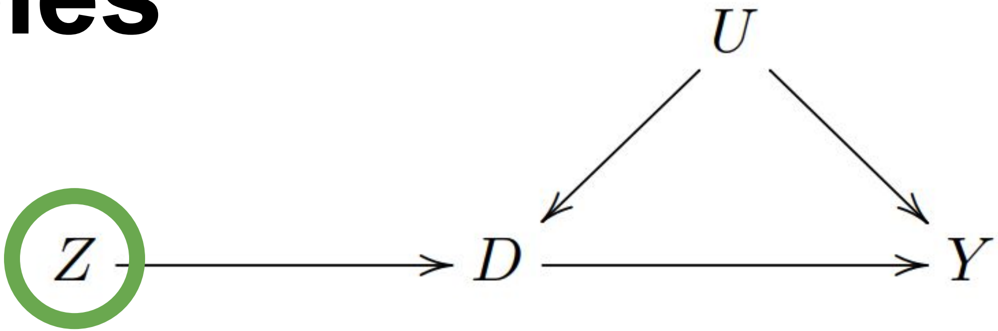
Understanding Instrumental Variables

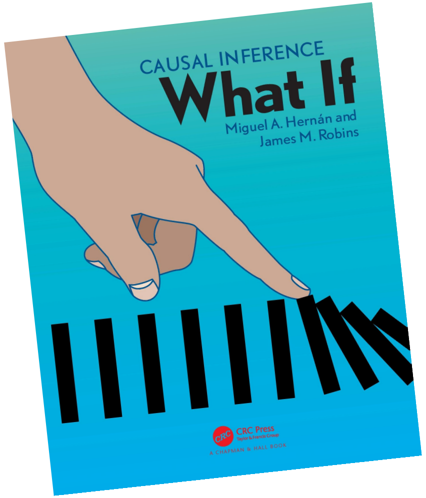


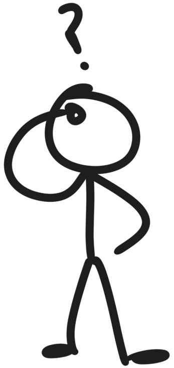
with the Potential Outcomes Framework

Mentor: Rui Wang | **Mentee:** Eric Leonen

- **Potential Outcomes Framework**
- **Encouragement Design**
- **Example: Job Search Intervention Study**
- **What’s an Instrument?**

The Potential Outcome Framework [1]





Some notation…

Units are indexed $i = 1, \ldots, n$

$D_i \in \{0, 1\}$ is the observed, binary **treatment** for unit i

Y_i is observed outcome for unit i

$Y_i \in \{Y_i(D = 0), Y_i(D = 1)\}$ are **potential outcomes** for unit i

X_i are covariates for unit i

Individual treatment effect (ITE) is the difference in potential outcomes under treatment vs. no treatment for a unit

$$\tau_i = Y_i(D=1) - Y_i(D=0)$$

Unfortunately, ITE is impossible to measure!

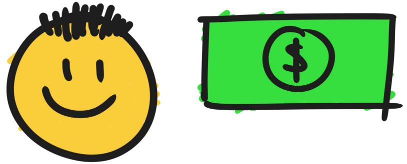
An **average treatment effect (ATE)** is the average treatment effect across the entire population

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

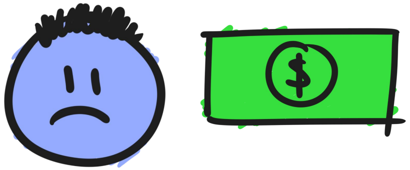
An example…

Does money cause happiness?

If $Y_i(D = 0)$ $Y_i(D = 1)$ then YES!

If $Y_i(D = 0)$ $Y_i(D = 1)$ then NO!

Assuming **strong ignorability** $\{Y(D=0), Y(D=1)\} \perp D \mid X$

and **positivity** $0 < P(D=1) < 1$

$$\tau = E[Y|D=1] - E[Y|D=0]$$

$$\hat{\tau} = \frac{1}{n_1}\sum_{i=1}^{n} Z_i Y_i - \frac{1}{n_0}\sum_{i=1}^{n}(1-Z_i)Y_i$$

Possible with a **completely randomized experiment (CRE)**

Unfortunately, a valid CRE *forces* treatment onto units

Encouragement Design [2]

Some more notation…

$Z_i \in \{0, 1\}$ is the assigned, binary treatment

$$D_i \in \{D_i(Z=0), D_i(Z=1)\} \quad \text{are potential, actual treatments}$$
$$= \{D_i(0), D_i(1)\}$$

$$Y_i \in \{Y_i(Z=0), Y_i(Z=1)\} \quad \text{are potential outcomes given } Z_i$$
$$= \{Y(0), Y(1)\}$$

Assuming **<u>“Independence”</u>** $Z \perp \{D(1), D(0), Y(1), Y(0)\}$

$$\tau_D = E[D(1) - D(0)] = E[D \mid Z = 1] - E[D \mid Z = 0]$$

$$\tau_Y = E[Y(1) - Y(0)] = E[Y \mid Z = 1] - E[Y \mid Z = 0]$$

Unfortunately, this is **<u>intention to treat (ITT)</u>**, not ATE

Not everyone does what they're told! There are **compliance groups**

$$U_i = \begin{cases} a, & D_i(0) = D_i(1) = 1 \\ c, & D_i(0) = 0, D_i(1) = 1 \\ d, & D_i(0) = 1, D_i(1) = 0 \\ n, & D_i(0) = D_i(1) = 0 \end{cases}$$

$$\tau_Y = \sum_u E[Y(1) - Y(0) \mid U = u] P(U = u)$$

Assuming **"Monotonicity"** $D_i(1) \geq D_i(0)$

and **"The Exclusion Restriction"** $Y_i(1) = Y_i(0)$ if $U_i \in \{a, n\}$

$$\tau_Y = E[Y(1) - Y(0) \mid U = c] P(U = c)$$

$$\tau_D = E[D(1) - D(0) \mid U = c] P(U = c) = P(U = c)$$

Assuming **"Relevance"** $\tau_D \neq 0$

We have **local average treatment effect (LATE)!** [3]

$$\tau_C = E[Y(D=1) - Y(D=0) \mid U = c]$$
$$= E[Y(1) - Y(0) \mid U = c]$$
$$= \frac{\tau_Y}{\tau_D}$$

The **Wald estimand:** $\tau_C = \dfrac{E[Y \mid Z = 1] - E[Y \mid Z = 0]}{E[D \mid Z = 1] - E[D \mid Z = 0]}$

Unfortunately, this is still not quite ATE

LATE is approximately ATE if compliers represent the population.
So, it may be helpful to characterize compliers' covariates X (4)

$$
\begin{aligned}
E[g(X) \mid U = c] &= \int g(x) f(x|U = c) dx && \text{def. of } E[\cdot] \\
&= \int g(x) \frac{P(U = c \mid X = x) f(x)}{P(U = c)} dx && \text{Bayes' theorem} \\
&= \frac{1}{P(U = c)} \int g(x) P(U = c \mid X = x) f(x) dx \\
&= \frac{E[g(X) P(U = c \mid X)]}{P(U = c)} \\
&= \frac{E[g(X)\{E[D \mid Z = 1, X] - E[D \mid Z = 0, X]\}]}{E[D \mid Z = 1] - E[D \mid Z = 0]} && \text{formula for } \tau_D \\
&= \frac{E[g(X) E[D \mid Z = 1, X]] - E[g(X) E[D \mid Z = 0, X]]}{E[D \mid Z = 1] - E[D \mid Z = 0]}
\end{aligned}
$$

Connection to TSLS…

An alternative way to estimate:

$\hat{\tau}_Y$ is the coefficient from $Y \sim Z$

$\hat{\tau}_D$ is the coefficient from $D \sim Z$

Wald estimator with TSLS

$$\hat{\tau}_C = \frac{\hat{\tau}_Y}{\hat{\tau}_D}$$

Now add control for the covariates X!

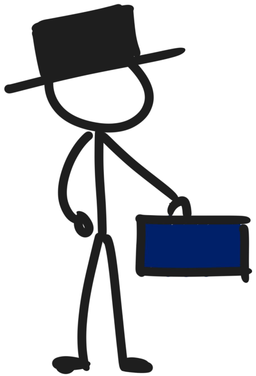
$\hat{\tau}_{Y,L}$ is the coefficient from $Y \sim Z + X$

$\hat{\tau}_{D,L}$ is the coefficient from $D \sim Z + X$

Lin Estimator (5)

$$\hat{\tau}_{c,L} = \frac{\hat{\tau}_{Y,L}}{\hat{\tau}_{D,L}}$$

Example: Job Search Intervention Study [6]



A Question

Does a job training program make unemployed people more effective at job searching?

Z_i is whether a person was randomly selected to participate

D_i is whether a person participated in the training program

Y_i is a person's efficacy (1-5) at job searching

X_i are a person's covariates (sex, age, married, race, education, and income)

Data

Source: **Job Search Intervention Study (JOBS II)**

- Subsample of the whole dataset includes 899 unemployed workers
- Workers randomly assigned to treatment and control:
  - **Treatment:** workshops that taught job-search skills and coping strategies
  - **Control:** no workshop
- Surveyed for job-search efficacy

Justification for Assumptions

- **Independence:** Treatment assignment was *random*
- **Relevance:** $\hat{\tau}_D \neq 0$
- **Exclusion Restriction:** Assignment shouldn't affect outcome
- **Monotonicity:** No defiers (nor any always takers)

Results

estimator	tau_Y	tau_D	estimate	std err	lower CI	upper CI
Wald (Formula)	0.067	0.62	0.109	0.081	-0.05	0.267
Wald (Regression)	0.067	0.62	0.109	0.081	-0.049	0.267
Lin	0.072	0.613	0.117	0.081	-0.041	0.276

* Standard errors estimated with bootstrapping (n = 10000 samples)

Are "Compliers" Representative?

☹ **Unfortunately,** compliers may not be representative

$$\hat{E}[\text{sex} \mid U = c] - \hat{E}[\text{sex}] = 0.088$$ Female

$$\hat{E}[\text{age} \mid U = c] - \hat{E}[\text{age}] = 14.974$$ Older

$$\hat{E}[\text{nonwhite} \mid U = c] - \hat{E}[\text{nonwhite}] = 0.019$$ White

$$\hat{E}[\text{educ} \mid U = c] - \hat{E}[\text{educ}] = 0.884$$ More education

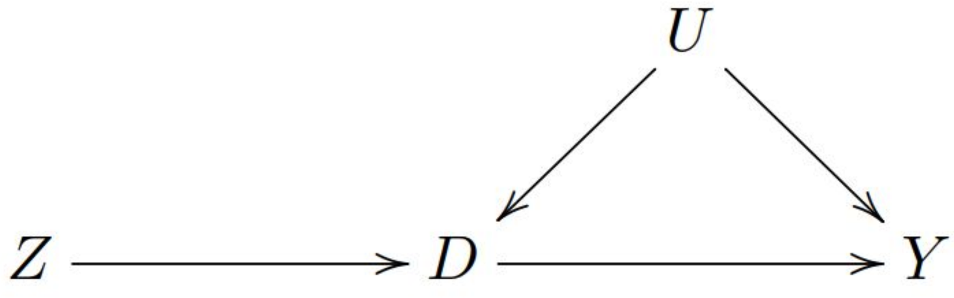
$$\hat{E}[\text{income} \mid U = c] - \hat{E}[\text{income}] = 0.925$$ More income

What’s an Instrument?

$Z =$ 

An **instrument** is a variable that is independent of potential outcomes and affects the observed outcome only by shifting the treatment in a single direction.

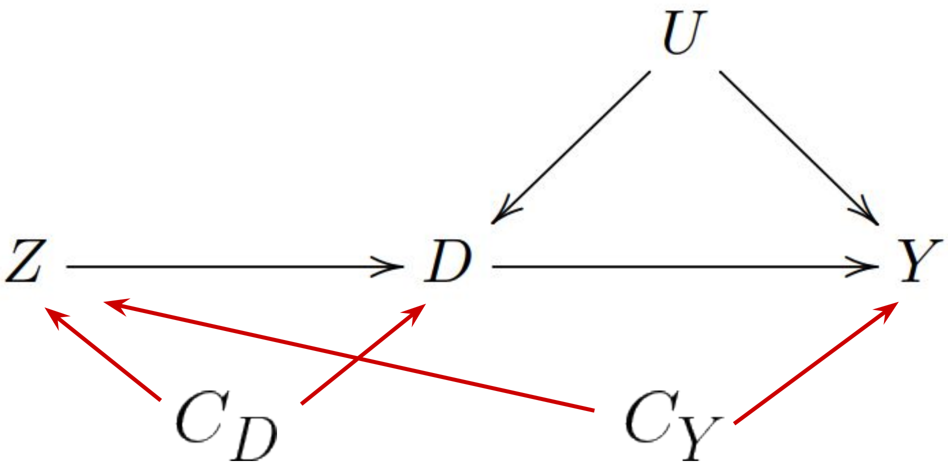
- Independence
- Relevance
- Exclusion Restriction
- Monotonicity



Independence

$$Z \perp \{D(1), D(0), Y(1), Y(0)\}$$

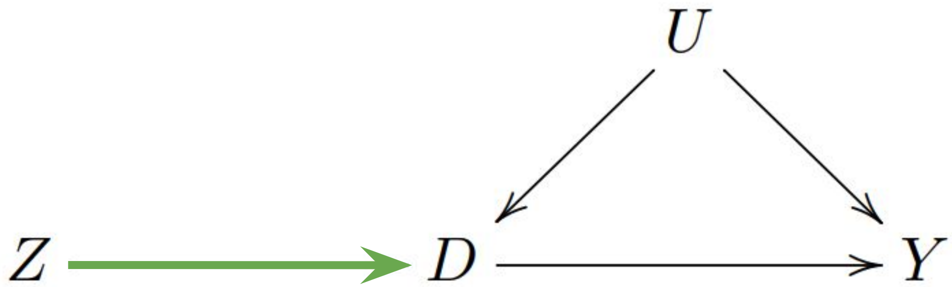
We cannot have:



Relevance

$$Z \not\perp D$$

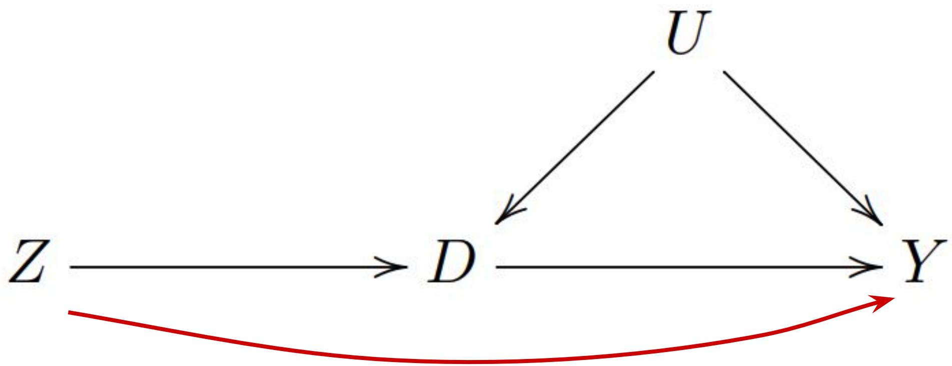
A weak correlation means a **weak instrument**



Exclusion Restriction

The instrument only affects the outcome through the treatment

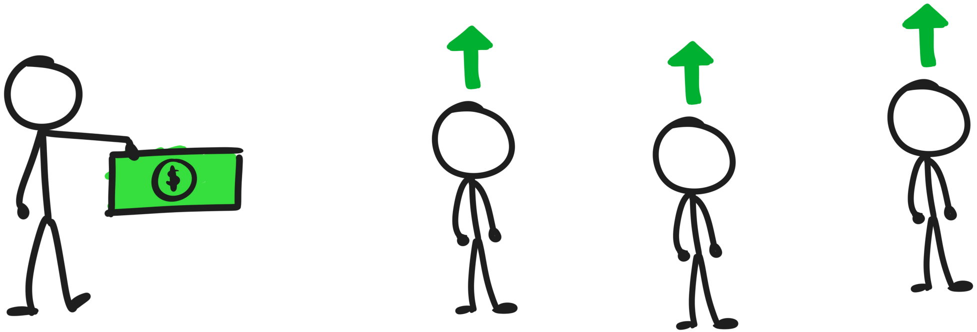
We cannot have:



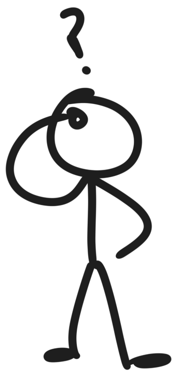
Monotonicity

The instrument affects the treatment for all units in the same direction

$$D_i(1) \geq D_i(0)$$



Questions?



References

(1) Hernán, Miguel A., and James M. Robins. Causal Inference: What If. Chapman & Hall/CRC, 2020.
(2) Ding, Peng. A First Course in Causal Inference. Chapman & Hall/CRC, 2023
(3) Imbens, Guido W., and Joshua D. Angrist. "Identification and Estimation of Local Average Treatment Effects." Econometrica, vol. 62, no. 2, 1994, pp. 467–75. JSTOR, https://doi.org/10.2307/2951620.
(4) Baiocchi, Michael et al. "Instrumental variable methods for causal inference." Statistics in medicine vol. 33,13 (2014): 2297-340. doi:10.1002/sim.6128
(5) Lin, Winston. "AGNOSTIC NOTES ON REGRESSION ADJUSTMENTS TO EXPERIMENTAL DATA: REEXAMINING FREEDMAN'S CRITIQUE." The Annals of Applied Statistics, vol. 7, no. 1, 2013, pp. 295–318. JSTOR, http://www.jstor.org/stable/23566512.
(6) Vinokur, A D et al. "Impact of the JOBS intervention on unemployed workers varying in risk for depression." American journal of community psychology vol. 23,1 (1995): 39-74. doi:10.1007/BF02506922