The Case for Randomized Controlled Trials in Evaluating Energy-Efficiency Programs

California's energy-efficiency programs have been well vetted—their results have been better evaluated and measured and their usefulness better verified than similar programs in many states. The state's analyses have provided critical insights into costs and benefits. Even so, many of these studies have used methodologies that inhibit calculation of each program's *true net energy savings*. The most accurate methodology to assess the impact of an energy efficiency program is a randomized controlled trial (RCT). As new energy efficiency programs are developed, their implementation could be structured to ensure that impacts can be assessed accurately using RCTs or similar methods.

Observational Studies: A Catch-All of Effects

Many empirical estimates of energy savings—analyses not based on engineering models—come from observational studies. Some of these studies compare energy usage before and after consumers invest in energy efficiency. This approach overlooks *unobservable* influences that may affect energy consumption. Around the time of an efficiency investment, for example, the number of people living in a house may have increased or a consumer may have changed jobs, necessitating a shorter commute and less time away from home. Both these changes would lead to more energy consumption, causing the study to understate the impacts of the efficiency investment.

To adjust for changes that can affect energy consumption, some studies use consumers who made no efficiency investment as a benchmark. But picking a truly comparable group is difficult. Consumers who invest in energy efficiency may differ systematically from those who don't. They may use their air conditioners or clothes driers less, or they may be more careful about turning off lights. Without a careful research design, investigators can't attribute observed changes in behavior to an efficiency program. Consumers might have made the changes regardless. Net savings estimates may thus be exaggerated.

The Power of Randomized Controlled Trials

Accurately measuring the impact of an energy-efficiency program requires that the research design involve a treatment group (people who receive the benefits of the energy efficiency program) and a control group (ones who don't). Ideally, these groups should be as similar as feasible—nearly identical, if possible—except for the fact that one participates in the program and the other doesn't. Good treatment and control groups enable an investigator to pinpoint the impact of a program and eliminate other influences. Crucially, experimental designs randomly allocate subjects—households, schools, etc. —to treatment and control groups. This increases the chances that any observed differences result from the energy-efficiency program being assessed. With random assignment, the researcher can construct an accurate counterfactual — "What would have happened in the absence of the program?" With a robust counterfactual, good experiments can show causality with a high degree of confidence.

Randomized controlled trials (RCTs) are a well-known form of experimental research. They're the standard, for example, among scientists evaluating the effectiveness of medical treatments. They've also been used to evaluate social programs, like anti-poverty efforts in the developing world. RCTs reduce the risks of biases that can invalidate a study. Selection bias, for example, is the risk that the people receiving an intervention differ from those in the control group. And confounding bias is the risk that an additional unappreciated factor is causing an outcome.

In medical research, randomized controlled trials have overturned observational studies—and probably saved lives. Observational studies, for example, suggested that hormone replacement therapy posed no health risks to women. But those studies couldn't fully control for all differences between women who took hormones and those who didn't. When researchers performed an RCT, they saw that women in the treatment group had a higher incidence of heart attacks and some cancers. The results were so decisive that the researchers ended the RCT early and publicized the results.

How to Design an RCT

A randomized controlled trial for an energy efficiency program can be designed in many ways, but the treatment and control groups should be identified at the outset. For new programs, the treatment group might be a randomly selected set of households within a utility's service territory participating in a pilot. The control group would then be households that did not participate. If the evaluation shows that a program produces net energy savings, it can be rolled out to all households.

For more targeted programs where participants must meet qualifying criteria and funding is limited, such as a free home weatherization program, a treatment group could be determined by randomly choosing sufficient qualifying applicants to exhaust the first year of funding for the program. The control group would then consist of all qualifying applicants who do not receive the treatment. Not everyone can receive the weatherization simultaneously, so this avoids potential claims of discrimination about whose home gets weatherized first. Random assignment of implementation dates to households permits the identification of verifiable treatment and control groups. Once the trial ends, everyone can receive weatherization.

Quasi-experimental Research Designs

If it isn't feasible to determine treatment and control groups at the outset, investigators can sometimes compensate and produce estimates of program impacts after the fact. But the results won't be as reliable as those from an RCT. These kinds of studies are called "quasi-experimental" designs, and their effectiveness depends on the particulars of the program and data available. Quasi-experimental studies assign households to treatment and control groups by a method other than random assignment. People might end up in a treatment group based on a cut-off date or where they live.

An example of a quasi-experimental design is a recent study of California's 20/20 electricity rebate program run by all three investor-owned utilities in the summer of 2005. In 20/20, eligible households were automatically enrolled and could save 20 percent on their summer electricity bills if they cut their 2005 summer consumption by 20 percent relative to the previous summer. To be eligible, a household had to have started its electric service by a specific date in 2004. The study compared households that had opened accounts in the 90 days before the cut-off date with those that opened accounts in the 90 days after it. The eligibility rule generated a near random assignment, and customers in both groups were likely to have similar patterns of electricity consumption. Since enrollment was automatic, there was no selection bias in the households in the treatment group. The study suggested that the program did little to reduce energy consumption, though its results were stronger in inland areas and among low-income consumers.

For more information on randomized controlled trials for energy and social science programs see:

- U.S. Department of Energy Smart Grid Investment Grant Technical Advisory Group Guidance Document #7, Topic: Design and Implementation of Program Evaluations that Utilize Randomized Experimental Approaches

- Evaluation, Measurement, and Verification (EM&V) for Behavior-Based Energy Efficiency Programs: Issues and Recommendations, by Annika Todd, Elizabeth Stuart, Steven R. Schiller, and Charles Goldman, Lawrence Berkeley National Laboratory, May 2012.