B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method used. Provide data on the number of entities (e.g., establishments, State and local governmental units, households, or persons) in the universe and the corresponding sample in tabular form. The tabulation must also include expected response rates for the collection as a whole. If the collection has been conducted before, provide the actual response rate achieved.

This effort will involve collecting data from entities (companies and/or their facilities) that have a certified occupational health and safety management systems (OHSMS). Certifying Bodies (CBs) maintain lists of the entities that they have certified. CEO's subcontractor for this work, ERG, has been in contact with the CBs to request access to their customers for this data collection. To date, a total of eight have provided ERG with preliminary assurance that their customer lists would be available for a survey effort. Additionally, ERG has maintained its own list of certified entities based on publicly available information (e.g., press releases, etc.). ERG will supplement the list from the CBs using its own information.

The eight CBs that ERG has had contact with have certified a total of approximately 900 entities. Based on ERG's knowledge of the overall OHSMS certification market, CEO expects that the eight CBs represent a majority of the certified entities in the United States. Thus, CEO has assumed that the total universe may be closer to 1,500 entities.[1] In addition to the 900 entities on the lists from the eight CBs ERG has had contact with, ERG expects at least 2 to 4 additional CBs may provide their customer lists, adding an estimated 100 more entities to the total. Thus, CEO expects the total number of entities available on lists that can be used for the survey to be approximately 1,000. ERG plans to send the survey to all entities on these lists (discussed below).

ERG expects a response rate of 20 percent or more. The reason for the response rate assumption is that these certified entities will be contacted prior to the survey by the CB and the survey will be associated with each entity's certification. Thus, ERG expects the existing business relationship between the CB and the certified entities will result in at least a 20 percent response rate. A number of other DOL survey efforts also provide some guidance in developing a response rate estimate. The 2014 Employer Survey of the Short-Time Compensation Program achieved a response rate of 35 percent for employers that had participated in the program and three percent for employers that had not participated (Balducchi et al, 2015). The 2012 Family Medical Leave Act (FMLA) Worksite Survey achieved a response rate of 21 percent (Abt Associates, 2014); the 2018 FMLA Worksite Survey achieved a response rate of 7 percent (Abt Associates, forthcoming). Finally, the 2018 Survey of Employer Policies achieved a response rate of 17 percent (Westat, 2020). We have assumed a response rate above those of the 2018 FMLA and 2018 Survey of Employer Practices since the population we are surveying has a relationship with the CBs. The 2014 Employer Survey of the Short-Time Compensation Program had a 35 percent response among program participants; our population could be considered similar to a set of program participants. Nevertheless, we have assumed only a 20 percent

---

[1] Currently, no organization in the U.S. or worldwide maintains a master list of certified entities, and DOL has not made efforts to determine the total number of entities with certified OHSMS's. Thus, no information is available on the potential universe size; our estimate is based on ERG's professional judgement, knowledge of the certification market, and communications with CBs.

response rate as a conservative approach.

Table 1. Universe and Sample Information

Group	Assumed Population	Sample Frame	Expected Response Rate [c]	Expected Respondent Size
Entities with a certified OHSMS	1,500 [a]	1,000 [b]	20%	200

[a] ERG assumption; see text for details.
[b] ERG has preliminary agreements with eight CBs who are willing to provide ERG a list of their OHSMS customers. Combined, the customer base of these eight CBs totals approximately 900 entities. ERG further expects an additional 2 to 4 CBs will agree to participate, adding an estimated additional 100 entities to the total.
[c] ERG has assumed that certified entities will have a strong incentive to participate in the survey based on the existing business relationship between them and their certifier.

Selection method

ERG plans to select all entities on the lists that are provided; no statistical sampling methods will be used. The reason for selecting all entities is to allow for responses from entities across a large number of industries and types (e.g., sizes). CEO's goal in the survey is to develop a comprehensive picture of these entities that will allow for understanding the demographics of the universe, their motivations for seeking certification, and the perceived and actual costs and benefits of certification. Thus, selecting the largest set of respondents that can be obtained from the lists that are provided will allow for a fuller picture of the types of entities that obtain certification.

The sampling unit will the person designated as being responsible for the OHSMS at the entity. ERG expects this to be the person designated as the contact person by the CB for each entity.

2. Describe the procedures for the collection, including: the statistical methodology for stratification and sample selection; the estimation procedure; the degree of accuracy needed for the purpose described in the justification; any unusual problems requiring specialized sampling procedures; and any use of periodic (less frequent than annual) data collection cycles to reduce burden.

Stratification, sample size, and precision and accuracy

The survey will involve selecting all entities on the lists that are provided to ERG. Thus, no random selection will be implemented, and the survey will not involve any form of stratification.

ERG calculated the likely precision from this collection using an estimate of the maximum possible variance for different question types that are used in the survey. The survey primarily relies on yes/no questions and on 5-point scale questions with some 4-point questions as well. In each case, we calculated the minimally detectable effect size that could be detected from a sample of 200 respondents using a 95 percent confidence level and 80 statistical power in the R programming language.[2] The effect size is a measure of the level of precision adjusted for sample variance.[3]

---

[2] For our calculations for 5- and 4-point scales, we used the *pwr.t.test()* R function and for the yes/no questions (proportions), we used the *pwr.2p.test()* function. Both are contained withing the *pwr* library in R. These functions implement the calculations detailed in the tables of Cohen (1988).
[3] For the 5- and 4-point scales where we rely on t-tests, the effect size is the difference between the sample mean and

For the 5- and 4-point scales, we used the following formula to calculate minimal detectable effect size (*d*):

$$d = \frac{\overline{x}_1 - \overline{x}_2}{s_x}$$

where *d* is the effect size, $\overline{x}_1$ and $\overline{x}_2$ are the two test values and s_x is the pooled standard deviation. We set $\overline{x}_2$ to be the mid-point of the scale and assumed that half of the respondents would select the lowest value and half would select the highest value.[4] With one half selecting the highest value and the other selecting the lowest value, the estimated standard deviation for the 5-point scale would be 2.005 and for the 4-point scale 1.504.[5] Using the R statistical package and the parameters identified above for power and confidence, the effect size (*d*) we can detect with a sample of 200 respondents is 0.1991. Calculating the value for $\overline{x}_1 - \overline{x}_2$ using our estimates for s_x and *d*, we find that the sample of 200 respondents will allow for detecting non-random differences between two test values when the actual difference is 0.399 units for the 5-point scale and 0.299 units for the 4-point scale (assuming 95 percent confidence and 80 percent power).

For the proportional questions, the minimal detectable effect size (*h*) is defined as:

$$h = 2\left[\sin^{-1}\left(\sqrt{p_1}\right) - \sin^{-1}\left(\sqrt{p_2}\right)\right]$$

where $\sin^{-1}$ is the arcsine function and p_1 and p_2 are test values. Following standard sample size calculation approaches, our assumed maximum variance involves setting a test value (p_2) equal to 0.5. The effect size in the R package is calculated to be 0.2801. Inserting those values into the equation above results in an estimate of 0.638 for p_1. Thus, for the proportional questions, we can detect a non-random difference between two test values when the actual difference between the values is 13.8 percentage points (assuming 80 percent power and 95 percent confidence).

CEO has determined that these levels of precision will more than satisfy the objectives of this study. Specifically, CEO is interested in better understanding the set of entities who have attained certification of their OHSMS to voluntary consensus standards, and their perceived and actual benefits and costs of certification.

Unusual Problems Requiring Specialized Sampling Procedures

None are required.

Periodic Data Collection Cycles

---

a test value divided by the sample variance. For the yes/no questions where we rely on proportional testing, the effect size involves a trigonometric transformation of the value (see text).

[4] This results in the largest possible estimate for the variance.

[5] For the 5-point scale, all respondents are 2 units from the mean value (i.e., 5 – 3 or 1 – 3). The variance is then calculated as the squared distance from the mean multiplied by the number of respondents (200) divided by the number of respondents minus one (199); the standard deviation of the sample is then the square root of that number. For the 4-point scale, each respondent would be 1.5 units from the mean (i.e., 4 – 2.5 or 1 – 2.5).

This request is for a one-time data collection.

Data Analysis Plan

ERG plans to perform two primary sets of analyses. First, ERG will develop a set of summary statistics for each question used in the survey. Second, ERG will perform analyses of outcome variables we have included in the survey. Specifically, the survey contains a set of agreement scale questions related to perceived outcomes associated with certified plans and a set of questions that asks for data on injuries and illnesses and workers' compensation outcomes. ERG will perform analyses on these three sets of outcomes by performing regression (linear and/or non-linear, as needed) analyses to relate these outcomes to other data collected as part of the survey.

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

Maximizing response rates

To maximize the response rate, ERG will perform the following activities:

- Follow good implementation design – ERG will send out pre-notification emails to all respondents followed by an email with the survey link. ERG will then send up to three reminder emails to those who have not responded.
- Document out of scope – ERG will document all bounce-back emails and indications by respondents they are not in scope for the survey.
- Leverage relationships with CBs – ERG will ask the CBs to "advertise" the survey prior to the send-out in order to make the in-scope entities aware of the effort and to validate the effort among those entities.

Finally, as we have noted above, entities with certified OHSMS's generally maintain an ongoing business relationship with the CBs: i.e., the CB is called on periodically to verify the certification and the entities are audited on a semi-regular basis.[6] Thus, we expect that the entities have a strong incentive to provide a response to the survey to be perceived as a "good actor" by the CB. Furthermore, organizations that have invested time and effort to get certified presumably would have some interest in providing input into a study of this type. ERG's discussions with the CBs suggest that the major concern of certified entities is maintaining the privacy of the data. In the survey design ERG has been sensitive to this concern. ERG will offer respondents an assurance that the data collected during this survey will remain private, which should obviate this apprehension.

Nonresponse bias analysis plan

---

[6] For example, to maintain certification to the ISO 45001 standard, an organization undergoes an initial certification audit by its CB. The certification is valid for three years, but the organization must undergo a "surveillance" audit each year following certification.

ERG will compare the respondents to the non-respondents based on data that are available from the CBs. This will include:

- CB source. ERG will be compiling the sampling frame from multiple sources and will assess response rates from different CB sources.
- Industry sector
- Geographic location
- Date of certification
- Number of employees

ERG will perform the nonresponse bias analysis by comparing the full set of respondents to all entries in the sampling frame, as well as performing non-response analyses among entities from specific CB sources.

ERG does not plan to compare the set of respondents to larger population characteristics since (1) there are no reliable data on the population of entities with certified OHSMS[7] and (2) other populations of entities (e.g., Census Bureau data) are likely to have different characteristics.

<u>Weighting for Nonresponse</u>

To account for the effect of nonresponse on the data analysis we perform, ERG will calculate nonresponse weights using a weighting class adjustment (Lohr, 1999; Chapter 8). ERG will the lists provided by the CBs to formulate classes based on either entity employment size or sector (e.g., two-digit NAICS code).[8] In weighting class adjustment, the response probability for any member of a class (θ_c for class c) is calculated as (Lohr, 1999):

$$\theta_c = \frac{\sum \text{ of weights for respondents} \in \text{class } c}{\sum \text{ of weights for selected sample} \in \text{class } c}$$

Given that we are selecting the population as the sample in this survey, the weight for each respondent and each member of the sample is one. Thus, the calculation of θ_c reduces to:

$$\theta_c = \frac{\text{Respondents} \in \text{class } c}{\text{Selected sample} \in \text{class } c}$$

The weight for each member of class c is then calculated as the inverse of θ_c.

As noted, we plan to use either employment size or sector to form the weighting classes. For employment size, ERG will use the standard BLS employment size classes (1 to 4 employees, 5 to 9 employees, 10 to 19 employees, 20 to 49 employees,50 to 99 employees, 100 to 499 employees, and 500 or more employees) or some compacted subset of those classes.[9] For the NAICS sectors, ERG will use two-digit NAICS codes as the basis for the size classes as a default. However, some three- or four-digit NAICS may have significant representation among

---

[7] As has been noted, a key outcome from this survey will be help better characterize this set of entities.

[8] The choice between the two will depend on the availability and quality of those data from the CBs.

[9] For example, given that we are interested in certified OHSMS, the smaller size classes may be less relevant, and we may want to combine some of the smaller size classes to have a meaningful number in each.

certified entities (e.g., hospitals, NAICS 622); in those cases, ERG would formulate a class specifically for those sectors.[10] Selecting between these two aspects will depend on (1) the quality of the data from the CBs (e.g., how complete and reliable are the data on each aspect?) and (2) the results of our nonresponse bias analysis discussed above.

ERG would apply the weights we form from the weighting class adjustment in calculating sample means and sample totals.

<u>Missing Data</u>

ERG will use a combination of both case wise and pairwise deletion to handle missing data. Pairwise deletion will be used in exploratory analyses and in assessing the overall characteristics of the collected data. This will include calculating summary statistics and in performing some basic correlation analyses. The choice of pairwise deletion in this case is to allow for use of the largest number of respondents in each analysis.

Case wise deletion will be used as the analysis proceeds to define a specific set of analyses that ERG will focus on in the final report. Specifically, we will remove respondents who do not provide data to questions that are included in our final analysis. Thus, the final analysis will include a set of respondents who have provided complete data to the questions we are analyzing.

ERG will also use case wise deletion when reporting on summary statistics for sets of questions that appear in the same question bank within the survey. For example, the survey includes 19 items related to outcome to be rated on an agreement scale. ERG will use case wise deletion for providing summary statistics to question banks such as these to ensure the same set of respondents are included across the items.

<u>Adequacy for intended uses</u>

This study represents a scoping study to some degree to fully understand what types of entities are certified, the reasons they became certified, and the associated perceived benefits and costs of certification. As such, we have decided to collect data from all entities on the lists to better understand the full set of entities who have certified OHSMS's, including identifying relevant sub-populations among the larger population. Also, as noted under B2 above, the estimated sample size of 200 will provide a decent level of precision for estimating mean values among 4- and 5-point scaled questions and a strong level of precision among yes/no questions.

ERG also calculated the impact of a worst-case response rate of 10 percent (i.e., 100 total responses). Under that worst-case assumption, CEO would be able to detect a difference between two test values when the difference between the values in the population is:

- 0.566 points for a 5-point scale

- 0.425 points for a 4-point scale

---

[10] For example, for NAICS 62, this may mean have a "NAICSC 622" class (hospitals) and a "all other NAICS 62" class.

6

- 19.3 percentage points for a difference in proportions.

Although these worst-case scenarios limit what can be learned, they still provide CEO with the ability to betted understand the types of entities that earn certification. As noted above, the purpose of the survey effort is to better understand the characteristics of this group of entities. Even at this worst-case scenario, the collected data would provide information to CEO on these entities.

4. Describe any tests of procedures or methods to be undertaken. Tests are encouraged as effective means to refine collections, but if ten or more test respondents are involved OMB must give prior approval.

As part of project ERG is undertaking for CEO, ERG will be visiting up to nine certified entities to assess their systems. During this time, ERG will also pre-test the instrument with those entities. ERG also expects to review the instrument with CBs to ensure it will be something that certified entities would be able to complete. We are not seeking approval on these interactions, however, since we are limiting the discussions with certified entities to no more than nine, and the information we are seeking from these entities is different than what we are seeking from the CBs. With the certified entities, we are conducting a pre-test, but from CBs we are seeking a document review.

5. Provide the name and telephone number of individuals consulted on the 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.

CEO has contracted with Eastern Research Group, Inc. (ERG) of Lexington, MA, to design the survey instrument, develop the sampling approach, implement the survey, and analyze the resulting data collected. The survey design team included the following individuals:

Dr. Lou Nadeau (781) 674-7316; lou.nadeau@erg.com
Jeff Cantin, (781) 674-7315; jeff.cantin@erg.com
Mike Seymour, (540) 454 1629, mseymour@icarusenv.com

References Cited

Abt Associates, 2014. *Family and Medical Leave in 2012: Technical Report,* Report prepared for the U.S. Department of Labor. https://www.dol.gov/sites/dolgov/files/OASP/legacy/files/FMLA-2012-Technical-Report.pdf.

Abt Associates, forthcoming. *FMLA 2018 Worksite Survey*. Report prepared for the U.S. Department of Lapor

Balducchi, D. E., Wandner, S. A., Goger, S., Miller, Z., Shetty, S., Agbayani, C., Eucogco, J. (2015). *Employer views about the short-time compensation program: A survey and analysis in four states* (Final report.). Washington, DC: U.S. Department of Labor, Employment and Training Administration.

Cohen, Jacob., 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Lohr, Sharon, 1999. *Sampling: Design and Analysis*, Duxbury Press, Pacific Grove, CA.

Westat, 2020. *Survey of Employer Policies on the Employment of People with Disabilities*, Report prepared for U.S. Department of Labor, Chief Evaluation Office (CEO) and the Office of Disability Employment Policy (ODEP). June.