“Challenges to Randomization: Noncompliance and Missing Data”

ICPSR 2022 Session 2

Jake Bowers & Tom Leavitt

July 26, 2022

Today

1. Agenda: One step away from easy to interpret experiments: non-random doses/compliance (Gerber and Green, 2012) Chapter 5, non-random missing data (Gerber and Green, 2012) Chapter 7.
2. Recap: We use statistics to **infer** about unobservable counterfactual quantities (functions of potential outcomes); we can estimate unobservable averages; we can test unobservable hypotheses; we can test unobservable hypotheses about averages.
3. Questions arising from the reading or assignments or life?

Today

1. Agenda: One step away from easy to interpret experiments: non-random doses/compliance (Gerber and Green, 2012) Chapter 5, non-random missing data (Gerber and Green, 2012) Chapter 7.
2. Recap: We use statistics to **infer** about unobservable counterfactual quantities (functions of potential outcomes); we can estimate unobservable averages; we can test unobservable hypotheses; we can test unobservable hypotheses about averages.
3. Questions arising from the reading or assignments or life?

Today

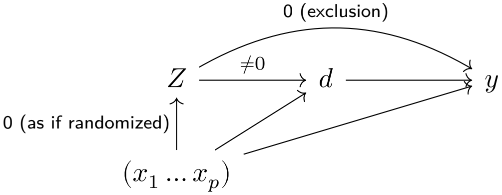
1. Agenda: One step away from easy to interpret experiments: non-random doses/compliance (Gerber and Green, 2012) Chapter 5, non-random missing data (Gerber and Green, 2012) Chapter 7.
2. Recap: We use statistics to **infer** about unobservable counterfactual quantities (functions of potential outcomes); we can estimate unobservable averages; we can test unobservable hypotheses; we can test unobservable hypotheses about averages.
3. Questions arising from the reading or assignments or life?

1. Causal effects when we do not control the dose
2. Hypothesis Tests about Complier causal effects
3. Learning about causal effects when data are missing

Defining causal effects I

Imagine a door-to-door communication experiment where some houses are randomly assigned to receive a visit. Note that we now use Z and d instead of T.

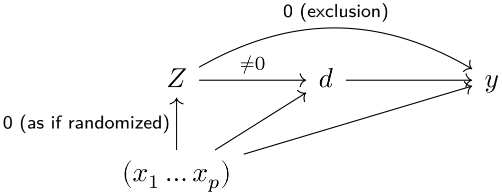
- Z_i is random assignment to a visit ($Z_i = 1$) or not ($Z_i = 0$).
- $d_{i,Z_i=1} = 1$ means that person i would open the door to have a conversation when assigned a visit.
- $d_{i,Z_i=1} = 0$ means that person i would not open the door to have a conversation when assigned a visit.
- Opening the door is an outcome of the treatment.



Defining causal effects I

Imagine a door-to-door communication experiment where some houses are randomly assigned to receive a visit. Note that we now use Z and d instead of T.

- Z_i is random assignment to a visit ($Z_i = 1$) or not ($Z_i = 0$).
- $d_{i,Z_i=1} = 1$ means that person i would open the door to have a conversation when assigned a visit.
- $d_{i,Z_i=1} = 0$ means that person i would not open the door to have a conversation when assigned a visit.
- Opening the door is an outcome of the treatment.



Defining causal effects I

Imagine a door-to-door communication experiment where some houses are randomly assigned to receive a visit. Note that we now use Z and d instead of T.

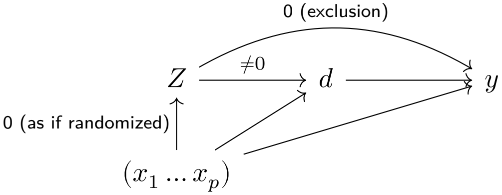
- Z_i is random assignment to a visit ($Z_i = 1$) or not ($Z_i = 0$).
- $d_{i,Z_i=1} = 1$ means that person i would open the door to have a conversation when assigned a visit.
- $d_{i,Z_i=1} = 0$ means that person i would not open the door to have a conversation when assigned a visit.
- Opening the door is an outcome of the treatment.

$$
\begin{tikzcd}
Z \arrow[r, "\neq 0"] \arrow[rr, bend left, "0 \text{ (exclusion)}"] & d \arrow[r] & y \\
(x_1 \dots x_p) \arrow[u, "0 \text{ (as if randomized)}"] \arrow[ur] \arrow[urr] & &
\end{tikzcd}
$$

Defining causal effects I

Imagine a door-to-door communication experiment where some houses are randomly assigned to receive a visit. Note that we now use Z and d instead of T.

- Z_i is random assignment to a visit ($Z_i = 1$) or not ($Z_i = 0$).
- $d_{i,Z_i=1} = 1$ means that person i would open the door to have a conversation when assigned a visit.
- $d_{i,Z_i=1} = 0$ means that person i would not open the door to have a conversation when assigned a visit.
- Opening the door is an outcome of the treatment.



Defining causal effects IV

We can simplify the ways in which people get a dose of the treatment like so (where d is lower case reflecting the idea that whether you open the door when visited or not is a fixed attribute like a potential outcome).

- Y : outcome ($y_{i,Z}$ or $y_{i,Z_i=1}$ for potential outcome to treatment for person i, fixed)
- X : covariate/baseline variable
- Z : treatment assignment ($Z_i = 1$ if assigned to a visit, $Z_i = 0$ if not assigned to a visit)
- D : treatment received ($D_i = 1$ if answered door, $D_i = 0$ if person i did not answer the door) (using D here because $D_i = d_{i,1}Z_i + d_{i,0}(1 - Z_i)$)

Defining causal effects IV

We can simplify the ways in which people get a dose of the treatment like so (where d is lower case reflecting the idea that whether you open the door when visited or not is a fixed attribute like a potential outcome).

- Y : outcome ($y_{i,Z}$ or $y_{i,Z_i=1}$ for potential outcome to treatment for person i, fixed)
- X : covariate/baseline variable
- Z : treatment assignment ($Z_i = 1$ if assigned to a visit, $Z_i = 0$ if not assigned to a visit)
- D : treatment received ($D_i = 1$ if answered door, $D_i = 0$ if person i did not answer the door) (using D here because $D_i = d_{i,1}Z_i + d_{i,0}(1 - Z_i)$)

Defining causal effects IV

We can simplify the ways in which people get a dose of the treatment like so (where d is lower case reflecting the idea that whether you open the door when visited or not is a fixed attribute like a potential outcome).

- Y : outcome ($y_{i,Z}$ or $y_{i,Z_i=1}$ for potential outcome to treatment for person i, fixed)
- X : covariate/baseline variable
- Z : treatment assignment ($Z_i = 1$ if assigned to a visit, $Z_i = 0$ if not assigned to a visit)
- D : treatment received ($D_i = 1$ if answered door, $D_i = 0$ if person i did not answer the door) (using D here because $D_i = d_{i,1}Z_i + d_{i,0}(1 - Z_i)$)

Defining causal effects IV

We can simplify the ways in which people get a dose of the treatment like so (where d is lower case reflecting the idea that whether you open the door when visited or not is a fixed attribute like a potential outcome).

- Y : outcome ($y_{i,Z}$ or $y_{i,Z_i=1}$ for potential outcome to treatment for person i, fixed)
- X : covariate/baseline variable
- Z : treatment assignment ($Z_i = 1$ if assigned to a visit, $Z_i = 0$ if not assigned to a visit)
- D : treatment received ($D_i = 1$ if answered door, $D_i = 0$ if person i did not answer the door) (using D here because $D_i = d_{i,1}Z_i + d_{i,0}(1 - Z_i)$)

Defining causal effects V

We have two causal effects of Z: $Z \to Y$ (known as δ, ITT, ITT_Y), and $Z \to D$ (known as ITT_D, p_c).
And different types of people can react differently to the attempt to move the dose with the instrument.

		$Z = 1$	
		$D = 0$	$D = 1$
$Z = 0$	$D = 0$	Never taker	Complier
	$D = 1$	Defier	Always taker

Defining causal effects VI

The $ITT = ITT_Y = \delta = \bar{y}_{Z=1} - \bar{y}_{Z=0}$.

But, in this design, $\bar{y}_{Z=1} = \bar{y}_1$ is split into pieces: the outcome of those who answered the door (Compliers and Always-takers and Defiers). Write p_C for the proportion of compliers in the study, p_A for proportion always-takers, etc... The proportions have to sum to 1. So, we have weighted averages:

$$\bar{y}_1 = (\bar{y}_1|C)p_C + (\bar{y}_1|A)p_A + (\bar{y}_1|N)p_N + (\bar{y}_1|D)p_D. \tag{1}$$

And $\bar{y}_0$ is also split into pieces:

$$\bar{y}_0 = (\bar{y}_0|C)p_C + (\bar{y}_1|A)p_A + (\bar{y}_0|N)p_N + (\bar{y}_0|D)p_D. \tag{2}$$

Defining causal effects VII

So, the ITT itself is a combination of the effects of Z on Y within these different groups. People who are compliers tend to be different types of people than people who are always takers: comparisons across types would raise questions about how to interpret the results — interpretations that would focus more on differences in types than in differences caused by Z. But, we can still estimate it because we have unbiased estimators of $\bar{y}_1$ and $\bar{y}_0$ within each type.

Learning about the ITT I

First, let's learn about the effect of the policy itself. To write down the ITT, we do not need to consider all of the types above. We have no defiers ($p_D = 0$) and we know the ITT for both Always-takers and Never-takers is 0.

$$\bar{y}_1 = (\bar{y}_1|C)p_C + (\bar{y}_1|A)p_A + (\bar{y}_1|N)p_N \tag{3}$$

$$\bar{y}_0 = (\bar{y}_0|C)p_C + (\bar{y}_0|A)p_A + (\bar{y}_0|N)p_N \tag{4}$$

Learning about the ITT II

First, let's learn about the effect of the policy itself. To write down the ITT, we do not need to consider all of the types above. We have no defiers ($p_D = 0$) and we know the ITT for both Always-takers and Never-takers is 0.

$$
\begin{aligned}
ITT =& \bar{y}_1 - \bar{y}_0 && (5)\\
=& ((\bar{y}_1|C)p_C + (\bar{y}_1|A)p_A + (\bar{y}_1|N)p_N) - && (6)\\
& ((\bar{y}_0|C)p_C + (\bar{y}_0|A)p_A + (\bar{y}_0|N)p_N) && (7)
\end{aligned}
$$

collecting each type together — to have an ITT for each type

$$
\begin{aligned}
=& ((\bar{y}_1|C)p_C - (\bar{y}_0|C)p_C) + ((\bar{y}_1|A)p_A - (\bar{y}_0|A)p_A) + && (8)\\
& ((\bar{y}_1|N)p_N - (\bar{y}_0|N)p_N) && (9)\\
=& \left((\bar{y}_1|C) - (\bar{y}_0|C)\right) p_C + && (10)\\
& \left((\bar{y}_1|A) - (\bar{y}_0|A)\right) p_A + \left((\bar{y}_1|N) - (\bar{y}_0|N)\right) p_N && (11)
\end{aligned}
$$

Learning about the ITT III

$$
\begin{aligned}
ITT =& \bar{y}_1 - \bar{y}_0 && (12)\\
=& ((\bar{y}_1|C)p_C + (\bar{y}_1|A)p_A + (\bar{y}_1|N)p_N) - && (13)\\
& ((\bar{y}_0|C)p_C + (\bar{y}_0|A)p_A + (\bar{y}_0|N)p_N) && (14)\\
=& ((\bar{y}_1|C)p_C - (\bar{y}_0|C)p_C) + ((\bar{y}_1|A)p_A - (\bar{y}_0|A)p_A) + && (15)\\
& ((\bar{y}_1|N)p_N - (\bar{y}_0|N)p_N) && (16)\\
=& ((\bar{y}_1|C) - (\bar{y}_0|C))p_C + ((\bar{y}_1|A) - (\bar{y}_0|A))p_A + && (17)\\
& ((\bar{y}_1|N) - (\bar{y}_0|N))p_N && (18)\\
=& (\text{ITT among compliers})(\text{proportion of compliers}) + (\text{ITT among always takers}) && (19)
\end{aligned}
$$

Learning about the ITT IV

And, if the effect of the dose can only occur for those who open the door, and you can only open the door when assigned to do so then:

$$((\bar{y}_1|A) - (\bar{y}_0|A))p_A = 0 \text{ and } ((\bar{y}_1|N) - (\bar{y}_0|N))p_N = 0 \tag{20}$$

And

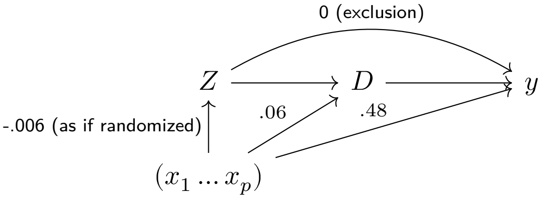
$$ITT = ((\bar{y}_1|C) - (\bar{y}_0|C))p_C = (CACE)p_C. \tag{21}$$

The complier average causal effect I

If we want to can learn about the the causal effect of answering the door and having the conversation why not just compare people who answer the door to people who do not?

The problem with this "as-treated" or "per-protocol" comparison is that this comparison is confounded by x: a simple $\bar{Y}|D = 1 - \bar{Y}|D = 0$ comparison tells us about differences in the outcome due to x in addition to the difference caused by D.

(Numbers below from some simulated data)



The complier average causal effect II

In actual data:

```
with(dat, cor(Y, x)) ## can be any number
with(dat, cor(d, x)) ## can be any number
with(dat, cor(Z, x)) ## should be near 0
```

And we just saw that, in this design, and with these assumptions (including a SUTVA assumption) that $ITT = ((\bar{y}_1|C) - (\bar{y}_0|C))p_C = (CACE)p_C$, so we can define $CACE = ITT/p_C$. That is, we can learn about the effect of answering the door without worrying about the bias from x (or any set of x's).

VERY COOL You can learn about the causal effect of a non-random intervention (deciding to open the door) without "controlling for" $x_1, x_2, ...$ in this case.

How to calculate the ITT and CACE/LATE I

Some example data (where we know all potential outcomes):

	ID	X	u	type	D_Z_0	D_Z_1	Y_D_0	Y_D_1	Y_Z_0	Y_Z_1	Z	D	Y
1	084	3	0.7773	Complier	0	1	0.7773	1.30418	0.7773	1.0408	0	0	0.7773
2	088	1	0.6207	Always-Taker	1	1	0.6207	1.14752	0.6207	0.8841	0	1	1.1475
3	058	1	-1.5785	Never-Taker	0	0	-1.5785	-1.05167	-1.5785	-1.3151	1	0	-1.5785
4	056	1	-1.2829	Complier	0	1	-1.2829	-0.75607	-1.2829	-1.0195	1	1	-0.7561
5	079	3	0.7853	Complier	0	1	0.7853	1.31215	0.7853	1.0487	1	1	1.3122
6	037	2	1.3539	Complier	0	1	1.3539	1.88072	1.3539	1.6173	1	1	1.8807
7	005	2	0.1644	Complier	0	1	0.1644	0.69122	0.1644	0.4278	0	0	0.1644
8	069	3	-1.1382	Complier	0	1	-1.1382	-0.61132	-1.1382	-0.8747	0	0	-1.1382
9	015	3	1.0140	Complier	0	1	1.0140	1.54087	1.0140	1.2774	1	1	1.5409
10	073	2	-0.5509	Complier	0	1	-0.5509	-0.02401	-0.5509	-0.2874	0	0	-0.5509
11	040	1	0.8188	Always-Taker	1	1	0.8188	1.34569	0.8188	1.0823	1	1	1.3457
12	081	2	0.3009	Complier	0	1	0.3009	0.82771	0.3009	0.5643	1	1	0.8277
13	042	2	0.8455	Complier	0	1	0.8455	1.37239	0.8455	1.1090	1	1	1.3724
14	098	2	-1.1724	Complier	0	1	-1.1724	-0.64554	-1.1724	-0.9090	1	1	-0.6455
15	052	4	0.1371	Always-Taker	1	1	0.1371	0.66397	0.1371	0.4005	0	1	0.6640

How to calculate the ITT and CACE/LATE II

The ITT and CACE (the parts)

```
itt_y <- difference_in_means(Y ~ Z, data = dat0)
itt_y
```

```
Design:  Standard
  Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
Z   0.5672     0.2173   2.611  0.01048    0.136   0.9985 96
```

```
itt_d <- difference_in_means(D ~ Z, data = dat0)
itt_d
```

```
Design:  Standard
  Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
Z     0.72     0.0701   10.27 3.13e-17   0.5809   0.8591 98
```

How to calculate the ITT and CACE/LATE III

All together:[1]

```
cace_est <- iv_robust(Y ~ D | Z, data = dat0)
cace_est
```

```
            Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
(Intercept) 0.003383     0.1737 0.01947 0.984502  -0.3413   0.3481 98
D           0.787846     0.2919 2.69880 0.008196   0.2085   1.3672 98
```

```
## Notice same as below:
coef(itt_y)[["Z"]] / coef(itt_d)[["Z"]]
```

```
[1] 0.7878
```

[1] works when $Z \rightarrow D$ is not weak see Imbens and Rosenbaum (2005) for a cautionary tale

Variance of IV estimator

- Recall that there exist analytic expressions for $\mathrm{Var}\left[\widehat{\mathsf{ITT}}_Y\right]$ and $\mathrm{Var}\left[\widehat{\mathsf{ITT}}_D\right]$
- We can conservatively estimate $\mathrm{Var}\left[\widehat{\mathsf{ITT}}_Y\right]$ and $\mathrm{Var}\left[\widehat{\mathsf{ITT}}_D\right]$ via $\widehat{\mathrm{Var}}\left[\widehat{\mathsf{ITT}}_Y\right]$ and $\widehat{\mathrm{Var}}\left[\widehat{\mathsf{ITT}}_D\right]$
- However, in general, there is no closed-form analytic expression for the variance of a random ratio
- We do not have an estimator for $\mathrm{Var}\left[\dfrac{\widehat{\mathsf{ITT}}_Y}{\widehat{\mathsf{ITT}}_D}\right]$ that is known to be unbiased, consistent or conservative
- Bloom (1984) proposed treating $\widehat{\mathsf{ITT}}_D$ as fixed
- Others use Delta method (Taylor series approximation), e.g., in `AER` or `estimatr` package in `R`

How do our estimators perform?

```
  inquiry estimand
1    CACE   0.5225
2   ITT_y   0.2613
3   ITT_d   0.7700
         estimator term estimate std.error statistic   p.value conf.low conf.high df outcome inqu
1        iv_robust    D   0.9211   0.29289     3.145 2.201e-03   0.3398    1.5023 98       Y    C
2   diff means ITT    Z   0.6447   0.21854     2.950 3.971e-03   0.2111    1.0784 98       Y   IT
3 diff means ITT_D    Z   0.7000   0.07074     9.895 2.041e-16   0.5596    0.8404 98       D   IT
4     per-protocol    D   1.0013   0.19758     5.068 1.898e-06   0.6092    1.3933 98       Y   IT
5     per-protocol    D   1.0013   0.19758     5.068 1.898e-06   0.6092    1.3933 98       Y    C
Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced

Warning in sqrt(diag(vcov_fit$Vcov_hat)): NaNs produced
```

Summary of Encouragement/Complier/Dose oriented designs:

- Analyze as you randomized, even when you don't control the dose you can learn something.
- The danger of per-protocol analysis: you give up the benefits of the research design (i.e. randomization)
- Variance calculations approximate (and can be untrustworth in small samples, with weak instruments, and in other cases where we would worry about consistency).

Summary of Encouragement/Complier/Dose oriented designs:

- Analyze as you randomized, even when you don't control the dose you can learn something.
- The danger of per-protocol analysis: you give up the benefits of the research design (i.e. randomization)
- Variance calculations approximate (and can be untrustworth in small samples, with weak instruments, and in other cases where we would worry about consistency).

Summary of Encouragement/Complier/Dose oriented designs:

- Analyze as you randomized, even when you don't control the dose you can learn something.
- The danger of per-protocol analysis: you give up the benefits of the research design (i.e. randomization)
- Variance calculations approximate (and can be untrustworth in small samples, with weak instruments, and in other cases where we would worry about consistency).

1. Causal effects when we do not control the dose
2. Hypothesis Tests about Complier causal effects
3. Learning about causal effects when data are missing

Hypothesis Tests about Complier causal effects

- We can test the sharp null hypothesis no effect among all units
- We know by random assignment that
  1. this test will have a type I error probability at least as small as α
  2. Will have power greater than α for a class of alternative hypotheses
- Under what conditions is a test of the sharp null of no effect among all units equivalent to a test of the sharp null of no effect among Compliers?
  1. Exclusion restriction
  2. No Defiers
  3. Non-zero proportion of Compliers
  4. Non-interference

Sharp null hypothesis testing example

Imagine that our observed data is as follows:

$\mathbf{z}$	$\mathbf{y}$	$\mathbf{y_c}$	$\mathbf{y_t}$	$\mathbf{d}$	$\mathbf{d_c}$	$\mathbf{d_t}$
1	14	?	14	0	?	0
0	22	22	?	0	0	?
1	21	?	21	1	?	1
1	36	?	36	1	?	1
0	23	23	?	0	0	?
0	12	12	?	1	1	?
0	25	25	?	1	1	?
1	27	?	27	0	?	0

Observed experimental data

The observed Difference-in-Means test statistic, $\hat{\bar{\tau}}\left(\mathbf{Z}, \mathbf{Y}\right)$, is 16.75.

Sharp null hypothesis testing example

We can represent the sharp null hypothesis of no effect for all units without hypothesizing about non-random compliance (this is like the ITT_Y in that both can be assessed safely in a randomized experiment).

$\mathbf{z}$	$\mathbf{y}$	$\mathbf{y_c}$	$\mathbf{y_t}$	$\mathbf{d}$	$\mathbf{d_c}$	$\mathbf{d_t}$	Principal stratum
1	14	14	14	0	?	0	Never Taker or Defier
0	22	22	22	0	0	?	Complier or Never Taker
1	21	21	21	1	?	1	Complier or Always Taker
1	36	36	36	1	?	1	Complier or Always Taker
0	23	23	23	0	0	?	Complier or Never Taker
0	12	12	12	1	1	?	Always Taker or Defier
0	25	25	25	1	1	?	Always Taker or Defier
1	27	27	27	0	?	0	Never Taker or Defier

Sharp null of no effect for all units

Sharp null hypothesis testing example

The null hypothesis of no effect among compliers under excludability (meaning only a complier in the treatment group can have a causal effect), no Defiers and nonzero proportion of Compliers assumptions:

$\mathbf{z}$	$\mathbf{y}$	$\mathbf{y_c}$	$\mathbf{y_t}$	$\mathbf{d}$	$\mathbf{d_c}$	$\mathbf{d_t}$	Principal stratum
1	14	14	14	0	0	0	Never Taker
0	22	22	22	0	0	?	Complier or Never Taker
1	21	21	21	1	?	1	Complier or Always Taker
1	36	36	36	1	?	1	Complier or Always Taker
0	23	23	23	0	0	?	Complier or Never Taker
0	12	12	12	1	1	1	Always Taker
0	25	25	25	1	1	1	Always Taker
1	27	27	27	0	0	0	Never Taker

Sharp null of no effect among Compliers

We don't need to know which of units 2 – 5 are Compliers, only that at least one of these 4 units is a Complier.

Excludability means that the effect must be 0 for all units who are not compliers (i.e. implying the sharp null).

Summary

- The sharp null of no effects is meaningful and can be tested in a randomized experiment using assignment to treatment and ignoring compliance.
- The assumptions of excludability, no defiers, and at least one complier mean that we can interpret the test of the sharp null of no effects as a test of the sharp null of no effects on compliers.

Summary

- The sharp null of no effects is meaningful and can be tested in a randomized experiment using assignment to treatment and ignoring compliance.
- The assumptions of excludability, no defiers, and at least one complier mean that we can interpret the test of the sharp null of no effects as a test of the sharp null of no effects on compliers.

1. Causal effects when we do not control the dose
2. Hypothesis Tests about Complier causal effects
3. Learning about causal effects when data are missing

Review of core assumptions from randomized experiments

1. Excludability: Potential outcomes depend only on assigned treatment (and not other factors)
2. Non-interference
3. Random assignment of subjects treatment

Review of core assumptions from randomized experiments

1. Excludability: Potential outcomes depend only on assigned treatment (and not other factors)
2. Non-interference
3. Random assignment of subjects treatment

Review of core assumptions from randomized experiments

1. Excludability: Potential outcomes depend only on assigned treatment (and not other factors)
2. Non-interference
3. Random assignment of subjects treatment

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups. (We have trouble analyzing as we randomized!)
- Dropping the missing observations brings us closer to per-protocol analysis.

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups. (We have trouble analyzing as we randomized!)
- Dropping the missing observations brings us closer to per-protocol analysis.

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups. (We have trouble analyzing as we randomized!)
- Dropping the missing observations brings us closer to per-protocol analysis.

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups. (We have trouble analyzing as we randomized!)
- Dropping the missing observations brings us closer to per-protocol analysis.

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups. (We have trouble analyzing as we randomized!)
- Dropping the missing observations brings us closer to per-protocol analysis.

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups. (We have trouble analyzing as we randomized!)
- Dropping the missing observations brings us closer to per-protocol analysis.

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups. (We have trouble analyzing as we randomized!)
- Dropping the missing observations brings us closer to per-protocol analysis.

Attrition (missing data on outcomes)

- Some units may have missing data on outcomes (= units attrit) when:
  - some respondents can't be found or refuse to participate in endline data collection.
  - some records are lost.
- This is a problem when treatment affects missingness.
  - For example, units in control may be less willing to answer survey questions.
  - For example, treatment may have caused units to migrate and cannot be reached
- If we analyze the data by dropping units with missing outcomes, then we are no longer comparing similar treatment and control groups. (We have trouble analyzing as we randomized!)
- Dropping the missing observations brings us closer to per-protocol analysis.

What can we do?

- Check whether attrition rates are similar in treatment and control groups.
- Check whether treatment and control groups have similar covariate profiles.
- Do not drop observations that are missing outcome data from your analysis.
- When outcome data are missing we can sometimes **bound** our estimates of treatment effects.

What can we do?

- Check whether attrition rates are similar in treatment and control groups.
- Check whether treatment and control groups have similar covariate profiles.
- Do not drop observations that are missing outcome data from your analysis.
- When outcome data are missing we can sometimes **bound** our estimates of treatment effects.

What can we do?

- Check whether attrition rates are similar in treatment and control groups.
- Check whether treatment and control groups have similar covariate profiles.
- Do not drop observations that are missing outcome data from your analysis.
- When outcome data are missing we can sometimes **bound** our estimates of treatment effects.

What can we do?

- Check whether attrition rates are similar in treatment and control groups.
- Check whether treatment and control groups have similar covariate profiles.
- Do not drop observations that are missing outcome data from your analysis.
- When outcome data are missing we can sometimes **bound** our estimates of treatment effects.

What can we do?

- But the best approach is to try to anticipate and prevent attrition.
  - Blind people to their treatment status.
  - Promise to deliver the treatment to the control group after the research is completed.
  - Plan ex ante to reach all subjects at endline.
  - Budget for intensive follow-up with a random sample of attriters.

What can we do?

- But the best approach is to try to anticipate and prevent attrition.
  - Blind people to their treatment status.
  - Promise to deliver the treatment to the control group after the research is completed.
  - Plan ex ante to reach all subjects at endline.
  - Budget for intensive follow-up with a random sample of attriters.

What can we do?

- But the best approach is to try to anticipate and prevent attrition.
    - Blind people to their treatment status.
    - Promise to deliver the treatment to the control group after the research is completed.
    - Plan ex ante to reach all subjects at endline.
    - Budget for intensive follow-up with a random sample of attriters.

What can we do?

- But the best approach is to try to anticipate and prevent attrition.
  - Blind people to their treatment status.
  - Promise to deliver the treatment to the control group after the research is completed.
  - Plan ex ante to reach all subjects at endline.
  - Budget for intensive follow-up with a random sample of attriters.

What can we do?

- But the best approach is to try to anticipate and prevent attrition.
  - Blind people to their treatment status.
  - Promise to deliver the treatment to the control group after the research is completed.
  - Plan ex ante to reach all subjects at endline.
  - Budget for intensive follow-up with a random sample of attriters.

Missing data on covariates is not as problematic

- Missing **background covariates** (i.e.,variables for which values do not change as a result of treatment) for some observations is less problematic.
  - We can still learn about the causal effect of an experiment without those covariates.
  - We can also use the background covariate as planned by imputing for the missing values.
  - We can also condition on that missingness directly: we could assess causal effects for the subgroup of those missing on income and compare it to the subgroup of those not missing on income.

Missing data on covariates is not as problematic

- Missing **background covariates** (i.e.,variables for which values do not change as a result of treatment) for some observations is less problematic.
  - We can still learn about the causal effect of an experiment without those covariates.
  - We can also use the background covariate as planned by imputing for the missing values.
  - We can also condition on that missingness directly: we could assess causal effects for the subgroup of those missing on income and compare it to the subgroup of those not missing on income.

Missing data on covariates is not as problematic

- Missing **background covariates** (i.e.,variables for which values do not change as a result of treatment) for some observations is less problematic.
  - We can still learn about the causal effect of an experiment without those covariates.
  - We can also use the background covariate as planned by imputing for the missing values.
  - We can also condition on that missingness directly: we could assess causal effects for the subgroup of those missing on income and compare it to the subgroup of those not missing on income.

Missing data on covariates is not as problematic

- Missing **background covariates** (i.e.,variables for which values do not change as a result of treatment) for some observations is less problematic.
    - We can still learn about the causal effect of an experiment without those covariates.
    - We can also use the background covariate as planned by imputing for the missing values.
    - We can also condition on that missingness directly: we could assess causal effects for the subgroup of those missing on income and compare it to the subgroup of those not missing on income.

Summary about Missing data and Experiments.

- Missing outcomes or missing treatment assignment (or missing blocking information) are all big problems. How might those with missing outcomes have behaved in treatment versus control? We don't know.
- Missing covariate information is not a problem: it is fixed, same proportion should be missing covariate information in both treatment and control conditions

Summary about Missing data and Experiments.

- Missing outcomes or missing treatment assignment (or missing blocking information) are all big problems. How might those with missing outcomes have behaved in treatment versus control? We don't know.
- Missing covariate information is not a problem: it is fixed, same proportion should be missing covariate information in both treatment and control conditions

References

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

Imbens, Guido W and Paul R Rosenbaum (2005). "Robust, Accurate Confidence Intervals with a Weak Instrument: Quarter of Birth and Education". In: Journal of the Royal Statistical Society: Series A (Statistics in Society) 168.1, pp. 109–126.