Methodology Report

Sugar Alert Study

Public Health Institute

September 2020

Table of Contents

I. Introduction

The Public Health Institute partnered with ICF International (hereafter, "ICF") to administer the Sugar Alert Study.[1] The purpose of the study was to collect data for evaluating San Francisco's sugar sweetened beverage advertising warning label ordinance. The study consisted of a baseline and three follow-up studies. The baseline study was conducted by telephone using random digit dialing techniques applied to sampling frames of landline and cell telephone exchanges. Concerns about low response rates—coupled with growing challenges related to traditional telephone surveys led PHI to change the sampling methodology and mode of administration. Data collection for the baseline study was via telephone and subsequent studies utilized a "push-to-web" data collection method, in which sampled households were sent an invitation via mail, text, and/or E-mail to complete an online Web questionnaire. Additional web completes were collected using a non-probability web panel. Questionnaires were completed in English, Spanish, or Chinese.

Table 1 provides the dates and modes for each wave of data collection.

Table 1. Data Collection Dates and Modes

Data collection wave	Dates	Data collection mode for newly recruited respondents	Data collection mode for repeat respondents
2016 Baseline	2016-2017	Telephone (landline and cell)	N/A
2017 First Follow-up	2017-2018	Mail-push-to-Web, Web Panel	Text, E-mail, and Mail-push-to-Web
2018 Second Follow-up	2018-2019	Mail-push-to-Web, Web Panel	Text, E-mail, and Mail-push-to-Web
2019 Third Follow-up	2019-2020	Mail-push-to-Web	Text, E-mail, and Mail-push-to-Web

This report details the methodology for the baseline and all follow-up studies.

[1] This project was carried out in compliance with International Standards Organization Research Standard 20252.

II. Sample Design

The target population for the Sugar Alert Study was adults living in households located in the Cities of San Francisco and San Jose. During the baseline study, an overlapping dual frame landline and cellular random digit dial (RDD) sample design was implemented. The dual-frame covered households with at least one landline telephone or at least one cell phone. Adults living in phoneless households were not covered by the dual-frame sample. Follow-up studies utilized three sample sources:

1) **Follow-up:** We followed up with respondents who agreed to participate during a past year,
2) **ABS Recruits:** Address-based sampling was utilized to recruit new respondents and replace those who did not agree to the follow-up studies, and
3) **Web Panel Recruits:** A non-probability web panel was also utilized to recruit new respondents.

In total, 9,904 completed questionnaires were collected. Table 2 provides the number of completed interviews by city and sample type.

Table 2. Completed Interviews by City and Study, and Sample Type

	Baseline		First Follow-up			Second Follow-up			Third Follow-up	
	Landline	Cell	Follow-up	ABS	Panel	Follow-up	ABS	Panel	Follow-up	ABS
San Francisco	402	876	306	401	618	818	356	38	916	283
San Jose or other	343	859	299	480	510	795	350	53	899	302
Total	745	1,735	605	881	1,128	1,613	706	91	1,815	585

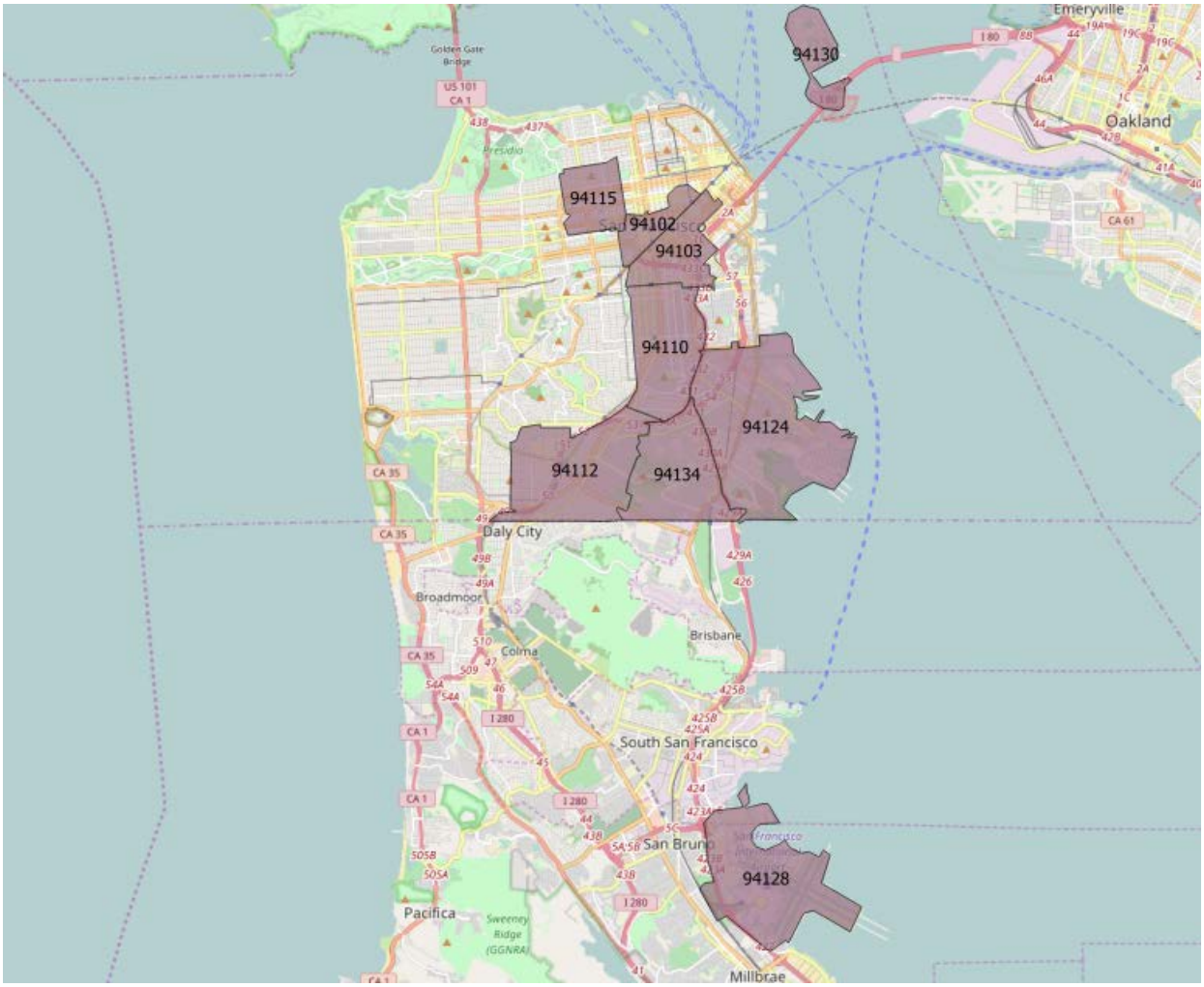
The study oversampled African American and Hispanic respondents. Table 3 presents the number of interviews achieved by city and race/ethnicity.

Table 3. Completed Interviews by City and Race/Ethnicity

City	African American	Hispanic	Total Completes
San Francisco	516	702	5,014
San Jose	191	892	4,890
Total	707	1594	9,904

The sample selection was targeted to zip codes with high African American and/or high Hispanic populations in both cities. The zip codes for the sample and the stratification are presented in Exhibits 1 and 2.

Exhibit 1. Zip Codes for Sampling in San Francisco

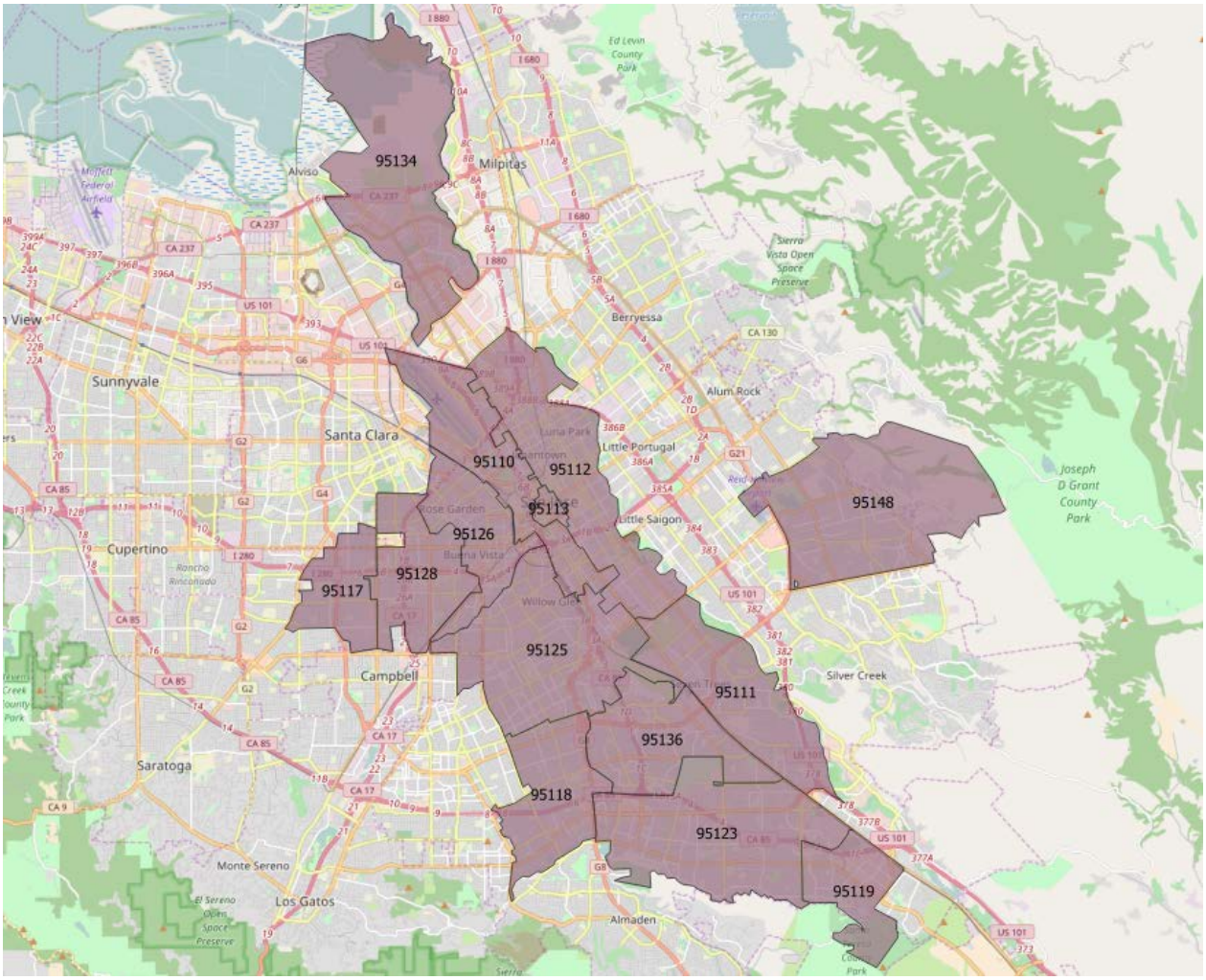


Stratum	Definition	Zip Codes
1	Zip codes with 10% or more African American	94124, 94128, 94130, 94102, 94115
2	Zip codes with <10% African American and 30% or more Hispanic	94103, 94134, 94110, 94112

In San Francisco, 5.5% of the population is African American and 15% of the population is Hispanic.[2] To achieve the race/ethnicity targets, the zip codes in San Francisco were selected based on having at least 10% African American population or at least 30% Hispanic Population. These zip codes covered an estimated 67% of the African American population and 63% of the Hispanic population. Zip codes that did not qualify based on high African American or high Hispanic populations were not included in the sample.

[2] American Community Survey, 2010-2014.

Exhibit 2. Zip Codes for Sampling in San Jose



Stratum	Definition	Zip Codes
1	Zip codes with 4% or more African American	95113, 95126, 95117, 95128, 95134, 95119
2	Zip codes with <4% African American and 35% or more Hispanic	95112, 95110, 95123, 95148, 95111, 95118, 95136, 95125

In San Jose, 2.9% of the population is African American and 33% of the population is Hispanic.[3] To achieve the race/ethnicity targets, the zip codes in San Jose were selected based on having at least 4% African American population or at least 35% Hispanic Population. These zip codes covered an estimated 67% of the African American population and 53% of the Hispanic population. Zip codes that did not qualify based on high African American or high Hispanic populations were not included in the sample.

[3] American Community Survey, 2010-2014.

1. Landline Sample – Baseline Study

The landline sample for the baseline study was a list-assisted sample based on the set of telephone exchanges associated with the selected zip codes. First, the exchanges where at least 40% of the directory listed households are geographically located in zip codes were included in the frame. Exchanges where less than 40% of the directory listed households are in the zip codes (meaning 60% are outside the zip codes) were discarded. Second, the exchange was assigned to a geographic stratum based on the one containing most of the listed households.

Similarly, the list-assisted landline RDD frame for San Jose was defined as the set of telephone exchanges associated with the selected zip codes. Exchanges where at least 40% of the directory listed households are geographically located in zip codes were included in the frame. Exchanges where less than 40% of the directory listed households are in the zip codes were discarded.

After identifying the telephone exchanges in each area, all possible telephone numbers were then divided into blocks (or banks) of 100 numbers.[4] Zero-blocks, or 100 blocks without any residential assignments, were excluded from the sampling frame. The remaining 100-blocks, those with at least one residential assignment (or 1+ blocks), were assigned to one of three strata based on the zip code. 1,000-blocks of telephone numbers were associated with zip codes by tallying the number of geocoded landline households in each zip code. The 1,000-block was assigned to the zip code with the most number of geocoded telephones.

All possible telephone numbers, both listed and unlisted, in 1+ blocks are eligible for selection through RDD with equal probability within the assigned stratum.

The landline sample was generated by ICF using Marketing Systems Group (MSG) Genesys software. The frame and sample sizes are presented Tables 5-7 in the weighting section.

2. Cell Phone Sample – Baseline Study

The cell phone sample for the baseline study was selected from a frame of cell phone numbers whose billing address is located in one of the target zip codes. This sample was selected by Survey Sampling International's (SSI) SmartCell®. SmartCell includes the billing address from the wireless provider. All cell phone users with a billing address in the target zip codes is eligible for selection. The frame and sample sizes are presented in Tables 5-7 in the weighting section.

[4] A hundred block is a set of 100 telephone numbers with the same area code, prefix, and first two digits of the suffix. A 1+block is a 100 block with at least one telephone number that is assigned to a residence. A 0-block is a 100 block with no residential assignments.

3. County Supplements – Baseline Study

During fielding of the baseline study, three supplement samples were selected to evaluate the operational efficiency of the targeted frames. Landline samples were selected in San Francisco and San Jose and a cell phone sample was selected in San Francisco. These samples were city wide and not targeted to the high minority zip codes.

The landline sample was based on telephone exchanges associated with all zip codes in San Francisco and San Jose. The cell phone sample was based on telephone thousand-blocks associated with rate centers located in Santa Francisco County.

4. Address Based Sample (ABS) - New Recruits

The source of the ABS frame for new recruits to the follow-up studies was the Computerized Delivery Sequence File (CDSF), a list of addresses that originates from the U.S. Postal Service. The frame included all residential addresses including city-style addresses, P.O. boxes,[5] and rural-route addresses. To maximize coverage of the population, our sampling frame included units identified by the USPS as central drop points, seasonal, and vacant.

For each wave, the frame and sample sizes for the strata in San Francisco and San Jose are presented in Tables 8-10 in the weighting section. San Francisco zip codes that were at least 4% African American were oversampled at a rate of 4:1 relative to the Hispanic zip codes. San Jose zip codes that were at least 10% African American were oversampled at a rate of 2:1 relative to the Hispanic zip codes.

5. Screening

The Sugar Alert study had the following eligibility requirements. Respondents were (1) adults aged 18 or older, and (2) residents of San Francisco or San Jose. Respondents who completed the interview in prior years also confirmed that they were the person who completed that interview, confirmed their age, and which community they live in.

III. Questionnaire Content

The Sugar Alert questionnaire was designed by PHI. Each question was selected based upon the defined objectives of the study. The survey content was divided into the following 14 sections:

1. Introduction and Screening
2. Beverage Intake Questionnaire
3. Personal and Household Characteristics
4. Health Consciousness

[5] We only include P.O. Boxes that are classified as the only way for a household to get mail (OWGM). Ninety-one percent of P.O. Boxes are registered to households that also receive residential mail delivery.

5. Warning Labels
6. Knowledge of Sugar Sweetened Beverage Health Risks
7. Attitudes towards Sugar Sweetened Beverages
8. Sugar Sweetened Beverage Intent
9. Attitudes towards Sugar Sweetened Beverage Policy
10. Race/Ethnicity
11. Education
12. Travel in and Out and Shopping
13. Income
14. Closing

IV. Data Collection Protocol

Data collection was conducted across four waves in 2016 through 2020.

Table 4. Data Collection Dates

Data collection wave	Dates
Baseline	November 2016 to March 2017
First Follow-up	December 2017 to January 2018
Second Follow-up	November 2018 to March 2019
Third Follow-up	November 2019 to January 2020

Data collection procedures for each wave are described in detail below.

1. Interviewing Protocol - Baseline Study

The baseline study was a computer-assisted telephone interview (CATI). A minimum of 15 call attempts were made for landline sample records and a 7 call attempts were made for cell sample records.

1.1 Landline Sample

The 15-attempt landline protocol is executed over three day-parts: weekdays, weekday evenings, and weekends, along with a set of anytime attempts.

- Monday - Friday: 9:00 a.m. - 5:00 p.m.: 3 attempts
- Monday - Friday: 5:00 p.m. - 9:00 p.m.: 3 attempts
- Saturday (10:00 a.m. - 9:00 p.m.) & Sunday (10:00 p.m. - 9:00 p.m.): 3 attempts
- Remaining 6 attempts at most productive time

1.2 Cell Sample

- The 7-attempt cell protocol was executed over the following dayparts:
- Monday - Friday: 9:00 a.m. - 5:00 p.m.: 2 attempts
- Monday - Friday: 5:00 p.m. - 9:00 p.m.: 2 attempts
- Saturdays & Sundays (10:00am – 9:00pm): 3 attempts

All times shown are respondent local time (Pacific Time).

Each sample record received a final outcome, or disposition for the survey. This final disposition was attained when:

- The respondent completed the interview;
- The telephone number was found to be invalid;
- The record reached the maximum number of required attempts per the protocols (as outlined above); or
- The respondent gave a final refusal.

The average interview length was 27.6 minutes.

2. Contacting Respondents - Baseline Study

The following protocols were followed when contacting households and potential respondents for the baseline study:

Treatment of No Answers. If a call to a sampled telephone number was not answered, the number was repeatedly called at different times, during daytime and evening hours, on different days of the week, in a pattern designed to maximize the likelihood of contact with a minimum number of calls. At least 15 contact attempts, over a 30-day period, were made to reach a sampled number. Once any contact was made at a residence, as many calls as necessary were made to reach the selected adult (within the permitted wave of data collection). For cell phone, the treatment of no answers was the same as that for landline, but was limited to seven call attempts.

Rings per Attempt. The telephone rang a minimum of five times on each attempt made on a record.

Busy Lines. Traditionally, a busy signal indicates a respondent may shortly be available to take a call. Therefore, busy lines were called back at least twice at 10-minute intervals. If the line was still busy after the third attempt, the number was assigned a "busy" disposition and called during the next shift.

Landline Respondent Selection. Once a household was contacted, an adult was selected for participation in the study. No interview was conducted if:

The adult was:

- Unavailable during the survey period;
- Unable or unwilling to participate; or,
- Did not speak English or Spanish well enough to be interviewed.

Or, a randomly sampled number yielded:

- A business;
- An institution/group quarters;
- A number belonging to a minor, teen or child; or,
- An other strictly non-residential space.

Language of Interviewing. Interviewing for the baseline study was conducted in English, Spanish, and Chinese.

Handling Refusals. Protocol for the Sugar Alert Study required two refusals to terminate the record from calling. Once a household or selected respondent initially refused participation, specially trained conversion interviewers contacted them, at least three days later, to encourage participation in the survey and if a second refusal was received, the record was dispositioned as a final refusal.

Leaving Messages. Answering machine, or voicemail messages, were left on the 1st, 4th and 9th attempts. The standardized voicemail message left for a potential survey respondent was:

> "Hello, my name is _____. We're doing a study on family nutrition in the Bay Area, funded by the National Institutes of Health. Qualified participants who complete the survey will receive a $10 amazon.com gift code. We will call again in the next few days to conduct

the interview. If you have any questions or would like to call us to take the survey, please call us at 1-844-212-7823 at your convenience. Thanks."

CallerID and IVR. An IVR (interactive voice response) line was maintained and staffed by ICF interviewers during the course of the study. This included a specific-toll free information line, 844-212-7823, to which respondents were directed to call back. This line was staffed with trained Sugar Alert Study interviewers who were available to answer respondent's questions. PHI provided ICF with a local, San Francisco-based telephone number that was displayed by ICF on each outbound call placed; the local telephone number was 510-379-1826. If the local 510-number was dialed back by a survey respondent, they were connected to our Sugar Alert Study IVR line, as described above.

3. Data Collection Protocol - Follow-up Respondents

Data collection protocol for respondents who agreed to participate during a previous study consisted of a series of successive communications via mailed letters, texts, and emails based on the respondent's preferred mode. All letters, texts, and emails were sent in the respondent's preferred language (English, Spanish, or Chinese). The letter included the URL for the survey and unique ID. The text and emails included a unique URL. Follow-up respondents were also sent a thank-you note after participation with the URL for a website where they could update their contact information between data collection waves.

4. Data Collection Protocol - ABS

Data collection protocol for ABS consisted of an invitation letter in English, Spanish, and Chinese, followed by a reminder postcard, and a final reminder postcard. The letter and postcards included the URL for the survey and a unique ID.

5. Data Collection Protocol - Non-Probability Web Panel

For the follow-up studies, we partnered with Dynata (formerly SSI), a sampling company which also provides mobile panel sample. A non-probability sample was selected from Dynata's current panel members. For each wave of data collection, ICF provided targets by city, age group, race and gender. Dynata then send invitations to complete the survey, filling the most difficult quotas first, and then opening invitations to respondents who fit the remaining age, gender, and race quotas. Respondents were redirected to ICF's programmed web questionnaire via unique web links.

6. Incentives

Respondents were offered a $10 Amazon.com gift code for completing the first questionnaire, $15 for the second, $20 for their third and fourth completion. In addition, $1 was included in mailed invitations sent to ABS records.

V. Weighting

Survey weights were computed for the baseline and the follow-up studies to correct for disproportionate sampling probabilities introduced by the sampling design and to correct for

differences in demographic characteristics of the sample versus the population, reducing the risk of nonresponse and coverage biases in substantive estimates that may be associated with those demographics. The combined weights for each wave that weights the sample to the total population of adults (18+) living in the target zip codes in San Francisco and San Jose.

1. Baseline Study

The baseline study was a dual-frame RDD sample so we calculated the weights in three steps: 1) calculating cell and landline design weights, 2) combining the cell phone and landline samples, and 3) population calibration (i.e. poststratification and raking).

Calculating Cell and Landline Design Weights

The first stage in the weighting involved creating sampling weights that correct for disproportionate probabilities of selection, *design weights*. The design weight for a sampling unit is the inverse of the probability that the particular unit is drawn into the sample.

Selection of the Telephone Number

The landline phone sample was selected by drawing n_L landline phone numbers from N_L numbers on the frame for each of the strata. The sample selection probability for stratum *s* was calculated as $\Pr(L)=(n_L/N_L)$ and the base weight as the inverse of the probability of selection, W1 = 1/Pr(*L*). The calculations are shown in Table 5.

Table 5: Base Weights for the Landline Sample

		Landline frame NL	Sample	Selection probability Pr(L)	Base weight W1
San Francisco	Zip codes with 10% or more African American	69,800	13,672	0.1959	5.11
	Zip codes with <10% African American and 30% or more Hispanic	339,500	39,871	0.1174	8.51
San Jose	Zip codes with 4% or more African American	59,500	19,774	0.3323	3.01
	Zip codes with <4% African American and 35% or more Hispanic	422,000	42,553	0.1008	9.92

The cell phone sample was selected from SSI's SmartCell, with the frame including all cell phone numbers with a billing zip code in one of the target zip codes. The landline phone sample was selected by drawing n_C landline phone numbers from N_C numbers on the frame for each of the strata. The sample selection probability for stratum *s* was calculated as $\Pr(C)=(n_C/N_C)$ and the base weight as the inverse of the probability of selection, W1 = 1/Pr(*C*). The calculations are shown in Table 6.

Table 6: Base Weights for the Cell Phone Sample

		Cell frame N_c	Sample	Selection probability Pr(C)	Base weight W1
San Francisco	Zip codes with 10% or more African American	77,155	11,000	0.1426	7.01
	Zip codes with <10% African American and 30% or more Hispanic	164,780	13,982	0.0849	11.79
San Jose	Zip codes with 4% or more African American	91,130	13,250	0.1454	6.88
	Zip codes with <4% African American and 35% or more Hispanic	268,225	10,250	0.0382	26.17

Selection of Household Member

For the landline sample, one member (18+ adult) from each household was randomly selected to take the survey. To account for the within household selection probability, we multiplied the weight by the number of eligible adults in the household (A). The number of adults was capped at 2 to reduce weight variability. There is no within household selection for the cell phone.

Combining the Cell Phone and Landline Samples

The sample design was a fully overlapping landline and cell phone dual frame, meaning those who have a landline and cell phone are eligible to be selected via either sample. To account for the overlap of dual-users selected in the cell sample and the dual-users selected in the landline sample, we use a composite weight.

The two samples are averaged based on a composite weight designed to optimize the variances of weighted estimates. The composite weight is a ratio of the effective sample sizes, c = $neff_1$/ ($neff_1$+ $neff_2$), where $neff = n/deff$ is the effective sample size; $deff = n \times \sum w_i^2 \times \left[\sum w_i\right]^{-2}$ is a measure of variability of respondent level weights (w_i) and n is the sample size for the survey. The landline design weight is multiplied by c, where 0 < c < 1 and the cell phone design weight by 1-c. Before averaging the landline and cell samples, we adjust each individually to match the estimated number of cell-only and landline population based on the estimated cell-only percentage, 44% in San Francisco and 45% in San Jose, from Marketing Systems Group (MSG). The MSG cell-only estimate is calculated by subtracting the estimated landline households from the estimated telephone households.

Table 7: Distribution of Landline and Cell Users

		Sample Size		MSG Population Estimate
San Francisco		Landline	Cell Phone	
	Cell-only		455	44%
San Jose	Landline	401	428	56%
	Cell-only		417	45%
	Landline	344	435	55%

Population Calibration

As the final weighting step, we post-stratified the combined sample into demographic categories and ratio adjusted the weights so that the final weighted sample matches the population with respect to those demographic characteristics. We used a raking algorithm that iteratively calibrates the weighted sample to the population on these dimensions:

1) City(SF; SJ) * Age (18-24; 25-34; 35-44; 45-54; 55-64; 65+) * Gender;
2) City(SF; SJ) * Gender * Education (Less than high school; HS grad; Some coll/Associates; Coll Grad; Grad school);
3) City(SF; SJ) * Race/ethnicity (Hispanic; non-Hisp white; non-Hisp black; non-Hisp Asian, non-Hisp all other);

The population controls are based on data from the 2015 American Community Survey (ACS) Summary File. The population totals represent the adult household population living in the target zip codes in San Francisco and San Jose.

2. First Follow-up Study

The sample sources for the first follow-up include repeat responses from the baseline (conducted as RDD telephone) as well as newly recruited cases from an ABS push-to-web and from an online web panel. The steps for combining and weighting the three sample sources include computing base weights for each sample source, calibrating each sample source to the demographic distributions of the San Francisco and San Jose, and combining samples were combined into one data file.

Base Weights

The base weight for interviews from the baseline was equal to the final weight from the baseline study. This weight retained the corrections for disproportionate sampling probabilities from the baseline study.

The base weight for interviews from the non-probability web panel was 1.

The base weight for ABS recruits was the inverse of the selection probability. The ABS sample was selected by drawing n_A landline phone numbers from N_A numbers on the frame for each of the strata. The sample selection probability for stratum *s* was calculated as Pr(A)=(n_A/N_A) and the base weight as the inverse of the probability of selection, W1 = 1/Pr(*A*). The calculations are shown in Table 8.

Table 8: Base Weights for the ABS Sample

		ABS frame N_A	Sample	Selection probability Pr(A)	Base weight W1
San Francisco	Zip codes with 10% or more African American	49,663	4,214	0.0849	11.79
	Zip codes with <10% African American and 30% or more Hispanic	77,461	3,286	0.0424	23.57
San Jose	Zip codes with 4% or more African American	59,499	4,863	0.0817	12.24
	Zip codes with <4% African American and 35% or more Hispanic	129,006	2,637	0.0204	48.92

Population Calibration

Each sample was post-stratified into demographic categories the base weights were ratio adjusted so that each weighted sample matched the population with respect to those demographic characteristics. The calibration for the repeat responses represents an adjustment for attrition from the baseline to the first follow-up. For the new samples, the calibration is an adjustment for differential response to the survey based on age, gender, race/ethnicity and educational attainment.

The same raking methodology as used in the baseline and first follow-up was used in the second follow-up sample calibration. The data was based on data from the 2016 American Community Survey (ACS). Due to small sample sizes in the 18-24 age range, non-Hispanic Black race group, and less than high school education, we collapsed cells.

Combining samples

The RDD and ABS samples were averaged by city based on a composite weight designed to optimize the variances of weighted estimates. The composite weight is a ratio of the effective sample sizes, c = $neff_1$/ ($neff_1$+ $neff_2$), where *neff* = *n*/*deff* is the effective sample size; $deff = n \times \sum w_i^2 \times \left[\sum w_i\right]^{-2}$ is a measure of variability of respondent level weights (w_i) and n is the sample size for the survey. The RDD weight was multiplied by c, where 0 < c < 1 and the ABS weight by 1-c.

This process was repeated to average the combined RDD/ABS sample with the panel sample. A composite weight c_1 was computed as the ratio of the effective sample sizes for the RDD/ABS and panel samples, The RDD weight was multiplied by c_1, where $0 < c_1 < 1$ and the panel weight by $1-c_1$.

3. Second Follow-up Study

The sample sources for the second follow-up include repeat responses from the baseline (conducted as RDD telephone), repeat responses from the ABS sample, as well as newly recruited cases from an ABS push-to-web and from an online web panel. The steps for combining and weighting the three sample sources include computing base weights for each sample source, calibrating each sample source to the demographic distributions of the San Francisco and San Jose, and combining samples were combined into one data file.

Base Weights

The base weight for interviews from the 2016 baseline was equal to the final weight from the baseline study. This weight retained the corrections for disproportionate sampling probabilities from the baseline study.

Similarly, the base weight for the ABS respondents entering the study in 2017 is their 2017 final weight.

The base weight for the new 2018 respondents and the repeat respondents from the non-probability web panel was 1.

The base weight for 2018 ABS recruits was the inverse of the selection probability. The ABS sample was selected by drawing n_A landline phone numbers from N_A numbers on the frame for each of the strata. The sample selection probability for stratum *s* was calculated as $Pr(A)=(n_A/N_A)$ and the base weight as the inverse of the probability of selection, $W1 = 1/Pr(A)$. The calculations are shown in Table 9.

Table 9: Base Weights for the ABS Sample

		ABS frame N_A	Sample	Selection probability Pr(A)	Base weight W1
San Francisco	Zip codes with 10% or more African American	50,253	3,060	0.0609	16.4
	Zip codes with <10% African American and 30% or more Hispanic	78,477	2,390	0.0305	32.8
San Jose	Zip codes with 4% or more African American	59,711	3,531	0.0591	16.9
	Zip codes with <4% African American and 35% or more Hispanic	129,826	1,919	0.0148	67.7

Population Calibration

The design has samples for 3 years and 3 modes. Each sample was separately post-stratified into demographic categories the base weights were ratio adjusted so that each weighted sample matched the population with respect to those demographic characteristics. The calibration was done separately for:

- 2016 RDD sample
- 2017 ABS sample
- 2018 ABS sample
- 2017 and 2018 panel (combined)

The same raking methodology as used in the baseline was used in the first follow-up. As with the baseline, the data was based on data from the 2015 American Community Survey (ACS). Due to small sample sizes in the 18-24 age range, non-Hispanic Black race group, and less than high school education, we collapsed cells.

Combining samples

The sample design was fully overlapping, meaning those who have a landline or cell phone were eligible to be selected via either the baseline RDD sample or the ABS sample in 2017 or 2018. Furthermore, panel members would have also been eligible for selection in the baseline RDD or ABS samples. The steps for combining the data together are:

1. Combine ABS 2017 and 2018: The two ABS samples were averaged by city based on a composite weight designed to optimize the variances of weighted estimates. The composite weight is a ratio of the effective sample sizes, c = $neff_1$/ ($neff_1$+ $neff_2$), where $neff = n/deff$ is the effective sample size; $deff = n \times \sum w_i^2 \times \left[\sum w_i\right]^{-2}$ is a measure of variability of respondent level weights (w_i) and n is the sample size for the survey. The 2017 ABS weight was multiplied by c, where 0 < c < 1 and the 2018 ABS weight by 1-c.
2. Combine the ABS sample with the RDD: The RDD and the 2017/2018 combined ABS samples were averaged by city using the same composite adjustment as described in step 1.
3. Combine the ABS/RDD with panel: Finally, the combined ABS/RDD samples were averaged with the panel respondents by city using the same composite adjustment as described in step 1.

4. Third Follow-up Study

The sample for the 2019 Sugar Alert was selected from respondents who responded to the 2018 survey supplemented with a new sample selected from ABS. The weights for the 2019 survey were calculated in three steps. First, we calculated base weights for the 2019 ABS sample as the inverse of the probability of selection. Second, we adjusted the 2018 repeat respondents for attrition by calibrating to match demographic distributions of the San Francisco and San Jose. We also calibrated the new ABS sample to match the demographic distributions. Finally, the repeat respondents from the 2018 sample were combined with the 2019 sample.

Base Weights

The base weight for 2019 ABS recruits was the inverse of the selection probability. The ABS sample was selected by drawing n_A landline phone numbers from N_A numbers on the frame for each of the strata. The sample selection probability for stratum *s* was calculated as $Pr(A)=(n_A/N_A)$ and the base weight as the inverse of the probability of selection, W1 = 1/Pr(*A*). The calculations are shown in Table 10.

Table 10: Base Weights for the 2019 ABS Sample

		ABS frame N_A	Sample	Selection probability Pr(A)	Base weight W1
San Francisco	Zip codes with 10% or more African American	50,574	3,257	0.0644	15.5
	Zip codes with <10% African American and 30% or more Hispanic	78,804	2,543	0.0323	31.0
San Jose	Zip codes with 4% or more African American	60,248	3,758	0.0624	16.0
	Zip codes with <4% African American and 35% or more Hispanic	130,693	2,042	0.0156	64.0

Population Calibration

The 2019 ABS respondents and the 2018 repeat respondents were separately post-stratified into demographic categories the ratio adjusted so that each weighted sample matched the population with respect to those demographic characteristics. The calibration was done separately for:

- 2018 Repeat respondents (start with 2018 final weight)
- 2019 ABS respondents (start with base weight)

The same raking methodology as used in previous studies was used in the third follow-up sample calibration. The population total were based on data from the 2017 American Community Survey (ACS). Due to small sample sizes in the 18-24 age range, non-Hispanic Black race group, and less than high school education, we collapsed cells.

Combining samples

As a last step, we combined the weighted 2019 ABS sample and the weighted 2018 Repeat sample. We combine the samples for each city by averaging based on a composite weight designed to optimize the variances of weighted estimates. The composite weight is a ratio of the effective sample sizes, c = $neff_1$/ ($neff_1$+ $neff_2$), where $neff = n/deff$ is the effective sample size; $deff = n \times \sum w_i^2 \times \left[\sum w_i\right]^{-2}$ is a measure of variability of respondent level weights (w_i) and

n is the sample size for the survey. The 2017 ABS weight was multiplied by c, where 0 < c < 1 and the 2018 ABS weight by 1-c.

VI. Response Rates

ICF assigned records to American Association for Public Opinion Research (AAPOR) defined disposition categories. Table 11 indicates the disposition categories and their definitions.

Table 11. Categories of Disposition Codes

Category	Description
I	Completed interviews.
R	Refusals and breakoffs. This category contains eligible respondents who refused to complete the survey or started the survey but ended before it was completed.
NC	Non-contact. This category contains respondents who were identified as eligible but were unable to be contacted during the duration of fielding.
O	Other eligible. This category contains eligible respondents but a language barrier or impairment prevented completion of the interview.
UH	Unknown if household. This category contains phone for which it was not possible to determine household status.
UO	Unknown other. This category contains phone numbers confirmed to be assigned to households, but it was impossible to determine whether an eligible adult lived there.

The response rate for the follow-up records and the ABS Recruits was calculated using AAPOR's RR3 formula. RR3 represents the percentage of completions among all eligible records in the sample, including records known to be eligible and an estimate of eligibility among records where status is unknown. The formula is:

$$RR3 = \frac{I}{(I) + (R + NC + O) + e(UH + UO)}$$

Where e represents the estimated proportion of eligible records where actual eligibility status cannot be determined. It is computed as the percent of eligible records divided by the sum of eligible and ineligible records.

Table 12. Disposition Categories and Response Rate

	Baseline		First Follow-up		Second Follow-up		Third Follow-up	
	Landline	Cell	Follow-up	New ABS Recruits	Follow-up	New ABS Recruits	Follow-up	New ABS Recruits
Interviews (I)	745	1,735	605	881	1,613	706	1,815	585
Refusals/ Breakoffs (R)	440	794	46	42	0	31	0	20
Non-Contact (NC)	43	111	0	0	0	0	0	0
Other (O)	19	18	0	0	0	0	0	0
Unknown if HH (UH)	24,496	41,844	1,552	14,535	2,191	9,834	2,191	12,747
Unknown Other (UO)	78	0	0	0	0	0	0	439
Ineligible (X)	86,808	7,221	0	0	0	0	0	0
e	0.014	0.269	1	1	1	1	1	1
Response Rate (RR3)	**46.7%**	**12.5%**	**28%**	**6%**	**42%**	**7%**	**42%**	**4%**