Interference

Ryan T. Moore

American University

2024-08-23

Interference

Interference

Interference

Rubin (1978)

SUTVA 2: No interference between units.

Does A's assignment affect B's potential outcomes?

Interference

- Describe violation of SUTVA in social pressure experiment.

Interference

- Describe violation of SUTVA in social pressure experiment.
- Does assg of MTurk survey respondent A to conjoint profile X affect response of MTurker B?

Interference

- Describe violation of SUTVA in social pressure experiment.
- Does assg of MTurk survey respondent A to conjoint profile X affect response of MTurker B?
- Does assg of election monitors to polling place A affect fraud at B?

Interference

- Describe violation of SUTVA in social pressure experiment.
- Does assg of MTurk survey respondent A to conjoint profile X affect response of MTurker B?
- Does assg of election monitors to polling place A affect fraud at B?
- Does assg of child A to vaccine affect health of B?

Interference

- Describe violation of SUTVA in social pressure experiment.
- Does assg of MTurk survey respondent A to conjoint profile X affect response of MTurker B?
- Does assg of election monitors to polling place A affect fraud at B?
- Does assg of child A to vaccine affect health of B?
- Does assg of A to GOTV door knock affect turnout of neighbor?

Interference

- Describe violation of SUTVA in social pressure experiment.
- Does assg of MTurk survey respondent A to conjoint profile X affect response of MTurker B?
- Does assg of election monitors to polling place A affect fraud at B?
- Does assg of child A to vaccine affect health of B?
- Does assg of A to GOTV door knock affect turnout of neighbor?
- Does assg of develpmnt £ to village A affect trust in B?

Interference

- Describe violation of SUTVA in social pressure experiment.
- Does assg of MTurk survey respondent A to conjoint profile X affect response of MTurker B?
- Does assg of election monitors to polling place A affect fraud at B?
- Does assg of child A to vaccine affect health of B?
- Does assg of A to GOTV door knock affect turnout of neighbor?
- Does assg of develpmnt £ to village A affect trust in B?
- Other common examples?

Examples of Interference in Field Experiments

Gerber and Green (2012)

- Contagion

Examples of Interference in Field Experiments

Gerber and Green (2012)

- Contagion
- Displacement

Examples of Interference in Field Experiments

Gerber and Green (2012)

- Contagion
- Displacement
- Communication

Examples of Interference in Field Experiments

Gerber and Green (2012)

- Contagion
- Displacement
- Communication
- Knowledge of others (disappointment/pride)

Examples of Interference in Field Experiments

Gerber and Green (2012)

- Contagion
- Displacement
- Communication
- Knowledge of others (disappointment/pride)
- Deterrence

Examples of Interference in Field Experiments

Gerber and Green (2012)

- Contagion
- Displacement
- Communication
- Knowledge of others (disappointment/pride)
- Deterrence
- Persistence

Examples of Interference in Field Experiments

Gerber and Green (2012)

- Contagion
- Displacement
- Communication
- Knowledge of others (disappointment/pride)
- Deterrence
- Persistence
- Memory

Bias under Interference

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

Bias under Interference

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- $Y_i(0)$ *unstable*, since depends on who assigned Tr

Bias under Interference

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- $Y_i(0)$ *unstable*, since depends on who assigned Tr
- $\overline{Y}_{\text{None}} = 70$.

Bias under Interference

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- $Y_i(0)$ *unstable*, since depends on who assigned Tr
- $\overline{Y}_{\text{None}} = 70$.
- Calc TEs relative to $\overline{Y}_{\text{None}}$ ("uniformity")

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:
  - If Mary, $\frac{100+50+90}{3} - 70 = 80 - 70 = 10$

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:
  - If Mary, $\frac{100+50+90}{3} - 70 = 80 - 70 = 10$
  - If Peter, $\frac{50+50+50}{3} - 70 = 50 - 70 = -20$

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:
  - If Mary, $\frac{100+50+90}{3} - 70 = 80 - 70 = 10$
  - If Peter, $\frac{50+50+50}{3} - 70 = 50 - 70 = -20$
  - If Limor, $\frac{70+50+90}{3} - 70 = 70 - 70 = 0$

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:
  - If Mary, $\frac{100+50+90}{3} - 70 = 80 - 70 = 10$
  - If Peter, $\frac{50+50+50}{3} - 70 = 50 - 70 = -20$
  - If Limor, $\frac{70+50+90}{3} - 70 = 70 - 70 = 0$
- So, $E(\text{Diff in Means}) = \frac{10-20+0}{3} = -\frac{10}{3}$

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:
  - If Mary, $\frac{100+50+90}{3} - 70 = 80 - 70 = 10$
  - If Peter, $\frac{50+50+50}{3} - 70 = 50 - 70 = -20$
  - If Limor, $\frac{70+50+90}{3} - 70 = 70 - 70 = 0$
- So, $E(\text{Diff in Means}) = \frac{10-20+0}{3} = -\frac{10}{3}$
- So, $E(\text{Diff in Means}) \neq ATE$

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:
  - If Mary, $\frac{100+50+90}{3} - 70 = 80 - 70 = 10$
  - If Peter, $\frac{50+50+50}{3} - 70 = 50 - 70 = -20$
  - If Limor, $\frac{70+50+90}{3} - 70 = 70 - 70 = 0$
- So, $E(\text{Diff in Means}) = \frac{10-20+0}{3} = -\frac{10}{3}$
- So, $E(\text{Diff in Means}) \neq ATE$
- Diff in means is *biased* for uniformity ATE!

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:
  - If Mary, $\frac{100+50+90}{3} - 70 = 80 - 70 = 10$
  - If Peter, $\frac{50+50+50}{3} - 70 = 50 - 70 = -20$
  - If Limor, $\frac{70+50+90}{3} - 70 = 70 - 70 = 0$
- So, $E(\text{Diff in Means}) = \frac{10-20+0}{3} = -\frac{10}{3}$
- So, $E(\text{Diff in Means}) \neq ATE$
- Diff in means is *biased* for uniformity ATE!

Bias under Interference: Uniformity Trial

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- True uniformity ATE: $\frac{(100-70)+(50-50)+(90-90)}{3} = 10$
- But, calculate all TEs relative to uniformity:
  - If Mary, $\frac{100+50+90}{3} - 70 = 80 - 70 = 10$
  - If Peter, $\frac{50+50+50}{3} - 70 = 50 - 70 = -20$
  - If Limor, $\frac{70+50+90}{3} - 70 = 70 - 70 = 0$
- So, $E(\text{Diff in Means}) = \frac{10-20+0}{3} = -\frac{10}{3}$
- So, $E(\text{Diff in Means}) \neq ATE$
- Diff in means is *biased* for uniformity ATE! ☹

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?
- Empirically, if randomly select winner, calc diff-in-means,

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?
- Empirically, if randomly select winner, calc diff-in-means,
  - If Mary, $100 - \frac{50+90}{2} = 100 - 70 = 30$

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?
- Empirically, if randomly select winner, calc diff-in-means,
  - If Mary, $100 - \frac{50+90}{2} = 100 - 70 = 30$
  - If Peter, $50 - \frac{50+50}{2} = 50 - 50 = 0$

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?
- Empirically, if randomly select winner, calc diff-in-means,
  - If Mary, $100 - \frac{50+90}{2} = 100 - 70 = 30$
  - If Peter, $50 - \frac{50+50}{2} = 50 - 50 = 0$
  - If Limor, $90 - \frac{70+50}{2} = 90 - 60 = 30$

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?
- Empirically, if randomly select winner, calc diff-in-means,
  - If Mary, $100 - \frac{50+90}{2} = 100 - 70 = 30$
  - If Peter, $50 - \frac{50+50}{2} = 50 - 50 = 0$
  - If Limor, $90 - \frac{70+50}{2} = 90 - 60 = 30$
- $E(\widehat{\text{Diff in Means}}) = 20$

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?
- Empirically, if randomly select winner, calc diff-in-means,
  - If Mary, $100 - \frac{50+90}{2} = 100 - 70 = 30$
  - If Peter, $50 - \frac{50+50}{2} = 50 - 50 = 0$
  - If Limor, $90 - \frac{70+50}{2} = 90 - 60 = 30$
- $E(\widehat{\text{Diff in Means}}) = 20$

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?
- Empirically, if randomly select winner, calc diff-in-means,
  - If Mary, $100 - \frac{50+90}{2} = 100 - 70 = 30$
  - If Peter, $50 - \frac{50+50}{2} = 50 - 50 = 0$
  - If Limor, $90 - \frac{70+50}{2} = 90 - 60 = 30$
- $E(\widehat{\text{Diff in Means}}) = 20$

$E(\widehat{ATE}) \neq ATE$, so, estimator *biased* under randomization.

Bias under Interference: Difference-in-Means

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What about empirical difference in means, $\widehat{\text{Diff in Means}}$?
- Empirically, if randomly select winner, calc diff-in-means,
  - If Mary, $100 - \frac{50+90}{2} = 100 - 70 = 30$
  - If Peter, $50 - \frac{50+50}{2} = 50 - 50 = 0$
  - If Limor, $90 - \frac{70+50}{2} = 90 - 60 = 30$
- $E(\widehat{\text{Diff in Means}}) = 20$

$E(\widehat{ATE}) \neq ATE$, so, estimator *biased* under randomization. ☹

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use “Limor honored” as reference “control”?

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use "Limor honored" as reference "control"?
- True "Limor" ATE: $\frac{(100-70)+(50-50)}{2} = 15$

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use "Limor honored" as reference "control"?
- True "Limor" ATE: $\frac{(100-70)+(50-50)}{2} = 15$
- Calculate diff-in-means TE's relative to "Limor honored"

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use "Limor honored" as reference "control"?
- True "Limor" ATE: $\frac{(100-70)+(50-50)}{2} = 15$
- Calculate diff-in-means TE's relative to "Limor honored"
  - If Mary, $80 - 70 = 10$

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use "Limor honored" as reference "control"?
- True "Limor" ATE: $\frac{(100-70)+(50-50)}{2} = 15$
- Calculate diff-in-means TE's relative to "Limor honored"
  - If Mary, $80 - 70 = 10$
  - If Peter, $50 - 70 = -20$

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use "Limor honored" as reference "control"?
- True "Limor" ATE: $\frac{(100-70)+(50-50)}{2} = 15$
- Calculate diff-in-means TE's relative to "Limor honored"
  - If Mary, $80 - 70 = 10$
  - If Peter, $50 - 70 = -20$
  - $E(\text{Diff in Means}) = \frac{10-20}{2} = -5$

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use "Limor honored" as reference "control"?
- True "Limor" ATE: $\frac{(100-70)+(50-50)}{2} = 15$
- Calculate diff-in-means TE's relative to "Limor honored"
  - If Mary, $80 - 70 = 10$
  - If Peter, $50 - 70 = -20$
  - $E(\text{Diff in Means}) = \frac{10-20}{2} = -5$
- And, $E(\text{Diff}\ \widehat{\text{in Means}}) = 20$

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use “Limor honored” as reference “control”?
- True “Limor” ATE: $\frac{(100-70)+(50-50)}{2} = 15$
- Calculate diff-in-means TE’s relative to “Limor honored”
  - If Mary, $80 - 70 = 10$
  - If Peter, $50 - 70 = -20$
  - $E(\text{Diff in Means}) = \frac{10-20}{2} = -5$
- And, $E(\widehat{\text{Diff in Means}}) = 20$

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use "Limor honored" as reference "control"?
- True "Limor" ATE: $\frac{(100-70)+(50-50)}{2} = 15$
- Calculate diff-in-means TE's relative to "Limor honored"
  - If Mary, $80 - 70 = 10$
  - If Peter, $50 - 70 = -20$
  - $E(\text{Diff in Means}) = \frac{10-20}{2} = -5$
- And, $E(\widehat{\text{Diff in Means}}) = 20$

With interference, diff-in-means under randomization biased.

Bias under Interference: Reference Control

Agent	Y if MaryH	Y if PeterH	Y if LimorH	Y if NobodyH
Mary	100	50	70	70
Peter	50	50	50	50
Limor	90	50	90	90
Mean	80	50	70	70

- What if use "Limor honored" as reference "control"?
- True "Limor" ATE: $\frac{(100-70)+(50-50)}{2} = 15$
- Calculate diff-in-means TE's relative to "Limor honored"
  - If Mary, $80 - 70 = 10$
  - If Peter, $50 - 70 = -20$
  - $E(\text{Diff in Means}) = \frac{10-20}{2} = -5$
- And, $E(\widehat{\text{Diff in Means}}) = 20$

With interference, diff-in-means under randomization biased. ☹

Notation

$$Y_i(z_i, \mathbf{z}) \quad \overset{?}{=} \quad Y_i(z_i, \mathbf{z}')$$

Notation

$$
\begin{aligned}
Y_i(z_i, \mathbf{z}) &\stackrel{?}{=} Y_i(z_i, \mathbf{z}') \\
Y_1\,(1, (1, 0, 0, 1, \ldots)) &\stackrel{?}{=} Y_1\,(1, (1, 1, 0, 0, \ldots)) \\
Y_0\,(0, (0, 0, 0, 1, \ldots)) &\stackrel{?}{=} Y_0\,(0, (0, 1, 0, 0, \ldots))
\end{aligned}
$$

Estimation in Presence of Interference

- Let T_i' = whether my friend is treated

Estimation in Presence of Interference

- Let T_i' = whether my friend is treated
- T_i = whether I (unit i) am treated

Estimation in Presence of Interference

- Let $T_i' =$ whether my friend is treated
- $T_i =$ whether I (unit i) am treated
- Then, $Y_{T_i' T_i}$ is my potential outcome using both assignments

Estimation in Presence of Interference

- Let T_i' = whether my friend is treated
- T_i = whether I (unit i) am treated
- Then, $Y_{T_i' T_i}$ is my potential outcome using both assignments

Estimation in Presence of Interference

- Let T_i' = whether my friend is treated
- T_i = whether I (unit i) am treated
- Then, $Y_{T_i' T_i}$ is my potential outcome using both assignments

Under balanced random assignment, **no** interference, naive diff in means has expectation

Estimation in Presence of Interference

- Let T_i' = whether my friend is treated
- T_i = whether I (unit i) am treated
- Then, $Y_{T_i' T_i}$ is my potential outcome using both assignments

Under balanced random assignment, **no** interference, naive diff in means has expectation

$$\frac{\bar{Y}_{11} + \bar{Y}_{01}}{2} - \frac{\bar{Y}_{10} + \bar{Y}_{00}}{2} = \bar{Y}_{01} - \bar{Y}_{00}$$

Estimation in Presence of Interference

- Let T_i' = whether my friend is treated
- T_i = whether I (unit i) am treated
- Then, $Y_{T_i' T_i}$ is my potential outcome using both assignments

Under balanced random assignment, **no** interference, naive diff in means has expectation

$$\frac{\bar{Y}_{11} + \bar{Y}_{01}}{2} - \frac{\bar{Y}_{10} + \bar{Y}_{00}}{2} = \bar{Y}_{01} - \bar{Y}_{00}$$

which is *unbiased* for ATE.

Estimation in Presence of Interference

- But, if

$$\underbrace{\bar{Y}_{11} - \bar{Y}_{01}}_{\text{spillover eff on Tr}} > \underbrace{\bar{Y}_{10} - \bar{Y}_{00}}_{\text{spillover eff on Co}}$$

then tend to overest $\bar{Y}_{01} - \bar{Y}_{00}$

Estimation in Presence of Interference

- But, if
$$\underbrace{\bar{Y}_{11} - \bar{Y}_{01}}_{\text{spillover eff on Tr}} \quad > \quad \underbrace{\bar{Y}_{10} - \bar{Y}_{00}}_{\text{spillover eff on Co}}$$
then tend to overest $\bar{Y}_{01} - \bar{Y}_{00}$
- If
$$\underbrace{\bar{Y}_{11} - \bar{Y}_{01}}_{\text{spillover eff on Tr}} \quad < \quad \underbrace{\bar{Y}_{10} - \bar{Y}_{00}}_{\text{spillover eff on Co}}$$
then tend to underestimate $\bar{Y}_{01} - \bar{Y}_{00}$.

Design for Estimating Interference

- Multilevel experiments can help estimate these effects

Design for Estimating Interference

- Multilevel experiments can help estimate these effects
  - Given individuals within communities

Design for Estimating Interference

- Multilevel experiments can help estimate these effects
  - Given individuals within communities
  - Randomly assign whether community gets 20% or 80% treatment

Design for Estimating Interference

- Multilevel experiments can help estimate these effects
  - Given individuals within communities
  - Randomly assign whether community gets 20% or 80% treatment
  - Randomly assign treatment to some individuals in communities

Design for Estimating Interference

- Multilevel experiments can help estimate these effects
  - Given individuals within communities
  - Randomly assign whether community gets 20% or 80% treatment
  - Randomly assign treatment to some individuals in communities
- "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments", Sinclair, McConnell, and Green (2012)

Design for Estimating Interference

OES project with IRS (US tax authority):

Design for Estimating Interference

OES project with IRS (US tax authority):

- assign preparers to "some client letters", "no client letters"

Design for Estimating Interference

OES project with IRS (US tax authority):

- assign preparers to “some client letters”, “no client letters”
- assign clients in “some client letters” preparers to letters/no letters

Diagnosis of Potential Interference

1. Block:

```r
data(x100)
block.out <- block(data = x100, groups = "g",
                   id.vars = c("id"),
                   block.vars = c("b1", "b2"))
```

2. Assign:

```r
assg.out <- assignment(block.out, seed = 157)
```

Diagnosis of Potential Interference

Diagnose interference after assgnmnt (1D, Linday et al. (2001))

3. Diagnose:

```
diagnose(assg.out, data = x100, id.vars = "id",
    suspect.var = "b1", suspect.range = c(0, 5))
```

```
##
## Units differing by at least 0 and no
## more than 5 on b1:
##
## Group: a
##    Unit 1  Unit 2  Difference
## 1  1073    1098        0
## 2  1002    1036        1
## 3  1016    1060        1
## 4  1039    1076        3
##
## Group: b
```

Further Examination of Design

4. Get block IDs:

```
createBlockIDs(assg.out, data = x100, id.var = "id")
```

```
##   [1] 29 17 14  5 17 33 35 10 21 41 39 45 32 49 36 12 18
##  [26]  6 37 31  4 11 20 16 47 28 48 12 23 18  2 19 48 14
##  [51] 32 11 40 15 29  8 23  1  9 13  3 24 26 28  3 50  5
##  [76]  2 25 25 26 43 16 46 35  1 44 45 50 37  7 30 10 38
```

Further Examination of Design

5. Get balance:

```
assg2xBalance(assg.out, x100, id.var = "id",
    bal.vars = c("b1", "b2"))
```

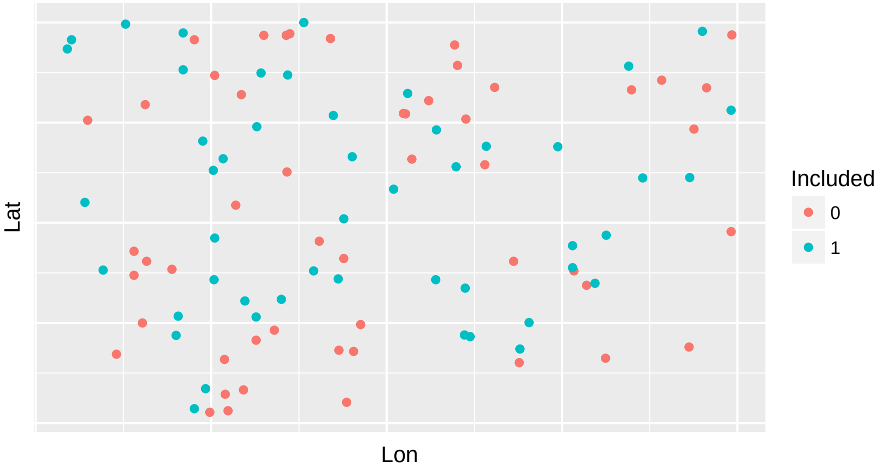
```
## $Group1
##      strata():   unstrat
##      stat      Treatment Control adj.diff std.diff    z
## vars
## b1                 -23.7     0.0    -23.7    -0.08   -0.
## b2                 28.6     0.0     28.6     0.10    0.
## ---Overall Test---
##         chisquare df p.value
## unstrat     0.161  2   0.922
## ---
## Signif. codes:  0 '***' 0.001 '** ' 0.01 '*  ' 0.05 '.
##
## $Group2
##      strata():   unstrat
```

Avoiding Potential Interference due to Proximity

Are units too near each other?

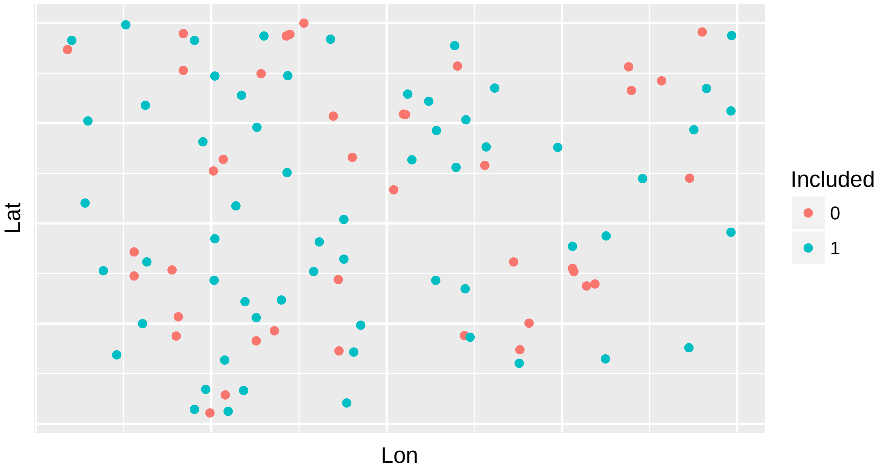
Avoiding Potential Interference due to Proximity

Are units too near each other?



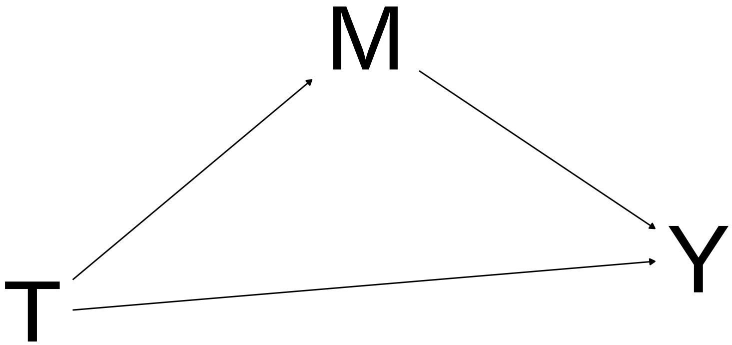
Avoiding Potential Interference due to Proximity

3000 iterations, max min distance:



Estimating Direct and Indirect Effects

- Different to mediation *direct*, *indirect* effects



Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

- Here,

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

- Here,
  - *direct*: from treating i

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

- Here,
  - *direct*: from treating i
  - *indirect*: from treating $j \neq i$

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

- Here,
  - *direct*: from treating i
  - *indirect*: from treating $j \neq i$

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

- Here,
  - *direct*: from treating i
  - *indirect*: from treating $j \neq i$

Think of this as different problem

- (Though, "effect of treatment through others" works?)

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

- Here,
  - *direct*: from treating i
  - *indirect*: from treating $j \neq i$

Think of this as different problem

- (Though, “effect of treatment through others” works?)
- (Common concept, but not really a mediating *variable*)

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

Table 1. Risk of cholera in recipients of killed oral cholera vaccines or placebo, by level of coverage of the bari during one year of follow-up, based on data from Ali et al. (2005)

Level of vaccine coverage	Target population	Vaccine recipients: Total	Vaccine recipients: Cases	Vaccine recipients: Risk per 1,000 population	Placebo recipients: Total	Placebo recipients: Cases	Placebo recipients: Risk per 1,000 population
>50%	22,394	12,541	16	1.27	6,082	9	1.47
41–50%	24,159	11,513	26	2.26	5,801	27	4.65
36–40%	24,583	10,772	17	1.58	5,503	26	4.72
28–35%	25,059	8,883	22	2.48	4,429	26	5.87
<28%	24,954	5,627	15	2.66	2,852	20	7.01

- Consider $> 50\%$ as A, $< 28\%$ as B

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

Table 1. Risk of cholera in recipients of killed oral cholera vaccines or placebo, by level of coverage of the bari during one year of follow-up, based on data from Ali et al. (2005)

Level of vaccine coverage	Target population	Vaccine recipients			Placebo recipients		
		Total	Cases	Risk per 1,000 population	Total	Cases	Risk per 1,000 population
>50%	22,394	12,541	16	1.27	6,082	9	1.47
41–50%	24,159	11,513	26	2.26	5,801	27	4.65
36–40%	24,583	10,772	17	1.58	5,503	26	4.72
28–35%	25,059	8,883	22	2.48	4,429	26	5.87
<28%	24,954	5,627	15	2.66	2,852	20	7.01

- Consider $> 50\%$ as A, $< 28\%$ as B
- Direct effect in B: $7.01 - 2.66 = 4.35$

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

Table 1. Risk of cholera in recipients of killed oral cholera vaccines or placebo, by level of coverage of the bari during one year of follow-up, based on data from Ali et al. (2005)

Level of vaccine coverage	Target population	Vaccine recipients: Total	Vaccine recipients: Cases	Vaccine recipients: Risk per 1,000 population	Placebo recipients: Total	Placebo recipients: Cases	Placebo recipients: Risk per 1,000 population
>50%	22,394	12,541	16	1.27	6,082	9	1.47
41–50%	24,159	11,513	26	2.26	5,801	27	4.65
36–40%	24,583	10,772	17	1.58	5,503	26	4.72
28–35%	25,059	8,883	22	2.48	4,429	26	5.87
<28%	24,954	5,627	15	2.66	2,852	20	7.01

- Consider $> 50\%$ as A, $< 28\%$ as B
- Direct effect in B: $7.01 - 2.66 = 4.35$
- Direct effect in A: $1.47 - 1.27 = 0.2$

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

Table 1. Risk of cholera in recipients of killed oral cholera vaccines or placebo, by level of coverage of the bari during one year of follow-up, based on data from Ali et al. (2005)

Level of vaccine coverage	Target population	Vaccine recipients: Total	Cases	Risk per 1,000 population	Placebo recipients: Total	Cases	Risk per 1,000 population
>50%	22,394	12,541	16	1.27	6,082	9	1.47
41–50%	24,159	11,513	26	2.26	5,801	27	4.65
36–40%	24,583	10,772	17	1.58	5,503	26	4.72
28–35%	25,059	8,883	22	2.48	4,429	26	5.87
<28%	24,954	5,627	15	2.66	2,852	20	7.01

- Consider $> 50\%$ as A, $< 28\%$ as B
- Direct effect in B: $7.01 - 2.66 = 4.35$
- Direct effect in A: $1.47 - 1.27 = 0.2$
- Indirect effect on unvaccinated: $7.01 - 1.47 = 5.54$

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

Table 1. Risk of cholera in recipients of killed oral cholera vaccines or placebo, by level of coverage of the bari during one year of follow-up, based on data from Ali et al. (2005)

Level of vaccine coverage	Target population	Vaccine recipients			Placebo recipients		
		Total	Cases	Risk per 1,000 population	Total	Cases	Risk per 1,000 population
>50%	22,394	12,541	16	1.27	6,082	9	1.47
41–50%	24,159	11,513	26	2.26	5,801	27	4.65
36–40%	24,583	10,772	17	1.58	5,503	26	4.72
28–35%	25,059	8,883	22	2.48	4,429	26	5.87
<28%	24,954	5,627	15	2.66	2,852	20	7.01

- Consider $> 50\%$ as A, $< 28\%$ as B
- Direct effect in B: $7.01 - 2.66 = 4.35$
- Direct effect in A: $1.47 - 1.27 = 0.2$
- Indirect effect on unvaccinated: $7.01 - 1.47 = 5.54$
- Total effect: $7.01 - 1.27 = 5.74$

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

Table 1. Risk of cholera in recipients of killed oral cholera vaccines or placebo, by level of coverage of the bari during one year of follow-up, based on data from Ali et al. (2005)

Level of vaccine coverage	Target population	Vaccine recipients: Total	Vaccine recipients: Cases	Vaccine recipients: Risk per 1,000 population	Placebo recipients: Total	Placebo recipients: Cases	Placebo recipients: Risk per 1,000 population
>50%	22,394	12,541	16	1.27	6,082	9	1.47
41–50%	24,159	11,513	26	2.26	5,801	27	4.65
36–40%	24,583	10,772	17	1.58	5,503	26	4.72
28–35%	25,059	8,883	22	2.48	4,429	26	5.87
<28%	24,954	5,627	15	2.66	2,852	20	7.01

- Consider $> 50\%$ as A, $< 28\%$ as B
- Direct effect in B: $7.01 - 2.66 = 4.35$
- Direct effect in A: $1.47 - 1.27 = 0.2$
- Indirect effect on unvaccinated: $7.01 - 1.47 = 5.54$
- Total effect: $7.01 - 1.27 = 5.74$
- Overall effect: $35/8479 - 25/18623 = 2.79/1000$

Estimating Direct and Indirect Effects

Hudgens and Halloran (2008) (on Ali et al. (2005))

Designing the randomized experiment:

Table 2. Illustrative example of a two-stage randomized placebo-controlled vaccine trial based on data from Ali et al. (2005)

		Vaccine recipients ($Z_{ij} = 1$)		Placebo recipients ($Z_{ij} = 0$)	
Group i	Group assignment S_i	Total $\sum_j Z_{ij}$	Cases $\sum_j Z_{ij} Y_{ij}(\mathbf{Z}_i)$	Total $\sum_j (1 - Z_{ij})$	Cases $\sum_j (1 - Z_{ij}) Y_{ij}(\mathbf{Z}_i)$
1	1	12,541	16	12,541	18
2	1	11,513	26	11,513	54
3	0	10,772	17	25,134	119
4	0	8,883	22	20,727	122
5	0	5,627	15	13,130	92

NOTE: Group assignment $S_i = 1$ (0) corresponds to 50% (30%) vaccine coverage.

Hudgens and Halloran (2008) (on Ali 2005)

Table 3. Estimates of population average direct, indirect, total, and overall effects per 1,000 individuals per year for data in Table 2

Effect	Parameter	Estimate	Estimated variance
Direct	$\overline{CE}^D(\psi)$	1.30	.856
Direct	$\overline{CE}^D(\phi)$	3.64	.178
Indirect	$\overline{CE}^I(\phi, \psi)$	2.81	3.079
Total	$\overline{CE}^T(\phi, \psi)$	4.11	.672
Overall	$\overline{CE}^O(\phi, \psi)$	2.37	1.430

- ψ: 50% coverage
- ϕ: 30% coverage

Application: Displacing Election Irregularities

Ichino and Schündeln (2012) in Ghana

Design:

- Blocks of 3 constituencies, select 1 for Tr, 2 for Co

Application: Displacing Election Irregularities

Ichino and Schündeln (2012) in Ghana

Design:

- Blocks of 3 constituencies, select 1 for Tr, 2 for Co
- Treatment constituencies: send observers to 25% of areas

Application: Displacing Election Irregularities

Ichino and Schündeln (2012) in Ghana

Design:

- Blocks of 3 constituencies, select 1 for Tr, 2 for Co
- Treatment constituencies: send observers to 25% of areas
- So, areas can be

Application: Displacing Election Irregularities

Ichino and Schündeln (2012) in Ghana

Design:

- Blocks of 3 constituencies, select 1 for Tr, 2 for Co
- Treatment constituencies: send observers to 25% of areas
- So, areas can be
  - Tr in Tr constituency

Application: Displacing Election Irregularities

Ichino and Schündeln (2012) in Ghana

Design:

- Blocks of 3 constituencies, select 1 for Tr, 2 for Co
- Treatment constituencies: send observers to 25% of areas
- So, areas can be
  - Tr in Tr constituency
  - Co in Tr constituency

Application: Displacing Election Irregularities

Ichino and Schündeln (2012) in Ghana

Design:

- Blocks of 3 constituencies, select 1 for Tr, 2 for Co
- Treatment constituencies: send observers to 25% of areas
- So, areas can be
  - Tr in Tr constituency
  - Co in Tr constituency
  - Co in Co constituency

Application: Displacing Election Irregularities

Ichino and Schündeln (2012) in Ghana

Design:

- Blocks of 3 constituencies, select 1 for Tr, 2 for Co
- Treatment constituencies: send observers to 25% of areas
- So, areas can be
  - Tr in Tr constituency
  - Co in Tr constituency
  - Co in Co constituency

- Where election observers are sent, smaller registration irregularities

Application: Displacing Election Irregularities

Ichino and Schündeln (2012) in Ghana

Design:

- Blocks of 3 constituencies, select 1 for Tr, 2 for Co
- Treatment constituencies: send observers to 25% of areas
- So, areas can be
  - Tr in Tr constituency
  - Co in Tr constituency
  - Co in Co constituency

- Where election observers are sent, smaller registration irregularities
- In nearby control areas, *larger* irregularities

Sobel (2006)

If interference, diff-in-means estimator (or regression coef)

- is **not** unbiased for ATE
- is difference:

(ITT for Tr group) – (indirect/spillover effect on Co group)

Rosenbaum (2007)

- Randomization inference gives valid coverage, even if interference

Null Hypotheses:

Rosenbaum (2007)

- Randomization inference gives valid coverage, even if interference

Null Hypotheses:

- No primary effect: $H_0 : Y_{bi\mathbf{z}} = Y_{bi\mathbf{z}'}$

Rosenbaum (2007)

- Randomization inference gives valid coverage, even if interference

Null Hypotheses:

- No primary effect: $H_0 : Y_{bi\mathbf{z}} = Y_{bi\mathbf{z}'}$
  - no effect of treatment itself on i

Rosenbaum (2007)

- Randomization inference gives valid coverage, even if interference

Null Hypotheses:

- No primary effect: $H_0 : Y_{bi\mathbf{z}} = Y_{bi\mathbf{z}'}$
  - no effect of treatment itself on i
- No effect: $H_0 : Y_{bi\mathbf{z}} = \widetilde{Y}_{bi}$

Rosenbaum (2007)

- Randomization inference gives valid coverage, even if interference

Null Hypotheses:

- No primary effect: $H_0 : Y_{bi\mathbf{z}} = Y_{bi\mathbf{z}'}$
  - no effect of treatment itself on i
- No effect: $H_0 : Y_{bi\mathbf{z}} = \widetilde{Y}_{bi}$
  - Outcome for i is same as if assigned Tr, but Tr withheld from *everyone* ("uniformity")

Rosenbaum (2007)

- Randomization inference gives valid coverage, even if interference

Null Hypotheses:

- No primary effect: $H_0 : Y_{bi\mathbf{z}} = Y_{bi\mathbf{z}'}$
  - no effect of treatment itself on i
- No effect: $H_0 : Y_{bi\mathbf{z}} = \widetilde{Y}_{bi}$
  - Outcome for i is same as if assigned Tr, but Tr withheld from *everyone* (“uniformity”)

Rosenbaum (2007)

- Randomization inference gives valid coverage, even if interference

Null Hypotheses:

- No primary effect: $H_0 : Y_{bi\mathbf{z}} = Y_{bi\mathbf{z}'}$
  - no effect of treatment itself on i
- No effect: $H_0 : Y_{bi\mathbf{z}} = \widetilde{Y}_{bi}$
  - Outcome for i is same as if assigned Tr, but Tr withheld from *everyone* ("uniformity")

Effect, but no Primary Effect:

- if exposing block gets everyone sick, then no primary effect of i getting directly exposed

Rosenbaum (2007)

- Randomization inference gives valid coverage, even if interference

Null Hypotheses:

- No primary effect: $H_0 : Y_{bi\mathbf{z}} = Y_{bi\mathbf{z}'}$
  - no effect of treatment itself on i
- No effect: $H_0 : Y_{bi\mathbf{z}} = \widetilde{Y}_{bi}$
  - Outcome for i is same as if assigned Tr, but Tr withheld from *everyone* ("uniformity")

Effect, but no Primary Effect:

- if exposing block gets everyone sick, then no primary effect of i getting directly exposed
- if news raises anxiety in HH, irrelevant if I saw news

Next: Causal Forests

References I

Ali, Mohammad, Michael Emch, Lorenz Von Seidlein, Mohammad Yunus, David A. Sack, Malla Rao, Jan Holmgren, and John D. Clemens. 2005. "Herd Immunity Conferred by Killed Oral Cholera Vaccines in Bangladesh: A Reanalysis." *The Lancet* 366 (9479): 44–49.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation.* New York, NY: WW Norton.

Hudgens, Michael G., and M. Elizabeth Halloran. 2008. "Toward Causal Inference With Interference." *Journal of the American Statistical Association* 103 (482): 832–42.

Ichino, Nahomi, and Matthias Schündeln. 2012. "Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana." *The Journal of Politics* 74 (1): 292–307.

Linday, Linda A., J. A. Tsiouris, Ira L. Cohen, Richard Shindledecker, and Robert DeCresce. 2001. "Famotidine Treatment of Children with Autistic Spectrum Disorders: Pilot Research Using Single Subject Research Design." *Journal of Neural Transmission* 108: 593–611.

Rosenbaum, Paul R. 2007. "Interference Between Units in Randomized Experiments." *Journal of the American Statistical Association* 102 (477): 191–200.

References II

Rubin, Donald B. 1978. "Multiple Imputations in Sample Surveys – A Phenomenological Bayesian Approach to Nonresponse."

Sinclair, Betsy, Margaret McConnell, and Donald P. Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments." *American Journal of Political Science* 56 (4): 1055–69.

Sobel, Michael E. 2006. "What Do Randomized Studies of Housing Mobility Demonstrate?: Causal Inference in the Face of Interference." *Journal of the American Statistical Association* 101 (476): 1398–1407.