Lecture VI
Problems with Experiments: Attrition

Stanislao Maldonado

Universidad del Rosario

Impact Evaluation

March 7th, 2016

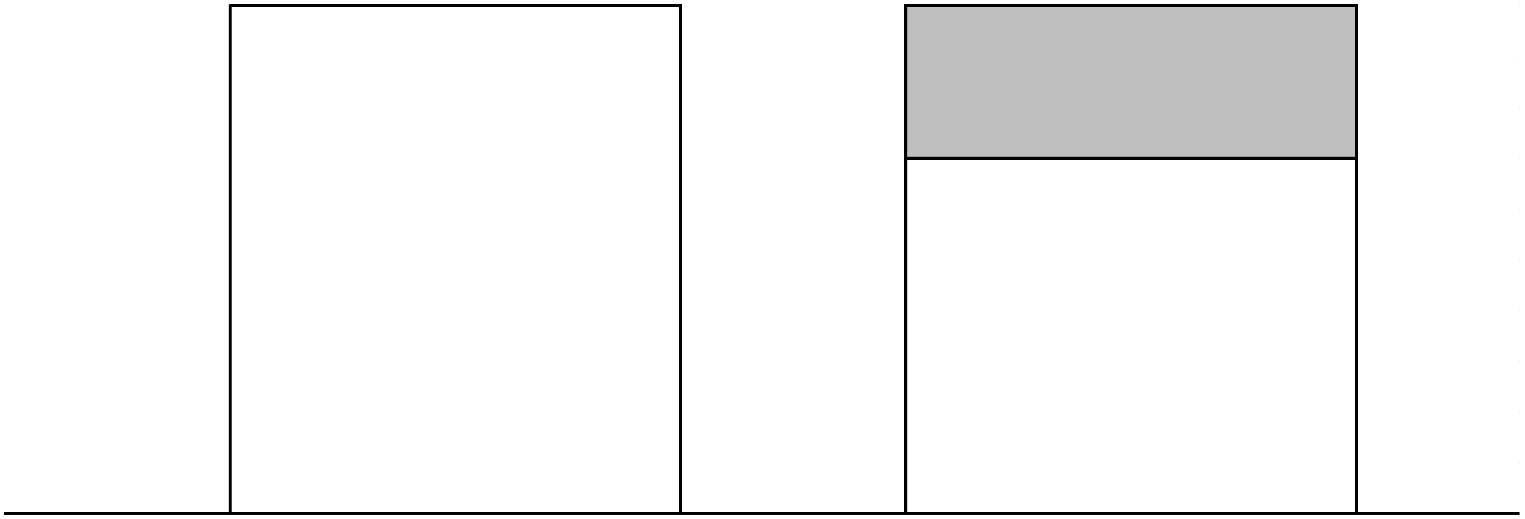
1. Introduction

- Experiments are subject to implementation issues that have profound consequences in the way they are analyzed and interpreted
- We focus on this lecture on attrition. Attrition occurs when outcome data are missing. Why?:
  - Subjects may refuse to cooperate with researchers
  - Researchers lose track of experimental subjects
  - Firms, organizations, or government agency block researchers' access to outcomes
  - Outcome variable may be intrinsically unavailable for some subjects
  - Researchers may deliberately discard observations

- What's wrong with attrition?

- Example: Campaign to increase children vaccination in poor areas of Bogota
  - 2,000 households randomly divided in two groups (1,000 treated)
  - Only 300 households in the control group were actually interviewed by the campaign

What to compare?



- Attrition forces the researcher to make assumptions about the statistical properties of missingness:
  - Option 1: Missingness is independent of potential outcomes
  - Option 2: Missingness is independent of potential outcomes conditional on observable covariates
  - Option 3: Missingness is independent of potential outcomes conditional on unobservable covariates
  - Option 4: Bounding of the parameter of interest

2. Attrition Bias: Formal Treatment

- Notation:
  - N units denoted by i
  - Two levels of treatment: Z=0 and Z=1 (perfect compliance)
  - Y is a measure of the outcome
  - r is outcome reporting: r=0 and r=1. Notice that:
    - Outcome reporting may or may not be random
    - SUTVA is assumed
- Outcome reporting: $r_i = r_i(Z_i)$
- Observed reporting outcome: $r_i = r_i(0)(1 - Z_i) + r_i(1)Z_i$

- Notice that observed r depends on Z. Since some subjects have the potential for attrition even if attrition is not observed:
  - r: reported outcome is the result of an past experiment
  - R: reported outcome of an hypothetical experiment
- Potential outcomes will become known based on whether r=0 or r=1: $Y_i(Z, r(Z))$
- To simplify, we assume that potential outcomes are unaffected by whether these outcomes are reported. Therefore:

$$Y_i(Z) = Y_i(Z, r(Z) = 1) = Y_i(Z, r(Z) = 0)$$

- The observed outcome:

$$Y_i = Y_i(0) + [Y_i(1) - Y_i(0)]Z_i \qquad \text{if } (r_i = 1)$$

$$Y_i = . \qquad \text{if } (r_i = 0)$$

- Recall that, with no attrition, a simple difference in means between treated and control units recovers ATE. With attrition this is not longer true.

- To see this, consider the **expected potential outcome under treatment** for those observable and those missing:

$$E[Y_i(1)] =$$

$$E[R_i(1)]E[Y_i(1)/R_i(1) = 1] + \{1 - E[R_i(1)]\}E[Y_i(1)/R_i(1) = 0]$$

- Also consider the **expected potential outcome under control** for those observable and those missing:

$$E[Y_i(0)] =$$

$$E[R_i(0)]E[Y_i(0)/R_i(0)=1]+\{1-E[R_i(0)]\}E[Y_i(0)/R_i(0)=0]$$

- ATE for all subjects can be expressed as:

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

$$E[R_i(1)]E[Y_i(1)/R_i(1)=1]+\{1-E[R_i(1)]\}E[Y_i(1)/R_i(1)=0]-$$

$$E[R_i(0)]E[Y_i(0)/R_i(0)=1]-\{1-E[R_i(0)]\}E[Y_i(0)/R_i(0)=0]$$

- Compare this to the case in which the analysis is restricted to subjects with no missing values:

$$E[Y_i(1)/R_i(1)=1]-E[Y_i(0)/R_i(0)=1]$$

- ATE for all subjects may be quite different for ATE for non-missing subjects
- When these two quantities are the same?

$$E[Y_i(1)/R_i(1)=1]=E[Y_i(1)/R_i(1)=0]$$

$$E[Y_i(0)/R_i(0)=1]=E[Y_i(0)/R_i(0)=0]$$

- This is true when missingness is not related with Y(0) and Y(1)

Example: Potential Outcomes under Attrition

Individual	Y(0)	Y(1)	R(0)	R(1)	Y(0)|R(0)	Y(1)|R(1)
1	3	8	0	1	.	8
2	3	7	0	1	.	7
3	8	10	1	1	8	10
4	7	8	1	1	7	8
5	6	6	0	0	.	.
6	5	8	1	1	5	8
7	6	6	1	0	6	.
8	6	7	1	1	6	7
Average	5.5	7.5	0.625	0.75	6.4	8

- The ATE for subjects with no missing values can be decomposed as follows:

$$E[Y_i(1)/R_i(1)=1]-E[Y_i(0)/R_i(0)=1]=$$

$$\underbrace{E[Y_i(1)/R_i(0)=1]-E[Y_i(0)/R_i(0)=1]}_{\text{ATE for subjects who report when untreated}}$$

$$+\underbrace{E[Y_i(1)/R_i(1)=1]-E[Y_i(1)/R_i(0)=1]}_{\text{Heterogeneity}}$$

- Key: ATE for respondents does not recover ATE for any meaningful group!!

3. Several Forms of Attrition Bias

- Consider a more general way to express outcomes to motivate different forms of attrition:

  $$Y_i = \varphi(\mathrm{Z}, \mathrm{X}, \mathrm{U})$$

- Where the outcome depends on the assignment Z, a set of observable covariates X and non-observable determinants U
- Z is assumed to be randomly assigned: $(\mathrm{A}.1)\ \ \mathrm{Z} \perp (\mathrm{X}, \mathrm{U})$
- Motivating example: impact of vouchers for private schooling (Z) on test scores (Y)
  - Relevant determinants such age and gender (X) and talent (U)

A. Missing completely at random (MCAR)

- Attrition is not related to with any observed or unobserved factor (Rubin 1976). Therefore:

  $(\mathrm{A}.2)\ R \perp (\mathrm{Z}, \mathrm{X}, \mathrm{U})$

- This implies that:

  $Y(1), Y(0) \perp R$

  $Y(1), Y(0), R \perp Z$

- In our example, taking the test is neither related to winning the lottery nor to any observable (age, gender) or unobservable (talent) dimension
- Therefore:

$$E[Y_i/Z_i = 1, R_i = 1] - E[Y_i/Z_i = 0, R_i = 1]$$
$$= E[Y_i/Z_i = 1] - E[Y_i/Z_i = 0] = \text{ATE}$$

- This assumption cannot be verified although indirect evidence can be gathered by the researcher:
  - If MCAR is true, then we should expect no systematic relationship between response and subject's background attributes
  - However, failure to reject the null does not imply that missingness is not related to potential outcomes

B. Missingness is related to Z

- Attrition is not related to with any observed or unobserved factor but it depends on treatment assignment Z:

$$(\text{A.3a})\ R = 1\{\zeta(Z, V) \geq 0\}$$

$$(\text{A.3b})\ V \perp (\text{X}, \text{U})$$

- V is a non-observable factor that does not depend on X and U
- By A.3a, Z shifts R such that: $\Pr(\text{Z} = 1 \mid R = 1) \neq \Pr(\text{Z} = 1)$
- Since Z also affects Y, it follows that: $E(Y \mid R = 1) \neq E(Y)$
- This does not affect identification since the distribution of (X,U) is not related to R

- Therefore:

$$Y(0), Y(1), R(0), R(1) \perp Z$$

$$Y(0), Y(1) \perp R \mid Z$$

- It holds that:

$$E[Y_i / Z_i = 1, R_i = 1] - E[Y_i / Z_i = 0, R_i = 1]$$

$$= E[Y_i / Z_i = 1] - E[Y_i / Z_i = 0] = \text{ATE}$$

C. Missingness is related to Z and X

- Attrition is not related to with any unobserved factor but it depends on treatment assignment Z and observed factors X:

$$(\text{A.4a})\ R = 1\{\zeta(Z, \text{X}, V) \geq 0\}$$

$$(\text{A.4b})\ V \perp (\text{X}, \text{U})$$

$$(\text{A.4c})\ \text{X} \perp Z \mid R = 1$$

- A.4a and A.4b similar to previous case, except that X now influences R. X and Z are independent conditional on response
- Attrition affects distributions of X and Z, but they are not related each other given A.4c.

- This implies that change in distribution of X is the same across treatment states
- Example: Age is positively related with participating in the test. Then, conditional on participating, age must change in the same manner for lottery winners and losers
- The experiment identifies ATE among respondents:

$$E[Y_i / Z_i = 1, R_i = 1] - E[Y_i / Z_i = 0, R_i = 1]$$
$$= E[Y_i(1) / R_i = 1] - E[Y_i(0) / R_i = 0] = \text{ATE}_{R=1}$$

- Imposing additional restrictions allow us to recover ATE. This is achieved by weighting respondents according to the likelihood that their observed characteristics appear in the population
- A propensity score for the likelihood of response can be defined: $p(Z, X) \equiv \Pr(R = 1 | Z, X)$
- A common support restriction is imposed for identification:

  $\Pr(R = 1 | Z = z, X = x) > \mathrm{c}$ for all $x \in \chi$, $z \in \{0,1\}$, $\mathrm{c} > 0$

- Under A1, A4 and the common support, it can be shown that:

$$ATE = E\left[\frac{R.Y}{p(1,X)} \mid Z = 1\right] - E\left[\frac{R.Y}{p(0,X)} \mid Z = 0\right]$$

- Weighting observations by the inverse of their respective **response propensity score** recovers ATE
- Inverse probability weighting (IPW) has been extensively applied in contexts where there are survey missing data (Horvitz and Thompson 1952)

Example: IPW to recover ATE

Individual	Y(0)	Y(1)	R(0)	R(1)	Y(0)|R(0)	Y(1)|R(1)	X
1	3	4	1	1	3	4	1
2	4	7	1	1	4	7	1
3	3	4	1	1	3	4	1
4	4	7	1	1	4	7	1
5	10	14	0	0	.	.	0
6	12	18	0	0	.	.	0
7	10	14	1	1	10	14	0
8	12	18	1	1	12	18	0
Average	7.25	10.75	0.75	0.75	6	9	0.5

- ATE=10.75-7.25=3.5
- We can use the expression for ATE and derive estimates for the inverse probability as follows:

$$E\left[\frac{1}{p(1, X=1)} \mid Z=1\right] = E\left[\frac{1}{1} \mid Z=1\right] = 1$$

$$E\left[\frac{1}{p(0, X=1)} \mid Z=0\right] = E\left[\frac{1}{1} \mid Z=0\right] = 1$$

$$E\left[\frac{1}{p(1, X=0)} \mid Z=1\right] = E\left[\frac{1}{2/4} \mid Z=1\right] = 2$$

$$E\left[\frac{1}{p(0, X=0)} \mid Z=0\right] = E\left[\frac{1}{2/4} \mid Z=0\right] = 2$$

- We use these weights to compute ATE as follows:

$$E\left[\frac{R.Y}{p(1, X)} \mid Z=1\right] = \left(\frac{1}{8}\right)\left(\frac{4}{1}+\frac{7}{1}+\frac{4}{1}+\frac{7}{1}+\frac{14}{0.5}+\frac{18}{0.5}\right) = 10.75$$

$$E\left[\frac{R.Y}{p(0, X)} \mid Z=0\right] = \left(\frac{1}{8}\right)\left(\frac{3}{1}+\frac{4}{1}+\frac{3}{1}+\frac{4}{1}+\frac{10}{0.5}+\frac{12}{0.5}\right) = 7.25$$

$$ATE = 10.75 - 7.25 = 3.5$$

- Some potential concerns with weighting:
  - If assumptions are incorrect, applying IPW may produce misleading estimates
  - IPW may increase sampling variability
  - If attrition is related to unobservables, IPW does not work without further restrictions

D. Missingness is related to unobservables

- We allow now for a non-zero correlation between U and V even after conditioning for Z and X
- An instrument I that affects R but without direct effect on Y is required to point-identification
- Example: Probability of taking the test is a function of unobserved motivation and ability, even after conditioning on treatment and observed characteristics (age, gender, etc.)
- Several approaches have been suggested:
  - Sample selection methods (Heckman 1976, DiNardo et al 2006)

- IPW methods (Huber 2012)
- Bounds (Horowitz and Manski 1998 and 2000, Lee 2009)

i. Sample Selection

- We introduce some additional notation. Let's define the latent variables:

$$Y^* = ZY^*(1) + (1-Z)Y^*(0)$$

$$R^* = \mu_R(X, Z, I) + V$$

- Where the counterfactual outcomes are:

$$Y^*(1) = \mu_1(X, U_1)$$

$$Y^*(0) = \mu_0(X, U_0)$$

- We observe:

$$Y = \begin{cases} Y^* & \text{if } R = 1 \\ . & \text{if } R = 0 \end{cases}$$

$$R = 1(R^* \geq 0)$$

- We can re-express the latent variables as follows:

$$Y^* = ZY^*(1) + (1-Z)Y^*(0) = \mu_0 + \theta Z + \alpha X + U$$

$$R^* = \mu_R(X, Z, I) + V = \mu_R + \beta Z + \gamma X + \delta I + V$$

- The standard textbook approach relies in strong distributional assumptions and assumes that (U,V) are distributed bivariate normal with mean zero and a variance matrix which is normalized to 1. However, as long as $\delta \neq 0$, normality is not critical for identification
- The standard two-step Heckman estimator implies running R on (Z,X,I), compute fitted values to construct the inverse Mill ratio (IMR), and regress Y on Z, X and IMR

DiNardo, McCrary and Sanbonmatsu (2006)

- Dinardo et al (2006) suggest an algorithm to deal with attrition based on the use of the two-step Heckman strategy:
  - Randomize individuals at baseline to Z=1 or Z=0
  - For each group Z, randomize individuals into subgroups based on instrument I (Ex: # visits for interviews or monetary transfers)
  - Based on I, employ different effort intensities to collect necessary outcome data
  - Run the following OLS regression implied by:

$$E[Y_i / Z, X, I, R_i = 1] = \mu_0 + \theta Z + \alpha X + kg(\Pr(R = 1 | X, Z, I))$$

Where k is the covariance between U and V. The values for IMR are the estimated values.

ii. Bounds on ATE

- Sample selection and IPW approaches relies in strong and unattractive assumptions about attrition process
- Instead of modelling attrition process, some scholars suggests to compute bounds for treatment effects imposing few assumptions about the filling of the missing potential outcomes
- Two popular approaches:
  - Extreme value bounds (Horowitz and Manski 1998,2000)
  - Lee bounds (Lee 2009)

Extreme Value Bounds

- Let's introduce some additional notation:

  $Y^L =$ Lowest possible value of outcome

  $Y^H =$ Greatest possible value of outcome

- Bounds are constructed assuming the "worst case" about missing data
- Bounds are given by:

$$\beta^H = P[R=1|Z=1].E[Y|Z=1] + \{1-P[R=1|Z=1]\}Y^H$$
$$-\left(P[R=1|Z=0].E[Y|Z=0] + \{1-P[R=1|Z=0]\}Y^L\right)$$

$$\beta^L = P[R=1|Z=1].E[Y|Z=1] + \{1-P[R=1|Z=1]\}Y^L$$
$$-\left(P[R=1|Z=0].E[Y|Z=0] + \{1-P[R=1|Z=0]\}Y^H\right)$$

- When the outcome have wide support, bounds can uninformative
- Example about bounds (see next page):
  - Assume that outcome range from 0 to 10
  - Assume subjects 2,3,5 and 7 were assigned to treatment and 1,4,6, and 8 to control

Example: Bounds on ATE

Individual	Y(0)	Y(1)	R(0)	R(1)	Y(0)|R(0)	Y(1)|R(1)
1	3	8	0	1	.	8
2	3	7	0	1	.	7
3	8	10	1	1	8	10
4	7	8	1	1	7	8
5	6	6	0	0	.	.
6	5	8	1	1	5	8
7	6	6	1	0	6	.
8	6	7	1	1	6	7
Average	5.5	7.5	0.625	0.75	6.4	8

- Without attrition:
  - Treatment: (7+10+6+6)/4=29/4=7.25
  - Control: (3+7+5+6)/4=21/4=5.25
- With attrition:
  - Treatment: (7+10+?+?)/4=?
  - Control: (?+7+5+6)/4=?
- Upper bound:
  - Treatment: (7+10+10+10)/4=37/4
  - Control: (0+7+5+6)/4=18/4
  - T-C=19/4

- Lower bound:
  - Treatment: (7+10+0+0)/4=17/4
  - Control: (10+7+5+6)/4=28/4
  - T-C=-11/4
- Notice that, although true ATE is contained in the set, bounds are too wide to be informative and contain 0

Lee's (2009) Bounds

- Let's introduce some additional notation. Recall R(Z) is the response as a function of treatment assignment Z. Consider the following cases:

$$\tau_{R_{11}} : R(0)=1, R(1)=1 \Rightarrow Y=y \text{ if } Z=0,\ Y=y \text{ if } Z=1$$

$$\tau_{R_{01}} : R(0)=0, R(1)=1 \Rightarrow Y=. \text{ if } Z=0,\ Y=y \text{ if } Z=1$$

$$\tau_{R_{10}} : R(0)=1, R(1)=0 \Rightarrow Y=y \text{ if } Z=0,\ Y=. \text{ if } Z=1$$

$$\tau_{R_{00}} : R(0)=0, R(1)=0 \Rightarrow Y=. \text{ if } Z=0,\ Y=. \text{ if } Z=1$$

- Selection into response is given by:

$$R = R(1)Z + R(0)(1-Z)$$

- Using some assumptions, it is possible to bound the treatment effect for individuals of type $\tau_{R_{11}}$:
  - Random assignment to treatment Z
  - Monotonicity:
    $$R(1) \geq R(0) \vee R(0) \geq R(1) \text{ for all i}$$
- As in the case of non-compliance, monotonicity rules out those who attrite in opposite direction
- Bounds are for ATE for the subpopulation of always reporters
- Assume that $R(1) \geq R(0)$, then:

- Observed treated individuals: $\tau_{R_{11}}, \tau_{R_{01}}$
- Observed control individuals: $\tau_{R_{11}}$

- In this case, Lee bounds are given by:

$$\beta^H = E[Y \mid Z=1, R=1, Y \geq \psi_{(p_0)}] - E[Y \mid Z=0, R=1]$$

$$\beta^L = E[Y \mid Z=1, R=1, Y \leq \psi_{(1-p_0)}] - E[Y \mid Z=0, R=1]$$

$$\psi_{(p_0)} = G^{-1}_{R=1,Z=1}(p_o)$$

$$p_o = \frac{P[R=1 \mid Z=1] - P[R=1 \mid Z=0]}{P[R=1 \mid Z=1]}$$

- **Algorithm**: Assume that 50% of treatment group remain and 40% control group remain
  - Observations for the more frequently observed group are trimmed (treatment group in this case)
  - We compute: $p_o = (0.5 - 0.4) / 0.5 = 0.2$
  - The upper bound is computed as follows:
    - Compute the mean for control group
    - Drop the lowest 20% outcomes for treatment group and calculate mean for remaining
    - Calculate the difference between the trimmed treatment group and the control group mean
  - Same logic applies to the lower bound