Abstract Title Page

Title:

Comparing Performance of Methods to Deal with Differential Attrition in Lottery Based Evaluations

Authors and Affiliations:

Gema Zamarro, PhD (gzamarro@uark.edu), The University of Arkansas, Fayetteville, AR
Kaitlin Anderson (kaitlina@uark.edu), The University of Arkansas, Fayetteville, AR
Jennifer Steele, PhD (steele@american.edu), American University, Washington, D.C.
Trey Miller, PhD (Trey_Miller@rand.org), RAND Corporation, Santa Monica, CA

Abstract Body

Background / Context:

Since its introduction by Angrist (1990) to evaluate the impact of military service on earnings, a growing literature has made use of lottery-based randomization in the hope to arise at causal effects of diverse educational programs (see, e.g. Rouse (1998); Angrist et al. (2002).; Hoxby and Rockoff (2005); Cullen, Jacob, and Levitt (2006); Hastings, Kane, and Staiger (2010); Abdulkadiroglu et al. (2009); Hoxby and Murarka (2009); Dobbie and Fryer (2009), Engberg et. al. (2014)) among others).

It is common for school districts around the country to use lotteries to determine access to oversubscribed educational programs. Then, those winning the lottery have the possibility of enrolling in the specific program while those non-placed would not have the option to participate in this program but would have multiple other outside options. By comparing average outcomes of lottery winners with average outcomes of those non-placed the hope is to arise at causal effects not affected by bias due to selection into the program.

However, it is not uncommon that students who are not placed by the lottery seek alternative options outside the district, e.g. by choosing a charter, private school, or moving to a different school district instead. For those who leave the school district, it is uncommon to have data of those students and this creates a missing data problem. In particular, if attrition rates differ considerably depending on the lottery status, this creates a differential attrition bias problem jeopardizing the identification of causal effects trough the randomization induced by the lottery. A unique feature of our study is that we were able to complement our school district dataset that suffer from high rates of attrition with State level data, having then an expanded dataset with much lower rates of differential attrition.

Two type of approaches have been used frequently in the literature to try to deal with differential attrition bias: inverse probability weighting methods (Hirano et al, 2003; Busso et al., 2014) and estimation of informative bounds for the treatment effects (Lee, 2009; Angrist et al., 2006). These two methods differ in the assumptions they make to arise at causal effects. Inverse probability weighting methods assume that we have enough observable information that would determine the decision to attrite from the sample. The idea is to weight observations in the data so weighted average characteristics of treated and control students look alike in key observable characteristics. On the other hand, bound estimation approaches (Lee, 2009; Angrist et al., 2006) relax the assumption that we have information on key variables driving attrition decisions and offer the estimate of potential bounds for the treatment effect of interest under, less strict, alternative assumptions about who those who attrite are (e.g. students leaving the district are those with potentially higher outcomes if they were to stay in the district).

Purpose / Objective / Research Question / Focus of Study:

The purpose of this study is to study the performance of different methods (inverse probability weighting and estimation of informative bounds) to control for differential attrition by comparing the results of different methods using two datasets: an original dataset from Portland Public Schools (PPS) subject to high rates of differential attrition, and the expanded PPS and state level dataset that does not suffer as much from differential attrition. The main research questions are:

1. *Do various methods (inverse probability weighting or estimation of informative bounds) adequately compensate for differential attrition in a random assignment evaluation?*

2. *How do various assumptions within these methods affect our results?*

The comparison of the results of estimates provided by the different methods described above on these two datasets will guide our recommendations on the most appropriate methods to be used to correct for common attrition problems.

Setting:

We use data from an evaluation of Dual Language Immersion programs in PPS. PPS uses lotteries to assign access to this program. The original study was subject to high rates of differential attrition, however additional state level data was obtained. With this new dataset, the amount of differential attrition is greatly diminished.

Data Collection and Analysis:

This study utilizes two datasets which will be compared using various methods:

1. PPS school district data: this data set experienced high levels of differential attrition. Attrition in the control group was about 24 percentage points higher than attrition in the treatment group.
2. PPS school district data supplemented with Oregon Department of Education (ODE) State level data. This will serve as the benchmark with which to compare the district-level data that experienced more differential attrition. Once the data was supplemented, differential attrition was reduced to only about 6 percentage points.

Population / Participants / Subjects:

Not applicable.

Intervention / Program / Practice:

Not applicable.

Significance / Novelty of study:

Differential attrition between treatment and control groups is a common problem in social experiments. Inverse probability weighting or bounding methods are used frequently to correct for this problem, but there is little evidence on how successful they are on correcting for differential attrition. We have a unique situation in which the PPS data experienced differential attrition at such a high rate that efforts were made to recover lost data from the ODE. For this reason, the current study is a unique opportunity to verify the effectiveness of these methods under various assumptions.

Statistical, Measurement, or Econometric Model:

In this study, we test the ability of two methods (inverse probability weighting and bound estimation) to correct for attrition bias. Descriptions of each of these two methods are below:

Inverse Probability Weighting:

This method attempts to estimate the average treatment effect for the treated (ATT), or the average effect for those in dual-immersion programs. In these weighting methods, we weight each observation in the control group in a way that creates a better counterfactual for the treatment group. Hirano et al. (2003) find that weighting by the inverse of a non-parametric estimate of the propensity (or probability of being treated) leads to an efficient estimate of the average treatment effect. One limitation on this method, however, is that weighting is most effective when overlap or common support between the treatment and control groups is good, but can perform poorly if overlap is poor (Busso et al., 2014).

Bound Estimation:

The use of bounds relaxes the assumption (in inverse probability weighting) that we have information on key variables driving attrition. Instead, in this method, we estimate potential bounds for the treatment effect of interest under, less strict, alternative assumptions about who those who attrite are (e.g. students leaving the district are those with potentially higher outcomes if they were to stay in the district). Lee (2009) bounds essentially trim the sample in two different ways (from the top of the distribution of test scores or from the bottom of the distribution of test scores) in such that the proportion of observed individuals is the same in the treatment and control groups. This methods makes relatively few assumptions, but in general will create larger bounds than the bounds Angrist et al. (2006) describe leading to the potential that in some situations the bounds are so big that become uninformative. From Angrist et al., (2006) we create estimates using both parametric and non-parametric bounds. These require more assumptions about who attrite, but generally create narrower bounds, relative to Lee bounds. Another potential advantage over Lee bounds is that rather than trimming the sample and losing observations Angrist bound estimates are based on the entire sample.

Usefulness / Applicability of Method:

Randomized controlled trials often have attrition issues, and in some cases, attrition rates can differ considerably between treatment and control groups. This differential attrition bias problem jeopardizes the identification of causal effects trough the lottery-induced randomization. This study will contribute to the research methods literature investigating the effectiveness and properties of two approaches often used to deal with differential attrition bias: inverse probability weighting methods and estimation of informative bounds for the treatment effects.

Either of these methods, if found to be effective at correcting bias due to differential attrition, are easily accessible using statistical packages such as Stata. Further, there are references available for details on applying inverse probability weighting (Stata Manual) and Lee (2009) bounds in Stata (Tauchmann, 2013).

Research Design:

Not applicable.

Findings / Results:

This study is still in the early stages.

Conclusions:

In this study, we evaluate the performance of inverse probability weighting and bounding methods (under various assumptions) to correct for differential attrition. This study represents a unique opportunity in which supplemental data was obtained from the Oregon Department of Education to recover missing student information. Therefore, we have a benchmark against which we compare the performance of these methods. The findings will be novel and relevant to many studies utilizing these methods to correct for the common issue of differential attrition in lottery-based experiments.

Appendix A. References

Angrist, J., Bettinger, E., & Kremer, M. (2006). Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia. *The American Economic Review*, 847-862.

Busso, M., DiNardo, J., McCrary, J. (2014). New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators. *The Review of Economics and Statistics* 96(5): 885-897.

Hirano, K., Imbens, G.W., & Ridder, G. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica* 71(4): 1161-1189.

Lee, D.S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies* 76: 1071-1102.

Stata Manual. *Teffects ipw – Inverse-Probability Weighting.* Retrieved from: http://www.stata.com/manuals13/teteffectsipw.pdf

Tauchmann, H. (2013). *Lee's treatment effect bounds for non-random sample selection – an implementation in Stata.* (SFB 823 Discussion Paper.) Source: https://www.statistik.tu-dortmund.de/fileadmin/user_upload/SFB_823/discussion_papers/2013/DP_3513_SFB823_Tauchmann.pdf

Steele, J.L, Slater, R.O., Zamarro, G., Miller, T., Li, J., & Burkhauser, S. *Effects of Dual-Language Immersion on Student Achievement: Evidence from Lottery Data.* Unpublished Manuscript.