Part B:

Supporting Statement

Voice in the Workplace Survey

Collection of Information Employing Statistical Methods

Supporting Statement
Voice in the Workplace Survey

B. Collection of Information Employing Statistical Methods

B.1.   **Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods 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 had been conducted previously, include the actual response rate achieved during the last collection.**

The goal of this study is to gauge the current level of workers' voice in the workplace and the factors affecting voice, specifically, voice relating to the laws administered and enforced by DOL's Occupational Safety and Health Administration (OSHA) and Wage and Hour Division (WHD). The universe for this study will consist of all "currently working" adults (18 years of age or older) residing in U.S. households in any of the 50 states or in the District of Columbia. Respondents who report that they are currently working full time or part time and are not self-employed will be considered eligible for this study. So, all adults with a current job, i.e., working for pay, will be included, and the group of self-employed adults will be excluded. Based on latest Current Population Survey (CPS) data, the estimated number (annual average 2010, i.e. average of 12 monthly estimates of 2010) of "currently employed" adults (i.e., working for pay and excluding self-employed adults) residing in U.S. households in any of the 50 states or in the District of Columbia is about 128 million.

For this study, a household-based telephone survey will be conducted to complete about 800 interviews for the pilot study and 5,400 interviews nationwide for the main study. The 5,400 interviews for the main study will include an oversample of working minority women (African American, Hispanic, Asian or American Indian). It is expected that roughly 1,760 of the 5,400 interviews will be completed with minority women. For the main study, about 4,000 interviews

will be completed based on a household based RDD (random Digit dialing) telephone sample consisting of both landline and cell phone numbers. This group of 4,000 interviews is likely to include around 350 interviews with minority women. In addition, another 1400 or so interviews with working minority women will be completed by screening the RDD sample for this targeted group. For both the pilot and the main study, the survey will consist of a core set of questions followed by two separate modules of questions—one each for OSHA and WHD—in which specific questions about each agency will be included. Respondents will respond to the core set of questions and then will be randomly assigned to one or the other module. The random assignment of questions to one or the other module will be done using CATI (Computer Assisted Telephone Interviewing) based software called SURVENT. As a result, the number of completed OSHA interviews (those containing responses to specific questions in the OSHA module) is expected to be around 400 for the pilot study and about 2,700 for the main study and this will also be the case with the WHD interviews. The set of core questions will be answered by everyone and so the number of completed interviews for the core questions will be 800 for the pilot study and 5,400 for the main study.

The primary goal of the pilot study is to test the survey instrument and the sample design to ensure that these are performing according to DOL requirements. Upon completion of the pilot study, a report will be prepared summarizing the findings. Changes, if necessary, in the survey instrument, sample design, or in any other aspect of the study will be made, and approved in accordance with PRA, before launching the main study. In order to minimize bias, both landline and cell phones will be included in the telephone sample. The selection of landline numbers will be based on list-assisted RDD (Random Digit Dialing) sampling of telephone numbers. The cell phone sample will be a simple random sample drawn from all dedicated exchanges for cell phones. 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.

This study has not been conducted previously and so there is no past response rate to refer to. The goal will be to maximize the response rate by taking necessary steps as outlined in section B3 on "Methods to maximize response rates."

The population parameter of primary interest will be the proportion of "currently working adults" in specific categories—for example, the "proportion of workers who are aware of their rights as a worker" or the "proportion of workers who think there are problems in their workplace with employees being exposed to health and safety hazards." The sample-based estimate (p) of the parameter representing an unknown population proportion (P) can be expressed as:

$$p = \frac{\sum_{i=1}^{n} WiYi}{\sum_{i=1}^{n} W\,i},$$

where Yi = 1 if the ith sampled respondent belongs to the category of interest (aware of their rights, for example) and 0 otherwise; W_i is the sample weight attached to the ith respondent and "n" is the number of completed surveys.

For this baseline study, these parameters (proportions or means) will be estimated at the overall national level, i.e., for all "currently working" adults in the U.S. The corresponding estimates at subgroup level (geographic regions, specific industries or groups of industries, etc.) may be computed and the precision associated with those estimates will depend on the resulting sample size (number of completed surveys) for these subgroups. No disproportional sample allocation by geographic region or by industry (of employment) is proposed to boost the sample size for any specific subgroup. As described above, the group of working minority women will be oversampled.

B.2.   Describe the procedures for the collection of information, including:

Statistical methodology for stratification and sample selection—The target population for this survey, as mentioned before, will consist of all U.S. adults who are currently working (full time

or part time and not self-employed) and living in households in any of the 50 states or in the District of Columbia. A telephone survey will be conducted to complete about 800 interviews for the pilot and about 5,400 interviews nationwide for the main study. In order to minimize bias, both landline and cell phones will be included in the telephone 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 working adults. Based on latest CPS data, the distribution of working adults across the four regions is as follows: 19% (Northeast), 22% (Midwest), 36% (South) and 23% (West). Using proportional sample allocation, the targeted number of surveys to be completed in each region is expected to be close to those proportions. For each group (OSHA or WHD), the targeted number of completes will be about 50% of the total allocated number for that region. Within each region, roughly 45 percent of the interviews will be done from the cell phone sample while the rest (55%) will be done from the landline sample. For the pilot study, the total sample of 800 will be equally split across the four census regions, i.e. about 200 interviews (about 100 each for OSHA and WHD) will be completed for each region. Within each region, roughly 40 percent of the interviews will be from the cell phone frame while the rest will be obtained from the landline frame.

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 proposed by Casady and Lepkowski (1993). This procedure uses the Telcordia frame that is generated by appending all 10000 four-digit suffixes (0000 to 9999) to the area code-prefix combinations. In view of cost and operational efficiency, this study will follow the truncated version of the Casady and Lepkowski (1993) method and will sample from 100-banks containing at least 1 listed residential number (1+). For within-household sampling, Gallup will use the "most recent birthday" method to randomly select one eligible person from all eligible adults in each sampled household. Following the "most recent birthday" method, the interviewer asks to speak with the eligible person in the household who most recently had a 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. The temporary/(pay as you go) phone numbers will be included in the cell phone frame. 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?"). For the cell phone sample, data will be collected from all respondents regardless of whether they also have access to a landline. A

respondent reached on a cell phone 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 100-banks (with at least one listed residential number) of the remaining telephone exchanges. 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 information) and will not be determined based on their telephone exchange.

For the oversampling of working women belonging to the minority group, necessary screening questions based on race/ethnicity will be asked. The RDD (Random Digit Dialing) telephone sample consisting of both landline and cell numbers will be screened to generate this oversample. In order to maximize the incidence rate for this group, certain telephone exchanges containing higher percentages of minority population will be oversampled. If needed, additional sample source generated from Gallup's G1K survey may be used. Gallup conducts a daily survey (called G1K survey) where about 1,000 interviews are completed daily nationwide using a full dual frame (landline and cell) telephone sample design. A significant amount of demographic and other information (including employment status) is available for the respondents of the G1K survey and a certain percentage of these respondents (those who are willing to participate in a follow-up survey) may be recontacted to oversample this group of working minority women.

Estimation procedure—Sample data will be weighted to generate unbiased estimates. Within each stratum (region), weighting will be carried out 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 procedure 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, 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 ith respondent in the hth weighting class (h=1, 2, 3, 4) will be calculated as follows:

$$W_{(landline,hi)} = (N_{hl}/n_{hl})(1/RR_{hl})(n_{cwa}/n_{ll})(\lambda^{I^{Dual}}) \qquad \text{for landline sample cases} \qquad (1)$$

$$W_{(Cell,hi)} = (N_{hc}/n_{hc})(1/RR_{hc})(1 - \lambda)^{I^{Dual}} \qquad \text{for cellular sample cases} \qquad (2)$$

where

   N_{hl}: size of the landline RDD frame in weighting class h

   n_{hl}: sample size from landline frame in weighting class h

   RR_{hl}: response rate in weighting class h associated with landline frame

   n_{cwa}: number of "currently working" adults in the sampled household

   n_{ll}: number of residential telephone landlines in sampled household

   I^{Dual}: indicator variable with value 1 if the respondent is a dual user and value 0 otherwise

   N_{hc}: size of the Cell RDD frame in weighting class h

n_{hc}: sample size from Cell frame in weighting class h

RR_{hc}: response rate in weighting class h associated with Cell frame

'λ' 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 λ. This adjustment of the weights for the dual users based on the value of the mixing parameter 'λ' will be carried out within each census region. For this study, the plan is to use a value of 'λ' 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. One or two additional values of the mixing parameter may be tested to see the impact on survey estimates. It is anticipated that the value of the mixing parameter will be close to 0.5.

It may be noted that equation (2) above for cellular sample cases doesn't include weighting adjustments for (i) number of "currently working" 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 eligible adults (n_{cwa}) and (ii) number of residential telephone lines (n_{ll}) 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 (RR_{hl} and RR_{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)

= (number of completed interviews) / (known eligibles + presumed eligibles)

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 = ER × 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 two 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 currently working full time or part time and not self-employed) and (ii) an estimated parameter for relative telephone usage (landline-only, cell only, cell mostly, other dual users).

As mentioned before, adults who are "currently working" full time or part time and are not self-employed will be eligible for this study. For the main study, the plan is to use the following variables for post-stratification weighting within each stratum (Census Region).

(i)   Employment status: employed full-time, Employed part-time, Temporary/Day/Seasonal worker
(ii) How are they paid: Salary, Hourly, Paid by unit produced or action performed, Daily
(iii) Industry: Agriculture and related industry, Non-agricultural industries

The target numbers for post-stratification weighting will be obtained from the latest available Current Population Survey (CPS) data. The selection of the variables for post-stratification weighting for the main study will be finalized after the pilot survey data are examined. It may be necessary to combine some of the levels (or categories) or add more levels for the post-stratification weighting variables mentioned above. The collapsing of categories for post-stratification weighting may become necessary particularly for the pilot survey data 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, the following question (Question D23C in the attached questionnaire) will be included in the survey.

D24C
QID:103424   Of all the telephone calls your household receives **(read 1-3)**?

1	All or almost all calls are received on cell phones
2	Some are received on cell phones and some on regular phones, OR
3	Very few or none are received on cell phones
4	(DK)
5	(Refused)

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.

The weighting procedure described above will be applicable for the core questions (i.e. questions that will be asked to every eligible respondent). An additional adjustment to the composite pre-weights (equations (1) and (2)) will be necessary for the non-core questions (those in the OSHA and the WHD modules). As noted before, these are questions that will be asked only to about half of all the eligible respondents. Every eligible respondent for the study will answer all the core questions. At that point, roughly half of these respondents will be assigned to the OSHA module and the remaining half will be asked to answer the questions in the WHD module. Initially, about equal number (50-50 split) of respondents will be assigned to each module. However, depending on the response rates of the respondents assigned to these modules, it may also be necessary to change that proportion later during the course of data collection to end up with roughly equal number of completed surveys for each module. As mentioned before, the random assignment of respondents to either OSHA or WHD module will

be implemented using a CATI (Computer Assisted Telephone Interviewing) based software (SURVENT). If a proportion 'p' of the respondents are directed to OSHA module (i.e., a proportion q=(1-p) of respondents to the WHD module), then the pre-weights (equations (1) and (2) above) for questions in OSHA (WHD) module will be multiplied by 1/p (1/q). The rest of the weighting steps (including post-stratification weighting) for these non-core questions will be the same as those described for the for the core set of questions. The final data set of completed surveys will therefore include three weighting variables: (i) a weight variable for the core set of questions (ii) a weight variable for the OSHA module questions and (iii) a weight variable for the WHD module questions. The choice of the weight variable for any particular analysis will depend on the specific requirements of that analysis. It may also be noted that the target data for post-stratification will include working adults that do not have access to telephone (non-telephone population). This will help minimize the coverage bias due to exclusion of the non-telephone working adults from the scope of this telephone based survey.

Degree of accuracy needed for the purpose described in the justification—We propose a sample size large enough to generate 5,400 completed telephone interviews at the national level for the core questions and about 2,700 completed surveys for questions in each of the two modules—one each for OSHA and WHD. The Statement of Work (SOW) specifies that the sample size should be large enough to determine whether the workers in the bottom half on the voice measure have at least 10 percentage points' greater likelihood of perceiving their workplaces as unsafe than do workers with higher levels of voice. The necessary sample size for a two-sample proportion test (one-tailed test) to meet this requirement can be derived as follows:

$$n = [\{z_{(1-\alpha)} \, \text{SQRT} \, (2p^*q^*) + z_{(1-\beta)} \, \text{SQRT}(p_1q_1 + p_2q_2)\} / \{p_2 - p_1)\}]^2 \quad (3)$$

where

n: sample size (number of completed surveys) required per group to achieve the desired statistical power
$z_{(1-\alpha)}$, $z_{(1-\beta)}$ are the normal abscissas that correspond to the respective probabilities
p_1, 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 β=.2 (i.e., with 80% power), α=.05 (i.e., with 5% level of significance), and p_1=.55 and 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 p_1 and p_2 were chosen. Formula (3) presented above is from Lemeshow, Stanley, et al. (1990). Adequacy of Sample Size in Health Studies. Wiley, Chichester, U.K.

The SOW also states that detecting smaller differences or differences across even smaller groups are desirable. Taking into consideration the overall objectives of this study, a larger sample size (roughly 2,700 completed interviews for each module) is proposed so that the sample size requirement is satisfied not only at the national level but also for a wide variety of subgroups that may be of special interest in this study.

The survey estimates based on core questions (for example, the proportion of workers who know where to access information about their rights as workers) will have a sample size of 5,400 and a precision (margin of error) of about ±1.3 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:

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

In a dual frame household-based RDD survey, some design effect is expected but the precision for survey-based estimates for most subgroups of interest are likely to have reasonable precision. For example, the sampling error associated with an estimate based on a sample size of 1,000 with a design effect of 1.25 will still be below ±3.5 points. For questions in OSHA or WHD modules, the sample size will be around 2,700 and the sampling error associated with an estimate for an unknown population proportion will be around ±1.9 points ignoring any design effect. With an anticipated design effect of about 1.5, the precision will be around ±2.3 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.

1. **Unusual problems requiring specialized sampling procedures**—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.

2. **Any use of periodic (less frequently than annual) data collection cycles to reduce burden**—DOL currently plans for only this data collection of the worker voice survey. In future, there may be needs for additional data collection to ensure the possibility of data trending, analyses, and measuring the Departmental actions and their relationship to changes in voice in a longitudinal manner.

B.3. **Describe methods to maximize response rates and to deal with issues of non-response. 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.**

The response rates for RDD (Random Digit Dialing) household based telephone surveys have been steadily declining over the past decade. Groves (2006) addresses issues relating to non-response rates and non-response bias in household surveys. Non-response need not always cause non-response bias in survey based estimates. A synthesis of research studies shows that non-response can often cause non-response bias. It is, therefore, important to make all possible efforts to maximize response rates to avoid potential non-response bias. However, non-response bias can still be present in spite of all efforts to achieve a higher response rate. For this study, a follow-up non-response bias study is therefore planned to examine the non-response pattern and its possible effects on key survey estimates. This section outlines the steps to be taken to maximize the response rates and also describes additional details of the

proposed non-response bias study. Information collected in this study will 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 and conversion techniques. Gallup will use 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. At the national level, the sampling error associated with estimates of proportions based on core questions is expected to be around 1.5 percentage points and those based on questions in the OSHA or WHD module around 2.2 points, ignoring design effect. For any subgroup of interest, the sampling error will depend on the sample size. 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 follow-up study will be conducted to examine non-response pattern and identify potential sources of nonresponse bias. 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 nonrespondents. The bias of an estimate can be expressed mathematically as follows:

$$Bias\ (y_r) = (1 - r)\ \{E\ (y_r - y_n)\}$$

where y_r is the estimated characteristic based on survey respondents only, "r" is the response rate, and so $(1 - r)$ is the nonresponse rate, y_n is the estimated characteristic based on the nonrespondents only, and E is the expectation for averaging over all possible samples.

Bias may therefore be caused by significant differences in estimates between respondents and nonrespondents further magnified by lower response rates. As described earlier in this section (B3), 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. The goal of the nonresponse bias study will be to identify sources of nonresponse bias on estimates and to identify potentially biased estimates.

The mode of data collection for this non-response follow-up study will also be telephone and a respondent can receive anywhere between 1 and 10 calls to complete an interview. The group of non-respondents will include: (i) Non-contacts (sampled cases where no human contact will be established during the main phase of data collection), and (ii) Refusals (sampled cases where a human contact will be established but an interview can't be completed). In order to represent each of these groups in the sample for the non-response bias study, random samples will be selected from each of these groups (strata) independently and the goal will be to complete a total of about 400 interviews from these two groups.

The analysis plan for the non-response bias study will involve comparing the respondents and the non-respondents on key variables (survey data on selected survey questions) of interest. The survey data on selected variables for the group of respondents will be compared to those obtained from the 400 or so completed interviews from non-respondents. In addition, the respondents to the main study will be split into two groups: (i) early or 'easy to reach' and (ii) late or 'difficult to reach' respondents. The total number of calls required to complete an interview will be used to identify these groups. 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 will also involve comparison of survey-based estimates of important characteristics relating to "currently working" population to external estimates. For example, the proportion of "full-time or part-time workers" or proportion of "workers who get paid hourly wages" may be selected for such comparison. The analysis can be done using nonresponse adjusted weights but can also be done using the base weights. This process will help identify estimates that may be subject to nonresponse bias.

All questions in the Outcome/Loyalty and Perceived Voice sections (questions 1 through 21of the attached survey) will be selected for comparing the respondents and the nonrespondents. In addition, the following questions from the OSHA and the WHD modules will be selected for this analysis:

For the WHD module: 22A, 22B, 22C, 24A-21D, 28A-28E, 41
For the OSHA module: 22, 23, 25A-25F, 29, 42

For each of these selected variables, the mean of the two groups (for example, respondents and non-respondents) will be compared based on a t-test using software SUDAAN. Let the mean (or, equivalently, the proportion of 1's for a 0-1 variable) of the two groups for a specific 0-1 variable (Y) based on survey data be denoted by p_1 and p_2, respectively. Then, p_1 can be written as:

$p_1 = \sum W_i y_i / \sum W_i$, where y_i is 1 if the value of variable Y for i^{th} respondent is 1 and '0' otherwise; W_i is the weight assigned to the ith respondent and the summation in both numerator and denominator is over all respondents in the sample. p_2 can be similarly defined. The t-statistic for testing the equality of means for those two groups (H_o: $P_1 = P_2$ vs. H_1:$P_1 \neq P_2$ where P_1 and P_2 are the corresponding population means) will be computed as:

$$t = (p_1 - p_2)/SE\ (p_1 - p_2)\ ,\ \text{where SE}\ (p_1 - p_2)\ \text{is the standard error or the estimated square-root of}$$
$$\text{the variance of}\ (p_1 - p_2)$$

In order to obtain the value of t-statistic (and the corresponding significance level or p-value), the main SUDAAN commands using the DESCRIPT procedure will be as follows:

```
                    PROC DESCRIPT DATA=XXXX FILETYPE=SAS DESIGN=STRWR;
nest region;
WEIGHT FINALWT;
class respondent_nonrespondent;
var  Y;
contrast respondent_nonrespondent = (1 -1)/name = "respondent vs. nonrespondent";
print nsum t_mean p_mean mean;
```

The respondent_non-respondent variable will contain two distinct values (0-1 for example) to identify the two groups (respondent or non-respondent) for each case in the data set. The VAR statement will include the variables for which the mean has to be compared between the two groups. For each selected variable included in the VAR statement, the hypothesis of equality of means will be rejected (or not) based on the p-value (less than 0.05 or not).

Analysis Plan – An outline of the data analysis plan is attached as a separate document.

B.4.   Describe any tests of procedures or methods to be undertaken.

As mentioned before, a pilot study will be undertaken before the main study is launched. The pilot will mimic the main study and will complete a total of 800 surveys, about 400 each for

OSHA and WHD modules. Upon completion of the pilot study, a report will be prepared summarizing the findings. Changes, if necessary, in survey instrument, sample design, or in any other aspect will be made before launching the main study. One of the main goals in the design and conduct of this study will be to minimize respondent burden.

A non-response bias study will be conducted after completion of the main study to examine the nature of non-response pattern and identify potential non-response bias. Details are already described in Section B3 above.

The contractor will prepare a public use data set for this study following standard procedures to protect every respondent's personal information from disclosure, as required by law. Gallup has in place a comprehensive system of interlocking procedures to protect the privacy, and anonymity of respondents and data security measures during data collection, processing and reporting after the conclusion of the survey. If necessary, the features of our system can be specially tailored to meet the standard requirements required of any Federal agency.

B.5.   Provide the name, affiliation (company, agency, or organization), 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.

Name	Agency/Company/Organization	Number Telephone
Camille Lloyd	Gallup	202.715.3188
Dr. Manas Chattopadhyay	Gallup	202.715.3179

Reference

Robert J. Casady and James, M. Lepkowski (1993). Stratified Telephone Survey Designs. *Survey Methodology*, 19, 103-113.

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.

Lemeshow, Stanley, et al. (1990). Adequacy of Sample Size in Health Studies. Wiley, Chichester, U.K.Groves, Robert M. (2006): Nonresponse Rates and Nonresponse Bias in Household Surveys, *Public Opinion Quarterly*, Special Issue 2006, Volume 70/Number 5: 646-675.