IMPACT EVALUATION METHODS

WHAT ARE THEY AND WHAT ASSUMPTIONS MUST HOLD FOR EACH TO BE VALID?

J-PAL
ABDUL LATIF JAMEEL POVERTY ACTION LAB
NORTH AMERICA

QUASI-EXPERIMENTAL METHOD

METHODOLOGY	DESCRIPTION	WHO IS IN THE COMPARISON GROUP?	THE METHODOLOGY IS ONLY VALID IF...
Pre-Post (Before-and-after)	Measure how program participants improved (or changed) over time.	Program participants themselves—before participating in the program.	The program was the **only** factor influencing changes in the outcome over time. If the program did not exist, outcomes would be the same before and after the study period.
Simple Difference	Measure the difference between program participants and non-participants after the program is completed.	Individuals who did not participate in the program (for any reason), but for whom data were collected after the program.	Non-participants and participants were equally likely to enter the program before it started. Non-participants are identical to participants, except they did not participate in the program.
Differences in Differences	Measure the before-and-after change in outcomes for the program participants, then subtract the before-and-after change in outcomes of the non-participants to find the *relative* change in outcomes for program participants.	Individuals who did not participate in the program (for any reason), but for whom data were collected both before and after the program.	If the program had not existed, the participants and non-participants would have experienced identical trajectories during the study period. Any differences in characteristics between the treatment and control group do not have more or less of an effect over time on outcomes.
Multiple Linear Regression	Compare participants to non-participants, and estimate the effects of the program by adjusting for observed characteristics (income, age, gender) that might explain differences in outcomes between participants and non-participants.	Individuals who did not participate in the program (for any reason), but for whom data were collected both before and after the program.	The characteristics that were *not included* (because they are unobservable/or have not been measured) either do not affect the outcome **or** do not differ between participants and non-participants.
Statistical Matching	Individuals who received a program are compared to similar individuals who did not receive it.	**Exact matching:** For each participant, at least one non-participant who is identical *along a selected list of known characteristics*. **Propensity score matching:** For each participant, a non-participant with the same likelihood of participating, as predicted by known characteristics such as age, gender, and occupation.	The characteristics that were *not included* (because they are unobservable or have not been measured) either do not affect outcomes **or** do not differ between participants and non-participants.

METHODOLOGY	DESCRIPTION	WHO IS IN THE COMPARISON GROUP?	THE METHODOLOGY IS ONLY VALID IF...
QUASI-EXPERIMENTAL METHOD			
Regression Discontinuity Design	Individuals are ranked or assigned a score based on specific, measureable criteria. A cutoff determines whether an individual is eligible to participate in the program. Participants who are just above the cutoff are compared to non-participants who are just below the cutoff.	Individuals who are close to the cutoff, but fall on the "wrong" side of that cutoff, and therefore do not get the program.	After adjusting for the eligibility criteria (and other observed characteristics), the individuals directly below and directly above the cut-off score are statistically identical. The cutoff criteria must have been strictly adhered to. The cutoff must not have been manipulated to ensure that certain individuals qualify for the program.
Instrumental Variables	Participation can be predicted by an almost random factor. This "instrumental" factor only affects the outcome by way of predicting whether an individual participates in the program (and participation affects the outcome). Currie and Yelowitz (1999)[1] used having children of different genders as an instrument for public housing; families with two children of different genders were entitled to larger units in public housing, making them more likely to participate in the program. The gender of children was basically random— and that random factor predicted participation in public housing.	Individuals who, because of this "instrumental" factor, are predicted not to participate and (possibly as a result) did not participate.	The "instrumental" factor predicts the outcome, and if it weren't for the "instrumental" factor's ability to predict participation, this "instrumental" factor would otherwise have no effect on the outcome.
EXPERIMENTAL METHOD			
Randomized Evaluation	Random assignment (e.g. a coin toss or random number generator) determines who may participate in the program so that those assigned to participate in the program are, on average, the same as those who are not, in both observable and unobservable ways. Since the participants and nonparticipants are comparable, except that one group received the program, any differences in outcomes result from the causal effect of the program.	Participants who are randomly assigned to not participate in the program. This is often called the "control" group.	Randomization "worked" and the two groups are statistically identical (on observed and unobserved factors). The effects of the treatment do not spill over to the control group. Any behavioral changes are driven by the program— not by the evaluation itself, or by the fact that the participants or non-participants are being studied. If outcome data are missing, data for the same types of individuals are missing from both the control and treatment groups.

[1]Currie, Janet, and Aaron Yelowitz. "Are public housing projects good for kids?." *Journal of public economics* 75, no. 1 (2000): 99-124.