \$25 or more before taxes, delivery fee, and delivery premium. Limit: one per customer/household. All standard customer terms and conditions apply. FreshDirect reserves the right to cancel or modify the offer at any time. Offer expires at 11:59pm ET, September 1, 2019 and will be removed from orders that are modified after this time. Offer is nontransferable. Void where prohibited. Offer is limited time only. ©2019 Fresh Direct, LLC.

# Example: Going to Office Hours

An important question could be: Does going to office hours increase our GPA?

- What could be an example of **selection bias**?
  - Remember that selection bias means differences in  $Y(0)$  for people that go to office hours vs those that don't.
- What could be an example of **heterogeneous return to treatment bias**?
  - Remember that heterogeneous return to treatment bias means that  $Y(1) - Y(0)$  is different for those that go vs those that don't go to office hours.

# How do confounders affect our causal estimate?

confounding variable



# What happens if we control by our confounders?

adjusting for confounders



# Next class

- Continue talking about **Observational Studies**
- Can we use **matching** for causal effects?
- Introduce **difference-in-differences**

