SEE Action

STATE & LOCAL ENERGY EFFICIENCY ACTION NETWORK

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

## Issues and Recommendations

April 3, 2014


Michael Li *U.S. Department of Energy*

Annika Todd *Lawrence Berkeley National Lab*

# Outline: EM&V of Behavior-Based EE Programs

- What is a behavior-based EE program?
- Why is evaluation of these programs hard?
- **How can we be confident that the energy savings are valid?**
- **What are key guidelines on best practice methods (and why are RCTs the gold standard)?**

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


Randomized controlled trial (RCT)

Regression discontinuity

Variation in adoption

Propensity score matching

Non-propensity score matching

Pre-post comparison

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

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

- 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

- 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 tri

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

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

# **Insights from Smart Meters:** Focus on Persistence of Savings from Home Energy Reports

Annika Todd, Michael Li
April 2014

# 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 persistence issues in Behavior-Based (BB) programs
  - Today - we show example analyses from two particular program rollouts that help answer key persistence questions

# 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)*

# Data description

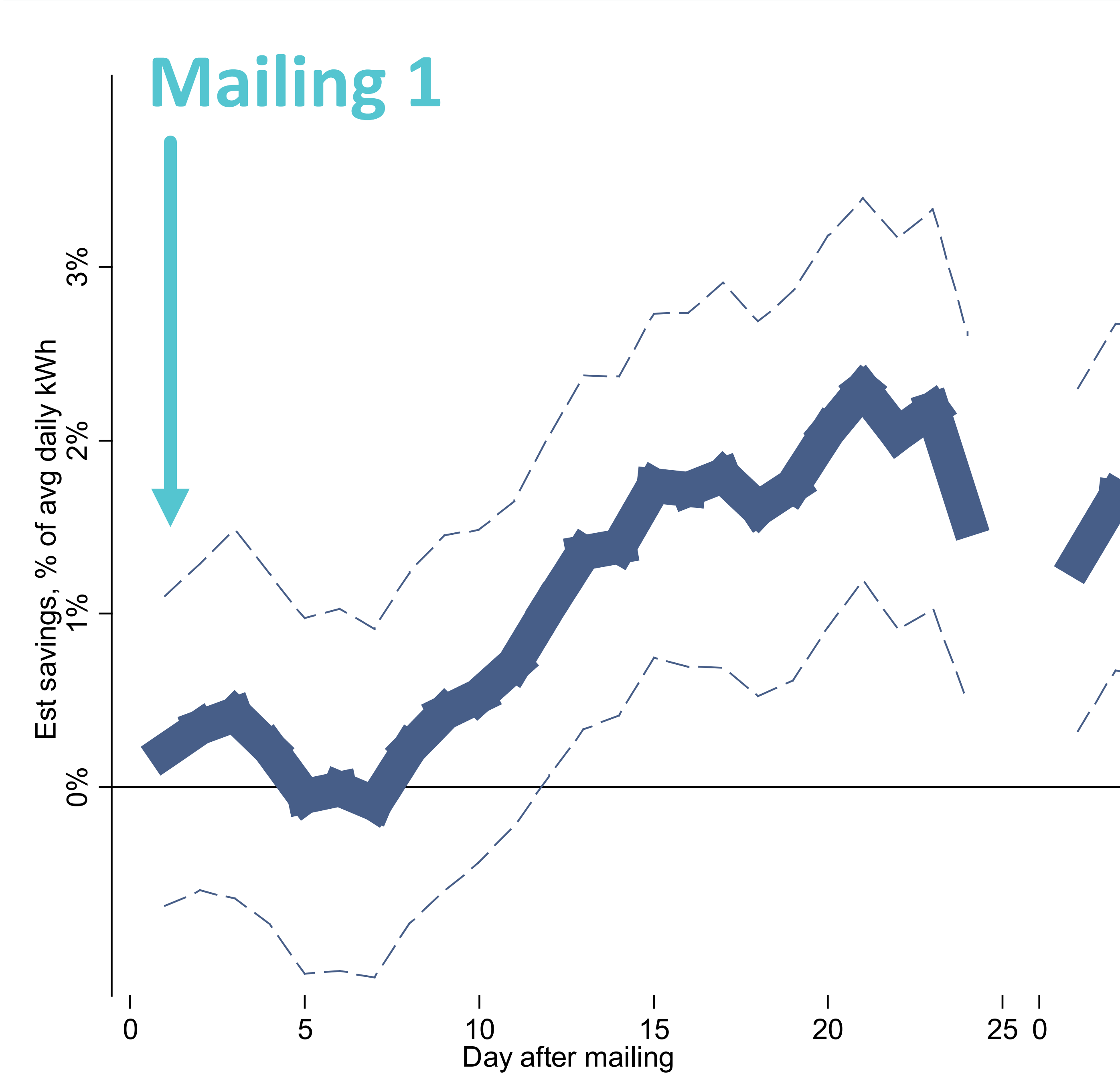
- HER program implemented as a "randomized controlled trial"
- Hourly electricity data from Pacific Gas & Electric's (PG&E) AMI system
- Two datasets from different rollouts ("waves")

| | # Treat | # Control | Launch Date | Hourly interval data available | PG&E baseline territory | Quartile of energy use |
|---|---|---|---|---|---|---|
| **Wave One** | **400,000** | **100,000** | Feb 2012 | Aug 1, 2012- Oct 31, 2012 | **P, Q, R, S, T, V, W, X, Y** | **Top 3 quartiles** |
| **Gamma Wave** | 72,300 | 72,300 | **Nov 2011** | **Nov 4, 2011- Aug 1, 2012** | **R**, S, **T**, **W**, X | All quartiles |

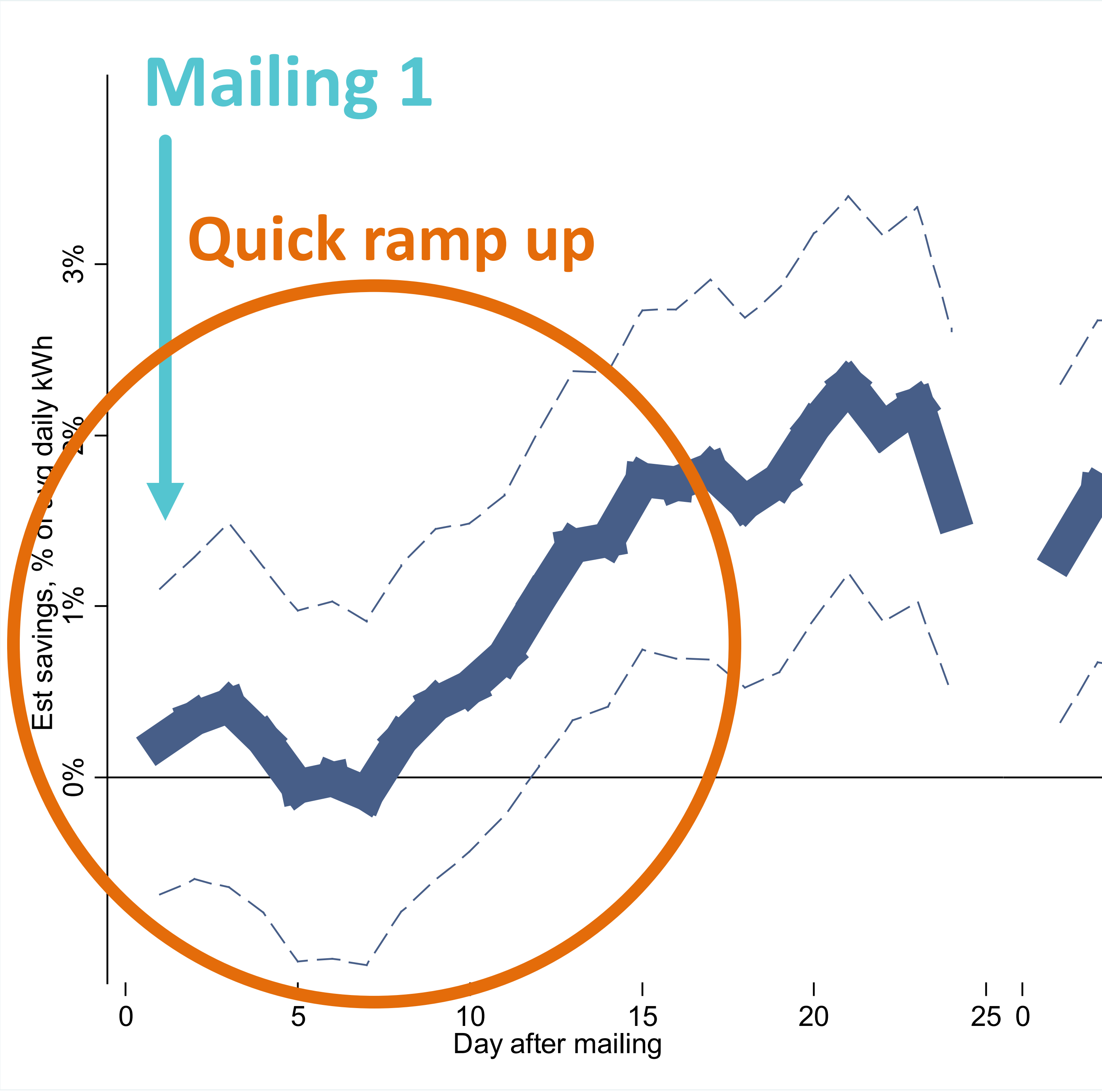
# Analysis using smart meter data can help answer key policy questions

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

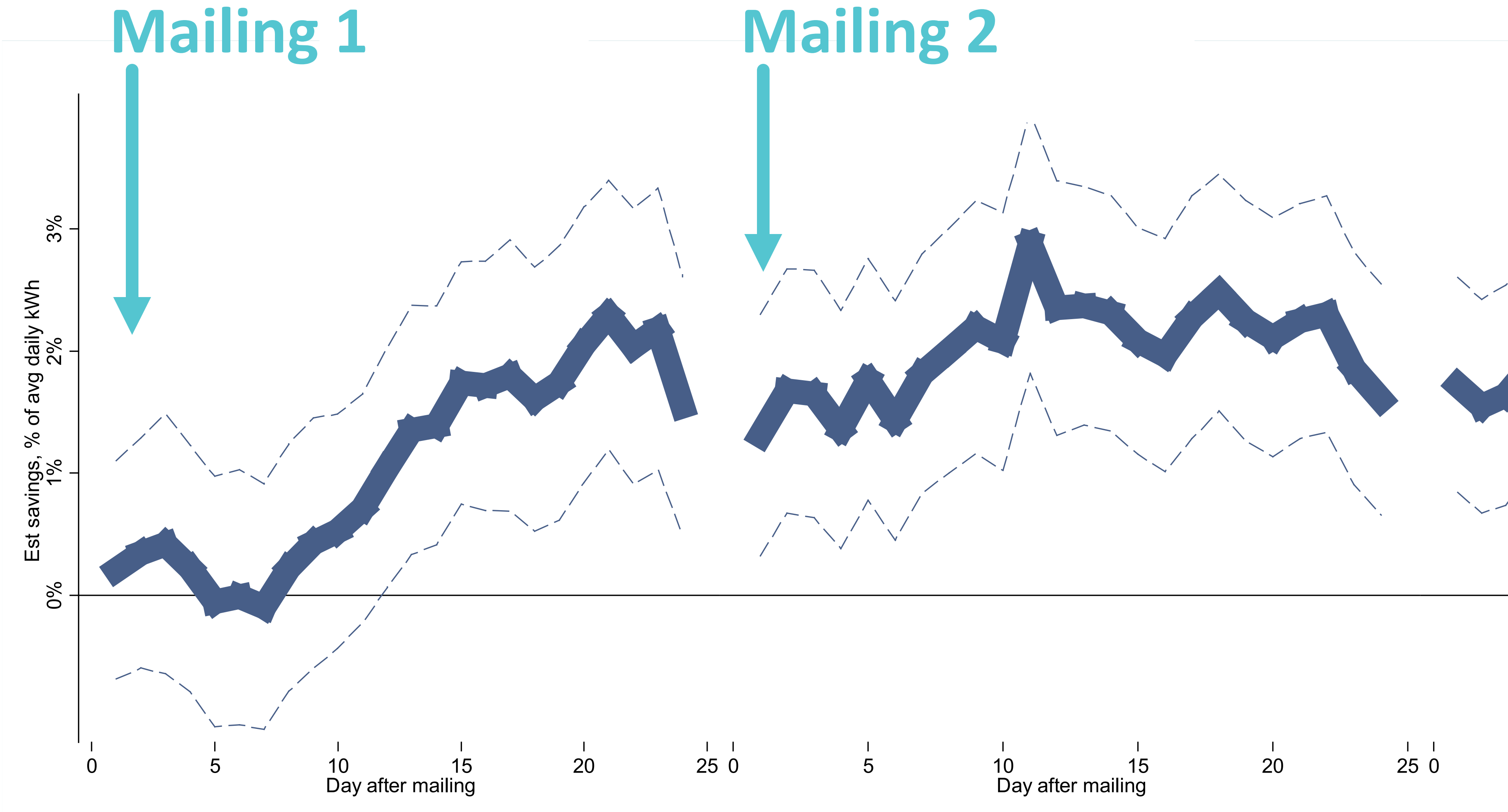
1. **What is the short-term persistence of savings?** ***(Results: savings within one-two weeks after first report mailed, stabilize after second report)***
2. **What is the long-term persistence of savings?** ***(Results: savings persist while mailings continue; savings decay after reports are discontinued)***
3. **What actions and characteristics are related to savings?** ***(Results: suggestive of AC – best guess: changing thermostat set point)***

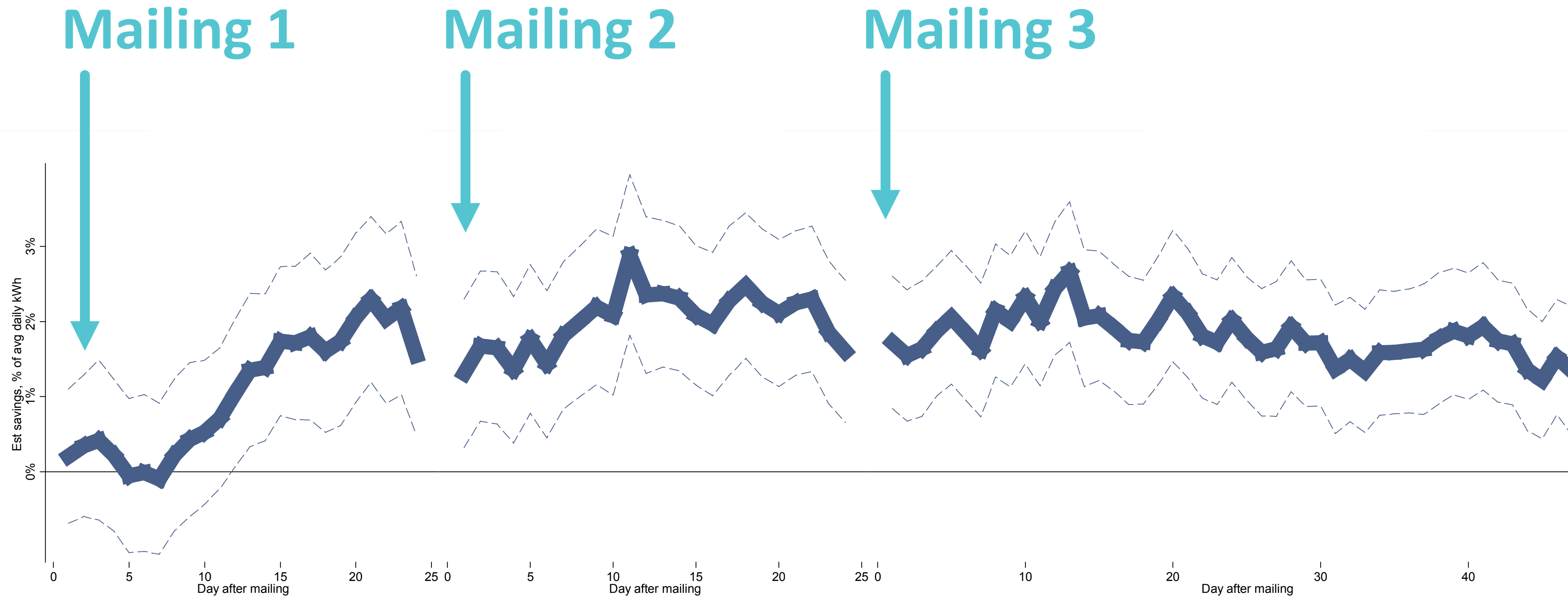


Mailing 1
Est savings, % of avg daily kWh
0%
1%
2%
3%
0
5
10
15
20
25
0
Day after mailing

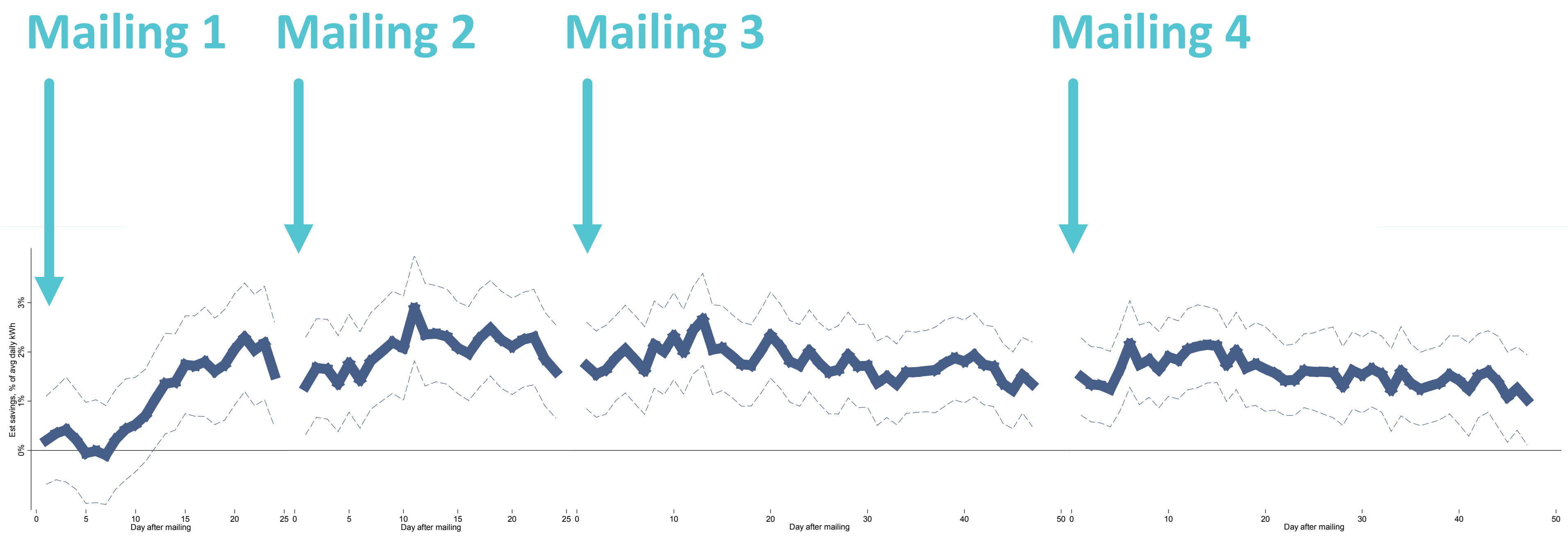


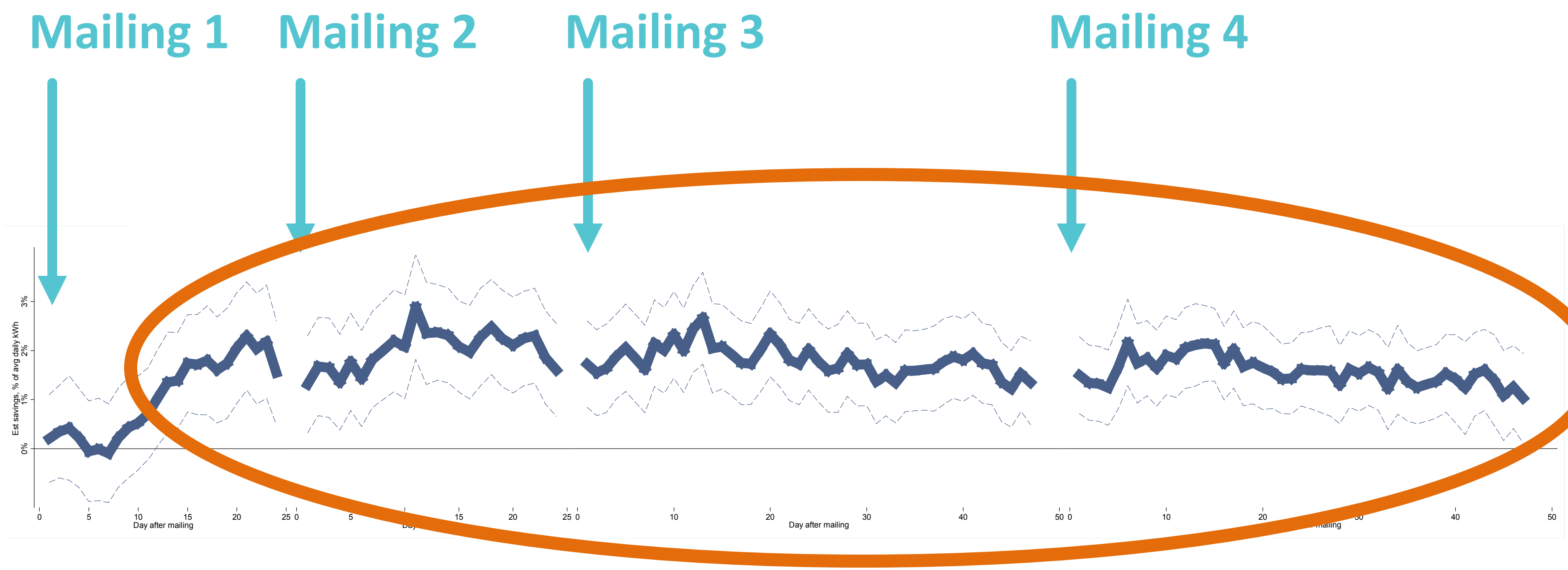
Mailing 1
Quick ramp up
Est savings, % of avg daily kWh
3%
2%
1%
0%
0
5
10
15
20
25
0
Day after mailing


Mailing 1
Mailing 2
Est savings, % of avg daily kWh
0%
1%
2%
3%
0
5
10
15
20
25
Day after mailing
0
5
10
15
20
25
Day after mailing
0

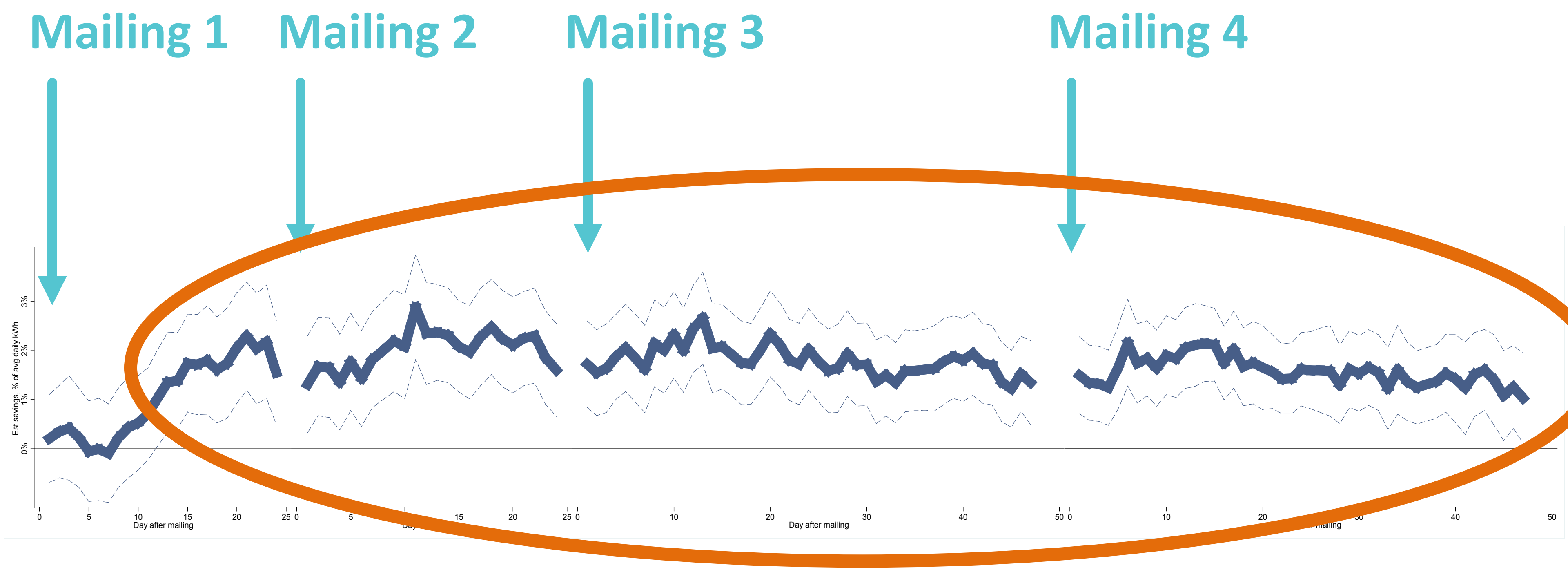


Mailing 1
Mailing 2
Mailing 3
Est savings, % of avg daily kWh
0%
1%
2%
3%
0
5
10
15
20
25
Day after mailing
0
5
10
15
20
25
Day after mailing
0
10
20
30
40
Day after mailing

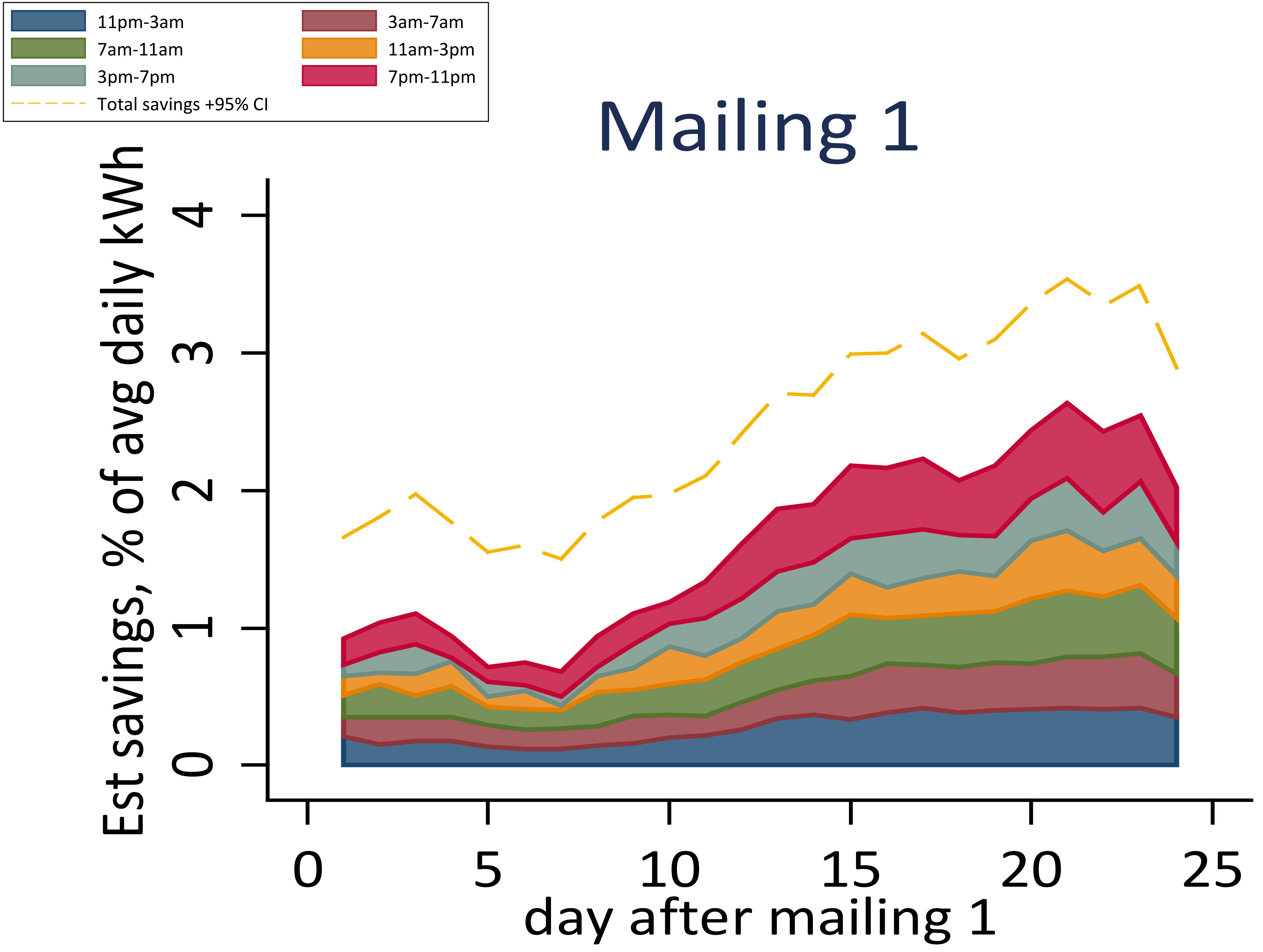

Mailing 1
Mailing 2
Mailing 3
Mailing 4
Est savings, % of avg daily kWh
0%
1%
2%
3%
Day after mailing
Day after mailing
Day after mailing
Day after mailing
0
5
10
15
20
25
0
5
10
15
20
25
0
10
20
30
40
50
0
10
20
30
40
50

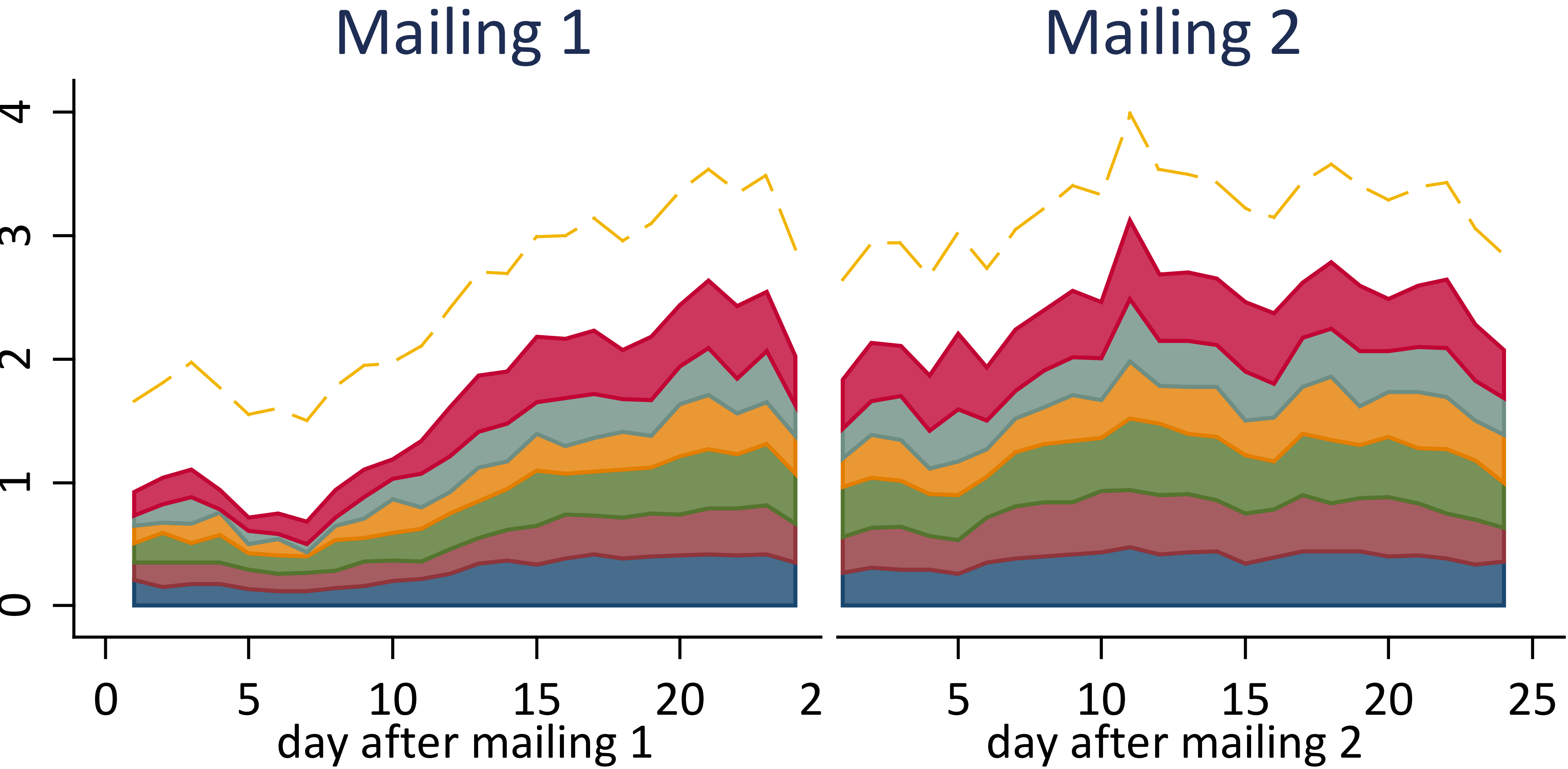


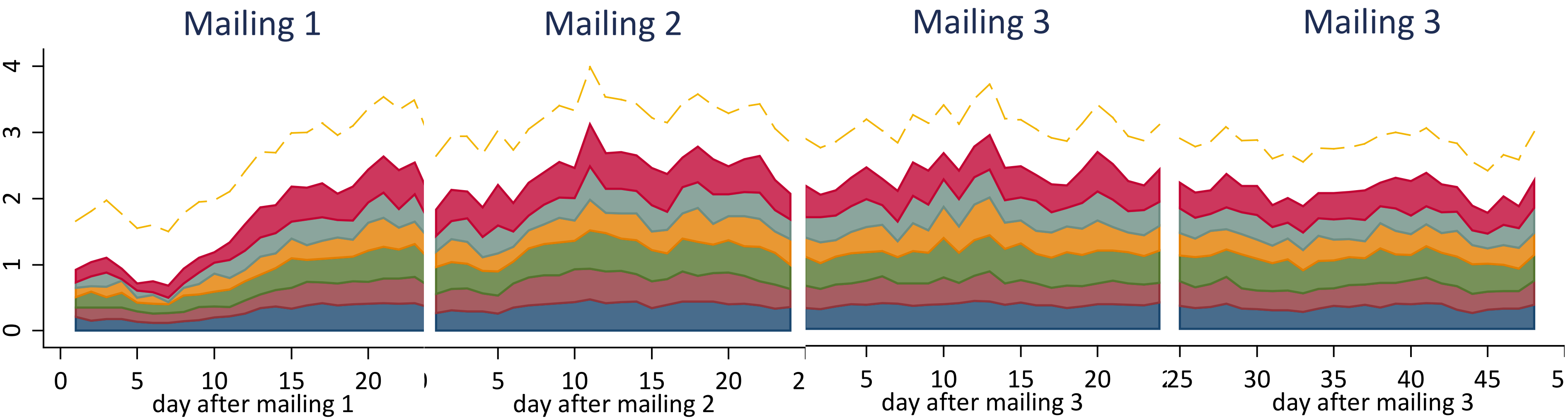

Savings continue between mailings (there are statistically significant savings every day)

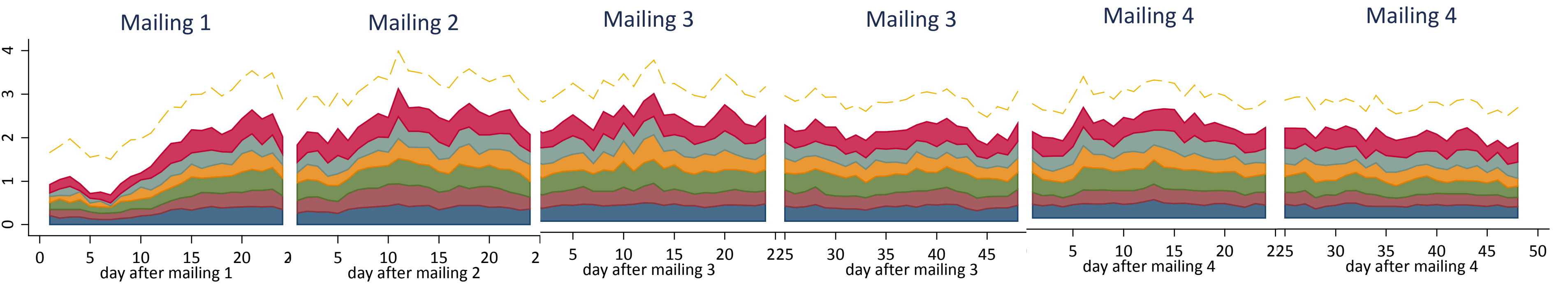


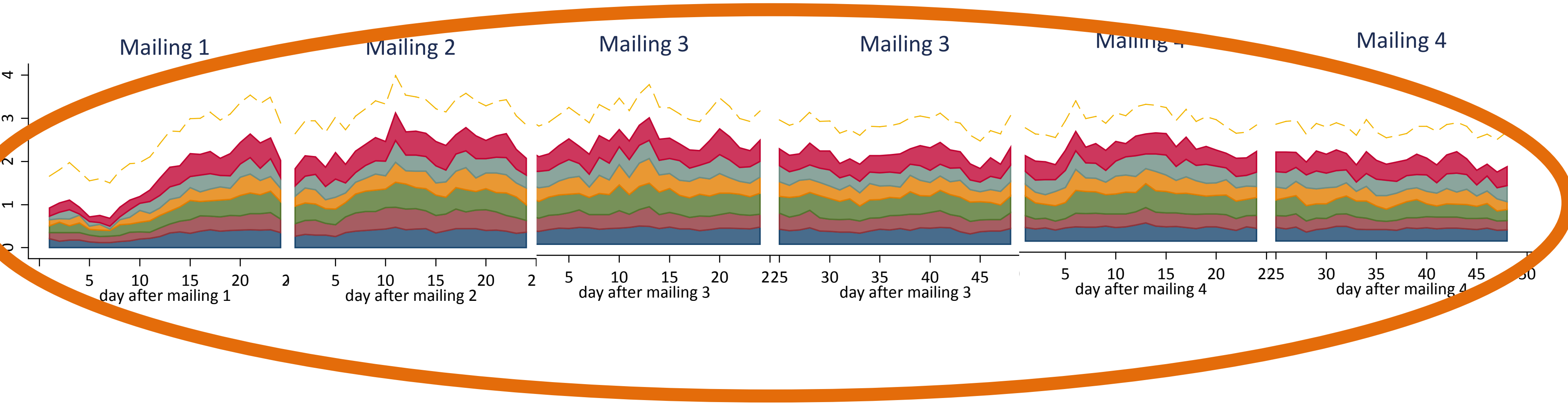

However the level of savings appears to vary somewhat


11pm-3am
3am-7am
7am-11am
11am-3pm
3pm-7pm
7pm-11pm
Total savings +95% CI
Mailing 1
Est savings, % of avg daily kWh
0
1
2
3
4
0
5
10
15
20
25
day after mailing 1


Mailing 1
Mailing 2
0
1
2
3
4
0
5
10
15
20
2
5
10
15
20
25
day after mailing 1
day after mailing 2


Mailing 1
Mailing 2
Mailing 3
Mailing 3
day after mailing 1
day after mailing 2
day after mailing 3
day after mailing 3


Mailing 1
Mailing 2
Mailing 3
Mailing 3
Mailing 4
Mailing 4
0
1
2
3
4
day after mailing 1
day after mailing 2
day after mailing 3
day after mailing 3
day after mailing 4
day after mailing 4




**No discernable pattern across hours**

# Results from analysis of short-term persistence

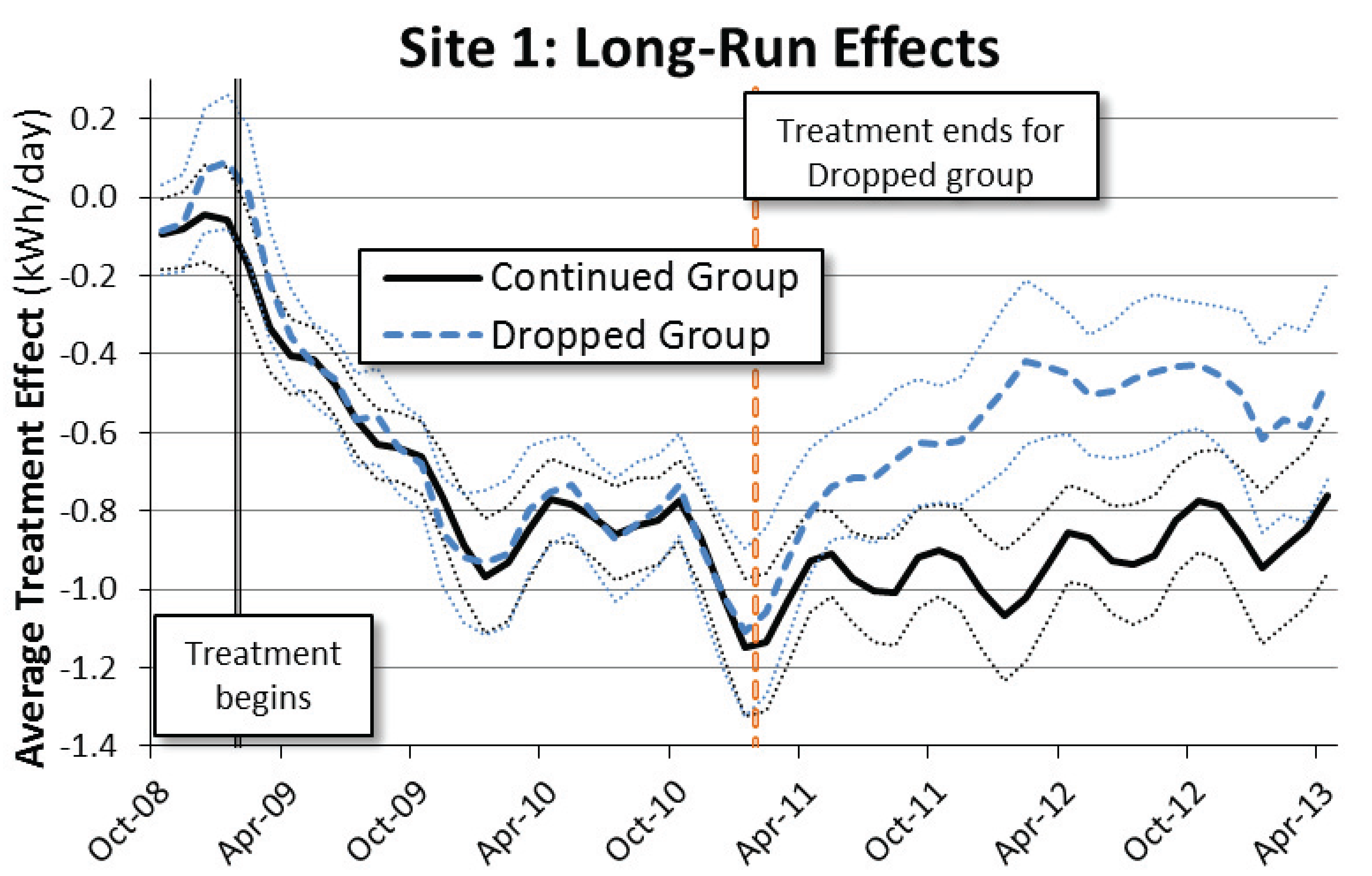
- **Savings continue between mailings (there are statistically significant savings every day)…but the level of savings appears to vary somewhat**
- **No discernable pattern across hours**
- **Quick ramp up rate**
- Suggests that the savings is driven by the initial customer reaction to the first mailing
- Suggests habitual or one-time savings actions rather than changes in installed equipment

# Analysis using smart meter data can help answer key policy questions

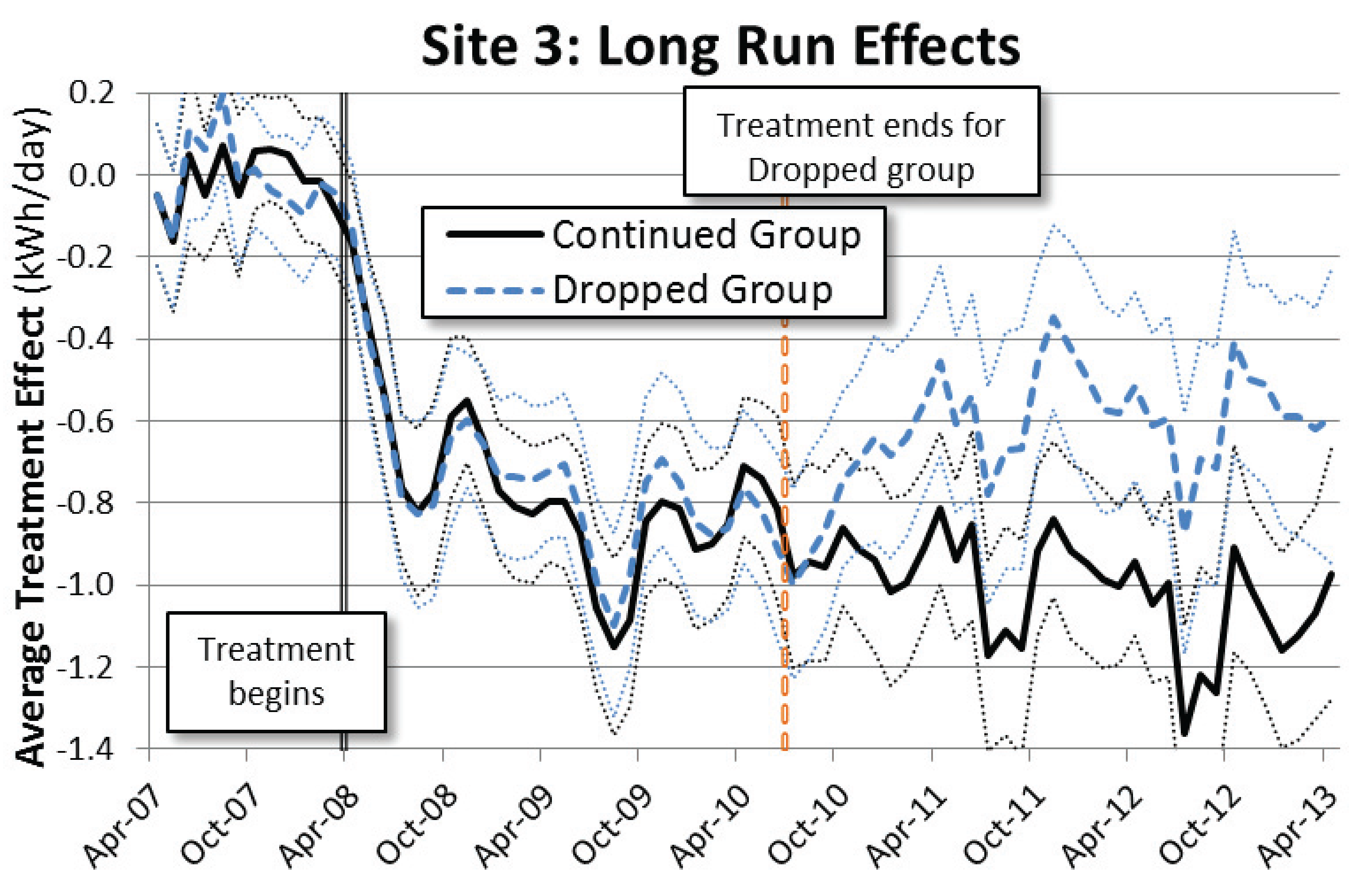
## Key policy questions for HER (and other BB) programs:

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)*

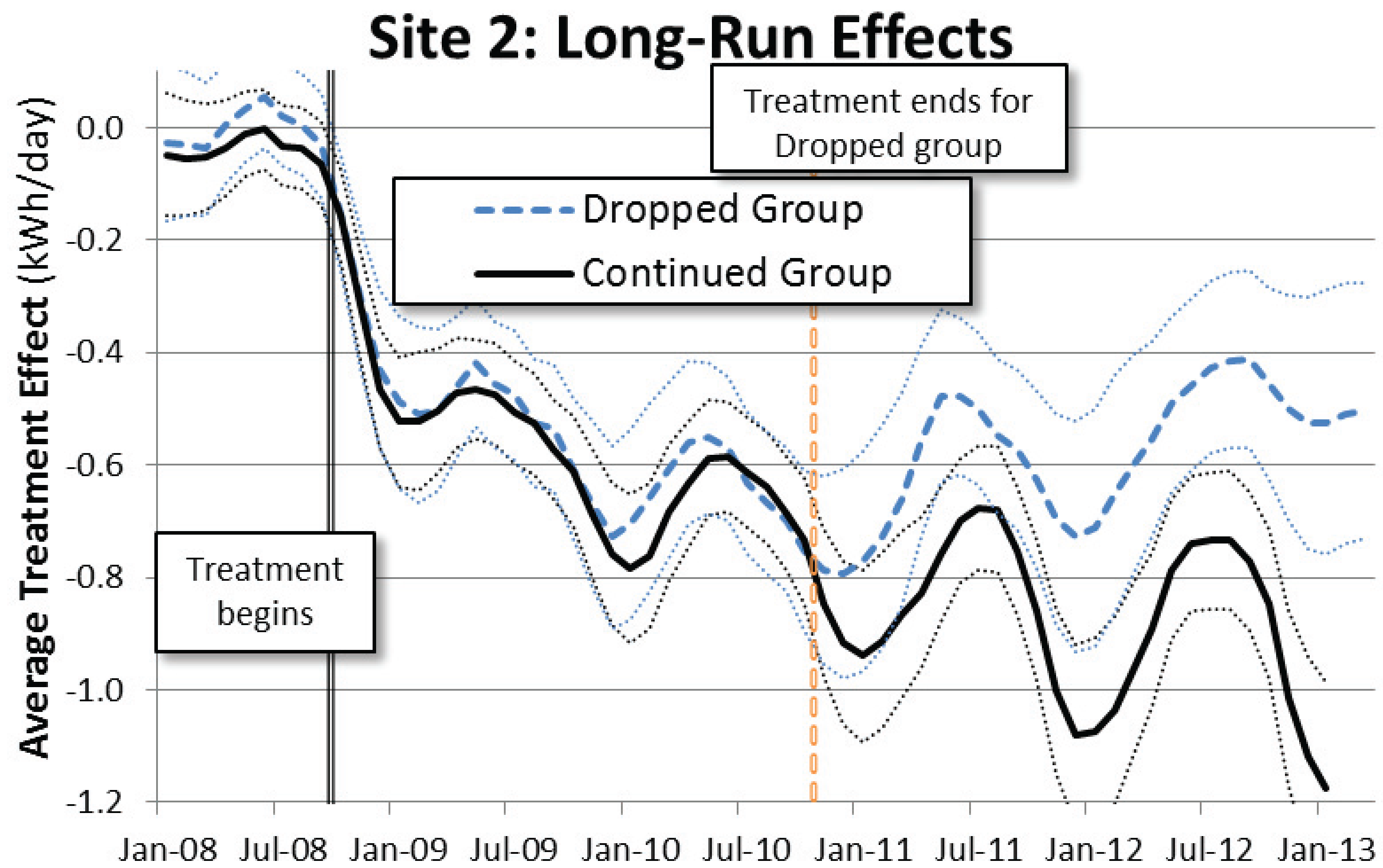
# Allcott and Rogers 2012



# Allcott and Rogers 2012



# Allcott and Rogers 2012



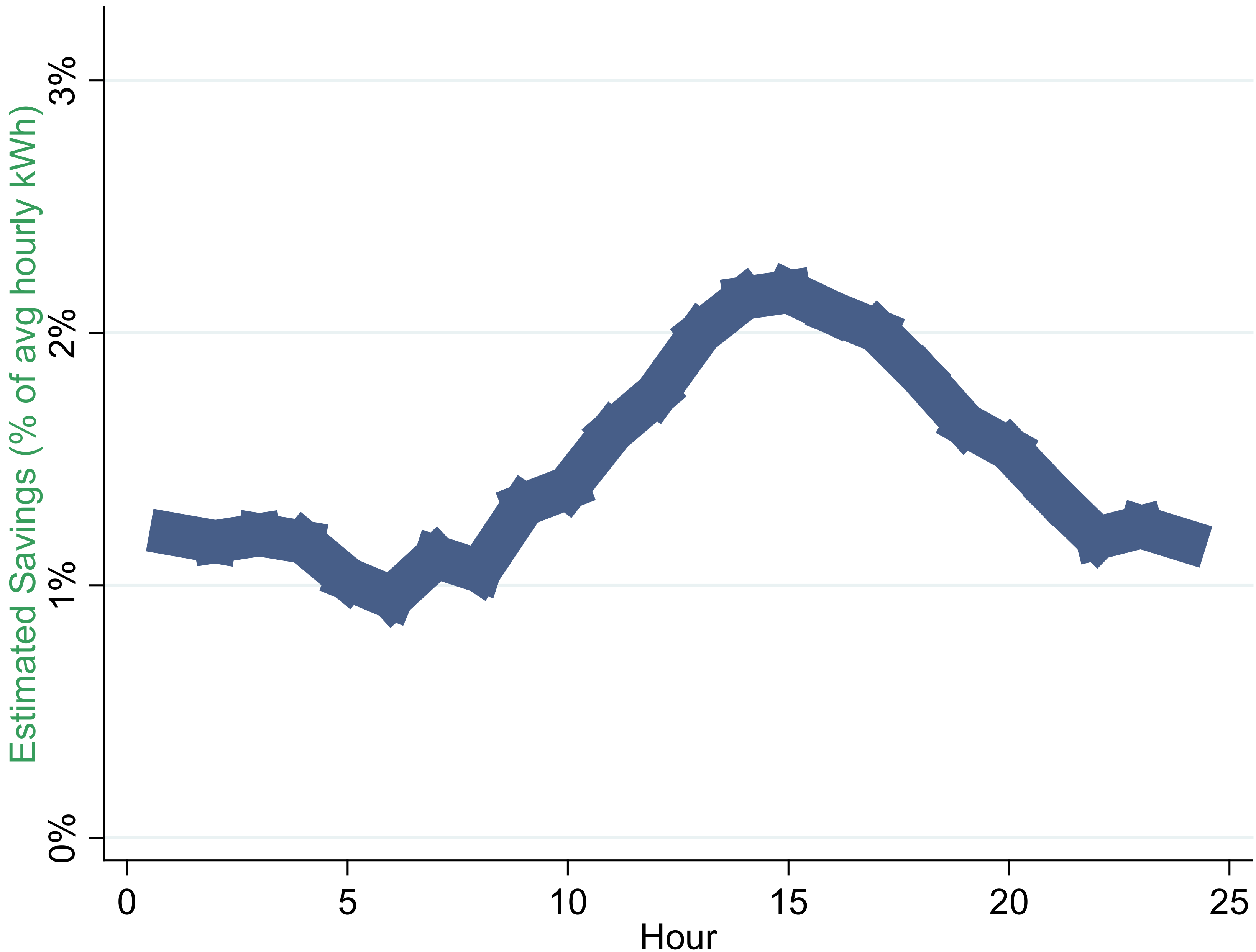
# Recent findings on long-term persistence

- While HER mailings continue, the savings persist
- When the mailings are discontinued, savings decay

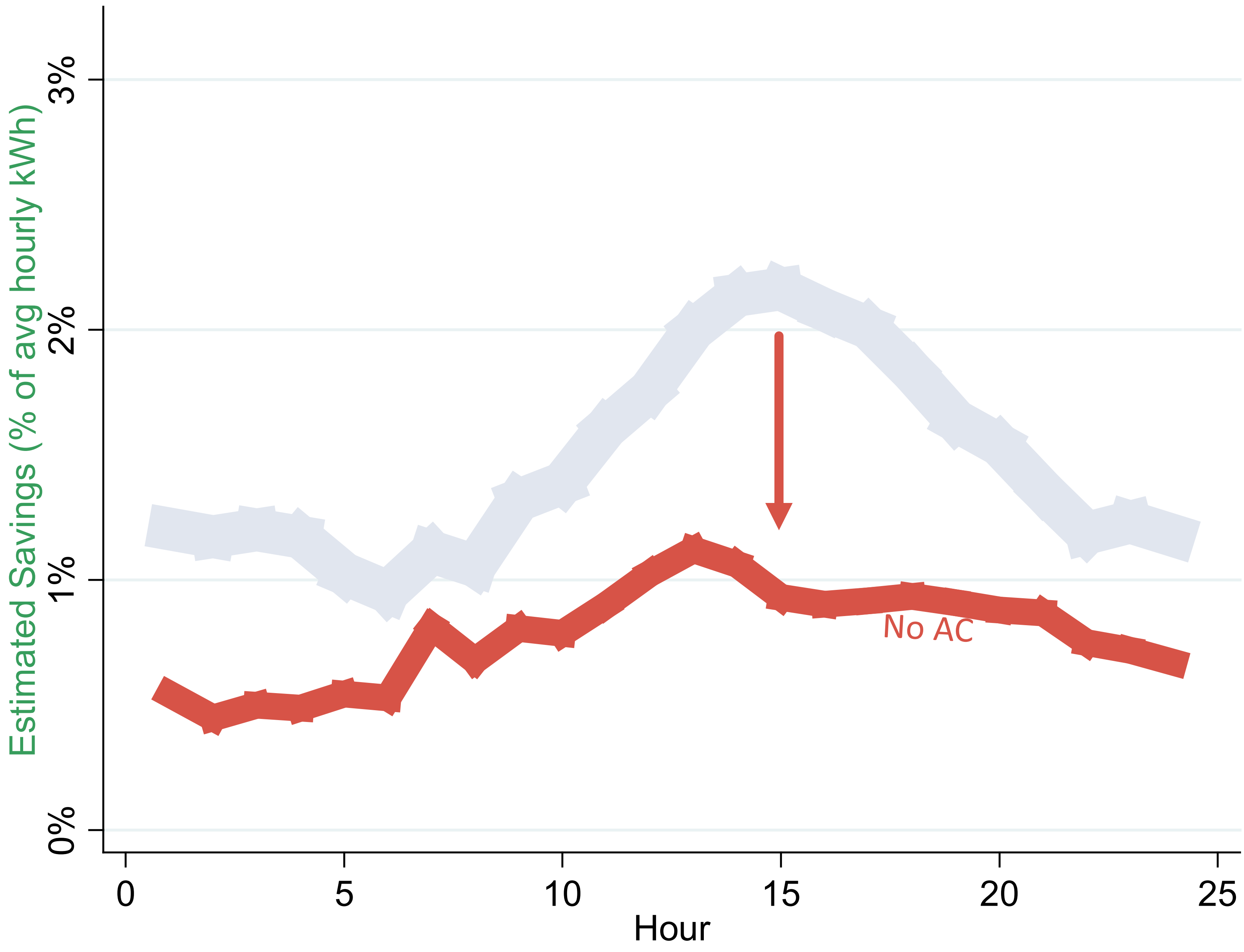
# Analysis using smart meter data can help answer key policy questions

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

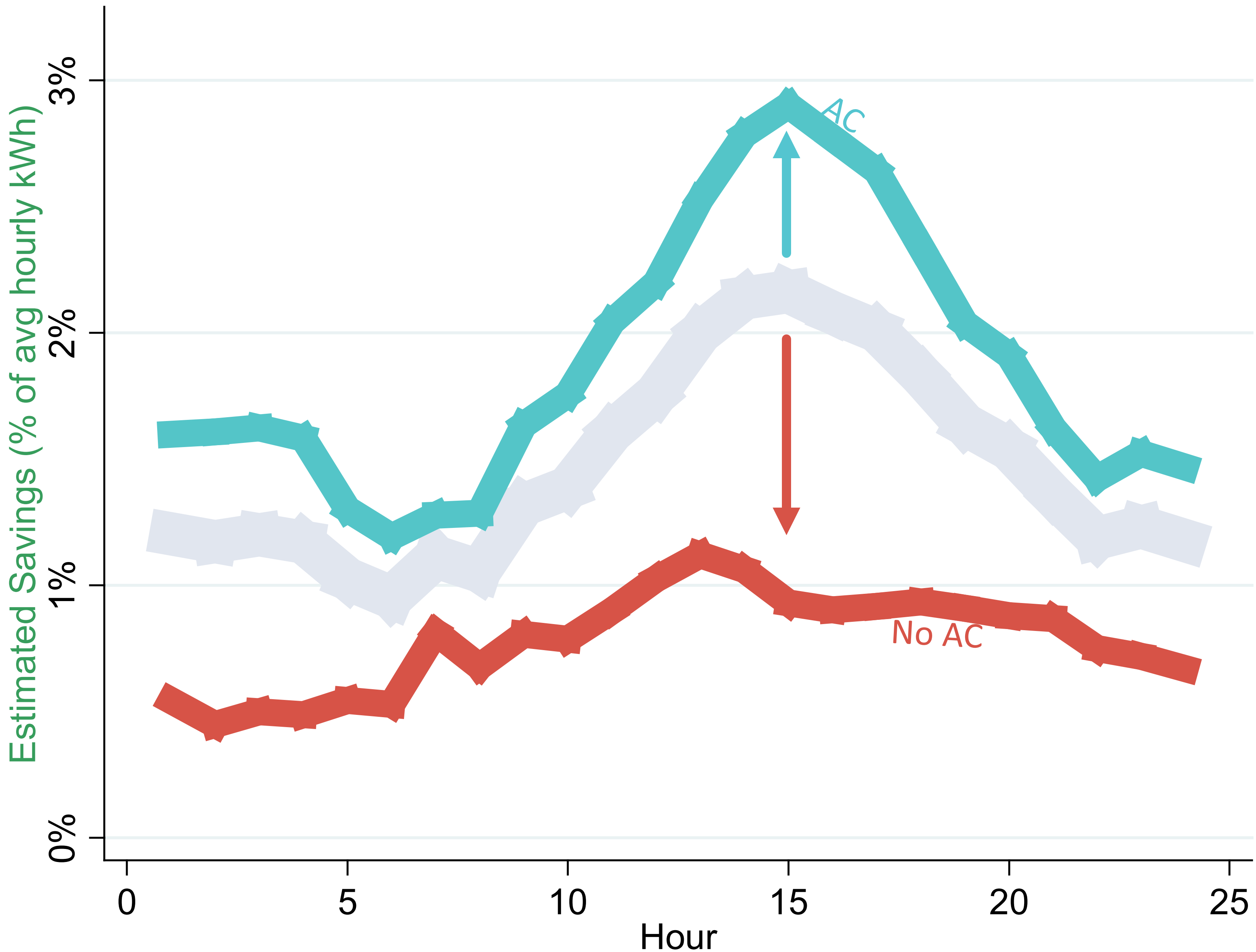
1. What is the short-term persistence of savings? *(Results: savings within one-two weeks after first report mailed, stabilize after second report)*
2. What is the long-term persistence of savings? *(Results: savings persist while mailings continue; savings decay after reports are discontinued)*
3. **What actions and characteristics are related to savings?** *(Results: suggestive of AC – best guess: changing thermostat set point)*



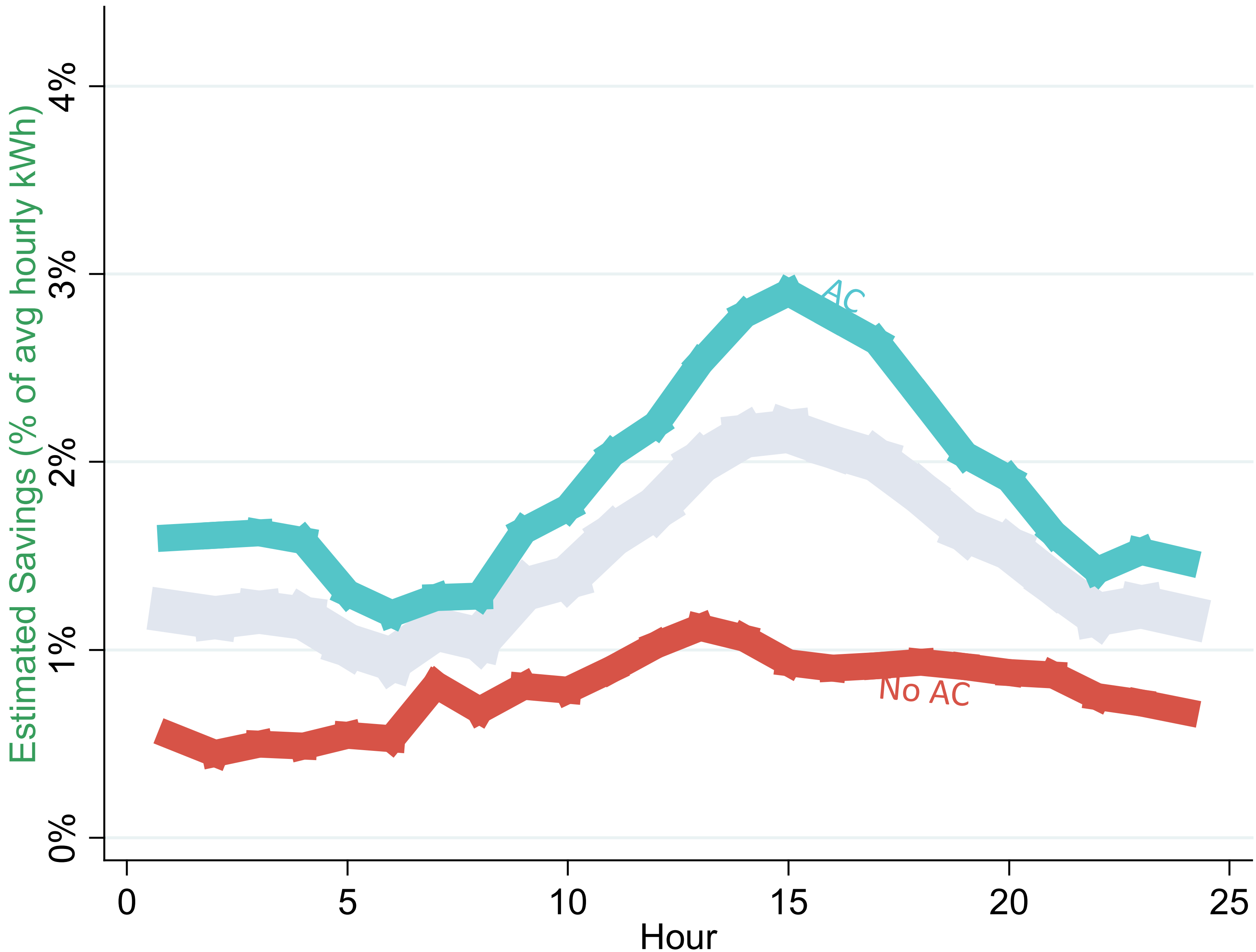
Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
0
5
10
15
20
25
Hour



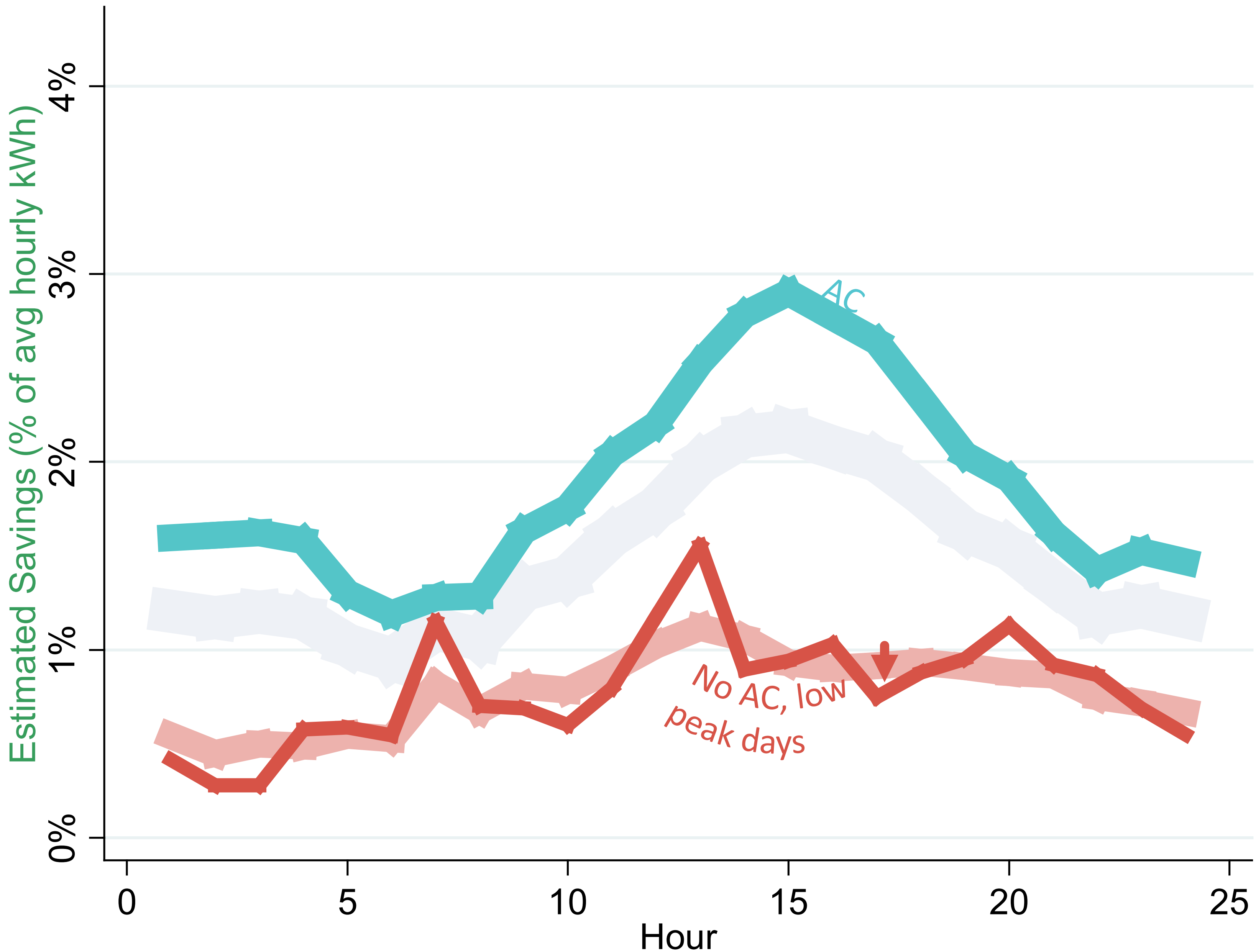
Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
0
5
10
15
20
25
Hour
No AC



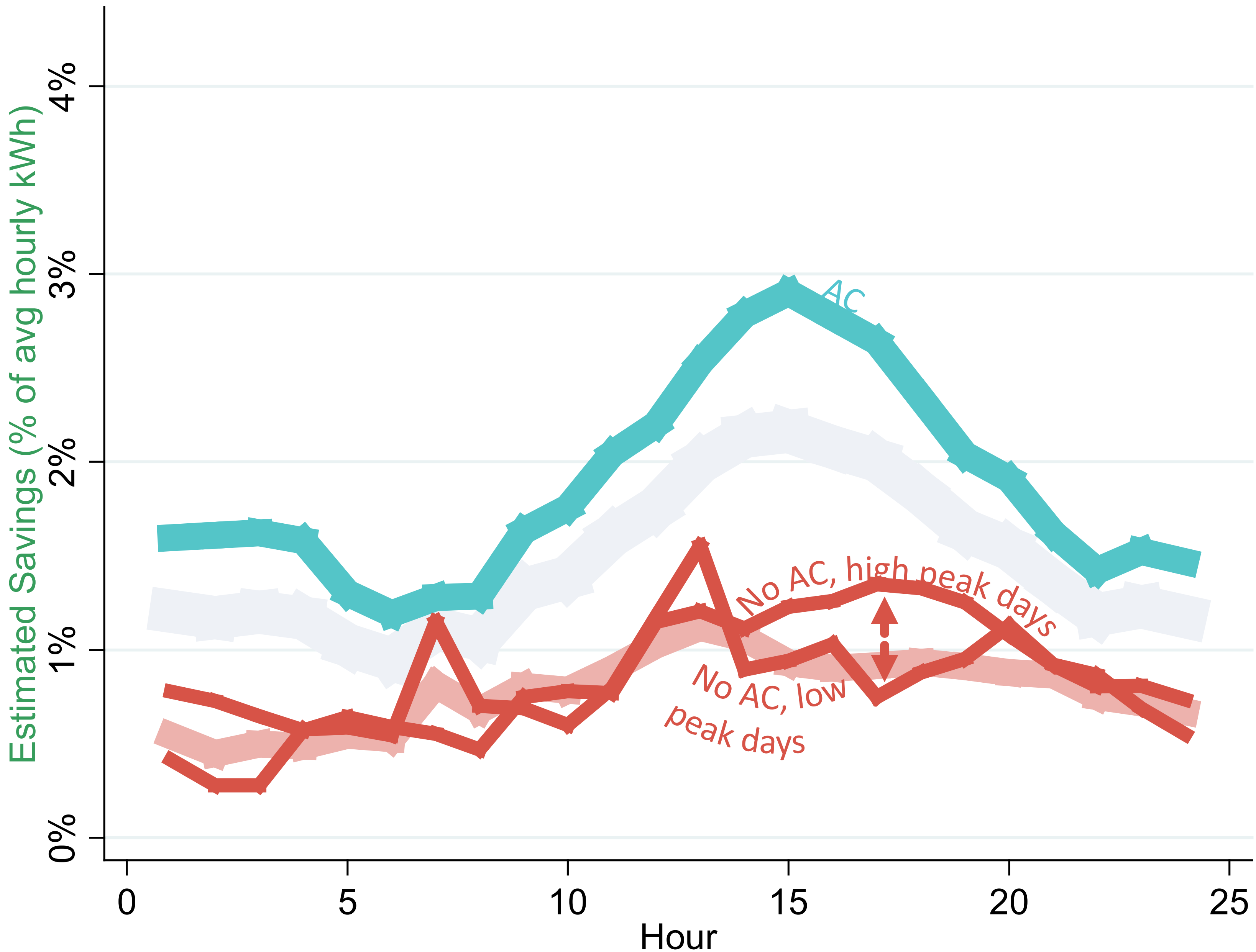
Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
AC
No AC
0
5
10
15
20
25
Hour



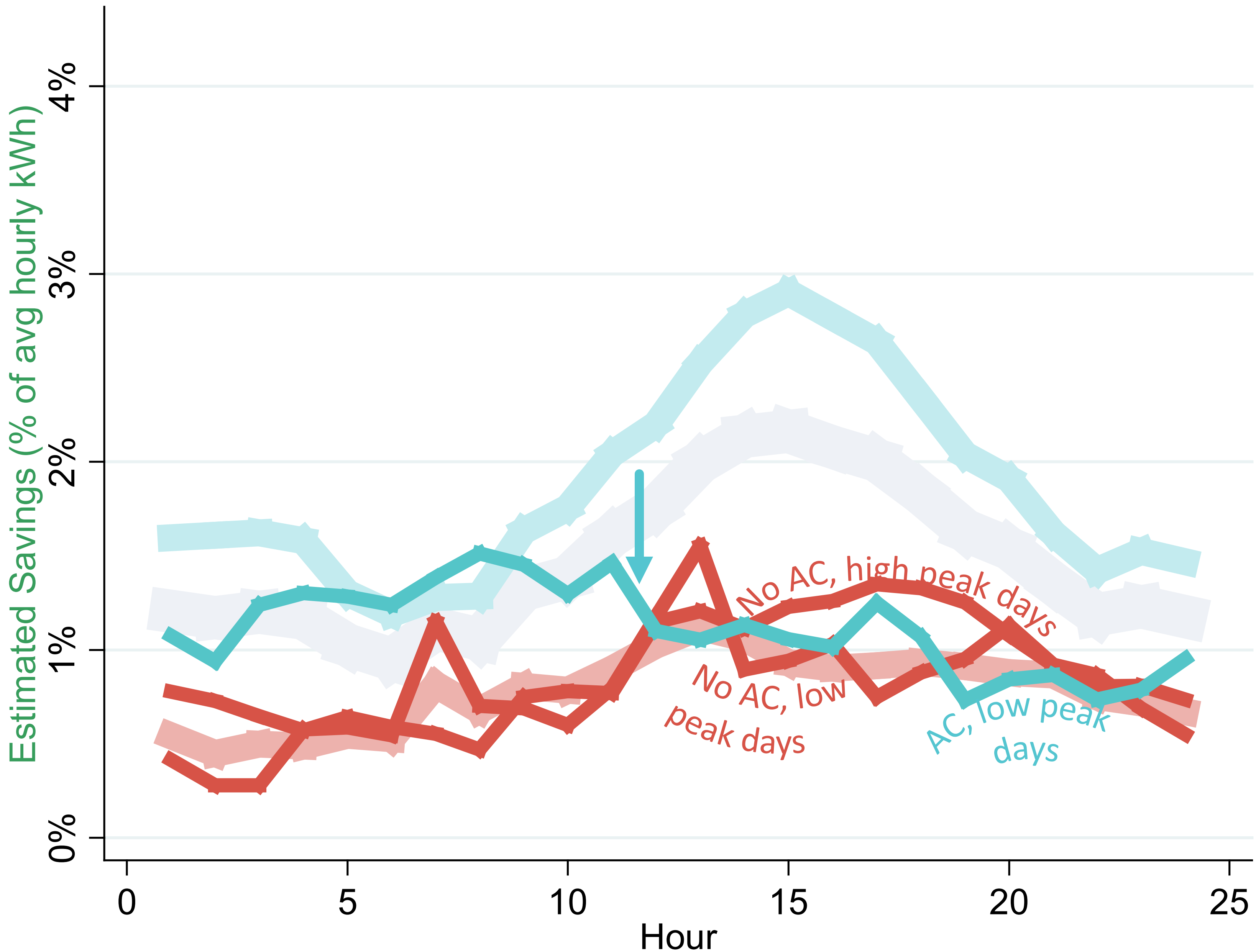
Estimated Savings (% of avg hourly kWh)
4%
3%
2%
1%
0%
AC
No AC
0
5
10
15
20
25
Hour



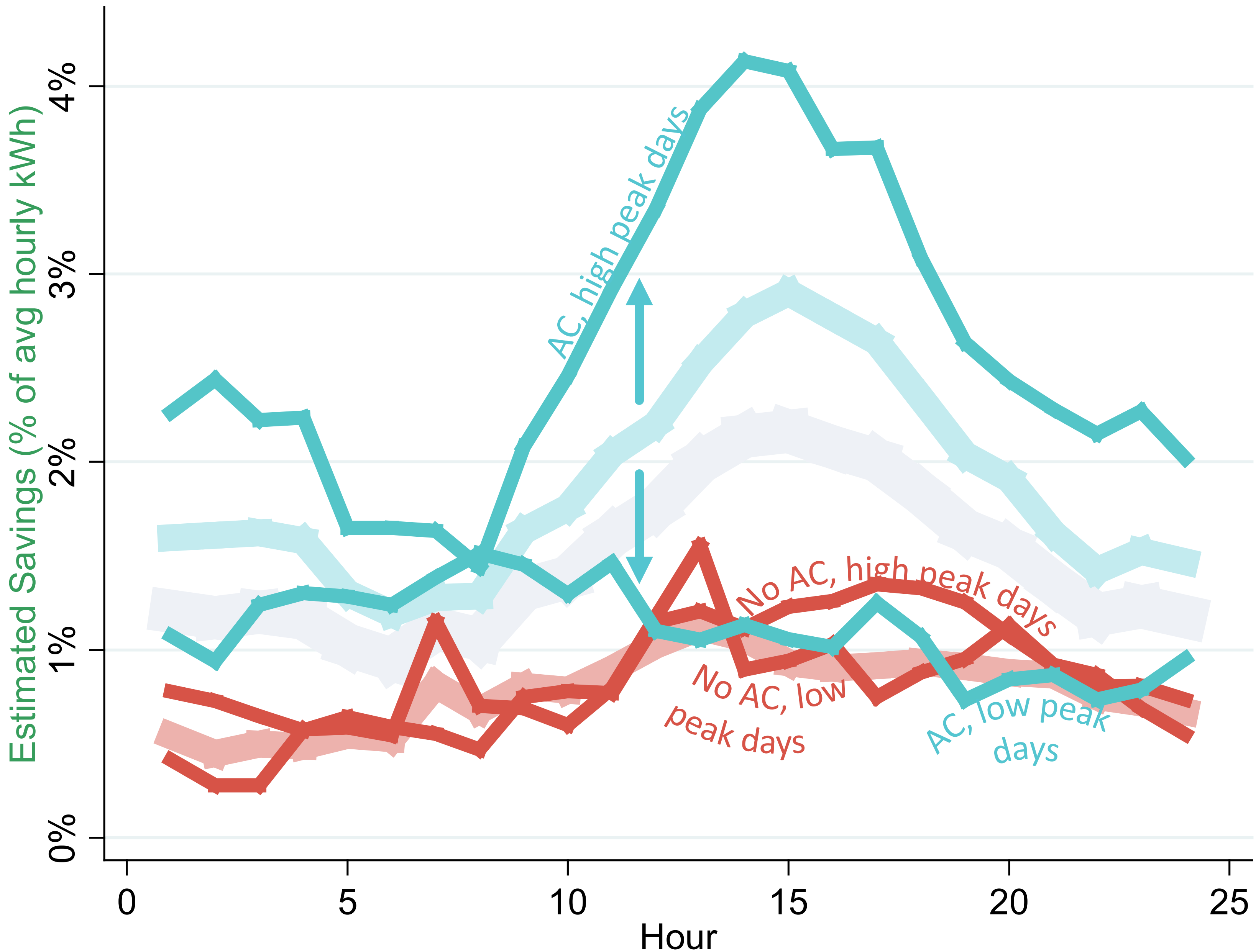
Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
4%
0
5
10
15
20
25
Hour
AC
No AC, low
peak days

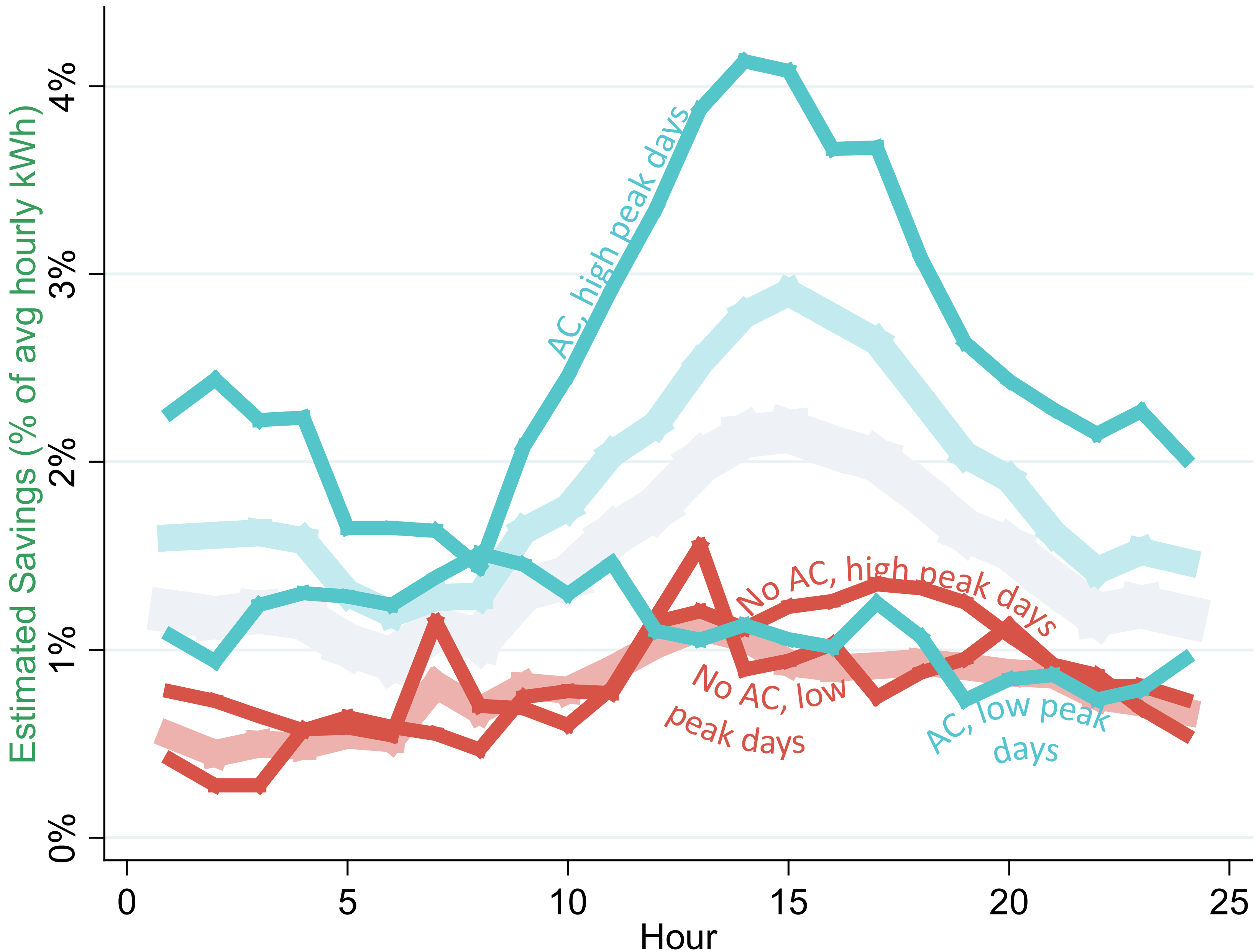


Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
4%
0
5
10
15
20
25
Hour
AC
No AC, high peak days
No AC, low peak days

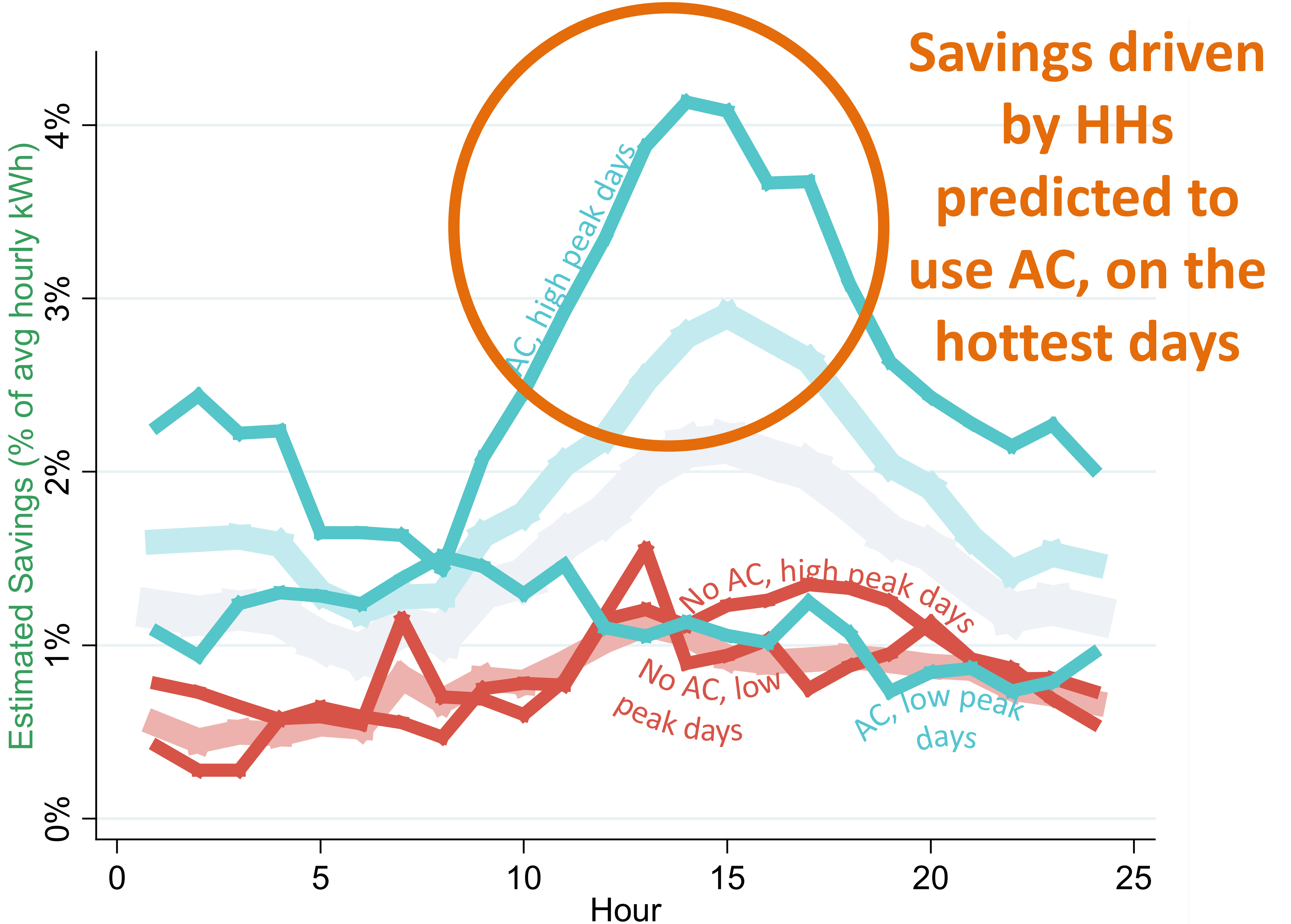


Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
4%
0
5
10
15
20
25
Hour
No AC, high peak days
No AC, low peak days
AC, low peak days


Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
4%
0
5
10
15
20
25
Hour
AC, high peak days
No AC, high peak days
No AC, low peak days
AC, low peak days



Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
4%
0
5
10
15
20
25
Hour
AC, high peak days
No AC, high peak days
No AC, low peak days
AC, low peak days


Savings driven by HHs predicted to use AC, on the hottest days
Estimated Savings (% of avg hourly kWh)
0%
1%
2%
3%
4%
0
5
10
15
20
25
Hour
AC, high peak days
No AC, high peak days
No AC, low peak days
AC, low peak days

# Results from analysis of actions and characteristics related to savings

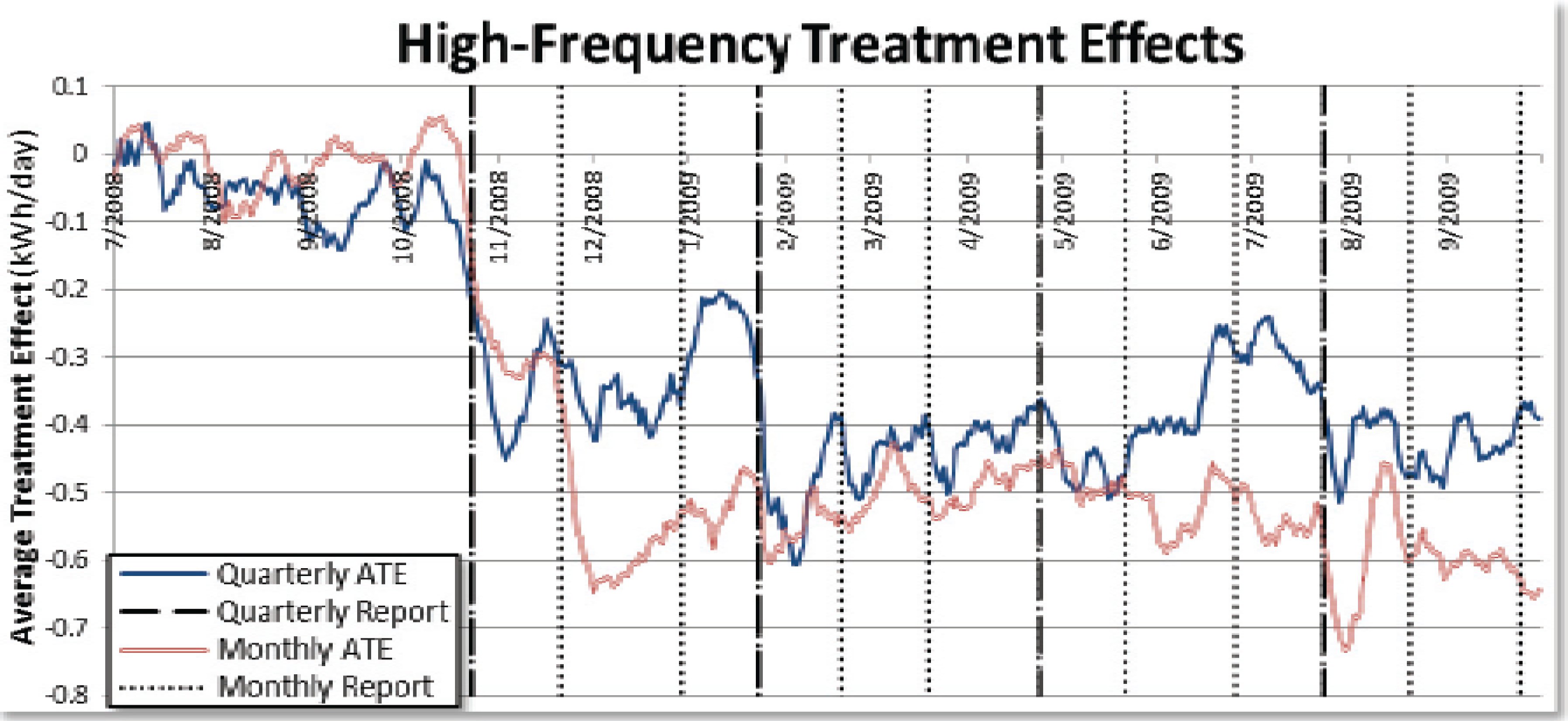
- **Savings are driven by:**
  - **Households predicted to use AC**
  - **During peak hours**
  - **On the hottest days**
- Suggests that the actions that drive the savings are related to AC - our best guess is that people are turning up the settings on their AC (for this particular program rollout)
- What does this mean for measure life / persistence? Do these habits last, or do they need constant reminders?

# Results and Conclusions

- Savings ramp up quickly, and persist while mailings continue
- Savings decay when mailings are discontinued
- For this example, our best guess is that much of the savings from these programs is driven by one-time or habitual actions (such as changing AC settings) rather than equipment purchase.
- Other research (Allcott 2014) shows that savings, actions that drive the savings, and the persistence of savings **vary widely** across utilities and across customer groups, and **are not at all predictable**

# Results and Conclusions

- **For the purposes of claiming savings, can we assign a pre-determined measure life / deemed savings to these programs?** ***(No –there are too many differences in savings, actions related to savings, and persistence across utilities and between households)***
- **For the purposes of cost-effectiveness and future program planning, should we assume that savings persist after mailings are discontinued?** ***(Yes – the savings likely persist after the program is discontinued)***
- **Key remaining question: what is the optimal program design? What is the optimal timing between mailings?**


High-Frequency Treatment Effects
Average Treatment Effect (kWh/day)
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
7/2008
8/2008
9/2008
10/2008
11/2008
12/2008
1/2009
2/2009
3/2009
4/2009
5/2009
6/2009
7/2009
8/2009
9/2009
Quarterly ATE
Quarterly Report
Monthly ATE
Monthly Report

- Examples of how to use Smart meter data to look at persistence to answer these questions
- \- Can we reliably measure the savings from behavior based programs?
  \- How long do savings last once the program stops?
  \- Does the length of the program impact how long the savings last?
  \- What should a regulator consider regarding the "measure life" of home energy reports?
- Part 1 – Measuring Savings Methodologies
- 1. Ways we recommend to measure savings
- 2. The pros and cons of each recommended method
- 3. Other industries that use these methods (i.e. drug development)
- 4. Methods not recommended
- 5. Reasons why they aren't recommended
- 6. How savings are being measured by PUCs (with some examples – we can maybe collaborate with OD on this)
- 
- Part 2 – What we know about measure life and persistence
- 1. Define persistence and measure life (OD might do this if they go first)
- 2. What you were attempting to learn with the smart meter data analysis
- 3. What you found out
- 4. What this tells or doesn't tell you about persistence and measure life

# Recent findings on prediction of savings

- **Can we create a model that accurately predicts program savings impacts for a behavior-based program implemented by utilities in a new area?**
- **Can program savings estimates from a behavior-based program be extrapolated to a new population (i.e., other utility sites)?**

# Hunt

- Analyze 14 nearly identical behavior-based, EE programs that utilized RCT (OPOWER pilots)

- Results: Predictive model NOT possible
  - OPOWER program savings impacts vary widely
    - Savings between 1.37% and 3.32%; average of 2.03%
    - Cost per annual saved kWh (to utility) ranges between 1.28¢/kwh and 5.36 ¢/kwh; average of 3.31 ¢/kwh
      - *CAG: Add ~15-25% to include EE pgm administrator and EM&V costs*
  - These differences can't be explained by observable variables
    - Household level demographics, weather, energy use
  - Characteristics of "early adopter" utilities differ significantly from non-partner sites – makes it more problematic to extrapolate ATE (i.e. % savings)

# Smart meter data enables new types of analysis

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

# Three analyses – each focusing on one key policy question

Key policy questions:

1. Do these programs have potential to provide peak-hour savings?
   **Analysis 1:** Estimate the hour-by-hour savings profile *(Wave One – late summer)*
2. What actions and characteristics are related to savings?
   **Analysis 2:** segment by customer characteristics to identify "high-savers" *(Wave One – late summer)*
3. What is the short-term persistence of savings?
   **Analysis 3:** segment across days after reports are mailed *(Gamma – winter and spring)*

# Outline

- Smart meter data enables many opportunities for new forms of analysis
- Purpose of this study: focus on one particular aspect of this analysis enabled by smart meters – what insights can we gain into Home Energy Report (HER) programs?
- Description of HERs, data, limitations of report
- Analyses and results
- **Conclusions and future research**

# Limitations of the report

- Limited data access
- Limited time period
- Only a few rollouts

# Summary

- Lots of smart meter data
- Opportunity for new types of analysis
- Today – one example of the value of this data
- Results:
  1. Potential for peak-hour savings
  2. Savings related to AC
  3. Savings increase within one week, then stable

# Policy Implications and Next Steps (discussion)

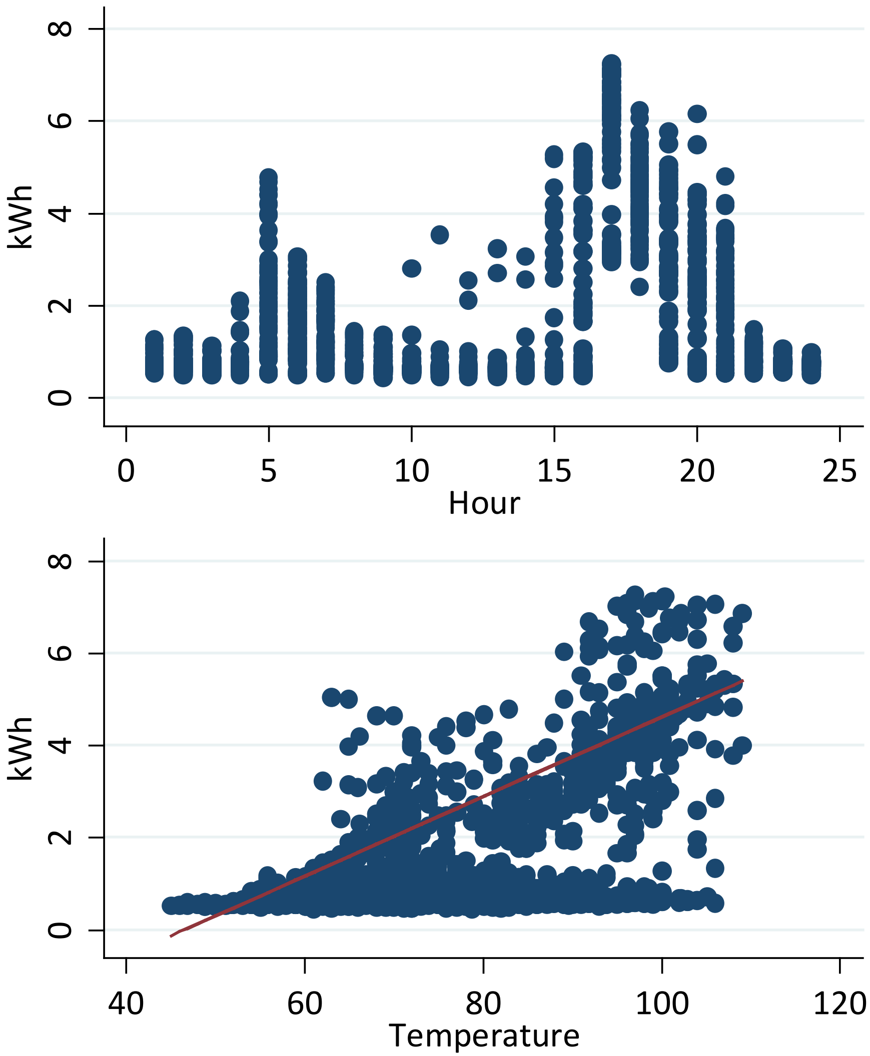
- Policy Implications:
  - Peak savings resource: possible integration of BB programs
  - Targeting:
    - If high total savings is the goal and money is no object, include all customers
    - If the goal is cost-effectiveness, target AC customers to get higher savings per dollar spent
  - EM&V implications for measure life, double counted savings
  - Implications for optimal, cost-effective program planning
- This is primary research – one of the first to look at hourly smart meter data for BB programs
  - Need to replicate this research in different locations, different situations, different years

# Future Research

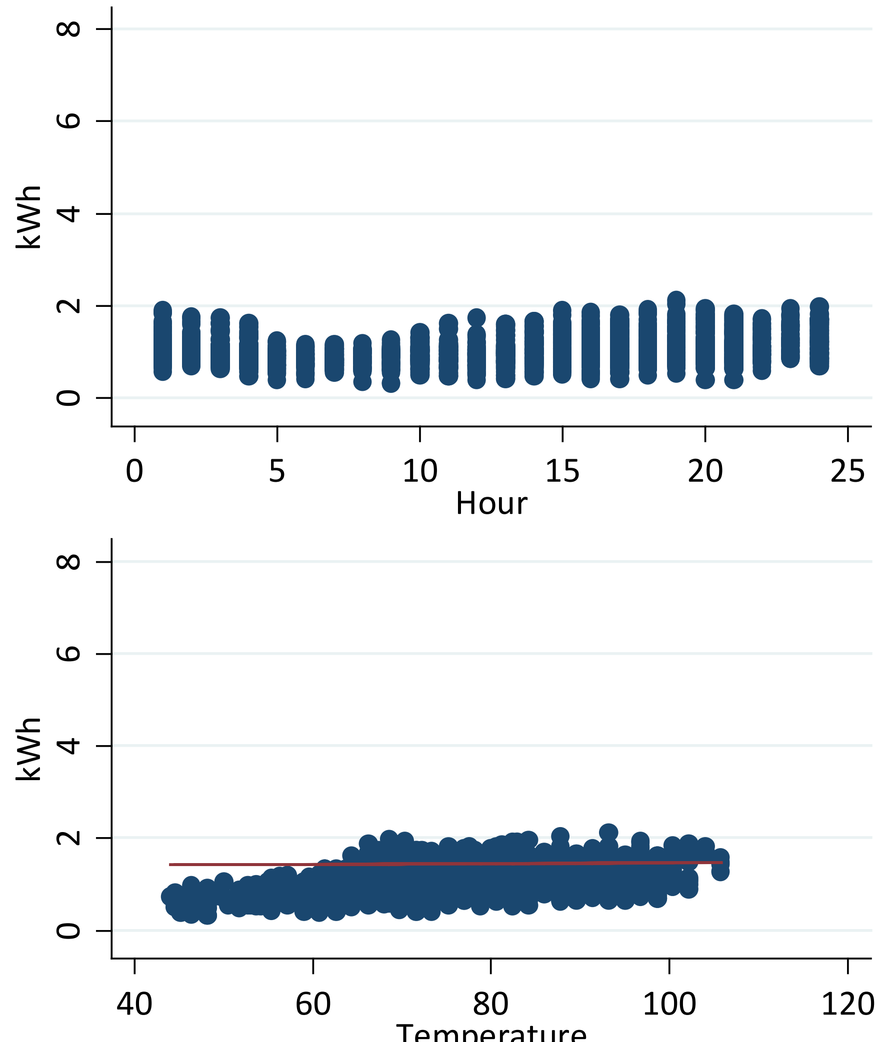
- Replicate in other situations, other programs
- Many other examples of the value of this data
- Future – a lot of potential research

# Which household has air conditioning?

## Household A

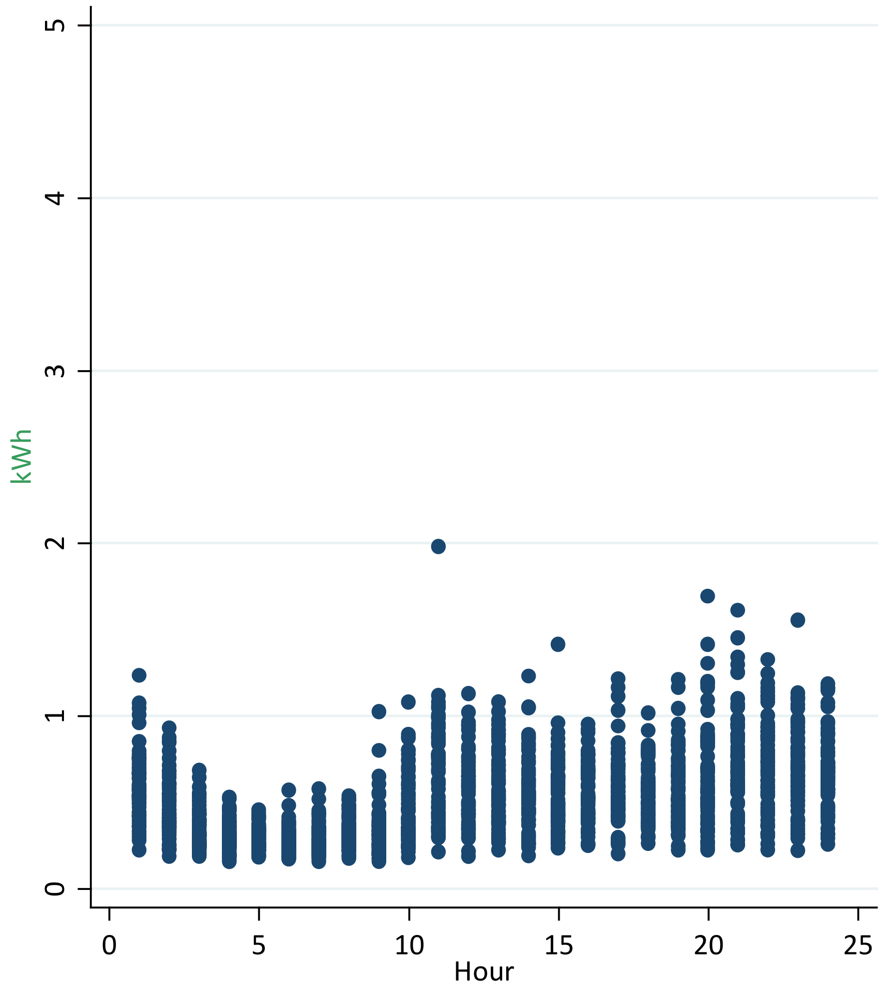



## Household B

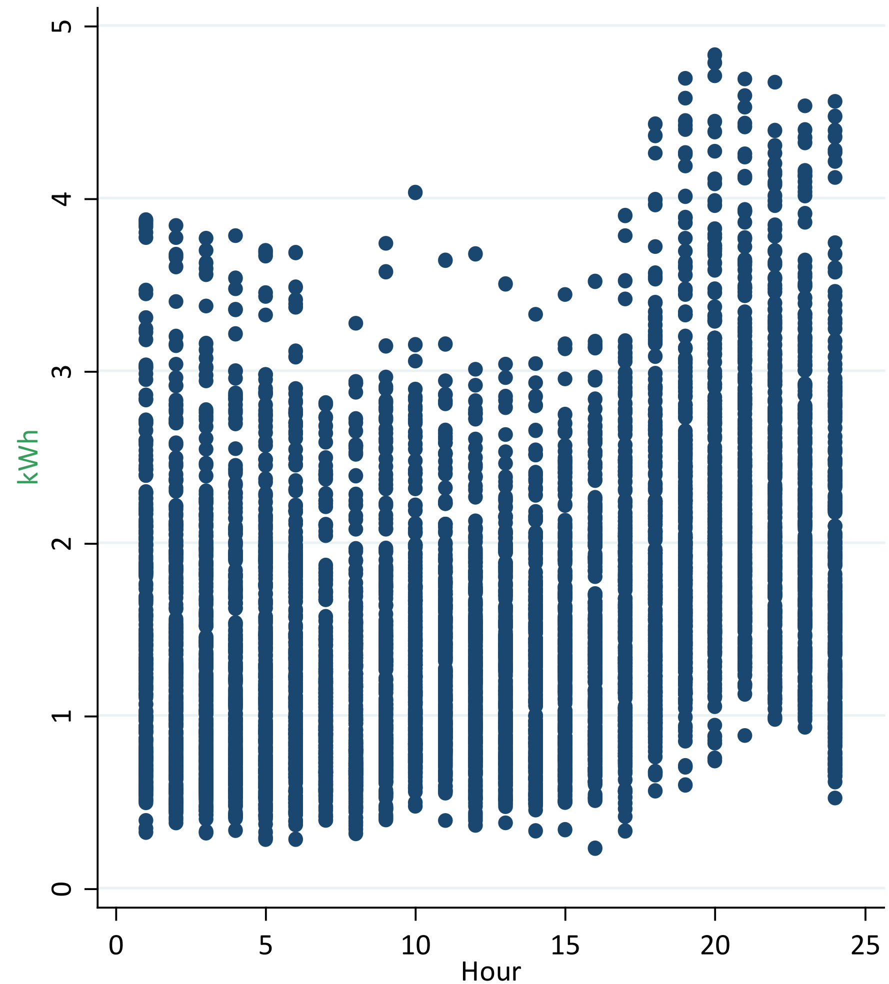


# Which household has a higher baseline usage?

## Household A




## Household B



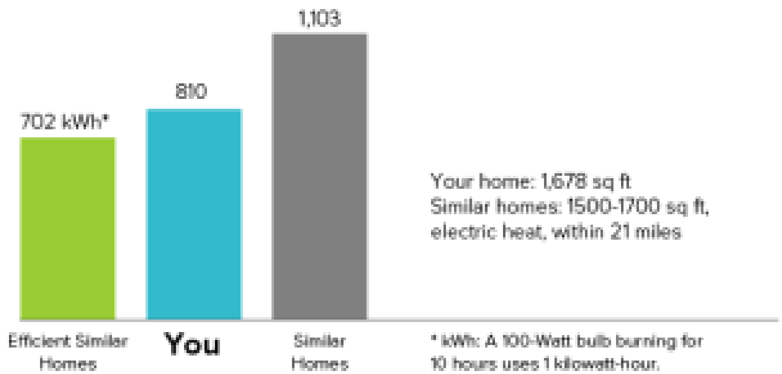
# Today

- Insights valuable for a range of energy programs

- Insights from smart meters: **focus on behavior-based (BB) programs**
  - Specifically: Home Energy Report (HER) programs
  - An illustrative example of the value of this analysis

# What is a HER program?

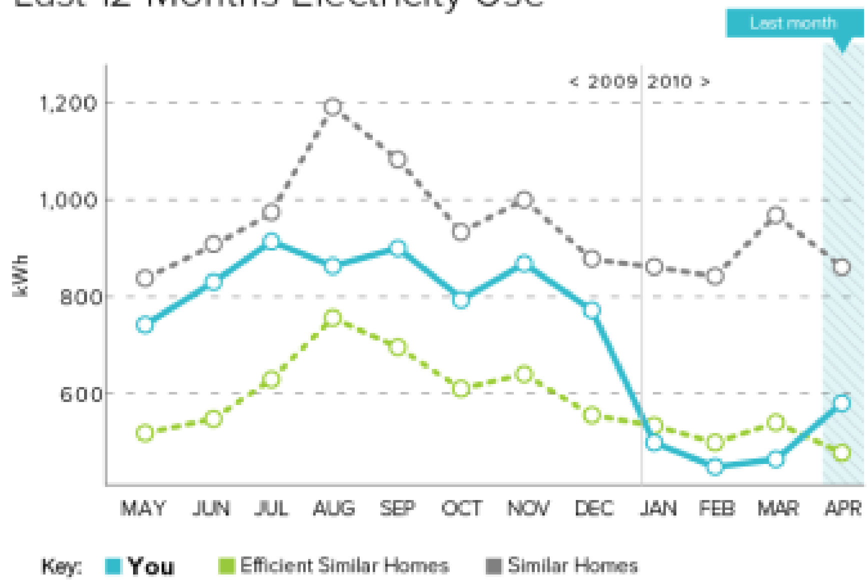
## Last Month Electricity Use




Your home: 1,678 sq ft
Similar homes: 1500-1700 sq ft, electric heat, within 21 miles

* kWh: A 100-Watt bulb burning for 10 hours uses 1 kilowatt-hour.

## Last 12 Months Electricity Use




**Welcome to your first home energy report.**
This report is part of a free program to help you save money and energy.

How you're doing:

Great

**Good**

Using more than average

We estimate that you could **save $150** each year.

**Turn over** for ways to save

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. **Do these programs have potential to provide peak-hour savings?** ***(Yes – for our dataset)***
2. **What actions and characteristics are related to savings?** *(Suggestive of AC – best guess: changing thermostat set point)*
3. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. **Do these programs have potential to provide peak-hour savings?** *(Yes – for our dataset)*
2. **What actions and characteristics are related to savings?** *(Suggestive of AC – best guess: changing thermostat set point)*
3. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. Do these programs have potential to provide peak-hour savings? *(Yes – for our dataset)*
2. **What actions and characteristics are related to savings?** ***(Suggestive of AC – best guess: changing thermostat set point)***
3. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

# Key policy questions for HER (and BB) programs

**Key policy questions:**

1. Do these programs have potential to provide peak-hour savings? *(Yes – for our dataset)*
2. What actions and characteristics are related to savings? *(Suggestive of AC – best guess: changing thermostat set point)*
3. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. Do these programs have potential to provide peak-hour savings? *(Yes – for our dataset)*
2. What actions and characteristics are related to savings? *(Suggestive of AC – best guess: changing thermostat set point)*
3. **What is the short-term persistence of savings?** ***(Savings within one-two weeks after first report mailed, stabilize after second report)***

# Outline

- Smart meter data enables many opportunities for new forms of analysis
- Purpose of this study: focus on one particular aspect of this analysis enabled by smart meters – what insights can we gain into Home Energy Report (HER) programs?
- **Description of data**
- Analyses and results
- Conclusions and future research

# Outline

- Smart meter data enables many opportunities for new forms of analysis
- Purpose of this study: focus on one particular aspect of this analysis enabled by smart meters – what insights can we gain into Home Energy Report (HER) programs?
- Description of HERs, data, limitations of report
- **Analyses and results**
- Conclusions and future research

# Three analyses – each focusing on one key policy question

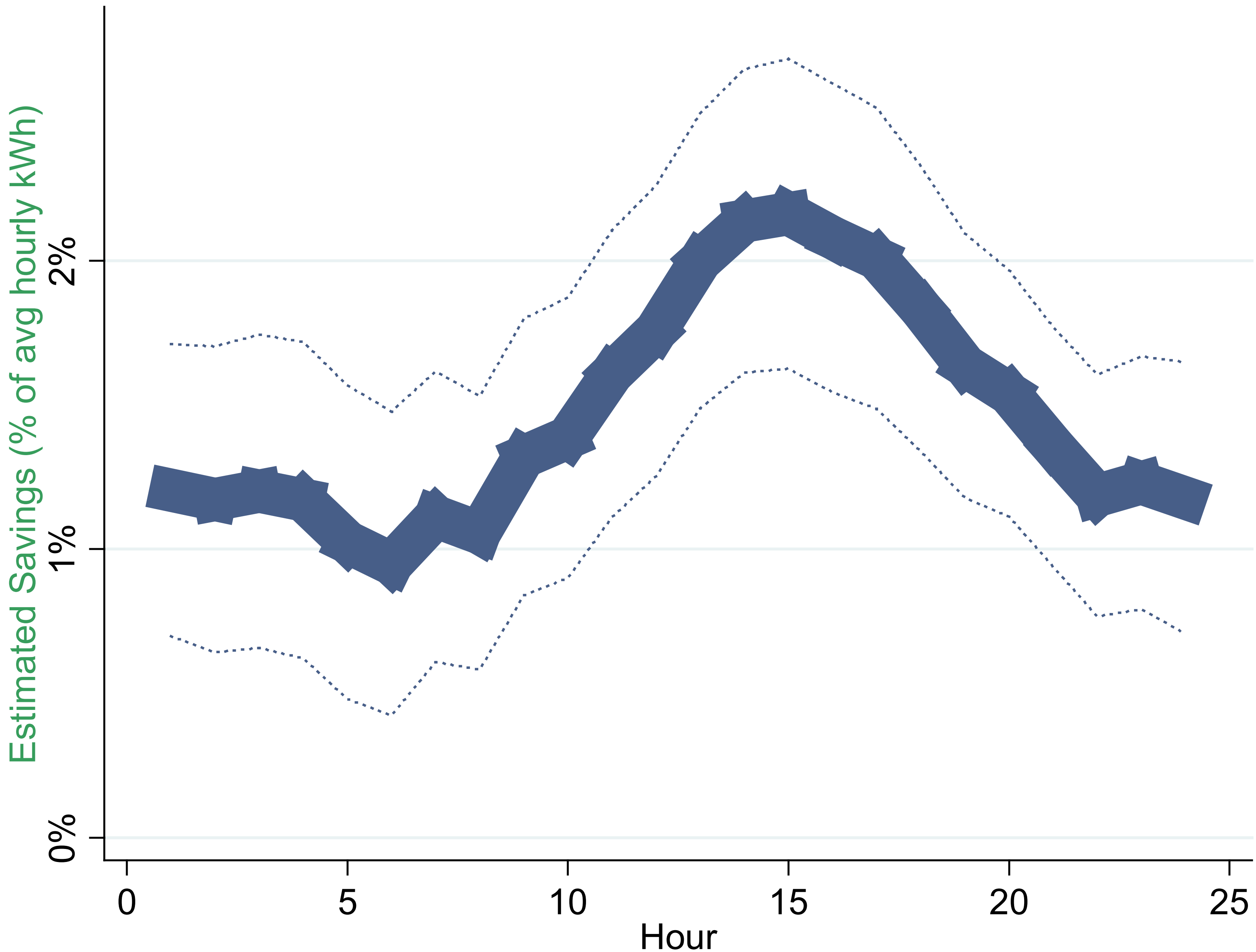
Key policy questions:

1. Do these programs have potential to provide peak-hour savings?
   **Analysis 1:** Estimate the hour-by-hour savings profile *(Wave One – late summer)*
2. What actions and characteristics are related to savings?
   **Analysis 2:** segment by customer characteristics to identify “high-savers” *(Wave One – late summer)*
3. What is the short-term persistence of savings?
   **Analysis 3:** segment across days after reports are mailed *(Gamma – winter and spring)*

# Three analyses – each focusing on one key policy question

Key policy questions:

1. Do these programs have potential to provide peak-hour savings?
   **Analysis 1:** Estimate the hour-by-hour savings profile *(Wave One – late summer)*
2. What actions and characteristics are related to savings?
   **Analysis 2:** segment by customer characteristics to identify "high-savers" *(Wave One – late summer)*
3. What is the short-term persistence of savings?
   **Analysis 3:** segment across days after reports are mailed *(Gamma – winter and spring)*


Estimated Savings (% of avg hourly kWh)
0%
1%
2%
0
5
10
15
20
25
Hour

# Three analyses – each focusing on one key policy question

Key policy questions:

1. Do these programs have potential to provide peak-hour savings?
   **Analysis 1:** Estimate the hour-by-hour savings profile *(Wave One – late summer)*

2. What actions and characteristics are related to savings?
   **Analysis 2:** segment by customer characteristics to identify “high-savers” *(Wave One – late summer)*

3. What is the short-term persistence of savings?
   **Analysis 3:** segment across days after reports are mailed *(Gamma – winter and spring)*

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. **Do these programs have potential to provide peak-hour savings?** *(Yes – for our dataset)*
2. **What actions and characteristics are related to savings?** *(Suggestive of AC – best guess: changing thermostat set point)*
3. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. **Do these programs have potential to provide peak-hour savings?** *(Yes – for our dataset)*
2. **What actions and characteristics are related to savings?** *(Suggestive of AC – best guess: changing thermostat set point)*
3. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. **Do these programs have potential to provide peak-hour savings?** *(Yes – for our dataset)*
2. **What actions and characteristics are related to savings?** *(Suggestive of AC – best guess: changing thermostat set point)*
3. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. Do these programs have potential to provide peak-hour savings? *(Yes – for our dataset)*
2. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

# Key policy questions for HER (and BB) programs

## Key policy questions:

1. Do these programs have potential to provide peak-hour savings? *(Yes – for our dataset)*
2. What actions and characteristics are related to savings? *(Suggestive of AC – best guess: changing thermostat set point)*
3. **What is the short-term persistence of savings?** *(Savings within one-two weeks after first report mailed, stabilize after second report)*

papers by Nolan et al. (2008) and Schultz et al. (2007). There is also a literature that studies Opower programs over shorter time horizons, including Allcott (2011), Ayres, Raseman, and Shih (2012), Costa and Kahn (2013), and a number of industry reports such as Ashby *et al.* (2012),

Integral Analytics (2012), KEMA (2012), Opinion Dynamics (2012), Perry and Woehleke (2013), and Violette, Provencher, and Klos (2009). Allcott and Mullainathan (2012) show that the average treatment effects in the first one to two years across the first 14 Opower sites range from 1.4 to 2.8 percent of electricity use. Relative to this literature, our contributions are clear. First, we document consumers' "action and backsliding" using high-frequency data.[4] Second, we study Opower's three longest-running programs over a relatively long time horizon. Third, we exploit the continued vs. discontinued treatment groups to measure both habituation and post-intervention persistence. Fourth, we bring together the high-frequency and long-run analyses to analyze how persistence and habituation affect cost effectiveness and optimal program design.

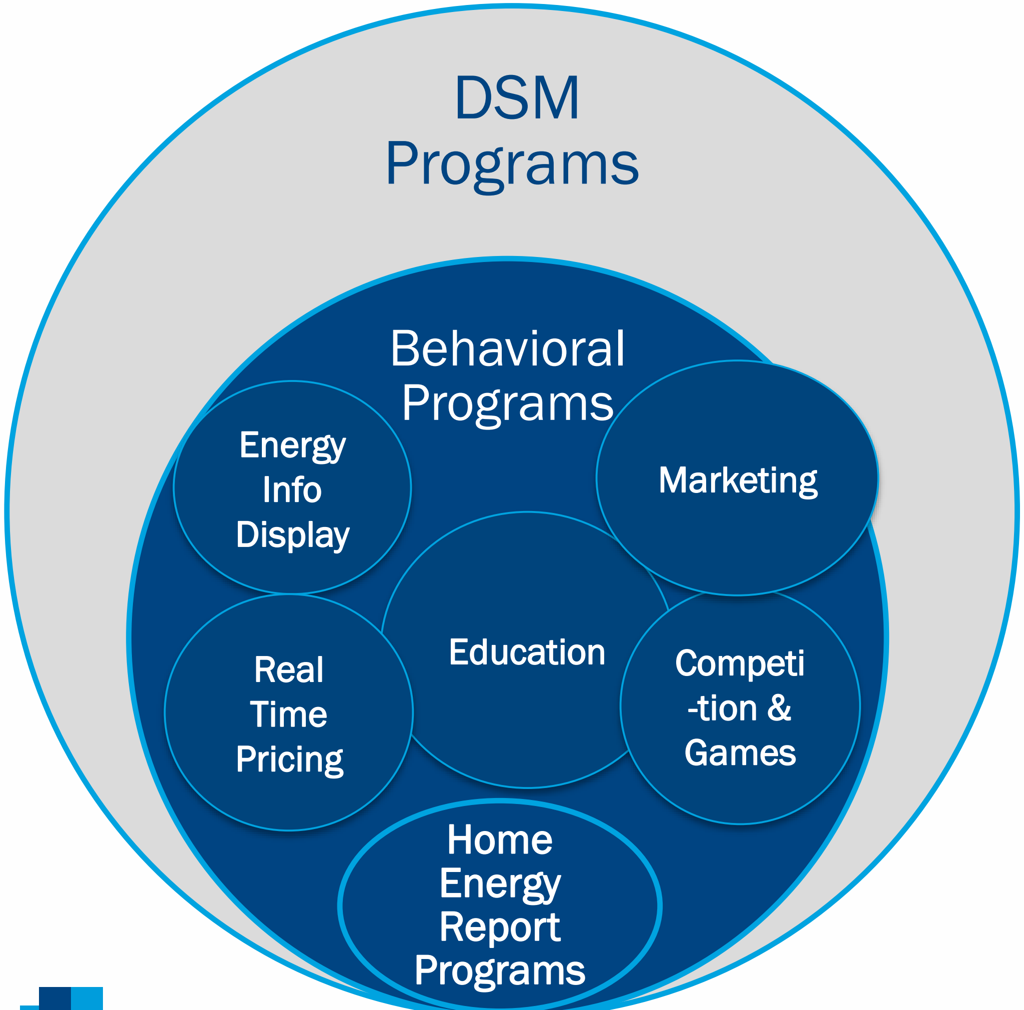
# EFFECTIVE USEFUL LIFE AND PERSISTENCE FOR BEHAVIORAL PROGRAMS

## Why it matters and what we know

April 3, 2014

Opinion Dynamics

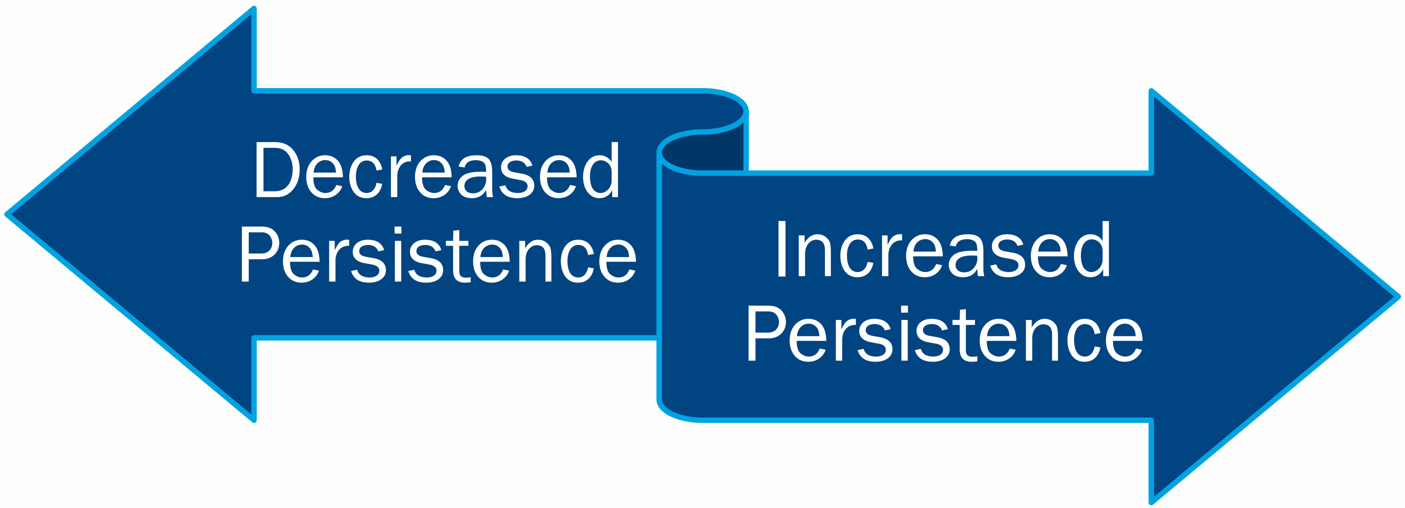
# How can we assess the cost-effectiveness of behavioral programs?




- Our cost-effectiveness framework isn't immediately transferable to behavioral programs –
  - Uncertainly about program actions & savings
  - Ongoing design (multi-year)
  - Unknown persistence
- Existing research is mainly for opt-out home energy reports, but discussion is generalizable to behavioral programs

Opinion Dynamics

# We suspect that behavioral programs may have an impact beyond the intervention period, but we don't know what it looks like



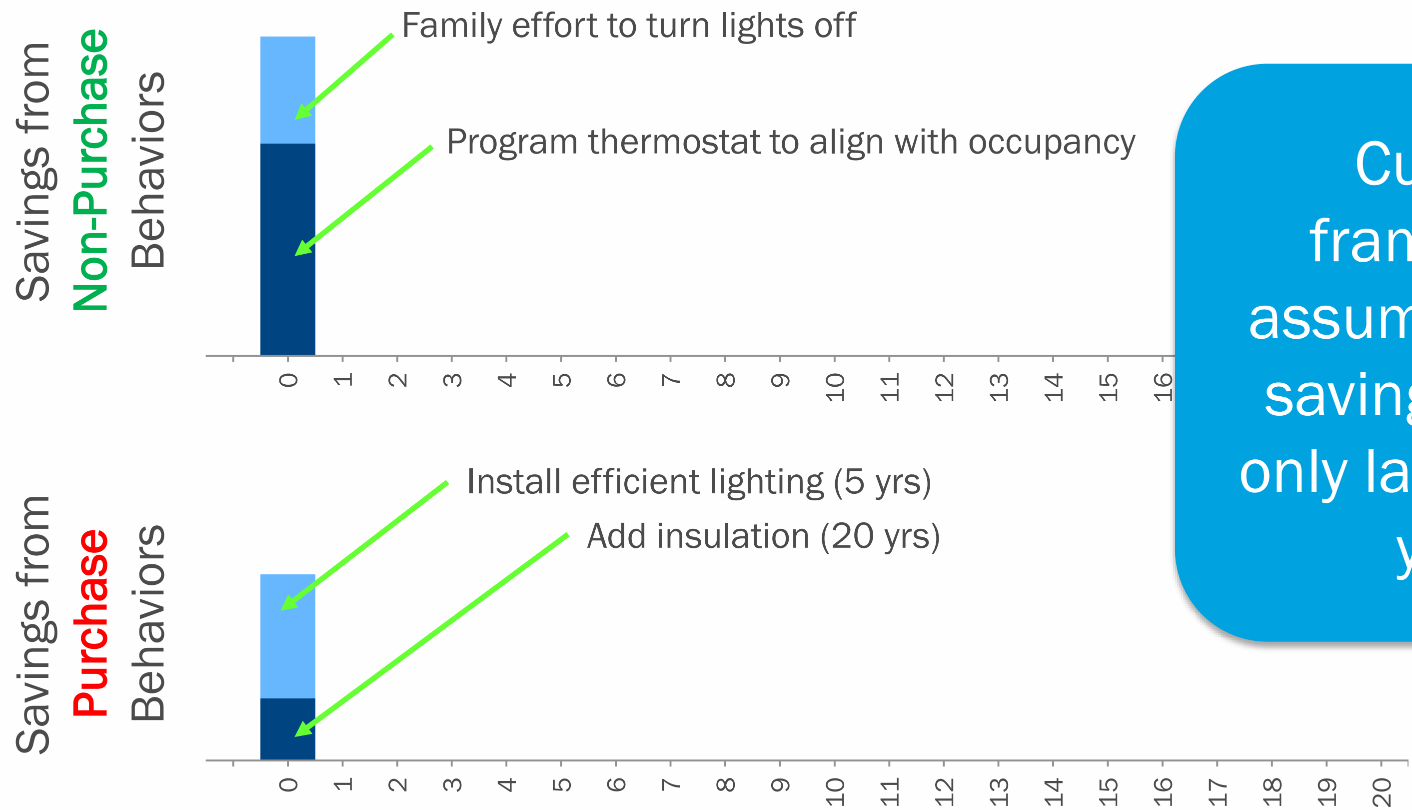

- Non-habituated behaviors

- Measure installations (esp. if measure life is longer)
- Habituated behaviors

# Defining Effective Useful Life and Persistence: Household Example

Illustrative Scenario

*Intervention: <u>One year</u> of home energy feedback and recommendations*

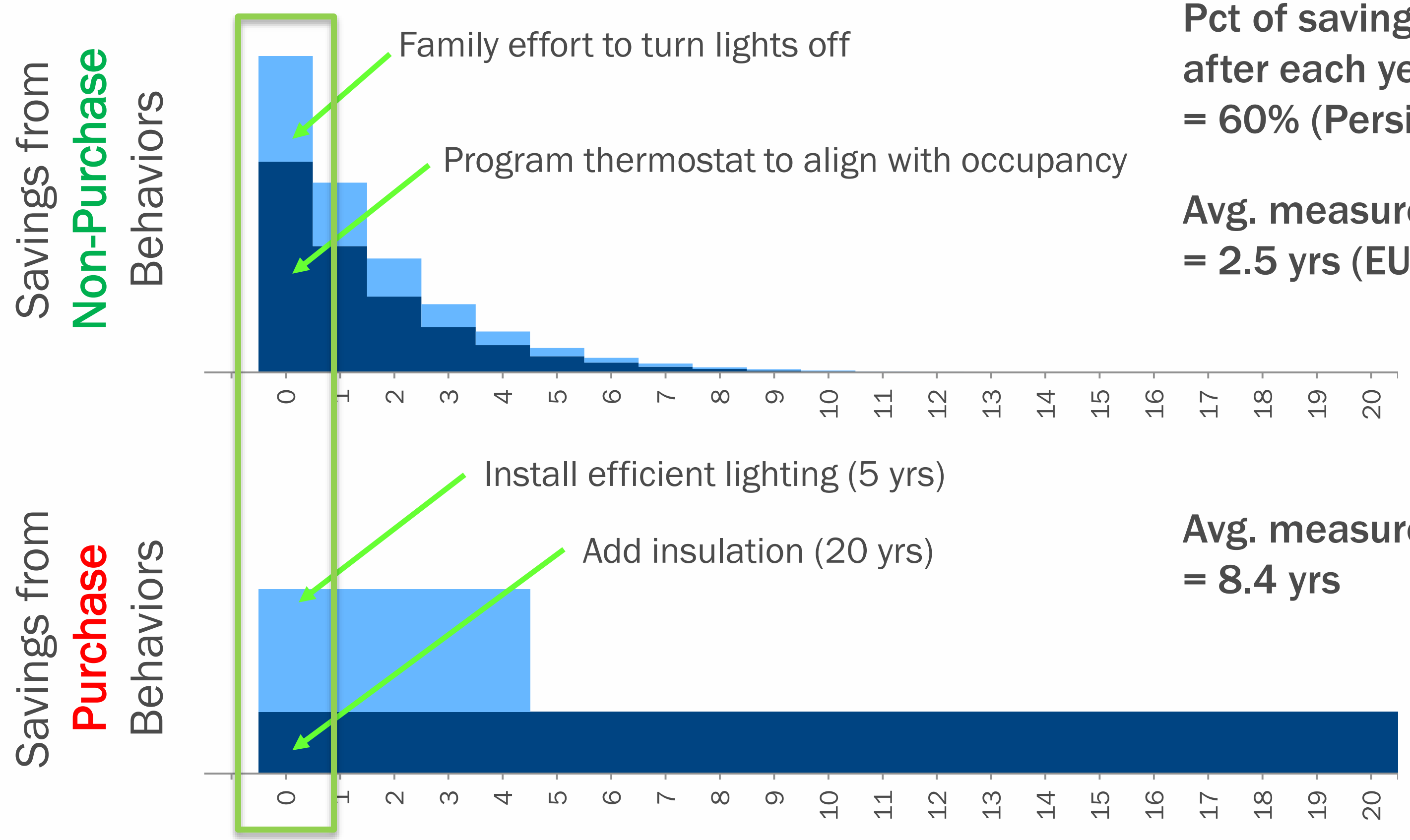



Currect framework assumes these savings would only last for one year

# Defining Effective Useful Life and Persistence: Household Example

Illustrative Scenario

*Intervention: One year of home energy feedback and recommendations*




**Pct of savings remaining after each year = 60% (Persistence)**

**Avg. measure life of behaviors = 2.5 yrs (EUL)**

**Avg. measure life of purchases = 8.4 yrs**

Opinion Dynamics

# What persistence research can tell us

Program Intervention Period

Post-program period

**Ongoing Persistence (a.k.a. "durability"):**
During program

- Evidence of actions that support long-run persistence
- Optimize program design

**Long-run persistence:**
After program stops

- How much savings will decline if we stop treatment
- How to assign an EUL
- Optimize program design

# What do we know about persistence?

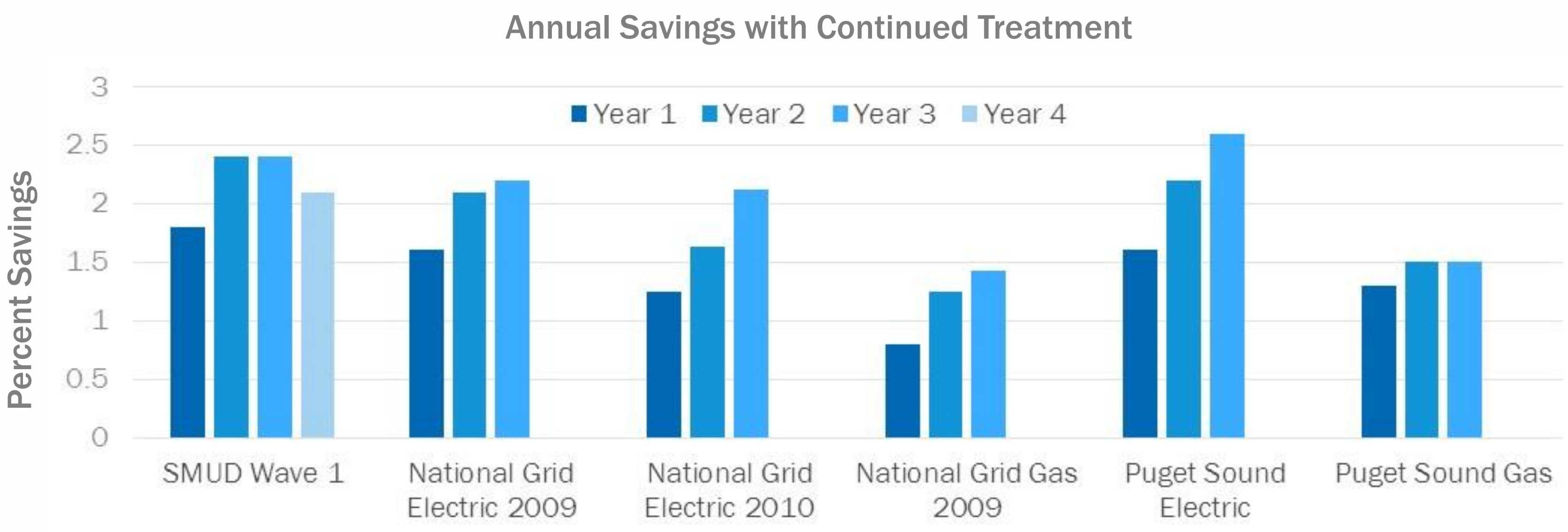
Opinion Dynamics

# Growing evidence to support ongoing and long-run persistence

1. Growing annual savings over multi-year programs
2. Continued but decaying savings after treatment stops
3. More stable savings between treatment events as intervention continues
4. “Actions” research showing mix of measures/behaviors

# 1. Savings tend to grow after year 1 in multi-year programs

- Participants "slow to habituate" [1] → may benefit from continued treatment
- …but what is the alternative?

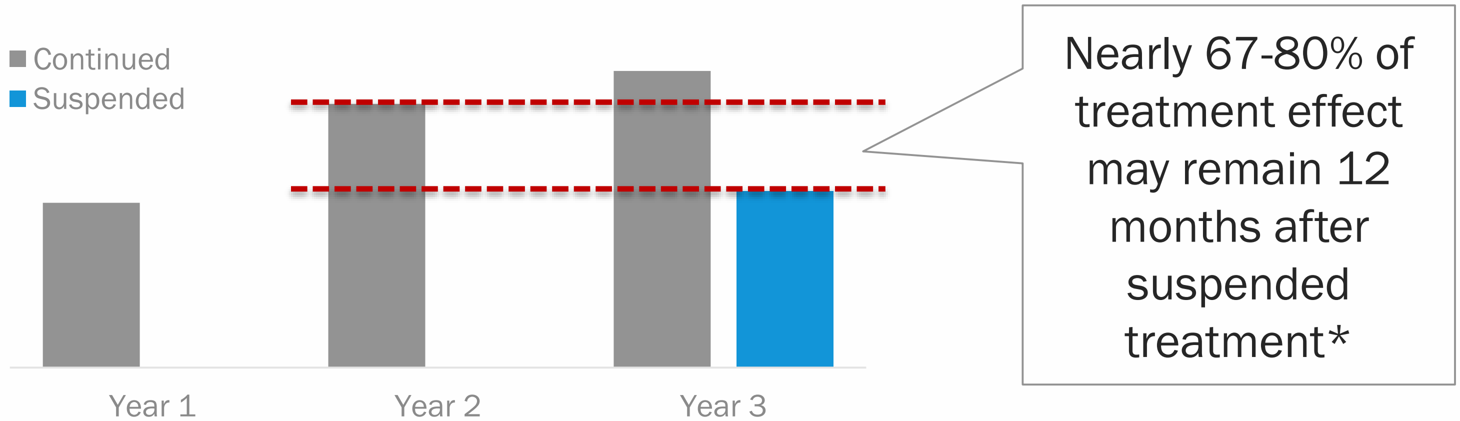



Opinion **Dynamics**

[1] Allcott & Rogers, 2013

# 2. Savings persist with decay after treatment stops

- A few long-term programs suspended treatment for some customers, and found **persistence with decay** after reports are suspended at year 2




- With our current assumption of zero persistence, **we're leaving a lot on the table**
- But decay rates differ by program and are likely **not predictable** ex ante

Opinion **Dynamics**

**See references at end of presentation for specific numbers*

# 3. As intervention continues, savings become more stable between periods

**Early reports:** Each report followed by noticeable reduction in consumption, that later "backslides"[1]

Sign of behavioral modification or change

**Later reports:** Short-run effect of each report not as noticeable; treatment effect more durable between cycles[1]

Evidence of habituated behaviors and/or installations

Knowledge of **how actions change** over time can help us optimize program design – e.g., when can we change frequency?

Opinion **Dynamics**

[1] Source: Allcott & Rogers, 2013

# 4. “Actions” research showing mix of measures and non-purchase behaviors

- Self-report survey research
  - Generally, a mix of behaviors and measures
  - **Different** results from **different** treatment groups and program designs


- Disaggregate **actions & end-uses** through smart meter / AMI data – Annika Todd’s presentation

By learning **what actions people are taking** we could improve **upfront** estimates of persistence

# What are the implications of this research for persistence, and what else do we need to know to move forward?

Opinion **Dynamics**

# In the current framework, programs may not be getting credit for the changes they are achieving

- **Our current assumption** **(zero persistence)** **is too conservative** in light of evidence that savings persist (with decay)

- Incorporating savings over time could change cost-effectiveness results, and potentially make **other designs** and/or **target audiences** more viable
  - Energy information and display – typically have higher costs
  - Targeting lower-usage customers (many current programs targeted to higher-usage)

Opinion **Dynamics**

## How can we incorporate this information going forward?

1. Need to develop appropriate framework for estimating cost-effectiveness.
   - A **one-year** program could have a first-year savings value, and a decaying savings value in subsequent years
   - A **three-year** program could have three years of growing savings, and decaying savings in subsequent years (decaying from year 3 savings)

   When making planning or CE projections, consider whether it will be a one-year or multi-year engagement

2. Need to develop a method for estimating a fair persistence value for each program

   May take more analysis to adjust and/or develop persistence estimates applicable to different programs

Opinion **Dynamics**

# What else do we need to know to pick a fair value?

- Conduct your own “interrupted” treatment? Extrapolate from current studies? Extrapolate from “early indicators” in your own program?
- Be careful “borrowing” assumptions from other jurisdictions
  - Just as first-year savings varies between programs, persistence varies too – Audience, length and intensity all matter

**Savings Magnitude**

- Target audience
- Behavioral “asks”
- Intervention strategy
- Frequency
- Duration

**Savings Persistence**

Consider similar factors

*Savings and persistence vary, and are inherently difficult to predict*

# What additional research can be done?

| Research Area | What it looks like | What it may tell us |
|---|---|---|
| "Stop treatment" experiments | • Stop for randomly-selected % of treatment pop. after 1, 2 or 3 years | • Long-run persistence under different program designs & different audiences |
| Assess actions taken | • Self-report research<br>• AMI analysis / disaggregation | • End-uses driving savings<br>• Potential for measure savings<br>• Bottom-up estimate (or adjustment) to persistence |
| Frequency & duration experiments | • Experiment with treatment frequency after x months (e.g., reduce to 2x / yr after 1 year)<br>• Observe difference in daily or monthly savings | • Help estimate decay rate<br>• Indication that habituated behaviors and/or measures have accumulated to a point where persistence effects could kick in |

# Persistence References


- Allcott, Hunt, and Todd Rogers (2012). "How Long Do Treatment Effects Last? The Persistence and Durability of a Descriptive Norms Intervention in Energy Conservation." Working Paper, Harvard University.
- Allcott, Hunt, and Todd Rogers (2013). “The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation.” National Bureau of Economic Research.
- Agnew, K., M. Rosenberg, B. Tannenbaum, B. Wilhelm (2012). Home Energy Reports: Power from the People. ACEEE Summer Study.
- Integral Analytics (2012). Impact & Persistence Evaluation Report: Sacramento Municipal Utility District Home Energy Report Program.
- KEMA (2012). "Puget Sound Energy’s Home Energy Reports Program: Three Year Impact, Behavioral and Process Evaluation." Madison, Wisconsin: DNV KEMA Energy and Sustainability.
- Opinion Dynamics (2012). "Massachusetts Three Year Cross-Cutting Behavioral Program Integrated Report." Waltham, MA: Opinion Dynamics Corporation.

# Please let us know what your persistence research is showing!


Amanda Dwelley

Associate Director

617-301-4629

adwelley@opiniondynamics.com


Visit us at www.opiniondynamics.com

Opinion **Dynamics**

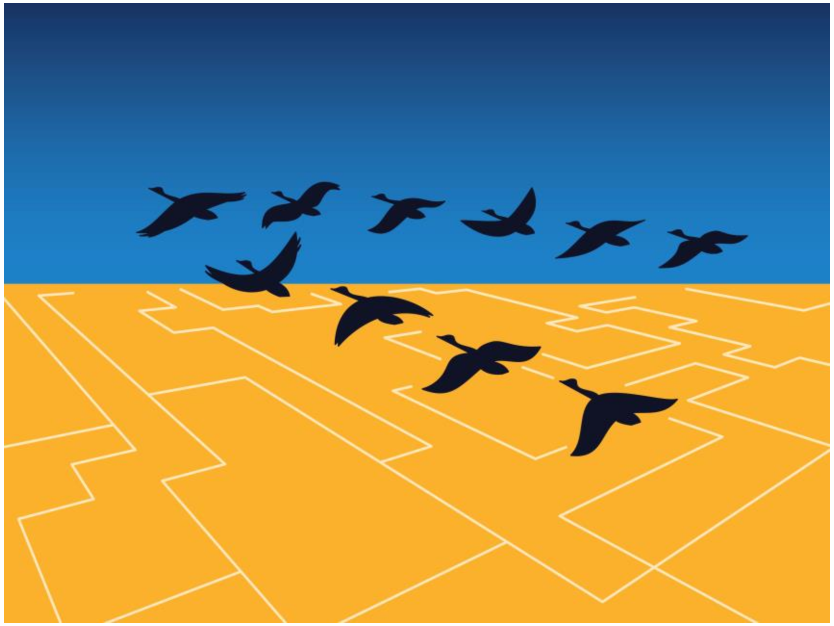
# FRAMING THE CHALLENGES ASSOCIATED WITH DETERMINING EFFECTIVENESS OF BEHAVIORAL PROGRAMS

April 3, 2014

Opinion Dynamics

# Guide to Presentation

- Review core challenges for program administrators regarding behavioral programs, their savings and effectiveness
  - Challenge #1: Defining Residential Behavioral Programs
  - Challenge #2:  Understanding What Actions Participants are Taking
- Set framework for presenters
  - Determining cost-effectiveness through measuring duration of savings and estimated useful life
  - Discuss method for estimating savings and short-term and long-term persistence

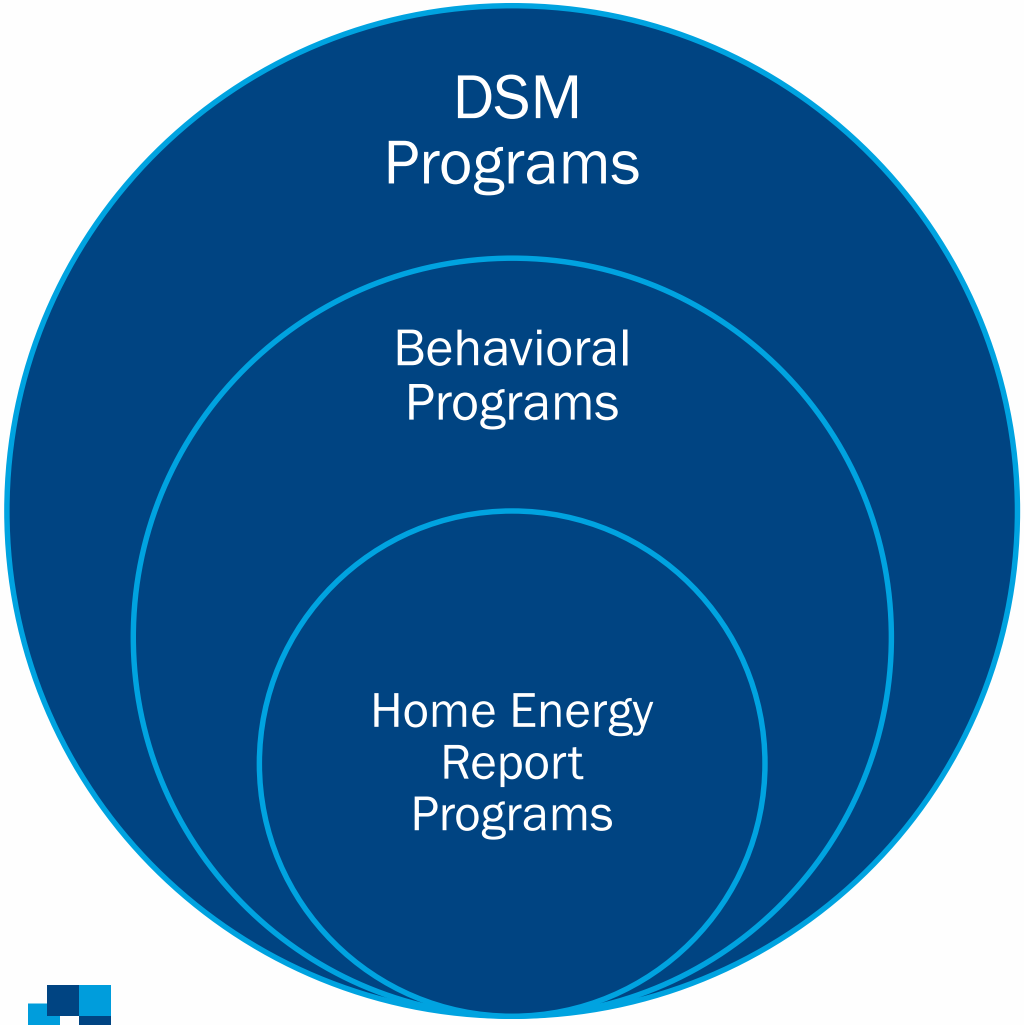


Opinion **Dynamics**

# Challenge #1: Defining Residential Behavioral Programs

Opinion **Dynamics**

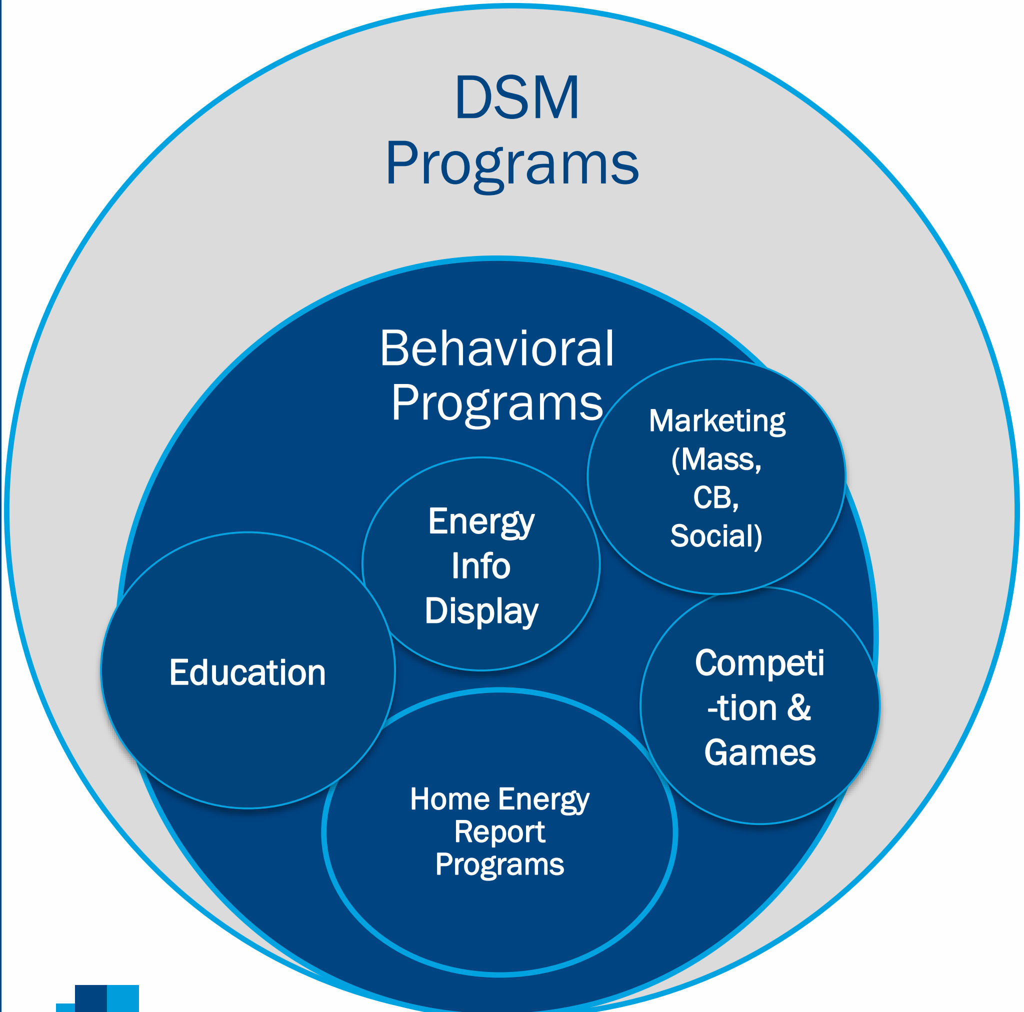
# There are varying definitions of residential behavioral programs, no agreed upon definition of behavior programs exists




- “All programs are behavior programs” (California Behavioral Whitepaper, "Paving the Way for a Richer Mix of Residential Behavior Programs")
- “Any type of energy efficiency program involves intervention to influence participant behavior. Even a standard rebate program is directed at influencing customer purchase behavior.” (ACEEE #108)

Opinion **Dynamics**

# Within “behavioral” programs, there are many intervention types and program designs




- Behavioral programs typically “encompass information, persuasion, and other non-price interventions” (Abrahamse et al 2005)
- Strategies include:
  - Feedback
  - Norms
  - Instruction
  - Commitment
  - Framing
  - Rewards / Gifts
  - Others

Opinion **Dynamics**

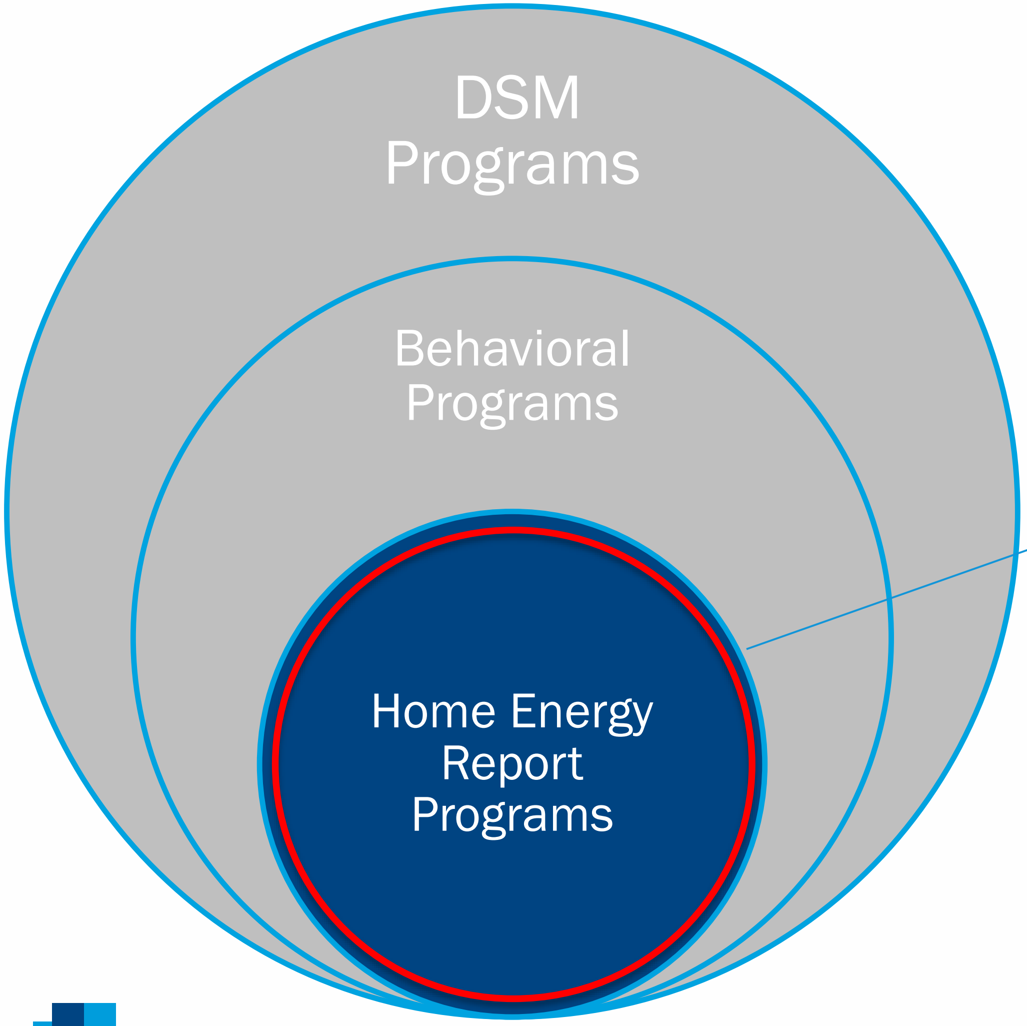
# These varied behavioral programs produce diverse ranges of savings, with each having more or less reliability depending on the estimation approach

| Program Type | Example Programs | Electric Savings Per Participant |
|---|---|---|
| **Home Energy Reports** | **Ameren Behavioral Modification Program, PG&E My Energy Program** | **1.4% to 2.8% annual kWh reduction per household** |
| Energy Information Display /HAN/IHD | Accelerated Innovations MYMETER™, CLC and SCE&G In-Home Display Pilots | 2.3% to 9.3% annual kWh reduction per household |
| Education & Training | Ohio Energy Project, LivingWise, DOE School Energy Program | 2.5% - 4.4% annual kWh reduction per household; 300–515 kWh per participant |
| Competition & Games | Team Power Smart (BC Hydro), Efficiency Vermont, CT energy Challenge | 1.9% annual kWh reduction per household; ~200 kWh per participant |
| Marketing (Community Based, Social Media, Mass Marketing) | MassSave®, Energy Upgrade California, Project Porchlight (One Change) | Not typically estimated |

Most evaluations are for Home Energy Report programs. These programs are typically RCT's that provide reliable savings estimates. Not all programs employ this research design.

Opinion **Dynamics**

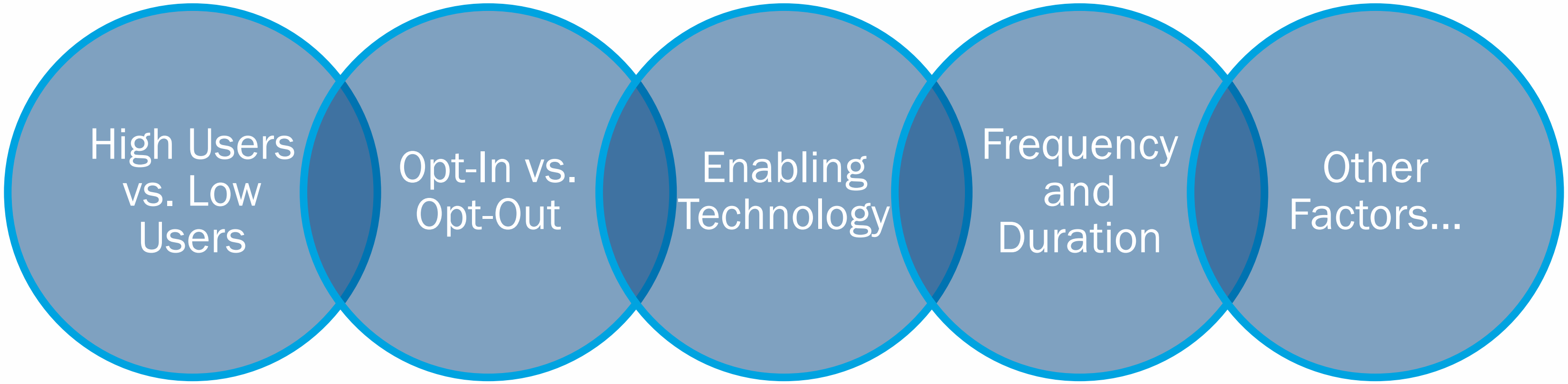
## This webinar focuses on Home Energy Report, as these programs have the most research available relevant to the discussion of energy savings over time




- Typical behavioral program is Home Energy Reports
- Offers paper based reports
- Provides normative comparisons, information, and usage history
- Delivered as a randomized controlled experiment

Opinion **Dynamics**

# Within Home Energy Report programs, variations in program delivery affect program savings




- For example:
  - Savings magnitude and persistence varies based on t**arget population** and **program model** (i.e., opt-in vs. opt-out)
  - **Frequency** and **duration** of behavior interventions has an impact on persistence of behavior (i.e., number of reports sent, enabling technologies, etc.)

# What does this mean for program administrators?

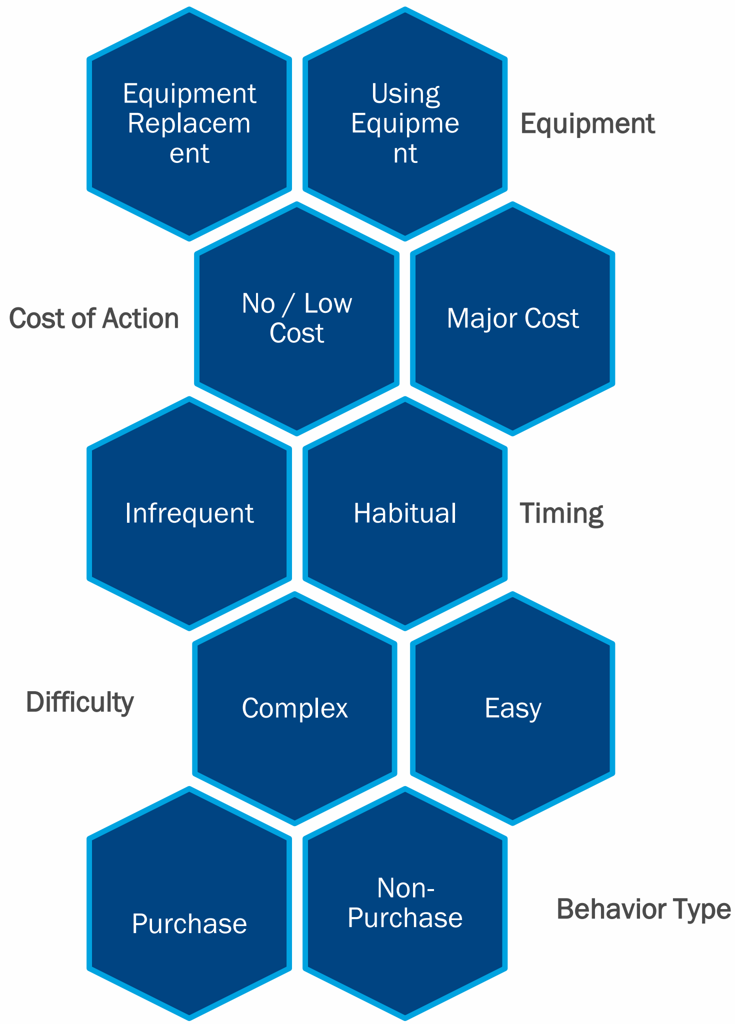
- There are many programs that could be considered “behavioral” and there is no standard definition for these programs
- Within “non price-driven” programs, there is a wide diversity of program designs, research designs, and magnitude of evaluation research
- *When considering offering behavioral programs, program administrators should consider:*
  - *Large ecosystem of program models that influence customer behaviors*
  - *Programs most appropriate to their intended audience*
  - *Range of savings that can be achieved (and best approach to measuring savings)*

# Challenge #2: Understanding what Actions Participants are Taking

Opinion Dynamics

# Savings can depend on what participants are “asked” to do

- Many frameworks exist to categorize the **types of actions** participants can take
  - Non-habituated behaviors
  - Habituated behaviors
  - Measure installations
- Most programs layer **multiple asks**
- Survey research indicates a range of self-reported activities
  - Varies substantially by program and region




Opinion **Dynamics**

# Uncertainty exists around the source of energy savings

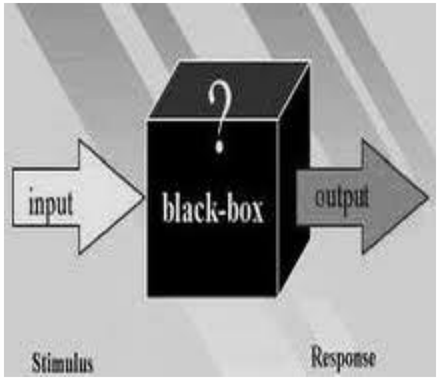
- It is unclear what specific actions are driving savings, and therefore how long those savings might persist
- For non-purchase behaviors, we do not know the breakdown between habituated and non-habituated behaviors
  - **Habituated Behaviors:** Once a behavior is internalized, it will persist without continued prompting from outside sources
  - **Non-Habituated Behaviors:** These can decay over time. We have no empirical evidence about the length of time they persist



Opinion **Dynamics**

# Savings estimation approaches typically do not characterize actions taken

- Most HER programs are evaluated through statistical analysis of billing records compared to control or comparison group (experimental or quasi-experimental design)
  - These evaluations produce one year aggregate annual savings results
  - This approach does not incorporate actions taken (i.e., equipment replacement/ equipment usage) into calculation, unless customers have participated in other DSM programs (i.e., removing double-counted actions)
- In some cases, self-report surveys are conducted but generally do not produce energy savings estimates




Opinion **Dynamics**

# Uncertainty about source of savings impacts how program administrators understand cost-effectiveness

- Program administrators (PAs) have been offering behavioral programs for a relatively short time and their fate as an effective program intervention depends on their associated costs and benefits
- As part of cost effectiveness calculations, PAs look to:
  - Effective Useful Life
  - Program costs
- However, EUL and incremental costs are less clear when considering actions taken by participants due to behavioral programs
- Current frameworks use a conservative persistence estimate (i.e., no savings after the first year)

Opinion **Dynamics**

# Measuring cost effectiveness is relatively straightforward for rebate programs, but less so for behavioral programs

| | *Known action and well-researched longevity*<br>**Rebate Programs** | *Unknown action and uncertain longevity*<br>**Behavioral Programs** |
| --- | --- | --- |
| Costs | ▪ Known incremental costs for purchases ✔ | ▪ Unknown incremental costs for purchases ✘ |
| Benefits: Savings | ▪ Known installation with known engineering factors ✔ | ▪ Known first year aggregate savings ✔ |
| Benefits: Measure Life | ▪ Average replacement cycle, known through market research ✔ | ▪ Unknown "habituation" / persistence within program period ✘<br>▪ How long people can maintain similar level of savings, or<br>▪ Rate at which first-year savings may decay |

Opinion **Dynamics**

# What does this mean for program administrators?

- There is uncertainty around the source of energy savings (i.e., actions taken)
- Uncertainty has implications in terms of cost-effectiveness and potential future program designs
- *Program administrators should consider the following long-term effectiveness issues when delivering a behavioral program:*
  - *What are the incremental costs to customers, and how does this affect program costs?*
  - *What is the "Effective Useful Life" or persistence of savings, and how does this impact program benefits?*
  - *What information is needed to optimize program delivery?*
- *Improving our understanding customer actions will help to inform these issues, and comprehensive CE tests could support optimized portfolio selection*

Opinion **Dynamics**

# Thank You!

Olivia Patterson

Project Manager

510-444-5050 ext. 191

opatterson@opiniondynamics.com

Visit us at www.opiniondynamics.com

Opinion **Dynamics**