

Large Population Level Energy Efficiency Program Design Method Checklist

A well-designed and carefully executed program design and measurement and verification (M&V method) is necessary to measure the true impact of a program on its participants. If done well (e.g., using randomized controlled trials), this approach can use metered data to estimate how much participants' energy consumption changed due solely to a program – free from the impact of other factors that may have influenced participants' energy usage. The estimated results will be the “net” savings of a program and make debates about net-to-gross ratios less cumbersome. With rigorous M&V results, the estimated impact of a program can be reported in the regulatory process and to current and prospective program participants.

The program implementer needs to make critical M&V-related decisions prior to finalizing program design and launching the program. When applying the most rigorous methodology, the implementer should work with an evaluator¹ to first select a representative sample from a target population, which is then split into a treatment group (customers who receive the energy efficiency program) and a control group (those who don't). While the target or representative sample can be very diverse, it is important that **the control and treatment groups** be as similar as feasible on average with the sole exception that one group participates in the program and the other does not. In an ideal setting, *randomly selecting a study sample* from a target population ensures representativeness and *randomly assigning treatment and control groups* enables an implementer to pinpoint the impact of a program and eliminate other influences. This allows implementer to conclude that any observed differences in consumption result from the energy-efficiency program being assessed.

With random assignment, the implementers can construct an accurate counterfactual: “What would have happened to energy consumption in the absence of the program?”

Rigorous design and M&V results will provide valuable information on which programs are working and which may not be having the intended impact. By carefully measuring the impact of a program, implementers can observe details about participants' responses to the program, thereby gaining valuable insights for future design changes.

Below is a checklist to guide a third-party energy efficiency implementer on the selection of the most rigorous and appropriate methodology for the desired program approach, and how the program should be designed to ensure it can be result in claims that can be independently evaluated by Commission staff using the identified method. The checklist provides a high-level overview of the main factors that need to be considered and decided in developing the M&V methodology. It is meant to guide the implementer and the implementer contracted evaluation expert in the proposal stage. The checklist also addresses questions about what data are to be collected.

CHECKLIST

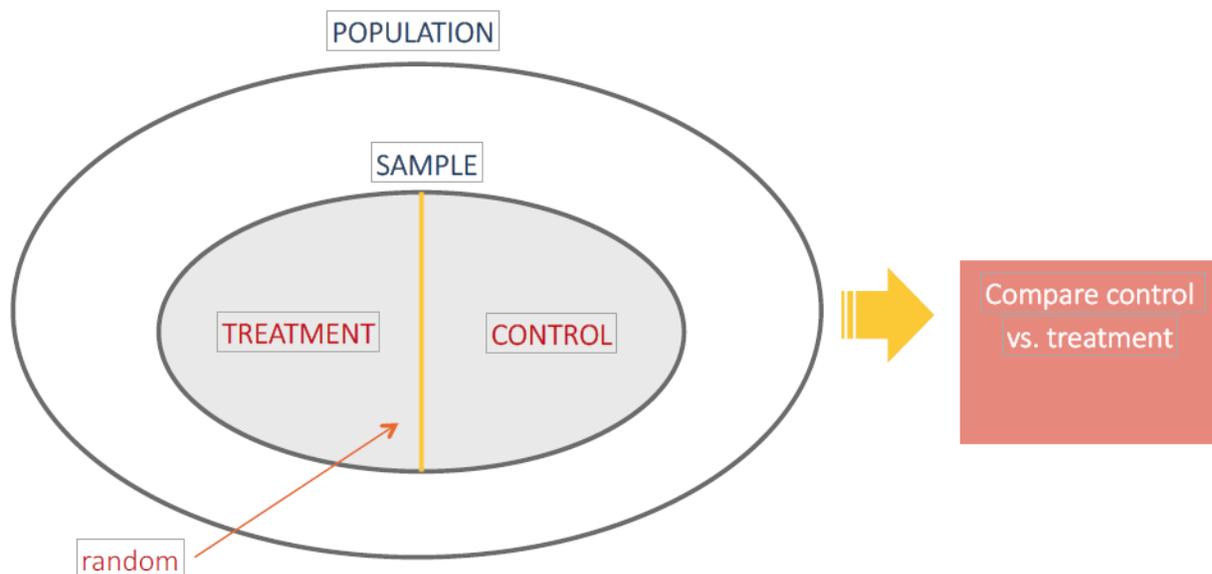
Step 1: Determine the appropriate design, measurement, and verification methodology to be used using the list below. Starting with the first and most rigorous methodology, explain your selection. If a less rigorous methodology is chosen, explain why the more rigorous methodologies were not chosen. More information on each methodology is provided in the link for each below.

¹ See decision D.18-01-004 for more rules for engaging evaluators.
<http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M205/K560/205560586.PDF>

- A. [Randomized Controlled Trial \(RCT\)](#) - RCTs are a well-known form of experimental research. They are the standard among scientists evaluating the effectiveness of treatments and have been used to measure results of social programs (education, health, poverty, etc.). RCTs reduce the risks of biases that can invalidate a study, such as having skewed results because those who enroll in a program are systematically different than those who don't.
- B. [Randomized Encouragement Design \(RED\)](#) – Randomly assign an inducement to participate in a program, e.g., rebates, information, within the treatment group.
- C. [Recruit and Delay/Deny](#) - Encourage participants from the treatment group to participate but either phase in their participation in stages or subject actual treatment to a lottery.
- D. [Regression Discontinuity \(RDD\)](#) - When a threshold determines who receives treatment, RDDs compare those just above and just below the threshold.
- E. [Differences-in-Differences](#) - Compares groups over two time periods: before either group receives the treatment, and after one group receives the treatment.

Step 2: Specify the population that the program is targeting. Who will be eligible for the program in the long run?

Step 3: Determine how the representative sample will be selected from the target population. The figure below illustrates the difference between the target population and the sample for the program.



Step 4: What is the strategy for identifying treatment and control groups within the program sample? The treatment group of the sample will receive the program and the control group will not, at least through the duration of the study. If the evaluation approach does not include a treatment and control group, please explain why it is not feasible.

Step 5: Perform statistical power tests to ensure that the proposed evaluation sample is sufficiently large to detect economically meaningful impacts.

Step 6: Describe how utility metered data will be used in the evaluation. Verify that metered data are capturing the information your program requires. For example, verify that the meter data are matched to the site /location where the program is expected to have an impact.