BEHAVIORAL PROGRAM METHODS & OUTCOMES

PRESENTED TO:

STREAMSAVE

NOVEMBER 15, 2022

PRESENTED BY:

ADAM THOMAS
ADM ASSOCIATES, INC.

ABOUT ADM

- Founded in 1979
- Evaluation of energy efficiency, load management, and decarbonization programs
- Consulting services for:
  - Regulatory agencies
  - Utilities
  - Research foundations
- Evaluation studies for 200+ cohort-years of energy behavioral interventions
- Studies included the first behavioral energy pilot in the United States: Sacramento Municipal Utility District, 2009

ABOUT ME

- Principal Consultant
- Leading research design and execution for energy efficiency & load management interventions:
  - Led 50+ energy behavioral intervention impact studies
  - Head of a business unit evaluating impacts for ~$1.5 billion USD in utility and government-sponsored programs annually

EXPERIMENTAL DESIGN BEST PRACTICES

Program Design – Opt-out Randomized Control Trial

- Select customer group, randomly assign to treatment / control group
- Typical cohort:
  - 50,000 treatment
  - 20,000 control
- Conduct validity testing
- Send messaging
- Measure impacts

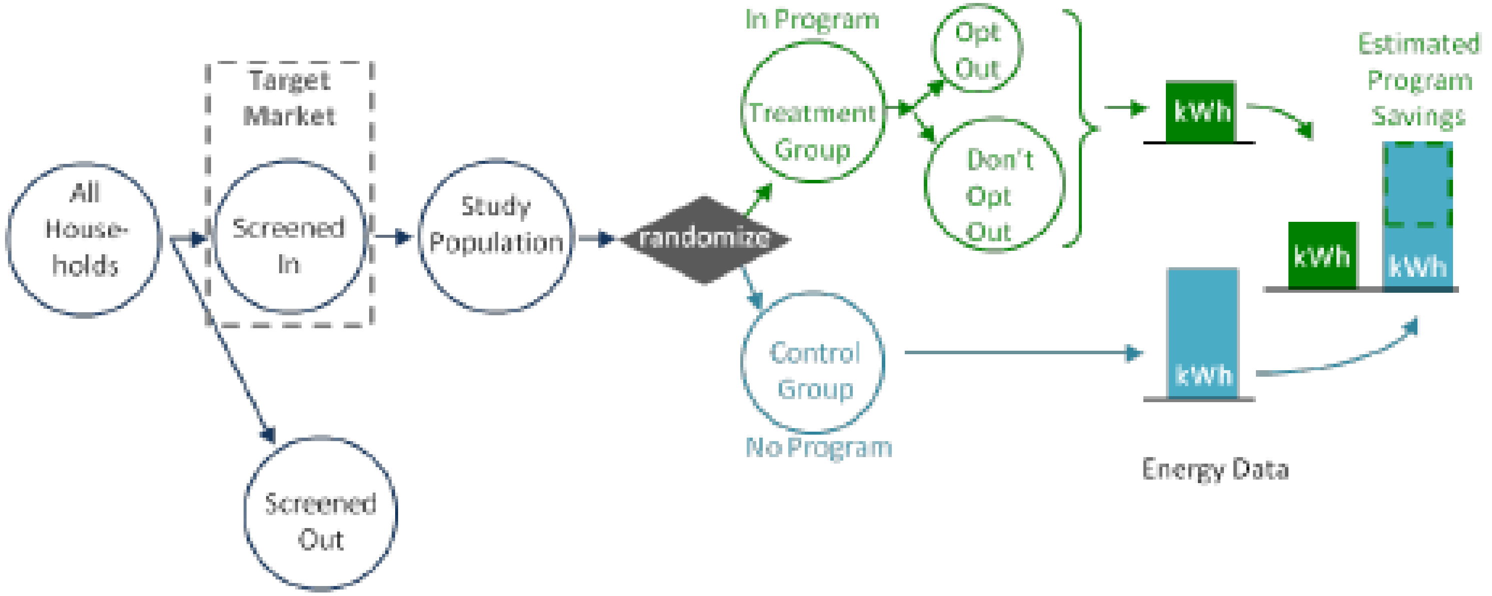


Figure 1. Illustration of RCT with opt-out program design

Source; the Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures. Chapter 17: Behavioral Protocol

Program Design – Opt-in Randomized Control Trial

- Recruit interested/willing customers
- Randomize into treatment and control, selecting a subset that receives no intervention after opting in
- Conduct validity testing
- Send messaging
- Measure impacts

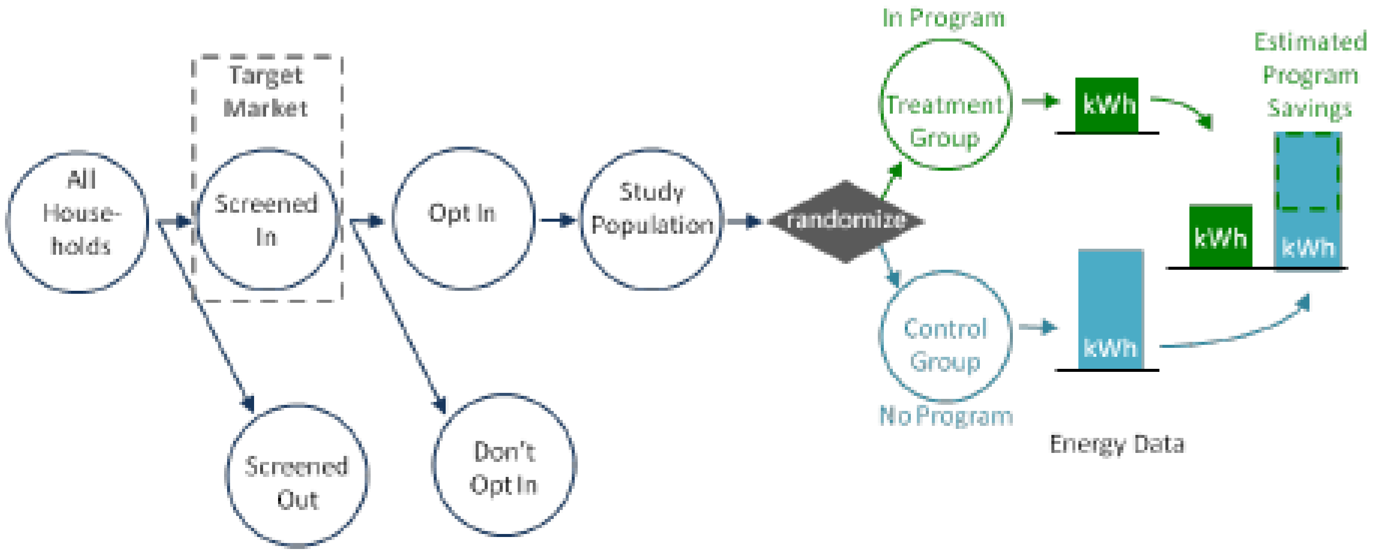


Figure 2. Illustration of RCT with opt-in program design

Source; the Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures. Chapter 17: Behavioral Protocol

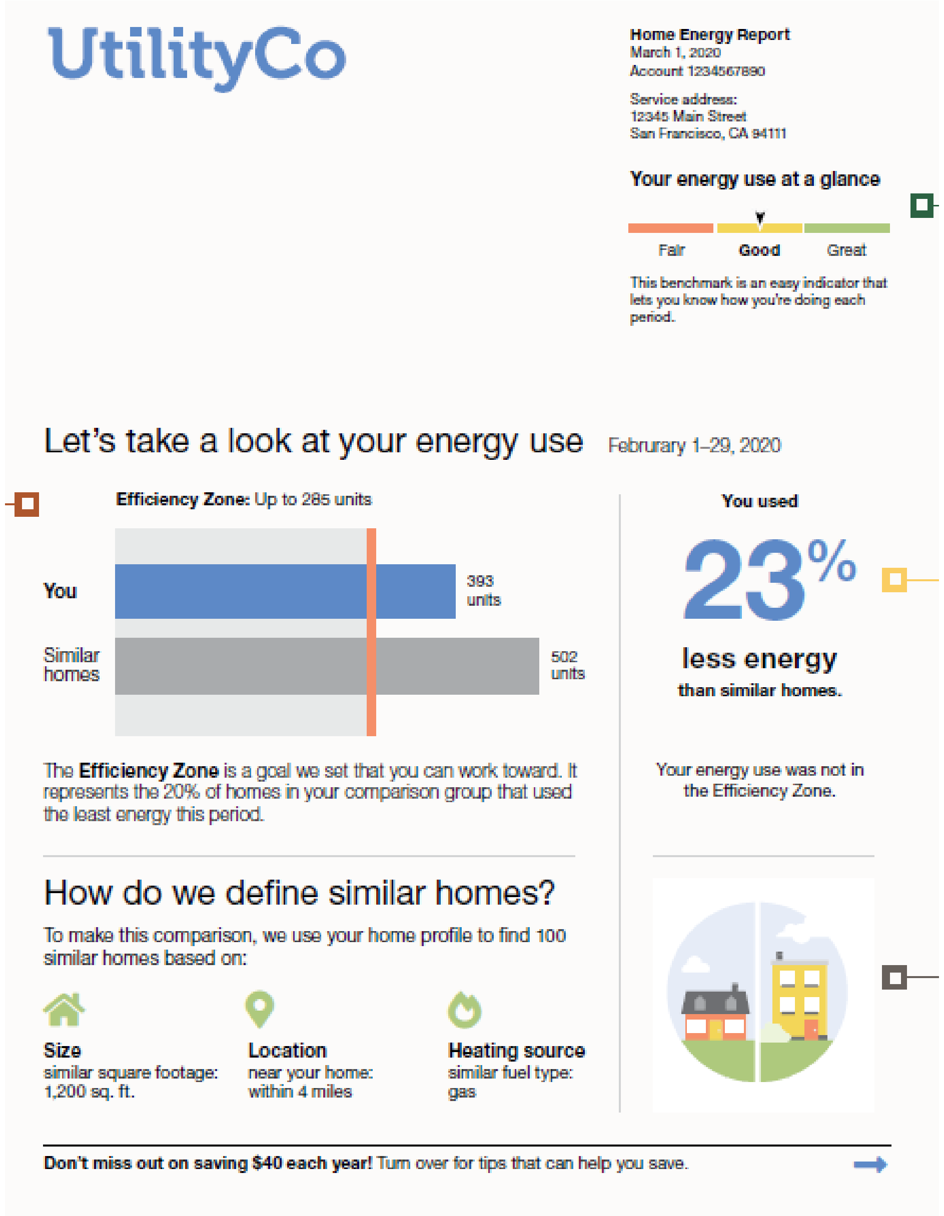
Program Design Comparisons

Opt-out RCT	Opt-in RCT	No RCT
• Internally & externally valid • Measurable results, robust to exogenous shocks • Lower satisfaction due to unwilling recipients • Highest volume of treatment households possible	• Internally valid • Externally invalid • Measurable results, robust to exogenous shocks • Higher satisfaction due to willing recipients • Lower volume of treatment households	• Internal validity is problematic • Quasi-experimental methods required for comparison group • Bias may be unmeasurable / unable to be mitigated

COMMON INTERVENTION TYPES

SOCIAL NORMING INTERVENTION

- Usage compared to neighbors
  - *"You use more energy than 86 of your 100 closest neighbors!"*
- Usage compared to self
  - *"You used 15% less energy than during the same time last year!"*
- Supplemented with online audit tools – users engage and add more home data
- Annual Impacts:
  - Electricity: 1% - 3% of annual
  - Natural Gas; .5% - 1.5% of annual

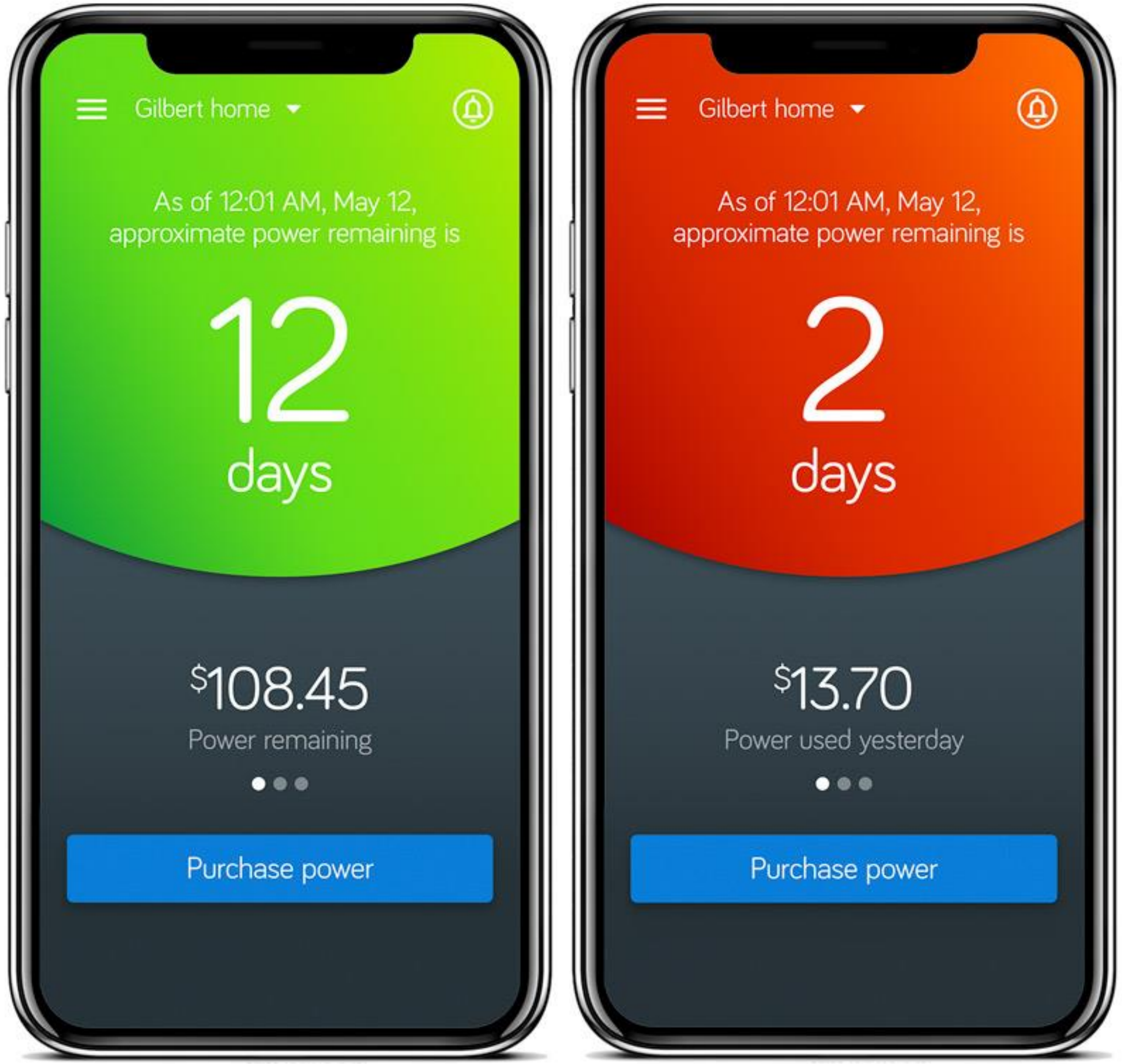


https://www.oracle.com/a/ocom/docs/industries/utilities/utilities-opower-home-energy-reports.pdf

BILLING PRE-PAY

- Customers “pre-buy” discounted energy, and face higher costs for exceeding “pre-bought” energy

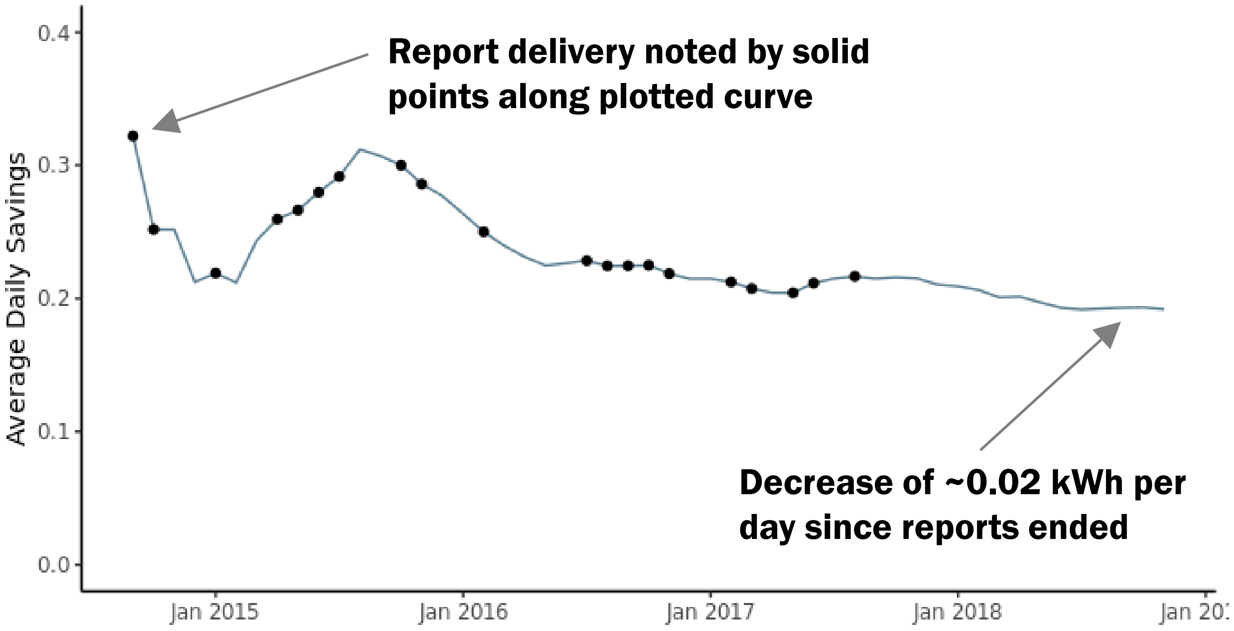
- Recommended purchase amount established to target ~5% energy reduction



https://www.srp.net

SAVINGS TRENDS

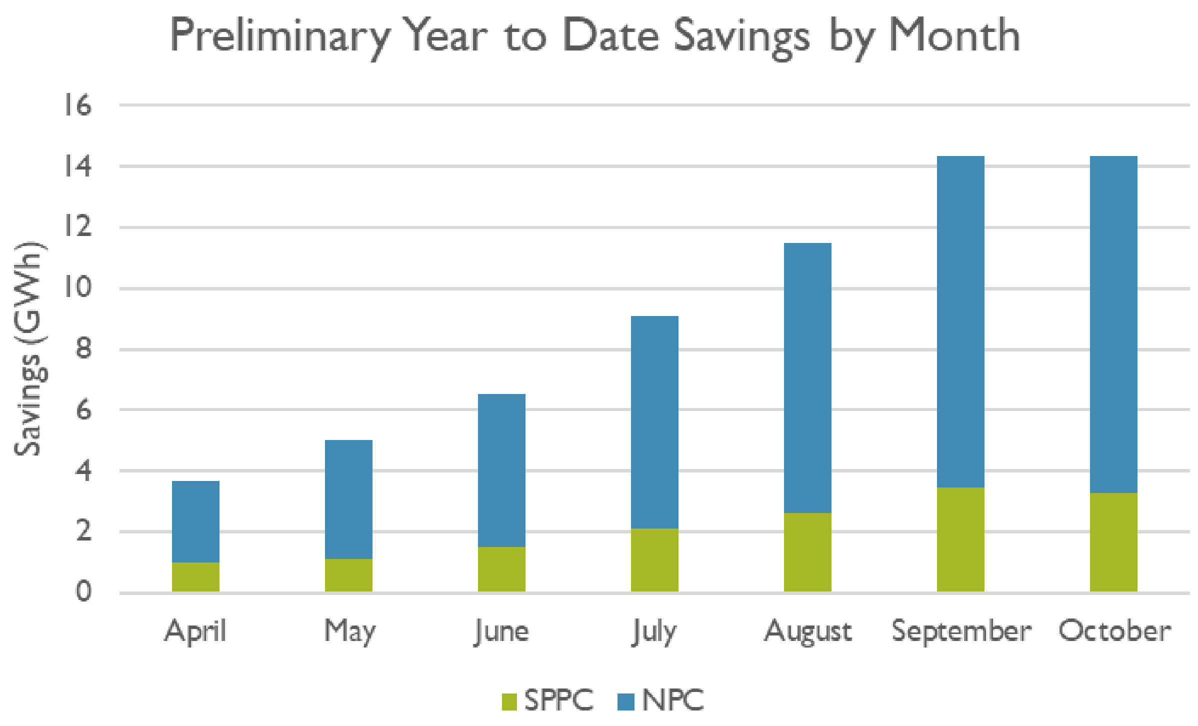
Example of Average Daily Savings Trends Through Time



- We have seen evidence of savings having an initial "spurt" when reports are first sent
- Typically, savings will increase and reach a peak around 11 to 15 months before eventually decaying through time

SAVINGS TRENDS

Example Preliminary Savings Trends During the Program Year



- Savings tends to increase throughout the program year as more reports are sent *and* as data from the summer season is incorporated into the regression.

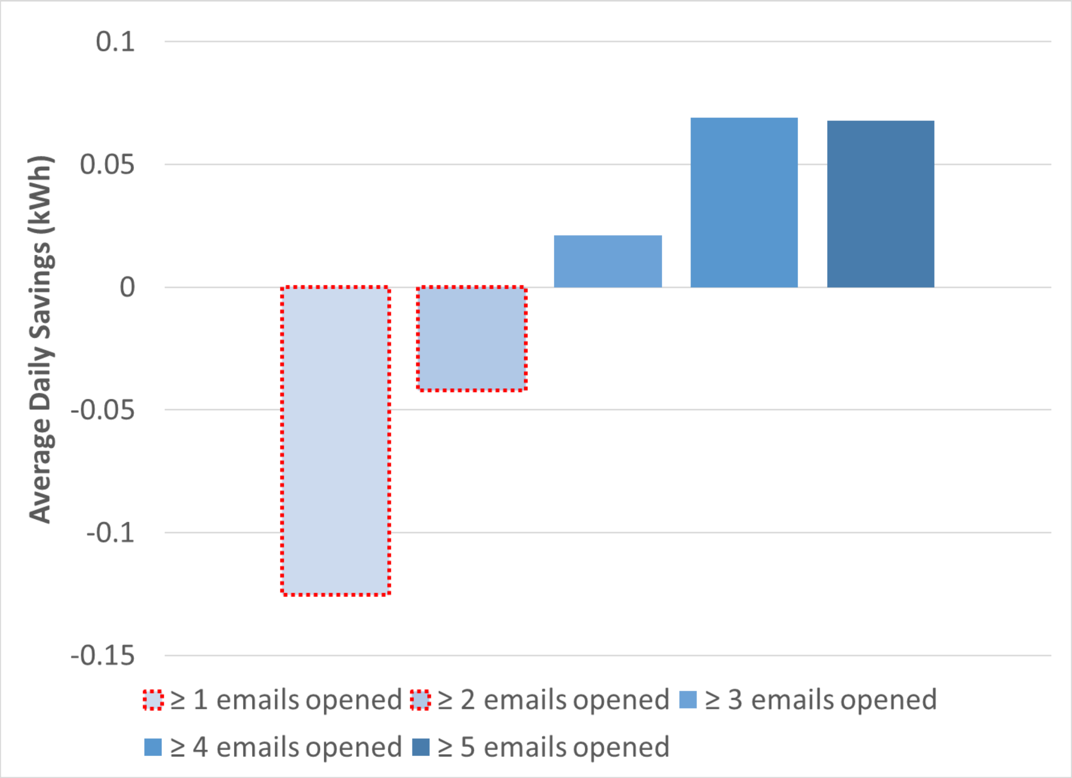
DELIVERY METHOD – PHYSICAL MAIL VS. EMAIL

- Trade-off of energy impacts vs. carbon-intensity

- Physical mail produces higher savings, countered by carbon-intensity of transportation
  - Open-rates for emails in continuous decline

- Optimal choice based on:
  - Existing preferences for billing method
  - Carbon-intensity of power supply vs. postal system
  - Priority of kWh vs. GHG

DELIVERY METHOD - EMAILED



- Customers that open 3 or fewer messages annually will not save energy
  - Results include zero/insignificant and statistically significant *increases* in usage
- Customers that open 4+ emailed reports show statistically significant energy savings

EVALUATION STEPS

EVALUATION STEPS

1. Data Prep	Cleaning, true-up & calendarization
2. Validity / Equivalency Testing	Testing control group for matching to treatment group in the pre-treatment period
3. Propensity Score Matching (PSM)	Develop control groups for cohorts with invalid control groups
4. Estimate Model Impacts	Calculating bi-monthly & annual impacts based on regression model output
5. Remove Double-Counted Savings	Account for participation in other energy efficiency programs
6. Estimate Program Attrition	Tallying opt-outs / drop-outs / move-outs

DATA PREPARATION:

TRUE-UP

$$Adjusted\ usage = \sum_{i}^{n} Billed\ usage \times \frac{Billing\ days_m}{\sum_i^n Billing\ days}$$

Where:

i = First estimated bill in a sequence of estimated bills leading to a metered bill.

n = A metered bill providing an adjustment factor for preceding estimated bills.

m = The billing month of interest.

$Billed\ usage$ = The total kWh billed in a monthly bill.

$Billing\ days$ = The total number of days in a monthly bill’s billing period.

DATA PREPARATION:

CALENDARIZATION

$$Monthly\ usage_m = \sum_{i}^{n} \left(Adjusted\ usage_i \times \frac{Month\ days_i}{Billing\ days_i} \right)$$

Where:

i = First bill containing the month of interest.

n = Last bill containing the month of interest.

m = The month of interest.

$Monthly\ usage$ = The calendarized monthly usage for a given month.

$Month\ days$ = The number of days belonging to the month of interest in a billing period.

$Billing\ days$ = The number of days in a billing period.

DATA CLEANING:

APPROPRIATE EXCLUSIONS

Erroneous Reads	• Single-day read periods
Outliers	• Standard outlier diagnostics – 3 standard deviations typical
Insufficient reads	• < 9 months pre-data • < 9 months post-data • Service disconnections

DATA CLEANING:

INAPPROPRIATE EXCLUSIONS

Negative Reads

- Don't exclude corrective entries
- Use them to correct prior month usage

Estimated reads

- Distribution isn't random.
- Estimated reads more likely for rural customers – exclusion damages internal & external validity

Opt-outs

- Treatment customers that opt out must be kept in analysis
- Their habits/attitudes are also reflected in the control group

EXAMPLE PROGRAM

EXAMPLE STUDY COHORT – UTILITY WITH 5 SEPARATE PROGRAM WAVES

WAVE	TREATMENT CUSTOMERS		CONTROL CUSTOMERS	
	ORIGINAL COHORT	EOY 2019	ORIGINAL COHORT	EOY 2019
Wave 1	16,851	10,239	16,762	9,704
Wave 2 **Wave 3**	34,246	6,020	14,427	5,688
Wave 4	57,662	15,543	23,044	14,471
New Movers	34,437	17,835	34,436	16,017

- Four "standard" waves
  - Customers with longer time in residence / lower volatility
- One piloted "New Movers" wave
  - Customers that move frequently. Low-income / renters, etc.

VALIDITY TESTING RESULTS

Failing Validity Testing

Pre-Period Month	Treatment Group Average Daily Usage (kWh/day)	Control Group Average Daily Usage (kWh/day)	P-value	Statistically Significant Difference
Apr 2016	30.53	27.43	<0.001	*
May 2016	27.50	24.72	<0.001	*
Jun 2016	25.20	22.71	<0.001	*
Jul 2016	24.06	21.75	<0.001	*
Aug 2016	23.84	21.62	<0.001	*
Sep 2016	25.14	22.85	<0.001	*
Oct 2016	29.46	26.86	<0.001	*
Nov 2016	36.31	32.96	<0.001	*
Dec 2016	39.68	36.11	<0.001	*
Jan 2017	40.12	36.50	<0.001	*
Feb 2017	40.93	37.52	<0.001	*
Mar 2017	41.30	37.80	<0.001	*

- Statistically significant differences if p-value < 0.05
- Waves can fail validity testing for multiple reasons:
  - Poor design / randomization
  - Degradation over years as customers drop out of the program

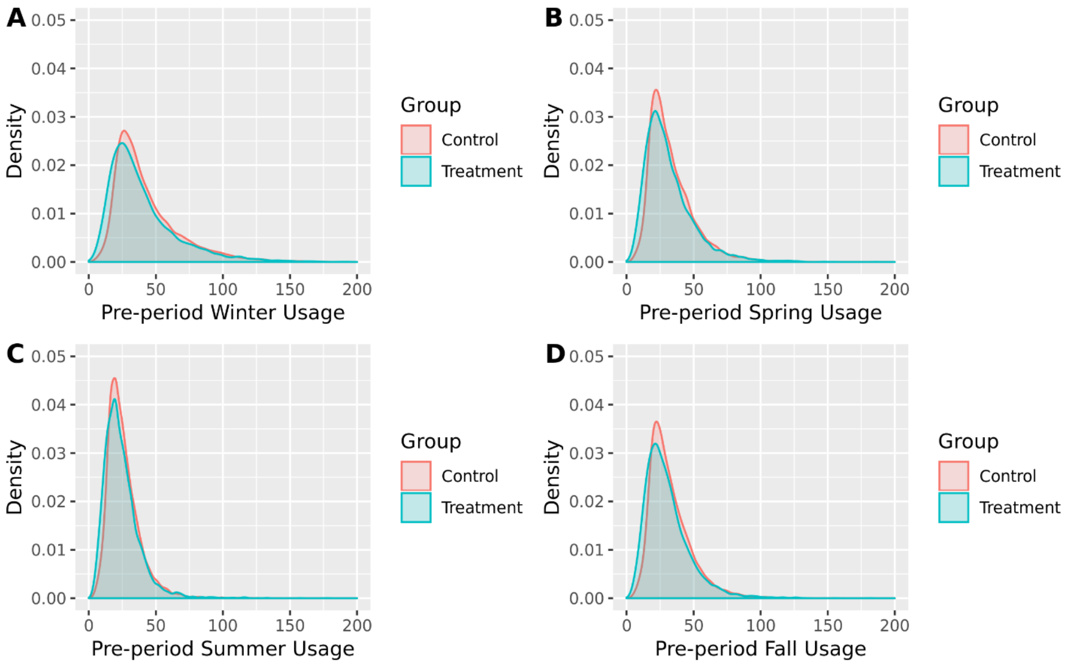
CORRECTING VALIDITY TESTING FAILURES:

PROPENSITY SCORE MATCHING (PSM)

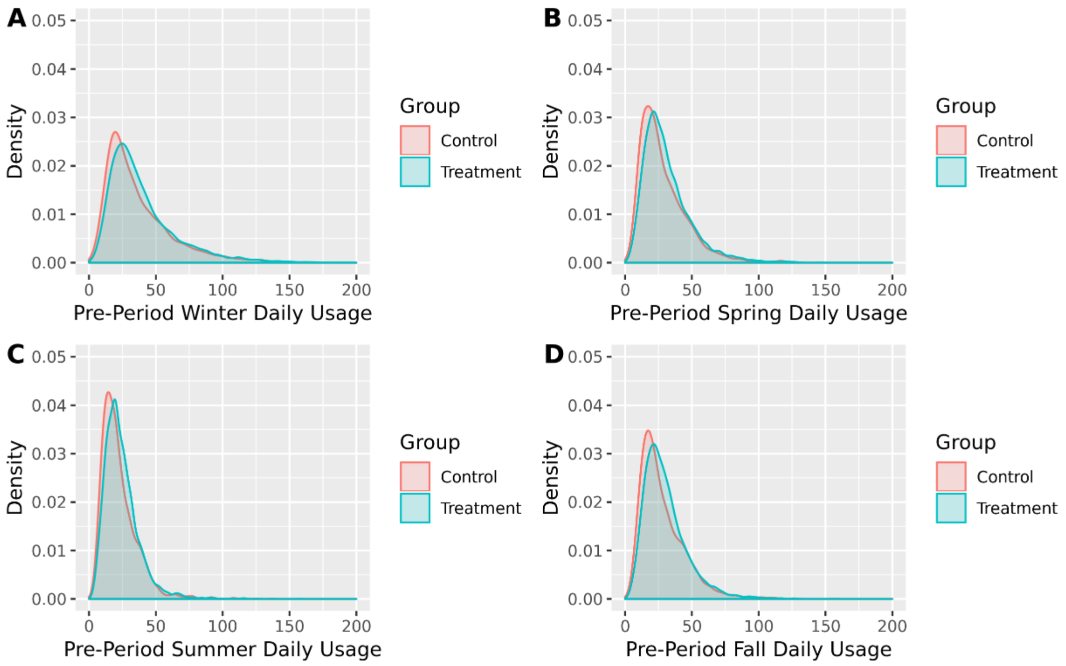
- PSM matches treatment customers to the most similar nonparticipant household
  - Based on customer billed consumption in baseline period
  - Verified with statistical difference testing
- Match households on known characteristics:
  1. Pre-period spring usage
  2. Pre-period summer usage
  3. Pre-period fall usage
  4. Pre-period winter usage
  5. Geography

PROPENSITY SCORE MATCHING RESULTS

Before Matching

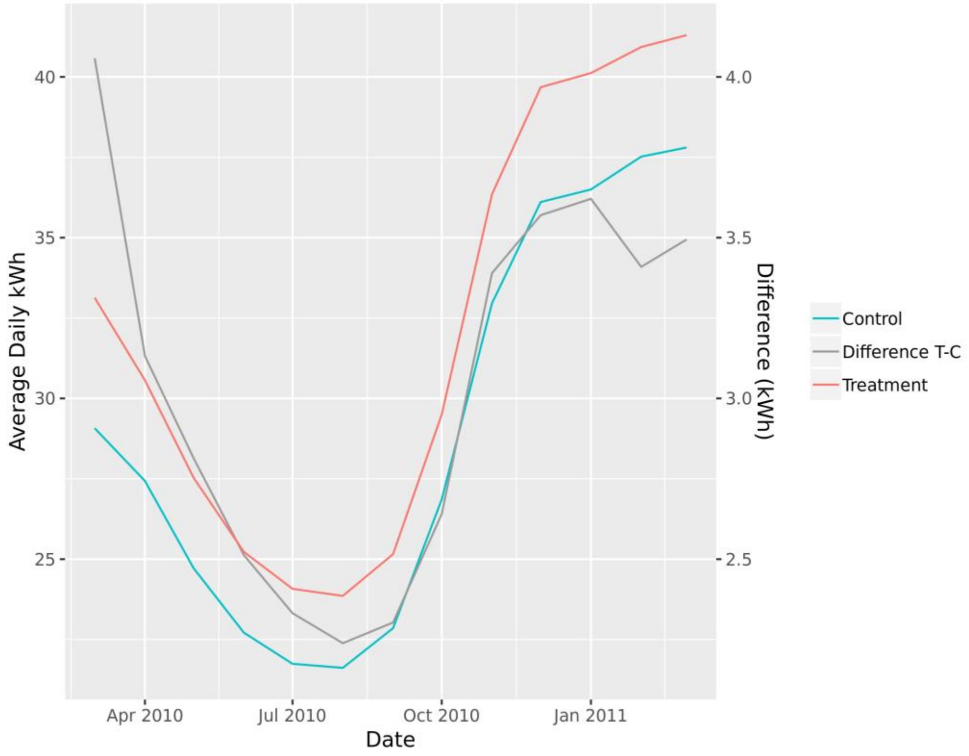


After Matching

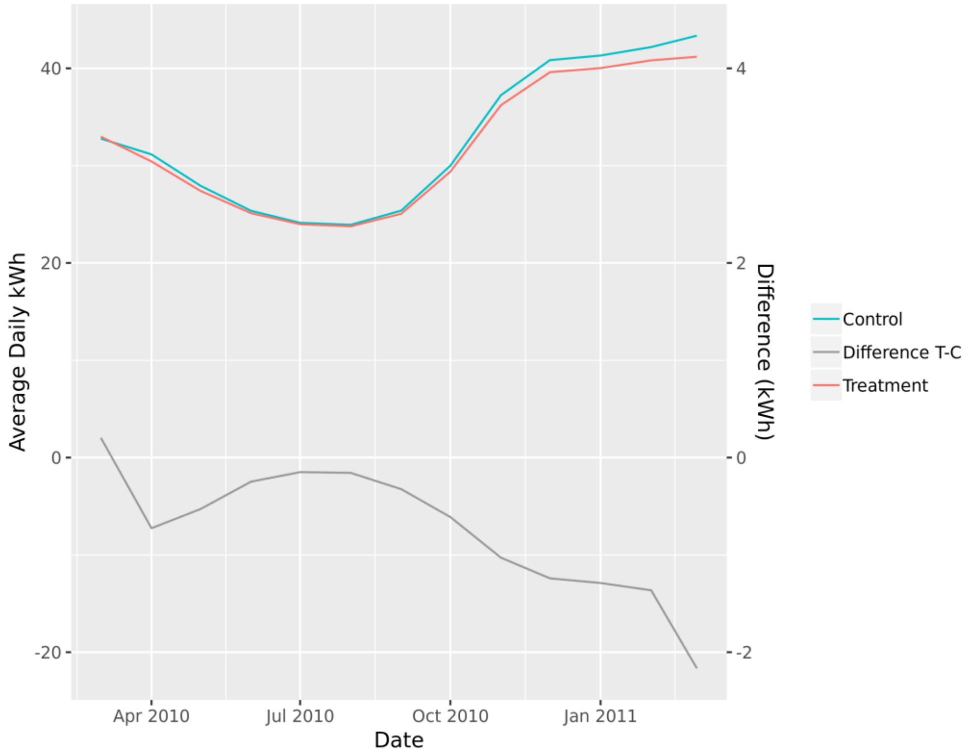


PROPENSITY SCORE MATCHING RESULTS

Before Matching



After Matching



PROPENSITY SCORE MATCHING RESULTS

Validity Testing after PSM

Pre-Period Month	Treatment Group Average Daily Usage (kWh/day)	Control Group Average Daily Usage (kWh/day)	P-value	Statistically Significant Difference
Apr 2016	30.43	31.15	0.1082	-
May 2016	27.40	27.93	0.1735	-
Jun 2016	25.11	25.36	0.4813	-
Jul 2016	23.98	24.13	0.6699	-
Aug 2016	23.76	23.92	0.6590	-
Sep 2016	25.06	25.38	0.3635	-
Oct 2016	29.38	29.99	0.1554	-
Nov 2016	36.21	37.24	0.0794	-
Dec 2016	39.60	40.84	0.0648	-
Jan 2017	40.03	41.32	0.0596	-
Feb 2017	40.82	42.18	0.0970	-
Mar 2017	41.19	43.36	0.0715	-

- Group passes monthly validity testing after propensity score matching of ad-hoc control group
- Statistically significant differences if p-value < 0.05
- Maintains measurability, but decreased randomization makes control group less robust to exogenous shocks

LINEAR REGRESSION MODELING

Regression Model Specification: Difference-in-Difference

$$ADC_{it} = \alpha_0 + \beta_1(Post)_{it} + \beta_2(Post \times Month)_{it} + \beta_3(Treatment \times Post)_{it} + \beta_4(Treatment \times Post \times Month)_{it} + \varepsilon_{it}$$

Variable	Parameter	Interpretation
Post	B1	Average daily usage in the post-period
Post*Month	B2	Average daily usage in month *i*
Treatment*Post	B3	Average daily usage for the treatment group in the post-period
Treatment*Post*Month	B4	Average daily usage in month *i* in the post-period

- Fixed-effects: unique customer intercept terms for unobserved heterogeneity

LINEAR REGRESSION MODELING:

ALTERNATIVE SPECIFICATION – POST-ONLY MODEL WITH PRE-USAGE CONTROLS

Regression Model Specification: Post-Only

$$ADC_{it} = \alpha_0 + \beta_1(Treatment)_{it} + \beta_2(PreUsage)_i + \beta_3(PreUsageSummer)_i + \beta_4(PreUsageWinter)_i + \beta_5(MM)_t + \{Vector\ of\ Month\ \&\ Usage\ Interactions\} + \varepsilon_{it}$$

- Post-only with pre-usage controls
- Requires same dataset, can produce lower standard errors in some instances
- Standard evaluation approach compares D-in-D and Post-only specifications

LINEAR REGRESSION RESULTS

Regression Results

Coefficient	Estimate	Std Error	P-Value	5%	95%
Treatment*Post	-3.06	0.14	<0.001	-3.29	-2.84
Treatment*Post*February	1.07	0.18	<0.001	0.77	1.37
Treatment*Post*March	1.46	0.18	<0.001	1.16	1.76
Treatment*Post*April	2.06	0.18	<0.001	1.76	2.36
Treatment*Post*May	3.00	0.18	<0.001	2.70	3.30
Treatment*Post*June	3.75	0.18	<0.001	3.45	4.05
Treatment*Post*July	4.16	0.18	<0.001	3.86	4.46
Treatment*Post*August	3.97	0.18	<0.001	3.66	4.27
Treatment*Post*September	3.02	0.18	<0.001	2.72	3.32
Treatment*Post*October	2.08	0.18	<0.001	1.78	2.38
Treatment*Post*November	1.33	0.18	<0.001	1.03	1.63
Treatment*Post*December	0.21	0.18	0.245	-0.09	0.51

Adjusted R2: 0.6826

*Additional terms omitted from table for brevity

DOUBLE-COUNT ANALYSIS

$$Double\ Count = \left(\frac{Other\ Program\ kWH_{Trt}}{\#customers_{Trt}} - \frac{Other\ Pgrm\ kWh_{ctrl}}{\#customers_{ctrl}}\right) \times \#customers_{Trt}$$

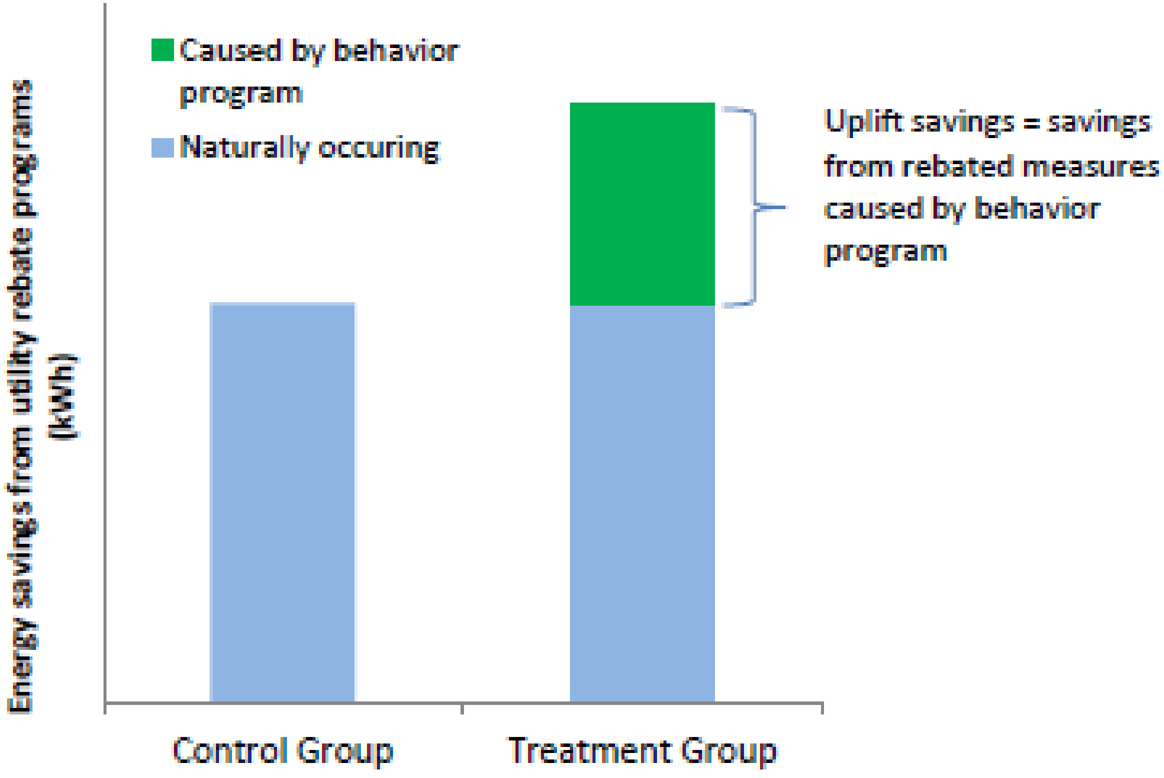
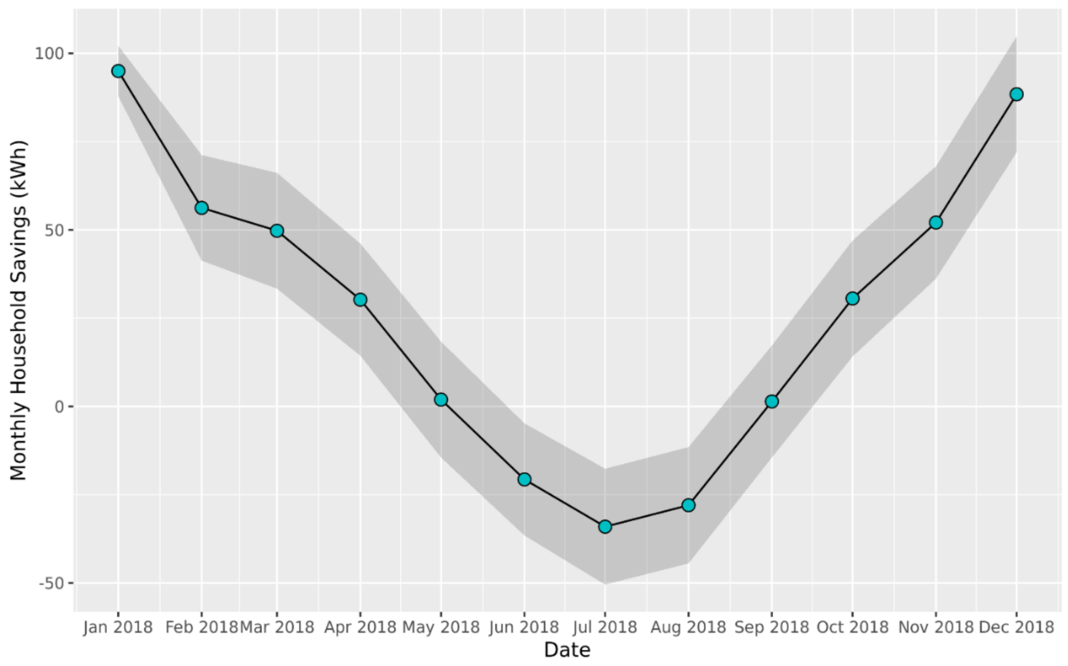


Figure 5. Calculation of double-counted savings

- Savings can be "double-counted" with energy impacts from other program interventions
- US policy: behavioral programs "last in line" for savings claim
- Net-out per-customer impacts from other programs, subtracted this total from behavioral program results
- Typically results in a < 1% reduction in program-level energy impacts

LINEAR REGRESSION RESULTS

Monthly Household Savings



Monthly Savings

Month	Impact Before Double Count	Double Counted Savings	Impact After Double Count	Percent Savings
January	94.98	-1.05	96.02	7.77%
February	56.23	-0.76	56.99	5.18%
March	49.75	-0.66	50.40	4.59%
April	30.22	-0.72	30.94	3.41%
May	1.93	-0.73	2.66	0.25%
June	-20.68	-0.68	-19.99	-2.95%
July	-34.06	-0.76	-33.30	-4.76%
August	-27.99	-0.85	-27.14	-3.92%
September	1.40	-0.83	2.23	0.19%
October	30.57	-0.88	31.45	3.49%
November	52.07	-0.52	52.60	5.06%
December	88.40	-0.48	88.88	7.38%
Total	322.82	-8.93	331.75	2.93%

LINEAR REGRESSION RESULTS

Calendar Year 2018 Program Savings

Wave	Weighted Customers	Annual Household Savings (kWh)	Annual Household 5% CI (kWh)	Annual Household 95% CI (kWh)	Program Savings (kWh)	Program Savings 5% CI (kWh)	Program Savings 95% CI (kWh)
Wave 1	9,961	331.75	368.81	276.83	3,304,735	3,673,840	2,757,624
Wave 2	10,648	218.36	285.59	124.91	2,325,150	3,040,996	1,330,038
Wave 3	13,724	382.01	426.03	332.20	5,242,634	5,846,828	4,559,102
Wave 4	38,827	271.29	304.59	244.07	10,533,217	11,826,182	9,476,661
New Movers	17,731	170.83	295.92	295.92	3,029,005	5,246,956	5,246,956
Total	**90,891**	**268.83**	**1,681.94**	**1,274.94**	**24,434,742**	**29,634,801**	**23,370,380**

LINEAR REGRESSION RESULTS

Calendar Year 2019 Program Savings

Wave	Weighted Customers	Annual Household Savings (kWh)	Annual Household 5% CI (kWh)	Annual Household 95% CI (kWh)	Program Savings (kWh)	Program Savings 5% CI (kWh)	Program Savings 95% CI (kWh)
Wave 1	9,413	339.07	382.58	285.83	3,191,646	3,601,199	2,690,508
Wave 2	10,040	238.15	315.76	150.58	2,391,015	3,170,294	1,511,818
Wave 3	12,937	409.95	446.04	346.75	5,303,406	5,770,400	4,485,893
Wave 4	36,059	272.58	301.63	237.69	9,829,207	10,876,578	8,570,906
New Movers	15,878	160.23	272.21	18.32	2,544,203	4,322,261	290,929
Total	**84,328**	**275.82**	**1,718.23**	**1,039.17**	**23,259,477**	**27,740,734**	**17,550,054**

ATTRITION ANALYSIS

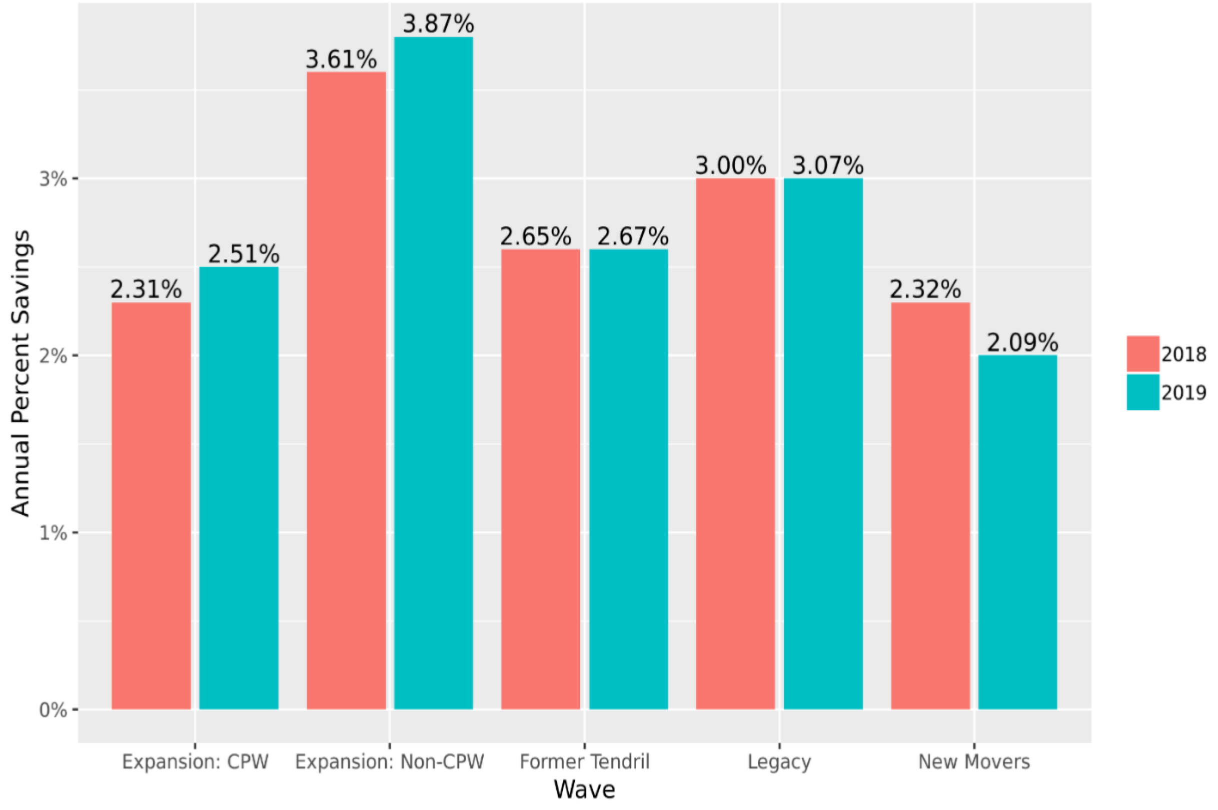
Calendar Year 2018 Moveout Rates by Wave

Wave	Treatment Customers	Control Customers	Treatment Moveout Customers	Control Moveout Customers	Treatment Moveout Percent	Control Moveout Percent
Wave 1	16,756	17,100	552	575	3.29%	3.36%
Wave 2	8,540	3,337	423	156	4.95%	4.67%
Wave 3	25,490	11,000	877	385	3.44%	3.50%
Wave 4	56,966	22,774	3,039	1,172	5.33%	5.15%
New Movers	34,325	34,366	6,890	6,728	20.07%	19.58%

- Moveout rates for each wave range between 3% and 6% with the exception of the New Movers wave
- New Movers wave attrition at 20% due to behaviors of targeted customers
  - These customers do not reside at a household for an extended amount of time

IMPLICATIONS OF EVALUATION APPROACH & RESULTS BY WAVE

- All evaluated waves displayed average annual electric savings of between 2% and 4% of annual billed use
- The New Movers wave displays the lowest savings at 2.3% and 2.1% likely due to shortened stay at residence
- HER programs are known to display larger savings effects as exposure to reports increases



SAVINGS PERSISTANCE / EFFECTIVE USEFUL LIFE

EFFECTIVE USEFUL LIFE IS A POLICY QUESTION

- Behavioral programs have persistence > 1 year
- Policy-makers may establish 1-year measure life for simplicity
  - **Benefit:** Creates a straightforward framework for establishing funding levels / program goals
  - **Drawback**: May result in repeated claiming of savings already induced

$$Lifetime\ kWh = 1st\ yr.kWh + \sum_{t=2}^{\infty} 1st\ yr.kWh + (1-\theta)^{t-1} \times (1-\lambda)^{t-1}$$

$Where,$

$t = year\ t$

$\theta = Savings\ degredation\ rate\ - \%\ decline\ in\ savings\ 1\ year\ after\ reports\ discontinued$

$\lambda = Program\ attrition\ rate$

$$Series\ Convergence: \frac{1st\ yr\ kWh}{\theta + \lambda - (\theta - \lambda)}$$

CALCULATING LIFETIME IMPACTS

$t = year\ t$

$\theta = Savings\ degredation\ rate\ - \%\ decline\ in\ savings\ 1\ year\ after\ reports\ discontinued$

$\lambda = Program\ attrition\ rate - \%\ of\ customers\ that\ drop\ out\ of\ treatment\ cohort\ annually$

$$Series\ Convergence: \frac{1st\ yr\ kWh}{\theta + \lambda - (\theta \times \lambda)}$$

<u>Wave 1 Example:</u>
25% degradation
3.29% attrition,
2,609,508 first-year kWh savings:

$Lifetime\ kWh: \frac{2{,}690{,}508}{25\%+3.29\%-(25\%\times3.29\%)}$=9,795,242

$$Measure\ LIfe = \frac{9{,}795{,}242}{2{,}690{,}508} = 3.64\ years$$

CALCULATING LIFETIME IMPACTS - EXAMPLE

CONCLUSIONS

CONCLUSIONS

- Behavioral programs consistently provide measurable electricity and natural gas savings
- Projects to be largest residential energy savings intervention in the United States, since lighting market has been transformed and codes & standards requiring LEDs have been adopted
- Savings are measurable through statistical analysis of billing data
- Measurability requires careful planning – Randomized Control Trial remains the “gold standard” for design

QUESTIONS?

Adam Thomas
Principal Consultant
ADM Associates, Inc.
+1 (916) 363-8383
adam@admenergy.com