Section 7

Instrumental Variables

Sooahn Shin

GOV 2003

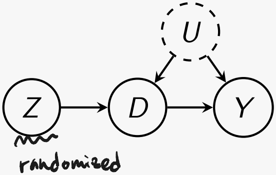
Oct 21, 2021

Overview

- Logistics:
  - **Pset 6 released!** Due at 11:59 pm (ET) on Oct 27
- Today's topics:
  1. Noncompliance in randomized experiments
  2. IV in observational studies using TSLS

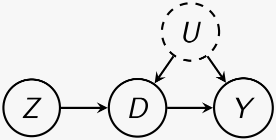
Noncompliance in randomized experiments

- Motivation: What if there is unmeasured confounding?
- In randomized experiments: when treatment assignment is randomized but cannot intervene treatment uptake.
  - $\rightsquigarrow$ **noncompliance** (one- or two-sided)
- DAG example:



Noncompliance in randomized experiments

- Motivation: What if there is unmeasured confounding?
- In randomized experiments: when treatment assignment is randomized but cannot intervene treatment uptake.
  - $\rightsquigarrow$ **noncompliance** (one- or two-sided)
- DAG example:



- Estimand: LATE = ITT effect on the outcome for compliers

$D_i(1) = 1$
$D_i(0) = 0$

$$\mathrm{ITT}_{Y,\mathrm{co}} = \frac{1}{n_{\mathrm{co}}} \sum_{i:C_i=\mathrm{co}} Y_i(1, D_i(1)) - Y_i(0, D_i(0))$$

Z_i $D_i(Z_i)$ $\rightarrow$ $Y_i(1,1) - Y_i(0,0)$

  - Q: Why not ITT_Y?

$\times$ $ITT_Y = ITT_{Y,co}\,\pi_{co} + ITT_{Y,nc}\,\pi_{nc}$

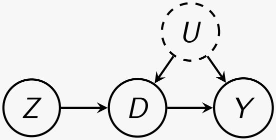
π_{co}: prop. comp.

$D_i(0) = 0$ $D_i(1) = 0$

$$\frac{1}{n_{nc}} \sum_{i:C_i=nc} Y_i(1, \underset{0}{\underset{\|}{D_i(1)}}) - Y_i(0, \underset{0}{\underset{\|}{D_i(0)}})$$

Noncompliance in randomized experiments

- Motivation: What if there is unmeasured confounding?
- In randomized experiments: when treatment assignment is randomized but cannot intervene treatment uptake.
  - $\rightsquigarrow$ **noncompliance** (one- or two-sided)
- DAG example:



- Estimand: LATE = ITT effect on the outcome for compliers

$$\mathrm{ITT}_{Y,\mathrm{co}} = \frac{1}{n_{\mathrm{co}}} \sum_{i:C_i=\mathrm{co}} Y_i(1, D_i(1)) - Y_i(0, D_i(0))$$

  - Q: Why not ITT_Y?
- Example (one- or two-sided)? Identification? Estimation?

Noncompliance (one-sided)

- One-sided example:
  canvass assignment (Z_i) - canvass recieved (D_i) - turn out (Y_i)
  - $D_i(0) = 0 \forall i$
  - $D_i(1) = 0$ or 1 depending on compliance type

Noncompliance (one-sided)

- One-sided example:
  canvass assignment (Z_i) - canvass recieved (D_i) - turn out (Y_i)
  - $D_i(0) = 0 \forall i$
  - $D_i(1) = 0$ or 1 depending on compliance type
- Compliance type by $D_i(1)$:
  - $D_i(1) = 1$: Compliers. If assigned to canvassing, I **would** recieve it.
  - $D_i(1) = 0$: Noncompliers. If assigned to canvassing, I **would not** receive it.
  - Q: Can we identify this by observing Z_i and D_i? What can we know (hint: ITT_D)?

$D_i(0)=0$

Z_i	D_i	
1	1	$\rightsquigarrow D_i(1)=1 \rightsquigarrow$ compliers
1	0	$\rightsquigarrow D_i(1)=0 \rightsquigarrow$ non-comp
0	0	$\rightsquigarrow D_i(0)=0 \rightsquigarrow D_i(0)=0$, $D_i(1)=? =0$ or 1

$$ITT_D = \underbrace{ITT_{D,co}}_{=1}\pi_{co} + \underbrace{ITT_{D,nc}}_{=0}\pi_{nc}$$

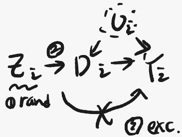
$$= \pi_{co}$$

$$ITT_{D,co} = \frac{1}{n_{co}} \sum_{i:C_i=co} \underbrace{D_i(1)}_{=1} - \underbrace{D_i(0)}_{=0} = \frac{n_{co}}{n_{co}} = 1$$

$$ITT_{D,nc} = \frac{1}{n_{nc}} \sum_{i:C_i=nc} \underbrace{D_i(1)}_{=0} - \underbrace{D_i(0)}_{=0} = 0$$

Noncompliance (one-sided)

- One-sided example:
  canvass assignment (Z_i) - canvass recieved (D_i) - turn out (Y_i)
  - $D_i(0) = 0 \forall i$
  - $D_i(1) = 0$ or 1 depending on compliance type
- Compliance type by $D_i(1)$:
  - $D_i(1) = 1$: Compliers. If assigned to canvassing, I **would** recieve it.
  - $D_i(1) = 0$: Noncompliers. If assigned to canvassing, I **would not** receive it.
  - Q: Can we identify this by observing Z_i and D_i? What can we know (hint: ITT_D)?
- Assumptions:
  1. Randomization of Z_i
  2. Presence of some compliers $\pi_{co} \neq 0$ "ITT_D
  3. Exclusion restriction $Y_i(z, d) = Y_i(z', d)$ (i.e., Z_i only affects Y_i through D_i)



Noncompliance (one-sided)

- Identification: LATE Theorem under the previous assumptions

$$\tau_{\text{LATE}} = \text{ITT}_{Y,co} = \frac{\text{ITT}_Y}{\text{ITT}_D}$$

$$\text{ITT}_Y = \text{ITT}_{Y,co}\,\pi_{co} + \underbrace{\cancel{\text{ITT}_{Y,nc}}^{\,0}}\,\pi_{nc}$$

$$\underbrace{\text{ITT}_D} = \underbrace{\pi_{co}}$$

$$\text{ITT}_{Y,co} = \frac{\text{ITT}_Y}{\text{ITT}_D}$$

$$\frac{1}{n_{nc}} \sum_{i:C_i = nc} Y_i(1, D_i(1)) - Y_i(0, D_i(0))$$

$$\parallel$$

$$\underbrace{Y_i(1,0)} - \underbrace{Y_i(0,0)}$$

$$Z_i \to \boxed{D_i} \to Y_i$$

$$Y_i(D_i = 0) - Y_i(D_i = 0)$$

$$= 0$$

Noncompliance (one-sided)

- Identification: LATE Theorem under the previous assumptions

$$\tau_{\text{LATE}} = \text{ITT}_{Y,co} = \frac{\text{ITT}_Y}{\text{ITT}_D}$$

- Point estimation:
    - **Wald** or **IV estimator**

$$\widehat{\tau}_{iv} = \frac{\widehat{\text{ITT}}_Y}{\widehat{\text{ITT}}_D} \quad \begin{matrix} \leftarrow \text{Diff-in-means} \\ \leftarrow \quad \text{"} \end{matrix}$$

        - It is biased, but consistent for τ_{LATE}

Noncompliance (one-sided)

- Identification: LATE Theorem under the previous assumptions

$$\tau_{\text{LATE}} = \text{ITT}_{Y,co} = \frac{\text{ITT}_Y}{\text{ITT}_D}$$

- Point estimation:
  - **Wald** or **IV estimator**

$$\widehat{\tau}_{iv} = \frac{\widehat{\text{ITT}}_Y}{\widehat{\text{ITT}}_D}$$

    - It is biased, but consistent for τ_{LATE}
  - Equivalent to the **TSLS estimator** under binary instrument and binary treatment
    1. Regress D_i on Z_i and get fitted values $\widehat{D}_i$
    2. Regress Y_i on $\widehat{D}_i$ and get the slope
    - Intuitively, TSLS retains only the variation in D_i that is generated by the instrument Z_i in the first stage.
    - $\rightsquigarrow$ use `AER::ivreg()` in practice.

$D_i \sim Z_i \rightsquigarrow \hat{D}_i$

$Y_i \sim \hat{D}_i \rightsquigarrow \hat{\tau}_{2SLS} = \hat{\tau}_{IV}$

Noncompliance (one-sided)

- Variance estimation:
  - **Wald** estimator: Use delta method to find the asymptotic variance

$$\mathbb{V}[\widehat{\tau}_{iv}] \approx \frac{1}{\mathsf{ITT}_D^2}\mathbb{V}\left[\widehat{\mathsf{ITT}}_Y\right] + \frac{\mathsf{ITT}_Y^2}{\mathsf{ITT}_D^4}\mathbb{V}\left[\widehat{\mathsf{ITT}}_D\right] - 2\frac{\mathsf{ITT}_Y}{\mathsf{ITT}_D^3}\mathsf{cov}\left[\widehat{\mathsf{ITT}}_Y, \widehat{\mathsf{ITT}}_D\right]$$

  - **TSLS** estimator: **Don't** use SEs from second step (see MHE section 4.6.1 2SLS Mistakes) $\rightsquigarrow$ use `ivpack::robust.se()` in practice.

Noncompliance (two-sided)

- Two-sided example:
  encouragement (Z_i) - treatment (D_i) - outcome (Y_i)
  - Or, testing habitual voting (Coppock and Green 2016):
    GOTV canvassing (2006) - turn out (2006) - turn out (2008)
  - LATE: Habitual voting for those who would vote iif they are contacted by a canvasser in this election

Noncompliance (two-sided)

- Two-sided example:
  encouragement (Z_i) - treatment (D_i) - outcome (Y_i)
  - Or, testing habitual voting (Coppock and Green 2016):
    GOTV canvassing (2006) - turn out (2006) - turn out (2008)
  - LATE: Habitual voting for those who would vote iif they are contacted by a canvasser in this election
- Compliance type by $(D_i(0), D_i(1))$:
  - (0,1): Complier
  - (1,1): Always-taker
  - (0,0): Never-taker
  - (1,0): Defier
  - Q: Can we identify this by observing Z_i and D_i? What can we know (hint: ITT_D)?

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	co / nt	nt / de
$D_i = 1$	at / de	co / at

$D_i(0)=0$ $D_i(0)=1$ $D_i(1)=0$ $D_i(1)=1$

$\text{ITT}_D = \pi_{co} - \pi_{de}$

Noncompliance (two-sided)

- Two-sided example:
  encouragement (Z_i) - treatment (D_i) - outcome (Y_i)
  - Or, testing habitual voting (Coppock and Green 2016):
    GOTV canvassing (2006) - turn out (2006) - turn out (2008)
  - LATE: Habitual voting for those who would vote iif they are contacted by a canvasser in this election
- Compliance type by $(D_i(0), D_i(1))$:
  - (0,1): Complier
  - (1,1): Always-taker
  - (0,0): Never-taker
  - (1,0): Defier
  - Q: Can we identify this by observing Z_i and D_i? What can we know (hint: ITT_D)?
- Assumptions: 1-3 from the previous setup, and
  4. Monotonicity: $D_i(1) \geq D_i(0), \forall i$ (no defiers)
  - Q: What does exclusion restriction/monotonicity imply in words?

Noncompliance (two-sided)

- Same identification result:

$$\tau_{\text{LATE}} = \text{ITT}_{Y,co} = \frac{\text{ITT}_Y}{\text{ITT}_D}$$

- Same estimation as before.

$$\boxed{ITT_Y = \underbrace{ITT_{Y,co}\,\pi_{co}}} + \underbrace{\cancel{ITT_{Y,at}}}_{=0\ (e.r)}\,\pi_{at} + \underbrace{\cancel{ITT_{Y,nt}}}_{=0\ (e.r)}\,\pi_{nt} + \cancel{ITT_{Y,de}}\,\underbrace{\pi_{de}}_{=0\ (mon)}$$

$$\boxed{ITT_D = \pi_{co}} - \underbrace{\cancel{\pi_{de}}}_{=0}$$

Noncompliance (two-sided)

- Same identification result:

$$\tau_{\text{LATE}} = \text{ITT}_{Y,co} = \frac{\text{ITT}_Y}{\text{ITT}_D}$$

- Same estimation as before.
- Further issues:
  - What if exclusion restriction/monotonicity is violated? Can we still use IV estimand for LATE? [Pset 6 Q2]
  - Detecting weak instruments? [Pset 6 Q3 (c)]   F-test

$Z_i \rightarrow D_i \rightarrow Y_i$

$D_i = \alpha + \gamma Z_i + \varepsilon_i$

$H_0: \gamma = 0$

In R: Wald estimator

$= \frac{\widehat{ITT_Y}}{\widehat{ITT_D}}$

```r
# *Recall what we did in Neyman's approach*
my_data # data includes Z, D, and Y

# Proportion of compliers (using ITT_D)
pi_co <- mean(my_data$D[my_data$Z == 1]) - mean(my_data$D[my_data$Z == 0])

# Compute ITT's
ITT_Y <- mean(my_data$Y[my_data$Z == 1]) - mean(my_data$Y[my_data$Z == 0])
ITT_D <- mean(my_data$D[my_data$Z == 1]) - mean(my_data$D[my_data$Z == 0])
# (ITT_D = pi_co)

# TODO 1: Compute Wald estimator
Wald_est <- NULL
```

In R: Wald estimator

```r
# TODO 2: Compute variance
# TODO 2-1: Compute variance terms using neyman estimator
Var_ITT_Y_est <- NULL
Var_ITT_Y_est <- NULL

# Compute covariance term
# demean
demeaned_y <- my_data$Y[my_data$Z == 1] - mean(my_data$Y[my_data$Z == 1])
demeaned_d <- my_data$D[my_data$Z == 1] - mean(my_data$D[my_data$Z == 1])
# denominator
denom <- sum(my_data$Z)*(sum(my_data$Z) - 1)
Covar_est <- (demeaned_y %*% demeaned_t)/denom

# TODO 2-2: Compute the estimate of the formula in p.6
Var_Wald_est <- NULL
```

In R: TSLS estimator

$Y \sim D \mid Z$

treat inst.

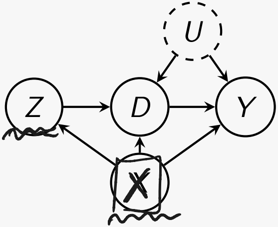
U

$Z \rightarrow D \rightarrow Y$

```
ivmodel <- AER::ivreg(Y ~ D | Z, data = my_data)
ivpack::robust.se(ivmodel)
```

IV in observational studies using TSLS

- Motivation: What if there is unmeasured confounding?
- In observational studies: In case where
  - treatment is not randomized and there exist unmeasured confounder;
  - can find instrumental variable;
  - "exogenous covariates" ($\mathbf{X}_i$): may exist observable confounders between Z_i, D_i, and Y_i
- DAG example:



IV in observational studies using TSLS

- TSLS is the classical approach to IV
  - w/o covariates

$$D_i = \hat{\delta} + \gamma Z_i + \eta_i$$
$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

$D_i \sim Z_i \to \hat{\gamma}$

$Y_i \sim \underbrace{\hat{\gamma} Z_i}_{\hat{D}_i} \to \hat{\tau}_{2SLS}$

$= \boxed{\frac{Cov(Y_i, Z_i)}{Cov(D_i, Z_i)}}$

$\tau = \frac{Cov(Y_i, D_i)}{V(D_i)}$

IV in observational studies using TSLS

- TSLS is the classical approach to IV
  - w/o covariates

$$D_i = \delta + \gamma Z_i + \eta_i$$
$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

  - w/ covariates

$$D_i = \delta + \gamma Z_i + \mathbf{X}_i' \beta_d + \eta_i$$
$$Y_i = \alpha + \tau D_i + \mathbf{X}_i' \beta_y + \varepsilon_i$$

IV in observational studies using TSLS

- TSLS is the classical approach to IV
  - w/o covariates

$$D_i = \delta + \gamma Z_i + \eta_i$$
$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

  - w/ covariates

$$D_i = \delta + \gamma Z_i + \mathbf{X}_i' \beta_d + \eta_i$$
$$Y_i = \alpha + \tau D_i + \mathbf{X}_i' \beta_y + \varepsilon_i$$

- Recall the four canonical IV assumptions.
- Suppose we have binary treatment and binary instrument.
  - w/o covariates, TSLS estimand (τ) = LATE (τ_{LATE}) and TSLS estimator = Wald estimator

IV in observational studies using TSLS

- TSLS is the classical approach to IV
  - w/o covariates

$$D_i = \delta + \gamma Z_i + \eta_i$$
$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

  - w/ covariates

$$\left\langle \begin{array}{l} D_i = \delta + \gamma Z_i + \boxed{\mathbf{X}_i'}\beta_d + \eta_i \\ Y_i = \alpha + \tau D_i + \boxed{\mathbf{X}_i'}\beta_y + \varepsilon_i \end{array} \right)$$

Ch4. MHE

τ_{LATE}

- Recall the four canonical IV assumptions.
- Suppose we have binary treatment and binary instrument.
  - w/o covariates, TSLS estimand (τ) = LATE (τ_{LATE}) and TSLS estimator = Wald estimator
  - w/ covariates, we need $\boxed{\textbf{constant effects}}$ so that TSLS estimand (τ) = LATE (τ_{LATE})
    - Otherwise, τ is an odd weighted function of causal effects $\tau \neq \tau_{LATE}$ QoI

In R: TSLS estimator w/ exogenous covariates

(~ FE by X3

+ factor(X3) + factor(X3)

```
ivmodel <- AER::ivreg(Y ~ X1 + X2 + D | X1 + X2 + Z, data = my_data)
# Or, equivalently: Y ~ X1 + X2 + D | . -D + Z
ivpack::robust.se(ivmodel)
# For clustered error: use ivpack::cluster.robust.se()
```