Evaluation of Residential Behavior-Based Programs

September 21, 2016

EM&V Webinars Facilitated By:
Lawrence Berkeley National Laboratory
https://emp.lbl.gov/emv-webinar-series

With Funding From:
U.S. Department of Energy's Office of Electricity Delivery and Energy Reliability-
Electricity Policy Technical Assistance Program

In Collaboration With:
U.S. Environmental Protection Agency

National Association of Regulatory Utility Commissioners
National Association of State Energy Officials

Introduction

◆ LBNL is supported by the U.S. Department of Energy to conduct non-classified research, operated by the University of California

◆ Provides technical assistance to states—primarily state energy offices and utility regulatory commissions

The presentation was funded by the U.S. Department of Energy's Office of Electricity Delivery and Energy Reliability-National Electricity Delivery Division under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231.

Technical Assistance

◆ LBNL's provides technical assistance to state utility regulatory commissions, state energy offices, tribes and regional entities in these areas:

- ❑ Energy efficiency (e.g., EM&V, utility programs, behavior-based approaches, cost-effectiveness, program rules, planning, cost recovery, financing)

- ❑ Renewable energy resources

- ❑ Smart grid and grid modernization

- ❑ Utility regulation and business models (e.g., financial impacts)

- ❑ Transmission and reliability

- ❑ Resource planning

- ❑ Fossil fuel generation

◆ Assistance is independent and unbiased

◆ LBNL Tech Assistance website: https://emp.lbl.gov/projects/technical-assistance-states

◆ US DOE Tech Assistance gateway: http://energy.gov/ta/state-local-and-tribal-technical-assistance-gateway

Webinar Series

◆ Webinars designed to support EM&V activities for documenting energy savings and other impacts of energy efficiency programs

◆ Funded by U.S. DOE in coordination with EPA, NARUC and NASEO

◆ Audience:

- ❑ Utility commissions, state energy offices, state environment departments, and non-profits involved in operating EE portfolios
- ❑ Particular value for state officials starting or expanding their EM&V
- ❑ Evaluation consultants, utilities, consumer organizations and other stakeholders also are welcome to participate

◆ For more information (upcoming and recorded webinars, EM&V resources) see:

- ❑ https://emp.lbl.gov/emv-webinar-series
- ❑ General Contact: EMVwebinars@lbl.gov

Series Contact:
Steve Schiller
Senior Advisor, LBNL
SRSchiller@lbl.gov

Next Webinar

◆ Transmission and distribution (T&D) efficiency programs and their EM&V – Scheduled for October

◆ More webinars coming for November and beyond…

Today's Webinar

Residential behavior-based (BB) programs use strategies grounded in the behavioral and social sciences to influence household energy use. These programs have unique evaluation challenges. Today we will cover:

◆ **The Basics**

  ◆ Description of BB programs

  ◆ Basics and options for evaluating BB programs,

  ◆ BB program resources

  ◆ Example processes and results from BB program evaluations

  C. Anna Spurlock, Senior Scientific Engineering Associate, Berkeley Lab

  Annika Todd, Principal Scientific Engineering Associate, Berkeley Lab

◆ **Lessons learned from a regulators point of view when evaluating BB programs**

  Bill Saxonis, Utility Supervisor/Evaluation, New York Department of Public Service

◆ **Q&A with panelists**

Lawrence Berkeley National Laboratory
Environmental Energy Technologies
Division **Behavior Analytics**
Providing insights that enable evidence-based, data-driven decisions

Behavior-Based programs and Evaluation:
New (and old) methods

Annika Todd, Anna Spurlock

September 2016

Outline

- What is a BB program?

- Why is rigorous evaluation important, and why is it hard?

- Best evaluation practices
  - Resources: SEE Action, UMP 2014, UMP forthcoming, LBNL
  - Randomized controlled trials (RCTs) are "gold standard"

- Recent advances in evaluation
  - EMV 2.0: High frequency data, new analytics
  - Compare RCTs to other methods in the real world

Outline

- **What is a BB program?**

- Why is rigorous evaluation important, and why is it hard?

- Best evaluation practices
  - Resources: SEE Action, UMP 2014, UMP forthcoming, LBNL
  - Randomized controlled trials (RCTs) are "gold standard"

- Recent research
  - EMV 2.0: High frequency data, new analytics
  - Compare RCTs to other methods in the real world

What is a behavior-based program?

Behavior-based programs are those that utilize strategies intended to affect consumer energy use behaviors in order to achieve energy and/or peak demand savings.

- Traditionally, BB programs include outreach, education, competition, rewards, benchmarking and/or feedback elements.

What is a behavior-based EE program?

- Example 1: Comparing your energy use with your neighbors

- Example 2: Providing real-time information and feedback about energy use

- Example 3: Goal setting and reward points per kWh saved

What is a behavior-based program?

Behavior-based programs are those that utilize strategies intended to affect consumer energy use behaviors in order to achieve energy and/or peak demand savings.

- **But** – isn't *everything* behavior? Don't all programs rely on people changing energy decisions?
  - LED programs – people have to actually install and use them!
  - Pricing programs – people have to change their behavior in some way! (Re-program thermostat, move dishwasher to later)

What is a behavior-based EE program?

- Example 1: Behavioral Demand Response
  - Compare usage to neighbors *during peak hours* in order to reduce peak-load
    - Behavior
    - Demand Response
    - Also EE! Spillover – reduce during off-peak hours

- Example 2: Critical Peak Pricing, higher prices during peak hours on critical event days
  - Demand Response: reduce peak load
  - Behavior: re-program thermostat
  - Energy Efficiency: spillover – reduce during off-peak hours and non-critical days

What is a behavior-based EE program?

- Lines are blurred between EE, DR, and behavior

- Adopt a more holistic approach, where programs aren't siloed into BB, EE, DR?

Outline

- What is a BB program?

- **Why is rigorous evaluation important, and why is it hard?**

- Best evaluation practices
  - Resources: SEE Action, UMP 2014, UMP forthcoming, LBNL
  - Randomized controlled trials (RCTs) are "gold standard"

- Recent research
  - EMV 2.0: High frequency data, new analytics
  - Compare RCTs to other methods in the real world

Why is rigorous evaluation important, and why is it hard?

→ It is very important to accurately evaluate the effectiveness of programs

- Cost effectiveness
- For planning purposes – want to select optimal program portfolios
- For validly claiming energy savings

Why is evaluation of these programs hard?

- Strong problem of **"Selection Bias"**: households that join (e.g., opt-in, screened) are fundamentally different
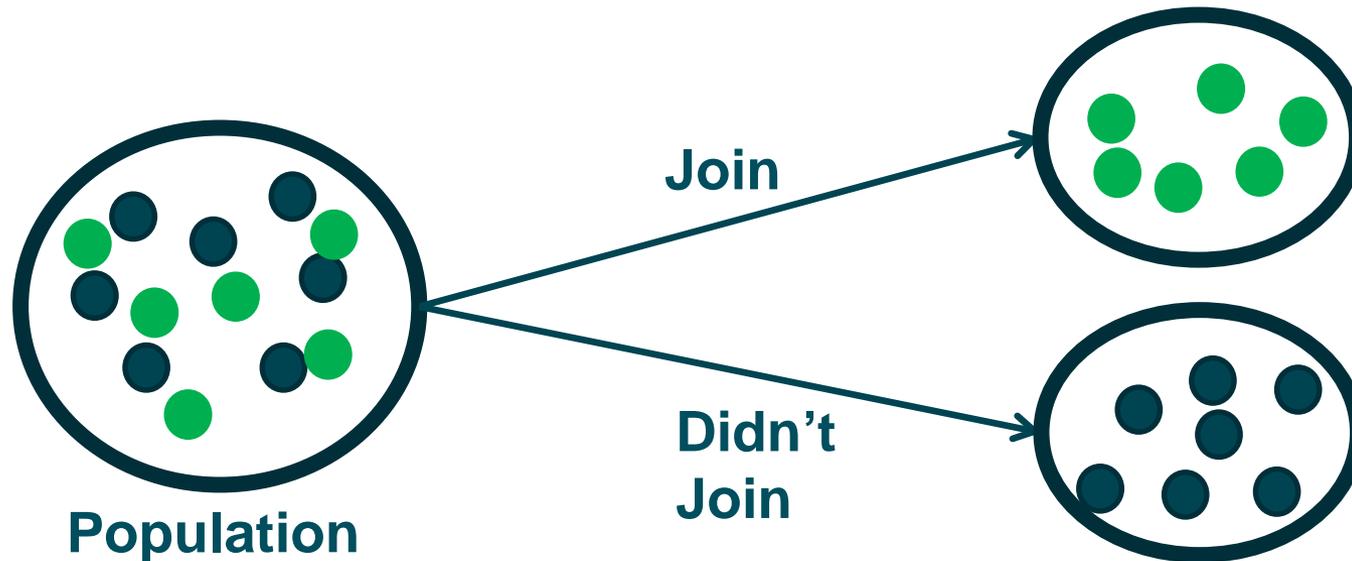
- Observed differences *might* be due to program, but might just be a difference between groups
- Selection bias can skew the results of the evaluation

Why is evaluation of these programs hard?

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

- Implement programs that are not cost effective
- Screening out programs that may be cost effective

What are some evaluation choices?

Randomized:

- **Randomized controlled trials**

Quasi-experimental:

- Regression Discontinuity
- **Propensity score matching**
- Other matching, including Variation in Adoption
- **Pre-post comparison, e.g. Baseline Method**

Many others….how can you tell which to use?

Outline

- What is a BB program?

- Why is rigorous evaluation important, and why is it hard?

- **Best evaluation practices**
  - Resources: SEE Action, UMP 2014, UMP forthcoming, LBNL
  - Randomized controlled trials (RCTs) are "gold standard"
  - Other, "quasi-experimental" methods when RCTs aren't feasible

- Recent research
  - EMV 2.0: High frequency data, new analytics
  - Compare RCTs to other methods in the real world

What resources are available for best practices in evaluation?

#1: SEE Action report, 2012

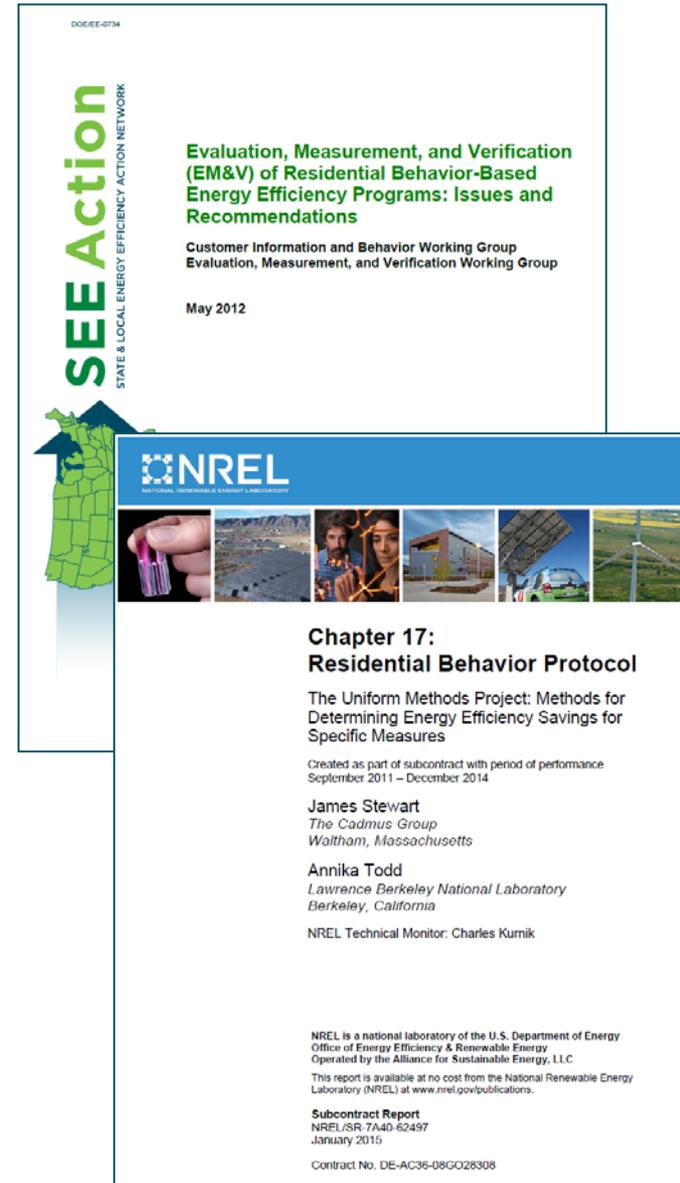"EM&V for Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations"

#2: UMP, 2014

"The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures. Chapter 17: Residential Behavior Protocol"

#3: New UMP coming soon!

#4: Technical Assistance from LBNL

→ *Many* states / regulators have adopted these protocols

DOE/EE-0734

SEE Action
STATE & LOCAL ENERGY EFFICIENCY ACTION NETWORK

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

Customer Information and Behavior Working Group
Evaluation, Measurement, and Verification Working Group

May 2012

NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Chapter 17:
Residential Behavior Protocol

The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures

Created as part of subcontract with period of performance September 2011 – December 2014

James Stewart
The Cadmus Group
Waltham, Massachusetts

Annika Todd
Lawrence Berkeley National Laboratory
Berkeley, California

NREL Technical Monitor: Charles Kurnik

NREL is a national laboratory of the U.S. Department of Energy Office of Energy Efficiency & Renewable Energy Operated by the Alliance for Sustainable Energy, LLC.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Subcontract Report
NREL/SR-7A40-62497
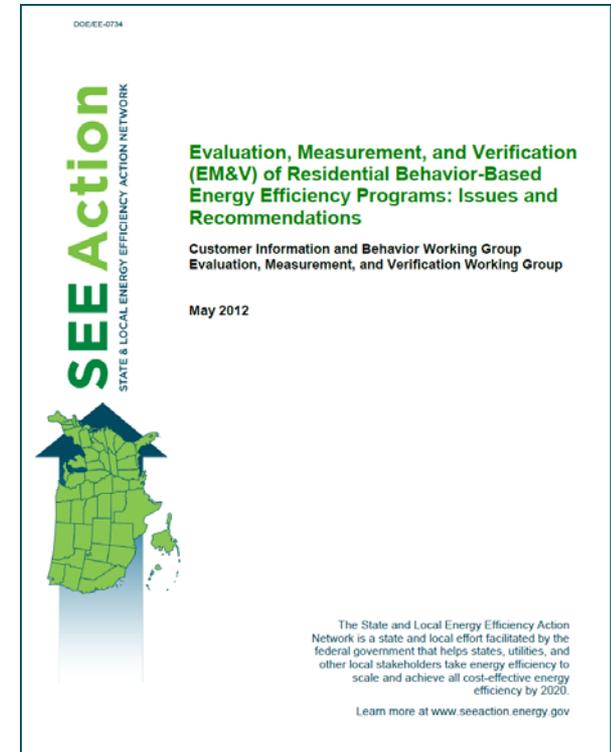January 2015

Contract No. DE-AC36-08GO28308

What resources are available for best practices in evaluation?

#1: SEE Action report, **2012**

"EM&V for Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations"

- Goal: a high degree of confidence that program impact estimates are valid

→Key recommendation: "Gold Standard" of evaluation is….

DOE/EE-0734

SEE Action
STATE & LOCAL ENERGY EFFICIENCY ACTION NETWORK

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

Customer Information and Behavior Working Group
Evaluation, Measurement, and Verification Working Group

May 2012

The State and Local Energy Efficiency Action Network is a state and local effort facilitated by the federal government that helps states, utilities, and other local stakeholders take energy efficiency to scale and achieve all cost-effective energy efficiency by 2020.

Learn more at www.seeaction.energy.gov

Use a Randomized Controlled Trial (when possible)!

Key recommendation 1: use a randomized controlled trial (RCT)

★★★★★ Randomized controlled trial (RCT)

★★★★☆ Regression discontinuity

★★★★☆ Variation in adoption

★★★☆☆ Propensity score matching

★ Not Advisable Non-propensity score matching

★ Not Advisable Pre-post comparison

★★★★★ Randomized controlled trial (RCT)

★★★★☆ Regression discontinuity

- **Primary recommendation** – a program that is designed as a RCT results in:
  - Transparent, straightforward analysis
  - Robust, accurate, valid program impact estimates
  - **High degree of confidence in program evaluation**
  - RCTs are the gold standard

Key recommendation 1: use a randomized controlled trial (RCT)

★★★★★ Randomized controlled trial (RCT)

★★★★☆ Regression discontinuity

★ Variation in adoption

- Why is designing a program as a (RCT) so important?
    - RCT means that households are assigned to the program randomly (as opposed to household choice or screening criteria)
    - Solves selection bias

Key recommendation 1: use a randomized controlled trial (RCT)

★★★★★ Randomized controlled trial (RCT)

★★★★☆ Regression discontinuity

Variation in adoption

- RCTs have many different forms

- Can be used for Opt-in, Opt-out programs

Key recommendation 1: use a randomized controlled trial (RCT)

★★★★★ Randomized controlled trial

★★★★☆ Regression discontinuity

★★★★☆ Variation in adoption

★★★☆☆ Propensity score matching

Not Advisable Non-propensity score matching

Not Advisable Pre-post comparison

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

What resources are available for best practices in evaluation?

#2: UMP, 2014

"The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures. Chapter 17: Residential Behavior Protocol"

- Key recommendation, **same as SEE Action 2012: RCTs**

Chapter 17: Residential Behavior Protocol

The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures

Created as part of subcontract with period of performance September 2011 – December 2014

James Stewart
The Cadmus Group
Waltham, Massachusetts

Annika Todd
Lawrence Berkeley National Laboratory
Berkeley, California

NREL Technical Monitor: Charles Kurnik

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Subcontract Report
NREL/SR-7A40-62497
January 2015

Contract No. DE-AC36-08GO28308

What resources are available for best practices in evaluation?

#3: New UMP coming soon!

- Key recommendations **still the same….RCTs**

- We are updating with latest research:
  - LBNL research comparing RCTs with other methods, based on real world evidence
  - More detailed guidance about impacts of BB programs on EE programs
  - New methods using smart meter data.

What resources are available for best practices in evaluation?

#4: Technical Assistance from LBNL

- We love program design and evaluation – EE, BB, and DR

    – Tricky implementation? We are creative!

    – Want options? We know RCTs are not always feasible!

    – Just want to brainstorm? Nothing we'd rather do

Outline

- What is a BB program?

- Why is rigorous evaluation important, and why is it hard?

- Best evaluation practices
  - Resources: SEE Action, UMP 2014, UMP forthcoming, LBNL
  - Randomized controlled trials (RCTs) are "gold standard"

- **Switch gears: Recent** advances in evaluation
  - **EMV 2.0: High frequency data, new analytics**
  - Compare RCTs to other methods in the real world (Anna)

Lawrence Berkeley National Laboratory
Environmental Energy Technologies Division Behavior Analytics
Providing insights that enable evidence-based, data-driven decisions

EM&V 2.0?

Behavior Analytics: combining behavioral theories with cutting-edge data science analytics

Team: Sam Borgeson, Dan Fredman, Ling Jin, Sid Patel, Anna Spurlock, Annika Todd (ETA, LBNL), Alex Sim, John Wu (CRD, LBNL), Taehoon Kim, Dongeun Lee, Jaesik Choi (UNIST, South Korea)

Data "explosion" in energy

- Smart meters, thermostats, appliances, cars

- Linked to other time and location-specific information (temperature, census, satellite)

- Provide vast, constantly growing streams of rich data

Smart meter data enables many possibilities for new types of analysis

- What can we do with this data?
- Many possibilities!

- These data have the potential to provide tremendous value to a wide range of energy policies
- One example: use it to examine issues in Behavior-Based (BB) programs
  - Example: Home Energy Report – compare usage to neighbors

What is a HER program?
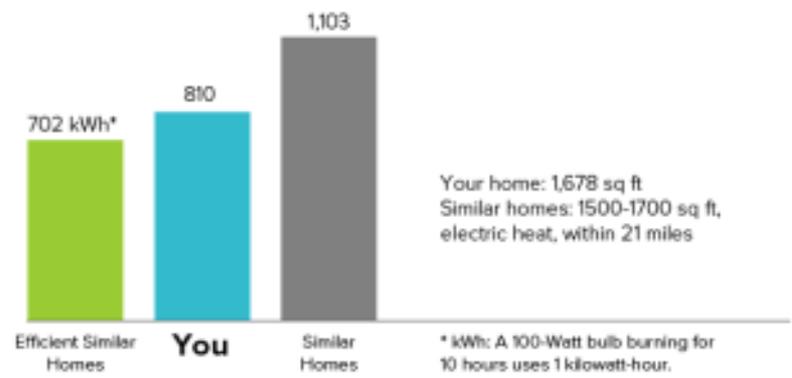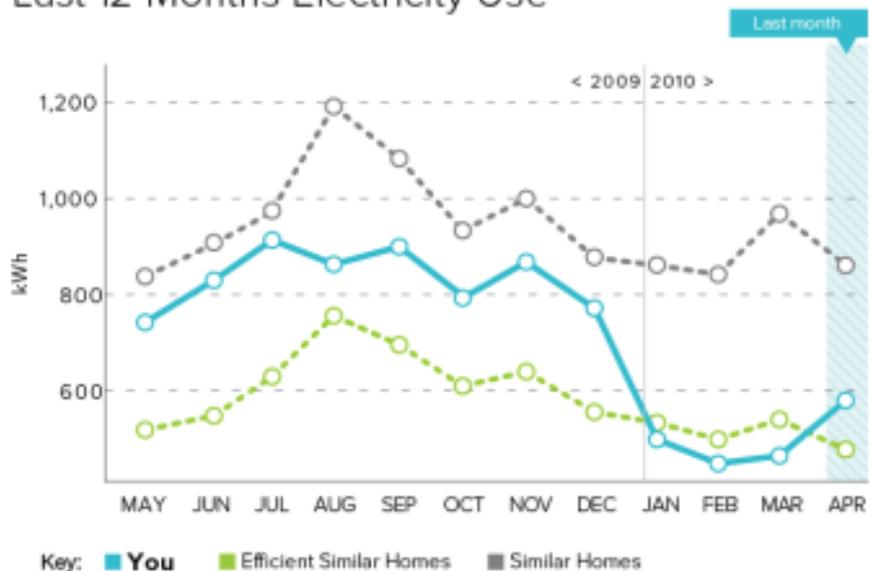
Analysis using smart meter data can help answer key policy questions

Key policy questions for HER (and other BB) programs:

1. **What is the short-term persistence of savings?** *(Results from our study: savings within one-two weeks after first report mailed, stabilize after second report)*

2. What is the long-term persistence of savings? *(Results: savings persist while mailings continue; savings decay after reports are discontinued)*

3. What actions and characteristics are related to savings? *(Results: suggestive of AC – best guess: changing thermostat set point)*
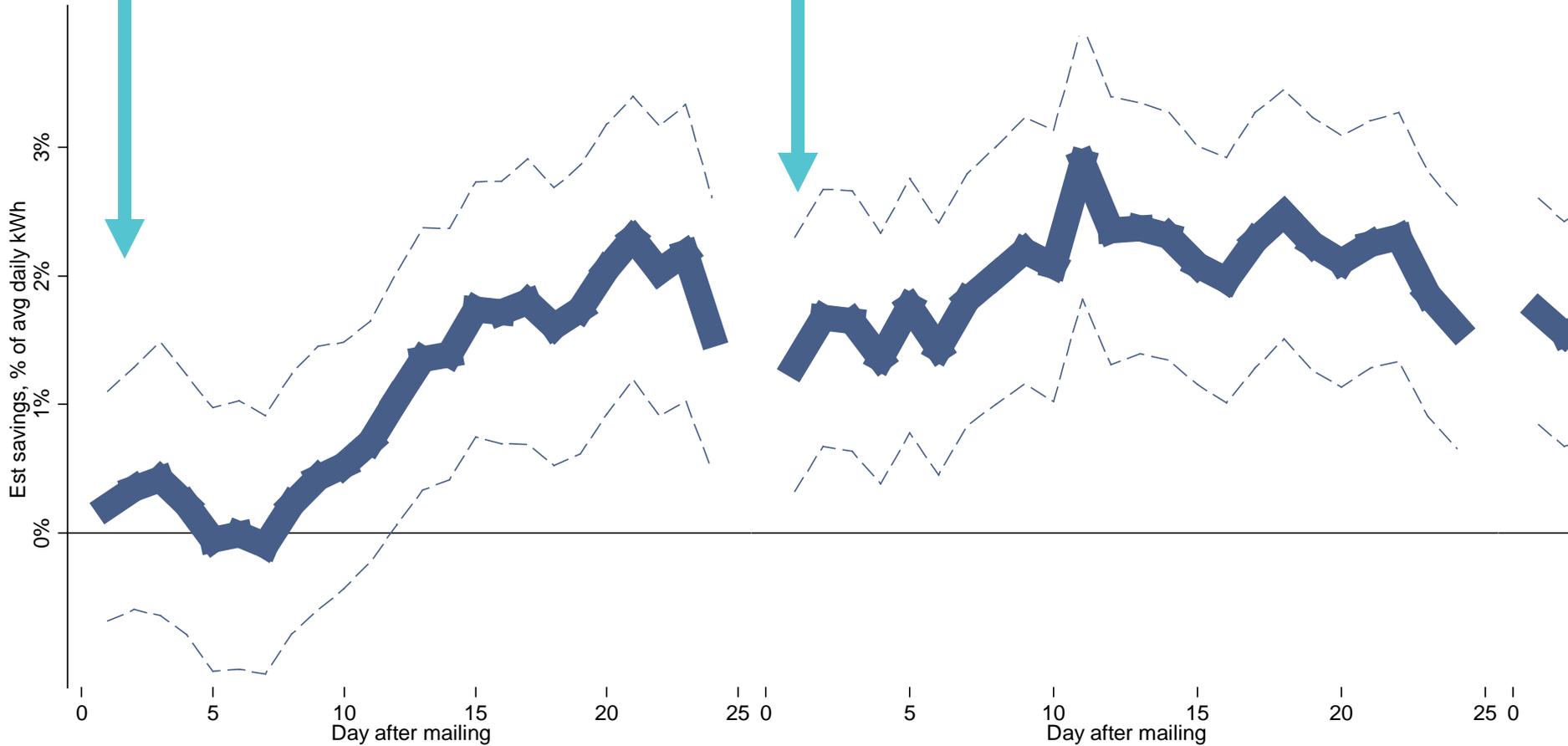
Mailing 1

Quick ramp up
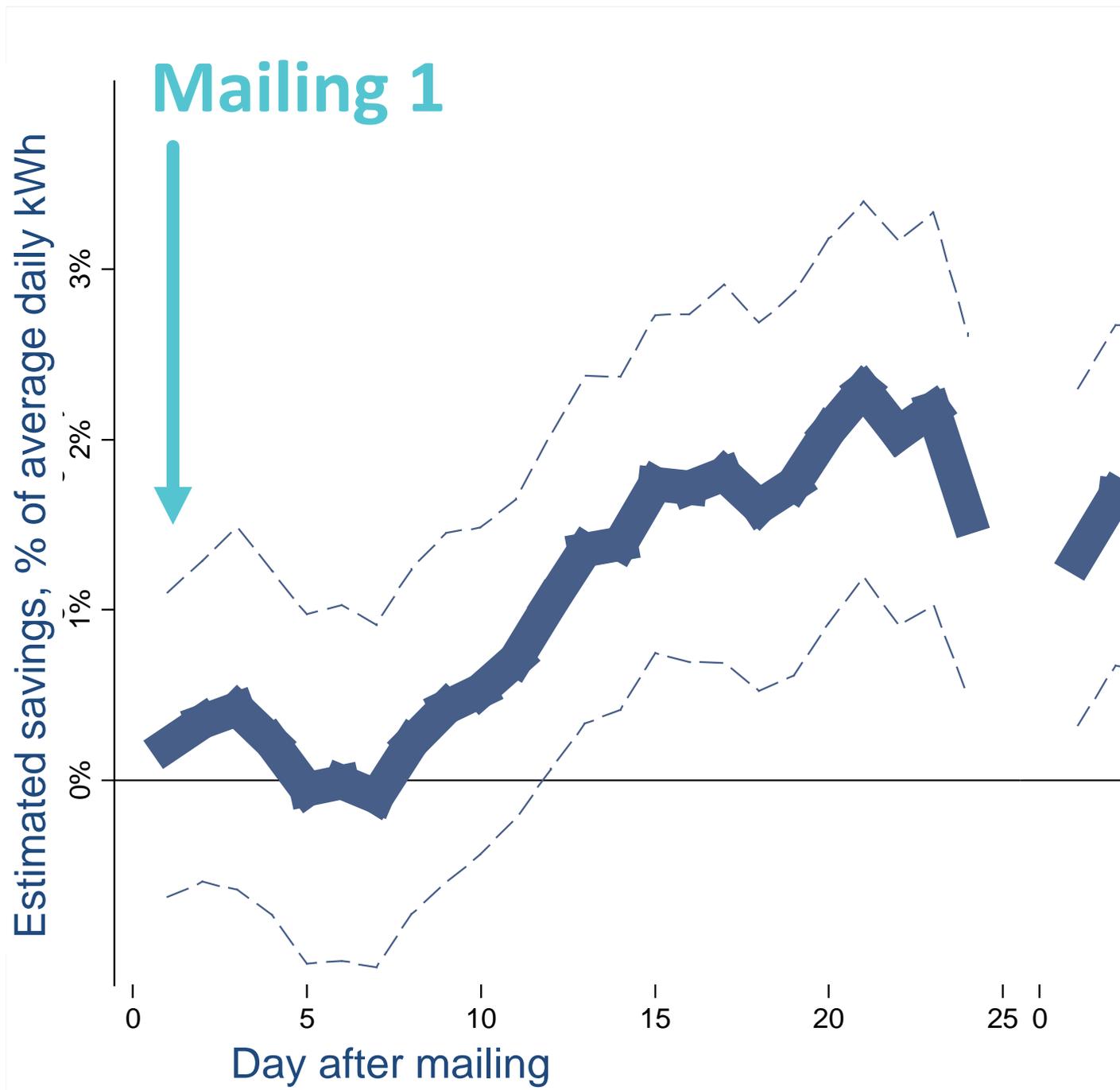
Suggests habitual or one-time savings actions rather than changes in installed equipment

Est savings, % of avg daily kWh

Day after mailing
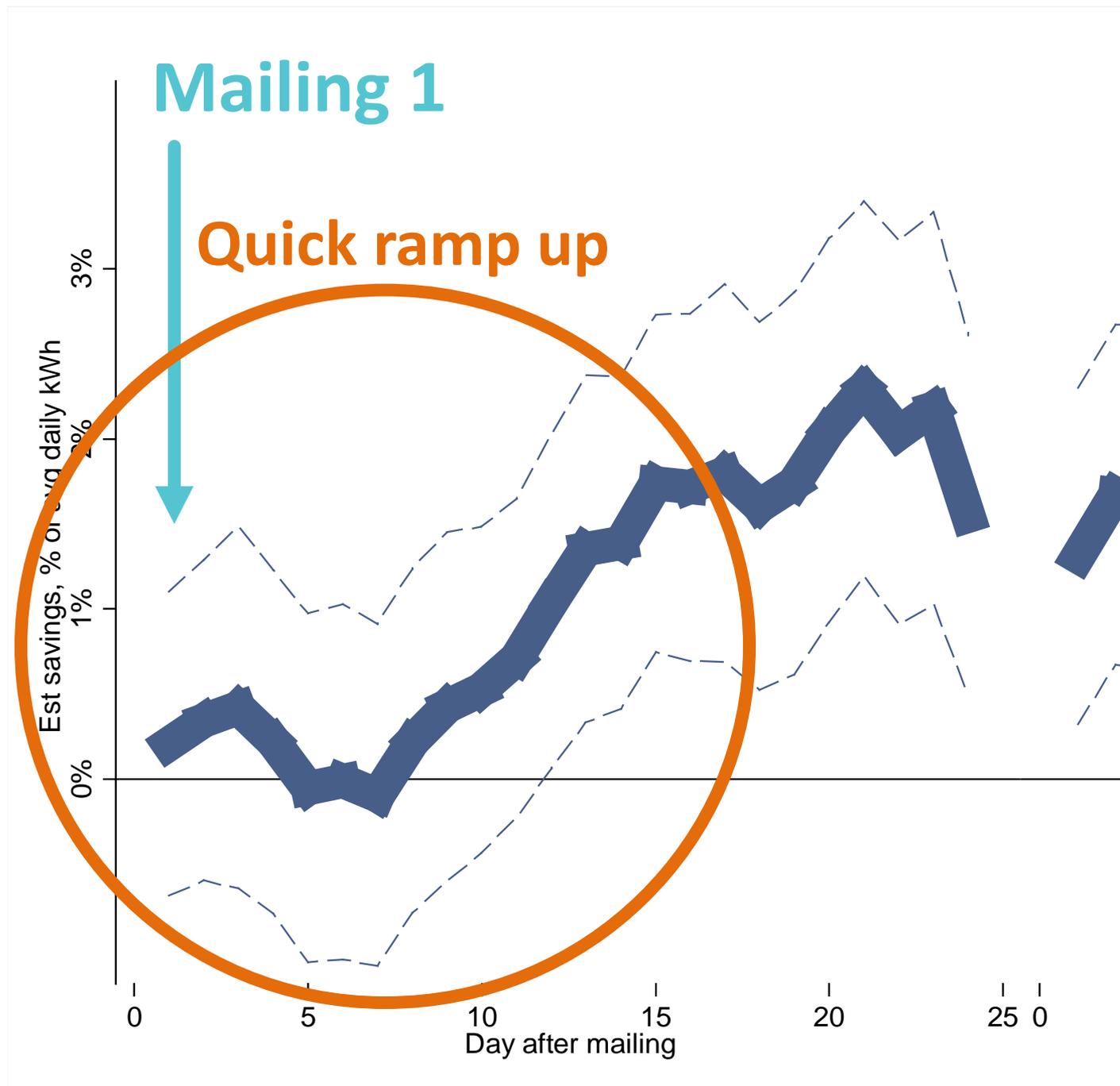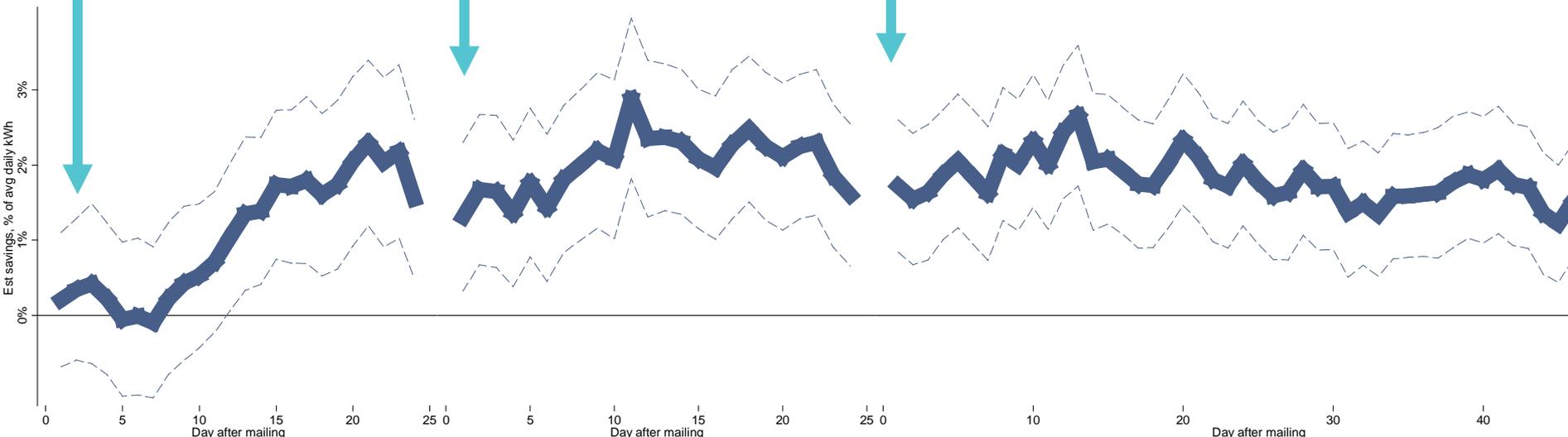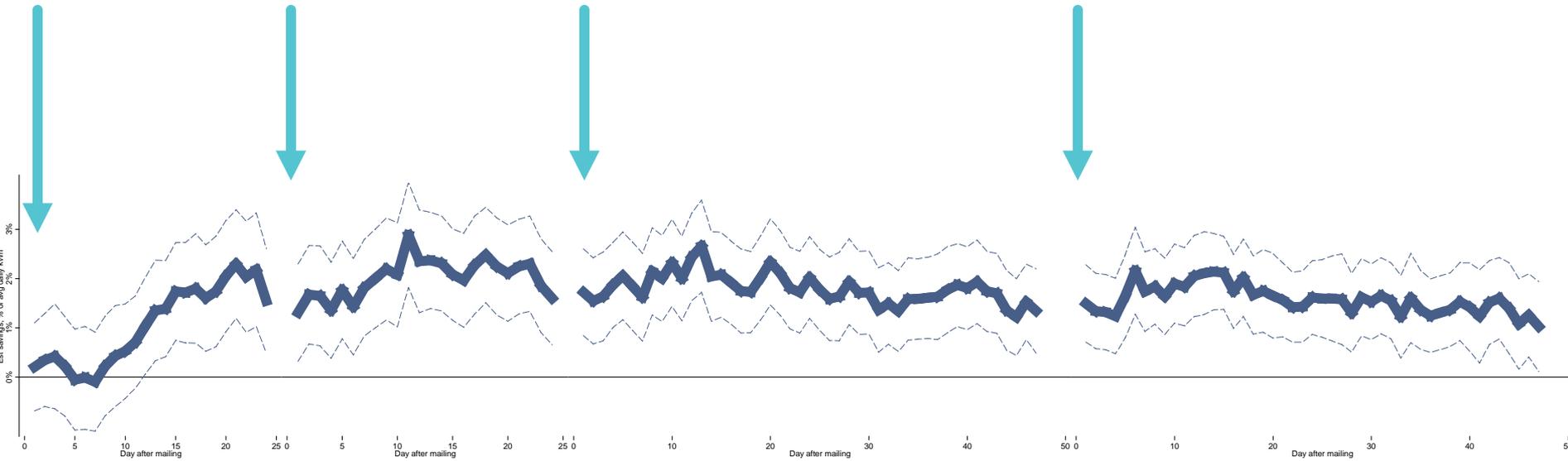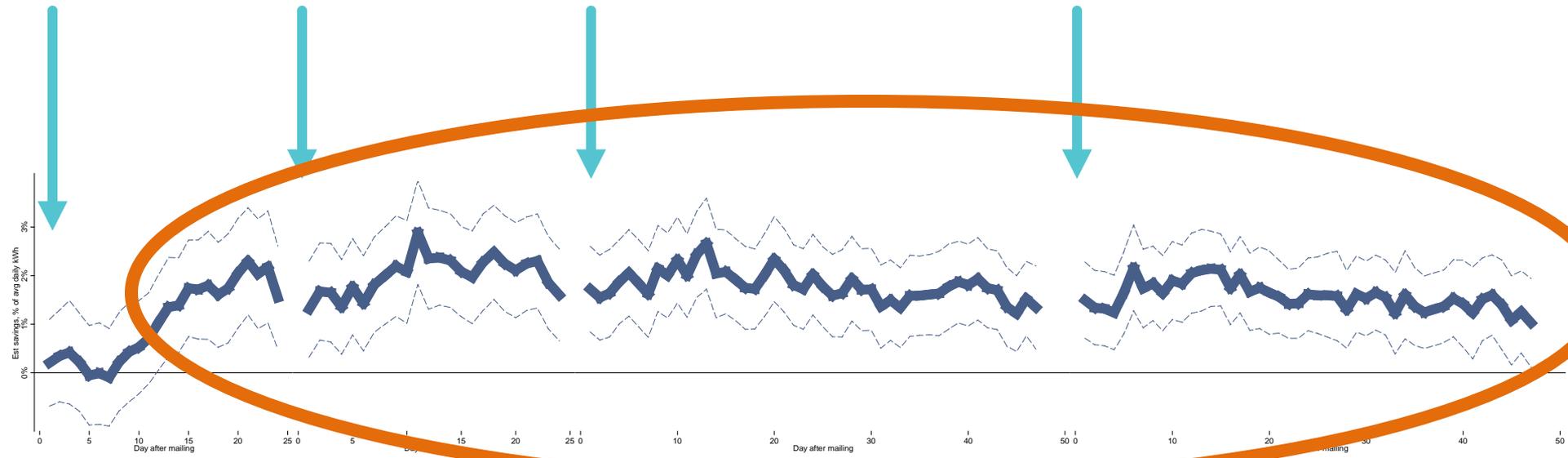
Mailing 1 Mailing 2 Mailing 3 Mailing 4

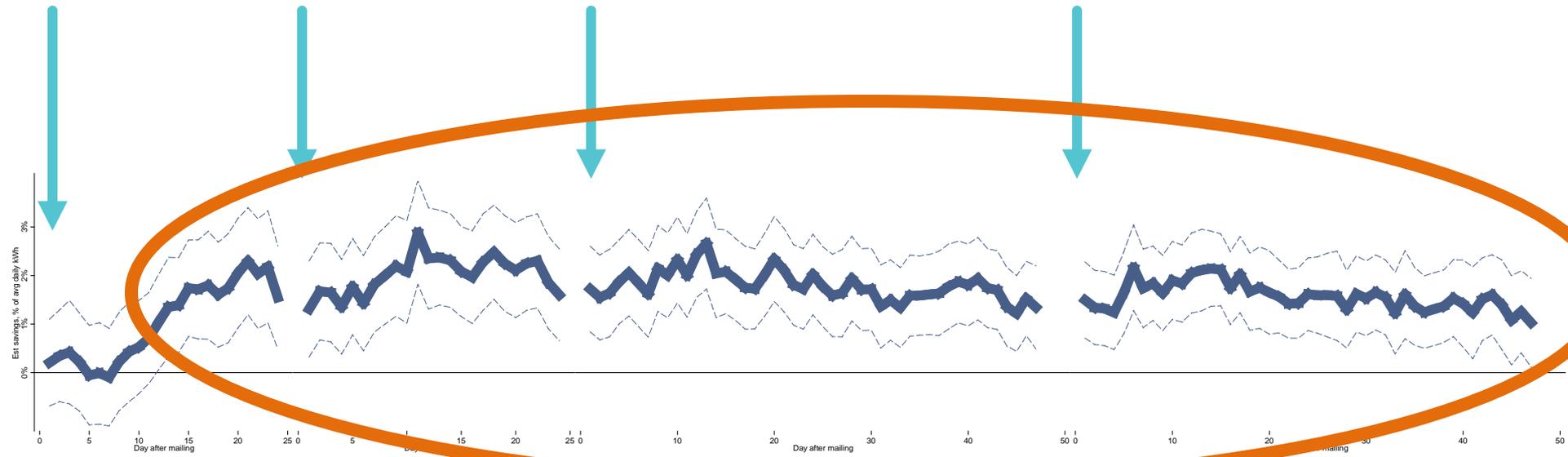Mailing 1   Mailing 2   Mailing 3   Mailing 4

Savings continue between mailings (there are statistically significant savings every day)

Mailing 1 **Mailing 2** **Mailing 3** **Mailing 4**

However the level of savings appears to vary somewhat

Analysis using smart meter data can help answer key policy questions

Key policy questions for HER (and other BB) programs:

1. **What is the short-term persistence of savings?** *(Results: savings within one-two weeks after first report mailed, stabilize after second report)*

2. **What is the long-term persistence of savings?** *(Results from recent example: savings persist while mailings continue; savings decay after reports are discontinued)*

3. **What actions and characteristics are related to savings?** *(Results: suggestive of AC – best guess: changing thermostat set point)*

Analysis using smart meter data can help answer key policy questions

Key policy questions for HER (and other BB) programs:

1. **What is the short-term persistence of savings?** *(Results: savings within one-two weeks after first report mailed, stabilize after second report)*

2. **What is the long-term persistence of savings?** *(Results: savings persist while mailings continue; savings decay after reports are discontinued)*

3. **What actions and characteristics are related to savings?** *(Results: suggestive of AC – best guess: changing thermostat set point)*

Outline

- What is a BB program?

- Why is rigorous evaluation important, and why is it hard?

- Best evaluation practices
  - Resources: SEE Action, UMP 2014, UMP forthcoming, LBNL
  - Randomized controlled trials (RCTs) are "gold standard"

- **Recent research**
  - EMV 2.0: High frequency data, new analytics
  - **Compare "Gold Standard" RCTs to other methods in the real world (Anna)**

Lawrence Berkeley National Laboratory
Environmental Energy Technologies
Division **Behavior Analytics**
Providing insights that enable evidence-based, data-driven decisions

Contact us:

Annika Todd atodd@lbl.gov
Anna Spurlock caspurlock@lbl.gov

SMART GRID INVESTMENT GRANT
CONSUMER BEHAVIOR STUDY ANALYSIS

Go for the Silver? Evidence from field studies quantifying the difference in evaluation results between "gold standard" randomized controlled trial methods versus quasi-experimental methods

Patrick Baylis[†], Peter Cappers[*], Ling Jin[*], C. Anna Spurlock[*], Annika Todd[*]

* Lawrence Berkeley National Lab
† Stanford Center for Food Security and the Environment

Motivation

- **Randomized controlled trials (RCTs) are widely viewed as the "gold standard" of evaluation.**
  - Widely used in applied research fields such as health and public policy
  - Requires forethought and planning in program implementation, and may not always be possible

- **Analysis of the effect of DR and energy pricing programs have largely been conducted through non-RCT ("quasi-experimental") methods.**
  - Can be applied after a program has taken effect

Is Pricing "Behavior-Based"?

- **Pricing, like other programs that have more traditionally been categorized as "behavior based," are similar in that:**
  - Success from the program depends on people changing their behavior
    - In a pricing program the change is induced by a monetary incentive
    - In a behavior-based program the change is induced by a message, feedback, nudge or other similar intervention
  - Success of the program does not rely on the performance of a particular piece of equipment or improved insulation, etc.
  - Principals relevant for evaluation of both pricing and behavior-based programs are the same.

Evaluating a behavior-based or pricing program – Need a Comparison Group!

- **In order to evaluate a pricing or behavior-based program, the key piece of information that is needed is a "counterfactual."**
  - A counterfactual (comparison group, baseline, etc.) is a way of defining what would have happened had a particular program not been put in place.
  - In a Randomized Controlled Trial (RCT), the validity of the counterfactual comes from the randomization.
    - Two identical groups of people are compared to each other, one of which faced a new type of pricing (or message, nudge, etc.) and one of which did not.
  - But what if you don't, or can't, have an RCT?

How evaluation is done without an RCT

If an RCT is not done, and so a randomized control group is not available to use as a counterfactual, there are two common methods used:

1. **Propensity Score Matching (P-Score):**
   - Compare the self-selected treated group to people in an untreated group based on observable characteristics that make them look the most similar.

2. **Difference-in-Differences (DID):**
   - Compare the self-selected treated group to a randomly selected control group that didn't experience the program.

3. **In the case of an event-based program: create a baseline**
   - Compare a customer's use in the non-event days to their use during events. In this way customers are used as their own counterfactual, or baseline.

Concerns with non-randomized methods

- **Selection bias:**
  - Generally people will opt in (or self-select) into a program if they anticipate the program will be better for them than if they did not join.
  - This may mean that those people are systematically different, in ways that would affect their performance on the program, as compared to those who chose not to join.
  - Comparing a self-selected treatment group to an invalid counterfactual group would confound the difference in their behavior due to the program with differences in their behavior simply due to their underlying differences.

Concerns with non-randomized methods

- **Spillover**
  - In the case of an event-based program, when a customer's non-event usage is being used as a baseline for their event usage, the effect of the program is combined with any potentially changes to that customer's behavior that might spill over onto non-event days:
    - Habit formation, changing thermostat settings, new equipment purchased in response to the program, etc.

Data Used

- **DOE Smart Grid Investment Grant (SGIG) Consumer Behavior Studies (CBS)**

- **Sacramento Municipal Utilities District (SMUD)**
  - Tested different time-based pricing types, recruitment methods (opt-in vs. default) and an in-home energy information display.
  - Treatment in effect during the summers of 2012 and 2013.

Compared RCT to non-RCT evaluation methods for:

1. **Overall program impact evaluation**

    – Estimating peak period energy savings on average across all critical peak event days.

2. **Baseline methods used to generate household-specific savings on event days individually**
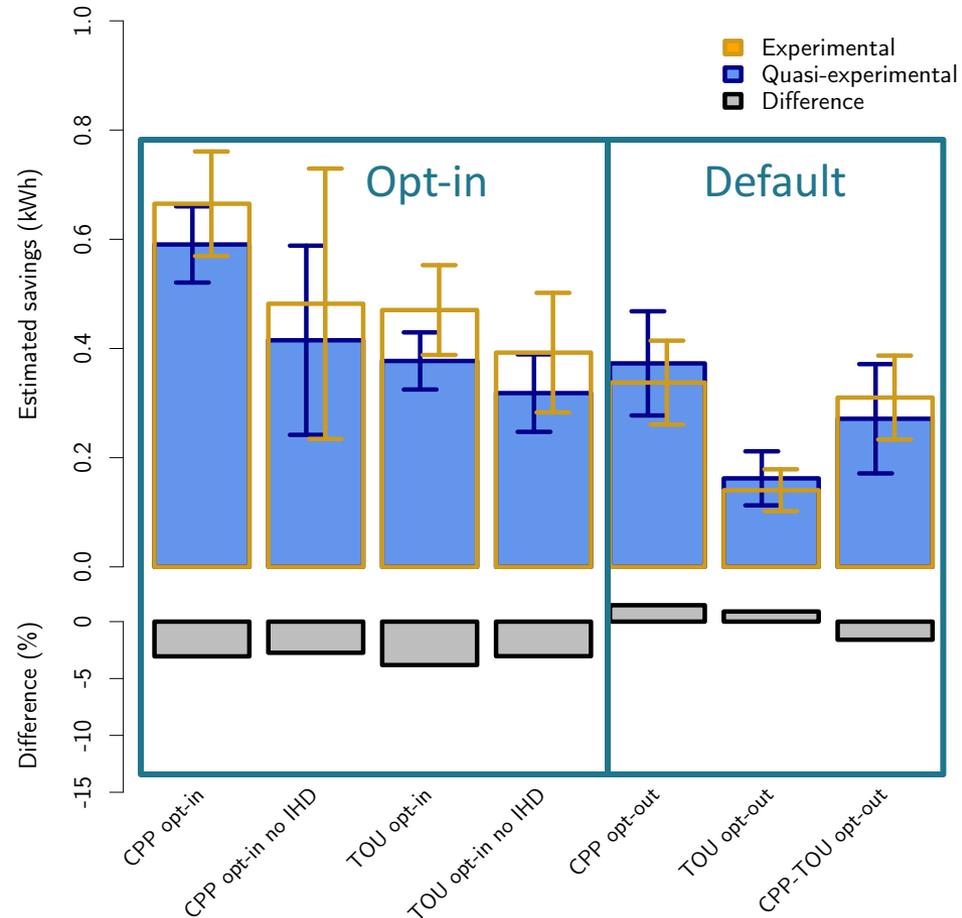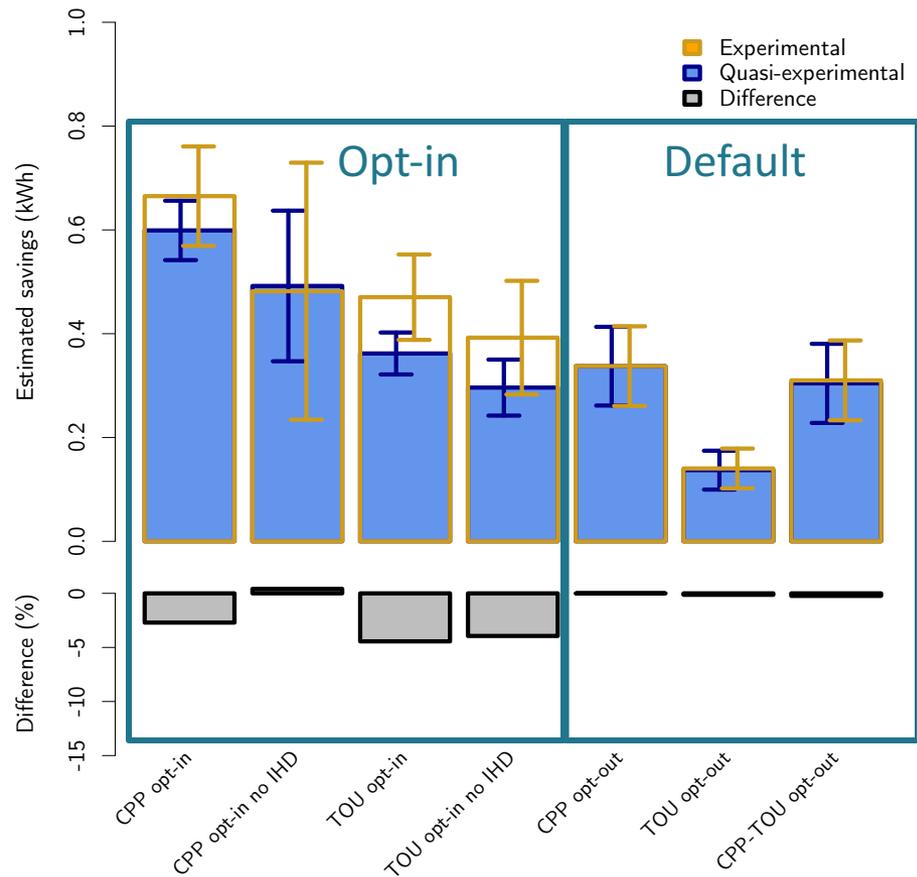
    – Four-in-five day baseline with a weather adjustment

      • Average consumption from four highest consumption days in the the last 5 non-event business days, with an adjustment based on off-peak use in baseline and event days

1. Evaluation Results: DID & P-Score

DID

Propensity Score

1. Evaluation Results

- **Non-RCT methods tend to underestimate treatment effect, particularly for opt-in treatments**
  - Bias as much as 5 percentage points (i.e., estimated effect would have been 15% of average peak period consumption when effect was actually 20%)
 - Average Absolute Difference: 1.7%-2.4%

1. Evaluation Results

- **Biases are more pronounced in opt-in vs opt-out designs**
  - Highly suggestive of the role of selection bias
  - Opt-in achieved at most 20% enrollment, opt-out was 90%
  - This means default treatment group resembles control group to a much greater extent than the opt-in treatment group.

2. Baseline Results: Estimate Comparison
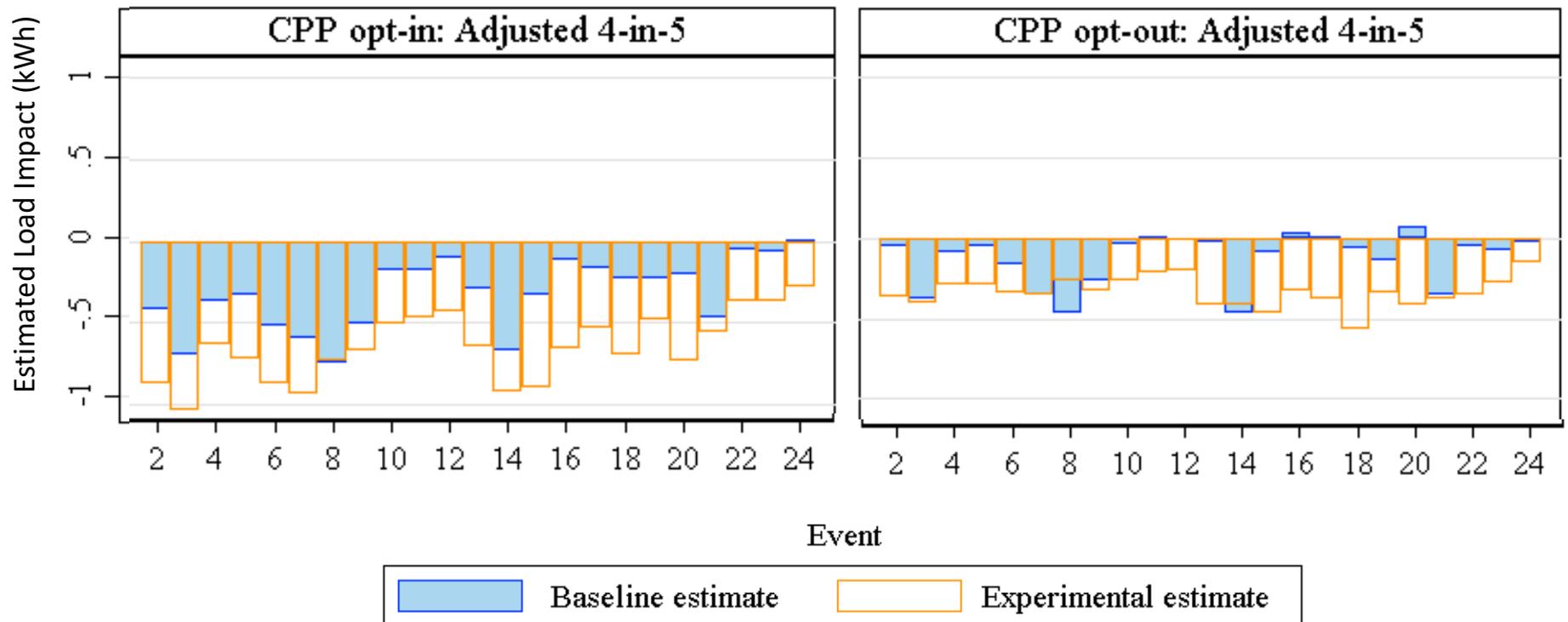
2. Baseline Results

- **Savings based on non-RCT baseline method underestimated treatment effect**

  – With differences as much as 20 percentage points in some cases (meaning a true effect of 25% of peak consumption would have been estimated at only 5%).

  – Another way to look at it is that the 4-in-5 adjusted baseline generated estimates that were 39-46% of the RCT estimates.

2. Baseline Results

- **Results strongly suggest the role of spillover effects in causing bias**
  - Customers change behavior on non-event hours and non-event days as a result of the rate because of habitual behavior change, reprograming equipment (e.g., thermostats), or new equipment investments.
  - These changes are caused by the rate and cause consumption in the "baseline" hours to be lower.
  - Therefore a comparison of event consumption to baseline consumption underestimates the true effect of the rate.

Lawrence Berkeley National Laboratory
Environmental Energy Technologies
Division Behavior Analytics
Providing insights that enable evidence-based, data-driven decisions

Contact us:

Annika Todd atodd@lbl.gov
Anna Spurlock caspurlock@lbl.gov

Evaluating Residential Behavior-Based Programs: Lessons Learned

The New York Experience

Bill Saxonis- New York State Department of Public Service-
9/21/2016

Evaluating Residential Behavior-Based Programs
A Vision in Transition

❑ EEPS-Energy Efficiency Portfolio Standard (2008)

❑ REV (Reforming the Energy Vision)-the goal is to establish New York as a leader in the transition to a clean energy economy (ongoing effort began in 2014)

❑ REV proposes a comprehensive "regulatory overhaul" that will reinvent how we produce, deliver and consume energy

❑ Clean Energy Standard (2016) - 50 percent of the state's power will come from renewable sources by 2030

Evaluating Residential Behavior-Based Programs

"The farther backward you can look,
the farther forward you are likely to see."
Winston Churchill

Evaluating Residential Behavior-Based Programs Data Challenge

❑ Privacy concerns, cybersecurity

❑ Release of personally identifiable customer utility data to a third party is generally prohibited

❑ The Commission issued an Order allowing customer data to be used for behavioral programs because the programs represented "a rate-payer funded utility function" (Case 07-M-0548, et al., 12/03/2010)

❑ Data related issues remain controversial within and beyond the world of evaluation

Evaluating Residential Behavior-Based Programs Evaluation Policy

❑ New York Evaluation Guidelines (2008 with periodic updates)

❑ Recognition that behavior-based programs represent a new program design and evaluation protocols used for more traditional energy programs may not be optimal

❑ Adopted the SEE Action "Issues and Recommendations" report (May 2012) on behavior-based programs to provide guidance and encourage best practices

❑ New York's Evaluation Guidelines are currently undergoing a compressive review and update to better align with "REV"

SEE Action
STATE & LOCAL ENERGY EFFICIENCY ACTION NETWORK

Evaluating Residential Behavior-Based Programs Results From Four Utilities

National Grid

Electric Savings–2.43%

Gas Savings–0.82%

Central Hudson

Electric Savings–2.45%

Gas Savings-1.44%

Rochester Gas & Electric (RG&E) and New York State Electric and Gas (NYSEG)

Initial results show savings of less than 3% (report not finalized)

Evaluating Residential Behavior-Based Programs Interesting Findings and Recommendations

❑ Customers generally were satisfied with the programs

❑ Important to keep the customer interested (e.g., engagement with web site, targeted recommendations)

❑ Need to place more emphasis on "cross-selling" other utility energy programs

❑ Interest in better understanding certain customer categories (e.g., retiree households, customers who do not directly pay utility bills)

Evaluating Residential Behavior-Based Programs Evaluation and Programmatic challenges

❑ **Retail Competition Program (1999):**

Nearly 75% of the respondents reported that the bills savings were less than expected or they were unable to determine if there were any savings
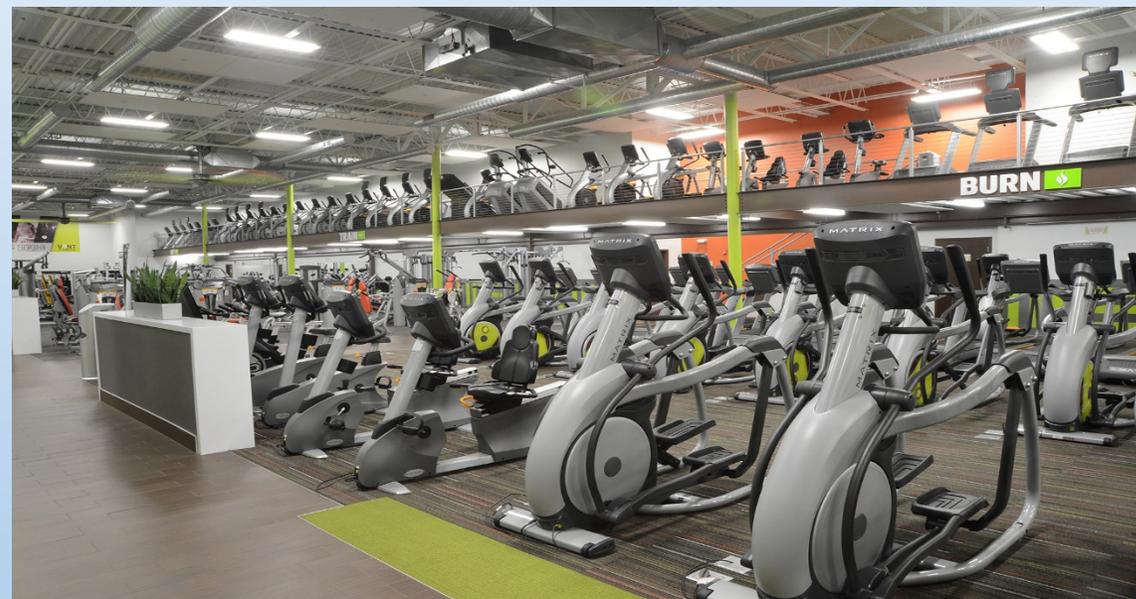
❑ **RG&E/NYSEG Behavioral Program (2015):**

Only about 30% perceived that their bills were lower

❑ Small evaluation problems can result in misleading results when the savings are small (less than 3%)

Evaluating Residential Behavior-Based Programs

❑ Possible problems - missing data, sample contamination

❑ Three keys to good evaluation-
plan, plan, plan

❑ We often wonder about
persistence…do the savings last?

Evaluating Residential Behavior- Based Programs

❑ Central Hudson and National Grid will continue to offer residential behavioral programs. RG&E and NYSEG will not

❑ Influenced by the evaluation results, both National Grid and Central Hudson are placing increased emphasis on customized energy tips, improved cross program promotion and enhanced web interaction. National Grid is adding a rewards program

Evaluating Residential Behavior-Based Programs

Evaluating Residential Behavior- Based Programs

*"A problem is a chance for
you to do your best."*

Duke Ellington (1899-1974)
American genius

Evaluating Residential Behavior-Based Programs

Thank you for your attention!

A few references:

NYS Department of Public Service:

http://www.dps.ny.gov/

Central Hudson Evaluations

http://www.savingscentral.com/programevaluations/

National Grid Evaluations:

https://www9.nationalgridus.com/EnergyEfficiencyReports.asp

Bill Saxonis

william.saxonis@dps.ny.gov

Discussion/Questions

For more EM&V information see:

- Webinars: https://emp.lbl.gov/emv-webinar-series

- For technical assistance to state regulatory commissions, state energy offices, tribes and regional entities, and other public entities see: https://emp.lbl.gov/projects/technical-assistance-states

- Energy efficiency publications and presentations – financing, performance contracting, documenting performance, etc. see: https://emp.lbl.gov/research-areas/energy-efficiency

- New Technical Brief - *Coordinating Demand-Side Efficiency Evaluation, Measurement and Verification Among Western States: Options for Documenting Energy and Non- Energy Impacts for the Power Sector* https://emp.lbl.gov/publications/coordinating-demand-side-efficiency

From Albert Einstein:

"Everything should be as simple as it is, but not simpler"

"Everything that can be counted does not necessarily count; everything that counts cannot necessarily be counted"