Southern California Edison

Summer Reliability Program

Measurement & Verification Plan

Version 2.0

June 2023

Table of Contents

1. Introduction

Background

On July 30, 2021, Governor Newsom signed an emergency proclamation to "free up energy supply to meet demand during extreme heat events and wildfires that are becoming more intense and to expedite deployment of clean energy resources this year and next year." In the Governor's July 30, 2021 Emergency Proclamation, all energy agencies, including the California Public Utilities Commission, were directed to act immediately to achieve energy stability during this emergency.

In response to the Governor's Emergency Proclamation, on August 6, 2021, the assigned Administrative Law Judge (ALJ) sent an e-mail ruling to parties in R.13-11-005, seeking input on actions that the Commission could take, specific to energy efficiency (EE) and reliability, to help support the Governor's Proclamation and the Commission's overall goals. After receiving comments on the ruling from the parties, on December 8, 2021, the Commission issued the Decision, which orders the IOUs to take actions to prepare for potential extreme weather in the summers of 2022 and 2023.

On February 7, 2022, Southern California Edison (SCE) submitted Advice Letter 4715-E for SCE's Market Access Program to the California Public Utilities Commission (Commission or CPUC). The Advice Letter received CPUC approval on March 9, 2022.

Program Description

The Decision authorizes a two-year Market Access Program (MAP), marketed as the Summer Reliability Program (SRP) by Southern California Edison (SCE), that is being funded by $150 million allocated among Pacific Gas and Electric Company (PG&E), San Diego Gas & Electric Company (SDG&E), and SCE, to deliver peak and/or net peak demand savings using the normalized metered energy consumption (NMEC) method of measuring energy and peak demand savings in residential and commercial buildings for 2022-2023.

This measurement and verification (M&V) plan provides technical details regarding the estimates of energy savings that underpin both sets of compensation. It is important to recognize that the Population NMEC procedures used to settle with TradePros are the same procedures SCE will use to report program performance. The program-level achievements are simply the sum of the performance estimates across TradePro groups.

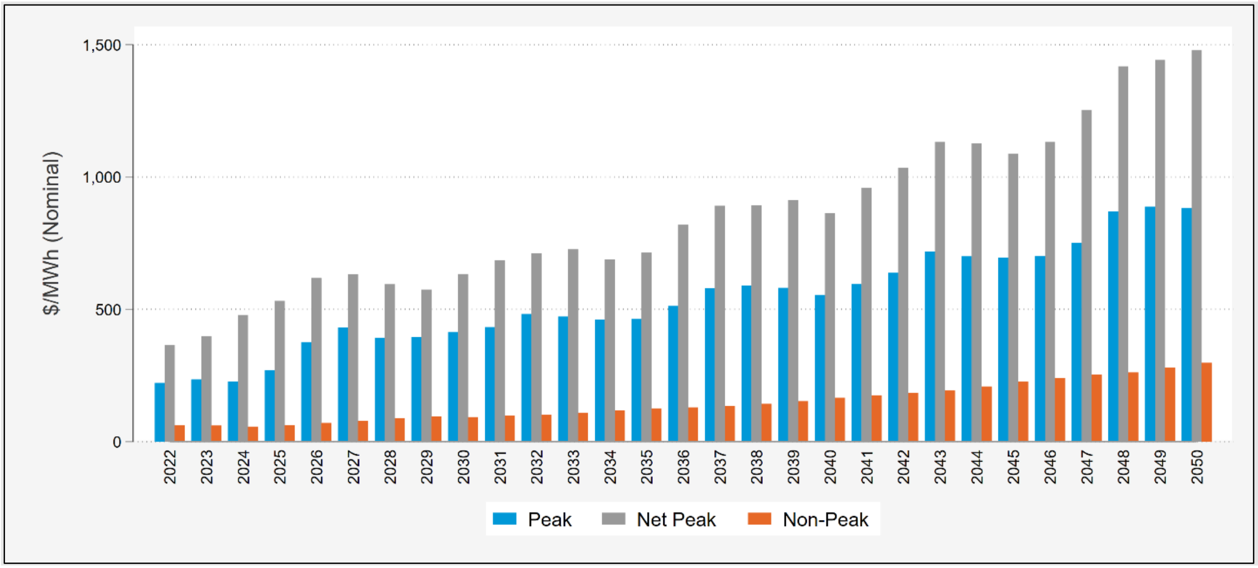
The primary objective of the Summer Reliability Program is to deliver peak and net peak demand savings during the summers of 2022 and 2023. To clearly signal this goal to the market, SCE will show voided cost tables by period and EUL to incentivize projects and measures that (a) deliver peak and net peak reductions and; (b) are long-lived (EUL). SCE classifies individual hours across the year into three separate categories: Peak, Net Peak, and Non-Peak hours:[1]

- **Peak hours**: Hours between 4 p.m.-7 p.m. on business days between June 1 to September 30
- **Net Peak hours**: Hours between 7 p.m.-9 p.m. on business days between June 1 to September 30
- **Non-Peak hours**: All other hours will be considered Non-Peak.

[1] SCE has elected to define peak as 4 p.m.-7 p.m. instead of 4 p.m.-9 p.m. so that the peak and net peak periods are mutually exclusive.

Figure 1 shows the difference in value to the system across these three categories over time. The avoided costs come from the CPUC's 2021 Avoided Cost Calculator[2] and are an average across SCE's eight climate zones. Energy savings during the Peak and Net Peak periods provide 3-8 times the grid value of Non-Peak hours, on average, and the SRP program compensation structure reflects this price signal to the market.

Figure 1: Average Avoided Cost by Year and Period – SCE Climate Zones (Nominal)



2 https://www.ethree.com/public_proceedings/energy-efficiency-calculator/

Summary of Key M&V Plan Elements

Table 1: M&V Plan Overview

M&V Consideration	Planned Approach
Settlement Enrollment Group Definition (Population)	All projects completed by a TradePro within a four-month period. The three periods are defined as: • February through May • June through September • October through January
Analytical Method	Individual premise regression. The model used is a seasonal Time of Week Temperature (TOWT) model that includes 168 hour-of-week dummy variables, a temperature spline, and one or more granular profiles which act as a synthetic control[3].
Evaluation, Measurement and Verification Consultant	SCE has retained kW Engineering and Demand Side Analytics to develop and implement this M&V plan and build out the settlement platform
Calculation Software	Stata 16.1
Data Collection Strategies	Upfront capture of typical efficiency attributes: • project location (contract number) • project start and completion date • equipment type, quantity, capacity, and specifications • project cost Back-end consolidation of participant meter data, performance estimates, and compensation
Performance Metrics	• Aggregate peak kW savings • Aggregate net peak kW savings • Annual kWh savings • Weighted Average EUL • Total System Benefits
Determination of Net Savings	Analytical method directly estimates net savings. If gross savings are required a default net-to-gross ratio will be applied to reverse engineer gross savings.
Weather normalization	Settlement and reporting will be based on actual ex-post measurement of savings during the 2022-2023 observation period without weather normalization. Regression models developed using data from the baseline period will be used to predict participant loads during the performance period.

3 Abadie, Alberto. 2021. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature*, 59 (2): 391-425.

Total System Benefits Calculation	2021 ACC values by climate zone averaged by (1) climate zone (2) year (3) month (4) business/non-business day. Capacity value is spread across Monday-Saturday within hours of the same month.
Reporting	TradePro-specific reports issues with compensation performance at the end of each four-month period. Aggregate reports to SCE and CPUC that encompass performance across all TradePros.

2. Methods, Savings, Eligibility, and Compensation Structure

Suitability of NMEC Methods

SCE's design of its Summer Reliability Program follows the CPUC's guidance for a Market Access Program which will enhance grid reliability during the summers of 2022 and 2023 as noted in the introduction to this M&V plan. Importantly, the Decision specifies that a population-level NMEC approach be used for the measurement of energy savings and the resultant settlement and reporting.

While one approach to developing a program evaluation strategy is to base it on the measures, delivery approach, market segments and other key factors that drive savings estimation, the approach for the Summer Reliability Program has been specified by the Decision to take the form of population NMEC. As such, this M&V plan will focus on this approach.

The ability to measure energy savings accurately using population NMEC methods depends on four key components:

1. **The effect or signal size** – The effect size is most easily understood as the percent change in energy use following the intervention. It is easier to detect large changes than it is to identify small ones.
2. **Inherent data volatility or background noise** – The more volatile the load, the more difficult it is to detect small changes. Non-routine events effectively add noise to the data.
3. **The ability to filter out noise or control for volatility** – Statistical models – no matter how simple or complex – are tools to reduce noise (or unexplained variation) and allow the effect or impact to be more easily detected.
4. **Sample/population size** – The full participant population is analyzed. Regardless, it is easier to precisely estimate average impacts for a large population than for a small population because individual customer behavior patterns "smooth out" and offset individual customer volatility across large populations.

With these considerations in mind, the Summer Reliability Program has been designed to be compatible with population NMEC measurement methods, as shown in Figure 2, below.

Figure 2: Program Design Elements to Increase NMEC Suitability

Effect Size	• Target measures and business types with a large expected reduction during the peak and net peak windows
Data Volatility	• Focus outreach on industry types with weather-sensitive loads and consistent operating schedules • Prohibit participation by sites with behind-the-meter solar and batteries
The Ability to Filter out Noise	• Conduct all modeling on hourly data • Leverage granular comparison group profiles as explanatory variables
Population Size	• Educate Trade Pros on the benefits of aggregation for settlement frequency and reduced uncertainty • Provide warnings about settlement risk for cohorts with limited participants.

Eligibility and Permissible Measures and Projects

The Summer Reliability Program is open to residential and select non-residential customers who receive SCE electricity and/or electricity distribution services and pay into the Public Purpose Program (PPP) surcharge. These requirements include:

Commercial Customer Requirements

- Project site must be located in SCE's service area
- Bundled customers (that is, receiving electric service from SCE), or
- Unbundled customers (that is, receiving electric generation service from SCE)
- The customer must pay the Public Purpose Program (PPP) surcharge on the electric meter in which the energy-efficient equipment is being proposed:
  - This is broken down on the bill where the costs that make up the Delivery Charges are shown
- Site is not currently participating in, and has not participated in the past 12 months, in a CPUC ratepayer-funded energy efficiency program:
  - Eligibility of master-metered facilities will be determined case by case with SCE, based on the expected savings at the facility level

- 12 consecutive months of energy usage data is available
- Customers must be individually metered, no sub-metering
- The customer is individually metered and has adequate pre-installation billing history on an SCE meter
- The project site does not have on-site generation (that is, solar, thermal energy storage, battery storage, etc.) and customer agrees not to install any during the 12-month performance period
- Fuel substitution measures are not eligible due to their potential electrical load building nature
- Natural gas measures are also not eligible due to the program's focus on electrical load reduction only
- SCE may also elect to flag sites as ineligible based on recent energy efficiency participation if the recent project is expected to confound the population NMEC measurement.
- Tenants must have authorization from the property owner or property management company to implement the upgrades
- The customer must agree to provide all required documentation and access to the facility for project-related audits, inspection or data gathering by SCE or by the CPUC
- Commercial customers with a maximum demand of ≥ 500 kW (that is, TOU-8) will require pre-approval
- The energy efficiency project will reduce at least 5% of the customer's metered annual electrical usage at the project site
- Projects (combined measures) must have a weighted average EUL of ≥ 1 year
- Projects must be fully installed, commissioned, and have an error-free installation report submitted no later than March 31, 2024, and
- Per Decision 16-08-019, Industrial NMEC projects are not eligible outside of Strategic Energy Management (SEM).

Residential Customer Requirements

- Must be a qualifying single-family or manufactured mobile home
- The residence must have its own service account
- The energy efficiency project will reduce at least 3% of the customer's metered annual electrical usage at the project site
- Renters must have authorization from the property owner or property management company to implement the upgrades
- The project site does not have on-site generation (that is, solar, thermal energy storage, battery storage, etc.) and customer agrees not to install any during the 12-month performance period:
    - SCE may also elect to flag sites as ineligible based on recent energy efficiency participation if the recent project is expected to confound the population NMEC measurement.
- The customer is individually metered and has adequate pre-installation billing history on an SCE meter:

    - Eligibility of master-metered facilities will be determined case by case with SCE, based on the expected savings at the facility level, and

- The customer must agree to provide all required documentation and access to the facility for project-related audits, inspection or data gathering by SCE or by the CPUC.

Projects must be fully installed, commissioned, and have an error-free installation report submitted no later than March 31, 2024.

Qualifying Energy Efficiency Measures

The Summer Reliability Program accepts a wide variety of energy-saving projects for residential and non-residential sites. All measures must meet the following criteria:

1. Equipment retrofits, weatherization, and add-on equipment
2. Retrocommissioning (RCx) measures
3. Must exceed baseline energy performance by a minimum of 3% (5% for commercial)
4. Must operate at least five years
5. Must be permanently installed
6. Cannot overlap with other energy efficiency programs funded by SCE, including Statewide Programs
7. Existing equipment must be decommissioned and removed

Table 2, below, shows the commercial building types that SCE plans to target with SRP outreach efforts, and Table 3 lists some target measures within these building types. This list is not exhaustive and other sectors may be added upon pre-approval.

Table 2: Target Building Types

Building Type
Supermarkets and Other Grocery (except Convenience) Stores
Limited-Service Restaurants
Department Stores
Drugs and Druggists' Sundries Wholesalers
New Car Dealers
Pharmacies and Drug Stores
Gasoline Stations with Convenience Stores
Warehouse Clubs and Superstores
Convenience Stores
Other Grocery and Related Product Wholesalers
Home Centers
Indoor Horticulture

Table 3: Examples of Eligible SRP Measures (Subject to Administrator Approval)

Lighting Measures	• Interior LED Product Retrofits • Exterior LED Retrofits • Lighting controls systems
HVAC	• Efficient Electric Hot Water Heat Pump Retrofits • Controls and energy management systems for HVAC
Refrigeration	• Controls and energy management systems for Refrigeration • High efficiency evaporator fan motors • Retrofit of refrigerated/freezer cases and doors
Weatherization	• Cool Roof

Estimated Savings

The project savings forecast calculations are essential to assess the viability of the project. They serve as an important guide to the metered data analysis and to ensure that the project's energy use is within acceptable tolerance levels towards the projected energy savings. At the project application stage, the package must provide a clear, detailed, all-inclusive, and defensible explanation of the energy savings and demand reduction calculation methodology that incorporates a weighted EUL methodology. The project application must explain all assumptions and provide fully reviewable calculations. The project application should reference relevant DEER, EM&V, CPUC, and use pre-approved preferred program administrator calculation tools. Supporting attachments should be embedded or referenced in the Attachments and References sections. If any measures are taken directly from or created with READI, either embed the READI export or indicate the DEER Measure ID.

Load shapes for estimated savings are used for portfolio lifecycle cost analysis of a measure's energy savings over one year. A load shape is a set of fractions summing to unity, with one fraction per hour (or other time period). Multiplying a savings value by the load shape value for any particular hour yields the energy savings for that particular hour. If possible, use DEER load shapes, which are hourly. The ideal load shape for net benefits estimates would represent the difference between the base case and measure case.

Effective Useful Life

The weighted average EULs should comprise the best available estimate of the relative contribution of different measures to total savings, based on available data. A weighted average EUL across measures will be determined for each participating site during upfront project processing. Service providers should consult with SCE about the approach to calculating weighted average EUL and provide their calculations and the data used.

Weighted average EUL example:

- Measure 1: 100,000 kWh savings, 10-year EUL
- Measure 2: 200,000 kWh savings, 3-year EUL

The EUL of the bundle would be (100,000 * 10 + 200,000 * 3) ÷ (100,000 + 200,000) = 5.33 years.

The same fundamental weighted average calculation across projects will be used for settlement with TradePros based on TSB and reporting of program-level impacts. The site-level EUL determined during upfront project processing will be used for these performance calculations because the population NMEC method does not quantify the EUL of the savings. The weighting of EULs across sites will be done via site-level savings determined at the meter.

Compensation Calculation

SRP is designed to be a 'break-even' program with the costs of delivering the program equal to the grid benefits it generates. The compensation amount is based on the expected Total System Benefits generated by the energy savings measures installed. SCE will pay ~65% of the estimated TSB in compensation, with the other ~35% withheld to cover the administrative costs of delivering the program. The total amount of this performance-based compensation will ultimately depend on the kWh savings achieved, by costing period, as measured according to the approved M&V Plan and the EUL of the measures installed.

The SRP Implementation Plan and Section 4 of this M&V Plan provides additional detail on the settlement procedures.

3. Settlement Risk, Accuracy Assessment and Population NMEC

Considerations

As noted above, the performance component of SRP compensation and the performance claims for the program will be based on population-level NMEC methods, consistent with Version 2.0 of the *Rulebook for Programs and Projects Based on Normalized Metered Energy Consumptio*n. NMEC methods rely on a comparison of energy consumption at the revenue meter during the pre- and post-intervention period. Regression models with weather and time variables help to explain variability in energy consumption and isolate the effect of the intervention. Adding a comparison group in population NMEC regressions improves the accuracy of the savings estimates as it can control for non-weather related exogeneous changes in energy use.

Population Definitions and Other Upfront Considerations

For SRP, a group or settlement group of projects will be defined as all projects completed by a given TradePro during a calendar quarter. The projects need not come from the same sector, climate zone, or industry type because each participants' modeling will be done independently from one another.

Table 4, below, indicates the relationship between sites, settlement groups and performance estimate risks associated with population NMEC in the residential sector. Table 5 shows the same information for the commercial sector. The table values represent the relative precision, or the expected margin of error divided by the effect size. This metric is referred to as Fractional Savings Uncertainty (FSU) in the NMEC rulebook. A group expected to save 5,000 MWh with a margin of error of ±3,000 MWh would have a margin of error of ±60% and a 95% confidence interval that the measured savings would fall between 2,000 MWh and 8,000 MWh. From a TradePro's perspective that means the performance payment amount can potentially vary from 40% to 160% of the actual value of the savings delivered due to measurement error. Values can be color coded to ensure correct interpretation:

- **Green** cells indicate limited settlement risk (error not more than half of the effect size)
- **Yellow** indicates moderate risk (error is no more than the effect size)

- **Red** indicates high measurement risk (expected margin of error is greater than the effect size and may not be detected at all via population NMEC methods).

Table 4: Settlement Risk as a Function of Effect Size and Population Size (Residential)

Number of Sites	3% Effect	5% Effect	10% Effect	15% Effect
5	676%	406%	203%	135%
10	499%	299%	150%	100%
25	320%	192%	96%	64%
50	224%	134%	67%	45%
100	156%	94%	47%	31%
250	96%	58%	29%	19%
500	48%	29%	15%	10%

Table 5: Settlement Risk as a Function of Effect Size and Population Size (Commercial)

Number of Sites	3% Effect	5% Effect	10% Effect	15% Effect
5	614%	369%	184%	123%
10	672%	403%	202%	134%
25	508%	305%	152%	102%
50	288%	173%	87%	58%
100	191%	115%	57%	38%
250	104%	62%	31%	21%
500	79%	48%	24%	16%

This type of lookup table will convey the settlement risk associated with a given number of SRP projects and expected percent savings that will be communicated to the market. We plan to make clear to TradePros that the measurement and settlement uncertainty is symmetric, meaning the likelihood of over- or under-compensation is the same. The values in Table 4 are based on bootstrapped standard errors using SCE residential AMI data for approximately 40,000 participants in the Smart Energy Program on non-DR days.

Savings estimates for the program-level results for SRP have the benefit of aggregation across both TradePros and groups. Although the ultimate number of participants, average effect size, and distribution across sectors cannot be known in advance, the NMEC Rulebook requires programs evaluated via population NMEC to have no more than 25% Fractional Savings Uncertainty at the 90% confidence level. For both the commercial and residential sectors, this metric is achieved with an average effect size of 10% and a total of 500 participants. Our testing procedures suggest this program-level requirement will be met if SCE achieves projected levels of participation, since increasing the number of participants to 1,000 would ultimately allow for a smaller average effect size. The primary uncertainty concerns for SRP lie in the precision of estimates for subgroups (TradePro groups) as the benefits of aggregation to the program level are little consolation to a TradePro who perceives their settlement calculations represent underpayment for the savings delivered.

Figure 3 and 4, below, focus exclusively on the peak period uncertainty component of the settlement uncertainty since performance during the summer peak and net peak periods are expected to represent a large portion of the final settlement payment. Figure 3 shows the distribution of errors in peak period average hourly impacts for various group sizes in the residential sector across modeling methods. Figure 4 shows the distribution of percent error in peak period average hourly impacts for various group sizes in the commercial sector across segmentation strategies. Both figures depict the same trend: the average peak and net peak performance estimate for groups made up of a small number of sites, which will be far less precise than a group made up of several hundred sites.

Figure 3: Reduced Uncertainty with Increased Number of Sites (Residential)

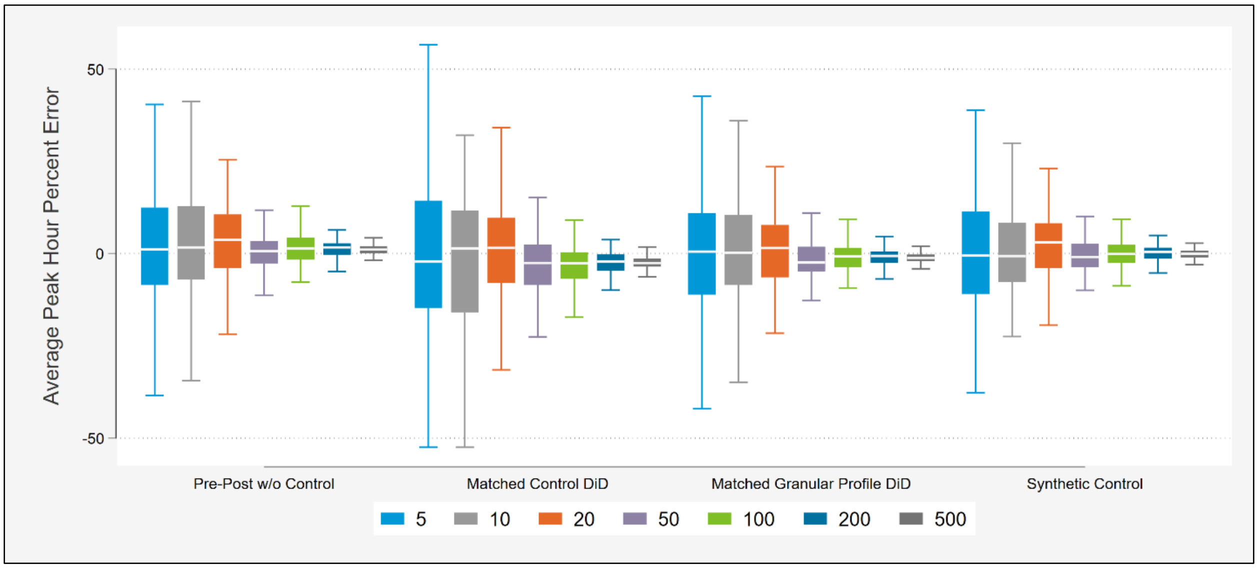
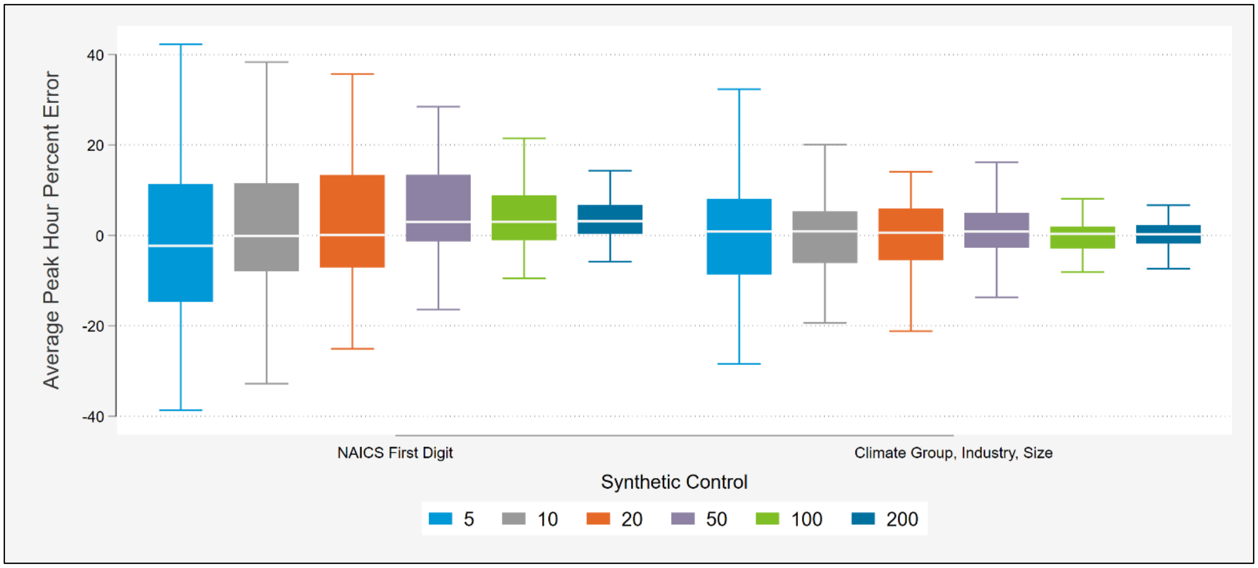


Figure 4: Reduced Uncertainty with Increased Number of Sites (Commercial)



Basis for Proposing the Recommended Analytical Method

Using both residential and commercial customer data as noted below, SCE and its M&V contractors, kW Engineering and Demand Side Analytics, conducted an accuracy assessment of likely population NMEC methods to ensure that the methods chosen (as described in the Analytical Methods section) can accurately quantify impacts for the proposed program. The key questions that were addressed as part of the assessment included:

- Which method is least biased overall?
- Which method has the smallest distribution of error across customers?
- Which method is most precise for individual customers at peak and net peak hours?
- Which method has the smallest distribution of aggregate error for a group?

The accuracy assessment's primary goal was to assess various population NMEC methods and segmentation strategies. In this framework, which is graphically depicted below, pseudo-participants were selected, and their energy consumption was estimated using the candidate models. (Pseudo-participants are customers who are selected to look like likely SRP participants but have not had any energy efficiency intervention during the evaluation period).

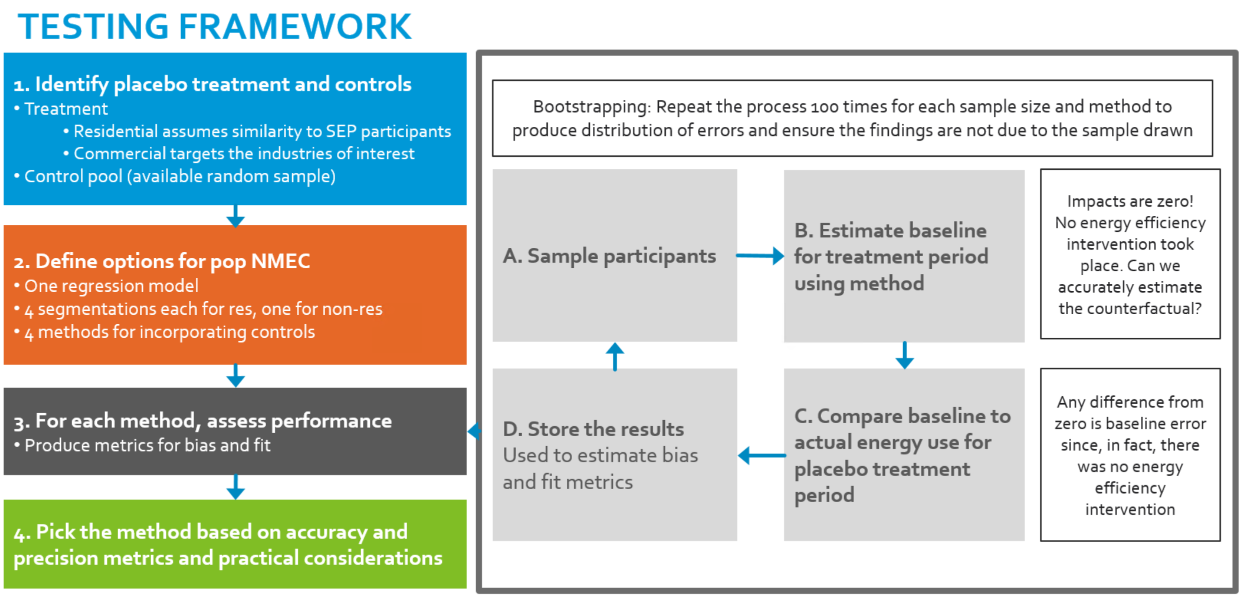
The lack of a true intervention is critical for modeling purposes because that allows for comparison between the estimated baseline and the true observed consumption during the post-treatment period. Instead of quantifying impacts in the post-treatment period, this approach quantifies the baseline method error. This M&V plan provides throughout an overview of our approaches as well as detailed discussions, including an attachment that presents slides containing substantial detail on the methods and segmentation tested.

In addition to these statistics of interest, qualitative results were also important in decision making. Specifically, the final proposed method should be straightforward to compute, transparent to stakeholders, and applicable to a wide range of customer segments.

Framework for Accuracy Assessment

The overall approach for an accuracy assessment is shown in Figure 5, below. The pseudo-participant population is identified up front, as are the models and segmentation strategies to be tested. Each model is then run for each pseudo-participant, where the regression on hourly consumption is estimated on a full year of pre-treatment data and then predicted for the post-treatment year. Because these are pseudo-participants, there is no expected program impact. Any difference between the baseline and observed consumption in the reporting period is modeling error. Once errors have been quantified for each participant, a bootstrapping exercise can be conducted to quantify errors across groups of different customer sizes. Finally, errors are aggregated in a variety of ways to answer the questions listed above.

Figure 5: Accuracy Assessment Framework



The assessment's quantification of errors determined which approaches are appropriate. In general, accuracy assessments tend to quantify performance across two main dimensions: accuracy and precision. Accuracy refers to the tendency of a method to overstate or understate performance on average and is typically summarized by a percent error value that can represent the following:

- **Per-participant error:** By how much does the baseline overstate or understate the true consumption in the reporting period?
- **Population error:** After aggregating all participants observed and baseline consumption together, by how much does the baseline overstate or understate the true consumption in the reporting period?
- **Group error:** The bootstrap procedure randomly pulls samples of participants from the overall participant population. The bootstrap pulls 100 iterations of each sample size, which replicates how each method accounts for differences in participant mix. The error at this level is calculated after aggregating the observed and baseline consumption from all sampled participants in that sample size and iteration, and then assessing by how much the baseline over or understates the true consumption in the reporting period.

The other dimension of interest is precision – the ability of the model to accurately predict consumption in a given hour. This metric, typically expressed as the root mean squared error across hours, can be computed at the participant level, population level, or group level. A summary of these metrics and their implications for the SRP program are shown in Table 6.

Table 6: Metrics of Accuracy and Their Implication for Program Performance

Question	Metric	Interpretation	Implications for Program Operations
Which method is least biased overall?	% Error	The average error for the entire pseudo-population. The tendency of the method to over- or under-predict the true value	The least biased method should minimize the chances of SCE over- or under-paying for the grid benefits generated by the program.
Which method has the smallest distribution of error across customers?	% Error per participant (min, 10th percentile, 25th percentile, median, 75th …	The average error for each pseudo-participant. The distribution summarizes how well each method can model disparate participants	Minimizes the likelihood of any given cohort receiving a spurious payment.
Which method is most precise for individual customers at peak and net peak hours?	Participant-level RMSE across periods of interest	Represents the average noise of each model in a typical hour for each pseudo-participant. A measure of how well a model can consistently mimic observed loads	Minimizes the likelihood of any given cohort receiving a spurious payment.
Which method has the smallest distribution of aggregate error for a cohort?	Distribution of aggregate cohort error across bootstrapped samples	Represents how precise a method can be when accounting for different sample sizes and cohort makeup. This simulates the effect of different samples – for a given set of participants, how much error should be expected?	Assesses how robust the method is to small sample sizes, which translates to settlement risk for the implementers.

Pseudo-Participants, Models, and Segmentation Tested

Because actual participants in the SRP program are not likely to be a perfectly random sample of the population, it is important to select pseudo-participants that represent the likely groups of program participants when conducting an accuracy assessment. For example, TradePros may target specific segments of the residential and commercial populations for participation, and it will be important to incorporate this selection effect into the accuracy assessment. While in practice it is impossible to know who will participate in any program prior to its commencement, the assessment was conducted on groups of customers in SCE's territory that are likely to be targeted based on the current understanding of the program:

- **Residential:** Pseudo-participants were selected to mimic the distribution of customers in SCE's Smart Energy Program, a residential smart thermostat demand response program. For this analysis, 1,000 pseudo-participants were sampled to test the modeling approaches and segmentation strategies.
- **Commercial:** Pseudo-participants were selected by over-sampling large customers and customers in segments identified by the program team as being likely targets based on NAICS codes. Testing for this sector fell into two stages: (1) which model frameworks performed best and; (2) did the proposed segmentation and matching perform acceptably given the recommended framework:
  - **Modeling:** Given the SCE commercial interval data in hand, 500 pseudo-participants were sampled to test the modeling approaches with a simple segmentation strategy of first digit NAICS code.
  - **Segmentation & Matching**: 2,000 pseudo-participants were sampled to test the segmentation strategies using the synthetic control modeling approach.
- **Timing:** All pseudo-participants were assigned a pseudo-treatment date of October 1, 2020. The baseline period for the analysis ran from October 1, 2019 through September 30, 2020 and the reporting period ran from October 1, 2020 through September 30, 2021.

Modeling Approaches

The models tested for this analysis are described in detail in the accompanying slides. In brief, however, the test included four modeling frameworks, as shown below:

1. **Pre-post:** The regression model, a temperature time-of-week (TOWT) model, run on an individual pseudo-participant with no control group.
2. **Difference-in-differences with individual matched control:** For each pseudo-participant, a single control customer was selected based on Euclidian distance matching[4] within a given segment. The same model as described in Option 1 was then run on the pseudo-participant and the selected matched control, both with the same pseudo-treatment date. The reporting period error of the matched control customer was subtracted from the reporting period error of the pseudo-participant to compute the difference-in-differences (DiD) error.
3. **Difference-in-differences with matched granular profile:** The same procedure as Option 2 was performed on each pseudo-participant, the only difference being that instead of using an individual matched control to do the differencing, a profile of aggregated non-participant consumption was used. This granular profile is constructed by sampling 100 or more non-participants within a given segment and then producing an average load profile (8760) for the sampled non-participants during the baseline and reporting periods. The granular profile was matched to the pseudo-participant based on having the same segmentation characteristics. More detail about granular profile production is provided below.
4. **Synthetic control:** The regression model described in Option 1 was modified to include multiple granular profiles as right-hand-side (RHS) variables. Granular profiles included as part of this "synthetic control" matched the participant based on climate zone and solar status but included all granular profiles for a given segmentation strategy.

Segmentation Strategies

The final parameter tested as part of this assessment is the segmentation strategy. Segmentation is important as it ensures that the comparison groups — however they are incorporated into the final model — correctly account for exogenous changes in consumption as intended. Four segmentation strategies were tested for the residential sector and one was tested on the commercial sector:

- **Residential 1:** Climate zone groups, solar status, quartiles of annual usage within climate zone groups and solar status. The quartiles were selected such that within a climate zone group and solar status, each quartile had 25% of the *premises*.
- **Residential 2:** Climate zone groups, solar status, bins of annual usage within climate zone groups and solar status. The bins were selected such that within a climate zone group and solar status, each quartile had 25% of the *consumption*.

---

[4] Matching was done on annual consumption (1 value), average pre-treatment monthly consumption (12 values), and average summer daily load shape (24 values)

- **Residential 3:** Climate zone groups, solar status, quartiles of summer peak hour usage[5] within climate zone groups and solar status. The quartiles were selected such that within a climate zone group and solar status, each quartile had 25% of the *premises*.
- **Residential 4:** Climate zone groups, solar status, quartiles of maximum demand[6] within climate zone groups and solar status. The quartiles were selected such that within a climate zone group and solar status, each quartile had 25% of the *premises*.
- **Commercial 1 (Modeling test):** First digit of NAICS code. No solar customers were included in this assessment for commercial, due to the relatively low penetration of solar among commercial premises.
- **Commercial 2 (Segmentation test):** Climate zone groups, industry, and bins of customer size (small or medium/large based on annual consumption). No solar customers were included in this assessment for commercial, due to the relatively low penetration of solar among commercial premises.

Commercial Segmentation Feasibility Assessment

As described above, supplemental testing in the commercial sector was performed to validate a more granular segmentation strategy than the first digit of a customer's NAICS code, and to include more customers than the 500 initially tested. Since synthetic control methods performed reasonably well for the commercial sector under the broad segmentation strategy of (1) first digit NAICS code, this method was employed to test the feasibility of an additional segmentation strategy. This second approach tested the feasibility of matching customers to more granular industry types (19 total) rather than the nine available using the first digit of the NAICS code. We tested the feasibility for the more granular segmentation of (2) climate zone group, industry, and customer size, leveraging data in hand.

The results, described below, suggest that the larger-scale test with more granular industry segmentation improved the performance of the synthetic control approach for commercial customers.

Assessment Results and Recommendations

Results for the accuracy assessment described in the prior sections are found in Table 7 to Table 12. These metrics were all conducted out of sample in the pseudo-reporting period — October 1, 2020 through September 30, 2021. A subset of key summary statistics for the accuracy assessment results are presented below and can be interpreted as:

- **Full Population % Error:** The overall percentage by which the baseline over (positive) or under (negative) states the true value. A value of 5% means that the baseline consumption is 5% higher than the observed consumption for the period of interest. This value is computed across the entire pseudo-participant population. While values can range from -100% to 100%, a value close to 0% is preferred.
- **Full Population CVRMSE:** The hourly noise for the model at the aggregate level. This value is computed by aggregating hourly consumption across all pseudo-participants, then calculating the

---

5 Average consumption between 4pm-9pm

6 Defined as the 98th percentile of usage at any point. This percentile was chosen to minimize the incorrect binning of customers based on spurious or outlier values.

hourly-level RMSE, normalized by the observed kWh. While values can range from 0 to 1, a value closer to 0 is preferred.

- **Average Participant RMSE:** The hourly noise for the average participant. This value is computed by calculating the hourly-level RMSE for each pseudo-participant, then taking the average. This value is not normalized by the observed kWh and gives a measure of the average error expected by the model. A value close to 0kWh is preferred.
- **Percentile of Average Participant Error:** The three values in this column are the 5th, 50th, and 95th percentile of customer aggregate error. These values are computed by calculating the average percent error for each pseudo-participant, and then finding the Xth percentile. A value of -20% in the 5th percentile means that 5% of pseudo-participants have an average percent error of less than -20%, while a value of 10% in the 95th percentile means that 95% of pseudo-participants have an average percent error smaller than 10%. The 50th percentile represents the median percent error. While values can range from -100% to 100%, a value close to 0% is preferred.
- **Percentile of Average Group Error:** The three values in this column are the 5th, 50th, and 95th percentile of settlement group aggregate error. These values are computed by aggregating all the consumption in each period for all customers in one of 100 iterations of a sample of 50 pseudo-participants, then calculating the average percent error for each group. A value of -20% in the 5th percentile means that 5% of groups of this sample size have an average percent error of less than -20%, while a value of 10% in the 95th percentile means that 95% of groups of this sample size have an average percent error smaller than 10%. The 50th percentile represents the median percent error. While values can range from -100% to 100%, a value close to 0% is preferred.

Table 7: Residential Accuracy Assessment Results: (1) Quartiles of Annual Usage

	Full Population		Avg. Participant	Percentile of Average Participant Error			Percentile of Average Group Error		
	% Error	CVRMSE	RMSE	5th	50th	95th	5th	50th	95th
Net Peak									
Matched Control DiD	1.65	0.03	1.41	-58.97	1.53	108.35	-8.94	0.83	9.57
Matched Granular Profile DiD	-1.52	0.03	1.00	-40.17	-1.83	71.43	-8.38	-1.81	4.59
Pre-Post w/o Control	1.21	0.08	1.00	-35.61	1.15	77.20	-5.23	1.03	7.23
Synthetic Control	-0.40	0.02	0.98	-35.86	-1.25	73.55	-6.68	-0.58	4.60
Non-Peak									
Matched Control DiD	1.19	0.05	1.04	-107.81	-0.16	95.63	-7.90	0.33	17.27
Matched Granular Profile DiD	-1.25	0.05	0.74	-64.90	-1.91	66.09	-7.56	-0.79	4.39
Pre-Post w/o Control	-0.47	0.22	0.78	-60.77	-2.63	67.71	-6.48	0.14	4.94
Synthetic Control	-0.68	0.04	0.73	-65.54	-1.95	63.36	-6.68	-0.06	4.72
Peak									
Matched Control DiD	-1.05	0.04	1.66	-112.18	-3.04	154.07	-13.33	-2.39	10.43
Matched Granular Profile DiD	-2.42	0.04	1.18	-80.09	-3.54	106.17	-10.80	-3.56	6.97
Pre-Post w/o Control	1.02	0.12	1.20	-74.24	-0.13	116.95	-7.17	0.65	10.25
Synthetic Control	0.02	0.03	1.15	-71.05	-1.92	116.66	-7.52	-0.81	8.27

Table 8: Residential Accuracy Assessment Results: (2) Bins of Annual Usage

	Full Population		Avg. Participant	Percentile of Average Participant Error			Percentile of Average Group Error		
	% Error	CVRMSE	RMSE	5th	50th	95th	5th	50th	95th
Net Peak									
Matched Control DiD	0.43	0.03	1.40	-60.54	0.16	108.35	-9.68	0.20	7.74
Matched Granular Profile DiD	-0.02	0.03	1.00	-36.28	-0.43	77.09	-6.80	-0.34	6.51
Pre-Post w/o Control	1.21	0.08	1.00	-35.61	1.15	77.20	-5.23	1.03	7.23
Synthetic Control	-0.32	0.03	0.98	-35.39	-0.43	72.29	-6.79	-0.71	5.73
Non-Peak									
Matched Control DiD	0.52	0.05	1.05	-108.46	-1.15	93.22	-8.83	-0.16	15.04
Matched Granular Profile DiD	-0.46	0.05	0.74	-60.19	-1.48	64.53	-6.99	-0.11	5.31
Pre-Post w/o Control	-0.47	0.22	0.78	-60.77	-2.63	67.71	-6.48	0.14	4.94
Synthetic Control	-1.39	0.05	0.73	-81.20	-2.17	61.59	-8.01	-1.32	5.40
Peak									
Matched Control DiD	-1.86	0.04	1.66	-138.00	-3.38	157.98	-14.17	-1.68	10.15
Matched Granular Profile DiD	0.28	0.03	1.18	-71.31	-0.71	114.66	-8.57	-0.60	10.37
Pre-Post w/o Control	1.02	0.12	1.20	-74.24	-0.13	116.95	-7.17	0.65	10.25
Synthetic Control	0.31	0.03	1.15	-71.86	-0.82	124.08	-7.56	-0.55	9.40

Table 9: Residential Accuracy Assessment Results: (3) Bins of 4-9pm Usage

	Full Population		Avg. Participant	Percentile of Average Participant Error			Percentile of Average Group Error		
	% Error	CVRMSE	RMSE	5th	50th	95th	5th	50th	95th
Net Peak									
Matched Control DiD	1.43	0.03	1.40	-60.93	0.68	104.26	-7.64	0.41	10.58
Matched Granular Profile DiD	0.32	0.03	1.00	-39.26	-0.82	74.43	-6.11	0.23	7.01
Pre-Post w/o Control	1.21	0.08	1.00	-35.61	1.15	77.20	-5.23	1.03	7.23
Synthetic Control	-0.18	0.03	0.98	-36.82	-1.02	71.20	-6.57	-0.85	5.53
Non-Peak									
Matched Control DiD	-0.20	0.05	1.05	-103.62	-1.85	105.60	-11.29	-0.62	13.74
Matched Granular Profile DiD	-0.17	0.05	0.75	-79.24	-1.52	67.19	-6.32	0.04	5.50
Pre-Post w/o Control	-0.47	0.22	0.78	-60.77	-2.63	67.71	-6.48	0.14	4.94
Synthetic Control	0.36	0.04	0.73	-62.03	-1.55	72.86	-6.45	0.73	6.77
Peak									
Matched Control DiD	-1.37	0.04	1.66	-133.57	-2.79	174.17	-13.95	-2.67	9.38
Matched Granular Profile DiD	0.49	0.04	1.19	-84.61	-3.40	115.25	-7.24	-0.14	9.95
Pre-Post w/o Control	1.02	0.12	1.20	-74.24	-0.13	116.95	-7.17	0.65	10.25
Synthetic Control	-0.21	0.03	1.15	-69.93	-2.49	117.97	-7.89	-0.96	7.60

Table 10: Residential Accuracy Assessment Results: (4) Bins of Maximum Demand

	Full Population		Avg. Participant	Percentile of Average Participant Error			Percentile of Average Group Error		
	% Error	CVRMSE	RMSE	5th	50th	95th	5th	50th	95th
Net Peak									
Matched Control DiD	-0.17	0.03	1.41	-58.40	-0.38	105.66	-9.86	-1.35	7.79
Matched Granular Profile DiD	-0.65	0.03	0.99	-36.41	-0.95	73.05	-7.32	-0.79	5.52
Pre-Post w/o Control	1.21	0.08	1.00	-35.61	1.15	77.20	-5.23	1.03	7.23
Synthetic Control	-0.28	0.02	0.98	-35.83	-0.48	71.67	-6.48	-1.13	5.34
Non-Peak									
Matched Control DiD	0.85	0.05	1.05	-117.63	-1.48	92.10	-9.64	-0.17	12.90
Matched Granular Profile DiD	-0.45	0.05	0.74	-63.92	-1.84	69.71	-7.00	-0.18	5.34
Pre-Post w/o Control	-0.47	0.22	0.78	-60.77	-2.63	67.71	-6.48	0.14	4.94
Synthetic Control	-0.69	0.04	0.73	-67.30	-1.52	69.23	-7.11	-0.37	5.24
Peak									
Matched Control DiD	-2.44	0.04	1.67	-129.61	-3.04	157.03	-15.63	-2.56	10.36
Matched Granular Profile DiD	-1.18	0.04	1.17	-73.22	-1.50	114.25	-10.02	-2.45	8.38
Pre-Post w/o Control	1.02	0.12	1.20	-74.24	-0.13	116.95	-7.17	0.65	10.25
Synthetic Control	-0.13	0.03	1.15	-72.64	-1.64	108.91	-7.41	-0.93	8.82

Table 11: Commercial Accuracy Assessment Results: (1) NAICS Code Matching

	Full Population		Avg. Participant	Percentile of Average Participant Error			Percentile of Average Group Error		
	% Error	CVRMSE	RMSE	5th	50th	95th	5th	50th	95th
Net Peak									
Matched Control DiD	-0.73	0.02	125.34	-92.37	-1.22	280.28	-18.02	4.09	21.47
Matched Granular Profile DiD	1.30	0.02	98.31	-202.08	3.68	579.64	-8.86	0.23	22.37
Pre-Post w/o Control	4.50	0.05	86.71	-46.16	-2.01	100.59	-6.44	4.03	23.75
Synthetic Control	1.41	0.02	100.89	-83.89	-3.24	96.95	-11.23	0.94	25.30
Non-Peak									
Matched Control DiD	0.16	0.02	115.85	-82.68	0.36	132.31	-11.25	3.58	16.43
Matched Granular Profile DiD	1.82	0.04	99.02	-316.76	-3.39	115.79	-6.73	1.31	18.70
Pre-Post w/o Control	7.75	0.10	90.85	-22.86	2.00	108.98	-0.68	7.95	22.86
Synthetic Control	-0.92	0.04	98.94	-46.40	-3.37	83.98	-10.13	0.08	16.15
Peak									
Matched Control DiD	0.09	0.02	139.76	-86.69	0.39	288.69	-17.96	4.22	22.67
Matched Granular Profile DiD	2.34	0.03	108.51	-191.80	4.56	343.52	-8.77	1.82	21.49
Pre-Post w/o Control	5.32	0.06	96.77	-44.94	-2.84	98.13	-6.44	5.33	23.03
Synthetic Control	3.57	0.04	110.91	-70.44	-1.83	117.85	-10.76	2.98	25.46

Table 12: Commercial Accuracy Assessment Results[7]: (2) Climate Zone Group, Industry, and Customer Size

	Full Population		Avg. Participant	Percentile of Average Participant Error			Percentile of Average Group Error		
	% Error	CVRMSE	RMSE	5th	50th	95th	5th	50th	95th
Net Peak									
Synthetic Control	0.15	0.01	2.30	-56.27	1.00	106.86	-10.98	0.86	11.23
Non-Peak									
Synthetic Control	-1.86	0.02	2.24	-37.14	-0.63	41.88	-8.79	-1.28	6.00
Peak									
Synthetic Control	0.06	0.01	2.52	-52.55	1.72	90.73	-6.01	0.23	10.14

7 As discussed above, this second-stage feasibility assessment only included the recommended approach from the commercial modeling assessment – a synthetic control approach with multiple granular profiles mapped to each participant.

Several observations emerge from this assessment. In brief:

1. Adding a control group significantly improves precision in essentially all cases, relative to a pre-post model
2. For the residential sector, modeling with synthetic controls was found to be the best option
3. Segmentation using quartiles of annual usage (option 1) for residential performed the best across dimensions of interest
4. The ability to test commercial segmentation and profiles was initially limited due to time and sample constraints, which prompted us to use additional IOU data for supplemental feasibility testing
5. Synthetic control methods performed reasonably well for the initial segmentation testing in the commercial sector. Through the additional feasibility testing, we find they perform better given access to more granular profiles and a larger and more representative commercial pseudo-population.

Based on these findings, the recommendations for population NMEC methods to be implemented for the SRP program are:

1. Use a common method for residential and non-residential sectors — that is, synthetic controls.
2. Rely on segmentation of solar status, climate zone groups, and quartiles of annual consumption for the residential sector. Although NEM sites are not eligible for SRP initially, we have elected to preserve NEM profiles in case eligibility rules are modified.
3. Rely on segmentation of climate zone groups, industry, and customer size for the non-residential sector.
4. Rely on a back-up method for non-residential as follows:
   a. Produce the results using synthetic controls.
   b. Assess the accuracy of synthetic controls at the site level.
   c. If the site-level CVRMSE is above 0.5, run the backup option: a matched control group with difference-in-differences.
   d. If the site level CVRMSE for the back-up option is lower, use those results instead.

4. Analytical Methods

Modeling of participant baselines in the reporting period will be accomplished using hourly AMI meter reads for the participant’s site, weather data, and the incorporation of a comparison group that will account for exogeneous changes in energy consumption. The specific steps to produce estimates of program energy savings are as follows:

1. For each participant in the group, ensure that a full year of baseline and reporting period hourly consumption interval data is available, along with hourly weather data. Each participant should also have one year of pre-baseline data, referred to in this section as the testing period, to ensure model validity. The testing, baseline, and reporting periods together comprise the analysis period.
2. Remove any data in a blackout period in between the start and end of the measure installation. The baseline period is defined as the 365 days prior to the installation start and the reporting period is defined as the 365 days after the installation end.

3. Construct the regression variables. These are defined in more detail below, but include seasonal indicators, hour-of-week indicators, temperature characteristics and control customer consumption as described in the prior section.
4. For each participant, estimate the regression model during the baseline period. This model is a seasonal time-of-week and temperature (TOWT) model. This model includes variables for each hour of the week, a temperature spline, and one or more granular profiles that act as a synthetic control group.
5. For each participant, predict usage during the reporting period. This is the counterfactual consumption: a representation of what the participant would have done if they had not enrolled in the program.
6. Remove customers based on screens based on model accuracy during the testing period.
7. The difference between the counterfactual and the observed usage is the impact of the program
8. Aggregate the results to the annual total kWh savings, the total kWh savings in the peak period and the net peak period.

The regression model used for determining participant performance will be run at the individual participant premise level. Eligible participants must have at least one full year of interval data available prior to the installation of the program measures, and program savings are estimated until the end of the first-year post-installation. Participants must have corresponding hourly weather data for their premises, which can be mapped to the appropriate CALMAC weather station data using the participant's zip code. Each participant in the group will be modeled individually.

Program savings will be aggregated to the TradePro group level. A group is defined as all participants that enroll in the program within a given calendar quarter through a specific implementer (TradePro). The procedure described in the following paragraphs define how savings are estimated and reported for a specific group, though in practice a TradePro may have multiple groups of participants active at any given point.

The regression specification used for estimating participant impact is based on the time of week temperature (TOWT) model developed by LBNL[8]. There are five components to the regression, which is run on the hourly participant consumption data:

1. The regression constant term, representing the average base consumption for the participant.
2. Hour-of-week fixed effects. There are 7 x 24 = 168 dummy variables that capture deviations from the base consumption in each hour of the week.
3. Temperature spline. Between one and seven bins of temperature, with cut points for each temperature bin set algorithmically to ensure sufficient coverage.

[8] Quantifying Changes in Building Electricity Use, with Application to Demand Response Johanna L. Mathieu, Phillip N. Price, Sila Kiliccote, Mary Ann Piette Lawrence Berkeley National Laboratory April 2011

4. Granular profiles[9]. These are average hourly consumption profiles for a sample of non-participants in similar segments to the participant. The role of the granular profile is to capture information about non-weather characteristics of each date-hour that may influence participant energy consumption. Excluding these granular profiles from the model result in a simple pre-post model.
5. The error term.

The exact specification is shown in Equation 1:

Equation 1: Seasonal Time of Week Temperature Model

$$kWh_{p,t} = \alpha_p + \sum_{i=1}^{168}(\beta_i * I_{i,t}) + \sum_{b=1}^{b=[2,7]}(\gamma_b * B_{b,t}) + \sum_{g=0}^{n}(\delta_g * GP_{g,t}) + \varepsilon_{p,t}$$

Table 13: Definition of Equation Terms

Symbol	Interpretation
$kWh_{p,t}$	The observed kWh consumption for participant p in date-hour t
α_p	The constant for participant p
β_i	The coefficient representing the base energy consumption for hour-of-week i, above or below the participant average
$I_{i,t}$	A dummy variable for each hour-of-week i. Equal to 1 when date-hour t is in that hour-of-week, and 0 otherwise
γ_b	The coefficient representing the marginal consumption associated with a one-degree change in outdoor temperature for temperature bin b
$B_{b,t}$	The value of temperature bin b. The construction of temperature bins is described in more detail below.
δ_g	The coefficient representing the marginal effect of one kWh change in the control group granular profile g.
$GP_{g,t}$	The average consumption of the granular control group profile g in date-hour t.
$\varepsilon_{p,t}$	The error term for participant p in date-hour t

The temperature spline is comprised of between one and seven temperature bins that relate outside air temperature to participant consumption. A spline model splits temperature from a single value into ordered bins that correspond to the degrees Fahrenheit (or Celsius) that fall in that bin. As examples, the temperatures in Table 14, below, can be represented as temperature bins in the following manner:

[9] For the residential sector, the granular profiles included in the regression are the four profiles of customers in each quartile of annual consumption. Participants are assigned granular profiles that match the participant's climate zone. For example, participants in CZ 6 are only assigned the four profiles of customers that are also in CZ 6. The commercial granular profiles are assigned to those with similar industry segmentation and size in a given climate zone.

Table 14: Relationship Between Temperature and Spline Temperature Bins

Temperature	B_1	B_2	B_3	B_4	B_5	B_6	B_7
Condition (F)	**< 30**	**30-45**	**45-55**	**55-65**	**65-75**	**75-90**	**> 90**
25F	25						
47F	30	15	2				
65F	30	15	10	10			
83F	30	15	10	10	10	8	
101F	30	15	10	10	10	15	11

To ensure that the relationship between temperature and consumption can be robustly estimated, there must be sufficient data in each temperature bin. To that effect, the number of bins used in the regression are modified dynamically by algorithmically removing cut points between the bins. The procedure for this pruning is described in further detail in Section 3.9 of the CalTRACK methods[10]. In brief, the procedure involves:

1. Count the number of hours in each temperature bin B_1 through B_7
2. If any of bins B_1 through B_6 have fewer than 20 observations in that range, combine the observations in that bin with the next highest bin:
    a. For example, if bin B_2 (30-45F) had 17 observations and bin B_3 (45-55F) has 30 observations, combine B_2 and B_3 to create one bin from 30-55F with 47 observations
3. If B_7 has fewer than 20 observations, combine it with the next lowest bin until the 20-observation criteria is met
4. Continue pruning the bins until each bin contains at least 20 observations.

An example of this pruning procedure is shown in Figure 6, below.

10 http://docs.caltrack.org/en/latest/methods.html

Figure 6: Pruning of Temperature Bins

Starting		Iteration 1		Iteration 2		Termination	
Bin Conditions	Count	Bin Conditions	Count	Bin Conditions	Count	Bin Conditions	Count
< 30	5	<45	35	<45	35	<45	35
30-45	30						
45-55	50	45-55	50	45-55	50	>45	59
55-65	5	55-75	6	>55	9		
65-75	1						
75-90	2	> 75	3				
> 90	1						
Total Bins	7	4		3		2	
Count	94	94		94		94	

The final element in this Seasonal TOWT model are the granular profiles. These represent the average granular (8760) consumption of a group of non-participants. Participants are matched to the correct granular profile(s) based on having similar segmentation, as described in the previous section. The regression will have one or multiple granular profiles added as explanatory (right-hand-side) variables. This approach is called a synthetic control and relies on exploiting the correlations that exist between participant loads and nearby similar customers. These customers experience similar economic conditions and other unobserved conditions that may influence energy use. This correlation does not have to be positive to yield useful information, though in practice it is often easiest to understand the intuition for this approach with positive correlations. For example, if July 4th falls on a Thursday, many residential premises may have altered consumption on Friday, July 5th or even earlier in the week as households take vacation. Including granular profiles of other residential customer segments in the specification will show this change in consumption during the holiday week. Without the inclusion of the granular profiles, this information would not be observable in the model and the observed change in consumption would be misattributed to the effect of program participation.

The weighting of the granular profiles in any given site's regression model of hourly energy consumption will be determined via the regression process. Each model also contains weather and time variables to explain the variation in energy consumption. As an explanatory variable, consider two hypothetical restaurant participants which each leverage a granular profile of hundreds of non-participating restaurants. One restaurant's loads patterns might be highly correlated with the patterns of the granular profile in its pre-intervention period. In this case, the coefficients determined via regression would rely heavily on the explanatory power of the granular profile in the predictive model. If the second restaurant's loads were not explained well by the hourly loads of the granular profile, the coefficients determined via regression would rely more heavily on the time-of-week and weather variables in the predictive model. The specific mathematical relationship between the participating site's loads and the granular profile is determined individually via regression. This is analogous to the way each site's model includes weather variables, but the specific relationship between electric demand and weather is different for each participant.

The regression model is estimated independently for each season[11] in the training period, and then predicted for that season in the reporting period. The predicted hourly consumption in the reporting period is called the counterfactual consumption. These values represent what the consumption would have been had the premises not participated in SRP. Savings in the reporting period are simple summations of the hourly impacts by period of interest.

Because all participants must have at least one year of pre-installation data and settlement occurs at the end of the first-year post-installation, all participants will have the same number of peak, net peak and non-peak observations in the reporting period. Total kWh savings in each period can simply be summed across participants and hours in that group-period. Average participant peak or net peak kW savings can be computed by averaging the hourly kWh readings across all hours in the respective peak period.

Data Preparation

SCE will establish a monthly data transfer procedure with the M&V team which includes both project/measure package data from implementation and AMI data for modeling. Prior to modeling, DSA will prepare the participant load data for analysis according to the data structure required to implement the selected modeling approach:

- **Weather Station**: Merge hourly weather data from one of the CALMAC weather stations. [12] Weather station mapping and data sufficiency will follow Section 2.4.1 of the CalTRACK Technical Appendix.
- **Define the "blackout" period, Baseline, and Reporting periods**: Using the project completion data collected during implementation, create a buffer period in either direction that is not part of the baseline or performance period. The 365 days prior to the beginning of the buffer are the baseline period. The 365 days following the buffer are the reporting period.
- **Merge the granular control profiles**: Based on the characteristics of the participant, merge one or more granular profiles by date and hour.

There are several important mechanical considerations regarding the granular profiles that we believe are important to call out in this plan.

- The definition and composition of the profiles will be defined in advance, but the profiles themselves must be maintained as new meter data becomes available.
  - Because the baseline period model is fitted with the granular profile as an explanatory variable, the prediction of counterfactual energy consumption in the performance period requires the profile data be available for the performance period.
- DSA will document the profile definitions and which accounts make up each profile before compensation for the first round of project groups are issued. Hourly AMI data for members of the synthetic control group will be transferred along with participant load data on a regular cadence.

[11] Seasons are defined as: Summer – June through September. Winter – December through March. Shoulder – April, May, October, November.

12 **http://calmac.org/weather.asp**

- The SCE accounts that make up the synthetic control group profiles will need to be monitored for SRP participation, account closure, or other significant changes such as adoption of solar or batteries.
    - We plan to select alternate members for each profile to use as replacements when these types of changes occur. The selection of alternate profile candidates will be based on substituting 1:1 the existing granular profile candidate based on the size of the customer and the segmentation (customer size, industry, and climate zone as described in the customer segmentation strategy).

Comparison Group Segmentation

SCE and Demand Side Analytics will use synthetic controls on the right-hand side to model SRP impacts for both residential commercial participants. The residential segmentation strategy for developing granular profiles has been tested and finalized and is shown in Table 15, below. There will be 32 distinct segments (4 x 2 x 4 combinations), each composed of a minimum 500 non-participant accounts.

Table 15: Residential Granular Profile Segments

Climate Zone Group	Solar Status	Annual Consumption Quartile
Coastal (CZ5 and CZ6)	NEM	Bottom 25%
Mild (CZ8 and CZ9)	Non-NEM	25%-50%
Hot (CZ10)		50%-75%
Very Hot (CZ13-CZ16)		Top 25%

Table 16, below, shows the segmentation approach for the commercial sector. DSA has requested interval data to SCE to develop and operationalize these profiles.

Table 16: Commercial Granular Profile Segments

Climate Zone Group	Size	Industry Type
Coastal (CZ5 and CZ6)	Small	Agriculture and Pumping
Mild (CZ8 and CZ9)	Medium/Large	Automotive and Repair
Hot (CZ10, CZ13-CZ16)		Education
		Fitness and Esthetic Services
		Full-Service Restaurants and Bars
		Gas Stations and Convenience Stores
		Government-Institutional
		Grocery
		Health
		Limited-Service Restaurants
		Lodging and Entertainment
		Manufacturing
		Miscellaneous
		Office
		Property Management
		Religious
		Retail
		Transportation, Communications, and Utilities
		Warehouse and Wholesale

While all combinations of the dimensions shown in Table 16 suggests a total of 114 distinct profiles, not all industry types require, or can support, six distinct profiles due to limited weather sensitivity or variation in size. The final number[13] of non-residential granular profiles will be 102.

Any given participant's regression model will not include all profiles on the right-hand side. We will use a subset of these distinct profiles for each site, based on:

1. **Same Industry, Cross-Climate Zone and Size:** A "Small, Coastal Lodging" site would have all other Lodging profiles included on the right-hand side (up to 6 profiles for the other climate zones and customer sizes)
2. **Related Industry, Same Climate Zone and Size:** Additional right-hand side variables will be based on pre-defined industry correlations within a climate zone and customer

[13] The industries that will be consolidated across customer size are: Agriculture & Pumping, Gas Stations and Convenience Stores, Personal Care Services, and Warehouse and Wholesale. These industries will have 3 available granular profiles, one for each climate zone that is comprised of customers of all sizes.

size. For example, consumption in Lodging, Gas Stations, and Full-Service Restaurants within a climate zone are expected to be highly correlated as these industries are related to each other.

The choice of industries shown in Table 16 above represent segmentation that performed well in the synthetic control method; that is, leveraging information about consumption patterns in a given industry to predict consumption in another. However, *it is important to note that not all of the industry segments listed above necessarily represent target or eligible SRP participants.* As discussed in the methodology section described above, granular non-participant profiles rely on correlations in consumption patterns between industries to better predict participant consumption in the performance period. For that reason, industry granular profiles of ineligible SRP segments, for example, Agriculture and Pumping, may still be used in the estimation of eligible SRP participant savings.

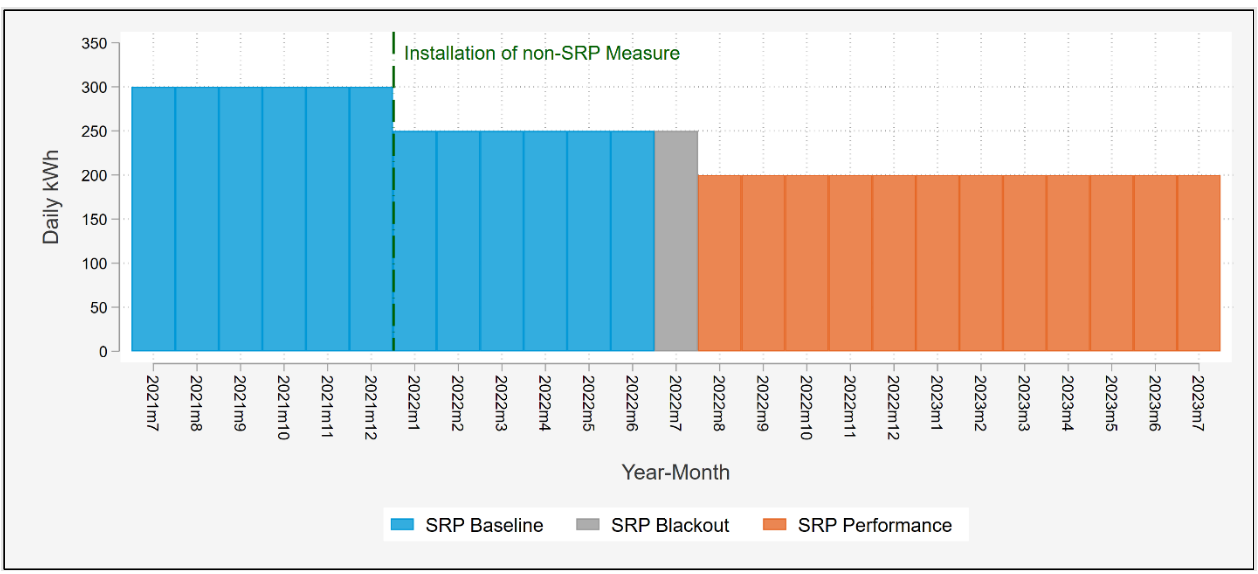
Dual Participation on other EE and DR Programs and Incremental Savings

SRP is designed to deliver incremental savings to SCE's existing portfolio of energy efficiency and demand response programs. The program design centers on compensating projects for the grid value these SRP projects deliver. This requires processes to prevent over-payment or under-payment due to dual participation. Along with other project completion details, SCE will pass the NMEC modeling team information on current demand response program enrollments and any energy efficiency measures completed in the twelve months prior to SRP participation. Participants will not be eligible for any new energy efficiency rebates during their twelve-month performance period.

Recent Energy Efficiency Participation and Incremental Savings

The threat associated with allowing customers with recently completed EE projects to participate in SRP is that the regression model of consumption will overstate the counterfactual if it is estimated on data prior to the non-SRP measure installation. The accounts selected to make up the synthetic control group profiles will not have prior EE participation by design. Consider the simplified example shown below in Figure 7, where a hypothetical participant completes an EE project outside of SRP in January 2022 and an SRP project in July 2022.

Figure 7: Recent EE Participation Example



Using the twelve months of data prior to SRP participation, we would estimate a baseline of 275 kWh per day. However, this is biased upward by the inclusion of six months of data from prior to the installation of the non-SRP measure which saves 50 kWh per day or 18,250 kWh annually. In this simplified example, the appropriate baseline for the site is 250 kWh per day. SCE plans to address this bias via a downward adjustment to the predicted baseline using the following steps. The procedure mirrors handling of a non-routine event in the baseline period for site-level NMEC, as follows:

1. Determine whether each day in the performance period requires adjustment. In the example shown in Figure 7, days in July, August, September, October, November, and December require adjustment.
2. Determine the 8760 load shape of the non-SRP measures based on DEER or eTRM profiles and spread the claimed kWh savings over the year.
3. Multiply the 8760 load shape from Step 2 by the adjustment flag (0,1) to arrive at the hourly adjustment.
4. Subtract the calculated adjustment from Step 3 from the predicted baseline determined via NMEC.
5. Compute hourly impacts as the difference between the adjusted baseline and metered consumption during the performance period.

Enrollment in SCE Demand Response Programs and Incremental Savings

SCE offers full-suite, supply-side demand response programs for both the residential and commercial sectors. The addition of the statewide Emergency Load Reduction Program (ELRP) in 2022 will further increase DR saturation among potential participants. While it would be cleaner from an M&V standpoint to disallow dual participation in SRP and DR programs, SCE believes this would block off an excessive portion of the market and make it difficult to achieve the targeted participation levels. Our proposed approach to avoid SRP from including DR impacts is as follows:

- Exclude DR event hours from the baseline model and performance period. If DR events begin or end mid-hour exclude the entire hour.
- Also exclude the hour prior to and following DR events to account for pre-cooling, or post-event snapback which lead to DR participants having higher load than they would otherwise.
- Set SRP performance equal to 0.0 kWh during any excluded hour in the performance period.
    - This ensures that sites dually enrolled in SRP and DR cannot receive compensation from the SRP for DR reductions during the summer of 2022 or 2023. This will be communicated clearly to TradePros as it has bearing on the settlement calculation. This approach to avoid double payment likely affects less than 40 hours a year and avoids needless complexity.
- It also means that DR events in the baseline period do not bias participant baseline up or down unfairly.
- It ensures baselines that measure the incremental change in daily peak, net load peak, and off-peak energy use.

Weather Normalization

The analytical methods described above include a series of explanatory variables to capture time and weather effects in the mathematical model of energy consumption. This relationship will be modeled in the baseline period and predicted for the reporting period using the actual reporting period weather conditions. We do not plan to estimate separate regression models for the reporting period and perform parallel predictions against normalized weather conditions when determining claimable savings [14]. This decision is based on four factors:

- The primary objective of the program is to deliver peak and net peak reductions during the summers of 2022 and 2023. Measuring performance and settling with TradePros based on delivered impact during these periods of interest removes a layer of complexity and presents a clearer signal to the market.
- Not running separate models allows for faster reporting. If a separate mathematical model of energy consumption is required for the performance period, it is imperative to wait for adequate coverage of independent variables before estimating impacts. Under our proposed approach, we can measure savings as soon as the performance period begins and show cumulative sums of impacts at regular intervals.
- The difference between predicted savings under the weather conditions in the weather-normalized scenario and actual conditions is unlikely to be materially different. So, running separate models increases costs while likely not increasing benefits.
- We are already "smoothing" the avoided costs used to compute Total System Benefits. The CPUC's Avoided Cost Calculator loads significant capacity value on a small number of hours based on loss of load probabilities. Averaging the avoided costs by month, hour, and business day distributes the value more evenly and mitigates the risk to TradePros and SCE that actual weather conditions during the summers of 2022 and 2023 will be misaligned with ACC assumptions.

Determination of Net Savings

Use of a granular profile as a synthetic control group in the NMEC procedure obviates the need for a separate assessment of free ridership and spillover and directly estimates net savings. The indices of non-participating accounts capture both exogenous effects, like the COVID-19 pandemic, and the effect of energy efficiency purchases and behaviors SRP participants would have taken absent program intervention. If estimates of gross savings are needed for reporting, SCE will divide the net savings by a default population NMEC net-to-gross ratio of 0.85 to reverse engineer gross savings values.

Trade Professional Settlement

The final settlement with program implementers for a group of projects will be equal to ~65% of the Total System Benefits generated by the group. Consistent with the intent of the Governor's emergency declaration this compensation structure rewards TradePros heavily for savings generated during the peak and net peak periods. As a result, the load shape of the NMEC-based savings estimate is a key driver of the final compensation amount. Figure 8, below, visualizes the 8760 load shapes for two measures in the DEER catalog. The residential HVAC measure savings are highly concentrated in the summer months, while the non-residential indoor lighting shape has limited seasonality and generally follows business hours on weekdays.

[14] We believe that this approach provides TradePros and SCE with accurate and timely estimates of savings based on observed weather and can additionally produce weather-normalized claimable savings at the end of the program's operations. If required to estimate weather-normalized payable savings, SCE may elect to do so.

Figure 8: DEER 8760 Load Shapes for Two Sample Measures

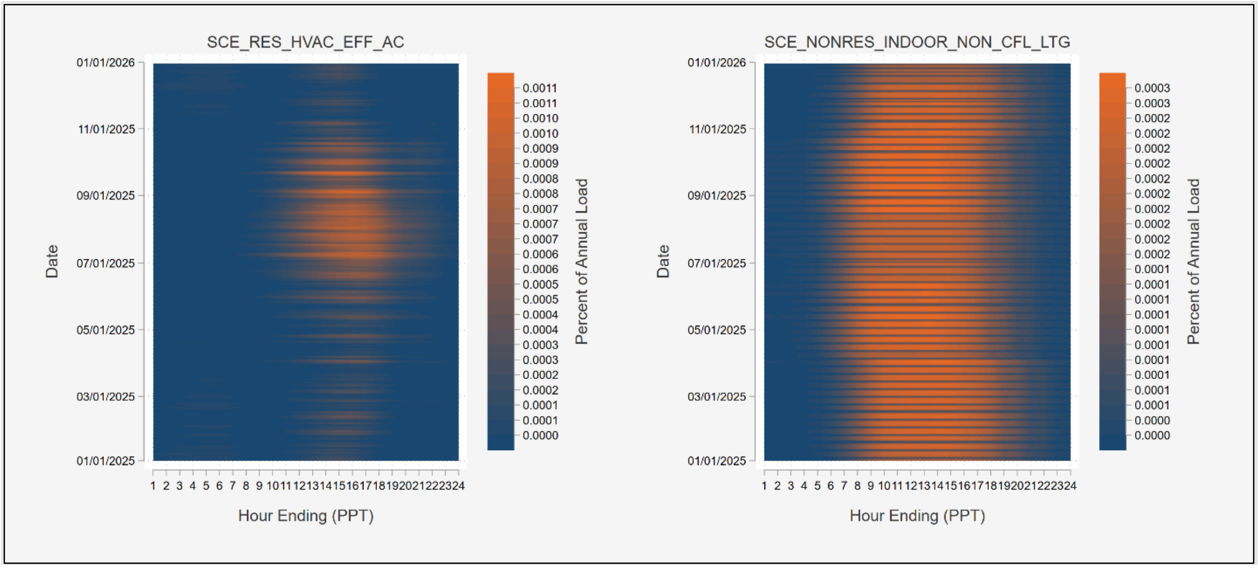


Table 17, below, shows the annual kWh savings by period for a hypothetical commercial lighting measure that saves 1 MWh annually using the load shape visualized on the right side of Figure 8, above. The TSB generated is the equal to the present value of the lifetime avoided costs generated using a 7.68% nominal discount rate. The avoided costs in each year of the EUL are the product of the kWh savings in each costing period and avoided costs for the corresponding period. In this example, we use the full 8760 load shape and avoided costs. Less than 5% of the annual savings for the indoor commercial lighting measure occur during the high-value peak and net peak periods.

Table 17: Total System Benefits (2022) by EUL (CZ13) – Non-Res Lighting

Period	Annual kWh Saved	EUL = 2	EUL = 5	EUL = 10	EUL = 15	EUL = 20
Peak	39.2	$15	$40	$85	$122	$154
Net Peak	15.3	$8	$27	$54	$76	$96
Non-Peak	945.4	$125	$259	$505	$713	$902
Total	**1,000**	**$149**	**$327**	**$643**	**$912**	**$1,151**

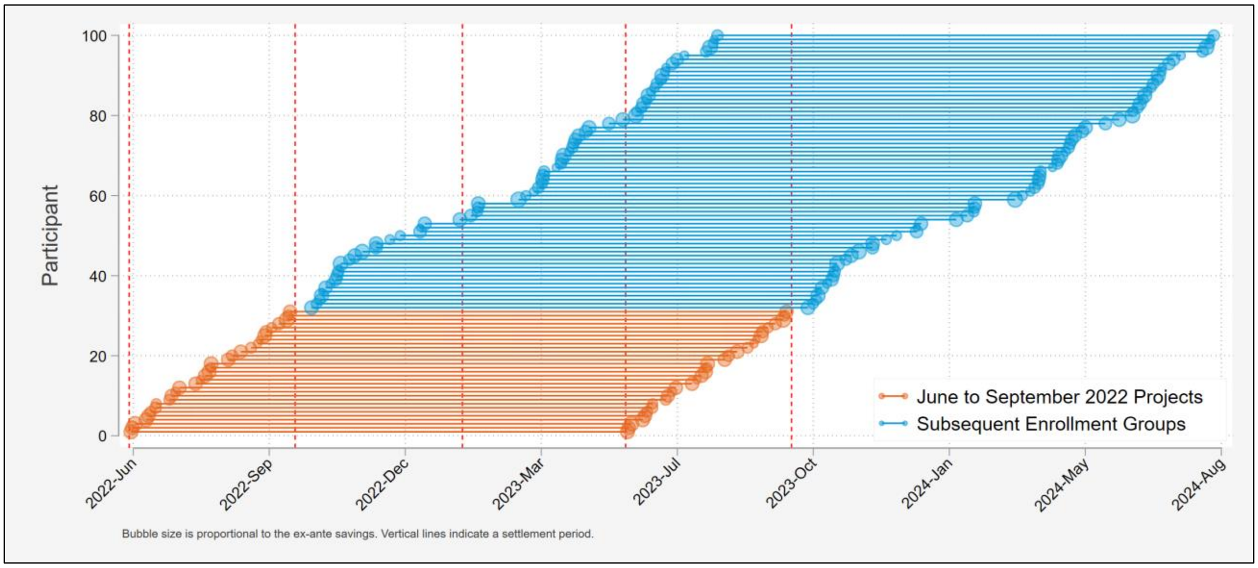
For any measure, TSB will increase with EUL. The growth rate of the avoided costs over time helps to counteract the discount rate, but there is still a diminishing return from each additional year of measure life. Table 18, below, shows the same information for a residential HVAC measure that also saves 1 MWh annually. Almost 20% of the annual savings occur during the high-value peak and net peak periods, so the TSB generated per unit of energy saved is much higher.

Table 18: Total System Benefits (2022) by EUL (CZ13) – Residential HVAC

Period	Annual kWh Saved	EUL = 2	EUL = 5	EUL = 10	EUL = 15	EUL = 20
Peak	147.9	$56	$149	$316	$454	$568
Net Peak	44.3	$20	$72	$149	$214	$269
Non-Peak	807.8	$168	$346	$692	$967	$1,203
Total	**1,000**	**$244**	**$566**	**$1,157**	**$1,635**	**$2,040**

Figure 9 illustrates the back-end performance-based settlement procedure. All completed projects by a given TradePro within a 4-month period (orange) will be analyzed and paid as a group.

Figure 9: Hypothetical Settlement Group



Once the performance estimates are finalized for a given group, SCE will issue compensation for the settlement. Settlement will lag the close of the performance period approximately 60-90 days to allow for validation and transfer of meter data to the modeling team and analysis. Table 19, below, shows the program lifecycle for the hypothetical group of projects from Figure 9. Because projects are installed continuously from June to September 2022, the summer performance period when peak and net peak performance is determined is split between the summer of 2022 and the summer of 2023. Compensation for summer performance, which is likely highest based on the avoided cost structure, are shaded green.

Table 19: Sample Performance Payment Cadence

Period	Group Activity	Performance Payment
June – September 2022	Project Installation and beginning of performance measurement	None
October 2022-January 2023	First full period of performance measurement	Based on population NMEC estimates from June – September 2022
February – May 2023	Second full period of performance measurement	Based on population NMEC estimates from October 2022-January 2023
June – September 2023	Final period of performance measurement	Based on population NMEC estimates from February – May 2023
October 2023-January 2024	None	Based on population NMEC estimates from June – September 2023

Compensation for performance in each period “p” is equal to the total settlement amount earned. The “Total Compensation Earned” term is equal to ~65% of the Total System Benefits demonstrated during the period of analysis.

While SCE and its M&V contractors will estimate 8760 load impacts and the CPUC’s avoided costs are on an 8760 basis, a “collapsed” set of impacts and avoided costs will be used to compute TSB and settle with TradePros. The rationale for using collapsed avoided costs and impacts is to smooth out the avoided costs to mitigate risk to SCE and TradePros. Consider the granular avoided costs for Climate Zone 14 in 2026 shown in Figure 10, below. Value is highly concentrated in a small number of hours which shoulder the generation, transmission, and distribution capacity value.

Figure 10: Load Duration Curve of Avoided Costs – CZ14

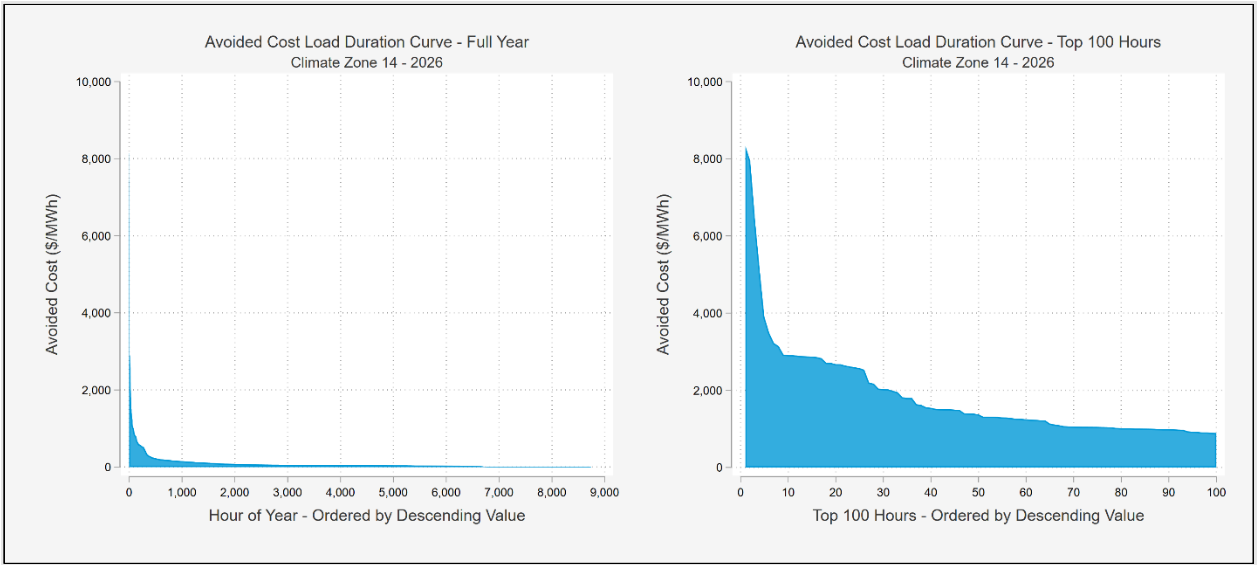
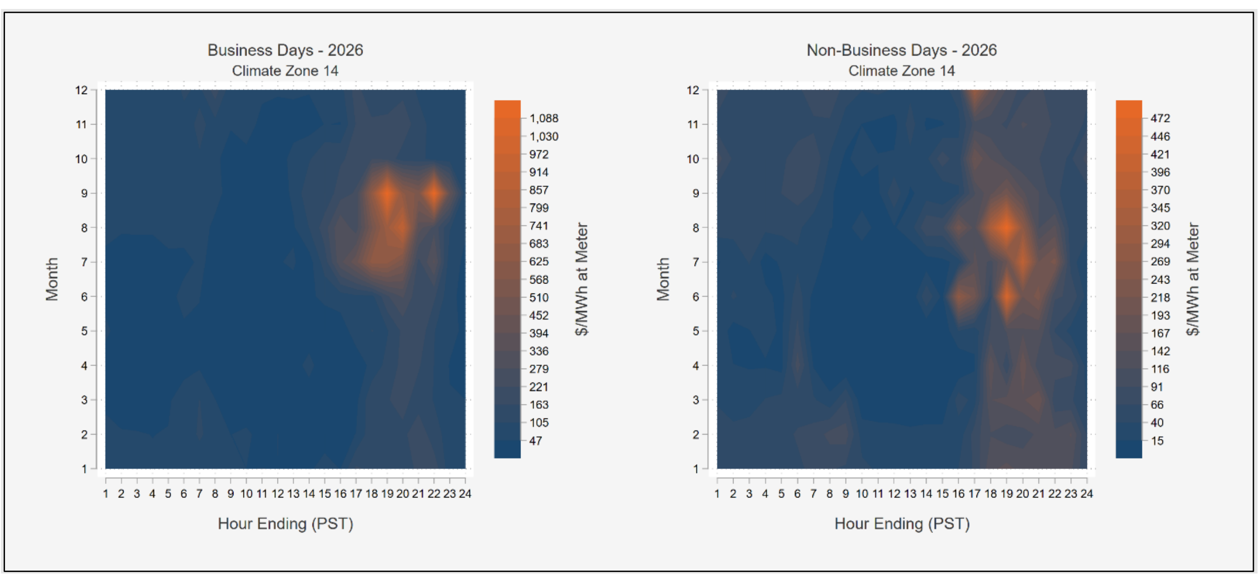


Figure 11, below, shows an example of the granular avoided costs averaged by three dimensions. The result is a profile with 576 values per climate zone and year. Prior to collapsing, SCE moved any generation, transmission, or distribution capacity value assigned to a Sunday or a holiday in the ACC onto Monday-Saturday for consistency with CPUC Resource Adequacy definitions.

1. Month of the year (n=12)
2. Hour of the day (n=24)
3. Separated into business day and non-business day (n=2)

Figure 11: Proposed Avoided Cost Array



This approach limits the value placed on any single hour while still preserving the price signals intended for a Market Access Program.

To-Code Savings

Population NMEC procedures rely on the condition of the participating facility during the 12 months prior to project completion to develop a baseline. To the extent that participants replace equipment that does not meet current codes and standards, those "to code" savings will be credited by SRP. This feature of SRP is expected to unlock some savings potential in the commercial sector and help allow TradePros to ramp up the program quickly by servicing "stranded efficiency" opportunities. SCE does not plan to disaggregate the estimated savings in to-code and above-code savings.

Transparency and Replicability

One of the useful features of using granular profiles rather than individual matched controls to estimate impacts is that the aggregate profiles can be shared without violating California's 15:15 rule regarding data privacy. SCE plans to publish the granular profiles used for SRP settlement at regular intervals. Armed with participant load data, the appropriate weather records, and the necessary granular profiles TradePros, the CPUC and its evaluation contractors should replicate our savings estimates and compensation calculations exactly. SCE will publish the "smoothed" avoided costs used for settlement once the above questions about the 2021 ACC have been resolved. Since comparison group methods, and particularly synthetic controls, are a new method for many stakeholders, DSA will make available applied examples with sample data and code in a Jupyter Notebook format.

5. Data Collection and Reporting

Data collection for purposes of M&V and settlement falls into three primary categories:

1. **Population characteristics and load data for development of the granular control group profiles (one-time).** Our testing and determination of the recommended segmentation and modeling strategy was based on large volumes on AMI data that Demand Side Analytics had available from the load impact evaluations they perform. While the residential testing data set includes hundreds of thousands of SCE customers, by nature it includes a disproportionate share of DR customers and the non-DR participants are a stratified sample designed to mirror the DR participants. The data request needed for selection of the accounts that comprise the final granular profiles for settlement was issued to SCE in May 2022 and fulfilled in June 2022.
2. **Project Completion Information (continuous).** As TradePros complete projects and submit them through the iEnergy platform, SCE collects a robust set of information about the participating customer, the efficient equipment installed, the expected energy savings, the date the work was completed. etc.
   - Along with the project information, customer characteristics, and other meta data associated with initial project completion package, SCE will extract and transfer the last 12 months of hourly AMI data for the new set of participating sites.
   - Because meter data is used to estimate program performance and determine compensation for TradePros, participant-to-meter correspondence is critical to program success. SCE will validate all participant information, service account and contract number mapping.
3. **Ongoing transfers of hourly load data (monthly)**. Includes all accounts that make up the granular profiles as well as all participants that have not reached the end of their 12-month performance period. SCE will establish a regular data transfer process to the

analysis team once AMI data has been processed and finalized. We expect the transfer to the kW/DSA team will occur approximately 45 days after the end of the month. For example, August interval data will be transferred for analysis by October 15th.

As noted earlier, SCE will broadly report in alignment with settlement periods:

Table 2020: Reporting Period Cadence

Period	Group Activity	Reporting Activity
June – September 2022	Project Installation and beginning of performance measurement	None (not sure if this is correct)
October 2022-January 2023	First full period of performance measurement	Population NMEC estimates from June – September 2022
February – May 2023	Second full period of performance measurement	Population NMEC estimates from October 2022-January 2023
June – September 2023	Final period of performance measurement	Population NMEC estimates from February – May 2023
October 2023-January 2024	None	Population NMEC estimates from June – September 2023

SCE and its TradePros will preserve all customer, project and load data for sharing with the CPUC, upon request, for evaluation or other purposes.

CPUC staff and its evaluation contractors (as appropriate) will also have SFTP site to view reporting outputs, except their credentials will provide access to reporting materials for all TradePros. SCE may also build an online dashboard, available to the public, where high-level metrics by climate zone and sector can be viewed. For example:

- Cumulative energy savings at the meter
- Peak and net peak demand savings acquired
- Expected lifetime energy savings and TSB based on completed measurement activities to date, and
- Compensation to date.

SCE plans to publish performance estimates, compensation calculations, and share supporting data with TradePros via a secure FTP site. Each TradePro will have separate credentials and only be able to access their own reporting materials.

The final savings claims will be substantiated by an M&V Report, developed by kW Engineering and Demand Side Analytics, consistent with the methods described in this M&V Plan.

Compensation for performance in each period “p” are equal to the total settlement amount earned. The ‘Total Compensation Earned’ term is equal to ~65% of the Total System Benefits demonstrated during the period of analysis.

The final savings claims will be substantiated by an M&V Report, developed by kW Engineering and Demand Side Analytics, consistent with the methods described in this M&V Plan.

6. M&V Data Requirements

Request	Detail	Purpose/notes
1. Population Sample for Granular Profile Development	Includes characteristics and load data a. Account numbers (SA number, CT number) b. Rates and effective dates of rates for 2020-2023 c. Net metering status, date that NEM status became effective, and installed capacity d. NAICS or SIC industry codes, if applicable e. Zip Code, Climate Zone, SubLAP, and A/B-Bank f. DR enrollment information (program and enrollment date) g. Hourly load data January 2021 to March 2022 (kWh in and kWh out)	Stratification scheme and sample size by stratum will be finalized with SCE in May 2022
2. Customer characteristic file for participants:	For each account that completes an SRP project between June 1, 2022 and September 30, 2023: a. Account numbers (SA number, CT number) b. Rates and effective dates of rates for 2020-2023 c. Net metering status, date that NEM status became effective, and installed capacity d. NAICS or SIC industry codes, if applicable e. Zip Code, Climate Zone, Sublap, and A/B-Bank f. DR enrollment information (program and enrollment date) g. Date and savings estimate for any EE measure installed on site during the baseline year	Customer characteristics will be used to: ▪ identify participants, ▪ map to control profile ▪ produce results by segment
3. Project Information	a. TradePro	

Request	Detail	Purpose/notes
	b. Measure(s) installed c. Quantity of measures installed d. Measure details (capacity, efficiency, wattage) e. Key dates (project start, project completion, approval date) f. Savings forecast and weighted average EUL g. Upfront compensation payment	
4. Hourly or 15-minute interval data for participant and site selected to be part of the granular control profiles	a. Account numbers (Customer Account Number, Service Account Number, etc.) b. Date c. Hour/Interval d. kW - IN e. kW – OUT (if applicable) f. QC code, if applicable	Interval data will be used to estimate energy and demand impacts
5. Weather data for relevant stations from June 1, 2020 to May 31, 2024	a. Station ID b. Station Name c. Date d. Hour e. Temperature (dry bulb) f. Humidity	http://calmac.org/weather.asp Weather data will be used to model energy use
6. DR Event data June 1, 2020 to May 31, 2024 for all DR programs	a. Program name b. Event date c. Event start d. Event end e. Dispatch sub LAP or group called	▪ Please include all commercial programs/rates so we can account for dual enrollments