Field Experiments with Individual-Level Randomization: A Risky Venture?

Sarah Humpage

University of Minnesota, Department of Applied Economics

ABSTRACT

In recent years, field experiments have become a popular tool in development economics. Randomized controlled trials (RCTS) are prized for generating unbiased estimates of a program's causal treatment effect, but this method does not address the problem of statistical power. Statistical power is greatly reduced in RCTs when compliance to treatment assignment is imperfect. This research explores why imperfect compliance is likely to be a problem in individual-level randomized experiments of center-based childcare programs. Drop-out in the treatment group is likely because families' demand for preschool is unknown when the sample is constructed, and because this demand is likely to change over time as households experience shocks and as they learn about the center. Non-compliance in the control group arises when children access the program being studied or access alternative preschool programs. This paper uses a recent evaluation of the Hogares Comunitarios childcare program in Guatemala to illustrate the challenges inherent in experimental evaluations of center-based childcare and offers strategies to identify situations in which studies are more likely to succeed.

THE ROLE OF RCTs

RCTs have become increasingly prominent in development and behavioral economics in recent years. They are prized for generating unbiased estimates of a program's causal treatment effect under simple assumptions.

However, RCTs are generally more costly than non-experimental methods that rely on existing data sources. They are also more risky, as individuals may not comply to their random treatment assignment.

RESEARCH ON EARLY CHILDHOOD INTERVENTIONS

Two randomized evaluations of the impact of preschool in the US have been influential: the Perry Preschool Study (Schweinhart et al., 2005) and the Abecedarian Study (Masse & Barnett, 2002). These studies used individual-level randomization to demonstrate that attending preschool has significant short- and long-term benefits.

According to a recent systematic review of research on early childhood development, most research to date has relied on non-experimental methods (Engle et al., 2011). Experimental methods have been limited to either evaluating variations in program characteristics, or to interventions that are not center-based, such as home visits.

The lack of experimental evidence on the impact of center-based care may be due to difficulties in obtaining sufficient statistical power to detect program effects.

RISKS INHERENT IN RANDOMIZED EVALUATIONS OF CENTER-BASED CARE

"Imperfect compliance" to treatment assignment occurs in individual-level randomized trials when an individuals assigned to the treatment group (to participate) choose not to participate, or drop out, or when individuals in the control group gain access to the program.

Non-compliance in the treatment group is likely because participants' real demand for the program is either unknown or unstable at baseline.

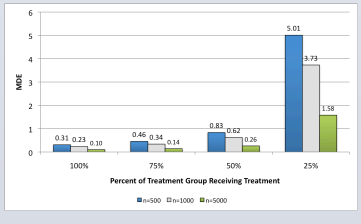
- At baseline, participant demand is unknown to researchers and participants alike because preschool is an "experience good". Participants only have full information once they have experienced it.
- Participant demand is unstable because their need for child care, or the availability of alternative child care arrangements may change.

For individuals in the control group, imperfect compliance occurs if program officials do not exclude them from the program. By the nature of randomization, demand in the control group is equal to demand in the treatment group on average.

CONSEQUENCES OF IMPERFECT COMPLIANCE ON STATISTICAL POWER

The minimum detectable effect size (MDES) measures an estimator's statistical power. This is the smallest effect size that an estimator will be able to detect, given the sample size, the sampling and randomization design, and other parameters. It is also a function of compliance to treatment. A small MDES is desirable as it indicates that the estimator is capable of detecting even small effects.

The graphic below shows how MDES increases as non-compliance in the treatment group increases, holding non-compliance constant in the control group. As it increases, the effect of non-compliance overwhelms the effect of sample size.



CASE STUDY: HOGARES COMUNITARIOS

In 2009, 989 children in 100 communities in Guatemala agreed to participate in an experimental evaluation of the Hogares Comunitarios day care program. Researchers put the 989 study participants into random order. Program officials offered the entire (limited) quantity of spots in the *hogares* to those study participants that appeared first on the list. The remaining study participants were not to be offered enrollment and comprised a control group.

The study was canceled because of non-compliance to treatment assignment in both the treatment and control groups. Because of the non-compliance, the MDES increased to an unacceptable level.

Table 1: MDES by Compliance

	% Enrolled		
	Treatment	Control	MDES
Perfect compliance	100%	0%	0.28 SD
At project end	30%	20%	1.67 SD

CAUSES OF NON-COMPLIANCE

Children in the **treatment group** dropped out or failed to enroll for the following reasons:

- **Unknown & unstable demand for the program.**
  - Some families were never interested in enrolling in day care. They signed up because they thought they would benefit from participating in the study.
  - Some families dropped out because the child did not like the day care. The families did not know their demand for the program until they had experienced it.
  - Parents' employment situations changed, changing their demand for day care during the course of the study.
- **Families can drop out easily.**

Children in the **control group** gained access to the program for the following reasons.

- **There were unanticipated openings in the day cares due to high drop-out in the treatment group.**
- **Day care providers wanted to maintain full enrollment at all times.**
  - They offered spaces to children from the control group when convenient.
- **Providers overrode the random assignment if they felt a child from the control group was needier than a child from the treatment group.**
- **In some cases, providers had not understood what following the random assignment involved.**

RECOMMENDATIONS

Randomized evaluations with the following characteristics may be more likely to succeed.

Desirable program characteristics:

- The program did not exist before the study. This reduces the likelihood that members of the control group will try to gain access to the program.
- Demand for the program exceeds its capacity. This way, an excluded group will already exist. The experimental evaluation only requires that this group be created through random assignment.
- The control group is unlikely to enroll in similar programs (possibly because there are none nearby). This is true of pilot programs that did not previously exist.
- There is an effective exclusion mechanism in place, and few gatekeepers determining who accesses the program. This is easier to manage.

What the researcher can do:

- Assess potential participants' demand for the program at baseline through a simple survey.
- Follow up with individuals that leave the program to boost retention.
- Offer some benefit to the control group that is unlikely to influence outcomes of interest.
- Ensure that program officials, or "gatekeepers", understand the importance of the experimental design and how it works.

REFERENCES

Engle, P., L. Fernald, H. Alderman, J. Behrman, C. O'Gara, A. Yousafzai, M. Cabral de Mello, M. Hidrobo, N. Ulkuer, I. Ertem, S. Iltus and the Global Child Development Sterring Group. 2011. Strategies for Reducing Inequalities and Improving Developmental outcomes for young children in low-income and middle-income countries. The Lancet 378: 8 Oct., 2011.

Masse, L.N. and W.S. Barnett. 2002. *A Benefit Cost Analysis of the Abecedarian early childhood intervention.* New Brunswick, NJ: National Institute for Early Education Research.

Schweinhart, L., J. Montie, Z. Xiang, W. S. Barnett, C. R. Belfield, and M. Nores. 2005. The High/Scope Perry Preschool Study Through Age 40: Summary, Conclusions and Frequently Asked Questions. High/Scope Press.

ACKNOWLEDGMENTS

This work was made possible with funding from the Inter-American Development Bank. The Hogares Comunitarios Study received funding from the Norwegian Trust Fund for Consulting Services and the Inter-American Development Bank.

CONTACT

Email: humpa087@umn.edu

Phone: +1-651-592-9693