Impact Evaluation Methodology Crib Sheet

Methodology	Description	When this might be the appropriate design for you	Who are you comparing	Assumptions for the design to be viable	Required Data
Individual-level RCT	In this experimental design, individuals from a sample of your population of interest are randomly assigned to either a treatment group or a comparison group. Note that the comparison group can be another treatment group rather than a pure control group, either treatment and control or treatment A and treatment B. Differences in the outcomes of interest between the treatment and comparison groups can be causally attributed to treatment status. Individuals can also be assigned to more than two groups if the researcher is interested in comparing multiple treatments. *Randomization is the most rigorous impact evaluation design methodology.*	When it is possible to randomize the offering of a program, and it is possible to offer the program to individuals (rather than groups of individuals e.g. schools, villages).	Individuals in one group (treatment A) with individuals in another group (control or treatment B)	Randomization is the most rigorous impact evaluation design methodology so there no strong assumptions. Assignment to either group is random; there are no spillovers in treatment between groups.	Outcome data for treatment and comparison groups. Baseline data can be helpful but is not required.
Cluster-level RCT	Similar to the individual RCT design, cluster-level RCT designs assign subjects to two separate treatment groups (again, no pure controls are necessary). Earlier, individuals receive treatment, but for cluster RCTs, groups, or clusters, of individuals are randomized to treatment groups. As such in a given cluster, all individuals are randomized to the same group. Randomization is the most rigorous impact evaluation design methodology, although	When it is possible to randomize the offering of a program, however, it is either not logistically possible to offer to random individuals or you are worried about individuals within the same group being influenced by the program even if they don't receive it. In the	Clusters of individuals in one group (treatment A) with clusters of individuals in another group (control or treatment B)	Randomization is the most rigorous impact evaluation design methodology so there no strong assumptions. Assignment to either group is random; there are no spillovers in treatment between groups.	Outcome data for treatment and comparison groups. Baseline data can be helpful but is not required.

Methodology	Description	When this might be the appropriate design for you	Who are you comparing	Assumptions for the design to be viable	Required Data
	randomizing clusters implies a larger sample size will be required.	case, you can randomize at a higher level (e.g. schools, villages).			
Encouragement Design	In this variant of the RCT, one group is randomly assigned to receive an "encouragement" to take up the intervention and another group is randomly assigned to not receive the encouragement. The idea is that the group assigned to receive the encouragement will be more likely to take up the intervention than those in the group assigned not to receive the encouragement. Randomization is the most rigorous impact evaluation design methodology, although an encouragement design will require a much larger sample size than an individual-level RCT.	When you do not have direct control over whether people take up a program or you cannot logistically exclude people from the program.	All individuals assigned to the encouragement group to individuals assigned to the non encouragement group. Note that this design your findings are only relevant for the type of individual that responds to that type of encouragement, not the average population.	The encouragement should only influence an individual's likelihood to take up the program, not the outcomes you are trying to measure.	Outcome data for treatment and comparison groups. Baseline data can be helpful but is not required.
Phase-in Design	In this variant of the RCT, the program has a phased roll-out and individuals or clusters are randomized into which phase they will determine the program.	When you want all individuals/areas to receive the program by the end of the study period.	Individuals in one group (treatment A) with individuals in another group (control or treatment B) in each phase.	Implementers are able to adhere to the randomized roll-out. Additionally, individuals should not change their behavior in expectation that they will receive the program in the future.	Outcome data must be collected for all treatment groups across each phase. Phases must be equal in length so that data collection happens at equally-spaced intervals.
Statistical Matching	In this quasi-experimental design, a comparison group is constructed by identifying individuals and matching them to individuals in the treatment group based on observable characteristics.	When you are not able to randomize the program but have idenfitied a group of individuals that are similar to those who	Individuals in the treatment group to individuals in the matched comparison group.	Individuals in the treatment group must be statistically similar to those in the constructed	Both baseline data and outcome data is required for this design. Baseline data is needed to

Methodology	Description	When this might be the appropriate design for you	Who are you comparing	Assumptions for the design to be viable	Required Data
		have received the program (and have baseline data on them).		comparison group both on observable and unobservable characteristics. In other words, there should be nothing that would influence outcomes of the two groups differently besides participation in the program.	match individuals as they should be matched on characteristics measured prior to the start of the program.
Difference-in-difference	This quasi-experimental design compares the changes in outcomes before and after a program between two treatment groups (no pure control necessary) at two different times.	You are not able to randomize the program but you have identified a group of individuals that are similar to those that have received the program (and have baseline data on them). While the two groups might not be similar on average (as they would be for Statistical Matching), you expect that their outcomes would have changed at the same rate without the program.	Two similar populations, one receiving an intervention and another serving as a comparison group, both before and after the intervention.	This design requires a strong assumption, called the parallel trends assumption: without the intervention, the outcomes of the two groups would have changed at the same rate over time.	At a minimum baseline and endline data is needed for the two groups. However, it is strongly encouraged to have multiple measures of pre-program data to assess the parallel trends assumption.
Regression Discontinuity	This quasi-experimental design leverages a cut-off point used to determine whether individuals are eligible for a program. Individuals right below the	Your program has eligibility criteria for inclusion and there is a	The statistically similar groups who lie either directly above or	Individuals directly above and directly below the cut-off are	Outcome data as well as the selection score for

Methodology	Description	When this might be the appropriate design for you	Who are you comparing	Assumptions for the design to be viable	Required Data
	cut-off are compared to individuals right above the program.	continuous selection score that determines whether an individual receives a program or not.	directly below the cut-off point.	statistically similar. Individuals cannot self-select into being above or below the cut-off.	each individual. Baseline data is useful to verify that individuals above and below the cut-off are statistically similar.
Instrumental Variables	In this quasi-experimental design, participation in an intervention can be predicted by a random factor or instrumental variable. This factor also needs to be unrelated to the outcomes of interest.	When you have identified a factor that predicts participation in the program but does not itself influence the outcome of interest.	Individuals who participated because of this random factor to individuals who did not participate to those who are predicted not to participate because of this same factor.	This design requires a strong assumption called the exclusion restriction. Not only do you need to identify a factor that predicts well whether individuals take up the program, this factor cannot be related to outcomes of interest.	Outcome data and data on the instrumental variable.

This document was made possible by the generous support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of the Humanitarian Assistance Evidence Cycle (HAEC) and do not necessarily reflect the views of USAID or the U.S. Government.