



SEE Action

STATE & LOCAL ENERGY EFFICIENCY ACTION NETWORK

Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations

June 25, 2012

Michael Li *US Department of Energy*

Annika Todd *Lawrence Berkeley National Lab*



SEE Action

STATE & LOCAL ENERGY EFFICIENCY ACTION NETWORK

This information was developed as a product of the State and Local Energy Efficiency Action Network (SEE Action), facilitated by the U.S. Department of Energy/U.S. Environmental Protection Agency. Content does not imply an endorsement by individuals or organizations that are part of SEE Action working groups, or reflect the views, policies, or otherwise of the federal government.

Outline: Evaluation of Behavior-Based Programs

- What is a behavior-based energy program?
- Why is evaluation of these programs hard?
- Why is *designing* a program as a “randomized controlled trial” (RCT) so important?



What is a behavior-based energy program?

- Programs that affect the way that consumers use energy without using traditional methods, such as prices and rebates
- Instead, use simple psychological levers or information to change behavior
- Example 1: Comparing your energy use with your neighbors
- Example 2: Providing real-time information about energy use
- Other examples:
 - Competitions, rewards: Turning energy use into a game
 - Education / Outreach: Information about energy behavior
 - Display of feedback: Simplify / Framing



What are the potential benefits and concerns?

- Potential Benefits
 - In theory, potentially cheap to implement and result in significant energy savings → cost effective
 - As a result, increasingly being adopted nationwide
- Potential Concerns
 - In reality, these programs are relatively new and evidence of energy saving effects is unclear
 - Potential for unsubstantiated claims



Why is evaluation crucially important?

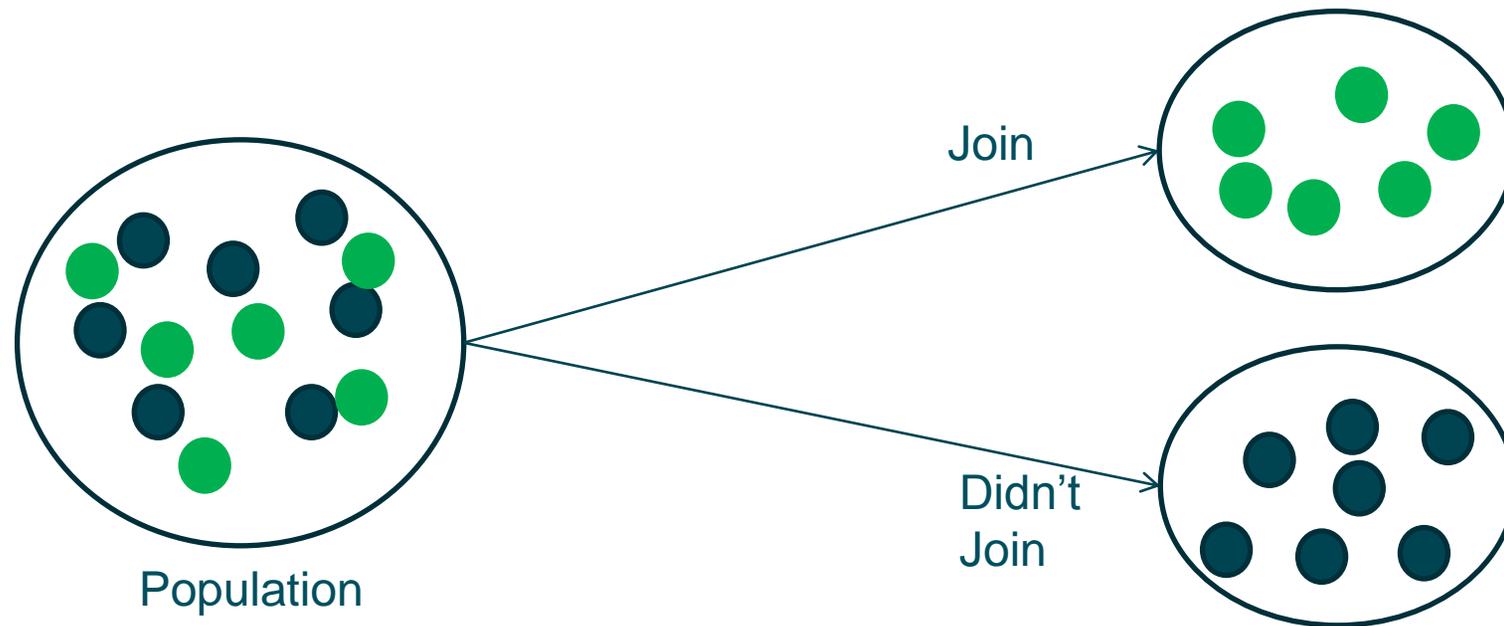
→ It is very important to measure effect of these programs

- Need to gain information about how well different types of programs work
- Are the estimates energy savings valid for utilities to claim savings?



Why is evaluation of these programs hard?

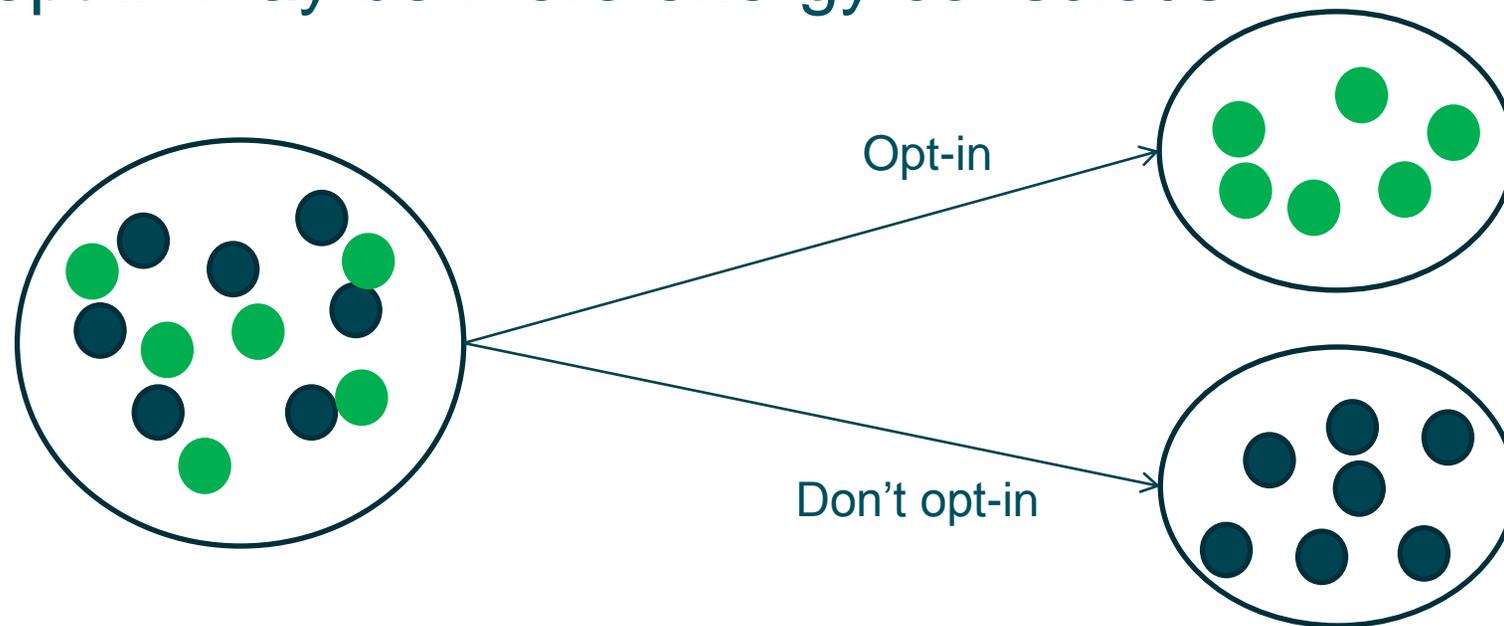
- Strong problems of “selection bias”: households who join (choice, screening) are fundamentally different



- Observed differences might be due to program; might difference between groups
- Selection bias can skew the results of the evaluation

Why is evaluation of these programs hard?

- Energy programs “selection bias”: households who opt-in may be more energy conscious



- Observed difference in energy use might be due to the program; but might be difference between groups

Why is evaluation of these programs hard?

- It may be more difficult to measure the impact of behavior-based programs correctly (in contrast to other programs such as appliance rebates)
 - Impacts vary significantly between households
 - Within a household, hard to disentangle changes in overall energy usage between program, other factors
 - Savings are relatively small: often 1-5%, so if an evaluation is biased, large implications



Why is evaluation of these programs hard?

→ **Bad evaluation could lead to bad policy decisions**



SEE Action Report

- “Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs”
- Provides guidelines and best practices for
 - Program design
 - Program analysis and evaluation given design
 - Provides rankings for different methods
- Target audiences:
 - Senior managers responsible for overseeing and reviewing efficiency program designs and evaluations
 - Practitioners, evaluation professionals, and staff responsible for reviewing efficiency program designs and evaluations



Why is *designing* a program as a randomized controlled trial (RCT) so important?

- Primary recommendation – a well designed, RCT program results in:
 - Transparent, straightforward analysis
 - Robust, accurate, valid program impact estimates
 - High degree of confidence in program effectiveness
- Why?
 - RCT means that households are assigned to the program randomly (as opposed to household choice or screening criteria)
 - Solves selection bias



Why is *designing* a program as a randomized controlled trial (RCT) so important?

- If RCTs are not feasible, recommendations for acceptable “quasi-experimental” methods
 - More opaque, difficult, complex analysis
 - Quasi-experimental methods try to correct for selection bias
 - Lower degree of confidence in validity of savings estimates



Other Key Recommendations

- **Problem:** Potential conflicts of interest
 - **Recommendation:** Third-party evaluator transparently defines and implements program evaluation, assignment to control and treatment groups, data selection
- **Problem:** The same savings may be claimed by two programs (e.g., a behavioral program & appliance rebate program both claim savings from appliances)
 - **Recommendation:** Estimate and account for this “double counted savings” overlap to the extent possible by comparing control to treatment groups



Recommendations for the Future

- The hope is that in the future, we will have conclusive evidence about the effectiveness of different types of behavior-based programs
- Move away from RCTs
- We are not yet at this point...



Questions?

- Main point: evaluation of behavior-based programs is difficult, but if the program is designed in the right way (using a RCT) then we can be confident that the evaluation of the program's energy savings is valid
- Many guidelines and technical recommendations in the report:
 - SEE Action website, www.seeaction.energy.gov
 - Lawrence Berkeley National Lab website: behavioranalytics.lbl.gov

Mike Li: Michael.Li@hq.doe.gov

Annika Todd: atodd@lbl.gov



www.seeaction.energy.gov