RECURVE

SHAPE THE FUTURE OF ENERGY

DRAFT: Comparison Groups for the Coronavirus Era

Plan for Design, Testing, and Execution

July 31, 2020

Context

The COVID-19 pandemic is having major impacts on energy consumption with some hard hit areas reporting grid-level reductions in electricity usage of 10 - 20%. However, impacts vary dramatically across sectors (residential vs commercial) and sub-sectors (hospitals vs. restaurants).

In MCE data Recurve has observed clear and abrupt impacts of COVID-19 on in-sector and between-sector consumption patterns. High-level are given in the Figure 1, which shows differentially privatized[1,2] measurements of average load shape impacts to the Residential (left) and Commercial (right) sectors.[3] The CalTRACK 2.0 methods implemented via the open source OpenEEmeter[4] have been used in these measurements and will be the foundation of this research. On average residential usage has risen significantly, while commercial consumption has fallen. These differences are most readily apparent during mid-day periods, presumably reflecting many customers staying home from work in response to stay-at-home mandates and a sense of social responsibility.

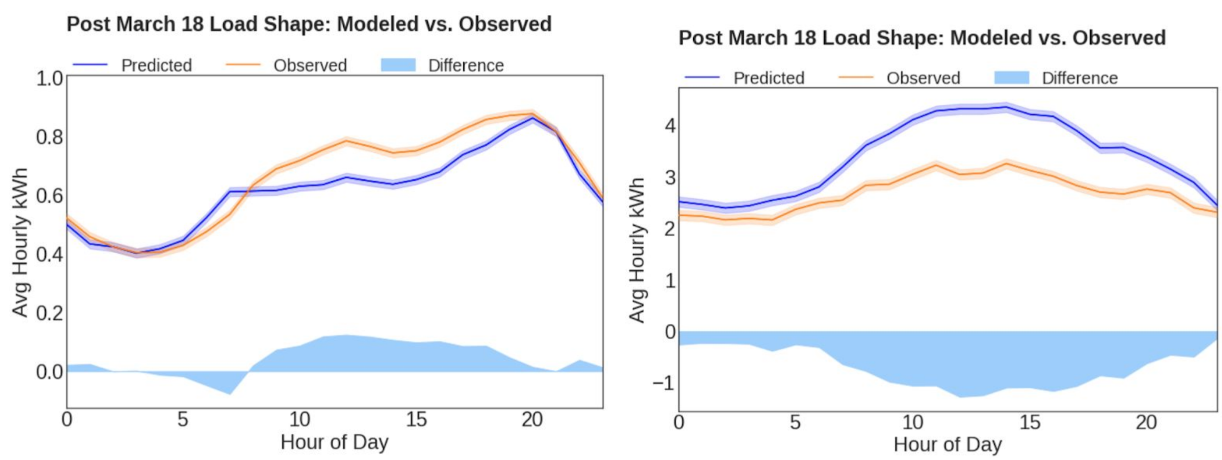

Figure 1: Average load shape impacts of COVID-19 on Residential (left) and Commercial (right) sectors of MCE's customer base.

A wide variety of impacts are observed for different sub-sectors of MCE's commercial customer base. Figure 2 shows several examples of both "essential" and "non-essential" business types.

[1]See for example *Google COVID-19 Community Mobility Reports: Anonymization Process Description (version 1.0)*, A. Aktay, S. Bavadekar, G. Cossoul et al. (2020). https://arxiv.org/abs/2004.04145
[2] Recurve's experience with differential privacy stems from the Energy Data Vault project, which is supported by the U.S. Department of Energy with the goal of enabling the secure sharing of energy data. https://www.energy.gov/eere/buildings/energy-data-vault
[3] These figures, along with those of Figure 2 were presented recently at the August 5th California Energy Efficiency Coordinating Committee meeting. https://www.caeecc.org/caeecc-documents
[4] See www.caltrack.org for full documentation

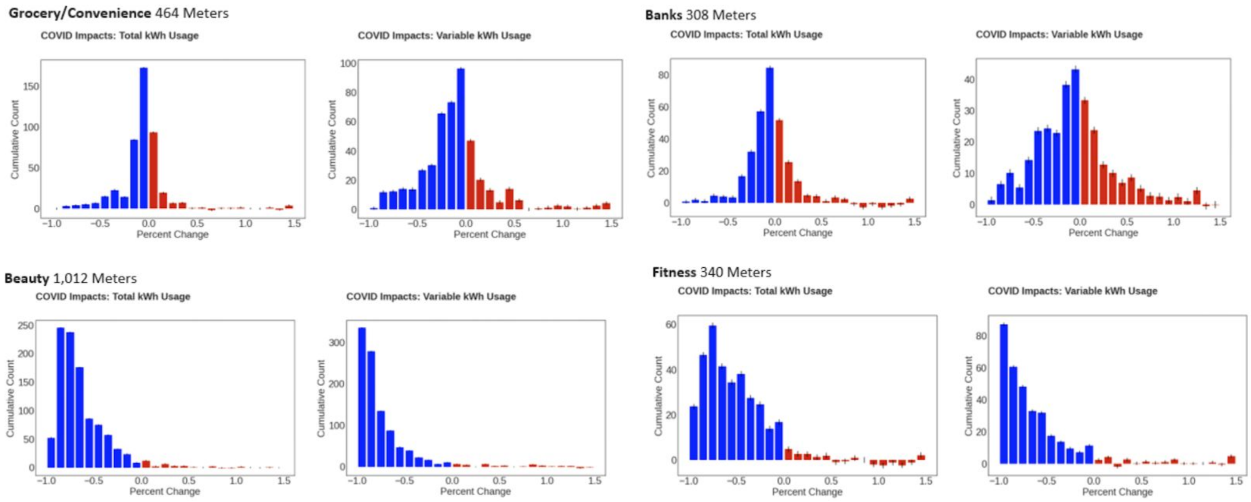

Figure 2: Histogram plots showing the total and variable load impacts experienced from COVID for several subsectors of MCE's commercial customer base. From top left clockwise: Grocery/ Convenience, Banks, Fitness, and Beauty.

These results make it clear that individual customer-level responses vary to a high degree, even within sectors that exhibit either high or low total impacts from COVID.

The COVID-related changes in energy consumption have important implications for the accurate measurement of programs based on Normalized Metered Energy Consumption (NMEC). Strategies to account for COVID over time must address the following:

- The COVID-19 crisis is occurring in the **reporting period** of some existing programs and in the **baseline period** of some future programs.
- Adjustments for COVID-19 will need to be made to ensure fair and reasonable payments to implementers and to best isolate the impacts of the program.
- Different jurisdictions will have different data available, both in number of customers and in the interval of consumption (AMI vs. monthly for example).

The ability to quickly bring sound and convincing strategies to the table will be essential to maintaining confidence in and credibility of NMEC programs and measurements. The development and deployment of consistent, clear, and replicable analyses will be key to keeping programs on track.

We will need to be prepared for several situations:

- Existing in-field programs (reporting period) vs. future programs (baseline period)
- The spectrum from broad-based residential programs that serve many thousands of customers to non-residential programs that serve tens or hundreds of buildings and focus on a particular customer type or application
- Hourly vs daily or monthly data

In general, Recurve has identified three possibilities to address COVID in the measurement of NMEC programs:

1. Maintain course, which may be most appropriate if the most important question remains changes in consumption regardless of outside trends.
2. Treat COVID-19 as a large non-routine event and institute a blackout period. When markets return to stable patterns, weather data could be applied to reporting period models to assign savings and payments for the duration of the COVID-19 NRE.
3. Incorporate comparison group measurements to isolate and eliminate impacts of COVID-19

The first two options are fraught with issues that won't be detailed here. However, comparison groups have long been a mainstream approach to adjust for exogenous factors and non-routine events. Therefore, this study plan will focus on work needed to prepare for the third option, which is viewed as a best practice by many in the industry yet will need further development to meet the challenges of COVID-19 and deploy in a routine tracking fashion to enable continuation and launch of NMEC and pay-for-performance programs.

Data Summary

To test and learn from the widest variety of important scenarios, an ideal dataset would have the following properties:

- Large number of meters
- Interval (15-minute or hourly) meter data
- Meter data covering a long period of time (several years)
- Regular updates to meter data (monthly or shorter)
- Electric and gas data
- Representation of different sectors
- Reliable and complete metadata
- Variety of climate regions

Fortunately, with the generous partnership of MCE and support from the U.S. Department of Energy, Recurve has a rich and regularly-updated dataset for residential and non-residential customers that can be used to test and develop reliable and open-source methods. Importantly this dataset contains formatted AMI data for MCE's population dating back to 2017. Along with the meter readings, the MCE data includes a number of metadata fields that can be used to catalog customer segments, which is a critical aspect of formulating relevant tests as will be discussed below. MCE's customer base is spread across a fairly diverse geographic region, largely north and east of San Francisco as can be seen in Figure 3.

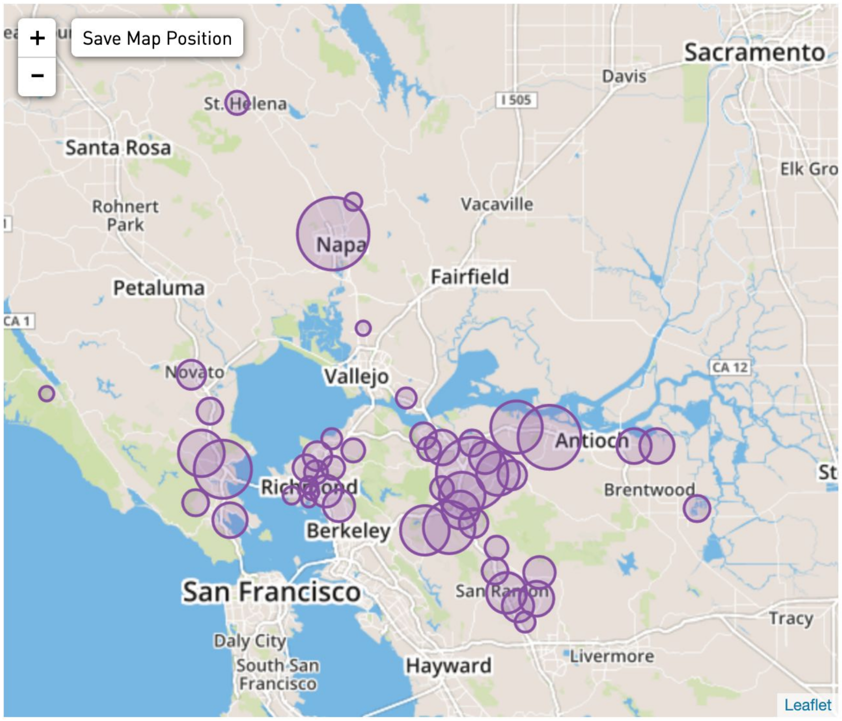

Figure 3: Map showing the density of MCE meters.

Table 1 results from an analysis of recent MCE data[5] and shows changes in consumption by sector computed during the COVID Period. Here we have used the Hourly CalTRACK 2.0 methods to project forward a counterfactual that represents predicted usage expected in the absence of COVID.

Post March 18 Percent Change in Usage by Sector; City = All

SECTOR	CNT 2019	TOTAL KWH	BASELOAD KWH	VARIABLE KWH
Commercial	28518	-21%	-10%	-33%
Residential	37833	11%	5%	15%
Government	337	-9%	-10%	-9%
Industrial	1682	-14%	-11%	-17%
Agriculture	337	7%	-2%	13%

Table 1: COVID impacts by sector: Analysis run on a sample of MCE metersBaseload kWh is an estimate of the "always on" component of electricity consumption and Variable kWh is an estimate of the electricity consumption that varies throughout the course of a typical day. A much more detailed description of these concepts and how they are calculated is provided in Appendix A.

The timeframe under consideration as the "COVID Period" is March 19 - May 8, 2020. This stems from the fact that California entered a statewide stay-at-home mandate on March 19. However, different locales throughout the state, including parts of MCE's service territory were under restrictions before this date. To ensure a model not contaminated by COVID

[5] The "4013" data originates at PG&E and the associated file contains many metadata fields

responses, and therefore appropriate to serve as a counterfactual to gauge COVID impacts, the timeline shown in Figure 4 has been used to produce the initial assessment of COVID impacts given in Table 1 and Table 2 below.

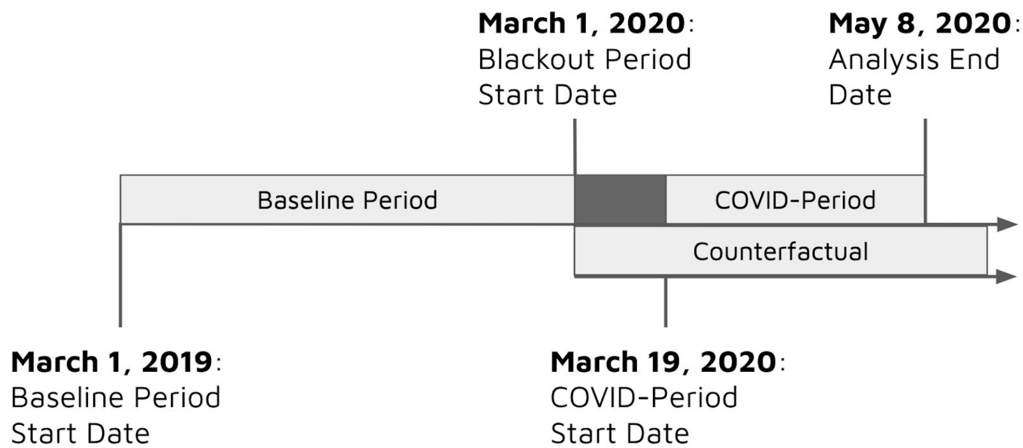

Figure 4: Metering timeline for an initial assessment of COVID-Impacts

Roughly 90% MCE meters are associated with the residential sector. The commercial and industrial sectors are grouped together in the raw data and we have broken them out via a NAICS code filter (any non-residential meter associated with '3*' is labeled as Industrial).

Table 2 digs deeper into the commercial sector by "NAICS" code and description.[6] The table shows NAICS groupings, which are detailed in the appendix, along with high level statistics on changes in energy consumption observed for the COVID-period (March 19 - May 8, 2020).

[6] https://www.census.gov/eos/www/naics/ *"The North American Industry Classification System (NAICS) is the standard used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the U.S. business economy."* (April, 2020)

NAICS GROUP	CNT 2019	TOTAL KWH	BASELOAD KWH	VARIABLE KWH
Beauty	1028	-70%	-26%	-83%
Fitness	346	-63%	-46%	-75%
Schools	129	-42%	-22%	-61%
Hotels/Lodging	407	-41%	-38%	-45%
Offices	1267	-33%	-13%	-51%
Churches/Religious	601	-31%	-17%	-44%
Resturants/Bars	1884	-30%	-16%	-42%
Medical_Offices	1095	-28%	-6%	-43%
Retail	1551	-28%	-12%	-39%
Real_Estate	2773	-17%	-5%	-27%
Automotive	964	-14%	-12%	-16%
Administrative/Civil	2707	-14%	-6%	-22%
Gas_Stations	185	-12%	-7%	-22%
Construction/Contractors	967	-11%	-2%	-19%
Banks	308	-7%	-3%	-14%
Unassigned	51869	-6%	-4%	-9%
Grocery/Convenience	473	-6%	-1%	-13%
Warehousing/Postal	153	0%	-3%	8%

Table 2: Commercial MCE meters by NAICS group along with COVID Impacts

These groupings will serve as a foundation for experimental design in the testing of comparison group strategies.

Experimental Design

Phase I: Comparison Groups

We now turn our attention to the design and development of Comparison groups. Importantly, the comparison group techniques that result from this effort must enable NMEC programs *as they are planned and executed*. In that sense these methods may not be equivalent to those applied in a typical impact evaluation context in which the program has run its course and there is ample time and data to select an "ideal" comparison group. One immediate consequence of this approach is the inability to use certain common comparison group methods such as future participants, which requires knowledge of who will participate one or more years down the road. Individual meter matching algorithms may be possible to execute as the program progresses, but provide little basis for forecasting, a critical element for most programs. Stratified sampling, however, can be conducted based on baseline-period data alone and on a forecasted basis. This method of comparison group selection is well-established in many fields of study and will be the focus of the research.

Keeping the purpose of supporting meter-based programs in mind, we outline the core questions that should be addressed by this research before defining specific methodological steps.

Core Questions

1. **What is the simplest formulation of comparison groups to provide reliable results for most programs?**
    a. Is stratified sampling demonstrably superior to random sampling?
    b. On what basis or bases should stratified sampling be conducted? Is there a single scheme that can be reasonably used for all customer types?
    c. How should a customer's energy consumption, both pre- and post-COVID, be incorporated into comparison group selection?
    d. Are different stratification schemes needed/beneficial for the commercial vs. residential sectors?
    e. Are different stratification schemes needed/preferable for different groups of treatment customers, which may vary by demographics, business type or other factors?
    f. Are different stratification schemes needed/preferable when hourly data are available (compared to daily/monthly data)?
    g. Should we be sampling at a building- rather than meter-level? What are the implications of either approach?

2. **Practical implementation questions**
    a. On what basis if any should outliers be identified and removed in both the treatment and comparison groups?
    b. What should define "baseline" and "reporting" periods for comparison group meters? How should different vintages of comparison groups be formulated?
    c. How many meters are needed to formulate reliable and stable comparison groups?

High Level Research Strategy

The success of this research will hinge on our ability to specify the elements of comparison group selection and tracking that can be standardized and to also identify where some flexibility may be needed. The primary research strategy will be composing "treatment" samples as random selections of customers from specific groups within MCE's customer base, and then assigning comparison groups. Unlike most applications of comparison groups, these "treatment" groups will not differ due to program participation. Instead of measuring program impacts, the objective will be to gauge the degree to which divergence is observed between "treatment" and comparison samples before and during the COVID period. Results will inform the ability of particular comparison group strategies to represent potential program populations and accurately account for COVID impacts. Because the "treatment" groups will not differ from the comparison group counterparts because of program participation, they will afford a clear test of the ability of comparison groups to track impacts due only to COVID and other exogenous factors.

An ongoing statistical match between "treatment" and comparison samples can be considered good evidence that such an approach would perform well when incorporated into the context of a program that serves a similar customer base. Where we observe significant divergence the opposite will be true. When we observe divergence will also provide important information and must be considered in the experimental design of this phase; depending on the specific "treatment" and comparison groups we may observe a good fit throughout, a high degree of bias throughout, or a good fit in the pre-COVID period but divergence that arises only during the COVID-period.

Sample Selection:

i. **General Considerations**

The samples for metering established in Phase I provide the foundation for "treatment" and comparison samples in this phase. However, before a more detailed discussion on sampling, we take a step back to describe the nature of in-field programs that will ultimately need comparison groups.

Demand-side programs do not serve a random selection of customers across a utility service territory that can be easily matched by a blind pull of non-participant data. Instead, programs are driven by utility/regulatory goals, contractors, and customer relationships. This manifests with programs that retrofit lighting in gas stations and liquor stores, programs that serve low-income customers or customers with old equipment or inefficient usage patterns, programs that renovate restaurant kitchens, or programs that install new HVAC systems in larger office buildings among a myriad of other possibilities. The point being that care must be taken to ensure that comparison groups are reasonably reflective of the treatment customers they are designed to emulate.

With this in mind, there are important limitations that will originate with utility data. First, many utilities have only monthly data available. Therefore, despite MCE offering 15-minute or hourly AMI data, it will be important that we also investigate situations where only monthly data are available. However, for these monthly tests the hourly data will provide us important insights to the nature of the bias that we may observe with the monthly samples.

There may also be cases where the hourly data reveals biases between comparison and "treatment" groups that are not observable with monthly data alone.

Second, oftentimes we will be presented with program data that sheds more light on customer type than is likely to be available in utility metadata, which can be sporadic and unreliable. Given this, for the "treatment" samples, **it will be important to formulate comparison groups using only meter data** and what can be discerned therefrom.

ii. COVID Considerations: Addressing Selection Bias

Recurve anticipates that an important element of comparison group selection will be the degree to which a customer's energy consumption has changed in response to COVID. A program may recruit customers who are impacted differently than the general population due to COVID-19. For instance, a program serving restaurants will bring forth a treatment group that differs from the general population of commercial buildings because restaurants have been impacted more than the commercial sector as a whole (see Table 2). Further, within the restaurant sub-sector, the program may see participation from customers who are impacted to a greater or lesser degree than the general population of restaurants. Such misalignments between treatment and comparison samples contribute to what is known as selection bias whereby establishment of comparison groups via random sampling is insufficient to isolate and remove non-program impacts.

With the potential for selection bias due to COVID, Recurve anticipates that it will be important to ensure that the degree to which participating customers have been impacted by COVID is captured and reflected in the comparison group. More methodological details on how this will be attempted are provided below.

Research Plan

With several critical questions, this research will not succeed without a systematic approach geared toward isolating the individual sources of uncertainty. Understanding how comparison group performance changes as data becomes more granular, the comparison pool grows larger, and sampling schemes become more finely tuned will be key to nailing down the best approach and understanding structural limitations that should be expected.

With this and the above considerations in mind Recurve will implement a study plan in three parts:

Part I. Test and Refine Comparison Group Selection Methods

In this step Recurve will design, code, and test stratified sampling schemes to ensure that unbiased samples can be formed and/or to understand the tradeoffs between finer binning and higher dimensions of stratification. Recurve will also develop specific parameters based on customers' usage patterns that can be used directly as the basis for stratification.

Definitions of usage characteristics Recurve will utilize for stratification parameters are given in Appendix C.

In stratified sampling, a *comparison group* is derived from a larger *comparison pool* in order to match the distribution of one or more characteristics observed in the treatment group. As an example, Figure 5 shows the result of stratified sampling that Recurve recently conducted as part of a heat pump analysis in partnership with the Sacramento Municipal Utility District (SMUD).[7] In this case, Recurve binned the treatment sample into ten equal groups defined by the customers' non-temperature dependent usage. The comparison pool (left-hand figure) was then resampled to a stratified comparison group (right-hand figure) to ensure that an equal number of the non-participating customers was also included in each bin. The resulting comparison group was then used to estimate the impacts of heat pump electrification.

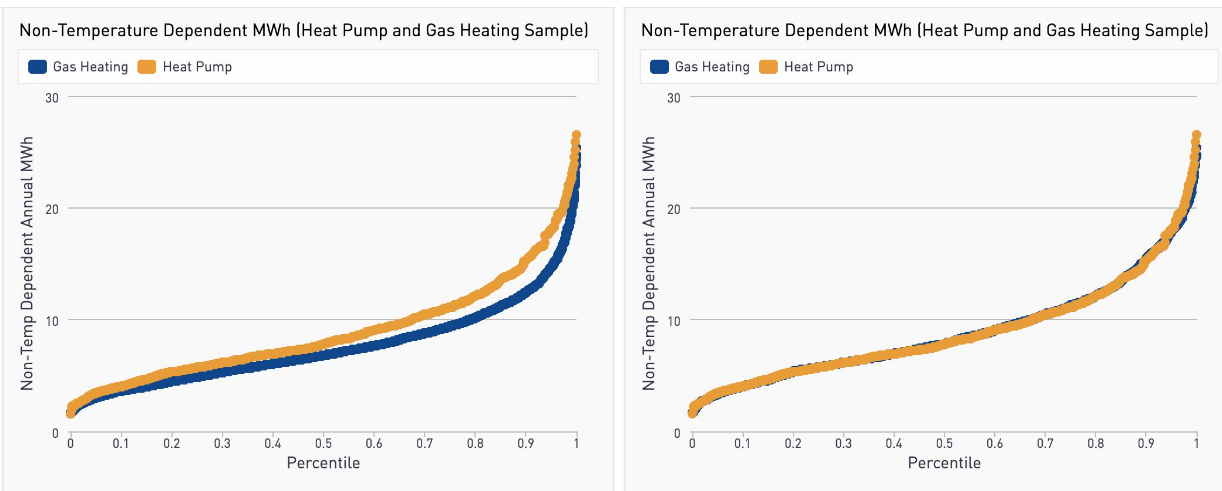

Figure 5. Temperature-independent load by customer for the gas heating sample (blue) and the heat pump sample (orange). The left graph shows the full gas heating sample and the right graph shows results after stratified sampling of the gas heating sample.

While this example reveals the general strategy behind stratified sampling, accounting for COVID may require more sophisticated schemes. In particular, Recurve anticipates that stratification on more than one dimension will likely be needed. At a minimum it is likely that sampling based on both pre-covid usage characteristic(s) and response to COVID will be needed to define appropriate comparison samples.

Part II. Pre-COVID: Test and Refine Comparison Group Selection Methods

With the mechanics of stratified sampling completed, Recurve will turn attention to initial testing of comparison group performance. In order to set the foundation for the strategies

[7] A.M. Scheer et. al, *Electrification: Meter Data Analysis of Grid Impacts and the Opportunity for Efficiency*, ACEEE Summer Study, 2020.

that will be required to account for COVID impacts, we will begin testing by focusing on the pre-COVID time period.

This research step will be critical to identify the usage characteristics that can be most reliably used for stratification and the underlying divergence between treatment and comparison groups that should be expected. Where poor performance is observed, approaches can be weaned out before incorporating COVID impacts in Part III.

Figure 6 shows the metering timeline that will be utilized for this phase.

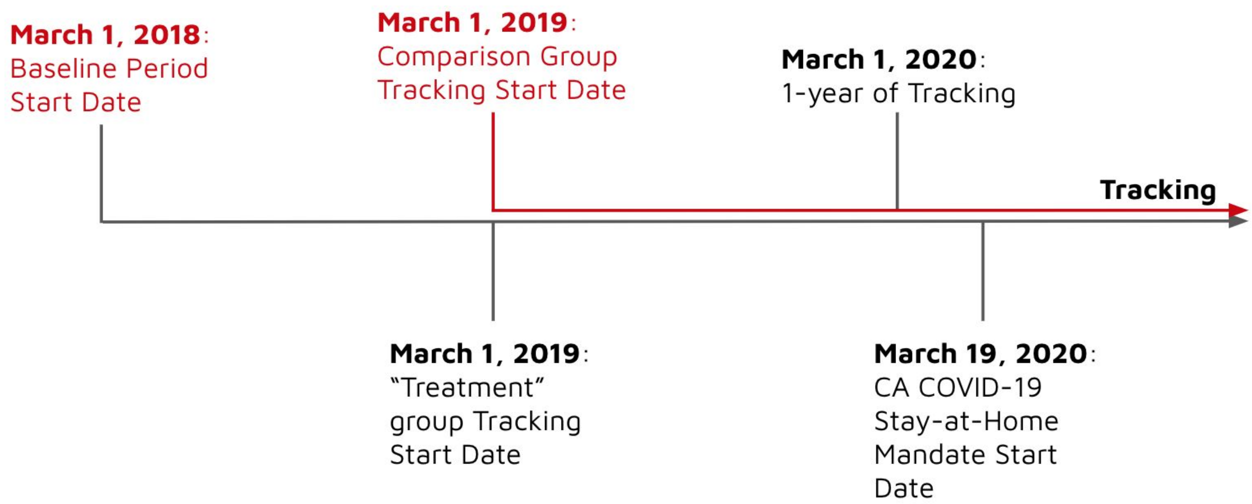

Figure 6. Metering timeline for pre-COVID comparison group testing.

The important stages of this timeline are as follows:

March 1, 2018 - Feb. 28, 2019: Computation of Sampling Parameters and Selection of Comparison Groups:

The period from March 1, 2018 - February 28, 2019 will be used for the computation of possible sampling parameters. Based on this year of data and CalTRACK modeling, "treatment" and associated comparison groups would be selected.

March 1, 2019 - Feb. 29, 2020: Pre-COVID Comparison Group Performance:

Independent tracking of "treatment" and comparison groups will then begin on March 1, 2019. A year of observation then brings us to the precipice of the COVID-19 disruption. At this point we will have an important gauge of how well our comparison group strategies are performing through relatively stable conditions.

March 1, 2019 - March 18, 2020: Transition Period:

During this timeframe California was transitioning from a period of little to no disruption from COVID to a full statewide stay-at-home restriction. Because of the ambiguity of this timeframe, it is not included as part of the analysis

March 19, 2020 onward: COVID Period:

At this point no attempt will have been made to account for COVID in comparison group selection. Nevertheless, it will be informative to understand how comparison groups selected based on pre-covid data alone succeed or fail to represent the corresponding treatment groups. Therefore tracking results during this timeframe will be conducted.

Initial Sample Selection

As detailed above and in Appendix B, Recurve has cataloged MCE's commercial customer base into a number of high-level "NAICS groups" that correspond to business types. Among these NAICS groups are a variety of customer segments including Retail, Grocery, Restaurants and many more. To start, we will select a portion of one of these NAICS groups (for instance 650 non-solar office buildings with annual usage between 5 and 250 MWh) as the "treatment" sample. We will then build a series of comparison group tests in which each step adds an element that allows us to build knowledge toward our key questions:

1. Comparison group built from random sample of commercial customers
2. Comparison group from stratified sampling only on baseline period Annual usage
    a. 2 bins
    b. 4 bins
    c. 8 bins
    d. 12 bins
3. Comparison group from multidimensional stratified sampling - In this step Recurve will conduct sampling based on multiple stratification parameters. It will be beneficial to sample based on parameters that are largely uncorrelated in order to represent fundamentally different aspects of customer usage that are of interest. For hourly testing it will likely be useful to include one or more load shape parameters. The following are examples of features that could be used in conjunction in a two-dimensional stratified sampling scheme (see Appendix C for more information on specific parameters):
    - Hourly: Annual Usage, Percent Cooling
    - Hourly: Annual Usage, Percent Summer Peak
    - Monthly: Summer Usage, Winter Usage, Shoulder Usage

In each of these tests Recurve will compute T-test and KS-test results to assess the degree to which the stratified samples are equivalent to the treatment group counterparts. This assessment will be done at inception as well as for pre-covid and post-covid time periods.

a. Annual kWh
b. Seasonal (or Monthly) kWh
c. Avg. Daily Load Shape
d. Avg. Weekly Load Shape

e. Avg Hourly Load Shape

Need to test multiple comparison group selections for same treatment groups to determine sampling uncertainty.

Additional Samples

With so many iterations possible on both “treatment” and comparison samples, it will be important to cover as many cases and as many shades as possible without creating too much complexity, cost, or organizational burden. In an attempt to strike a reasonable balance, the following nine “treatment” samples are proposed in Table 3 along with the number of meters in MCE service territory.

Category	Meters Available	Treatment Sample	Treatment Sampling Scheme	Usage Criteria
Residential	~450,000	5,000	High Peak Users	< 50 MWh/year
Residential	~81,000	5,000	CARE Customers	< 50 MWh/year
Administrative/Civil	3,527	1,500	Random by NAICS	< 250 MWh/year
Resturants/Bars	2,124	1,000	Random by NAICS	< 250 MWh/year
Retail	1,996	750	Random by NAICS	< 250 MWh/year
Offices	1,431	650	Random by NAICS	< 250 MWh/year
Beauty	1,176	500	Random by NAICS	< 250 MWh/year
Grocery	594	220	Random by NAICS	<2,500 MWh / year
Hotels/Lodging	471	200	Random by NAICS	<2,500 MWh / year

Table 3: Proposed “treatment” groups and meter counts

Part III - Incorporation of COVID Impacts

With the techniques that succeeded in Part II as a foundation, this task will be devoted to layering in the impacts from COVID as a stratification parameter and gauging comparison group performance in the post-COVID timeframe. In order to carry this out effectively, the COVID timeframe itself must be split into a sample selection period and a performance period. The metering timeline anticipated is provided in Figure 4.

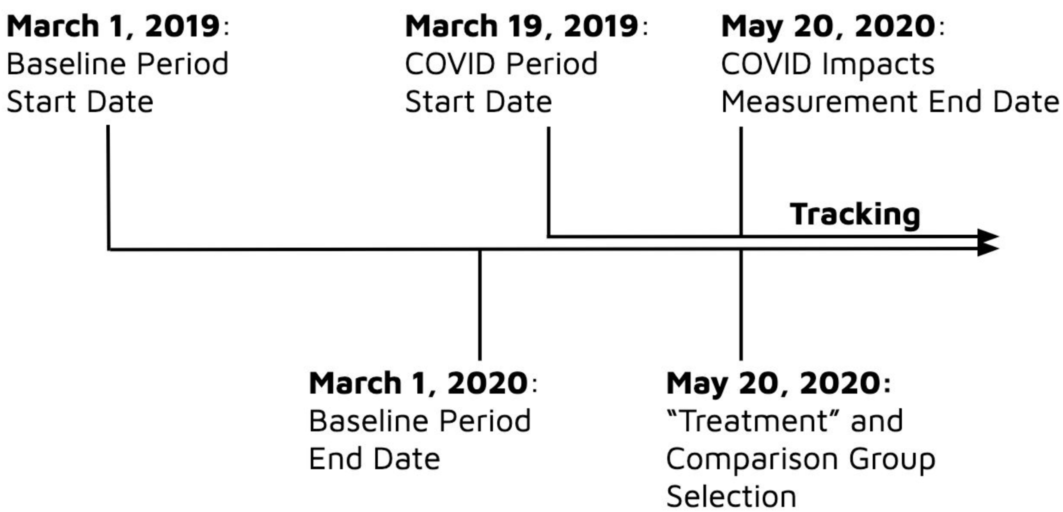

Figure 4. Metering timeline for COVID comparison group testing.

The important stages of this timeline are as follows:

March 1, 2019 - Feb. 28, 2020: Computation of Sampling Parameters and Selection of Comparison Groups:

The period from March 1, 2019 - February 28, 2020 will be used for the computation of possible sampling parameters. Based on this year of data and CalTRACK modeling, "treatment" and associated comparison groups would be selected.

March 19, 2020 - May. 19, 2020: Computation of COVID Impacts

In this two-month time period the percentage change in usage due to COVID will be computed for all customers. This metric will be used alongside those developed in Phase II as a stratification parameter.

May 20, 2020 - TBD: COVID Comparison Group Performance

This time period will be devoted to gauging the performance of comparison groups to account for COVID impacts.

Covid impacts will be determined on an individual-meter basis by projecting forward the baseline CalTRACK model for the March 19 - May 19 period and comparing this counterfactual to observed meter values. For each customer the percentage change in consumption during the COVID period will then be used in conjunction with the other stratification parameters from Part II for a stratification scheme that accounts for customer-response to COVID. A Similar suite of diagnostics will be conducted with the most promising strategies recommended for use in real program M&V.

Appendix A: Baseload and Variable Load Estimates

There are many definitions of "baseload" and many ways to estimate it. Here we are defining baseload as the electricity consumption constant across a 24 x 365 timeframe. This definition can be applied at a meter, building, sector, or grid level. Once baseload is determined, variable load can simply be calculated as the difference between total usage and baseload. The concepts of baseload and variable load are illustrated schematically in Figure A1.

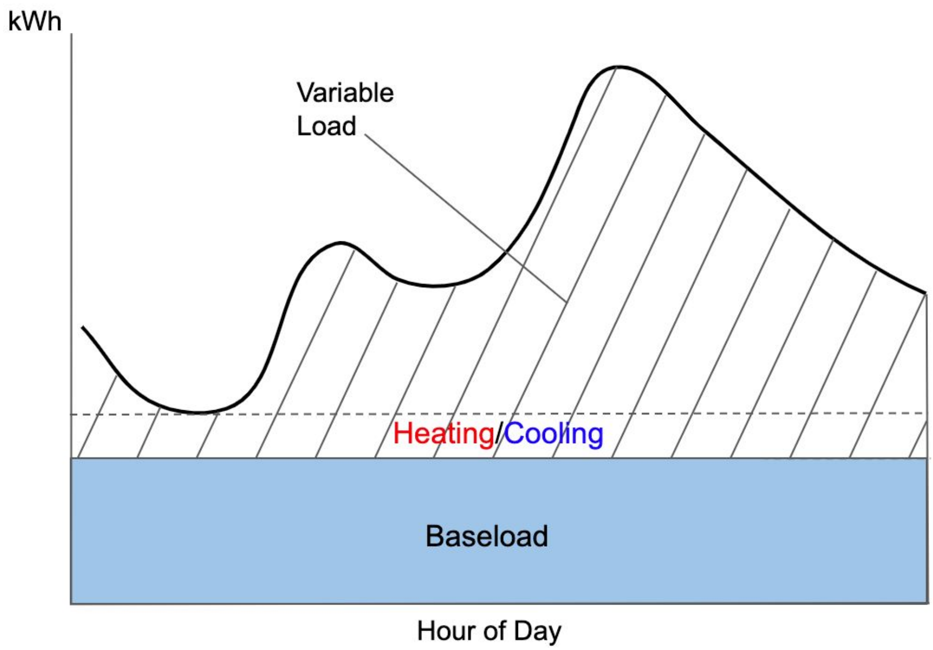

Figure A1: A schematic of baseload and variable load. Recurve estimates average daily baseload by assessing low-usage hours during days with little or no temperature dependent consumption. Variable load is then taken as total load minus baseload.

With AMI data one can make a crude estimate of baseload by simply taking the average of usage during the overnight hours, or by finding the minimum usage during a day or specified time period and extrapolating. Recurve has found that while valuable for certain cases, these methods often yield spurious results at an individual-customer level. A particular weak point is the inclusion of heating and cooling loads in the baseload calculation, which makes it difficult to successfully isolate usage from true "always on" end uses.

Recurve has devised a method to estimate baseload that is more robust using the outputs of both the CalTRACK daily and hourly models. To understand this calculation it is first helpful to revisit certain aspects of the CalTRACK daily model. In CalTRACK daily calculations on electric data four models are fit to each meter trace. These models correspond to those illustrated in Figure A2.

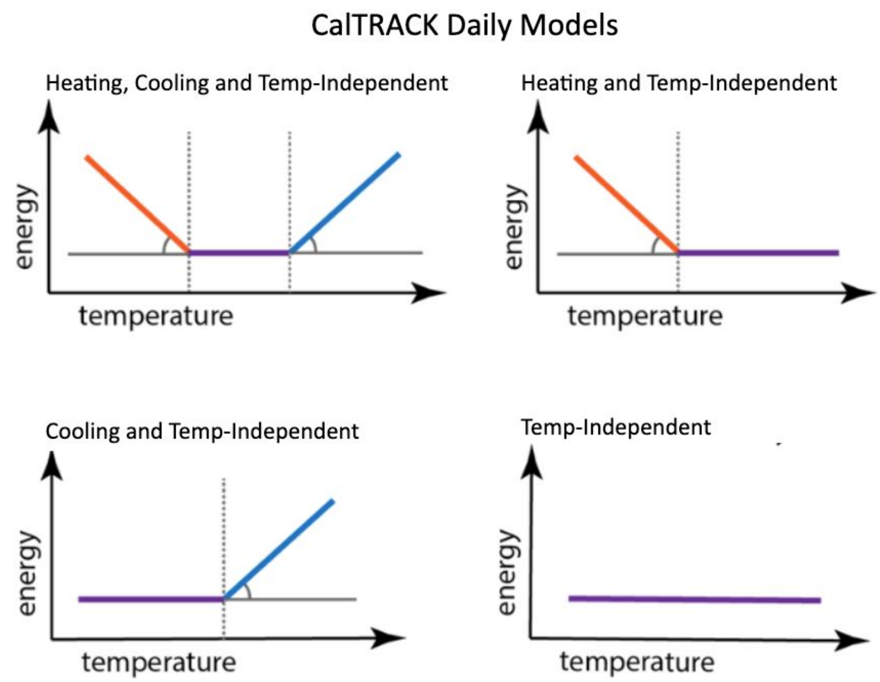

Figure A2: Schematics of the four possible CalTRACK daily models.

Clockwise from bottom right, CalTRACK fits a temperature-independent model and then models with only cooling, only heating, and both cooling and heating as additional temperature dependent components. After computing the optimal parameters for each of these models (a process involving determining the optimal balance point temperature and slope to model the temperature-dependencies of load), the final CalTRACK daily model for that meter is chosen as the best fit among the four models.

When this process is complete one can disaggregate temperature dependent usage (heating and cooling) from non-temperature dependent usage. However, the non-temperature dependent usage constitutes both baseloads along with discretionary loads like interior lighting, cooking, and many plug load devices among others. Therefore another step is needed to isolate baseload and for this the CalTRACK hourly outputs are essential.

In cases for which the CalTRACK daily model is assigned as temperature independent (bottom right of Figure 11), there is less risk of mixing variable heating and cooling loads into the baseload calculation as the observed usage was best fit via a model with no temperature dependence. In this case, for each day of the year, Recurve rank orders the hourly usage, discards the lowest hour (to avoid spurious results from short outages or other oddities) and takes the average of hours 2 - 4 as baseload for that day. Averaging baseload then across a 365 day period produces an estimate of annual average baseload.

For CalTRACK daily models that have temperature dependence, the situation is more complicated. We take the Heating, Cooling, and Temp-Independent model as an example.

Where this model is assigned there are three general possibilities shown schematically in Figure A3.

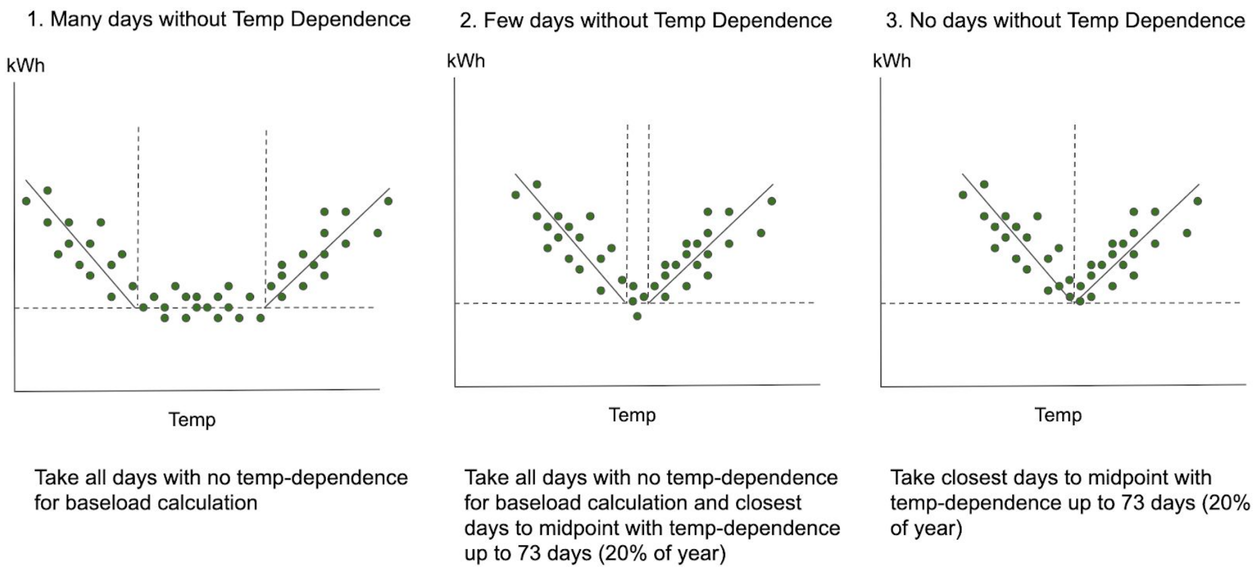

Figure A3: Illustrations of three possibilities for the CalTRACK daily cooling, heating, and temperature-independent model. Left: Many days exist between balance point temperatures. Middle: Few days exist between balance point temperatures. Right: No days exist between balance point temperatures.

In the leftmost scenario shown in Figure A3, many days (at least 20% of all days) are observed with mean temperatures between the heating and cooling balance point temperatures (indicated via vertical dashed lines). In this case, Recurve estimates baseload with the hourly usage of only those days for which the CalTRACK daily model assigns no temperature dependence to usage.

In the middle scenario, very few days are present with mean temperature between the heating and cooling balance point temperatures (between 0 and 20% of days). In this case, Recurve computes baseload using those days and then pulls in days with temperatures closest to the midpoint between balance points until 20% of all days are included. This corresponds to a minimum of 73 days in a 365-day period. Finally, in the right hand scenario, the heating balance point temperature is equal to the cooling balance point temperature. This can occur when buildings have simultaneous heating and cooling loads. In these cases, Recurve computes baseload with the 20% of days with mean temperature closest to the heating/cooling balance point.

For the other CalTRACK daily models (cooling only and heating only), a similar process is followed of calculating baseload where the daily model outputs indicate no temperature dependence and/or using the 20% of days closest to the balance point temperature.

Because the CalTRACK daily model serves an essential purpose in the computation of baseload, and therefore variable load, Recurve has modeled the partial year consisting of the

COVID period (March 19 - May 8). Again 20% of days are taken as a minimum to compute baseload, now corresponding to 10 days. The change in variable load is then calculated as the difference between variable load computed for the baseline period and COVID-period respectively.

Appendix B: NAICS Group Roll Ups and Meter Counts

Administrative/Civil

813312 Environment, Conservation-Wild
813410 Civic and Social Organizations
813910 Business Associations
813930 Labor Unions-Similar Labor Org
813990 Other Similar Organizations (no
921190 Other Genl Government Support

Automotive

423120 Auto Supply-Part Whsle
441100 Automobile Dealers
441110 New Car Dealers
441120 Used Car Dealers
441228 Motorcycle, ATV, and All Other Motor Vehicle Dealers
441310 Automtv Parts -Accessory Store
441320 Tire Dealers
488410 Motor Vehicle Towing
811100 Automotive Repair Maintenance
811111 General Automotive Repair
811121 Automtv Body, Paint-Interior R
811198 Other Automtv Repair-Maintenance

Banks

522110 Commercial Banking
522120 Savings Institutions
522130 Credit Unions

Beauty

812113 Nail Salons
812112 Beauty Salons
812199 Other Personal Care Services
812111 Barber Shops

Churches

813110 Religious Organizations

Construction/Contractors

230000 Construction
231531 General Contractor (4)
236110 Residential Bldg Construction
236115 Sgl-Fam Home Bldr-exc Op Bldrs
236118 Residential Remodelers
236220 Commercl&Instit Bldg Consruct
237210 Land Subdivision
237310 Highway/Street/Bridge Construc
238000 Specialty Trade Contractors
238150 Glass and Glazing Contractors
238160 Roofing Contractors
238210 Electrical Contractors
238220 Plumbing, HVAC Contractors
238320 Painting&Wall Cover Contractor
238330 Flooring Contractors
238990 Othr Specialty Trade Contractr

Fitness

713940 Fitness-Rec Sports Centers
611620 Sport & Recreation Instruction

Gas Stations

447100 Gasoline Stations
447110 gas Stations with Convenience
447190 Other Gasoline Stations

Grocery/Convenience Stores and Pharmacy

445100 Grocery Stores
445110 Supermarkets-Othr Grocery (not
445120 Convenience Stores
445310 Beer, Wine, and Liquor Stores
446110 Pharmacies and Drug Stores

Hotels/Lodging

721191 Bed-and-Breakfast Inns
721110 Hotels (except Casino)-Motels

Medical Offices

621210 Offices of Dentists
621310 Offices of Chiropractors
621320 Offices of Optometrists
621340 Ofcs of Physical, Occupational
621399 Ofcs of Othr Health Practrs
621400 Outpatient Care Centers
621491 HMO Medical Centers
621492 Kidney Dialysis Centers
621498 Other Outpatient Care Centers
621111 Ofcs of Physicians (not Mental

Offices

524200 Agencies, Brokerages-Othr Insu
541110 Offices of Lawyers
541210 Account Tax Preparation, Bookk
541211 Ofcs Certif Public Accountants
541213 Tax Preparation Services
541310 Architectural Services
541320 Landscape Architectural Services
541330 Engineering Services
541410 Interior Design Services
541430 Graphic Design Services
541511 Custom Computer Program Svcs
541512 Computer Systems Design Svcs
541611 Administrative Mgmnt-Gen'l Mgm
541618 Other Mgmnt Consulting Svcs
541620 Environmental Consulting Svcs
541690 Othr Sci-Tech Consult Svc
541850 Display Advertising
561110 Office Administrative Services
561439 Othr Business Service Centers
561499 Other Business Support Service
561710 Exterminating-Pest Control Svc
561720 Janitorial Services
561730 Landscaping Services
561990 All Other Support Services
611119 Unif School District Ofc

Real Estate

530000 Real Estate & Rental & Leasing

531000 Real Estate
531110 Lessors of Resid Bldgs-Dwelllin
531120 Lessors of NonResid Bldgs (not
531122 Offices Single/Multi Tenant(4)
531130 Lessors of Miniwarehouses-Self
531150 Comb Real Est Devlpr Contr(4)
531210 Ofcs of Real Estate Agent-Brkr
531310 Real Estate Property Managers

Restaurants/Bars

722410 Drinking Place (Alcoholic Bev)
722511 Full-Service Restaurants
722000 Food Services-Drinking Places
722513 Limited-Service Restaurants
722515 Snack and Nonalcoholic Beverage Bars

Retail

448110 Men's Clothing Stores
447190 Women's Clothing Stores
448140 Family Clothing Stores
448210 Shoe Stores
446120 Cosmetics, Beauty Supply-Perfu
448190 Other Clothing Stores
442110 Furniture Stores
446130 Optical Goods Stores
453310 Used Merchandise Stores
448100 Clothing Stores
448310 Jewelry Stores
451120 Hobby, Toy, and Game Stores
453320 Gift, Novelty,-Souvenir Stores
453110 Florists
451110 Sporting Goods Stores
451211 Book Stores
452319 All Other General Merchandise Stores.
453220 Gift, Novelty,-Souvenir Stores
453920 Art Dealers
453991 Tobacco Stores
453998 Othr Misc Store Retailers (not
453910 Pet and Pet Supplies Stores
531121 Otr Shop Centrs - Retl Sale(4)
442210 Floor Covering Stores
442299 Other Home Furnishings Stores

443142 Electronics Stores
444130 Hardware Stores
444190 Other Building Material Dealer
444220 Nursery and Garden Centers

Schools

611111 Elementary School - Private (4)
611113 Secondary Schools - Private (4)
611691 Exam Preparation and Tutoring

Warehousing and Postal

491110 Postal Service
493100 Warehousing and Storage
493110 General Warehousing & Storage
493190 Other Warehousing and Storage

Appendix C: Stratification Parameter Definitions

Parameters in blue indicate metrics that require hourly AMI data.

Usage Magnitude Metrics - The following parameters represent total building usage during a specified period of time (annual, summer peak etc.), or for a certain type of load (cooling, baseload etc.).

- **Annual_kWh** - Total Annual kWh usage
- **Summer_kWh** - Total Summer kWh usage (Summer = June - September)
- **Summer_Peak_kWh** - Total Summer peak kWh usage (Summer Peak = June - September, 4 - 9 pm)
- **Winter_kWh** - Total Winter kWh usage (Winter = January, February, November, December)
- **Winter_Morn_kWh** - Total Winter Morning kWh usage (Winter Morning = January, February, November, December, 6 - 10 am)
- **Shoulder_kWh** - Total Shoulder kWh usage (Shoulder = March - May and October)
- **Shoulder_Midday_kWh** - Total Shoulder Midday kWh usage (Shoulder Midday = March - May and October, 9 am - 3 pm)
- **Cooling_kWh** - A customer's temperature-dependent usage (warm weather) as determined by CalTRACK daily methods
- **Heating_kWh** - A customer's temperature-dependent usage (warm weather) as determined by CalTRACK daily methods

- **Baseload_kWh** - A customer's average baseload (always on) consumption. This is computed by first finding days where the CalTRACK daily model does not detect temperature-dependent usage. For these days hours are rank-ordered according to kWh usage. The average usage of hours ranked 2 - 4 are then averaged across the year and extrapolated to 8,760 hours.
- **Discretionary_kWh** - An estimate of a customer's non-temperature dependent, non-baseload usage that is independent of temperature.

Specific Load Shape Characteristics - The following parameters are intended to gauge specific features in a customer's load shape

- **Evening_Ramp** - The difference in a customer's average hourly usage during hour 18 (6 - 7 pm) compared to hour 14 (2 - 3 pm), averaged across the year.
- **Evening_Ramp_Ratio** - The ratio between a customer's average Evening Ramp and average hourly usage
- **Shoulder_Midday_RestofDay_Ratio** - The ratio between a customer's Shoulder-month usage during the midday hours to the remaining Shoulder-month usage

Normalized Metrics - These parameters are dimensionless quantities that can highlight how and when a customer is using energy relative to their total total (or seasonal) consumption.

- **Pct_Summer_Peak** - The percentage of a customer's annual usage that occurred during summer peak
- **Pct_Winter** - The percentage of a customer's annual kWh usage that occurred in the winter
- **Pct_Winter_Morn** - The percentage of a customer's Winter usage that occurred during the morning hours
- **Pct_Baseload** - The percentage of a customer's annual usage that is comprised of baseload
- **Pct_Discretionary** - The percentage of a customer's annual usage estimated to be discretionary
- **Pct_Cooling** - The percentage of a customer's annual usage determined by CalTRACK daily to be temperature dependent (warm weather)
- **Pct_Heating** - The percentage of a customer's annual usage determined by CalTRACK daily to be temperature dependent (cold weather)

- **Summer_Shoulder_Ratio** - The ratio of a customer's shoulder period kWh usage to summer kWh usage. For this filter Shoulder = April, May, October.