

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

B.1. Respondent Universe and Sampling Methods

The proposed survey will include a nationally representative survey of households within the United States. Eligible respondents must be 18 years of age or older, and be considered one of the heads of the household. The questions asked in the survey require knowledge regarding duration of residency, the age of the house and the equipment installed within the house. This necessitates that the respondent be a person who makes major decisions within household. A head of household will be considered a person living or staying in the home in whose name the house or apartment is owned, being bought, or rented. A probabilistic multistage sampling approach will be utilized to select sample units for this survey. A probabilistic sampling method will allow a random selection of units with a calculable probability of selection of each unit in the target population.

Considering the three-stage sampling approach, the primary sampling units are metropolitan areas, the second sampling units are US Census Tracts, and the elementary sampling units are US Households. At the first stage, we considered US metropolitan areas as primary sampling units (PSU), and select a random sample of 24 metropolitan areas among the 389 at the first stage (<https://www2.census.gov/programs-surveys/metro-micro/geographies/reference-files/2015/delineation-files/list1.xls>, accessed on 12/04/2017 at 3:04PM). Then, we considered US Census Tracts as secondary sampling units (SSU). More precisely, we selected a random sample of Census tracts in each of the 24 metro areas selected at the first stage, as well as a random sample of Census tracts in non-metropolitan areas in the proximity of metro areas. Lastly, we selected a random sample of occupied households through a random walk methodology in each Census tract selected at the second stage for survey fielding.

The end-objective is to ensure that the ultimate survey sample includes a majority of completes from housing units within metropolitan areas and a quantifiable minority from non-metro areas. The respondents' universe is, therefore, a random sample of US households selected from a random sample of Xensus tracts selected from a random sample of US metropolitan areas.

B.1.1 Sampling Frame

The sampling frame will consist of occupied housing units within metropolitan and non-metropolitan areas. Records from the latest housing surveys from the US Census Bureau will be used to identify Metropolitan Areas and Census Tracts.

Table 3. Summary of Sampling Stages and Respective Sampling Frames

Sampling Stage	Sampling Unit	Population Size	Sample
# 1	US Metropolitan Areas	388	24
# 2	US Census Tracts	74,002	192
# 3	US Housing Units	118,860,065	1,055

The original sampling methodology utilized an address-based sample (ABS) list of residential addresses to determine its sample frame. The revised methodology employs a random walk door-to-door knocking strategy to pre-notify and recruit households in each Census tract. As such, any occupied household within a chosen tract is eligible to participate in the study. Research teams

will select different parts of a tract to focus their recruitment efforts ensuring a mix of areas/neighborhoods, household types and demographic characteristics of respondents recruited to participate.

B.1.2 Sampling Approach

EurekaFacts team will adopt a proportional multistage sampling approach to select housing units for the SCOA survey. To ensure that housing units in metropolitan areas as well as those not in metropolitan areas are included in the survey, we consider the following steps:

1. At the first stage, a random sample of 24 metropolitan areas were selected among the 388 Census tracts as primary sampling units. The sample was stratified by Census Region and then stratified by metro area population size (those with a population of 1 million or more and those with less than 1 million), ensuring the number of Primary Sampling Units (PSUs) selected for each region is proportionate to the number of occupied housing units (OHUs).
2. At the second stage, a random sample of residential census tracts was selected in proportion to the number of OHUs within each of the 24 metropolitan areas selected at the first stage. Furthermore, a random sample of additional census tracts within non-metropolitan areas located within the same state of each PSU were selected at this stage.
3. At the third stage, a random walk door-to-door methodology is instituted in each Census tract so field interviewers can recruit residents of occupied housing units for the in-home survey (full length and shortened versions). Each tract has a target quota in relation to the OHUs of other tracts randomly selected.

The end-objective is to ensure that our survey sample has a majority of completes from housing units within metropolitan areas and a quantifiable minority from non-metro areas. Each tract has a target quota in relation to the other tracts randomly selected. The fourth stage of sampling would require selection of additional tracts to either replace or supplement tracts where recruitment is far below the threshold of the tract's quota. For example, if a tract has too many empty units or residents with a language barrier, a new tract would be randomly selected from tracts of a similar size.

In total, this approach makes it possible to calculate the probability of selection of every sample units at every stage, as well as to reliably calculate design effect, sampling error, etc. and infer the findings to the housing units in the US with a calculable level of precision.

B.1.3 Sample Size

CPSC has available funding for a sample size of 1,185 households. This sample size is large enough to provide accurate estimates representing the diversity of housing and household types, attitudes and state laws on smoke alarms and CO alarms across the country. This sample size is large enough to yield a margin of error of +/-2.83% at 95% level of confidence for estimates of the survey results in a simple random sample.

The margins of error provided below are only indicative and based on assumptions of a statistical power of 80% (usual default value), a confidence interval of 95% and a target population size of

116,900,000 units¹¹. The actual margin of error that will be provided with survey results may slightly differ from the ones shown here, as the actual multistage sampling design will provide a slightly different margin of error overall. The specific margin of error for multistage sampling will be calculated when the actual sample is drawn.

Table 4. Margin of Error by Sample Size

Sample Size (N)	1,055
Statistical Power	80%
Confidence Interval	95%
Total Occupied Housing Units	116,900,000
Margin of Error (SRS*)	3.0%

*SRS: Simple Random Sample

For a total sample size of 1,055 households the sample distribution will be:

- Metropolitan Areas: 24
- Census Tracts per Metropolitan Area: Proportion to the number of OHUs within each of the 24 metropolitan areas for a total of 192 tracts
- Housing Units per Census Tracts in Metropolitan Areas: Minimum of 2; quota per tract will be proportionate to the combined OHUs of the other tracts to reach target completes
- Housing Units in non-Metropolitan Area: Minimum of 2; quota will be proportionate to the combined OHUs of the other non-metro areas to reach target completes
- Total Housing Units in Metropolitan Areas: 995
- Total Additional Housing Units in non-Metropolitan Areas: 190

B.2. *Procedures for the Collection of Information*

The original survey methodology employed a mixed-mode, multi-stage approach to data collection. The combination of mailed pre-notification letters, multiple calls to gauge interest and screen for eligibility, additional calls to coordinate and confirm the interview date, and finally the arrival of the survey team at a designated time proved to be too complex of a recruitment procedure. The drop off in interested participants increased with each step, which hurt overall response and cooperation rates.

The revised process condenses the timeline between initial contact and potential participation in the study. Once a Census tract is randomly chosen, random areas of the tract are selected for field teams to then randomly walk and place door hangers on residences. These door hangers act as a pre-notification to the household that mentions the study’s purpose, incentive, and website/phone number to learn more. Additional details from the original pre-notification letter are also now available online on a publicly available EurekaFacts website. The posting contains a detailed explanation of the study and its objectives, who are the sponsors, details on how their responses and PII will be kept secure, FAQ, and links to the posting of the Federal Register.

11. U.S. Census, *Quick Facts: Population Estimates*. (2016), <https://www.census.gov/quickfacts/table/PST045216/00>

In the following one to three days after respondents have received their pre-notification door hanger notifying them of the study, field teams will follow the same walking route and knock on doors to recruit for immediate participation in the study. Field teams will have the pre-notification letter, copy of the door hanger, and their official badges visible to provide credibility to our survey effort and encourage respondents' cooperation. Overall, this reduces the number of total contacts made with the public and diminishes burden. This also allows field teams to have more meaningful, efficient, and prompt interactions with the community.

If someone is interested in participating in the survey, the field team can immediately screen them at their door. Depending on the status of smoke alarm installation or its type, eligible respondents will be engaged to participate in either the full-length survey interview or shortened survey version.

Residents who have a smoke alarm, that is not connected to a central or security alarm that will notify the police or fire department, will be eligible for full-length interviews. If, however, they do not have a smoke alarm, or, if they do, but it is connected to a central or security alarm, the respondent will be eligible to participate in the shortened survey version. The survey instrument will be programmed on Qualtrics and will be administered via in-home interviews using a Computer Assisted Personal Interview (CAPI) format.

During the in-home interviews, a qualified two-member survey team will ask household residents questions related to smoke and CO alarms. The full-length interview will include testing of some (potentially all depending on the number) of their smoke and CO alarms. If any of the alarms do not work, we will offer to provide a new one to them free of charge. If, however, residents do not have a smoke alarm, they will receive a shorter version of the questionnaire. The survey interview will take between 20 - 60 minutes, depending on if the shortened version of the survey is administered (about 20 minutes) or the full-length survey and alarm testing interview is conducted (no more than 60 minutes). The individual data will not be identified to a specific person. Any data provided to the client or included in the report will be delivered in the aggregate form.

B.2.1 Statistical Methodology for Stratification and Sample Selection

The survey's sample selection, and sampling methodology is discussed earlier in Question 1. A probabilistic multistage sampling approach will be utilized to select sample units for this survey. A probabilistic sampling method will allow a random selection of units and an equal chance for inclusion in the survey. Considering that there are more than 134 million housing units in the nation, a simple random sampling, a systematic sampling, a stratified sampling, or a cluster sampling will be too costly and less likely to capture the diversity of the smoke alarms and CO alarms adoption, operability, and local jurisdictions laws and regulations. The proposed proportional multistage sampling approach will consider metropolitan areas as primary sampling units, residential census tracts within metropolitan areas as secondary sampling units, and housing units within those residential census tracts as the final sample units. At the first two stages, a simple random sampling will be applied to select units. The third stage utilizes a random walk methodology as a more practical strategy for gaining completes, yet all households in a tract are eligible to participate. This approach makes it possible to calculate the probability of selection of every sample units at every stage. As such, we will be able to reliably calculate

estimates and sampling error and infer the findings to the housings in the US with a calculable level of precision.

The American Community Surveys data from the U.S. Census Bureau was used to identify and quantify the number of U.S. Metropolitan Areas, the number of US Census Tracts overall and within each metropolitan area, and finally, the number of housing units overall and within each census tract. Those entities—Metropolitan Areas, Census Tracts and Housing Units—are considered as defined by “U.S. Census Bureau, Population Division, based on Office of Management and Budget, July 2015 delineations.”

B.2.2 Estimation Procedure

Estimates will be produced using standard survey estimation procedures. Survey estimates include estimates of operability of smoke alarms and CO alarms, estimates of percentages of households as well as subgroups with installed of smoke alarms and CO alarms, estimates of the proportions of respondents demonstrating hazard awareness. EurekaFacts will consider and compare different methods of variance estimation such as Taylor series approximation or various replication (Jackknife, Balanced Repeated Replication) methods.

B.2.3 Unusual Problems Requiring Specialized Sampling Procedures

We adopted an on-the-ground door-knocking sampling approach over the multi-stage mixed-mode approach originally designed. As previously described, the extremely low response rate made the methodology impractical for completing the study in a timely and efficient manner. The new design is ultimately more streamlined and provides a cost-effective, time-efficient, and flexible strategy for field teams to recruit for in-person research that requires researchers to be in the participant’s home. The differences in efficiency are demonstrated in the contrast between the fielding attempts using the original methodology vs. the revised methodology.

Table 5. Comparison of Original and Revised Sampling Designs

Methodology	Fielding dates	Total weeks	Response rate	Cooperation rate	Completes to quota ratio
ORIGINAL: Address-based sampling multi-mode recruitment approach	Jan. 1 - May 30, 2019	23 weeks	.09%	3.0%	15% (9 completes/ 59 quota)
REVISED: Door-to-Door household random walk method sampling approach	Dec. 21, 2019 - March 1, 2020	11 weeks	3.5%	17.4%	101% (130 completes/ 128 quota)
Change		-12 weeks	+3.41%	+14.4%	+86% pts.

Additionally, vacant households are excluded from this survey. The U.S. Census Bureau estimates that 12% of housing units in the US are vacant. Field teams will skip leaving door

hangers and knocking on residences that are clearly abandoned/vacant to maintain efficiency during recruitment.

B.3. Methods to Maximize Response Rates and Deal with Non-response

To maximize response rates, EurekaFacts' approach includes multiple compounding efforts. First, a pre-notification door hanger will be left on the door of a sample of households in a randomly selected area of a tract. The door hanger will briefly provide the purpose of the research, sponsors of the research, a telephone number and a website providing more detailed information and an FAQ. The door hanger is in public view of all the neighbors, so residents can see that it is a neighborhood-wide canvas and not just their home that was selected. This allows time for neighbors to discuss the legitimacy of the study, which builds confidence in residents when field teams approach their household to recruit them. Dropping a unique piece of literature followed by researchers at their doors just a few days later is designed to enhance response and cooperation rates, and thus data quality. EurekaFacts will also create social media posts advertising the study and target their appearance by ZIP code based on dates field teams will be in that area.

The two-member survey team will take additional steps to persuade residents to participate in the study. Both members will present their government-issued IDs and their official badges to confirm their identity and legitimacy. They will also be wearing high-visibility safety vests, so their presence is well announced to everyone in the neighborhood. The survey team interviewers will be trained in refusal conversion techniques to reassure the respondents of the legitimacy of the team's presence and gain cooperation. The team will also carry with them a letter printed on official letterhead with endorsements from the local fire department and/or CPSC, should they be needed.

In the SCOA survey, two types of non-response may occur. The first type is unit non-response, which occurs when data is not obtained for the sample unit (*i.e.*, a respondent chooses not to participate in the survey). The second type is item non-response, which occurs when a respondent fails to answer one or more of the survey questions. For unit non-responses, EurekaFacts anticipate a response rate below 80% and will therefore conduct a non-response bias analysis. We will use households' characteristics selected from the American Community Surveys data of the US Census Bureau to assess whether there exists a significant difference between households who responded to the survey and those who did not. If the non-response analysis reveals the necessity to weight the survey data, we plan to use a weighting scheme that makes respondents representative of their respective metropolitan area, as oppose to an overall weighting scheme. Information about households that do not participate in the study (time of day, tract, reason for not participating, demographic information about person who answered door – if applicable) will be captured on a screen-out capture form on the researcher's tablet. This will allow for tracking of the total number of residents spoken to, capturing screen-out data, and analyzing non-response by demographics and reasons given not to participate.

Since the survey is administered by interviewer, item non-response has low likelihood. Nonetheless, a survey entry will be considered as complete only when 80% of the survey questions have been answered, skip patterns excluded. Based on the initial non-response analysis

the appropriate approach to handling missing data will be identified and whether simple or multiple-imputation is required.

B.4. Test of Procedures or Methods to be Undertaken

The current survey content was developed by Vision 20/20 and Tridata, LLC, which specialize in fire safety and engineering, and reviewed by specialists with expertise with fire and CO safety, including staff from fire departments, Red Cross, subcontracted engineering firms, and CPSC staff. EurekaFacts reviewed instruments to ensure adherence to survey design guidelines and ensure an efficient and accurate data collection process which would minimize overall administration length and reduce the burden placed both on the respondents and the survey team. Three survey methodologists and two statisticians reviewed the survey and provided recommendations for ensuring relevance, clarity, and minimal response burden.

The full survey instrument underwent testing via cognitive interviews with 18 household residents (OMB Control Number 3041-0136). Cognitive testing was carried out to ensure that any questions that were misunderstood by respondents or that were difficult to answer would be improved prior to the survey fielding, and thus increase the overall quality of survey data and the accuracy of the study results. The respondents for the cognitive interviews were recruited using a slightly modified version of the survey screener (modified to meet the needs of recruiting for cognitive interviews). Recruitment was conducted using a mixed-mode methodology that combined posting on community websites such as Craigslist, and telephone outreach to follow-up specific leads generated by the advertising. The interview sessions lasted for approximately ninety minutes. The cognitive interviews were conducted with two groups of participants: 1) individuals who report having a smoke alarm that is not connected to a central alarm that may notify the police or fire department, and 2) individuals who do not have a smoke alarm installed, or if they do, their smoke alarm is connected to central or security alarm. Similarly, as planned during the actual survey fielding, during the cognitive pretesting the first group (with smoke alarm) was administered the full in-house version of the survey instrument and the second group (no smoke alarm or central alarm) was administered the telephone version of the survey. The cognitive interviews examined how well the questions performed and ensured that the material was clear and easy to understand among potential survey respondents. As part of this effort, the cognitive testing was designed to assess the question-response process in terms of respondent's comprehension, information retrieval process, judgment as to providing requested information, and perceived degree of ease or difficulty experienced in formulating accurate/correct responses to each question posed.

Overall the cognitive interviews demonstrated that the survey instrument did not pose any considerable challenges to the respondents. A majority of the survey questions were clear and easy to understand, and the response categories for multiple-choice questions were relevant. The testing identified minor misunderstanding or different ways of interpreting a few terms. Based on the pretesting results, participants' comments and suggested improvements of the question wording were identified and implemented. By utilizing this method prior to survey fielding, this helped to increase the overall quality of the survey data and the accuracy of the study results.

EurekaFacts continued to adjust and refine the instrument with CPSC to make it appropriate for fielding. After collecting the first 9 completes using the original methodology, EurekaFacts evaluated the overall process and proposed the change to the door-to-door random walk

methodology. A non-substantive change request to OMB was submitted by CPSC, and EurekaFacts adopted the new method with OMB approval.

After the first round of recruitment and data collection using the new methodology, EurekaFacts found no major issues and continued with the data collection effort. After the first 50 completes were collected, a brief analysis of selected questions was conducted to ensure data quality and instrument functionality; no changes were needed. Additionally, an internal debrief was conducted and lessons learned from those initial interviews were incorporated into the rest of the data collection effort and highlighted in the pilot report. All data collected will be incorporated into the overall sample of 1,055 cases.

B.5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Consulted on statistical aspects of the design and to collect and analyze the data:

Jorge Restrepo (240) 403-1636

Robert Suls (240) 403-1641

Djass Mbangdadji (240) 403-1640

Bohdana Sherehiy, PhD (240) 403-1637

EurekaFacts, LLC

Arthur Lee (301) 987-2008

Matthew Brookman (301) 987-2467

Consumer Product Safety Commission