# Supporting Statement for Paperwork Reduction Act Submissions 

OMB Control Number: 1660-0105
Title: Community Preparedness and Participation Survey
Form Number(s): FEMA Form 008-0-15

## General Instructions

A Supporting Statement, including the text of the notice to the public required by 5 CFR 1320.5(a)(i)(iv) and its actual or estimated date of publication in the Federal Register, must accompany each request for approval of a collection of information. The Supporting Statement must be prepared in the format described below, and must contain the information specified in Section A below. If an item is not applicable, provide a brief explanation. When Item 17 or the OMB Form 83-I is checked "Yes", Section B of the Supporting Statement must be completed. OMB reserves the right to require the submission of additional information with respect to any request for approval.

## Specific Instructions

## B. Collections of Information Employing Statistical Methods.

When Item 17 on the Form OMB 83-I is checked "Yes", the following documentation should be included in the Supporting Statement to the extent it applies to the methods proposed:

1. Describe (including numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection has been conducted previously, include the actual response rate achieved during the last collection. The expected rate for the colleaction as a whole is $9 \%-10 \%$; The 2016 survey response rate was $9 \%$.

Federal Emergency Management Agency (FEMA) Individual and Community Preparedness Division (ICPD) will collect preparedness information from the public via a telephone survey. This collection of information, which began in 2007, is necessary to
increase the effectiveness of awareness and recruitment campaigns, messaging and public information, community outreach efforts, and strategic planning initiatives. The household telephone survey will measure the public's knowledge, attitudes, and behaviors relative to preparing for all hazards and for specific hazards selected from among the following: tornado, hurricane, flood, earthquakes, wildfire, extreme heat, winter storms and extreme cold, power outages, pandemic flu, hazardous materials, nuclear, and terrorism.

The potential respondent pool includes the entire civilian non-institutionalized U.S. adult population residing in telephone-equipped dwellings or owning a cell phone. This population does not include adults in penal, mental, or other institutions; adults living in dormitories, barracks, or boarding houses; adults living in a dwelling without a telephone; and/or adults who do not speak English or Spanish well enough to be interviewed.

The annual survey will consist of 5,000 respondents; 1,000 for the general nationwide survey and 500 for each of the eight hazard specific oversamples. Although most of the hazard specific sample populations will be defined by counties (or fips codes), some will be defined by zip codes. The specific selection of the hazard profiles to include in the surveys may vary across different administrations of the survey. The eight hazards selected for the 2017 survey include Flood, Hurricane, Tornado, Wildfire, Earthquake, Extreme Heat, Winter Storm, and all hazards requiring evacuation in urban areas.

The telephone samples will include both landline and cell phones to minimize bias in survey based estimates. In each administration, nine independent telephone samples will be chosen to generate the targeted number of surveys for the national and for each of the eight hazard areas. For each sample, the selection of landline numbers will be based on list-assisted RDD (Random Digit Dialing) sampling of telephone numbers for the corresponding geographic area. The cell phone sample will be a simple random sample drawn from all dedicated exchanges for cell phones for the targeted areas. For respondents reached on a landline phone, one respondent will be chosen at random from all eligible adults within a sampled household. For respondents reached on a cell phone, the person answering the call will be selected as the respondent if he or she is otherwise found eligible.

The goal will be to maximize the response rate by taking necessary steps as outlined later in this document on "Methods to maximize response rates." The calculation of response rates will be based on AAPOR RR3 definition.

## 2. Describe the procedures for the collection of information including:

## -Statistical methodology for stratification and sample selection:

In each administration of the telephone survey, about 1,000 interviews nationwide and 500 interviews for each of the eight selected hazard areas will be completed. Samples will be independently drawn for the national and for each of the hazard profile areas
defined based on complete counties. In order to minimize bias, both landline and cell phones will be included in the all telephone samples.

For the National sample, the target population will be geographically stratified into four census regions (Northeast, Midwest, South, and West) and sampling will be done independently within each stratum (region). The definition of the four census regions in terms of states is given below.

Northeast: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New Jersey, New York, and Pennsylvania.

Midwest: Illinois, Indiana, Michigan, Ohio, Wisconsin, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota.

South: Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia, Alabama, Kentucky, Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, and Texas.

West: Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming, Alaska, California, Hawaii, Oregon, and Washington.

The sample allocation across the four census regions (Northeast, Midwest, South and West) will be based on proportional allocation i.e. the sample size allocated to any particular region will be roughly in proportion to the size of that region in terms of the estimated number of adults. Using proportional sample allocation, the targeted number of surveys to be completed in each region is expected to be close to those proportions. Within each region, roughly 70 percent of the interviews will be done from the cell phone sample while the rest (30\%) will be done from the landline sample. It may be noted that the actual number of completed surveys for each census region (and by landline and cell phone strata within each region) will depend on observed response rates and so they may not exactly match the corresponding targets. However, the goal will be to meet those targets to the extent possible by constant monitoring of the response rates and by optimally releasing the sample in a sequential manner throughout the data collection period.

Within each region, the sampling of landline and cell phones will be carried out separately from the respective sampling frames. The landline RDD (Random Digit Dialing) sample of telephone numbers will be selected (without replacement) following the list-assisted telephone sampling method. For within-household sampling, the contractor will use the "next birthday" method to randomly select one eligible person from all eligible adults in each sampled household. Following the "next birthday" method, the interviewer asks to speak with the eligible person in the household who will be celebrating the next birthday. This is much less intrusive than the purely random selection method or grid selection that requires enumeration of all household members to make a respondent selection.

The cell phone sample of telephone numbers will be drawn (without replacement) separately from the corresponding dedicated (to cell phones) telephone exchanges. For respondents reached on cell phones, there will not be any additional stage of sampling (as there is with the within-household sampling for landline sample). The person answering the call will be selected for the survey if he/she is found otherwise eligible. For both landline and cell phones, the geographic location of the respondent will be determined based on respondent's self-reported response to a question on location (like "what is your zip-code?" and what is your county?). Data will be collected from all respondents regardless of whether they have access to a landline, a cell phone or both. Respondents will be asked a series of questions to gather information on his/her use of telephone (cell only, landline only, or dual-user cell mostly and other dual users).

As mentioned above, the cell phone numbers will be sampled from the telephone exchanges dedicated to cell phones while the landline numbers will be sampled from all area code exchange combinations for the corresponding geographic area. It may be noted that due to continuous porting of numbers from landline to cell and cell to landline, some numbers from landline exchanges may turn out to be cell phones and conversely, some numbers sampled from the cell phone exchanges may actually be landline numbers. However, such numbers will be relatively rare and the vast majority of landline and cell phone numbers will be from the corresponding frames. The survey will also find out from the respondents if the number called is actually a landline or a cell phone number. It is also possible that an individual respondent may have a telephone number in one region while he/she may actually be living in another region. The physical location of respondents will therefore be based on their self-reported location information (for example, based on their self-reported zip-code or county information) and will not be determined based on their telephone exchange.

The hazard area samples, as mentioned before, will be selected independently following procedures similar to those used for the national sample described above. The target population for each hazard area survey will consist of groups of counties (or zip-codes) identified based on specific requirements for each hazard. The total number of entities in the collection and the corresponding sample are listed in Table 1. The counties (fips) identified for each hazard used for 2017 are presented in Table 2 below as an example. Specific samples by hazard may be updated annually to maintain equivalency of the hazard impacted populations by including counties impacted by more recent events.

Table 1: Targeted number of completed surveys by profile

| Profile Number | Profile Name | Number of entities in <br> collection |
| :--- | :--- | :--- |
| Profile 0 | NATIONAL | 1000 |
| Profile 1 | TORNADO | 500 |
| Profile 2 | FLOOD | 500 |
| Profile 3 | HURRICANE | 500 |
| Profile 4 | HEAT | 500 |


| Profile 5 | WILDFIRE | 500 |
| :--- | :--- | :--- |
| Profile 6 | EARTHQUAKE | 500 |
| Profile 7 | WINTER STORM | 500 |
| Profile 8 | ALL HAZARD URBAN <br> AREAS | 500 |
| TOTAL |  | 5000 |

Table 2: Example of Sample - Flood

| Full FIPS | State | County |
| :---: | :---: | :--- |
| 48351 | TX | Newton County |
| 29155 | MO | Pemiscot County |
| 21195 | KY | Pike County |
| 19111 | IA | Lee County |
| 05017 | AR | Chicot County |
| 18019 | IN | Clark County |
| 17133 | IL | Monroe County |
| 37019 | NC | Brunswick County |
| 51131 | VA | Northampton County |
| 40089 | OK | McCurtain County |
| 39007 | OH | Ashtabula County |
| 47139 | TN | Polk County |
| 01069 | AL | Houston County |
| 20091 | KS | Johnson County |
| 36089 | NY | St. Lawrence County |
| 28059 | MS | Jackson County |
| 72147 | PR | Vieques Municipio |
| 54031 | WV | Hardy County |
| 31165 | NE | Sioux County |
| 27133 | MN | Rock County |

Hazard samples will also be stratified by census region and so each of the four strata for any hazard will consist of counties identified for that hazard in that particular census region. The sample allocation across strata will be proportional to the size of the stratum derived as the estimated total adult population of the counties selected for that hazard in that particular stratum. As proposed for national sample, both landline and cell phone numbers will be included in hazard area samples and the total number of completed surveys will be roughly split equally between the landline and cell phone samples. For
within household selection of respondent, the most recent birthday method will be used. For cell phone sample, the person answering the phone will be selected for interview as long as he/she is otherwise found eligible for the survey.

## -Estimation procedure:

Each of the nine samples (the national and the eight hazard samples) will be weighted independently. Once those weights are finalized, the sample data consisting of the national sample (1,000 completed surveys) and eight hazard area level surveys (500 completes each) will also be combined (composite weighting) and then weighted (poststratified) to generate estimates for unknown populations parameters at various levels (national, regional or for other subgroups of interest).

For the National sample, weighting will be carried out within each stratum (region) to adjust for (i) unequal probability of selection in the sample and (ii) nonresponse. Once the sampling weights are generated, weighted estimates can be produced for different unknown population parameters (means, proportions etc.) for the target population and also for population subgroups.

The weighting for this study will be done following the basic approach described in Kennedy, Courtney (2007): Evaluating the Effects of Screening for Telephone Service in Dual Frame RDD Surveys, Public Opinion Quarterly, Special Issue 2007, Volume 71 / Number 5: 750-771. In studies dealing with both landline and cell phone samples, one approach is to screen for "cell only" respondents by asking respondents reached on the cell phones whether or not they also have access to a landline and then interviewing all eligible persons from the landline sample whereas interviewing only "cell only" persons from the cell phone sample. The samples from such designs are stratified, with each frame constituting its own stratum. In this study, however, a dual-frame design is proposed where dual users (those with access to both landline and cell phones) can be interviewed in either sample. This will result in two estimates for the dual users based on the two samples (landline and cell). The two estimates for the dual users will then be combined and added to the estimates based on landline-only and cell-only population to generate the estimate for the whole population.

Composite pre-weight- For the purpose of sample weighting of the national sample, the four census regions will be used as weighting adjustment classes. Following Kennedy, Courtney (2007), the composite pre-weight will be generated within each weighting class. The weight assigned to the $i^{\text {th }}$ respondent in the $h^{\text {th }}$ weighting class ( $\mathrm{h}=1$, $2,3,4$ ) will be calculated as follows:

$$
\begin{array}{ll}
\mathrm{W}_{\text {(landline, hii) }}=\left(\mathrm{N}_{\mathrm{hl}} / \mathrm{n}_{\mathrm{hl}}\right)\left(1 / \mathrm{RR}_{\mathrm{hl}}\right)\left(\mathrm{n}_{\mathrm{cwa}} / \mathrm{n}_{\mathrm{ll}}\right)\left(\lambda^{\text {IDual }}\right) & \text { for landline sample cases } \\
\mathrm{W}_{(\text {Cell, }, \mathrm{hi})}=\left(\mathrm{N}_{\mathrm{hc}} / \mathrm{n}_{\mathrm{hc}}\right)\left(1 / \mathrm{RR}_{\mathrm{hc}}\right)(1-\lambda)^{\text {IDual }} & \text { for cellular sample cases } \tag{2}
\end{array}
$$

where
$\mathrm{N}_{\mathrm{h}}$ : size of the landline RDD frame in weighting class h
$\mathrm{n}_{\mathrm{hl}}$ : sample size from landline frame in weighting class $h$
$\mathrm{RR}_{\mathrm{h}}$ : response rate in weighting class h associated with landline frame
$\mathrm{n}_{\mathrm{cwa}}$ : number of adults in the sampled household
$\mathrm{n}_{11}$ : number of residential telephone landlines in sampled household
$I^{\text {Dual }}$ : indicator variable with value 1 if the respondent is a dual user and value 0
otherwise
$\mathrm{N}_{\mathrm{hc}}$ : size of the Cell RDD frame in weighting class h
$\mathrm{n}_{\mathrm{hc}}$ : sample size from Cell frame in weighting class h
$\mathrm{RR}_{\mathrm{hc}}$ : response rate in weighting class h associated with Cell frame
' $\lambda$ ' is the "mixing parameter" with a value between 0 and 1 . If roughly the same number of dual users is interviewed from both samples (landline and cell) within each census region, then 0.5 will serve as a reasonable approximation to the optimal value for $\lambda$. This adjustment of the weights for the dual users based on the value of the mixing parameter ' $\lambda$ ' will be carried out within each census region. For this study, the plan is to use a value of ' $\lambda$ ' equal to the ratio of the number of dual users interviewed from the landline frame and the total number dual users interviewed from both frames within each region.

It may be noted that equation (2) above for cellular sample cases doesn't include weighting adjustments for (i) number of adults and (ii) telephone lines. For cellular sample cases, as mentioned before, there is no within-household random selection. The random selection can be made from all persons sharing a cell phone but the percentage of those sharing a cell phone is rather small and it will also require additional questionnaire time to try to capture such information. The person answering the call will be selected as the respondent if he or she is otherwise found eligible and hence no adjustment based on "number of eligible adults in the household" will be necessary. The information on the number of cell phones owned by a respondent could also be asked to make adjustments based on number of cell phones. However, the percentage of respondents owning more than one cell phone is expected to be too low to have any significant impact on sampling weights. For landline sample cases, the values for (i) number of adults ( $\mathrm{n}_{\mathrm{cwa}}$ ) and (ii) number of residential telephone lines $\left(\mathrm{n}_{\|}\right)$may have to be truncated to avoid extreme weights. The cutoff value for truncation will be determined after examining the distribution of these variables in the sample. It is anticipated that these values may be capped at 2 or 3 .

Response rate: The response rates $\left(\mathrm{RR}_{\mathrm{hl}}\right.$ and $\mathrm{RR}_{\mathrm{hc}}$ mentioned above in equations (1) and (2)), will be measured using the AAPOR (3) definition of response rate within each weighting class and will be calculated as follows:

RR = (number of completed interviews) / (estimated number of eligibles)

$$
\begin{aligned}
& \text { = (number of completed interviews) / (known eligibles + presumed } \\
& \text { eligibles) }
\end{aligned}
$$

It will be straightforward to find the number of completed interviews and the number of known eligibles. The estimation of the number of "presumed eligibles" will be done in the following way: In terms of eligibility, all sample records (irrespective of whether any contact/interview was obtained) may be divided into three groups: i) known eligibles (i.e., cases where the respondents, based on their responses to screening questions, were found eligible for the survey), ii) known ineligibles (i.e., cases where the respondents, based on their responses to screening questions, were found ineligible for the survey), and iii) eligibility unknown (i.e., cases where all screening questions could not be asked, as there was never any human contact or cases where respondents answered the screening questions with a "Don't Know" or "Refused" response and hence the eligibility is unknown).

Based on cases where the eligibility status is known (known eligible or known ineligible), the eligibility rate (ER) is computed as:
ER = (known eligibles) / (known eligibles + known ineligibles)

Thus, the ER is the proportion of eligibles found in the group of respondents for whom the eligibility could be established.

At the next step, the number of presumed eligibles is calculated as:

Presumed eligibles $=\mathrm{ER} \times$ number of respondents in the eligibility unknown group
The basic assumption is that the eligibility rate among cases where eligibility could not be established is the same as the eligibility rate among cases where eligibility status was known. The response rate formula presented above is based on standard guidelines on definitions and calculations of Response Rates provided by AAPOR (American Association for Public Opinion Research).

Post-stratification weight- Once the landline and cell samples are combined using the composite weight (equations (1) and (2) above), a post-stratification weighting step will be carried out, following Kennedy (2007), to simultaneously rake the combined sample to (i) known characteristics of the target population (adults 18 years of age or older) and (ii) an estimated parameter for relative telephone usage (landline-only, cell only, cell mostly, other dual users). The demographic variables to be used for weighting will include Age, gender, Race, Ethnicity (Hispanic/Non-Hispanic), and Education. The target numbers for post-stratification weighting will be based on the latest Nielsen-Claritas county or
zipcode demographic estimates. The collapsing of categories for post-stratification weighting may become necessary where the sample sizes are going to be relatively small.

The target numbers for the relative telephone usage parameter will be based on the latest estimates from NHIS (National Health Interview Survey). For the purpose of identifying the "cell mostly" respondents among the group of dual users, a question similar to the following question will be included in the survey.
Question: Of all the telephone calls your household receives (read 1-3)?

$$
\begin{array}{ll}
1 & \text { All or almost all calls are received on cell phones } \\
2 & \text { Some are received on cell phones and some on regular phones, OR } \\
3 & \text { Very few or none are received on cell phones } \\
4 & \text { (DK) } \\
5 & \text { (Refused) }
\end{array}
$$

Respondents choosing response category 1 (all or almost all calls are received on cell phones) will be identified as "cell mostly" respondents.

After post-stratification weighting, the distribution of the final weights will be examined and trimming of extreme weights, if any, will be carried out if necessary to minimize the effect of large weights on variance of estimates.

Each of the eight hazard samples will be weighted separately following procedures similar to those described above for the national sample. For the hazard samples, the weighting classes will be based on county (or groups of counties) depending on the definition of specific counties. Once each of the nine samples (the national and the eight hazard samples) are weighted separately, they will also be pulled together into one combined sample using composite weighting. The combined sample will then be poststratified to known characteristics of the target population (i.e. the national population) for this study.

## -Degree of accuracy needed for the purpose described in the justification:

We plan to complete about 5,000 completed telephone interviews per administration including about 1,000 interviews using a national level sample and around 500 interviews each for each of the hazard areas. The survey estimates of unknown population parameters (for example, population proportions) based on a sample size of 5,000 will have a precision (margin of error) of about $\pm 1.4$ percentage points at $95 \%$ level of significance. This is under the assumption of no design effect and also under the most conservative assumption that the unknown population proportion is around $50 \%$. The margin of error (MOE) for estimating the unknown population proportion ' P ' at the $95 \%$ confidence level can be derived based on the following formula:

MOE $=1.96 * \sqrt{P *(1-P) / n}$ where " n " is the sample size (i.e. the number of completed surveys).

In this survey, where the total sample size $(5,000)$ will include oversamples from four hazard areas and therefore may be subject to a relatively higher design effect. A design effect of 2, for example, will result in effective sample size of 2,500 and a margin of error around $\pm 2.0 \%$ at $95 \%$ confidence level. The sampling error associated with an estimate based on just the national sample size of 1,000 with a design effect of 1.25 will still be below $\pm 3.5$ points. For each of the hazard areas with about 500 completed interviews, an estimate for an unknown population proportion will have margin of error around $\pm 4.4$ points ignoring any design effect. With an anticipated design effect of about 1.25, the precision will be around $\pm 4.9$ percentage points. Hence, the accuracy and reliability of the information collected in this study will be adequate for its intended uses. The sampling error of estimates for this survey will be computed using special software (like SUDAAN) that calculates standard errors of estimates by taking into account the complexity, if any, in the sample design and the resulting set of unequal sample weights. The necessary sample size for a two-sample proportion test (one-tailed test) can be derived as follows:

$$
\begin{equation*}
\left.\mathrm{n}=\left[\left\{\mathrm{z}_{(1-\alpha)} \operatorname{SQRT}\left(2 \mathrm{p}^{*} \mathrm{q}^{*}\right)+\mathrm{z}_{(1-\beta)} \operatorname{SQRT}\left(\mathrm{p}_{1} \mathrm{q}_{1}+\mathrm{p}_{2} \mathrm{q}_{2}\right)\right\} /\left\{\mathrm{p}_{2}-\mathrm{p}_{1}\right)\right\}\right]^{2} \tag{3}
\end{equation*}
$$

where
n : sample size (number of completed surveys) required per group to achieve the desired statistical power
$\mathrm{Z}_{(1-\alpha)}, \mathrm{Z}_{(1-\beta)}$ are the normal abscissas that correspond to the respective probabilities $\mathrm{p}_{1}, \mathrm{p}_{2}$ are the two proportions in the two-sample test and $p^{*}$ is the simple average of $p_{1}$ and $p_{2}$ and $q^{*}=1-p^{*}$.

For example, the required sample size, ignoring any design effect, will be around 310 per group (top and bottom halves) with $\beta=.2$ (i.e., with $80 \%$ power), $\alpha=.05$ (i.e., with $5 \%$ level of significance), and $\mathrm{p}_{1}=.55$ and $\mathrm{p}_{2}=0.45$. The sample size requirement is highest when $p_{1}$ and $p_{2}$ are around $50 \%$ and so, to be most conservative, those values (. 55 and .45) of $\mathrm{p}_{1}$ and $\mathrm{p}_{2}$ were chosen. The proposed sample size will therefore meet the sample size requirements for estimation and testing statistical hypotheses not only at the national level but also for a wide variety of subgroups that may be of special interest in this study.
-Unusual problems requiring specialized sampling procedures: (e.g., special hard to reach populations, bias toward landline verses cell phone respondents, populations that need to be reached via other methods such as those who do not use telephones for religious reasons, large non-English speaking populations expected to be surveyed but only English questionnaires available, exclusion of elderly using computer response only, etc.)

Note: For surveys with particularly low response rates and a substantial suspicion of non-response bias, it may be necessary to collect an additional sub-sample of
completed surveys from non-respondents in order to confirm if non-response bias is present in the sample and make adjustments if appropriate.

Unusual problems requiring specialized sampling procedures are not anticipated at this time. If response rates fall below the expected levels, additional sample will be released to generate the targeted number of surveys. However, all necessary steps to maximize response rates will be taken throughout the data collection period and hence such situations are not anticipated.
-Any use of periodic (less frequent than annual) data collection cycles to reduce burden:
During each administration of the survey, independent samples will be drawn and so the probability of selecting the same respondent in multiple administrations will be quite low.

1. Describe methods to maximize response rates and to deal with issues of nonresponse. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

Methods to maximize response rates- In order to maximize response rates, Gallup will use a comprehensive plan that focuses on (1) a call design that will ensure call attempts are made at different times of the day and different days of the week to maximize contact rates, (2) conducting an extensive interviewer briefing prior to the field period that educates interviewers about the content of the survey as well as how to handle reluctance and refusals, (3) having strong supervision that will ensure that high-quality data are collected throughout the field period, (4) using troubleshooting teams to attack specific data collection problems that may occur during the field period, and (5) customizing refusal aversion techniques. A $5+5$ call design, i.e., a maximum of five calls will be made on the phone number to reach the specific person we are attempting to contact and up to another five calls will be made to complete the interview with that selected person.

Issues of Non-Response- Survey based estimates for this study will be weighted to minimize any potential bias, including any bias that may be associated with unit level nonresponse. All estimates will be weighted to reduce bias and it will be possible to calculate the sampling error associated with any subgroup estimate in order to ensure that the accuracy and reliability is adequate for intended uses of any such estimate. Based on experience from conducting similar surveys previously and given that the mode of data collection for the proposed survey is telephone, the extent of missing data at the item level is expected to be minimal. We, therefore, do not anticipate using any imputation procedure to handle item-level missing data.

Non-response bias Study and analysis- A nonresponse bias analysis will be conducted to examine the non-response pattern and identify potential sources of nonresponse bias. No additional follow-up data collection for the non-respondents is planned for this study. Hence the proposed non-response analysis will be based on survey data collected in the main survey.

Nonresponse bias associated with estimates consists of two factors-the amount of nonresponse and the difference in the estimate between the groups of respondents and non-respondents. Bias may therefore be caused by significant differences in estimates between respondents and non-respondents further magnified by lower response rates. As described earlier in this section, necessary steps will be taken to maximize response rates and thereby minimize the effect, if any, of lower non-response rates on non-response bias. Also, nonresponse weighting adjustments will be carried out to minimize potential nonresponse bias. However, despite all these attempts, nonresponse bias can still persist in estimates.

As part of the non-response analysis, the respondents will be split into two groups: (i) early or 'easy to reach' and (ii) late or 'difficult to reach' respondents. The call design for this survey, as mentioned before, will be $5+5$ and so a maximum of up to 10 calls may be made to each sampled phone number. The total number of calls required to complete an interview with a respondent will be used to identify these two groups - "early" and "late" respondents. This comparison will be based on the assumption that the latter group may in some ways resemble the population of non-respondents. The goal of the analysis plan will be to assess the nature of non-response pattern in this survey. Nonresponse bias analysis may also involve comparison of survey-based estimates of important characteristics of the adult population to external estimates. This process will help identify estimates that may be subject to nonresponse bias. If non-response is found to be associated with certain variables, then weighting based on those variables will be attempted to minimize non-response bias.

Note: Describe all possible actions you plan to take to maximize response including incentives, call-backs, follow up, survey length kept to a minimum to increase participation, letters urging the importance of their contribution to this data collection, etc.
4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.

The CATI survey will be tested with fewer than 10 respondents, prior to fielding, to ensure correct skip patterns and procedures.
5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

The information collection is conducted for the Individual and Community Preparedness Division by a contractor:

The representatives of the contractor who consulted on statistical aspects of design are:

Camille Lloyd
Research Director
Gallup Inc.
901 F Street NW
Washington, DC 20004
202-715-3188
Camille lloyd@gallup.com
Manas Chattopadhyay
Chief Methodologist
Gallup Inc.
901 F Street NW
Washington, DC 20004
202-715-3030
Manas chattopadhyay@gallup.com

