

Using Multisite Experiments to Study Cross-Site Variation in Treatment Effects

Howard S. Bloom, Stephen W. Raudenbush, Michael J. Weiss, & Kristin Porter

Multisite randomized trials are experiments where individuals are randomly assigned to alternative experimental arms within each of a collection of sites (e.g., schools). They are used to estimate impacts of educational interventions. However, little attention has been paid to using them to quantify and report cross-site impact *variation*. The present paper, which received the 2017 JREE Outstanding Article Award, provides a methodology that can help to fill this gap.

Why and how is knowledge about cross-site impact variation important?

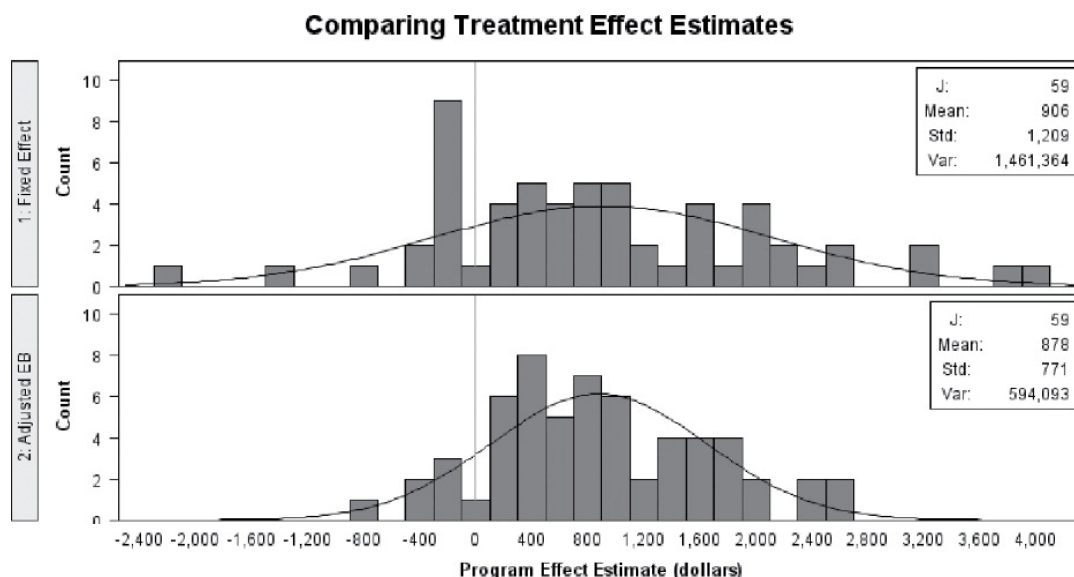
Policymakers need to know both the average impact of an intervention and its variation across settings to properly assess its likely benefits and risks in varied jurisdictions. For social scientists, large cross-site impact variation poses questions about varied implementations of the intervention and varied response to the intervention. Cross-site impact variation also affects sample size requirements for evaluation studies and the generalizability of their findings.

How should cross-site impact variation be quantified and reported?

At each site in a multisite experiment, the difference in the average outcome of program and control group members is an unbiased estimate of the mean impact of the intervention, at that site. However, reporting the cross-site variation in these *impact estimates* (e.g. through a frequency distribution or a standard deviation) can greatly *overstate* the corresponding cross-site variation in *true impacts*. This is because differences in *impact estimates* among sites reflect real cross-site differences in impacts PLUS estimation error. To address this issue, this paper presents a multilevel statistical method that minimizes several biases that can arise in multisite randomized trials. The article then describes how the approach operates, explores its assumptions, and illustrates its application.

How much difference can “doing it right” make?

To demonstrate how to estimate and report true cross-site impact variation, the paper uses data from individually-randomized studies of welfare-to-work programs. Impacts on 69,399 participants’ earnings at 59 welfare centers are estimated. The figure below illustrates the dramatic difference between the resulting cross-site distribution of impact estimates when we do (bottom) and do not (top) account for estimation error (a standard deviation of \$1,209 vs. \$771). For typical evaluations with much smaller site samples, this difference will be even more striking.



Full Article Citation:

Bloom, H. S., Raudenbush, S. W., Weiss, M.J., Porter, K. (2017). [Using Multisite Experiments to Study Cross-Site Variation in Treatment Effects: A Hybrid Approach with Fixed Intercepts and a Random Treatment Coefficient](#). *Journal of Research on Educational Effectiveness*, 10(2), 817-842, DOI: 10.1080/19345747.2016.1264518.