Observational Study Design: Selection on Observables

Northwestern Causal Inference Workshop

Brigham R. Frandsen

BYU

July 30, 2024

$15M

$$$

$15M

\$15M

\$500K

\$15M

$D = 1$

$Y(1)$

\$500K

$D = 0$

$Y(0)$

Potential outcomes

$\$15\text{M} - \$500\text{K} = \$14.5\text{M}$

$Y(1) - Y(0) =$ Treatment effect

$$$ $

$15M - $700K = $42.3M

Y(1) - Y(0) = Treatment effect

counterfactual

$\$43\text{M} - \$700\text{K} = \$42.3\text{M}$

$Y(1) - Y(0) =$ Treatment effect

What it would take to get individual-level treatment effects

1st Assumption: SUTVA

$15M $12M

$500k $85k

1st Assumption: SUTVA

$15M

$12M

$500k

$85k

1st Assumption: SUTVA

$15M

$12M

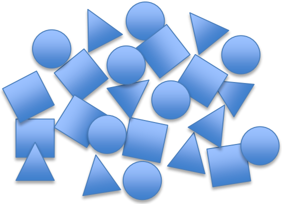
$500k

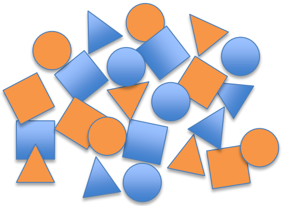
$85k

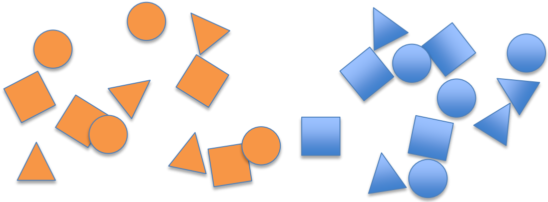
$15M - $700K = $42.3M

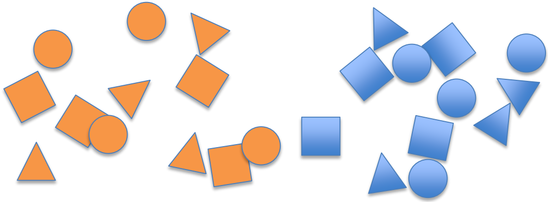
Y(1) - Y(0) = Treatment effect

counterfactual

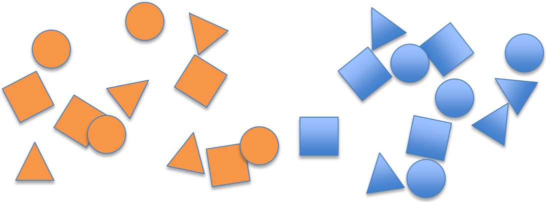






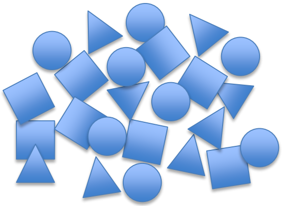


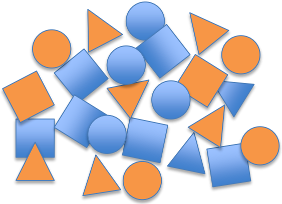
$E[Y(1)]$ $E[Y(0)]$

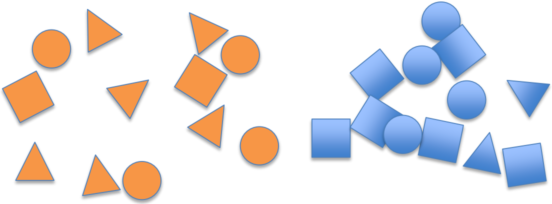


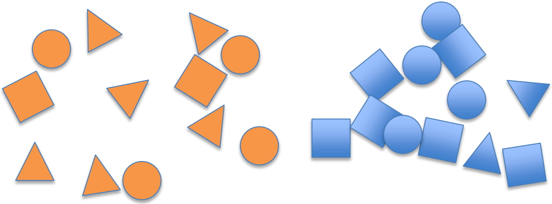
$E[Y(1)]$ - $E[Y(0)]$ = $E[Y(1) - Y(0)]$

ATE



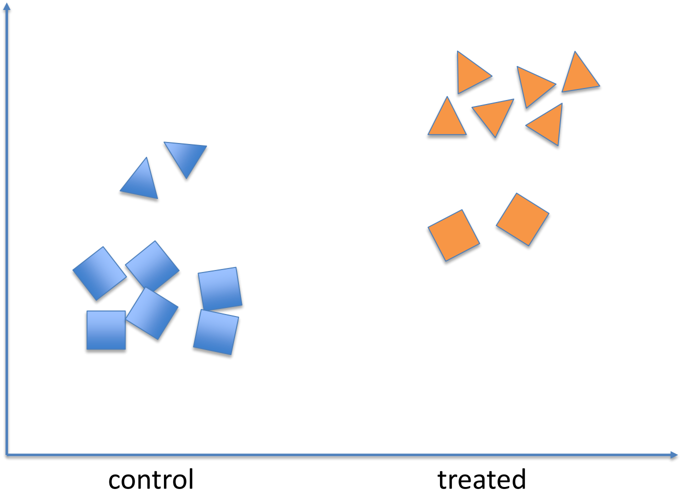


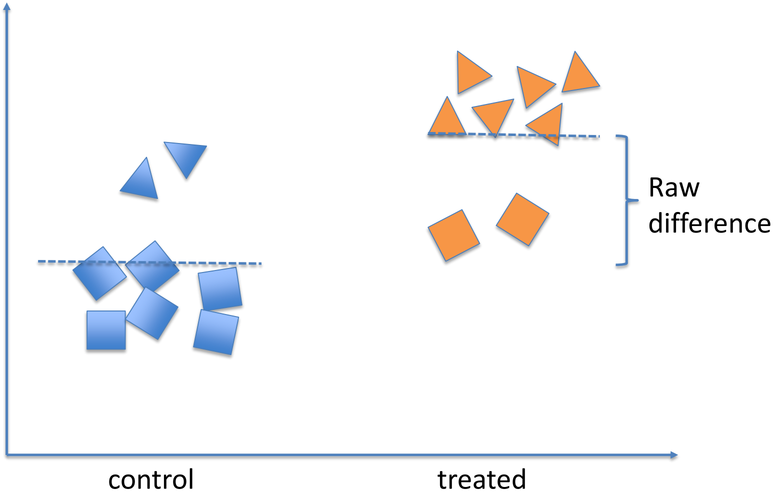


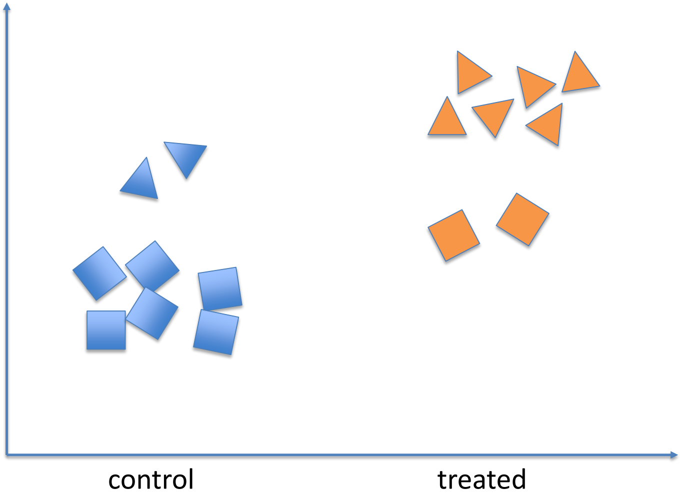


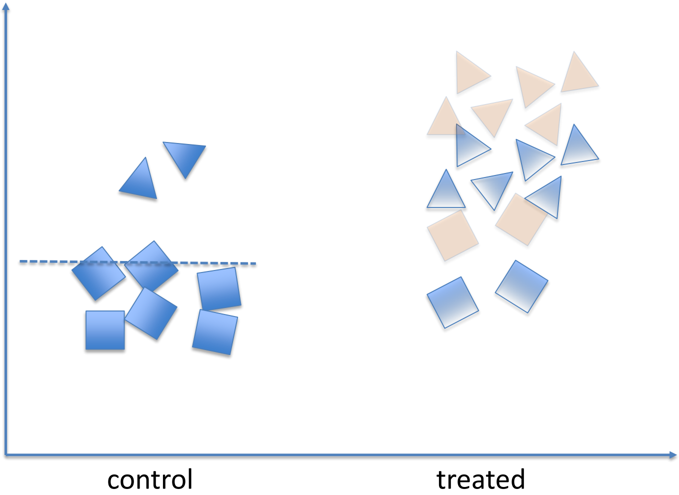
$$\underbrace{E[Y|D=1] - E[Y|D=0]}_{} = \overbrace{E[Y(1) - Y(0)|D=1]}^{\text{ATT}}$$

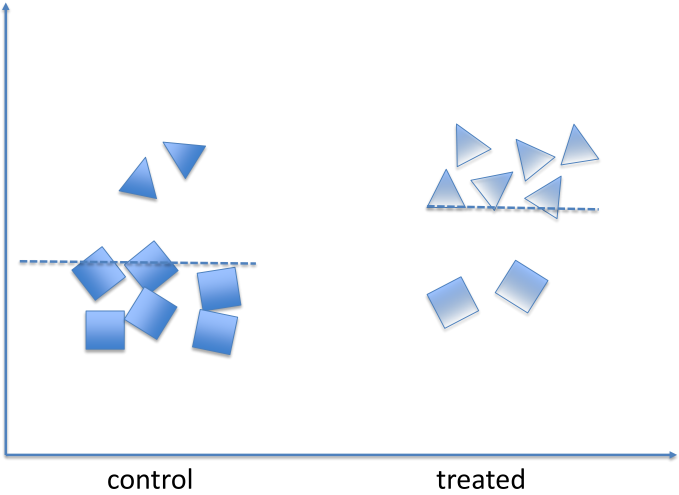
$$+ \underbrace{E[Y(0)|D=1] - E[Y(0)|D=0]}_{\text{selection bias}}$$

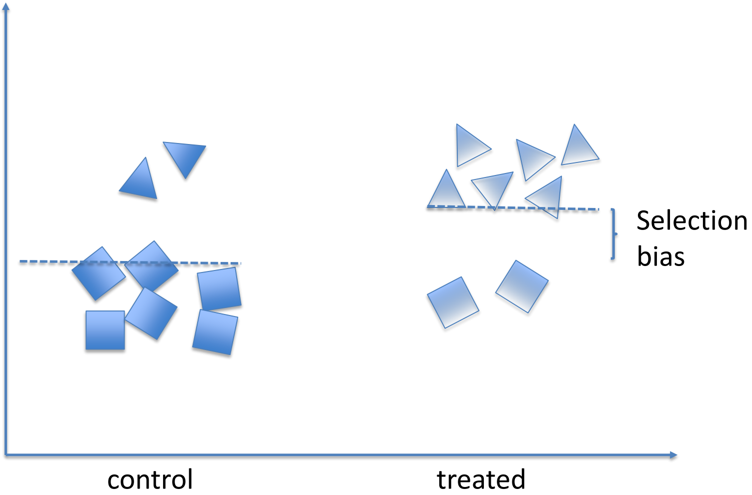

control
treated

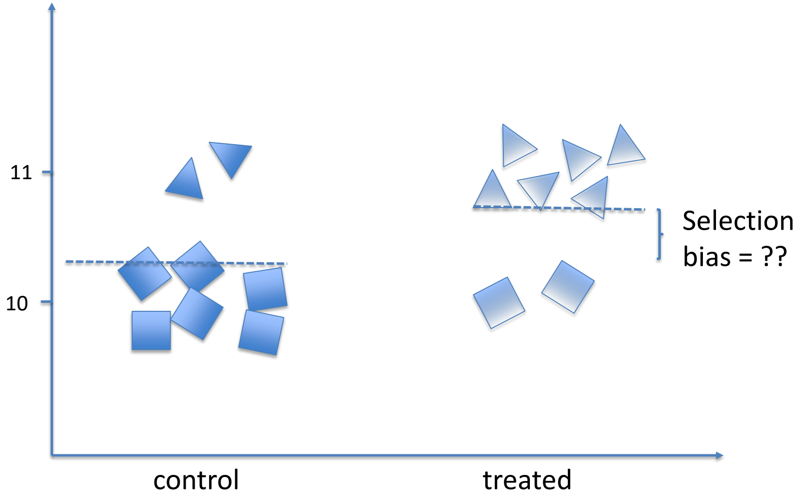

Raw
difference
control
treated

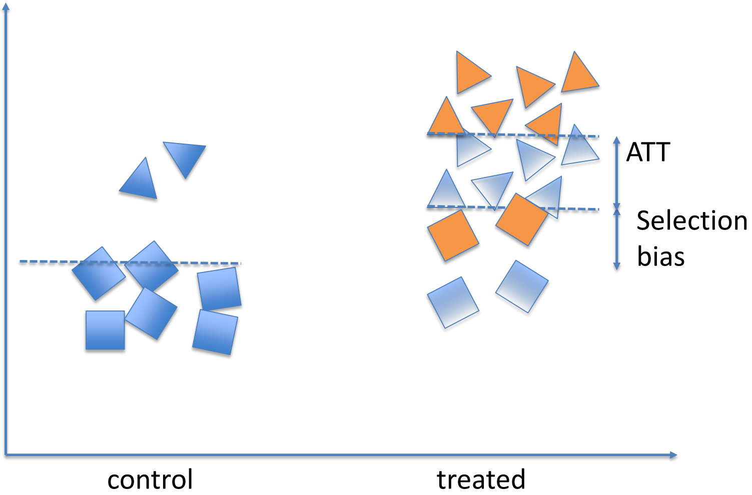

control
treated


control
treated

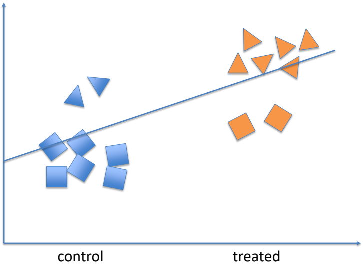

control
treated


Selection
bias
control
treated


11
10
Selection
bias = ??
control
treated


ATT
Selection bias
control
treated

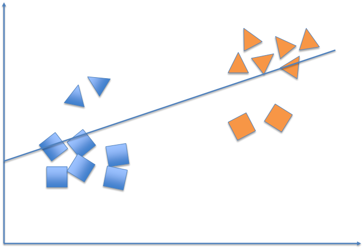
Aside: Selection bias by another name



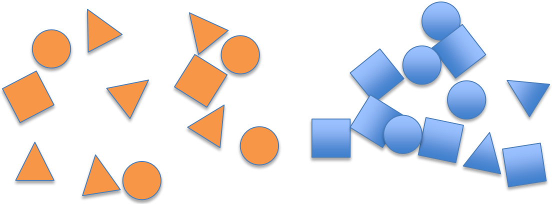
$$Y = \alpha + \beta D + \underbrace{\gamma S + \eta}_{\varepsilon}, \qquad Cov(D, \eta) = 0$$

$$\hat{b}^{OLS} \rightarrow \beta + \underbrace{\gamma \frac{Cov(S, D)}{Var(D)}}_{OVB}$$

Aside: Selection bias by another name



$$
\begin{aligned}
OVB &= \gamma \times \frac{Cov\left(S, D\right)}{Var\left(D\right)} \\
\text{Selection Bias} &= E\left[Y\left(0\right)|D=1\right] - E\left[Y\left(0\right)|D=0\right] \\
&= \underbrace{\left(E\left[Y\left(0\right)|S=1\right] - E\left[Y\left(0\right)|S=0\right]\right)}_{\gamma} \\
&\quad \times \underbrace{\left(\Pr\left(S=1|D=1\right) - \Pr\left(S=1|D=0\right)\right)}_{Cov(S,D)/Var(D)}
\end{aligned}
$$

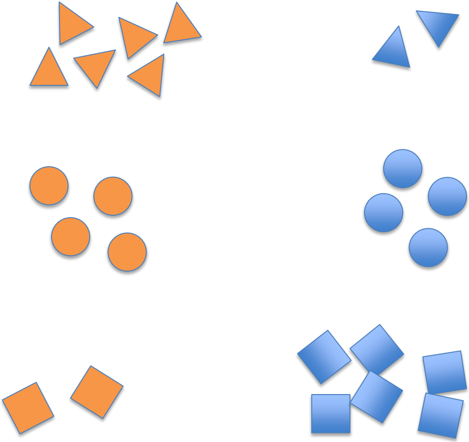


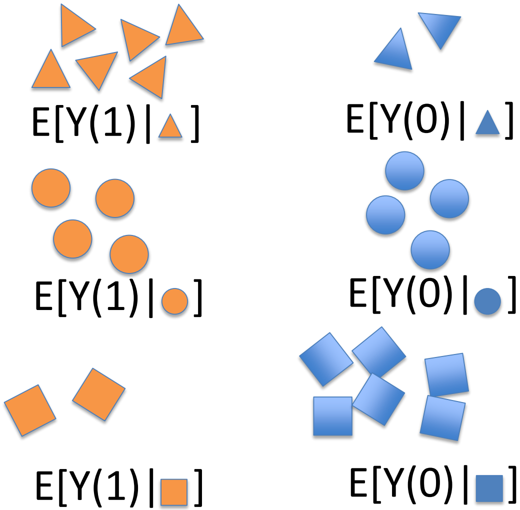
ATT

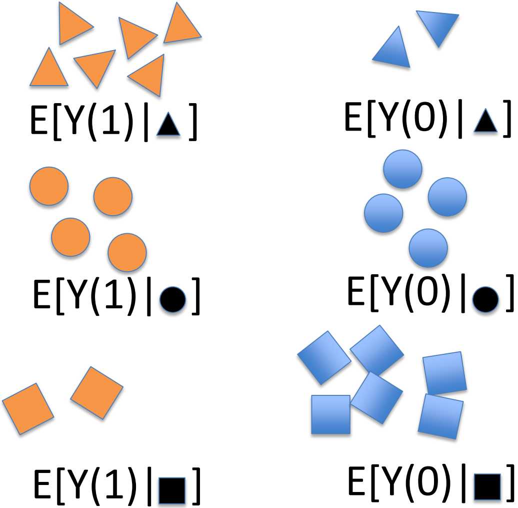
$$E[Y|D=1] - E[Y|D=0] = E[Y(1) - Y(0)|D=1]$$

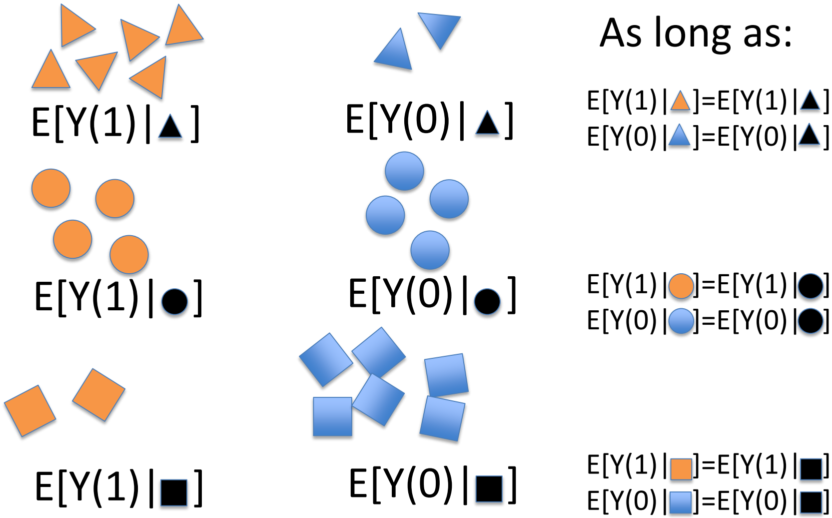
$$+ E[Y(0)|D=1] - E[Y(0)|D=0]$$

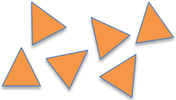
selection bias



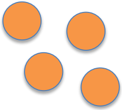

E[Y(1)|▲]
E[Y(0)|▲]
E[Y(1)|●]
E[Y(0)|●]
E[Y(1)|■]
E[Y(0)|■]

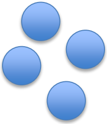

E[Y(1)|▲]
E[Y(0)|▲]
E[Y(1)|●]
E[Y(0)|●]
E[Y(1)|■]
E[Y(0)|■]


E[Y(1)|▲]
E[Y(0)|▲]
As long as:
E[Y(1)|▲]=E[Y(1)|▲]
E[Y(0)|▲]=E[Y(0)|▲]
E[Y(1)|●]
E[Y(0)|●]
E[Y(1)|●]=E[Y(1)|●]
E[Y(0)|●]=E[Y(0)|●]
E[Y(1)|■]
E[Y(0)|■]
E[Y(1)|■]=E[Y(1)|■]
E[Y(0)|■]=E[Y(0)|■]

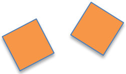


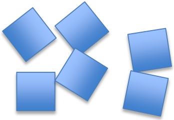
$$E[Y(1)|\blacktriangle] \quad - \quad E[Y(0)|\blacktriangle] = E[Y(1) - Y(0)|\blacktriangle]$$



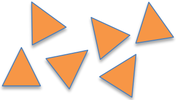


$$E[Y(1)|\bullet] \quad - \quad E[Y(0)|\bullet] = E[Y(1) - Y(0)|\bullet]$$

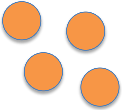
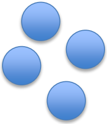




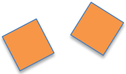
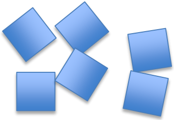
$$E[Y(1)|\blacksquare] \quad - \quad E[Y(0)|\blacksquare] = E[Y(1) - Y(0)|\blacksquare]$$

$$E[Y(1) \mid \blacktriangle] \quad - \quad E[Y(0) \mid \blacktriangle] = E[Y(1) - Y(0) \mid \blacktriangle]$$

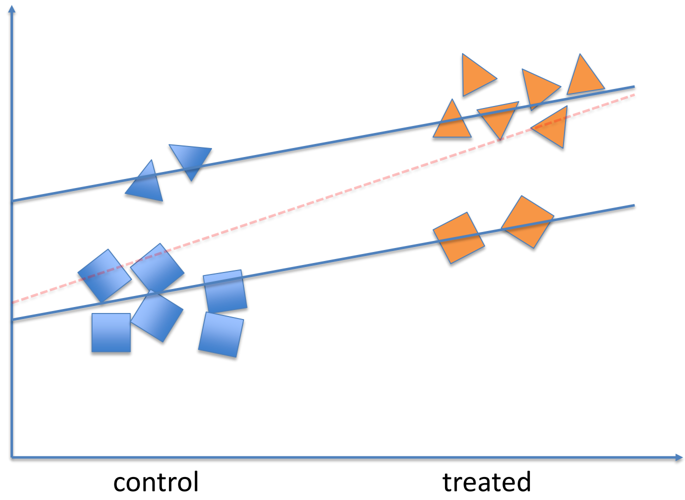
 

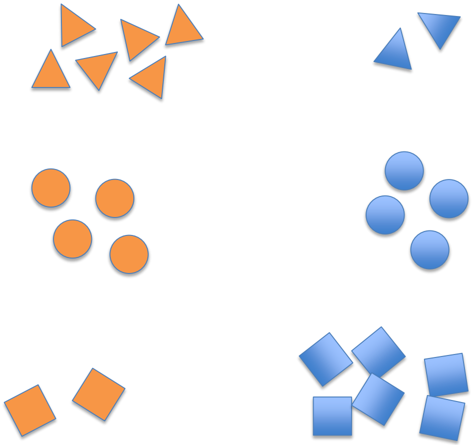
$$E[Y(1) \mid \bullet] \quad - \quad E[Y(0) \mid \bullet] = E[Y(1) - Y(0) \mid \bullet]$$

$$E[Y(1) \mid \blacksquare] \quad - \quad E[Y(0) \mid \blacksquare] = E[Y(1) - Y(0) \mid \blacksquare]$$

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

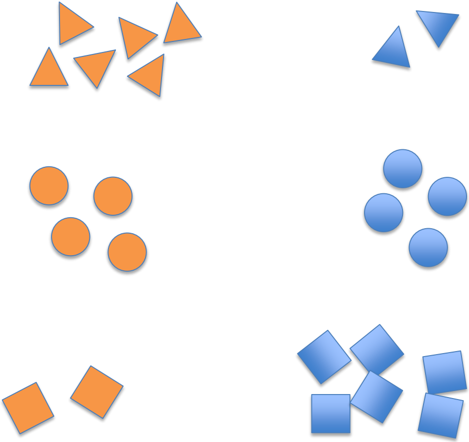

control
treated



$$Y(1),Y(0) \perp\!\!\!\perp D \mid \blacktriangle$$

$$Y(1),Y(0) \perp\!\!\!\perp D \mid \bullet$$

$$Y(1),Y(0) \perp\!\!\!\perp D \mid \blacksquare$$



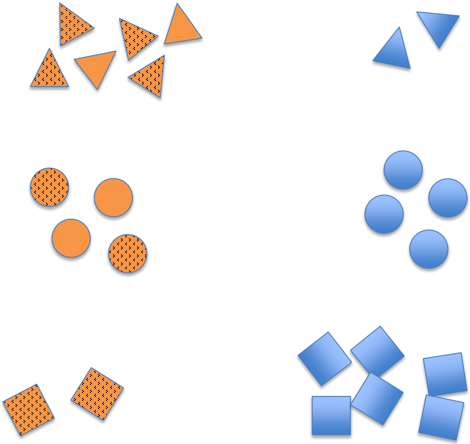
$$Y(1),Y(0) \perp\!\!\!\perp D \,|\, \blacktriangle$$

$$Y(1),Y(0) \perp\!\!\!\perp D \,|\, \bullet$$

$$Y(1),Y(0) \perp\!\!\!\perp D \,|\, \blacksquare$$

Selection on observables, “unconfoundedness”:

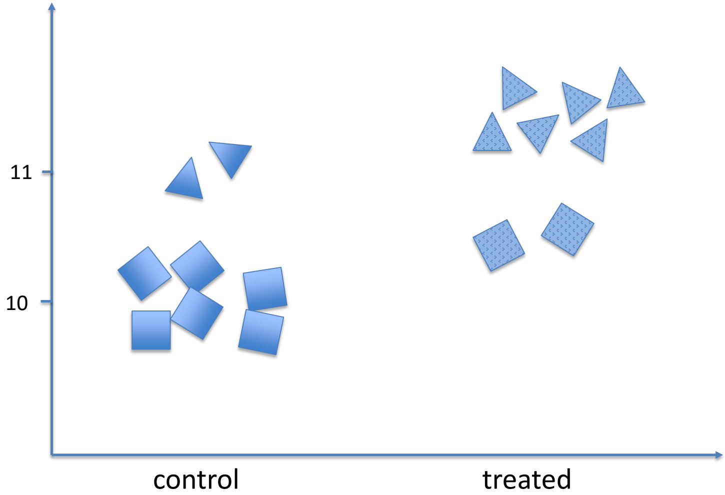
$$Y(1),Y(0) \perp\!\!\!\perp D \,|\, X$$

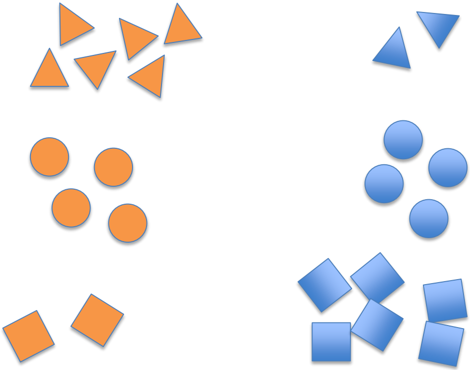


$$Y(1),Y(0) \not\perp\!\!\!\perp D \mid \blacktriangle$$

$$Y(1),Y(0) \not\perp\!\!\!\perp D \mid \bullet$$

$$Y(1),Y(0) \not\perp\!\!\!\perp D \mid \blacksquare$$

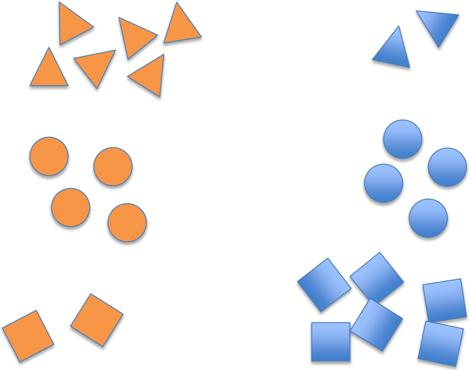

11
10
control
treated



$$\Pr(D=1 \mid \blacktriangle) = .75$$

$$\Pr(D=1 \mid \bullet) = .5$$

$$\Pr(D=1 \mid \blacksquare) = .25$$



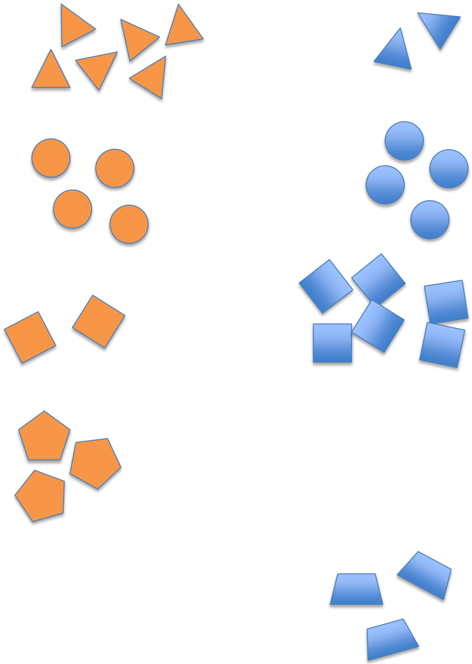
Pr(D=1|▲) = .75

Pr(D=1|●) = .5

Pr(D=1|■) = .25

Common support, "overlap":

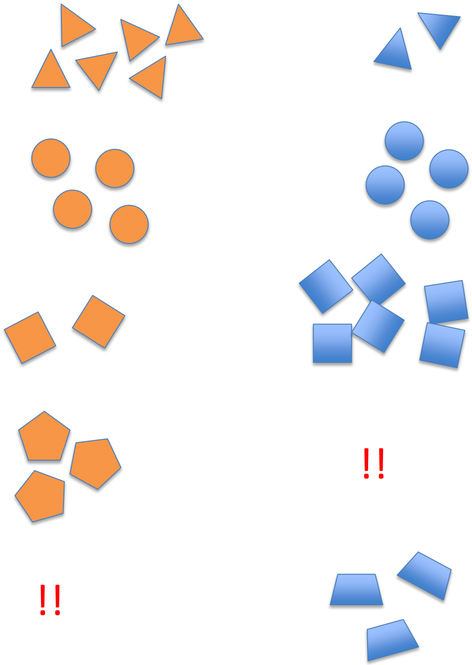
$$0 < \Pr(D=1|X) < 1$$



$\Pr(D=1|\blacktriangle) = .75$

$\Pr(D=1|\bullet) = .5$

$\Pr(D=1|\blacksquare) = .25$



$\Pr(D=1|\blacktriangle) = .75$

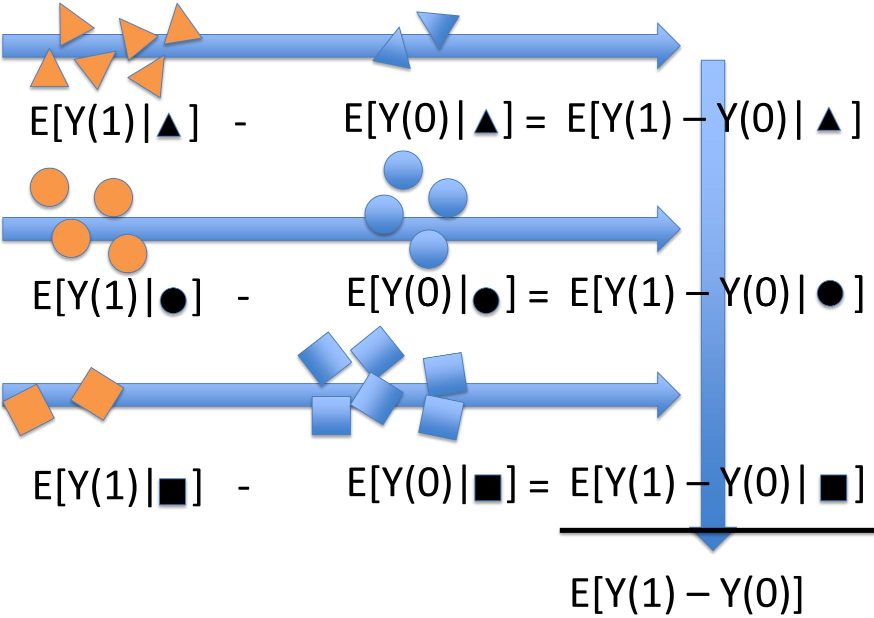
$\Pr(D=1|\bullet) = .5$

$\Pr(D=1|\blacksquare) = .25$

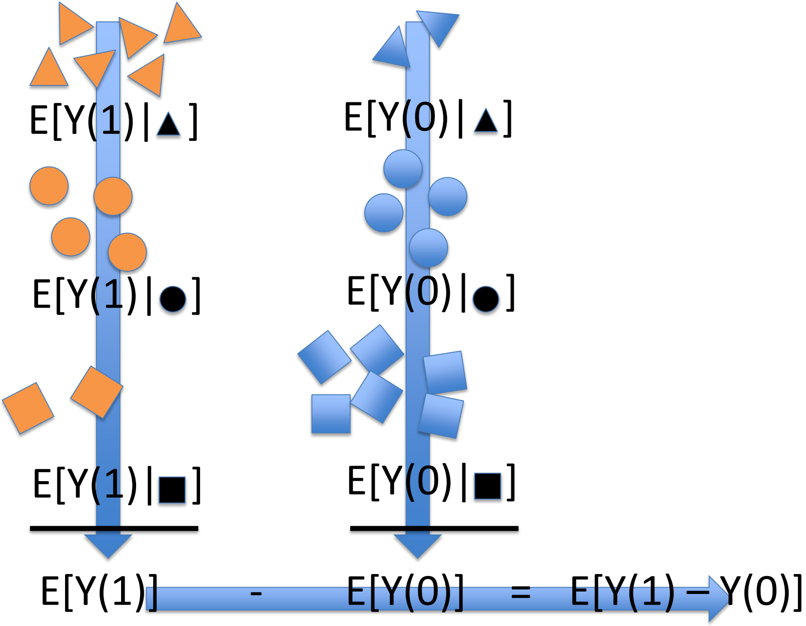
$\Pr(D=1|⬟) = 1$

$\Pr(D=1|⏢) = 0$

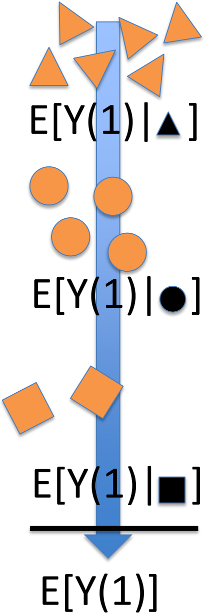
Matching, subclassification, p-score blocking



Inverse p-score weighting, regression



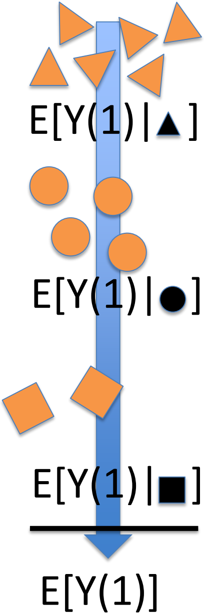
Inverse p-score weighting, regression



Inverse propensity score weighting:

$$
\begin{aligned}
\hat{\mu}_1 &= \bar{Y}_{1\blacktriangle} \frac{n_{1\blacktriangle} + n_{0\blacktriangle}}{n} + \cdots \\
&= \frac{1}{n} \left[n_{1\blacktriangle} \bar{Y}_{1\blacktriangle} \frac{n_{1\blacktriangle} + n_{0\blacktriangle}}{n_{1\blacktriangle}} + \cdots \right] \\
&= \frac{1}{n} \left[\sum_{i: D_i = 1, X_i = \blacktriangle} Y_i \frac{1}{\pi_{\blacktriangle}} + \cdots \right] \\
&= \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i D_i}{\pi_i}
\end{aligned}
$$

Inverse p-score weighting, regression



Regression:

$$
\begin{aligned}
\hat{\mu}_1 &= \bar{Y}_{1\blacktriangle} \frac{n_{1\blacktriangle} + n_{0\blacktriangle}}{n} + \cdots \\
&= \frac{1}{n} \left[\bar{Y}_{1\blacktriangle} \left(n_{1\blacktriangle} + n_{0\blacktriangle} \right) + \cdots \right] \\
&= \frac{1}{n} \left[\sum_{i:X_i=\blacktriangle}^{n} \bar{Y}_{1\blacktriangle} + \cdots \right] \\
&= \frac{1}{n} \sum_{i=1}^{n} \hat{Y}_i(1)
\end{aligned}
$$

Mathematical setup

D_i: treatment indicator
$Y_i(0)$: potential outcome if untreated
$Y_i(1)$: potential outcome if treated

Mathematical setup

D_i: treatment indicator
$Y_i(0)$: potential outcome if untreated
$Y_i(1)$: potential outcome if treated

Mathematical setup

D_i: treatment indicator
$Y_i(0)$: potential outcome if untreated
$Y_i(1)$: potential outcome if treated

Parameter of Interest (Average Treatment Effect)

$$E\left[Y_i(1) - Y_i(0)\right]$$

Mathematical setup

D_i: treatment indicator
$Y_i(0)$: potential outcome if untreated
$Y_i(1)$: potential outcome if treated

Parameter of Interest (Average Treatment Effect)

$$E\left[Y_i(1) - Y_i(0)\right]$$

- Main challenge:

$$E\left[Y_i | D_i = 1\right] \neq E\left[Y_i(1)\right]$$

$$E\left[Y_i | D_i = 0\right] \neq E\left[Y_i(0)\right]$$

Identification

Assumption (Stable Unit Treatment Value (SUTVA))

Let d be an $n \times 1$ vector of possible treatment assignments. For all d and d' such that $d_i = d_i'$, $Y_i(d) = Y_i(d') := Y_i(d_i)$.

Assumption (Selection on Observables)

$$Y_i(1), Y_i(0) \perp\!\!\!\perp D_i | X_i$$

Assumption (Common Support)

$$0 < \Pr(D_i = 1 | X_i) < 1 \textit{ almost surely}$$

Theorem (ATE Identification)

$$E[Y_i(1) - Y_i(0)] = E[E[Y_i | D_i = 1, X_i] - E[Y_i | D_i = 0, X_i]]$$

(matching!)

Identification

Assumption (Stable Unit Treatment Value (SUTVA))

Let d be an $n \times 1$ vector of possible treatment assignments. For all d and d' such that $d_i{=}d_i'$, $Y_i(d) = Y_i(d') := Y_i(d_i)$.

Assumption (Selection on Observables)

$$Y_i(1), Y_i(0) \perp\!\!\!\perp D_i | X_i$$

Assumption (Common Support)

$$0 < \Pr(D_i = 1 | X_i) < 1 \textit{ almost surely}$$

Theorem (ATE Identification)

$$E[Y_i(1) - Y_i(0)] = E[E[Y_i | D_i = 1, X_i]] - E[E[Y_i | D_i = 0, X_i]]$$

(regression!)

General guidelines for choosing covariates

Remember the formula for selection bias:

$$\begin{aligned}\text{Selection Bias} &= \left(E\left[Y(0)\,|X=1\right] - E\left[Y(0)\,|X=0\right]\right) \\ &\quad \times \left(\Pr\left(X=1|D=1\right) - \Pr\left(X=1|D=0\right)\right)\end{aligned}$$

- for which treated and control groups differ
- which affect outcomes
- not affected by treatment (beware of bad control!) example
- trade off between controlling for all confounding factors and curse of dimensionality
- possible resolution to tradeoff: control for propensity score

The role of the propensity score

Define the *propensity score* $\pi(X_i) := \Pr(D_i = 1 | X_i)$

- A sufficient control variable (Rosenbaum and Rubin 1983)

$$\text{Selection on Observables} \implies Y_i(1), Y_i(0) \perp\!\!\!\perp D_i | \pi(X_i)$$

The role of the propensity score

Define the *propensity score* $\pi(X_i) := \Pr(D_i = 1|X_i)$

- A sufficient control variable (Rosenbaum and Rubin 1983)

$$\text{Selection on Observables} \implies Y_i(1), Y_i(0) \perp\!\!\!\perp D_i|\pi(X_i)$$

- As a weighting strategy

$$E[Y_i(1)] = E\left[\frac{Y_i D_i}{\pi(X_i)}\right]$$

$$E[Y_i(0)] = E\left[\frac{Y_i(1-D_i)}{1-\pi(X_i)}\right]$$

Empirical Example: Return to Ivy League Education

- Research question: what is the financial return to attending an Ivy League college, relative to a less-prestigious college?

Empirical Example: Return to Ivy League Education

- Research question: what is the financial return to attending an Ivy League college, relative to a less-prestigious college?
- Data: College and Beyond Survey ($N = 14,238$, all attended a participating college)

Empirical Example: Return to Ivy League Education

- Research question: what is the financial return to attending an Ivy League college, relative to a less-prestigious college?
- Data: College and Beyond Survey ($N = 14,238$, all attended a participating college)
- D_i: binary indicator for attending Ivy-plus (Barnard, Columbia, Duke, Northwestern, Penn, Princeton, Stanford, Yale)

Empirical Example: Return to Ivy League Education

- Research question: what is the financial return to attending an Ivy League college, relative to a less-prestigious college?
- Data: College and Beyond Survey ($N = 14,238$, all attended a participating college)
- D_i: binary indicator for attending Ivy-plus (Barnard, Columbia, Duke, Northwestern, Penn, Princeton, Stanford, Yale)
- Y_i: personal income mid-career

Empirical Example: Return to Ivy League Education

- Research question: what is the financial return to attending an Ivy League college, relative to a less-prestigious college?
- Data: College and Beyond Survey ($N = 14,238$, all attended a participating college)
- D_i: binary indicator for attending Ivy-plus (Barnard, Columbia, Duke, Northwestern, Penn, Princeton, Stanford, Yale)
- Y_i: personal income mid-career
- X_i: ???

Empirical Example: Return to Ivy League Education

- Research question: what is the financial return to attending an Ivy League college, relative to a less-prestigious college?
- Data: College and Beyond Survey ($N = 14,238$, all attended a participating college)
- D_i: binary indicator for attending Ivy-plus (Barnard, Columbia, Duke, Northwestern, Penn, Princeton, Stanford, Yale)
- Y_i: personal income mid-career
- X_i: ???
  - SAT score

Empirical Example: Return to Ivy League Education

- Research question: what is the financial return to attending an Ivy League college, relative to a less-prestigious college?
- Data: College and Beyond Survey ($N = 14,238$, all attended a participating college)
- D_i: binary indicator for attending Ivy-plus (Barnard, Columbia, Duke, Northwestern, Penn, Princeton, Stanford, Yale)
- Y_i: personal income mid-career
- X_i: ???
  - SAT score
  - parents' education

Empirical Example: Return to Ivy League Education

- Research question: what is the financial return to attending an Ivy League college, relative to a less-prestigious college?
- Data: College and Beyond Survey ($N = 14,238$, all attended a participating college)
- D_i: binary indicator for attending Ivy-plus (Barnard, Columbia, Duke, Northwestern, Penn, Princeton, Stanford, Yale)
- Y_i: personal income mid-career
- X_i: ???
  - SAT score
  - parents' education
  - . . .

Estimation options

- Matching
- Weighting
- Regression
- Doubly-robust weighting + regression
- Propensity score subclassification + regression

Matching details (1)

1. For each observation i, find the m "nearest" neighbors in the opposite treatment group, $J_m(i)$
2. Impute $\hat{Y}_i(0), \hat{Y}_i(1)$ for each observation:

$$\hat{Y}_i(0) = \begin{cases} Y_i & , \quad D_i = 0 \\ \frac{1}{m}\sum_{j \in J_m(i)} Y_j & , \quad D_i = 1 \end{cases}$$

$$\hat{Y}_i(1) = \begin{cases} Y_i & , \quad D_i = 1 \\ \frac{1}{m}\sum_{j \in J_m(i)} Y_j & , \quad D_i = 0 \end{cases}$$

3. $\hat{\delta}^{\text{match}} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i(1) - \hat{Y}_i(0)\right)$

Matching details (2): What do we mean by "nearest"?

- simple Euclidean distance:

$$\sum_{l=1}^{k} (X_{li} - X_{lj})^2$$

- scaled Euclidean distance:

$$\sum_{l=1}^{k} (X_{li} - X_{lj})^2 / Var(X_l)$$

- Mahalanobis distance:

$$(X_i - X_j)' \Sigma_X^{-1} (X_i - X_j)$$

Weighting details

1. Estimate propensity score, e.g. `logit D X` $\rightarrow \hat{\pi}(X_i)$
2. Weight by inverse of propensity score to get average potential outcomes:
$$\hat{\mu}_1 = \frac{\sum_{i=1}^{n} Y_i D_i / \hat{\pi}(X_i)}{\sum_{i=1}^{n} D_i / \hat{\pi}(X_i)}$$
$$\hat{\mu}_0 = \frac{\sum_{i=1}^{n} Y_i (1 - D_i) / (1 - \hat{\pi}(X_i))}{\sum_{i=1}^{n} (1 - D_i) / (1 - \hat{\pi}(X_i))}$$
3. $$\hat{\delta}^{\text{IPW}} = \hat{\mu}_1 - \hat{\mu}_0$$

Regression details

1. Regress Y_i on X_i in $D_i = 0$ sample $\rightarrow \hat{\mu}_0(X_i)$
2. Regress Y_i on X_i in $D_i = 1$ sample $\rightarrow \hat{\mu}_1(X_i)$
3. $\hat{\delta}^{\text{reg}} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\right)$

Or, in one step:

$$Y_i = \delta D_i + X_i'\beta + D_i \times \left(X_i - \bar{X}\right)'\gamma + \varepsilon_i$$

Doubly-robust details

1. Estimate regression fits $\hat{\mu}_1(X_i), \hat{\mu}_0(X_i)$ and propensity score $\hat{\pi}(X_i)$ as above
2. $$\hat{\mu}_1^{\mathrm{DR}} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\mu}_1(X_i) + \frac{D_i(Y_i - \hat{\mu}_1(X_i))}{\hat{\pi}(X_i)}\right)$$
   $$\hat{\mu}_0^{\mathrm{DR}} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\mu}_0(X_i) + \frac{(1-D_i)(Y_i - \hat{\mu}_0(X_i))}{1-\hat{\pi}(X_i)}\right)$$
3. $$\hat{\delta}^{\mathrm{DR}} = \hat{\mu}_1^{\mathrm{DR}} - \hat{\mu}_0^{\mathrm{DR}}$$

- Advantage: consistent even if either propensity or regression functions misspecified
- But what if both are misspecified? (Kang and Schafer 2007)

Machine Learning + Double Robust (Chernozhukov, et al. 2018)

Divide sample into K folds and for each fold $k = 1, \ldots, K$:

1. Estimate regression fits $\hat{\mu}_1^k(X_i), \hat{\mu}_0^k(X_i)$ and propensity score $\hat{\pi}^k(X_i)$ using all data *except* fold k using favorite ML method (e.g. LASSO, random forest)
2. $$\hat{\mu}_1^k = \frac{1}{n_k} \sum_{i \in \text{fold } k} \left(\hat{\mu}_1^k(X_i) + \frac{D_i \left(Y_i - \hat{\mu}_1^k(X_i) \right)}{\hat{\pi}^k(X_i)} \right)$$
   $$\hat{\mu}_0^k = \frac{1}{n_k} \sum_{i \in \text{fold } k} \left(\hat{\mu}_0^k(X_i) + \frac{(1 - D_i) \left(Y_i - \hat{\mu}_0^k(X_i) \right)}{1 - \hat{\pi}^k(X_i)} \right)$$
3. $$\hat{\delta}^{\text{DR+ML}} = \frac{1}{K} \sum_{k=1}^{K} \left(\hat{\mu}_1^k - \hat{\mu}_0^k \right)$$

Propensity score subclassification + regression (Imbens and Wooldridge, 2007)

Poor man's double robustness (but works great!)

- Estimate propensity score, e.g. `logit D X` $\rightarrow \hat{\pi}(X_i)$

Propensity score subclassification + regression (Imbens and Wooldridge, 2007)

Poor man's double robustness (but works great!)

- Estimate propensity score, e.g. `logit D X` $\rightarrow \hat{\pi}(X_i)$
- Divide sample into J (e.g., $J = 5$) equal-sized blocks, sorting on $\hat{\pi}(X_i)$

Propensity score subclassification + regression (Imbens and Wooldridge, 2007)

Poor man's double robustness (but works great!)

- Estimate propensity score, e.g. `logit D X` $\rightarrow \hat{\pi}(X_i)$
- Divide sample into J (e.g., $J = 5$) equal-sized blocks, sorting on $\hat{\pi}(X_i)$
- For blocks $j = 1, \ldots, J$, estimate

$$Y_i = \alpha_j + \delta_j D_i + X_i' \beta_j + \epsilon_i$$

Propensity score subclassification + regression (Imbens and Wooldridge, 2007)

Poor man's double robustness (but works great!)

- Estimate propensity score, e.g. `logit D X` $\rightarrow \hat{\pi}(X_i)$
- Divide sample into J (e.g., $J = 5$) equal-sized blocks, sorting on $\hat{\pi}(X_i)$
- For blocks $j = 1, \ldots, J$, estimate

$$Y_i = \alpha_j + \delta_j D_i + X_i' \beta_j + \epsilon_i$$

- Final treatment effect:

$$\hat{\delta} = \frac{1}{J} \sum_{j=1}^{J} \hat{\delta}_j$$

Connection to multiple linear regression

Remember the familiar model:

$$Y_i = \delta D_i + X_i' \beta + \varepsilon_i$$

How does this relate to matching, propensity score weighting, etc.? (suppose X_i is saturated)

- Analogous assumptions:
  $E\left[\varepsilon_i | X_i\right] = 0 \sim Y_i\left(1\right), Y_i\left(0\right) \perp\!\!\!\perp D_i | X_i$
- Slightly different estimands:

$$\delta^{OLS} = \frac{E\left[\delta\left(X_i\right) Var\left(D_i | X_i\right)\right]}{E\left[Var\left(D_i | X_i\right)\right]}$$

$$\delta^{matching} = \delta^{IPW} = ATE$$

Inference

- Under suitable conditions the estimators above are approximately normally distributed in large samples:

$$\sqrt{n}\left(\hat{\delta} - ATE\right) \rightarrow_d N\left(0, V\right)$$

  with a variance V that can be estimated easily (formulas, bootstrap[1])

- Confidence intervals and hypothesis tests are then straightforward:

$$CI_{.95}\left(ATE\right) = \hat{\delta} \pm 1.96 \times \sqrt{V/n}$$

[1] except matching in some cases

Assessing the assumptions: Selection on observables

- Estimate "effects" on predetermined variables (besides those in X_i)
- Jointly test null of zero effect
- If reject, consider expanding X_i

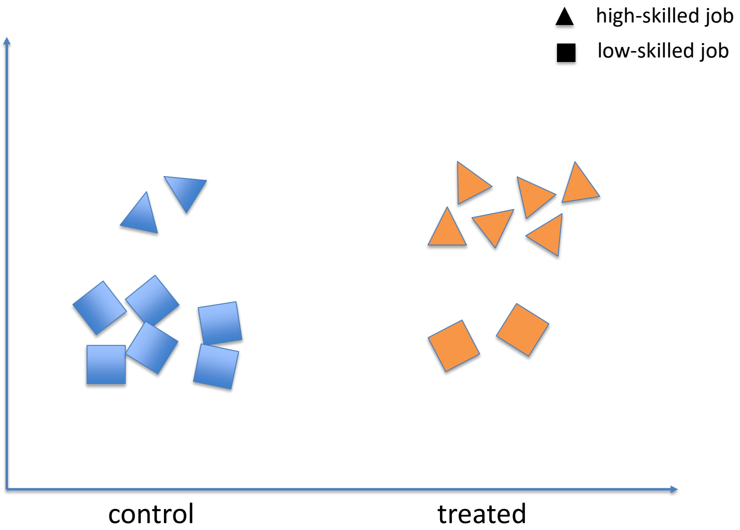
Assessing the assumptions: Common support

- Visually plot covariate distributions by treatment status
- Compute normalized imbalances (Imbens 2015):

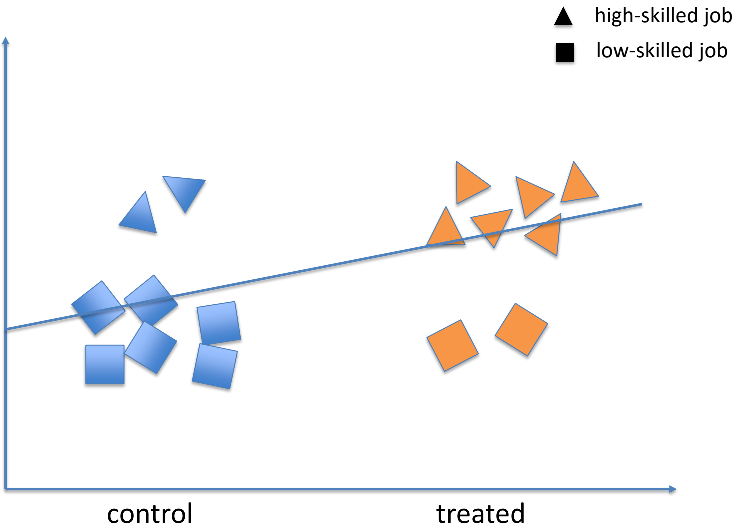
$$\frac{\bar{X}_1 - \bar{X}_0}{\sqrt{\left(\hat{S}_1^2 + \hat{S}_0^2\right)/2}}$$

- Examine estimated propensities
- Consider trimming observations with propensities very near zero or one

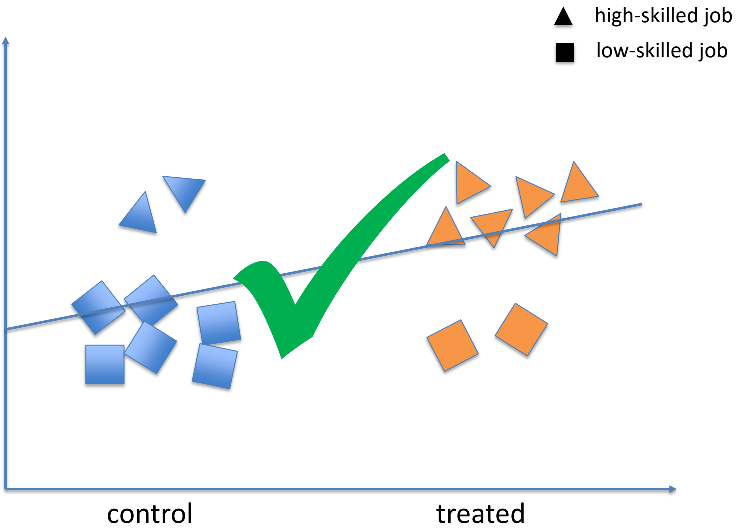
back



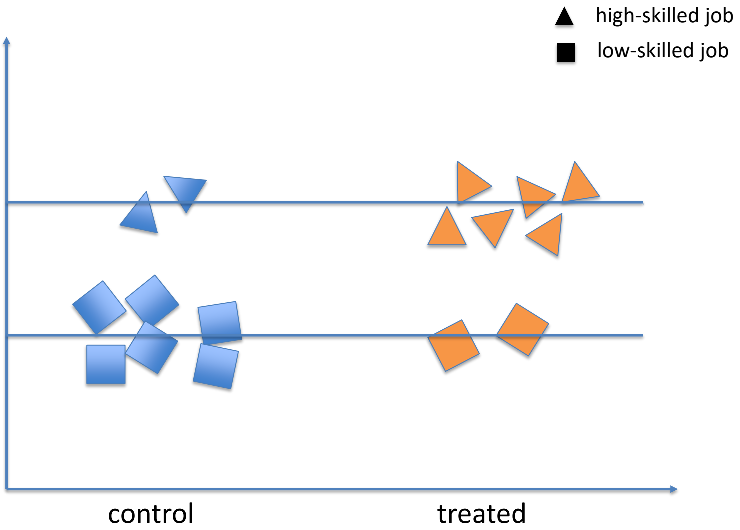
back

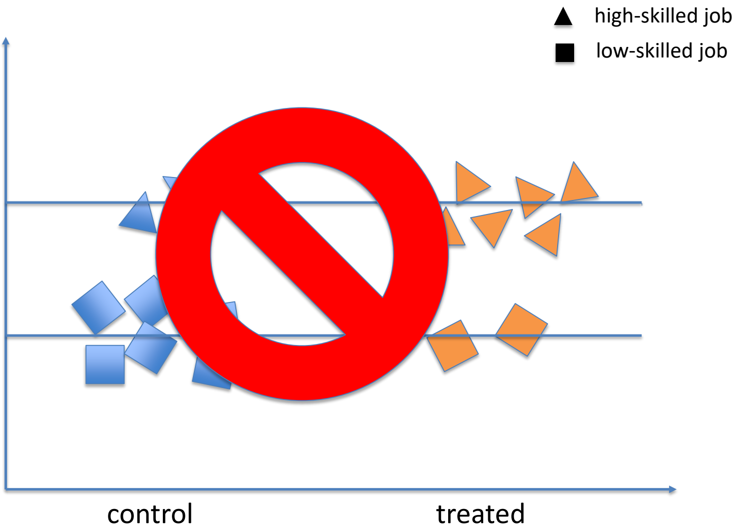


back



back




high-skilled job
low-skilled job
control
treated