Memorandum

To: San Joaquin Valley Data Gathering Project Working Group

From: Opinion Dynamics

Date: December 18, 2019

**Re:**   SJV DAC Data Gathering – Sampling Plan

Introduction

The San Joaquin (SJV) Disadvantaged Communities (DACs) Data Gathering Plan will collect baseline data to inform an economic feasibility study of various interventions intended to reduce energy costs and mitigate the use of fuels such as propane, wood, and diesel generators by residential customers residing in designated DACs in the SJV. In support of the Data Gathering Plan, Opinion Dynamics will conduct surveys, audits, and in-person interviews with households in 168 non-pilot DACs in the SJV. This document describes our sampling approach for these tasks.

Overview

The target population for this sampling plan is residential households (excluding multifamily homes from buildings with five or more units) located in the 168 DACs in the SJV where PG&E, SCE, and SoCalGas are not conducting pilot programs. The primary data collection activities will utilize a nested sampling design (Figure 1). We will initially complete quantitative surveys with 2,500 customers. We will complete in-home audits with 280 of the 2,500 customers who completed the survey. We will complete in-home interviews with 60 of the 280 customers who participate in the audit portion of the study.

Figure 1. Nested Sampling Approach

Quantitative Survey (n=2,500)

In-Home Audits (n=280)

In-Person Interviews (n=60)
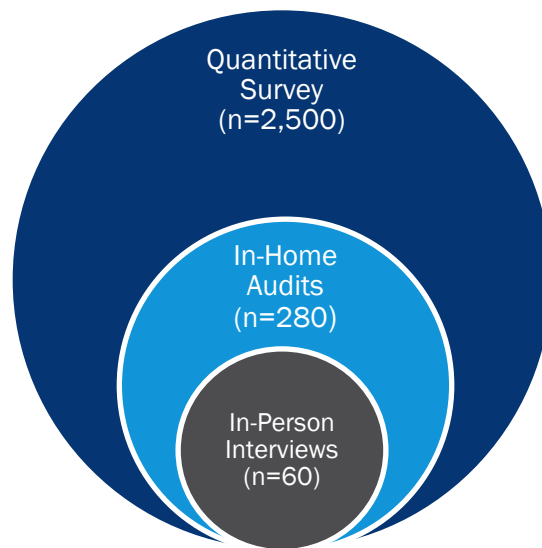
Profile of Households in San Joaquin Disadvantaged Communities

Before detailing the sampling approach, it is important to understand how the target population varies across metrics that are critical to the research effort.[1] Of particular importance is ensuring sufficient representation of homes that lack access to natural gas. Per CPUC Decision 18-08-019, the primary purpose of the Data Gathering Plan: "Is to collect the information needed to establish baseline conditions in the identified communities and to support an analysis of the economic feasibility of extending affordable energy options to these communities, in particular to dwellings that currently lack access to natural gas." The following describes population parameters that must be considered when developing a sampling plan that fulfills the data needs outlined in the Data Gathering Plan.

As seen in Figure 2, nearly all DACs have at least some access to natural gas. Only 7 DACs have no access at all and 29 have "low access" (where 44%-79% of households have access).

Figure 2. Number of DACs by Natural Gas Service Level
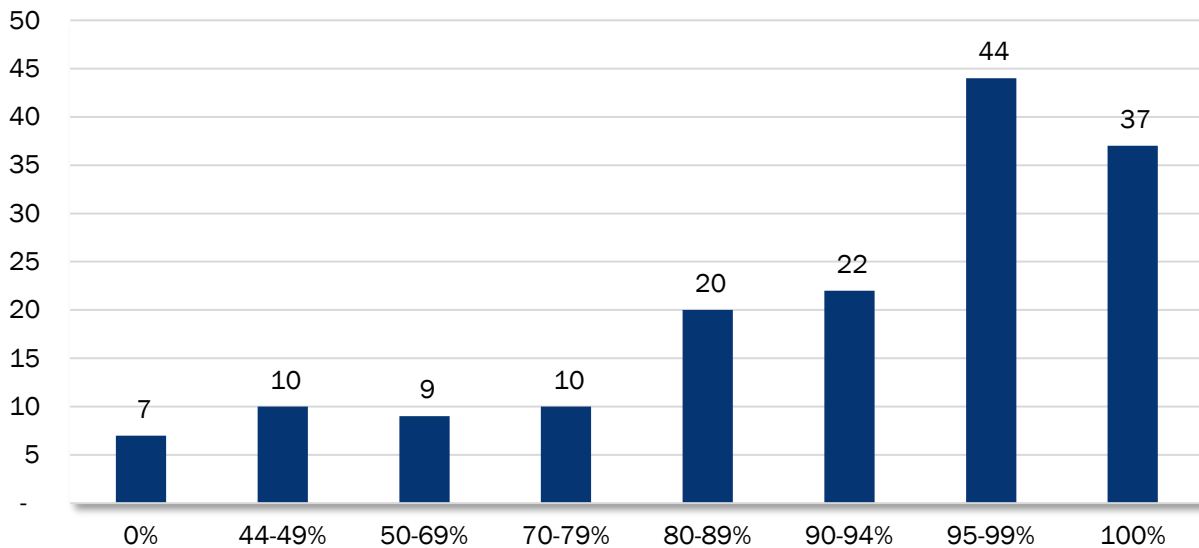
Further, access to natural gas becomes much more lopsided when we look at household-level access (Figure 3). According to the DAC list, the vast majority of homes in SJV DACs have natural gas access.

---

[1] For development of this sampling plan, our analysis included 159 of the 168 non-pilot communities. Due to insufficient data, we excluded the 9 non-pilot communities that were added the DAC list in 2018. Once this data gap is remedied, we will update our sampling approach accordingly.

Figure 3. Number of Households by Natural Gas Service Level
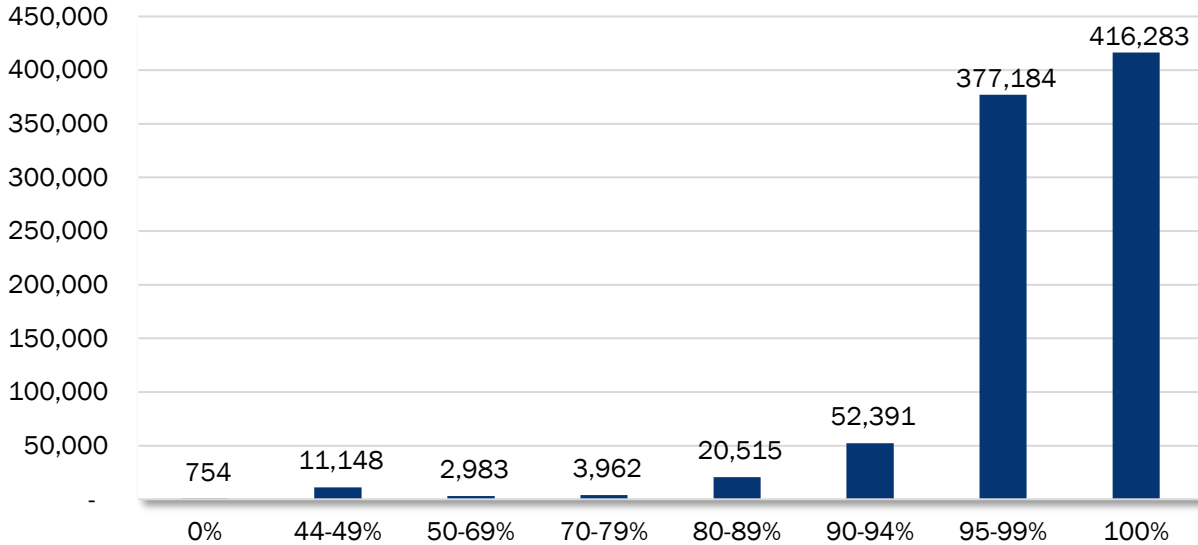
Given the disproportionate access to natural gas across and within SJV DACs, as well as the Decision's focus on "the expense incurred by low-income households that lack access to natural gas and must rely on electricity, propane or wood for space and water heating," we will use household-level fuel access as the primary sampling stratification criteria. Namely, we will rely on these two strata: 1) households with natural gas access (homes with a natural gas account from PG&E or SoCalGas), and 2) households without natural gas access (homes without a natural gas account from PG&E or SoCalGas).

According to the Decision, community size should be used as a "secondary grouping characteristic," as "the size of a community may constrain the economic viability of future energy options such as community solar." As seen in Table 1, 83 of the 159 SJV DACs are small communities [with less than 1,000 households]. Although small communities make up over half of all SJV DACs, households in small communities are only 3% of all households across the 159 SJV DACs. Thus, if we were to take a simple random sample of the population, we would achieve a sample with mostly households from large communities. In order to see representation of small communities, we will need to target community size in our sampling design.

Table 1. Distribution of Communities and Households by Community Size

Community Size	Communities		Households	
	N	%	N	%
Small	83	52%	28,341	3%
Medium	60	38%	183,974	21%
Large	16	10%	672,905	76%
Total	159	100%	885,220	100%

Relatedly, housing type varies by community size: small communities have more mobile homes, medium and large communities have more single family attached homes (Table 2).[2]

Table 2. Distribution of Households by Community Size and Housing Type

Community Size		Single Family - Detached	Single Family - Attached	Mobile Home
Small	%	73%	7%	20%
	N	22,565	2,260	6,128
Medium	%	82%	12%	5%
	N	147,571	21,898	9,727
Large	%	66%	31%	3%
	N	429,184	203,706	16,971

Quantitative Survey Sample Design

We will complete quantitative surveys with 2,500 residential households (excluding multifamily homes)[3] located in the 168 DACs in the SJV where PG&E, SCE, and SoCal Gas are not conducting pilot programs. Since there is considerable variation across communities and households, we will use a stratified random sampling approach.

As seen in the tables above, households in SJV DACs are predominantly natural gas-heated single-family homes located in large communities with "high" natural gas access. Thus, a simple random sample would result in few mobile homes, homes without access to natural gas, or households from small or medium DACs. A stratified random sample helps resolve this lopsidedness, thereby ensuring data collection fulfills the Data Gathering Plan objectives. Specifically, we will under-sample households with natural gas access, completing a maximum of 400 surveys with households with natural gas service to meet research objectives that focus on comparisons of households that use different fuel types (Table 3). We will complete the remaining surveys (2,100) with households lacking natural gas service.

Table 3. Stratified Random Sampling Design by Household Fuel Access

Household Fuel Access	Survey Sample Size Targets
Natural gas access	400
No natural gas access	2,100
Total	2,500

In addition, to ensure residents of small communities are represented, we will further stratify by community size and oversample households in small communities. This sample design ensures sufficient data in all

---

[2] Note that these census estimates exclude multifamily homes, which are out of scope for the Data Gathering Plan.
[3] The working group decided to exclude multifamily homes located in buildings with five or more units for two reasons: 1) they are a low incidence group that constitute a very small proportion of all residential households in SJV DACs, and 2) they are unlikely to use fuels other than natural gas or IOU-provided electricity, and are thus not the focus of this study.

community size groupings, thereby enabling statistical comparisons.[4] Table 4 details our overall sampling design.

Table 4. Stratified Random Sampling Design by Household Fuel Access and Community Size

Community Size	Household Fuel Access	Survey Sample Size Targets
Small	Natural gas access	150
	No natural gas access	700
Medium	Natural gas access	125
	No natural gas access	700
Large	Natural gas access	125
	No natural gas access	700

In-Home Audit Sampling

In-home audits will be sampled from the 2,500 survey respondents, targeting 280 completed audits. Due to smaller sample sizes, we will not cross stratification criteria (e.g.: targeting a certain number of audits from households without natural gas access within small communities). Instead, we will have multiple isolated stratification criteria that work in parallel (e.g.: the audit targets by natural gas access will be independent of targets for various community sizes).

Since homes not served by natural gas are of particular importance to the Data Gathering Plan, we will oversample homes lacking natural gas access. Table 5 exhibits our audit sampling approach by natural gas access.

Table 5. Projected Number of Audits by Household Fuel Access

Household Fuel Access	Audit Sample Size Targets
Natural gas access	50
No natural gas access	230
Total	280

---

[4] Further, this approach also provides adequate representation of mobile homes, as mobile homes are disproportionately located in small communities. Master metered mobile home parks present a unique challenge in that IOU address-based sampling will not effectively reach residents of these parks. Following further review of the incidence of mater metered mobile home parks, we may suggest a specialized, smaller quantitative survey effort with these residents.

For audit sampling purposes, we will collapse medium- and large-sized communities into a single stratum, targeting 155 completed audits across communities of either size (Table 6). We will target 125 completed audits with households from small communities, which is near maximum feasibility for that community size.

Table 6. Projected Number of Audits by Community Size

Community Size	Audit Sample Size Targets
Small	125
Medium/Large	155
Total	280

We will also set audit quotas for housing type (Table 7). Given the nested sampling approach, our survey sampling strategy (namely, oversampling small communities, which have a higher incidence of mobile homes) helps us ensure sufficient mobile home audit sample (as there is no reliable way to glean home type from IOU address data, so we must rely on achieving sufficient mobile home sample via the aforementioned oversamples). As previously mentioned, we will not be sampling multifamily households as they are not the focus of the study.

Table 7. Number of Audits by Housing Type

Housing Type	Audit Sample Size Targets
Single-family detached	180
Single-family attached	50
Mobile homes	50
Total	280

In-Person Interview Sampling

Continuing with the nested sample design, we will recruit a random stratified sample of 60 households for in-person interviews from the audit sample. Ideally, in-person interviews will take place during the audit site visit. Due to the overarching research questions for this study and the limited number of interviews being conducted, we will focus our interview sample on households without natural gas access. Table 8 exhibits our interview sampling plan, which is stratified by natural gas access. We will not have explicit housing type targets, but we will ensure sufficient mobile home representation as we schedule interviews. Due to small sample sizes, we will not further stratify the interviews by community size.

Table 8. Number of Interviews by Household Fuel Access

Household Fuel Access	Interview Sample Size Targets
Natural gas access	10
No natural gas access	50
Total	60

Confidence and Precision

The quantitative survey sampling plan will satisfy 95/2 confidence/precision for the overall SJV non-pilot DAC household population, at least 95/5 for each natural gas access/community size grouping, and 90/5 for alternative fuel households overall.[5] The audit sampling plan will satisfy 90/5 confidence/precision for the overall SJV non-pilot DAC household population and at least 90/12 for each sub-group of interest. Table 9 provides some examples of absolute precisions we would achieve when estimating proportions for select sample sizes similar to some outlined in this plan.

Table 9. Examples of Absolute Precisions by Various Sample Sizes (Proportion Estimation at 90% Confidence)

Proportion	Survey Sample Size			Audit Sample Size	
	n=750	n=350	n=275	n=95	n=45
0.1 or 0.9	0.018	0.026	0.030	0.051	0.074
0.2 or 0.8	0.024	0.035	0.040	0.068	0.098
0.3 or 0.7	0.028	0.040	0.045	0.077	0.112
0.4 or 0.6	0.029	0.043	0.049	0.083	0.120
0.5	0.030	0.044	0.050	0.084	0.123

---

[5] Minimum absolute precisions we would achieve when estimating proportions. Absolute precision estimations assume a proportion of 0.5.