Delivering Low-Income Whole-Home Retrofits: Consumption-Based Targeting for Deep Energy Savings

Bryce Dias, Pacific Gas and Electric Company; Roberta Brunkalla, CLEAResult Consulting; Josh Bode, Demand Side Analytics; Michael Logsdon, CLEAResult Consulting

ABSTRACT

Historically, low-income programs have relied on light-touch energy-efficiency measures and made them broadly available. Pacific Gas and Electric Company (PG&E) launched a novel pilot program aimed at low-income customers with the dual goal of testing innovations and achieving deep energy savings and decarbonization. How to best target and enroll low-income customers is a critical question. Customers with high energy or carbon savings potential are not always likely to participate (and vice versa).

PG&E, CLEAResult, and Demand Side Analytics (DSA) collaborated on a randomized controlled trial (an A/B test) to compare targeting based solely on adoption propensity (Group A) and based on both adoption propensity and energy use patterns extracted from smart meter data (Group B). The test compared the historical approach (Group A) to the innovation (Group B). CLEAResult used smart meter data and property data to develop a savings potential score based on seasonal heating and cooling consumption, annual loads, and energy use intensity.

The paper addresses several critical questions for program delivery:

- Does a targeting approach that combines usage patterns with adoption propensity help boost participation in deep energy savings and decarbonization among low-income households?
- Does the new targeting approach lead to more significant savings and carbon reduction?
- What are the bottlenecks in delivering deep savings and decarbonization for low-income households?
- How can this novel targeting approach be measured?

The paper will provide implementers, program administrators, and regulators with an increased understanding of how data-driven customer targeting could help deliver deep energy savings in the income-qualified space.

Introduction

The Energy Savings Assistance (ESA) program is a statutory[1] program funded through a public purpose surcharge that offers low-income households[2] weatherization, energy efficiency, and home upgrade services. California Public Utilities Commission Decision (D.) 21-06-015 authorized PG&E to implement the ESA Pilot Plus and Pilot Deep program (the Pilot) in 2022 to explore the feasibility of achieving persistent deep energy savings for low-income households.[3] PG&E is implementing the Pilot primarily in climate zone 12 (San Joaquin and Merced Counties predominantly) and limiting participation to households with single-family detached homes. By design, the Pilot serves as a testing ground for innovations.

Experimentation is at the very core of innovation. It is how we learn what works and what does not. Many pilots are implemented as program demonstrations and do not systematically vary or test design and delivery alternatives, and, as a result, do not always provide useful information for improving program performance (Sullivan, 2011). To determine what worked and what did not, it is necessary to be able to isolate specific components of key interventions—such as messaging, delivery channels, targeting, and incentive levels—and develop testable hypotheses.

Historically, low-income programs have relied on light-touch energy-efficiency measures and made them broadly available. The Pilot introduced several innovations that differ from PG&E's traditional low-income program. These innovations include:

- A focus on deeper savings. Specifically, the Pilot aims to include two saving tiers: "Plus," yielding 5–15% savings per home, and "Deep," yielding 15–50%.[4]
- A broader set of energy-efficiency and beneficial electrification measures.
- The use of a pay-for-performance payment structure instead of deemed savings.
- The use of new, data-driven targeting algorithms.
- Calibration of comprehensive home assessments and whole-home energy modeling to past usage.[5]
- Energy coaching.
- Post-installation monitoring using population-level normalized meter-based energy consumption (NMEC).

This paper focuses on testing one of these innovations: the use of new, data-driven targeting algorithms. Specifically, PG&E and its contractors CLEAResult and Demand Side Analytics designed and implemented a randomized controlled trial (an A/B test) to compare two distinct targeting algorithms:

[1] Public Utilities Code Section 2790(a).
[2] Customers in the participating utility's territory earning 250% of the Federal Poverty Level or below may qualify.
[3] D.21-06-015, p. 479.
[4] D.21-06-015 Attachment 2, p. 1.
[5] The Pilot utilizes Snugg Pro for estimating project modeled savings. Snugg Pro is home energy auditing software that allows for data collection and energy savings estimation.

- Targeting based solely on adoption propensity (the status quo).
- Targeting based on both adoption propensity and energy use patterns extracted from smart meter data (the innovation).

In its implementation plan, PG&E envisioned data-driven customer targeting as a novel means of selecting customers with high savings potential for participation in the Pilot. PG&E selected CLEAResult as the implementer in 2022, and CLEAResult proposed an advanced data-driven targeting approach. They have also been responsible for implementing those methods for the Pilot. PG&E selected Demand Side Analytics (DSA) to provide technical support for identifying program components for testing, designing randomized controlled trials, analyzing outcomes, and producing energy use intensity metrics (e.g., usage per square foot).

Targeting is central to program delivery and serves several purposes, by helping to:

- Identify customers most likely to participate.
- Identify customers with the greatest savings potential.
- Limit the number of customers contacted, thereby reducing the time and cost associated with assessing homes with less Pilot alignment and managing customer expectations.
- Select customers for home electrification who are most likely to see a neutral or positive bill impact post-treatment.

The Pilot's targeting methodology is designed around PG&E's Interval Data Analysis Tool (IDAT). IDAT analyzes smart meter data and produces a series of energy usage consumption features that indicate potential success in load management programs. The tool enables targeting while significantly minimizing the amount of smart meter data PG&E shares with its partners, protecting customer privacy, and reducing the risk of data loss. Potential uses of IDAT data include program targeting to identify the customers most likely to benefit from and enroll in a load management program. For the Pilot, the IDAT features were supplemented with property data to better understand customer savings potential.

A key element of innovation is that it is continuous and iterative. Innovations that are unsuccessful or inconclusive can either be discarded or modified and retested. In this case, in which PG&E, CLEAResult and DSA attempted to evaluate the accuracy of data-driven targeting techniques to reach customers with high savings potential, the evidence is inconclusive. Two observations may have contributed to this result: 1) the study was underpowered, or 2) some of the key assumptions about participation rates (and therefore sample sizes) were incorrect. Regardless, there is a pressing need to develop more effective methods for targeting customers with high savings potential to change behaviors and improve the delivery of energy-efficient and decarbonization technologies. Doing so will require systematic and controlled small-scale experiments such as the one presented in this paper.

Methodology

PG&E, CLEAResult, and DSA collaborated on a randomized controlled trial (an A/B test) to compare targeting based solely on adoption propensity (Group A) and based on both adoption propensity and energy use patterns extracted from IDAT data (Group B). This section provides details on the design of the randomized controlled trial and additional details on the two targeting algorithms tested (the interventions).

IDAT Feature Creation

PG&E's Interval Data Analysis Tool (IDAT) analyzes smart meter data and produces a series of energy usage consumption features of sites. The tool automates monthly data refreshes of 63 electric consumption features and 27 gas consumption features for approximately 5.5 million PG&E customer accounts. It was developed using R-based code from PG&E's previous research on targeting (Borgeson et al, 2018). For this research, DSA supplemented the IDAT features with property data to better understand customer savings potential (like calculating energy use intensity). The target population for the Pilot included customers with both electric and gas, and some with rooftop solar. Thus, gas and electric consumption were converted into MMBtu, and the estimated solar production was added back, to assess total energy usage (and energy use intensity) at a site.

Overall, the features that are available for targeting can be separated into four broad categories as shown in Figure 1. The data features are summaries of when, how, how much, and how efficiently each household uses electric and gas energy. It enables PG&E to provide information to vendors that is useful for targeting, while avoiding repeated transfers of large volumes of AMI data, which can pose security risks and also computational challenges. Due to the number of columns/features, the figure provides examples, rather a full list of all the features.

Descriptive summaries	AMI data analysis	Property data	AMI data analysis combined with property data
• Annual kWh • Annual therms • Loads coincident with peaking conditions	• Estimated summer cooling kWh • Estimated winter heating therms	• Property square footage • Year built • Lot size • Owner occupancy • Building type • Heating type	• Energy use intensity (EUI) • Heating energy use intensity • Cooling energy use intensity • EUI rankings overall and in comparison to similar buildings

Figure 1 Consumption features[6] available for customer targeting

[6] Summer cooling kWh is a feature computed by IDAT. It is the sum of cooling kWh (total daily kWh less baseload) for all days when cooling degree days (CDD) are greater than 0 (relative to local weather station temperature).

Randomized Controlled Trial Design

The primary challenge is to accurately detect changes in participation and energy savings while systematically eliminating plausible alternative explanations, including random chance. Did the new targeting approach lead to more significant savings and carbon reduction than targeting based on propensity scores alone?

The study was designed as a randomized controlled trial. In the design phase, DSA documented the hypothesis and the intervention tested, identified the data that would be collected and analyzed, identified the outcomes that would be analyzed, defined a meaningful effect, and established the sample sizes needed to detect a meaningful effect (i.e., statistical power analysis). The goal was to engage in science and leave little to no room for ambiguity regarding what data would be collected or how the data would be analyzed. Figure 2 summarizes the key design elements of the randomized controlled trial.

Power analysis provides the ability to detect a meaningful difference if one exists. Prior to conducting the randomized controlled trial, DSA conducted a power analysis that replicated the random assignment and analysis hundreds of times to quantify the ability to detect any difference in savings. The power analysis concluded that 300 participants were needed to detect an aggregate savings difference of 10 percentage points (e.g., 10% versus 20%) with 90% power.

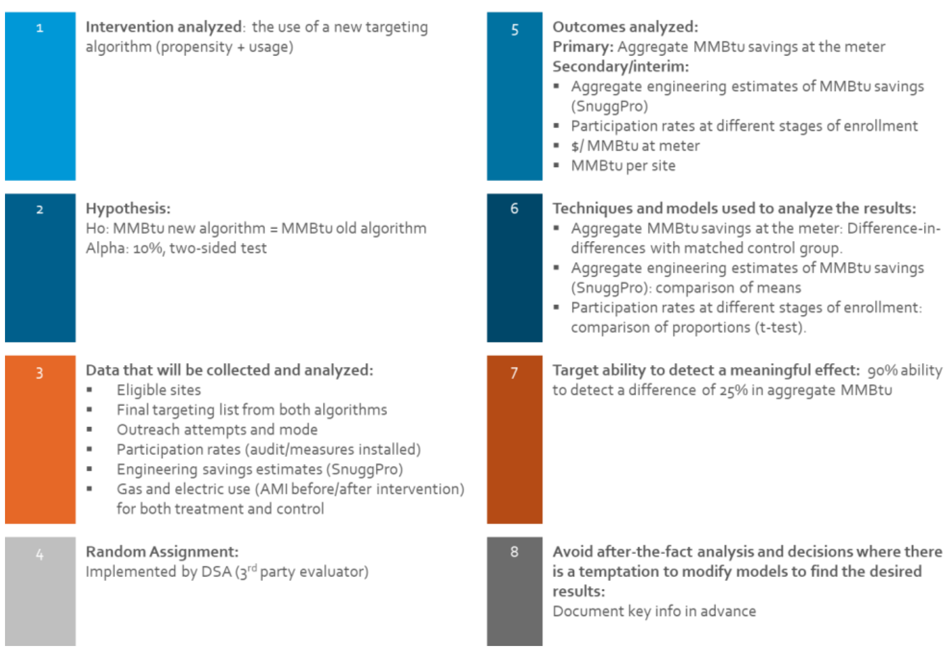

Figure 2: Key elements of randomized controlled trial design

Winter heating therms is computed by IDAT similar to summer cooling kWh. It is the sum of heating therms (total daily therms less baseload) for all days when heating degree days (HDD) are greater than 0 (relative to local weather station temperature).

Table 1 summarizes the power analysis sample recommendations and assumptions. PG&E could only control the targeting algorithm used. Ultimately, customers decided whether to apply and participate. Since the focus was on aggregate savings yield, the power analysis had to assume how enrollment rates, customer sizes, and energy savings would differ between the sites selected using the old and new targeting algorithms.

Table 1: Power analysis sample size recommendations

	Component	Old algorithm	New Algorithm
Input Assumptions	Assigned for selection	30,000	30,000
	Selected targeting List	6,000	6,000
	Assumed enrollment rates	3.00%	2.00%
	Assumed customer sizes (MMBtu/year)	50	75
	Assumed % savings (whole building)	10%	20%
Expected	Participants	180	120
	Aggregate MMBtu	900	1,800
	Statistical Power	92.70%	

For the randomized controlled trial:

- PG&E identified the sites that met the eligibility screening criteria. A total of 32,886 sites passed the screening (versus 60,000 in the plan), of which 21,702 were eligible for email.
- DSA randomly assigned 50% of the eligible sites with email (21,702) into two groups:[7]
    - Group A pool (10,853): used for propensity-only ranking and selection (old algorithm)
    - Group B pool (10,849): used for ranking based on propensity and high savings potential (new algorithm)
- DSA conducted checks to ensure the targeting pools were indeed randomly assigned.
- DSA identified the top 5,378 sites in the Group A pool using only the PG&E-developed propensity score.
- CLEAResult identified the top 5,378 sites in the Group B pool using a combination of savings potential estimates developed by CLEAResult and PG&E's propensity scores (new algorithm).
- Of the remaining sites, CLEAResult assigned 15,501 to a general marketing group, and the rest were excluded from the marketing campaign.
- Besides the application of the targeting algorithm, all other recruitment efforts for Group A and Group B were identical.

[7] DSA used proportional random sampling. Specifically, 50% of sites within each propensity score decile was randomly assigned to Group A or Group B.

Targeting Based on Propensity Scores Only (Old Algorithm)

For each marketing campaign and customer site, PG&E tracks who is sent recruitment materials, the mode by which they are recruited (email, direct mail, etc.), how many follow-up recruitment attempts are made, and whether the customer participated in the program (the outcome). The data is tracked for all programs and customer sites for three main reasons. First, it enables PG&E to coordinate recruitment efforts across multiple programs and marketing campaigns. Second, the data is also used to quantify the effectiveness of different recruitment tactics and the levels of marketing intensity. Third, and most relevant, the recruitment and participation data can be paired with customer characteristics to help us understand which customers are more or less likely to participate in specific programs. The likelihood of participation based on characteristics is commonly referred to as a propensity score.

PG&E utilizes a proprietary propensity model calibrated to the ESA program to predict which customers are most likely to meet the eligibility requirements and commit to participate in the program. The ESA propensity model generates granular scores, and customers are grouped into deciles based on their score on a scale of 1–10 (1 being best) relative to each other in any given campaign's population. The Q2 2023 Pilot campaign list was generated as a subset of the main ESA program's campaign, which occurred at the same time. Due to the Pilot's more selective screening criteria,[8] its Q2 list contained customers from every decile, though most sites were in the upper deciles.

Targeting Based on Savings Potential and Propensity Scores (New Algorithm)

The new algorithm was aimed at factoring in both savings potential and adoption propensity scores. With propensity scores already provided by PG&E, CLEAResult reviewed the available IDAT features and constructed a savings potential score. Given the importance of HVAC and shell impacts on energy usage in the selected climate zone (Scheer et al, 2017), we started estimating savings potential with HVAC as a percentage of the load, total annual load, and energy use intensity (EUI). These first principles strategy was necessary at the start, as the data to develop predictive models of energy savings based on IDAT features was not yet available, since relatively few customers had enrolled at the time of the randomized controlled trial.[9]

CLEAResult anticipated that most savings would need to come from shell and HVAC measures to achieve deep savings. The ideal candidate would be a home with large heating and/or cooling loads, especially relative to the square footage. For the consumption-based portion of the targeting score, five features were used:

- Summer cooling kWh % of total (inferred June–September cooling kWh divided by total June– September kWh). Both variables were from PG&E's IDAT

---

[8] Criteria included: geography (climate zones 11 and 12), PG&E customer with both gas and electric service for a minimum of 12 months, single-family detached home, and no participation in ESA within two years.

[9] Approximately 15 customers were enrolled between October 2022 (when the Pilot began outreach) and April 2023 (when the experimental design for Q2 targeting was defined). Initial outreach campaigns were deliberately small to enable observation of uptake and integration of feedback.

features. The cooling kWh was estimated by running individual customer regressions to isolate electric cooling loads, heating loads, and base loads.

- Total annual kWh
- Winter heating Therms % of total (inferred December–February heating Therms divided by total Therms December–February). Both variables were from PG&E's IDAT features. The heating therms were estimated by running individual customer regressions to isolate gas heating loads from base loads.
- Total annual Therms
- Total energy use intensity (total MMBtu/sq. ft). All of PG&E's residential accounts were linked to tax assessor property data allowing the estimating of energy use intensity.

These features were chosen to balance the expected savings percentage and total savings. If the score was only based on EUI or HVAC percentages, the highest scores would skew to the smallest houses with the lowest base loads. That might yield higher savings percentages but lower total savings. If the scores were based on total usage, it would skew the high scores to the biggest houses, which may not need this type of whole-home retrofit.

The concept was to establish a standardized scale for savings potential by converting the five consumption features into z-scores (calculated by subtracting the mean and dividing by the standard deviation). Next, the average of the scaled consumption features was used to develop a savings potential score that was divided into deciles. To align with PG&E's propensity score scale, 1 represented the highest and 10 the lowest potential savings. The 10 savings potential deciles were averaged with the 10 propensity deciles to calculate a composite score, which was ultimately used to identify the target population for Group B.

Figure 3 shows seasonal electric consumption by savings potential decile. Sites with the highest potential were similarly sized but had higher overall HVAC usage and thus were better candidates for HVAC, shell, and heat pump measures.

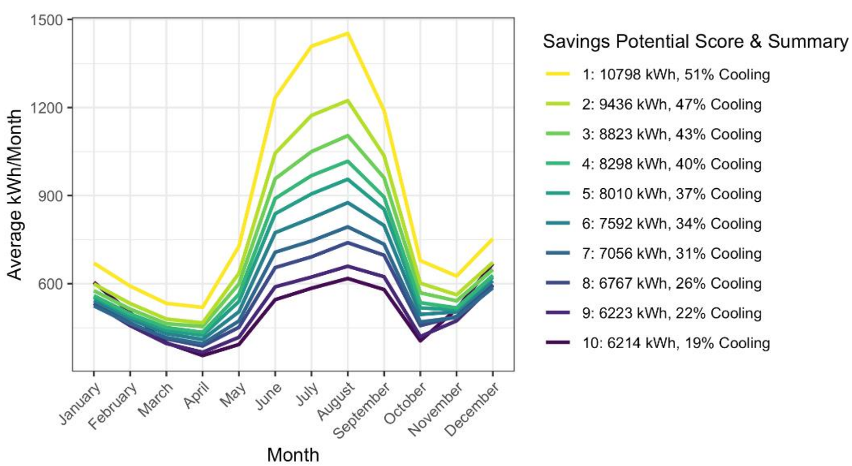

Figure 3: Seasonal electric consumption by savings potential decile

Results

Randomized Controlled Trial Results

The marketing campaign to test the old and new targeting algorithms was launched in April/May 2023, and installations occurred from May through December 2023. Because a full year of measured post-installation savings is not available yet, the initial analysis focused on the various participation stages and the energy savings modeled after the whole-home assessment.

Table 2 compares the groups *before* the eligible population was randomly assigned to the old and new targeting algorithms and shows both groups were randomly assigned. Figure 4 compares the distribution of propensity deciles. The groups assigned to each of the recruitment pools had no statistically significant difference, as expected, indicating that the population assigned for targeting by each algorithm were identical.

Table 2: Comparison before application of targeting algorithms

Variable/Feature	Group A Mean (n= 10,853)	Group B Mean (n = 10,849)	se	t	pval
kwh_tot_top10_days	446.3	445.7	3.25	-0.200	0.84
cooling_kwh	1,426.8	1,419.3	15.03	-0.500	0.62
heating_kwh	422.8	425.3	8.47	0.290	0.77
tot_therms_annual	10,280.4	10,262.3	60.49	-0.300	0.76
heat_therms	149.28	149.86	1.13	0.520	0.61
cvrmse_elec_hourly	0.940	0.940	0.01	0.260	0.80
cvrmse_elec_daily	0.400	0.400	0.00	0.680	0.50
cvrmse_gas_daily	0.690	0.690	0.00	0.050	0.96

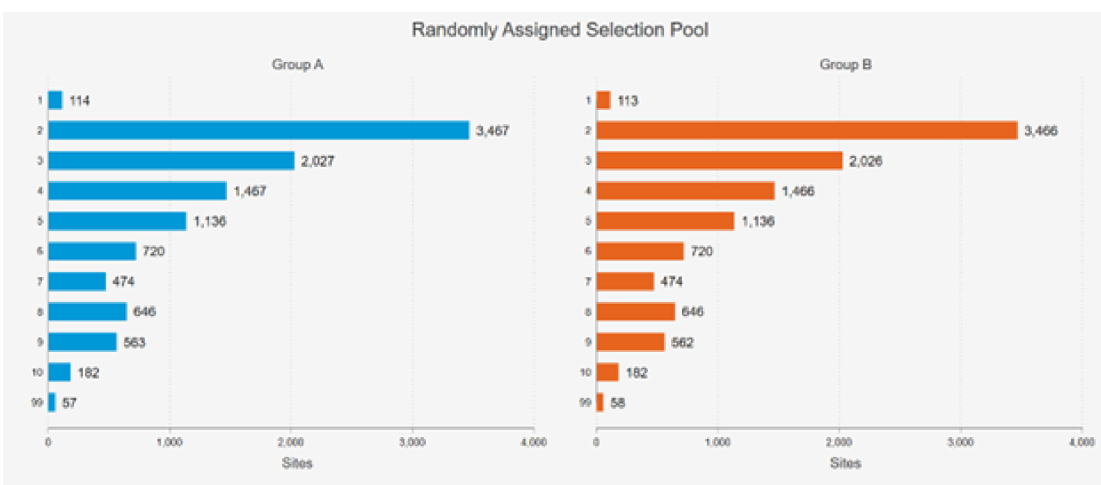

Figure 4: Distribution of propensity deciles b*efore* application of targeting algorithms

Table 3 compares the A/B groups after algorithms had been applied and the recruitment lists had been created. The algorithms were expected to introduce differences between the two groups. After the algorithms were applied, Group B had higher annual electric use, higher annual gas use, more heating loads, and more cooling loads than Group A. While Group A had a better propensity score, Group B had a better composite score, which combined both savings potential and adoption propensity.

Table 3: Comparison after application of targeting algorithms

Variable / Feature	**N1**	**N2**	**Group A Mean**	**Group B Mean**	**Std. Error**	**t-stat**	**P. value**
Propensity Decile	5378	5378	2.34	3.01	0.019	34.40	0.000
Combined Score (Savings potential + Propensity)	5378	4985	3.87	3.45	0.025	-16.60	0.000
Solar	5378	5378	14.2%	15.5%	0.007	1.98	0.048
Annual kWh	5378	5378	7,994.7	8,677.0	80.578	8.47	0.000
kWh on top 10 load days	5378	5378	450.0	515.8	4.176	15.74	0.000
Cooling kWh (Jun-Sep)	5378	5378	1,466.8	1,780.2	20.625	15.20	0.000
Heating kWh	5378	5378	417.6	389.8	10.973	-2.53	0.011
Annual Therms	5378	5378	423.7	478.3	3.432	15.89	0.000
Heating Therms (Add months)	5378	5378	144.2	179.6	1.486	23.81	0.000
cvrmse_elec_hourly	5374	5377	0.92	0.92	0.009	-0.51	0.609
cvrmse_elec_daily	5378	5378	0.39	0.39	0.004	-0.02	0.987
cvrmse_gas_daily	5378	5378	0.68	0.62	0.004	-14.09	0.000

Table 4 compares the two groups at various stages of the enrollment process. Figure 5 visualizes the difference in proportions and whether the difference is statistically significant. As expected, the group targeted using propensity score alone (Group A) had higher response rate, 14.2%, than the group targeted using the propensity score and savings potential (Group B). Beyond the initial two participation steps, however, the differences are not statistically significant. Overall, the rate of installation was substantially lower than the rates assumed in the study design and lower than the recommended sample sizes.

Table 4: Comparison of means for participation steps

Participation Step	**N (Group A)**	**N (Group B)**	**Proportion (Group A)**	**Proportion (Group B)**	**Difference**	**Std. Error of Difference**	**Z-stat**	**P. value**
Responded to recruitment	5,378	5,378	14.22%	12.05%	2.18%	0.0065	3.34	0.001
Initial contact/call	5,378	5,378	11.25%	10.00%	1.25%	0.0059	2.10	0.036
Pre-qualification call	5,378	5,378	4.95%	4.52%	0.43%	0.0041	1.04	0.296
Assessment created	5,378	5,378	0.73%	0.74%	-0.02%	0.0016	-0.11	0.910
Eligibility Docs.	5,378	5,378	0.48%	0.60%	-0.11%	0.0014	-0.79	0.430

Assessment complete	5,378	5,378	0.69%	0.63%	0.06%	0.0016	0.36	0.721
Treatment plan sent to PG&E	5,378	5,378	0.43%	0.30%	0.13%	0.0012	1.12	0.261
PG&E approved treatment plan	5,378	5,378	0.43%	0.28%	0.15%	0.0011	1.30	0.194
Review plan with customer	5,378	5,378	0.43%	0.28%	0.15%	0.0011	1.30	0.194
Contractor assigned	5,378	5,378	0.41%	0.28%	0.13%	0.0011	1.15	0.249
Scheduled installation	5,378	5,378	0.37%	0.22%	0.15%	0.0011	1.42	0.157
Measures installed	5,378	5,378	0.26%	0.19%	0.07%	0.0009	0.82	0.414
Passed QC	5,378	5,378	0.11%	0.07%	0.04%	0.0006	0.63	0.527

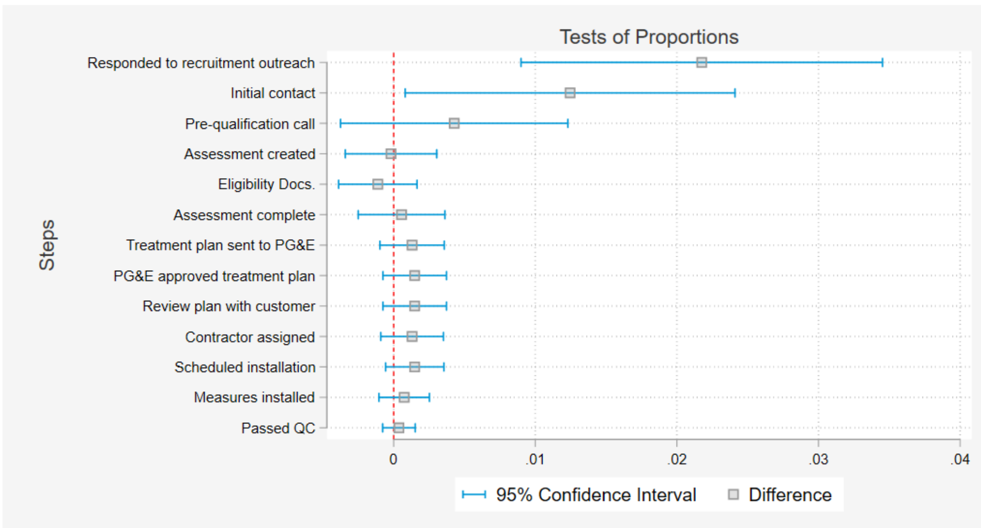

Figure 5: Test of Differences in proportions by participation step

Participation rates alone are half of the equation. We expected homes targeted using both savings potential and propensity scores (Group B) to be larger and have savings potential. Once enrolled, CLEAResult conducted comprehensive energy assessments at each home, utilizing energy modeling software to estimate energy savings. Figure 6 summarizes the modeled electric (kWh), gas (therms), and total (MMBtu) savings by site.

Figure 6: *Modeled average home savings by targeting algorithm*

On average, sites which included savings potential in the targeting algorithm had higher modeled energy savings. However, the differences were not statistically significant. The number of participants was smaller than the randomized control trial design. In addition, the percent difference in savings was also smaller than what the study was designed to detect.

Figure 7 shows the modeled aggregate savings, which combine both estimated savings per participant and the number of participants. Overall, the propensity only algorithm (Group A) delivered larger aggregate savings than the saving potential plus propensity score algorithm (Group B). However, the difference in aggregate savings is not statistically significant. Overall, the higher savings from sites targeted using the new algorithm (Group B) were not sufficiently large to outweigh the higher adoption rates among sites targeted based on the old algorithm (Group A).

Figure 7: Aggregate modeled energy savings from randomized control trial

Targeting based on savings potential and propensity still shows promise but needs to be refined. The initial savings potential scores were developed without the benefit of actual results from projects. In addition, the two groups were very similar. Too much of the targeting pool was selected, leading to insufficient differences in the composite scores (savings plus propensity).

The Benefit of Using Savings Potential Scores

In lieu of finding an advantage in Group B, we set out to determine if the savings potential score as originally conceived was useful at all. If the savings potential score was weak or uncorrelated with modeled savings, we would want to rethink the entire approach.

The savings potential score was in fact informative, especially for total BTUs saved. Figure 8 shows average total savings for each project in every propensity decile with a 95% confidence bound in grey. There is a threefold difference in modeled savings between the lowest and highest scores and the confidence bound is tight around the regression. The correlation with savings percentage was weaker. Figure 9 plots the average percent savings by decile and shows a weaker relationship between percent savings and propensity decile. When we developed the scores, we tried to balance total BTUs against savings percentage, and it appeared the score skewed toward selecting total BTUs.

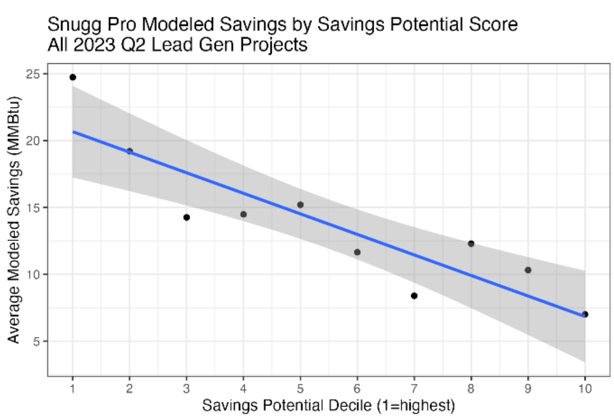

Figure 8: Modeled average savings by savings potential decile

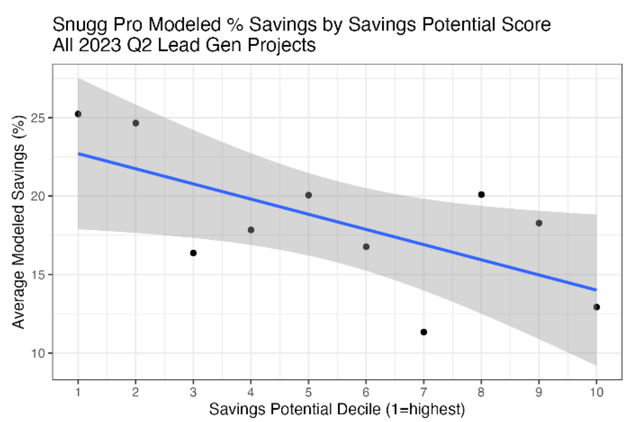

Figure 9: Modeled savings as a percentage of total home savings by savings potential decile

If there was such a strong relationship between savings potential score and modeled savings, then why did the early results show that Group B had only a small lead on savings per site? To answer that, we looked at the composition of propensity and savings potential scores within each group. Group A had an average savings potential score of 5.4, while Group B had an average savings potential score of 3.9, a difference of 1.5. When we projected those numbers onto the linear regression line of a scatterplot, the results were 13.7 and 16 MMBtu, respectively, or a 17% advantage in savings per site for Group B. Group A had an average propensity score of 2.3, while Group B's average score was 3. While we were able to limit Group A to propensity scores 1-3, we had to select scores 1-8 for Group B. The impact of higher Group A propensity scores is reflected in their higher response rates in Table 4.

Key Findings and Conclusions

Pacific Gas and Electric Company (PG&E) launched this Pilot with the aim of testing innovations and achieving deep energy savings for low-income customers. The goal of this project was to combine usage patterns with adoption propensity to help boost deep savings for low-income households. An important aspect of innovation is that it is continuous and iterative. Innovations that are unsuccessful should be discarded or modified and retested.

The key findings from the study were:

- Customers targeted based on propensity alone showed more interest (significant) and were more likely to complete the installation (not significant).
- Customers targeted based on propensity plus savings potential score showed higher savings per site (not significant), but lower interest rates (significant) and lower installation rates (not significant).
- The net effect is inconclusive. Targeting with the propensity plus savings potential score produced higher average energy savings, whereas targeting with propensity alone produced higher aggregate savings.

While the results of this project are inconclusive, there were lessons learned for the team, who will iterate on the approach for 2024.

- **Separate the operational tests from testing of innovations.** The Pilot was relatively early in operational maturity, leading to lower enrollments than what was necessary for the randomized control trial.
- **Do not over-screen if you are testing targeting algorithms.** We started with a much smaller pool to make selections from than what we had hoped for. Between Group A, Group B, and general marketing, we ended up marketing to almost the entire pool of candidates available post-screening (using propensity deciles 1-8 for Group B).
- **Savings potential scores appear to be correlated with real savings.** The savings potential score was developed in the absence of sufficient participation data. Despite that, they were heavily correlated to energy savings and percent savings based on a detailed model of the building.
- **There is a benefit to systematically testing innovations, even when the test fails.**

PG&E and CLEAResult have already implemented operational improvements to increase ultimate installation rates for the 2024 campaign. This includes increasing call center staff, releasing email and direct mail in small batches over six weeks, and recruiting a larger share of pre-qualified customers.

Deliberate targeting methods can yield beneficial program outcomes, including greater administrative and operational efficiency and greater customer satisfaction, and can support PG&E’s efforts to achieve new objectives, such as deep energy savings and electrification for low-income customers. We are also aware of the limitations inherent in data-driven methods. For instance, not all customers have a sufficient energy usage history to inform consumption-driven targeting, which is likely more prevalent among renters, who may relocate more often than homeowners. With this limitation in mind, and to promote equity by incorporating those who might otherwise be screened out of the targeting process, the Pilot is also implementing more traditional means of recruiting customers, such as canvassing and contractor-driven outreach.

The findings from this Pilot are intended to inform the future of PG&E's low-income customer programs, including whether, how, and to what extent PG&E can apply new targeting methods to the existing ESA program or integrate targeting into future ESA applications. This Pilot was designed to allow frequent experimentation, with the flexibility to incorporate lessons learned between experiments. While the experiment results highlighted in this study are inconclusive, PG&E, CLEAResult, and DSA are optimistic about the direction in which the Pilot is headed. The real innovation is having a dedicated laboratory for testing, refining, and retesting program design and implementation modifications, identifying ones that work, and incorporating them into the larger program. The dedicated testing ground and use of scientific tests leave us optimistic overall about uncovering more effective ways of delivering energy savings to customers in low-income, disadvantaged, and under-resourced communities.

References

Borgeson, S., and B. Gerke. 2018. Energy Efficiency Program Targeting: Using AMI Data Analysis to Improve At-the-meter Savings for Small and Medium Businesses. Oakland, CA: PG&E. https://www.calmac.org/publications/SMB_Targeting_Report_FINAL_8-10-18.pdf

Scheer A., S. Borgeson, K. Rosendo. 2017. Customer Targeting for Residential Energy Efficiency Programs: Enhancing Electricity Savings at the Meter. Whitepaper. www.pge.com/content/dam/pge/docs/about/doing-business-with-pge/Customer-Targeting-Final-Whitepaper.pdf.

Sullivan, M. 2009. Using Experiments to Foster Innovation and Improve the Effectiveness of Energy Efficiency Programs. California Institute for Energy and Environment and the California Public Utilities Commission's Energy Division. www.calmac.org/publications/Experimental_Design_White_PaperES.pdf

Sullivan, M., E. Vine, L. Lutzenhiser, C. Blumstein, and B. Miller. 2011. Experimentation and the Evaluation of Energy Efficiency Programs: Will the Twain Meet? International Energy Program Evaluation Conference. www.researchgate.net/publication/282253082_Experimentation_and_the_Evaluation_of_Energy_Efficiency_Programs_Will_the_Twain_Meet.