

Abstract Title Page

Title:

Biases in Estimating Treatment Effects Due to Attrition in Randomized Controlled Trials and Cluster Randomized Controlled Trials: A Simulation Study

Author(s):

Nianbo Dong
Peabody Research Institute, Vanderbilt University
nianbo.dong@vanderbilt.edu

Mark W. Lipsey
Peabody Research Institute, Vanderbilt University
mark.lipsey@vanderbilt.edu

Abstract Body

Background / Context:

Attrition occurs when study participants who were assigned to the treatment and control conditions do not provide outcome data and thus do not contribute to the estimation of the treatment effects. It is very common in experimental studies in education as illustrated, for instance, in a meta-analysis studying “the effects of attrition on baseline comparability in randomized experiments in education” (Valentine & McHugh, 2007) that found that 119 of 367 randomized education experiments reported student-level attrition. Shadish et al (1998) called attrition the Achilles’ heel of randomized experiments. Attrition reduces statistical power by decreasing sample size. It compromises external validity when those who do not contribute data are unrepresentative of the original sample and thus degrade the representativeness of those who remain in the sample (Orr, 1999).

Of greatest concern in experimental studies, however, is the potential for attrition to threaten internal validity (Campbell & Stanley, 1963; Little & Rubin, 1987; Shadish, Cook, & Campbell, 2001). In the Little & Rubin (1987) missing data framework, if attrition involves cases missing completely at random (MCAR: does not depend on observed or unobserved variables) or missing at random (MAR: conditional on observed variables) unbiased treatment effect estimates can be derived from the obtained data. However, if the attrition is not random under either of these definitions (is related to at least some extent to unobserved variables), it can bias the treatment effect estimate. Moreover, that bias can be large even when advanced statistical methods are used to address the attrition, e.g., multiple imputation (Foster & Fang, 2004; Puma, Olsen, Bell, & Price, 2009). Unfortunately for the internal validity of experimental studies, it is rarely the case that the factors that create attrition can be safely assumed to be represented among the observed variables or to be a virtually random process.

It should also be noted that attrition may not only bias the point estimate of the treatment effect but also the standard error of that estimate. Hence, it may distort the results of statistical significance tests and thereby threaten statistical conclusion validity.

We focus in this presentation on the threat attrition poses for internal validity. From that perspective, the critical questions for experimental studies are how much potential bias different levels and kinds of attrition might produce in the effect estimates and how much is too much to allow confidence in those estimates. The What Works Clearinghouse Procedures and Standards Handbook (version 2.0) set a standard of 0.05 standard deviation or less on the outcome variable as an acceptable level of bias (U.S. Department of Education, 2008). The overall attrition rates and differential attrition rates between treatment and control groups that could result in that much bias or more was estimated through a simulation study.

The team that did the work on attrition for the WWC handbook made a major contribution to understanding this issue. Remarkably, given how much awareness there is in the experimental research community of the problems associated with attrition, little systematic research has been done to provide a framework for appraising attrition and assessing its potential to bias effect estimates. The most notable gaps in the current literature on this topic include: (1) lack of a well-developed conceptual and statistical framework that makes the sources of attrition bias and their influence clear; (2) the potential of well-chosen covariates to reduce attrition bias; (3) the nature and influence of attrition in cluster randomized controlled trials; (4) attrition bias in the standard errors of effect estimates; and (5) how to use baseline information most effectively in estimating the potential attrition bias and what baseline information is most useful for that purpose. We

have explorations of these issues underway. The proposed presentation will report results to date on the first three issues in this list.

Purpose / Objective / Research Question / Focus of Study:

The main purpose of the study to be presented is to elaborate a model of the relationships between attrition and effect estimates and to use that model to guide Monte Carlo simulations that examine the sources and magnitude of attrition bias under various assumptions for randomized experiments and cluster randomized experiments.

Significance / Novelty of study:

This study contributes to methodology and practice for assessing attrition bias in randomized experiments and cluster randomized experiments in education. Its results provide guidance for identifying the conditions under which attrition bias might present a serious threat to internal validity and suggestions for ways to reduce that bias.

Research Design:

This is a simulation study that examines attrition bias in completely randomized controlled trials (RCT) and cluster randomized controlled trials (CRT). For each of these designs, this project involves three components: (1) specifying a model of attrition bias; (2) specifying a model for treatment effect estimates that can be linked to attrition bias; and (3) using these models to generate data to simulate different attrition situations and their influence on effect estimates.

Completely Randomized Controlled Trials (RCT)

1. Model of Attrition Bias

We start from the simplest model, i.e., the model without any covariates, which was used in the WWC Procedures and Standards Handbook. We will then move to a model with covariate.

Let a random variable, z , represent an individual’s latent propensity to respond. Assume z has a normal distribution, $z \sim N(0, 1)$. The response rate of the sample is p , $p = \frac{N_{respondent}}{N_{total}}$. An

individual is a respondent if her/his propensity value z exceeds a threshold:

$$(1) z > Q(z, 1 - p)$$

where Q is the quantile function of the normal distribution, i.e., the inverse of the cumulative distribution function. Given the response rate p , if z exceeds the value that corresponds to the percentile $(1 - p)$ of the normal distribution, then that individual responds on the outcome measure.

The outcome at follow-up, y , may be influenced by the propensity to respond, z , so can be modeled as follows with all other influences represented for the moment in the random term, e .

$$(2) y = \beta_z z + e \quad y \sim N(0, 1), z \sim N(0, 1), e \overset{iid}{\sim} N(0, \sigma_e^2)$$

The correlations of y and z are assumed to be: $\text{Corr}(y, z) = r_{yz}$, we will have $\beta_z = r_{yz}$.

When the treatment variable is included, we allow for the possibility of an interaction between treatment conditions and the propensity to respond and the model becomes:

$$(3) y = \beta_z z + \beta_1(TREAT) + \beta_2(TREAT)z + e$$

This model assumes that there is a treatment effect and that it may varies conditional on z . β_1 , is the average treatment effect when z equals to 0.

Expressed by the treatment status:

For the control group:

$$(4) y_c = \beta_z z_c + e_c \quad y_c \sim N(0, 1), z_c \sim N(0, 1), e \sim N(0, \sigma_c^2) \text{ iid}$$

The correlations of y_c and z_c are assumed to be: $\text{Corr}(y_c, z_c) = r_{yz}^c$, we will have $\beta_z = r_{yz}^c$.

For the treatment group:

$$(5) y_t = \beta_1 + (\beta_z + \beta_2)z_t + e_t \quad y_t \sim N(\bar{Y}_t, 1), \text{ where } \bar{Y}_t \text{ is the mean of } y_t, z_t \sim N(0, 1)^*, \\ e \sim N(0, \sigma_t^2) \text{ iid}$$

The correlations of y_t and z_t are assumed to be: $\text{Corr}(y_t, z_t) = r_{yz}^t$, we will have $\beta_z + \beta_2 = r_{yz}^t$.

Regarding the attrition model with covariate (e.g., pretest), we use the following model where x is the covariate :

$$(6) y = \beta_x x + \beta_z z + \beta_1(TREAT) + \beta_2(TREAT)x + \beta_3(TREAT)z + e$$

$$y \sim N(0, 1), x \sim N(0, 1), z \sim N(0, 1), e \sim N(0, \sigma_e^2) \text{ iid}$$

The correlations of y , x , and z are assumed to be:

$$\text{Corr}(y, x) = r_{yx}, \text{Corr}(y, z) = r_{yz}, \text{Corr}(x, z) = r_{xz}.$$

2. Model of Impact Estimate

First, we illustrate the sources of bias using the attrition model without covariate (Models 1-5). The treatment effect (δ) is the difference between the treatment and control group means on the outcome variable, y :

$$(7) \delta = \bar{Y}_t - \bar{Y}_c = \beta_1 + r_{yz}^t \bar{z}_t - r_{yz}^c \bar{z}_c = \beta_1 + (r_{yz}^t + \beta_2) \bar{z}_t - r_{yz}^c \bar{z}_c$$

For the whole sample without attrition, $\bar{z}_t = \bar{z}_c = 0$ and the treatment effect is

$$\delta = \bar{Y}_t - \bar{Y}_c = \beta_1.$$

Furthermore, according to Expression 1, for any given response rate p , we can find the threshold value for the propensity to respond, z_0 (or vice versa). The respondent sample consists of the people whose propensity to respond is above this threshold. The distribution of z , the propensity to respond, for those who do respond is a truncated standard normal distribution ($z \in (z_0, \infty)$).

Bias of the treatment effect estimate is defined as the difference between the treatment effect estimate derived from the full original sample, if that estimate were known, and that derived from the sample of those who actually respond:

$$(8) \text{Bias} = r_{yz}^t \bar{z}_t - r_{yz}^c \bar{z}_c, \text{ where } \bar{z}_t \text{ and } \bar{z}_c \text{ are for the respondent sample.}$$

Rearranging Expression 7, we have

* Note that the actual functions of propensity to respond could be different between the treatment and control groups. However, after standardization, they will be same, i.e., standard normal distribution.

$$(9) \text{ Bias} = (r_{yz}^t - r_{yz}^c) \left(\frac{\bar{z}_t + \bar{z}_c}{2} \right) + \left(\frac{r_{yz}^t + r_{yz}^c}{2} \right) (\bar{z}_t - \bar{z}_c).$$

From this formulation, we see that bias is the sum of two parts: $(r_{yz}^t - r_{yz}^c) \left(\frac{\bar{z}_t + \bar{z}_c}{2} \right)$ and $\left(\frac{r_{yz}^t + r_{yz}^c}{2} \right) (\bar{z}_t - \bar{z}_c)$. The first part concerns the difference of correlations between y and z for two groups and the average z of two groups, which is determined by the average response rate. The second part concerns the average correlation between y and z for two groups and the difference of mean z between the treatment and control groups, which is due to the differential response rates. In short, both the average and differential correlations between y and z for the treatment and control groups, and both the average and differential response rates between the treatment and control groups contribute to the overall bias, and the magnitude of that bias depends on the bias directions of these two parts.

The mean of z for the truncated standard normal distribution can be calculated using the formula below (Barr & Sherrill, 1999):

$$(10) E(z) = \frac{1}{\sqrt{2\pi}(1 - \phi(z_0))} e^{-z_0^2/2}, \text{ where } \phi(z_0) \text{ is the standard normal CDF (cumulative}$$

distribution function), and it equals to the attrition rate, i.e., $(1 - p)$.

There are many simple formulas to calculate $\phi(z_0)$ from z_0 (or vice versa). Based on Shah's approximation (1985, p. 80), we can calculate z_0 using the formula below:

$$(11) z_0 = -2.2 + 0.5\sqrt{40(1 - p) - 0.64}, \quad -2.2 < z_0 \leq 0 \text{ (equivalent to } 0.5 \leq p < 0.98 \text{)}.$$

Using Expressions 9, 10 and 11, we can easily calculate bias for any given parameters (p_t , p_c , r_{yz}^t , and r_{yz}^c).

To estimate the treatment effect and bias for the model with a covariate, we used simulation based on the model below:

$$(12) y = \beta_0^* + \beta_x^* x + \beta_1^* (TREAT) + \beta_2^* (TREAT)x + e^*$$

β_1^* , is the average treatment effect when x equals to 0. The average treatment effect at the average x is $(\beta_1^* + \beta_2^* \bar{x})$, where \bar{x} is the mean of x . Model 12 can produce an unbiased impact estimate for the complete sample because z is unrelated to the treatment status in the complete sample. However, Model 12 may produce a biased impact estimate for the respondent sample because z may be related to the treatment status in the respondent sample. If we set β_1 as 0, β_1^* and $(\beta_1^* + \beta_2^* \bar{x})$ estimated from the respondent sample are the overall biases of the treatment effect in raw score at x of 0 and the average, respectively.

3. Simulation Procedure

We used a macro program written in SAS 9.2 to generate and analyze data for the model with covariate. Based on the model of attrition bias, the parameters that we can manipulate include: response rate for the control group (p_c), response rate for the treatment group (p_t), r_{yx}^c , r_{yz}^c , r_{xz}^c ,

r_{yx}^t , r_{yz}^t , r_{xz}^t , and β_1 . To simplify, we set β_1 as 0. Given these specific parameters, we first calculated β_x , β_z , β_2 , β_3 , σ_c and σ_t . To create one master dataset containing the whole sample, which consists of 5,000 observations, we apply a Cholesky decomposition to create correlated random variables, x and z for any given r_{xz}^c (or r_{xz}^t). We created 2,500 observations for the treatment group and another 2,500 observations for the control group. The outcome variable, y , can be generated based on Model 6. The respondent sample includes those who satisfy Model 1. For each parameter combination, we generate 1,000 master datasets. These 1,000 datasets are analyzed using Model 12. The estimates of bias and *MSE* (mean square error) can be calculated.

Cluster Randomized Controlled Trials (CRT)

Attrition at level 1, level 2, and both levels will be examined using similar models (but taking the hierarchical structure into account) and simulations. These results will also be reported in the proposed presentation.

Findings / Results:

Completely Randomized Controlled Trials (RCT)

Tables 1 and 2 present the findings for bias calculated using Expressions 9, 10, and 11 for the model without a covariate. The parameters used in Table 1 are a subset of parameters used in Table A1 in the WWC Procedures and Standards Handbook. The overall attrition bias is identical with that reported there. Note that, as expected, when the response/attrition rate is same for treatment and control, the part 2 bias is 0. The overall attrition bias comes totally from Part 1 bias, i.e., bias due to the differential correlations between y and z for the treatment and control.

Table 2 presents the results of bias for $r_{yz}^t = r_{yz}^c$. The correlations are the average of r_{yz}^t and r_{yz}^c in Table 1. We can see that Part 1 bias is always 0 because of no differential correlations between y and z for the treatment and control groups. In addition, when the response rate is same, there is no attrition bias.

Table 3 presents the results of bias for the model with a covariate, x , assumed to be a pretest with $r_{yx} = r_{yx}^t = 0.7$ and $r_{xz} = r_{xz}^t = 0.3$. Comparing with Table 1, we find that bias was reduced when $p_c \neq p_t$, which means that including covariate in the model can reduce Part 2 bias.

Conclusions:

This study modeled the sources of attrition bias under various assumptions for completely randomized controlled trials (RCT) and (to be provided by the time of the SREE meeting) cluster randomized controlled trials (CRT). The overall bias is associated with both the overall attrition rate and the differential attrition rate, and both the overall and differential correlations between y and z for the treatment and control groups. In addition, these results show that bias can be reduced by including baseline covariates in the impact estimate model if those covariates are correlated with both the latent propensity to respond and the outcome variable.

Appendix A. References

- Bloom, H. S. (2006). The Core Analytics of Randomized Experiments for Social Research. MDRC Working Papers on Research Methodology. Available online at: <http://www.mdrc.org/publications/437/full.pdf>
- Barr, D. R. & Sherrill, E. T. (1999). Mean and Variance of Truncated Normal Distributions. *The American Statistician*, 53 (4), 357-361.
- Campbell, D. T., & Stantley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: RandMcNally.
- Foster, M. E., Fang, G. Y. (2004). Alternatives to Handling Attrition: An Illustration Using Data from the Fast Track Evaluation. *Evaluation Review*, 28: 434-464.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Orr, L. L. (1999). *Social experiments: Evaluating public programs with experimental methods*. Thousand Oaks, CA: Sage.
- Puma, M.J., Olsen, R.B., Bell, S.H., and Price, C. (2009). *What to Do When Data Are Missing in Group Randomized Controlled Trials* (NCEE 2009-0049). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Rosenbaum, P. R. (1991). Discussing hidden bias in observational studies. *Annals of Internal Medicine*, 115 (11): 901-5.
- Rosenbaum, P. R. (2002). *Observational studies*. 2nd ed. New York: Springer-Verlag.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.
- Shadish, W. R., Hu, X., Glaser, R. R., Kowancki, R., & Wong, S. (1998). A method for exploring the effects of attrition in randomized experiments with dichotomous outcomes. *Psychological Methods*, 3(1), 3-22.
- Shah, A. K., 1985. "A Simpler Approximation for Area Under the Standard Normal Curve," *The American Statistician*, 39(1), 80-80.
- Valentine, J. C., & McHugh, C (2007). The effects of attrition on baseline comparability in randomized experiments in education: A meta-analysis. *Psychological Methods*, 12(3), 268-282.
- U.S. Department of Education, What Works Clearinghouse (2008). *Procedures and Standards Handbook (Version 2)*, Washington: DC, December.

Appendix B. Tables and Figures

Table 1. Attrition Bias for $r_{yz(t)} \neq r_{yz(c)}$ (without covariate)

		Parameter	(1)	(2)	(3)	(4)	(5)	(6)
Pt	Pc	$R^2_{z(t)}$	0.075	0.10	0.15	0.20	0.30	0.50
		$R^2_{z(c)}$	0.05	0.05	0.05	0.15	0.20	0.20
		$r_{yz(t)}$	0.2738	0.3162	0.3873	0.4472	0.5477	0.7071
		$r_{yz(c)}$	0.2236	0.2236	0.2236	0.3873	0.4472	0.4472
0.90	0.90	Overall attrition bias	0.01	0.02	0.03	0.01	0.02	0.05
		Part 1 bias	0.01	0.02	0.03	0.01	0.02	0.05
		Part 2 bias	0.00	0.00	0.00	0.00	0.00	0.00
0.85	0.95	Overall attrition bias	0.05	0.06	0.08	0.08	0.10	0.14
		Part 1 bias	0.01	0.02	0.03	0.01	0.02	0.05
		Part 2 bias	0.04	0.04	0.05	0.07	0.08	0.09
0.70	0.70	Overall attrition bias	0.03	0.05	0.08	0.03	0.05	0.13
		Part 1 bias	0.03	0.05	0.08	0.03	0.05	0.13
		Part 2 bias	0.00	0.00	0.00	0.00	0.00	0.00
0.65	0.75	Overall attrition bias	0.06	0.09	0.13	0.09	0.12	0.21
		Part 1 bias	0.03	0.05	0.08	0.03	0.05	0.13
		Part 2 bias	0.04	0.04	0.05	0.06	0.07	0.09

Note. Entries are biases calculated using Expressions 9, 10, and 11. Part 1 and Part 2 biases were two terms in Expression 9. Pt and Pc are the response rates for the treatment and control group, respectively. The overall attrition bias may not equal the sum of the selection bias and omitted variable because of rounding.

Table 2. Attrition Bias for $r_{yz(t)} = r_{yz(c)}$ (without covariate)

Pt	Pc	Parameter	(1)	(2)	(3)	(4)	(5)	(6)
		$R^2_{z(t)} = R^2_{z(c)}$	0.06	0.07	0.09	0.17	0.25	0.33
		$r_{yz(t)} = r_{yz(c)}$.2487	.2699	.3055	.4173	.4975	.5772
		Overall attrition bias	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.90	Part 1 bias	0.00	0.00	0.00	0.00	0.00	0.00
		Part 2 bias	0.00	0.00	0.00	0.00	0.00	0.00
		Overall attrition bias	0.04	0.04	0.05	0.07	0.08	0.09
0.85	0.95	Part 1 bias	0.00	0.00	0.00	0.00	0.00	0.00
		Part 2 bias	0.04	0.04	0.05	0.07	0.08	0.09
		Overall attrition bias	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.70	Part 1 bias	0.00	0.00	0.00	0.00	0.00	0.00
		Part 2 bias	0.00	0.00	0.00	0.00	0.00	0.00
		Overall attrition bias	0.04	0.04	0.05	0.06	0.07	0.09
0.65	0.75	Part 1 bias	0.00	0.00	0.00	0.00	0.00	0.00
		Part 2 bias	0.04	0.04	0.05	0.06	0.07	0.09

Note. Entries are biases calculated using Expressions 9, 10, and 11. Part 1 and Part 2 biases were two terms in Expression 9. Pt and Pc are the response rates for the treatment and control group, respectively. The overall attrition bias may not equal the sum of the selection bias and omitted variable because of rounding.

Table 3. Attrition Bias with Covariate, x ($r_{yx} = r'_{yx} = 0.7$ and $r_{xz} = r'_{xz} = 0.3$)

		Parameter	(1)	(2)	(3)	(4)	(5)	(6)
Pt	Pc	$r_{yz(t)}$	0.2738	0.3162	0.3873	0.4472	0.5477	0.7071
		$r_{yz(c)}$	0.2236	0.2236	0.2236	0.3873	0.4472	0.4472
0.90	0.90	Overall attrition bias	0.01	0.02	0.03	0.01	0.02	0.05
0.85	0.95	Overall attrition bias	0.02	0.03	0.05	0.05	0.07	0.11
0.70	0.70	Overall attrition bias	0.02	0.05	0.08	0.03	0.05	0.13
0.65	0.75	Overall attrition bias	0.03	0.06	0.10	0.06	0.09	0.19

Note. $N = 5,000$. 1,000 replications. Pt and Pc are the response rates for the treatment and control group, respectively. ES is in terms of pooled standard deviation for the whole sample. The overall attrition bias may not equal the sum of the selection bias and omitted variable because of rounding.