Memorandum

To: Lucy Morris, PG&E and Ken Agnew DNVGL

From: Kelly Marrin, Abigail Nguyen, and Barb Ryan AEG

CC: Jia Huang, Henry Liu, and Oriana Tell PG&E

Date: December 20, 2018 (With minor text revisions by SCE on 01/22/2019)

**Re: Update: Developing Ex-Ante Statewide Estimates of Savings Based on PG&E’s Smart Thermostat Study**

## The Purpose of this Memo

The purpose of this memo is to describe our approach to developing ex-ante statewide estimates of savings based on PG&E’s Smart Thermostat Study and to present those estimates.

## Background

The PG&E Smart Thermostat study completed in 2017 focused on understanding the effect of this emerging technology by estimating the impacts of three different smart thermostats in residential homes within three climate zones in PG&E’s service territory (CEC zones 11, 12, and 13). The study was a randomized encouragement design (RED) trial in which several thousand homes were randomly assigned to the encouraged group, all of whom were offered a free, directly-installed smart thermostat, and two thousand similar homes were randomly assigned to a control group, none of whom were offered a thermostat. The study was (primarily) designed to focus on electric savings, but the analysis was customized to look for season-specific savings for each fuel type.

Three different smart thermostats were included in the study. A total of 2,207 customers had one of the three thermostats installed. After the first year of the study, all three thermostats tested achieved annual electric (cooling) savings ranging from 4% – 5%. One of the thermostats tested also achieved annual gas (heating) savings. Savings persisted in the second year of the study ranging from 1% – 5%.

## Approach

Below we describe our approach to developing a statewide ex-ante estimate based on PG&E’s Smart Thermostat Study (Year 1) results. First, we describe the modifications to the ex-post models and associated rationale. Then, we describe the final ex-ante model, and finally, the actual extrapolation approach and results.

### Modifications to the Ex-post Models

The regression models that were developed for the original analysis were developed with several goals in mind:

* Estimating ex-post impacts for three different thermostats while accounting for the RED design.
* Attaining reasonable comparison between the statistical Difference in Differences (DID) and the regression output.
* Estimating monthly impacts with a primary focus on summer savings for electric customers.

The original ex-post analysis goals do not fully align with the goals of an annual ex-ante estimate. Through collaboration with the CPUC’s ex-ante contractor, DNVGL, AEG understands the goals of the ex-ante analysis to be as follows:

* Develop a model that is appropriate for extrapolation to weather conditions outside of what participants experienced.
* Attain a single estimate for smart thermostats that is representative of the types of thermostats that PG&E customers are installing, rather than three separate estimates.
* Develop an annual estimate of savings without focusing on a specific season.

### Updated Ex-Ante Model

Based on the ex-ante analysis goals, AEG collaborated with DNVGL to develop an updated model that is more in alignment with ex-ante methods. AEG also updated the analysis to estimate the annual savings directly rather than estimate savings at a monthly level. The updated model has the following characteristics:

* The model maintains the full structure of the DID and weather elements regardless of significance which provides more stability for the estimates during extrapolation.
* The model employs a fixed effect approach.
* The model estimates the overall effect of the two thermostats which make up over 97% of PG&E’s current rebates jointly.

The final ex-ante model specification is below:

where:

= the kWh consumption of customer in month

= a fixed effect for each customer

= an indicator that takes on a value of for each customer when month is in the treatment period

= an indicator which takes on a value of customer is in the encouraged group

= total cooling degrees for customer in month

= total heating degrees for customer in month

= the error for customer in month

### Extrapolation Approach

Before describing the actual approach, we present several considerations which we believe provide important context for the estimates.

#### Important Considerations

In thinking about the most appropriate methodology for the ex-ante estimates of savings we considered the following:

* The PG&E Smart Thermostat Study was not designed to be representative of all Californians.
* The PG&E Smart Thermostat Study model was estimated using actual CDD and HDD from PG&E’s 23 weather stations, not CEC climate zone data.
* The results do not represent weather normal impacts, but actual ex-post impacts.
* Impacts from participants in PG&E hottest zones may not be representative of all CA customers, i.e. the weather relationships estimated by the ex-ante regression model may not be applicable across the entire state.
* In the second year of the study (which is excluded from this analysis), we noted the following anomalies compared to the first year:
  + The study participants experienced weather that was quite different with a more extreme summer
  + The samples eroded due to move outs and other account changes. The shrinking samples made the encouraged effect even harder to detect.
  + Based on our survey, both control group customers and encouraged group customers began installing smart thermostats, nearly 15% of each group installed thermostats independent of the program.
  + The savings patterns seemed to shift, with more savings being seen in the spring (and even winter) and less savings in the summer. These shifts in savings trends seemed counterintuitive.

#### Extrapolation Method

Given the considerations outlined above, and the remaining timeframe for completing the analysis, we believe a simple straightforward approach, which makes minimal additional assumptions based on the model parameters, is the most appropriate.

At a high level our proposed approach is as follows:

1. Estimate the average HVAC cooling + fan energy consumption for a single family home in each of CA’s 16 climate zones using the SFM DEER prototype[[1]](#footnote-1).
2. Generate a single annual savings estimate from the ex-ante model using Year 1 study data only. Year 2 data was not included due to the anomalies outlined above.
3. Use the ex-ante model to estimate the annual cooling baseline for the encouraged group and estimate the percent impact. When using the ex-ante model, the cooling baseline represents the ex-ante model’s estimate of weather responsive load based on the relationship between CDD and kWh.
4. Apply the percent impact from step 2 to the HVAC cooling + fan energy consumption for a single family home with CAC in each of California’s 16 climate zones based on the SFM DEER prototype.

The approach is described in more detail below:

1. Estimate the average HVAC cooling + fan energy consumption using the SFM DEER prototype.

This was done by enabling latest SFM DEER prototype using the MASControl3 software application under the following assumptions: All 16 Climates Zones; Median Vintage - 2007; Central HVAC with Gas Furnace <45kBtuh; SEER14; AFUE 0.80; DEER Thermostat Schedule 3; and CTZ2010 weather.

2. Use ex-ante model to generate an annual savings estimate.

Next, we use the ex-ante model to estimate the annual savings attributable to the smart thermostats. To do this we estimate the savings as a linear combination of the following parameters from the ex-ante model:

This estimate represents the annual savings for the encouraged group. We then used a weighted acceptance rate to estimate the savings of the treated group under the RED.

3. Use the ex-ante model to estimate an annual cooling baseline load and percent impact.

Next, we estimate the cooling baseline load for the encouraged group. This represents the ex-ante model’s estimate of weather responsive load absent the thermostat based on the estimated relationship between CDD and kWh. To estimate the cooling baseline load, we use the following linear combination of variables from the ex-ante model:

.

Next, we calculate an annual savings estimate in percent terms by dividing the annual savings by the cooling baseline. The percentage impacts are reflective of how pilot participants responded to the treatment but using percentages allows us to scale the impacts to represent PG&E participants in each climate zone.

4. Apply percentage impacts to SFM DEER HVAC + fan cooling consumption estimates for each CEC climate zone to estimate impacts.

The final step in the analysis is to apply the annual percentage savings from the ex ante model to the HVAC + fan cooling consumption for an average single family home with CAC in each of California’s 16 climate zones as estimated by the appropriate SFM DEER prototype. This provides us with an estimate of how other Californians would respond to the thermostats assuming they would respond similarly to PG&E’s program participants.

One key advantage of the proposed approach is that it does not require us to apply parameter estimates (relationships) across very different climate zones (inland vs. coastal vs. desert) which would in turn require us to assume that all customers have the same relationship with weather regardless of their climate. It does, however, assume that all customers in the state would have a response that would be similar to those that participated in the pilot.

## **Results**

In this section we present the results of the extrapolation analysis. First, we present the ex-ante model output, then we present the cooling baseline and finally the savings estimates.

### Ex-Ante Model output

In Figure 1 below we present the output of the updated ex-ante model from Stata. As with any standard modeling output the parameter estimates are presented on the left and the associated statistics on the right. It should be noted that Stata does not output a typical R-squared value. However, the “within” R-squared is analogous to an R-squared from a regular OLS model (without fixed effects) and the statistic *rho* represents the total fraction of variation captured by the fixed effects.

Figure 1 Ex Ante Model Output



### Ex-Ante Cooling Baseline Estimates vs. SFM DEER HVAC + fan Cooling Baseline Estimates

In this section, we first present the cooling baseline estimate from the ex-ante model, then we present the SFM DEER HVAC + fan cooling baseline estimates, and finally, we present a comparison of the two.

In Table 1 we present the ex-ante model’s estimate of the total baseline and the cooling baseline. We also present the ex-ante cooling baseline as a percent of the total. The model estimates that cooling load makes up between 3 and 48 percent of the total baseline load depending on the month of the year. The overall pattern makes sense given what we know about the climate zones included in the pilot with the hottest months of the year occurring in the mid-summer months of July and August.

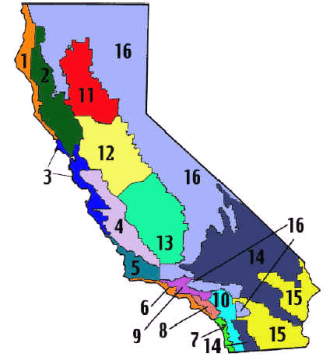
Table 1 Cooling Baseline Compared to Total Consumption Based on Ex-Ante Model

|  |  |  |  |
| --- | --- | --- | --- |
| Month | Total Baseline Consumption kWh | Cooling Baseline kWh | Cooling as a % of Total |
| November | 459 | 2 | 0% |
| December | 579 | 0 | 0% |
| January | 534 | 0 | 0% |
| February | 498 | 0 | 0% |
| March | 445 | 0 | 0% |
| April | 439 | 11 | 3% |
| May | 570 | 145 | 25% |
| June | 723 | 303 | 42% |
| July | 810 | 385 | 48% |
| August | 778 | 353 | 45% |
| September | 656 | 231 | 35% |
| **Annual** | **6,503** | **1,429** | **22%** |

In Table 2 we present the total consumption, and total HVAC + fan cooling consumption for each of the sixteen CEC climate Zones based on the SFM DEER prototype.[[2]](#footnote-2) Figure 2 also presents a map of the 16 zones for reference.

Table 2 Cooling Baseline Compared to Total Consumption Based on SFM DEER Prototype

Figure 2 California Climate Zones



|  |  |  |  |
| --- | --- | --- | --- |
| Climate Zone | Annual Cooling Load (kWh) | Annual Total Load (kWh) | Cooling as % of Total |
| CZ01 | 0 | 6,139 | 0% |
| CZ02 | 247 | 6,774 | 4% |
| CZ03 | 67 | 6,466 | 1% |
| CZ04 | 379 | 6,793 | 6% |
| CZ05 | 49 | 6,459 | 1% |
| CZ06 | 301 | 6,473 | 5% |
| CZ07 | 313 | 6,526 | 5% |
| CZ08 | 758 | 7,034 | 11% |
| CZ09 | 1,141 | 7,465 | 15% |
| CZ10 | 1,339 | 7,902 | 17% |
| CZ11 | 1,803 | 8,327 | 22% |
| CZ12 | 982 | 7,397 | 13% |
| CZ13 | 1,929 | 8,556 | 23% |
| CZ14 | 3,453 | 11,522 | 30% |
| CZ15 | 4,535 | 12,621 | 36% |
| CZ16 | 269 | 6,550 | 4% |

Recall from the introduction that the Pilot (PGE’s Year 1) focused on the three CEC climate zones 11, 12, and 13. If we compare the average total and cooling baselines across those three zones to the total and cooling baseline predicted by the model, we find that the estimates are quite close with the overall average proportion of cooling from the SFM DEER prototype being 19% and the average proportion of cooling coming from the ex-ante model being 22%. Table 3 presents the comparison.

Table 3 Proportion of Cooling Load SFM DEER Prototype vs. Ex-Ante Model

|  |  |  |  |
| --- | --- | --- | --- |
|  | Annual Total Load (kWh) | Annual Cooling Load (kWh) | Cooling as % of Total |
| DEER (11, 12, 13) | 8,093 | 1,571 | 19% |
| Ex-Ante Model | 7,095 | 1,559 | 22% |

### Ex-Ante Savings Estimates

In Table 4 we present the ex-ante model’s estimate of the savings as a percent of the ex-ante cooling baseline. The model shows that savings makes up about 11% percent of the total cooling baseline at an annual level.

Table 4 Cooling Baseline and Annual Savings as a Percent of Baseline

|  |  |  |  |
| --- | --- | --- | --- |
| Month | Cooling Baseline  (kWh) | Annual Savings  (kWh) | Savings as a % of Baseline |
| November | 2 | -5 | -278% |
| December | 0 | -15 | - |
| January | 0 | -11 | - |
| February | 0 | -7 | - |
| March | 0 | -3 | - |
| April | 11 | 0 | 3% |
| May | 145 | 19 | 13% |
| June | 303 | 41 | 13% |
| July | 385 | 52 | 14% |
| August | 353 | 48 | 13% |
| September | 231 | 31 | 13% |
| *11-Month Total* | |  | | --- | | *1,429* | | *157* | *11%* |
| **Annual** | **1,559** | **171** | **11%** |

In Table 5 below we present the ex-ante impact for each climate zone. The impact is calculated by applying the 11% impact predicted by the model to the annual cooling load in each zone. We present the nominal impact (kWh) and the impact as a percent of the total annual load below.

Table 5 Ex-Ante Impact by Climate Zone

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Climate Zone | Annual Cooling Load (kWh) | Annual Total Load (kWh) | Cooling as % of Total | % Impact (Cooling) | Savings (kWh) | % Impact (Total) |
| **CZ01** | 0 | 6,139 | 0% | 11% | **0** | **0%** |
| **CZ02** | 247 | 6,774 | 4% | 11% | **27** | **0%** |
| **CZ03** | 67 | 6,466 | 1% | 11% | **7** | **0%** |
| **CZ04** | 379 | 6,793 | 6% | 11% | **42** | **1%** |
| **CZ05** | 49 | 6,459 | 1% | 11% | **5** | **0%** |
| **CZ06** | 301 | 6,473 | 5% | 11% | **33** | **1%** |
| **CZ07** | 313 | 6,526 | 5% | 11% | **34** | **1%** |
| **CZ08** | 758 | 7,034 | 11% | 11% | **83** | **1%** |
| **CZ09** | 1,141 | 7,465 | 15% | 11% | **125** | **2%** |
| **CZ10** | 1,339 | 7,902 | 17% | 11% | **147** | **2%** |
| **CZ11** | 1,803 | 8,327 | 22% | 11% | **198** | **2%** |
| **CZ12** | 982 | 7,397 | 13% | 11% | **108** | **1%** |
| **CZ13** | 1,929 | 8,556 | 23% | 11% | **212** | **2%** |
| **CZ14** | 3,453 | 11,522 | 30% | 11% | **379** | **3%** |
| **CZ15** | 4,535 | 12,621 | 36% | 11% | **498** | **4%** |
| **CZ16** | 269 | 6,550 | 4% | 11% | **30** | **0%** |

1. Prototype based on the latest SFM DEER2019 (MASControl3) updates. [↑](#footnote-ref-1)
2. The prototype used in this analysis represent average energy consumption for single family homes per Climate Zone based on the latest DEER (MC3) documentation. The general assumptions include: SFM Residential prototype, Vintage 2007, DEER T-stat 3, Central HVAC with Gas Furnace < 45kBTUH, SEER 14 and CTZ2010 Weather. [↑](#footnote-ref-2)