OMB Approval No. 1225-0059

CUSTOMER SATISFACTION SURVEY AND CONFERENCE EVALUATION CLEARANCE FORM

A. SUPPLEMENTAL SUPPORTING STATEMENT

A.1. Title: **EBSA Participant Assistance Program Customer Survey**	
A.2. Compliance with 5 CFR 1320.5: Yes _X_ No _____	**A.3. Assurances of confidentiality:** No confidential data will be collected.
A.4. Federal cost: $544,397 Based on the cost for research contractor and for IT contractor for support	**A.5. Requested expiration date (Month/Year): 11/2012**
A.6. Burden Hour estimates: a. Number of Respondents: 6336 a.% Received Electronically 0% b. Frequency: Once c. Average Response Time: 8 minutes d. Total Annual Burden Hours: 845 hours	**A7. Does the collection of information employ statistical methods?** **X** Yes (Complete Section B and attach BLS review sheet).

A.8. Abstract:

This survey will collect customer satisfaction data for a sample of private citizens who call into the participant assistance program to ask about their private sector employer provided benefits such as pensions, retirement savings, and health benefits. Three types of callers will be queried: 1. Those who need benefit claim assistance 2. Those who have a valid benefit claim and 3. Those who have an invalid benefit claim. The detailed data collected will be used for internal management purposes. The overall satisfaction score will be used to establish a baseline customer satisfaction performance measure for the purposes of complying with the Government Performance and Results Act.

Program Official	Date	Departmental Clearance Officer	Date
Sharon S. Watson	01/25/2011	Michel Smyth	01/25/2011

B. SURVEYS AND EVALUATIONS 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 evaluate EBSA's Participant Assistance Program (PAP) and the universe will consist of participant inquiries (individuals who had contacted EBSA for assistance) handled by the 10 regional offices and the Office of Participant Assistance (OPA). Inquiries are made by telephone as well as using letters, emails, web-based inquiries and walk-in visits. For each of the 10 regional offices, the universe will include only telephone inquiries which accounts for at least 92 percent of all inquiries for any particular office. For all 10 offices put together, the number of telephone inquiries will be about 94 percent of the total number of inquiries handled by those 10 offices. For the OPA, the universe will consist of mail and web inquiries which account for about 81 percent of all OPA inquiries. These percentages are based on information on closed inquiries for FY2010. For purpose of sampling, the universe will be stratified into 11 strata (the 10 regional offices and the OPA as shown in Table 1 below) and sampling will be done independently within each office using a simple stratified sample design. Table 1 below provides the number of "closed" inquiries (web/mail inquiries for the OPA and telephone inquiries for the rest of the offices) during FY2010.

Table 1: Closed Inquiries by 11 offices during FY2010

Regional Office	Number of Inquiries
Atlanta	61,940
Boston	29,094
Chicago	22,226
Cincinnati	35,788
Dallas	29,219
Kansas City	32,728
Los Angeles	26,748
New York	29,343
Philadelphia	32,936
San Francisco	34,218
OPA	16,809
Total	351,049

Every two weeks, EBSA will send data on all of the participant inquiries identified as newly "closed" during the previous two weeks. Gallup will select samples of participant inquiry records on a bi-weekly

basis using a simple stratified sample design. Based on information from the past administration of this study, we anticipate an average response rate of 54% for each two-week period. DOL conducted this study in 2006 and achieved a response rate of 54 percent. In order to examine the issue of non-response bias, a non-response bias analysis will be conducted. The non-response bias analysis plan to be followed for this purpose is described in Section B3. In total for FY2011, there will be 18 data collection periods. The goal will be to complete a total of about 32 interviews from each of the 11 offices in each period so that a total of 576 (= 32 X 18) interviews can be completed for each office over the course of the 18 periods of data collection. The total number of interviews across all offices is therefore estimated to be around 6,336 (=576 X 11).

The population parameter of primary interest will be the proportion of customers in specific categories. For example, the "proportion of customers in the population who are satisfied with EBSA overall" or the "proportion of customers who think EBSA is a name they can trust" will have to be estimated based on survey data. The corresponding sample estimates will be computed based on responses to the survey. On a satisfaction question such as : "How satisfied are you with EBSA overall?", the proportion of satisfied customers may be estimated based on the proportion selecting one of the top two boxes on a 5-point likert scale. Customers will also be asked to indicate their level of agreement with statements like "EBSA is a name I can always trust" or "EBSA always delivers on what they promise" on a 5-point scale. The proportion of customers who select one of the top two boxes will provide an estimate of the corresponding population proportion. The sample based estimate (p) of the parameter representing an unknown population proportion (P) can be expressed as:

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

where Yi = 1 if the i^{th} sampled respondent belongs to the category of interest (satisfied, for example) and 0 otherwise; W_i is the sample weight attached to the i^{th} respondent and 'n' is the number of completed surveys.

These parameters (proportions or means) may have to be estimated at the overall EBSA level, for each of the 11 strata (offices) separately and possibly for other domains of interest within each stratum. For example, it may be of interest to generate similar estimates by the three Closure Types: (i) those who need benefit claim assistance (ii) those who have a valid benefit claim and (iii) those who have an invalid benefit claim. The bulk of the calls (about 80 to 90% on an average) will belong to the first

Closure type (those who need benefit claim assistance) and hence the number of completed surveys within each stratum for this subgroup will be large enough (around 500) to generate estimates of acceptable precision. Similar estimates for the other two Closure types can also be generated but the sample size within individual offices will be low and hence the estimates for these two Closure types may have to be generated at the overall EBSA level.

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

- **Statistical methodology for stratification and sample selection** – The universe of all telephone inquiries will be stratified by the 10 offices such that each office will be a separate stratum. For the OPA office, the universe will consist of all web and mail inquiries. The bi-weekly files to be transmitted from EBSA to the contractor for the purpose of sampling will contain variables necessary to link each inquiry to a particular office. Within each stratum, a simple random sample of specified size will be drawn independently once every two weeks. The initial sample size within each stratum for every two week period will be large enough to yield around 32 completed interviews. Based on an anticipated response rate of 54 percent, we will sample about 60 inquiries for each two week data collection period. That should on an average yield 32 (= 60 X .54) completed surveys. If the response rate varies during the data collection period or across different offices, the sample sizes will be adjusted accordingly to meet the target for the number of completed surveys.

- **Estimation procedure** – Sample data will be weighted to generate unbiased estimates for the target population subgroups. Within each stratum, weighting will be carried out to adjust for (i) probability of selection in the sample and (ii) non-response. Once the sampling weights are generated, weighted estimates will be produced for different unknown population parameters (means, proportions etc.) for the target population and also for population subgroups of interest. For the purpose of illustration, let us assume that we receive a total of 1,500 inquiries in a stratum (or regional office) in a particular two-week period and we select a random sample of size 60 from those 1,500 inquiries. Also, assume that 32 of those 60 sampled cases actually respond i.e. we get 32 completed surveys. The weight assigned to each of those 32 completed surveys will consist of two weighting factors: (i) selection probability weight (1500/60) and (ii) non-response weight (60/32). The first weighting factor is the inverse of the selection probability while the second factor is the ratio of the sample size and the number of completed surveys. The final weight will be the product of these two factors. In this specific example, the final weight assigned to each of those 32 completed cases will be (1,500/60) * (60/32) = 1,500/32. The sum of the weights of these 32 cases will add up to 1,500, the total number of inquiries for that two-week period.

Based on our previous experience in conducting this survey, we do not anticipate any "ineligible" cases in the sample. However, the weighting procedure can be easily adjusted to account for ineligible cases, if any. If, for example, 6 out of the 60 sampled cases turn out to be ineligible, the non-response weight factor will be equal to 54/32 and then the final weight assigned to each of the 32 completed surveys will be (1500/60)*(54/32). The sum of the weights for all 32 cases will then equal 32*(1500/60)*(54/32) = 1350, the estimated number of eligible cases in the population during that particular data collection period.

In terms of mathematical symbols, the weighting steps can be described as follows. Let N_{ij} and n_{ij} denote the population size (total number of cases received) and the corresponding sample size (number of cases sampled) for any particular office i (i=1,2, …,11) and for a specific two-week data collection period j (j=1,2,…,18). Also, let r_{ij} denote the number of responding units in the sample in i^{th} office and j^{th} data collection period. Then, the base-weight or the probability weight factor (W_{1ijk}) assigned to k^{th} sampled unit (k=1, 2, …, n_{ij}) will be derived as:

$$W_{1ijk} = N_{ij}/n_{ij} \ldots (1)$$

At the next step, the non-response adjustment factor (W_{2ijk}) will be derived as:

$$W_{2ijk} = \sum W1ijm * eijm \Big/ \sum W1ijm * dijm \ \ldots (2)$$

if the k^{th} unit (k=1, 2, …, n_{ij}) is a responding unit and 0 otherwise;
e_{ijm} = 1 if the m^{th} unit in the sample is eligible and 0 otherwise; d_{ijm} = 1 if the m^{th} unit is eligible and responds to the survey and 0 otherwise.
In the right hand side of equation (2) above, note that the summation in the numerator is over all sampled eligible cases whereas the summation in the denominator is over all selected eligible persons who actually respond to the survey.
The final weight (W_{ijl}) assigned to all r_{ij} responding units (in i^{th} office in j^{th} data collection period) will be the product of the two weighting factors:

$$W_{ijl} = W_{1ij} * W_{2ij} \ldots (3) \quad (l = 1, 2, \ldots, r_{ij}).$$

The simple random samples drawn during each data collection period within each office are likely to include proportional representation of cases (inquiries) by Closure types (those who need benefit claim-assistance, those who have a valid benefit claim, and those who have an invalid benefit claim) If response rates are found to vary significantly across Closure types, construction of non-response adjustment cells based on Closure types may be considered. However, for most strata (offices), the bulk of the calls (80 to 90%) are likely to belong to the first Closure type category (those who need benefit claim assistance) only and the number of calls belonging to the other two types (valid and invalid) will be small. Response rates among very small groups may vary significantly and thereby cause an increase in the variability of the weights. Collapsing of non-response adjustment cells, therefore, will be necessary. For this purpose, some collapsing rules will be developed after examining the nonresponse pattern. We anticipate using rules based on the (i) size of the cell and (ii) value of the non-response adjustment factor. Collapsing may be considered if (i) a cell has fewer than a certain number of cases (10, for example) or (ii) non-response adjustment factor exceeds a certain value (2, for example). For this study, given that bulk of the calls will belong to the first category only, we anticipate using the entire stratum as the non-response adjustment cell for most strata. In some strata, it may be possible to use two cells (those who need benefit claim assistance and Others). For OPA where the universe consists of all mail and web inquiries, non-response adjustment cells based on type of inquiry (mail or web) may also be considered. The non-response weighting adjustment procedure within each of these non-response adjustment cells will be the same as described above.

- **Degree of accuracy needed for the purpose described in the justification** – For each stratum (office), the total number of completed interviews over the 18 data collection periods will be around 576. For estimation of any unknown population proportion (P), for example, this will result in a margin of error of 4 percent at the 95% level of significance ignoring any design effect. 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 = $1.96 * \sqrt{P * \frac{1-P}{n}}$ where 'n' is the sample size (i.e. the number of completed surveys).

Under the most conservative assumption (P=0.5), the MOE for a sample size of 576 will be

$1.96 * \sqrt{\frac{.25}{576}} = 4.08\%$. The design effect within each stratum is expected to be minimal and hence the sampling error is likely to be in 4% to 5% range for estimates at the individual office level. It may be noted that, for any given office, the allocation of sample across all 18 data collection periods is not strictly

proportional because a fixed number of surveys (32) will be targeted in each two-week period. However, for the same office, the total number of cases (inquiries) is not expected to vary significantly over time and hence its impact on design effect for estimates based on data for several data collection periods should be minimal.

At the overall agency level, the total sample size will be around 6,336 and hence the MOE is expected to be around $1.96 * \sqrt{\frac{.25}{6336}} = 1.2\%$ under the assumption of no design effect. However, at the agency level, the disproportional sample allocation across different offices will contribute to the design effect. Based on approximate formula {design effect= (sample size)*(sum of squared weights)/ (square of the sum of weights)}, the design effect at the overall agency level is estimated to be around 1.12. After accounting for a conservative estimate of design effect equal to 1.25, the margin of error for estimates of population proportions at the overall agency level based on a sample size of 6336 will be about 1.4%. The sampling error of estimates for this survey may be computed using special software package 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.

- **Unusual problems requiring specialized sampling procedures and** – We don't foresee any unusual problems requiring specialized sampling procedures.

- **Any use of periodic (less frequently than annual) data collection cycles to reduce burden** – For this study, Gallup will sample once every two weeks and the data collection for every bi-weekly sample will be completed within the following two weeks. The data collection will therefore be a continuous process throughout the 9-month period but every respondent will be contacted/interviewed within two weeks after his/her inquiry is closed. This will be done to minimize the recall error and thereby increase the overall accuracy of the survey data. Once a respondent has completed a survey, he or she will not be contacted again for the FY2011 study.

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.

Methods to maximize response rates – In order to maximize response rates, Gallup will utilize a comprehensive plan that focuses on (1) a call design that will ensure call attempts are made at different times of day and different days of the week to maximize contact rates, (2) conducting an extensive interviewer briefing prior to the field period that educates them 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) utilizing 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 that we are attempting to contact and another up to 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 non-response. The bi-weekly files (sampling frames) to be received from DOL will contain some useful information (like Closure types) for all cases including the non-respondents. As mentioned in Section B2, response pattern by these variables will be examined and non-response adjustment cells, if found necessary, will be formed based on these variable and then ratio type adjustments will be carried out to correct for non-response. This will make the non-response weighting procedure quite effective in terms of minimizing non-response bias, if any. Based on our experience from conducting this survey 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.

As described above in Section B2, the sampling error associated with estimates of proportions at the individual office level is expected to be in the 4% to 5% range and that for the overall agency level is not likely to exceed 1.5% at the 95% level of confidence. For any other subgroup of interest (based on closure types, for example) the sampling error will depend on the sample size. Also, all estimates will be weighted to reduce bias. 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.
For this study, it will be possible to project the sample data to the target population. For collections based on the proposed sampling design, there will not be any collection that will not yield reliable data that can be projected to the universe studied.

Non-response Bias Analysis: A non-response bias analysis will be conducted to identify potential source of non-response bias. Non-response 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. 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 non-respondents only, and E is the expectation for averaging over all possible samples.

Bias may therefore be caused by lower response rate and/or by significant difference in estimates between respondents and non-respondents. As described earlier in this section (B3), necessary steps will be taken to maximize response rates and thereby minimize any non-response bias that may be caused by low response rates. Also, non-response weighting adjustments (refer to "Issues of non-response" above) will be carried out to minimize potential non-response bias. However, despite all these attempts, non-response bias can still persist in estimates. The goal of the non-response bias analysis will be to identify potential sources of nonresponse bias on estimates and to identify potentially biased estimates.

The non-response bias analysis will compare the "Early" respondents to "Late" respondents on selected key variables of primary interest. The basic assumption in such an approach is that later respondents to a survey are more similar to non-respondents than are earlier respondents. In this study, data collection will be conducted using telephone and a respondent can receive anywhere between 1 and 10 calls to complete an interview. Respondents will be divided into two groups (Early and Late respondents) based on the number of calls received. The exact definition of these two groups will be finalized after examining the distribution of the 'number of calls' needed to complete an interview for this study. Comparison of estimates (proportions or means of selected key variables like proportion of satisfied customers, for example) between these two groups will be carried out by testing the hypothesis of equality of proportions (or means). The analysis can be done using non-response adjusted weights but can also be done using the base weights. This process will help identify estimates that may be subject to non-response bias.

The selection of the key variables (or survey questions) for the comparison of "Early" and "Late" respondents described above will be done as follows. Based on the findings from the focus groups and regression analysis carried out in the previous round of this survey, the initial pool of possible candidate variables will include the following eight questions that are expected to be strong predictors of overall customer satisfaction. (Each of these eight questions uses a five-point scale, where 5 means strongly agree and 1 means strongly disagree and the respondent is asked to tell how much he or she agrees or disagrees with each statement as it applies to EBSA).

(i) EBSA treats me like a valued customer

(ii) EBSA is willing to work with me to make sure my needs are met

(iii) EBSA acts in a timely fashion

(iv) EBSA does what it says it will do

(v) EBSA services are available when I need them

(vi) EBSA is easy to reach

(vii) The information I receive from EBSA is clear and easy to understand

(viii) EBSA does its best to help me out

For this current round of this survey, we plan to fit a logistic regression model using the overall customer satisfaction question (Question: Taking into account all the information, products, and services you receive from them, how satisfied are you with EBSA overall?) as the dependent variable and the eight variables listed above as independent or predictor variables. For each of these questions using a 5-point scale, responses in the top two boxes (4 and 5) will be coded as '1' and a code of '0' will be assigned to responses in the bottom three boxes. So, a logistic regression model of the following form will be fit:

$$\text{Log}[P(Y=1)/P(Y=0)]=b_0 + b_1X_1+b_2X_2 \ldots + B_8X_8$$

[Y is the binary (0-1) coded variable corresponding to the overall satisfaction question i.e. Y=1 if the response to the overall customer satisfaction question is '4' or '5' and '0' otherwise; Similarly, X_1 through X_8 are the binary (0-1) coded variables corresponding to the eight questions listed above.]

The logistic regression model will be fit using software SUDAAN so that the sample design and the resulting sample weights can be taken into account. The main SUDAAN commands will consist of the following statements:

```
PROC RLOGIST data=XXXX FILETYPE=SAS DESIGN=STRWR;
NEST office_period;
WEIGHT FINALWT;
CLASS X1 X2 X3 X4 X5 X6 X7 X8;
REFLEVEL X1=0 X2=0  X3=0  X4=0  X5=0  X6=0  X7=0  X8=0;
MODEL Y = X1 X2 X3 X4 X5 X6 X7 X8;
```

As described in Section B1, sampling will be done independently within each two-week period within each office. For each two-week period (hereafter called 'period'), a simple random sample will be drawn from the list of all eligible cases (closed cases) for that period. So, the variable office_period (obtained by crossing the levels of regional offices and data collection periods) will represent all the strata. For example, if there are 11 offices and 18 data collection periods within

each office, the variable office_period will have 11X18= 198 levels and will represent all the strata. The design STRWR is proposed in the design statement based on the assumption that the sampling fraction within each stratum will be small (less than 10 percent). If that condition is not satisfied, the STRWOR (stratified without replacement) option will be used and necessary information (TOTCNT representing the total number of cases per stratum) will be included in the SUDAAN statements. The WEIGHT statement specifies the final weight variable. The CLASS statement defines the independent variables as categorical and REFLEVEL specifies the reference level for each of these variables. Finally, the MODEL statement specifies the dependent and independent variable for the logistic regression model to be fit.]

The variables found to be statistically significant (significance level less than .05) in the model (and the 'overall satisfaction' variable) will be selected as key variables for the purpose of comparison between 'early' and 'late' respondents. For each of these selected variables, the mean of the two groups ('early' and 'late' 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 'early' and 'late' respondents for a specific 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 the i^{th} respondent is 1 and '0' otherwise; W_i is the weight assigned to the i^{th} respondent and the summation in both numerator and denominator is over all 'early' 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:

$t=(p_1 - p_2)/SE\ (p_1 - p_2)$, where $SE\ (p_1 - p_2)$ is the standard error or the estimated square-root of 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 office_period;
WEIGHT FINALWT;
class early_late;
var  Y;
contrast early_late = (1 -1)/name = "respondent vs. nonrespondent";
print nsum t_mean p_mean mean;
```

The early_late variable will contain two distinct values (0-1 for example) to identify the two groups ('early' or 'late') 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).

Non-response bias analysis may also involve comparison of survey based estimates to known (i) Population Values and/or (ii) External Estimates that may be available. For this study, variables like Subject entry code, Closure types and Date open and Date closed (of inquiry) will be available for all cases on the sampling frame. It is, therefore, possible to compute the actual value of population parameters that are functions of these variables and compare those with the corresponding sample based (weighted) estimates. For example, there are several possible categories for Subject Entry Codes but those can be grouped into major categories like inquiries relating to (i) Health Care benefits (ii) Welfare benefits and (iii) Others. Based on population values, it will be possible to calculate the actual proportion (P_1, P_2 and P_3) of cases in each of those three categories. Those can then be compared to the corresponding survey estimates based on responses from the respondents only. Similar comparison can be made for proportions in each of the three Closure types or for proportions of inquiries requiring different amount of time to get resolved (derived based on Date Open and Date Closed variables). The hypothesis to be tested for such comparisons can be described as:

H_o: $P=P^*$ vs. H_1: $P\neq P^*$

where P^* is the computed population value for a certain proportion (or mean) and P denote the population parameter . For example, 'P' may denote the proportion of inquiries in the Health Care Benefits category, P* is the actual computed value of P based on population values and $p=\sum W_i y_i/\sum W_i$, where y_i is 1 if the value of Y (Inquiry subject type) for the i^{th} respondent is 1 (i.e., in the Health Care Benefits category) and '0' otherwise; W_i is the weight assigned to the i^{th} respondent and the summation in both numerator and denominator is over all respondents.

This test will be carried out based on a t-statistic = $(p - P^*)/SE(p)$, where 'p' is the corresponding sample based estimate of the population proportion 'P' and SE(p) is the standard error of p. The SUDAAN Design, Nest and Weight statements will be similar to those presented above for comparison between 'early' and 'late' respondents. The SE(p) will be computed using SUDAAN and then the t-statistic can be calculated using the formula given above based on the values of P* and p. The hypothesis will be rejected (or not) based on the observed significance level (less than 0.05 or not).

For the purpose of comparing survey based estimates with External estimates, the estimate of the proportion of satisfied customers (proportion of responses in the top two boxes on the overall satisfaction question) from the previous round of this survey can be used as an External estimate. The sample based estimate for the same proportion can be obtained from the current round of the survey and compared with that from the previous round. The comparison can be done based on a test of hypothesis of equality (as described above for comparison with population values). It is, however, noted that such comparisons may be subject to confounding factors (changes in methodology or actual change in satisfaction level over time etc.) and the results may not be influenced by non-response only. However, the results of such comparison along with other findings of the non-response bias analysis may be helpful for examining the issue of non-response bias.

The non-response bias analysis plan described above will be carried out at the full population level and there is no plan to carry out similar analyses for any sub-populations. There is a possibility that EBSA will not be able to conduct the nonresponse bias analysis in the first year of the study, however, it will ensure that the analysis is conducted in the future years of the study.

B.4. Describe any tests of procedures or methods to be undertaken. – None.

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
Dr. Alison Simon	Gallup Organization	202.715.3168
Dr. Manas Chattopadhyay	Gallup Organization	202.715.3179