Roadmap

- What controls to include?
- Proof of Unconfoundedness given propensity score

Good and bad controls

Which of these variables would you control for? Think identification and efficiency

$$\begin{tikzcd} & X_1 \arrow[dl] \arrow[d] \\ D \arrow[r] & Y \end{tikzcd} \qquad \begin{tikzcd} X_2 \arrow[d] & \\ D \arrow[r] & Y \end{tikzcd} \qquad \begin{tikzcd} & X_3 \arrow[d] \\ D \arrow[r] & Y \end{tikzcd} \qquad \begin{tikzcd} U_1 \arrow[dr, dashed] \arrow[dd, dashed] & & U_2 \arrow[dl, dashed] \arrow[dd, dashed] \\ & X_4 \arrow[dl] \arrow[dr] & \\ D \arrow[rr] & & Y \end{tikzcd}$$

Good and bad controls

Which of these variables would you control for? Think identification and efficiency

$$\begin{tikzcd} & X_1 \arrow[dl] \arrow[d] \\ D \arrow[r] & Y \end{tikzcd}$$

X_1 Necessary for identification

$$\begin{tikzcd} X_2 \arrow[d] & \\ D \arrow[r] & Y \end{tikzcd}$$

X_2 Bad for efficiency (but may be good for robustness)

$$\begin{tikzcd} & X_3 \arrow[d] \\ D \arrow[r] & Y \end{tikzcd}$$

X_3 Good for efficiency

$$\begin{tikzcd} U_1 \arrow[dd, dashed] \arrow[dr, dashed] & & U_2 \arrow[dl, dashed] \arrow[dd, dashed] \\ & X_4 \arrow[dl] \arrow[dr] & \\ D \arrow[rr] & & Y \end{tikzcd}$$

X_4 “Collider,” generates bias (but may serve as proxy for U_2 if U_2 affects D)

Good and bad controls

Which of these variables would you control for? Think identification and efficiency

$$\begin{tikzcd} D \arrow[r] \arrow[dr] & Y \\ & X_5 \arrow[u] \end{tikzcd} \qquad \begin{tikzcd} D \arrow[r] & Y \arrow[d] \\ & X_6 \end{tikzcd} \qquad \begin{tikzcd} D \arrow[r] \arrow[d] & Y \\ X_7 & \end{tikzcd}$$

Good and bad controls

Which of these variables would you control for? Think identification and efficiency

$$\begin{tikzcd} D \arrow[rr] \arrow[dr] & & Y \\ & X_5 \arrow[ur] & \end{tikzcd} \qquad \begin{tikzcd} D \arrow[r] & Y \arrow[d] \\ & X_6 \end{tikzcd} \qquad \begin{tikzcd} D \arrow[r] \arrow[d] & Y \\ X_7 & \end{tikzcd}$$

X_5 Bad control, generate bias

X_6 Bad control, generate bias

X_7 Exercise for you

Roadmap

- What controls to include?
- Proof of Unconfoundedness given propensity score

Proof

This proof is from Imbens (2005, RESTAT) (may skip in class).

Show $\mathbb{P}(D_i = 1 \mid Y_i(0), Y_i(1), \pi(\boldsymbol{X}_i)) = \mathbb{P}(D_i = 1 \mid \pi(\boldsymbol{X}_i)) = \pi(\boldsymbol{X})_i$ which implies that D_i and the potential outcomes are independent conditional on the propensity score

Proof

$$\begin{aligned}
\mathbb{P}(D_i = 1 \mid Y_i(0), Y_i(1), \pi(\boldsymbol{X}_i)) &= \mathbb{E}[D_i = 1 \mid Y_i(0), Y_i(1), \pi(\boldsymbol{X}_i)] \\
&= \mathbb{E}[\,\mathbb{E}[D_i = 1 \mid Y_i(0), Y_i(1), \pi(\boldsymbol{X}_i), \boldsymbol{X}_i] \mid Y_i(0), Y_i(1), \pi(\boldsymbol{X}_i)] \\
&= \mathbb{E}[\,\mathbb{E}[D_i = 1 \mid Y_i(0), Y_i(1), \boldsymbol{X}_i] \mid Y_i(0), Y_i(1), \pi(\boldsymbol{X}_i)] \\
&= \mathbb{E}[\,\mathbb{E}[D_i = 1 \mid \boldsymbol{X}_i] \mid Y_i(0), Y_i(1), \pi(\boldsymbol{X}_i)] \\
&= \mathbb{E}[\pi(\boldsymbol{X}_i) \mid Y_i(0), Y_i(1), \pi(\boldsymbol{X}_i)] = \pi(\boldsymbol{X}_i),
\end{aligned}$$

where 1st equality comes from D_i being a 0/1 variable, 2nd from LIE, 4th from CIA (conditional on $\boldsymbol{X}$), and fifth from definition of the propensity score