

2011 FAMILY MEDICAL LEAVE ACT (FMLA) EMPLOYER AND EMPLOYEE SURVEYS

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

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

Employee Survey:

Employees aged 18 or older who live in the United States, have a telephone (landline or cellular), and who are employed for pay in the private or public sector (self-employed are excluded) at any time during the 12 months prior to the interview will constitute the known respondent universe from which the sample for the 2011 Family and Medical Leave Act of 1993 (FMLA) Employee survey will be taken. According to the 2009 National Health Interview Survey, 98.7 percent of U.S. adults live in a household with landline or cellular telephone service. According to the March 2010 Current Population Survey, 57.0% of the entire adult population is employed (excluding self-employed). The estimated total size of the eligible respondent universe is thus 130,277,439 adults.

Exhibit 1. Respondent Universe and Sample Size for the 2011 FMLA Employee Survey

Number of persons in the universe covered by the data collection	Landline RDD sample	Cellular RDD sample	Total sample size
130,277,439	2,100	900	3,000

Respondents will be sampled through a dual frame, landline and cellular random digit dialing (RDD) telephone design. We project that 2,100 interviews will be completed with respondents sampled through the landline frame, and 900 interviews will be completed with respondents sampled through the cellular frame, for a total of 3,000 interviews. Numbers for the landline sample will be drawn with equal probabilities from active blocks (area code + exchange + two-digit block number) that contained one or more residential directory listings. The cellular sample will be drawn through systematic sampling from 1,000-blocks dedicated to cellular service according to the Telcordia database.

In the survey screener, interviewers will determine whether the household contains at least one person 18 years of age or older who has been employed (excluding self-employed) during the last 12 months. For all persons in the household meeting these criteria, the interviewer will attempt to determine if they have taken (or needed without taking) family or medical leave during the reference period.

Each eligible adult will be classified into one of three family or medical leave groups: leave-needer, leave-taker, or employed-only. One eligible person will be selected from the screened household for the extended interview.

Respondent selection will be conducted in three stages. Stage 1 determines from which family or medical leave group represented in the household the respondent will be selected. For households where multiple groups are represented, the leave-needer and leave-taker groups will be selected at a higher rate than the employed-only group because their incidence rates are significantly lower. Stage 2 sub-samples households headed for an employed-only interview. Past studies suggest that about 80 percent of U.S. workers belong to the “employed-only” group. Completing all of these cases would yield over 6,000 interviews but, consistent with the 1995 and 2000 surveys, only about 1,300 completes are needed for the analysis. We will, thus, administer the extended interview with a sub-sample of households in which the employed-only group was selected at Stage 1. Stage 3 selects a random adult from the family or medical leave group identified in Stage 1 as the extended interview respondent. The details of this algorithm are provided below.

STAGE 1: Select an FMLA group

- If all adults are of one family or medical leave group, that group is selected. Skip to STAGE 2.
- If the household has a leave-needer and a leave-taker, select the leave-needer group with 90 percent probability and the leave-taker group with 10 percent probability.
- If the household has a leave-needer and employed-only adult, select the leave needer-group with 90 percent probability and the employed-only group with 10 percent probability.
- If the household has a leave-taker and employed-only adult, select the leave taker-group with 90 percent probability and the employed-only group with 10 percent probability.
- If the household has a leave-needer, leave-taker and employed-only adult, select the leave-needer group with 80 percent probability, the leave-taker group with 10 percent probability, and the employed-only group with 10 percent probability.

STAGE 2: Subsample 20% of the households where the employed-only group was selected

- If the leave-needer group or leave-taker group was selected in Stage 1, skip to Stage 3.
- If the employed-only group was selected in Stage 1, select the household for the extended interview with 20 percent probability. If the household is not selected for the extended interview, terminate the call and assign final disposition as completed screener.

STAGE 3: Select a household member from the selected family or medical leave group

- Select a random person (with equal probability of selection) from the family or medical leave group selected in Stage 1.

The Stage 1 family or medical leave group selection rates reflect the fact that leave-needers and leave-takers are rare groups and we need to select them with a higher probability. The Stage 2 subsampling rate is based on the experience of the 2000 Employee Survey. From the survey report, we estimate that about 6,450 employed-only adults were identified in the screener but only 1,126 were interviewed (17.5 percent). We plan to implement a similar sub-sampling approach.

We will closely monitor the yield of extended interviews in each family or medical leave group throughout the field period. Depending on the yields observed in the early replicates, we may need to modify the group selection rates for later replicates. There are two reasons for this. The population incidences of the family or medical leave groups may have changed since 2000. This makes it difficult to anticipate the optimal group selection rates for households containing adults in multiple family or medical leave groups. The performance of the group selection rates will also depend on the distribution of households by family or medical leave group (e.g., How many leave-needers live with a leave-taker? How many live with an employed-only adult?). Unfortunately, no available data from the 2000 FMLA Employee Survey address this issue.

Based on the incidences reported in 2000 survey, we expect the study design to yield extended interviews with approximately 250 leave-needers, 1,440 leave-takers, and 1,310 employed-only adults. We will evaluate the results of the early sample replicates to determine whether the group selection rates discussed above are on pace to achieve these target sample sizes. If necessary, we will modify the group selection rates so that the data collection can be completed within the study budget and yield sufficient cases in each family or medical leave group.

The group selection rates will be pre-determined before each replicate is released. The rates may differ from one replicate to the next but will be uniform for all households within a given

replicate. We will document which selection rates are implemented for each replicate. This information will be incorporated into the weighting so that computations of the probabilities of selection are accurate for each replicate.

If the selected respondent is not the adult who responded to the screener, the interviewer will ask to speak with the selected respondent before administering the extended interview. If the selected respondent is present and available, the screener respondent would simply hand off the phone to the selected respondent. If such a handoff is not possible, the interviewer will ask for the date and time of day when the selected respondent will be available. Interviewers will also inquire as to the best phone number to reach the selected respondent. This procedure will be implemented for both the landline and cell phone samples.

While within household selection and resulting handoffs are quite common in landline surveys, they are less common in cell phone surveys. Traditionally, residential landlines have been viewed as a point of contact for the entire household. Cell phones, by contrast, are commonly viewed as a personal device, though some sharing does occur. Studies have demonstrated that within household selection procedures can be implemented for cell phone samples though, not surprisingly, response rates are lower when trying to handoff to another person in the household (AAPOR, 2010). For the Employee Survey, we propose to screen all adult household members in the cell sample as well as the landline sample because of the very low incidence rate of leave takers and leave needers. We view the challenge of a handoff as more manageable than the inefficiency of excluding a leave needer or leave taker in a household simply because their spouse or partner's cell phone was sampled and not their own (for example).

In 2000, the overall response rate to the FMLA Employee Survey was 58.3 percent. Given the continuing decline in survey response rates, the response rate to the 2011 Employee Survey is expected to be lower. A reliable estimate of the change in national RDD response rates over time comes from Curtin, Presser, and Singer (2005) who report that the response rate to the Survey of Consumer Attitudes (conducted by the University of Michigan Survey Research Center) declined, on average, 1.5 percentage points per year between 1996 and 2003. Applying this rate to the 2011 FMLA Employee Survey, the estimate for the response rate would be approximately 42 percent. This assessment is based on the response rate formula that was used in the previous surveys, which will need to be modified somewhat to account for the dual frame design. Every effort, within the specifications of the study, will be made to exceed this expectation.

Employer Survey:

The potential respondent universe for the 2011 FMLA Employer Survey consists of all private-sector business establishments excluding self-employed without employees, government and quasi-government units (federal, state, and local governments, public educational institutions, and post offices). As in the 1995 and 2000 surveys, the sampling frame will be created from the Dun's Market Identifiers (DMI) file, which provides all essential frame information (e.g., employee size, NAICS code, contact information) for 15.2 million private business establishments. This DMI file is considered the most comprehensive commercially available business list.

Given the detailed nature of the questionnaire, (which includes some questions that may require reference to company administrative records) it will be necessary to identify the human resources director or the person responsible for the company's benefits plan. Therefore, the employer survey includes two components: the first component includes a telephone call to sampled establishments to: 1) determine the eligibility of the establishment/business; and 2) determine the name of the person who is the most appropriate to complete the survey questionnaire given the nature of the information requested. This person is referred to as the "key informant." Please see Attachment C for the Employer Survey screener.

Sampled establishments meeting one or more of the following three criteria will be treated as ineligible for the survey. These include: 1) telephone recruitment efforts cannot confirm that the establishment is open/in business during the field period; 2) an employee (not necessarily the key informant) reports that it is not a private sector business; or 3) the establishment owner is self-employed without employees.

We will use a stratified sampling procedure to select the sample of 1,800 establishments. The sampling strata are defined by the cross-classification of employment size and North American Industry Classification System (NAICS) grouping. The employment size classes are: (1) 1-10; (2) 11-24; (3) 25-49; (4) 50-99; (5) 100-250; (6) 251-999 (7) 1,000+. These size classes reflect the employer size breaks used in the 1995 and 2000 survey reports. While these classes, as defined in the DMI file, will be used for sampling, all establishments will be asked the exact number of employees in the survey. NAICS codes will be combined to create four industry groups as follows:

Exhibit 2. NAICS Groupings for 2011 FMLA Employer Survey

Group	NAICS Codes
NAICS Group I	Manufacturing (31-33)
NAICS Group II	Retail Trade (44-45)
NAICS Group III	Information (51); Professional, Scientific, and Technical Services (54); Administrative Support; Waste Management and Remediation Services (56); Educational Services (61); Health Care and Social Assistance (62); Arts, Entertainment, and Recreation (71); Other Services, except Public Administration (81)
NAICS Group IV	Agriculture, Forestry, Fishing, and Hunting (11); Mining (21); Construction (23); Utilities (22); Transportation and Warehousing (48-49); Wholesale Trade (42); Accommodation and Food Services (72); Finance and Insurance (52); Real Estate and Rental and Leasing (53); Public Administration (92); Management of Companies and Enterprises (55)

This design yields 28 sampling strata (7 size categories times 4 NAICS groups). To arrive at the sample size for each stratum, we first allocate the sample to employment size classes proportional to the square root (\sqrt{X}) of the aggregate number of employees working for establishments in the class (X). This is the optimal allocation for computing per employee estimates from an employer survey. Given that most tabulations from the Employer Survey will be weighted to a “per employee” basis, this is the appropriate approach. This allocation method allows establishments with a large number of employees to be selected at a higher rate than establishments with fewer employees.

To demonstrate the advantage of the \sqrt{X} -proportional allocation, it may be helpful to contrast it with the simple N -proportional allocation, where N refers to the number of establishments in the stratum. Under N -proportional allocation, all establishments would be selected with the same probability. An establishment of three employees would have the same probability of being selected as an establishment with 30,000 employees. When the purpose of the study is to make inference to the population of *establishments*, N -proportional allocation would be the preferred approach. In the FMLA Employer Survey, however, many key variables are related to *employee* size (e.g., number or percentage of employees covered by the FMLA). Large

establishments are more important for estimating such variables than smaller establishments. That is, the variance in an employment related variable is concentrated more in the strata of large establishments than in the strata of smaller establishments. To reduce the variance (i.e., increase the precision) of an estimate of such a variable, the strata of large establishments is sampled with a higher probability than the probability used for the strata of smaller establishments. This way, the influence of large establishments on the variance of the estimate can be reduced.

The \sqrt{X} -proportional allocation is one way of over-sampling larger establishments. The square root has the desirable effect of dampening the influence of the X value. Consequently, the over-sampling under \sqrt{X} -proportional allocation is less severe than it would be under X -proportional allocation.

Under this \sqrt{X} -proportional allocation, we would expect to complete interviews with about 243 establishments in the 25-49 employee size group and about 178 interviews with establishments in the 50-99 employee size group. These two size classes are important because the 2000 FMLA Employer Survey report made a point of comparing FMLA covered establishments with 50-99 employees to non-covered establishments with 25-49 employees. The expected sample sizes need to be increased in order to support a reliable comparison. In 2000 these two size classes were oversampled, and approximately 300-400 establishments were interviewed in each class.¹ For the 2011 Employer Survey we will address this in an analogous fashion by first conducting the \sqrt{X} -proportional allocation with a slightly reduced total sample size, and then increasing the sample size in these size classes to 300. As shown in the far right column of Exhibit 3, this method will yield approximately 300 interviews in each of the 25-49 and 50-99 size classes and still yield an overall sample size of 1,800 establishments.

The second step in allocating the Employer Survey sample is to determine the sample sizes of the NAICS groups within each employment size class. For this task, we propose to use the simple proportional allocation method, which allocates sample proportional to the numbers of establishments in the NAICS groups within the size class. Exhibit 3 shows how this two-step allocation algorithm will work. The final allocation will be based on the most up-to-date frame

¹ The exact number of cases in the 25-49 and 50-99 size classes in the 2000 survey cannot be determined exactly from the public dataset. The “number of employees” variable was excluded, presumably to protect establishment confidentiality. Based on several coarse categorical variables on employee size that are in the dataset, we estimate that

totals available for the DMI, as of early 2011. The figures in Exhibit 3 are presented as an illustration of the process.

Exhibit 3. Allocated Sample Sizes for the 2011 FMLA Employer Survey

Size class	NAICS group				Total
	NAICS I	NAICS II	NAICS III	NAICS IV	
1-10	13	36	157	100	306
11-24	9	24	107	68	208
25-49	12	35	154	98	300
50-99	12	35	154	98	300
100-250	11	32	138	88	268
251-999	9	26	111	71	217
1,000+	8	24	103	66	201
Total	75	212	923	591	1,800

The response rates to the FMLA Employer Survey in 1995 and 2000 were 73.2 percent and 65.0 percent, respectively. If one were to extrapolate the change in response rates from the previous rounds, the estimate for the 2011 response rate would be approximately 47 percent. As discussed above, this is an inexact approximation considering that only two historical data points are available. Since the 2011 protocol for the Employer Survey is similar to that implemented in the previous rounds, this estimated response rate of 47 percent is the best available approximation. Every effort, within the specifications of the study, will be made to exceed this expectation (including a new internet response option).

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

- **Statistical methodology for stratification and sample selection;**
- **Estimation procedure;**
- **Degree of accuracy needed for the purpose described in the justification;**
- **Unusual problems requiring specialized sampling procedures; and**
- **Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

Employee Survey:

approximately 300-400 establishments were interviewed in each class.

In addition to a base weight reflecting the probability of selection, weights will include: 1) a sampling frame integration weight; 2) a non-response adjustment for people who are included on the household roster but do not complete the interview; and 3) a post-stratification adjustment to independent population controls by age, gender, education, race, Hispanic ethnicity, region, and employment status. Estimates of sampling variance reflecting the variance in the weights will be made using replication methods. The weighting and estimation procedures are explained in detail in Attachment D.

Regarding the accuracy of estimates, using variance estimates from the 2000 FMLA surveys, we report calculations of the expected accuracy of estimates from the 2011 survey. One continuous variable in both the 2000 and 2011 questionnaires is the total number of days that the respondent took off from work for the longest leave in the reference period. In 2000, the mean for this variable was 31.06 days. The variance of the mean is 5.2959 when using appropriate complex survey procedures versus 3.2382 under the false assumption of a simple random sample design. Taking the ratio of these variances, the design effect for that estimated number of days taken off work is 1.6355.

In 2000, the responding sample size for this question was 1,254, and the actual standard error of the mean was 2.301. In 2011, we will be fielding a large sample. We expect to interview about 1,443 leave-takers who will answer this question. Given this larger sample size, and assuming similar population variance and design effect, we expect the standard error for the estimated number of days taken off work in the longest leave to be 2.145.

The dual-frame RDD design that we will implement in 2011, however, will likely yield a slightly larger design effect. Adults living in cell phone only households will be somewhat underrepresented in the survey (15% of the sample but 24.9% of the population) and, thus, will need to be weighted up. Also, the mixing factor applied to respondents in dual service households (cell and landline) will increase the design effect slightly. The larger survey sample size in 2011 will help to offset this larger design effect to some extent. Assuming a design effect of 1.8000 rather than the 1.6355 observed for this estimate in 2000, the expected standard error for the number of days taken off work is 2.251.

A similar analysis can be done for estimated proportions. Both the 2000 and 2011 questions contain the question, "Did you receive pay for any part of your leave?" In 2000, some 64.8% of leave takers reported receiving pay for at least part of their leave. The standard error computed using complex survey software is 1.821, and the design effect is 1.846. If we assume a slightly

larger design effect in 2011 (2.000) and account for the larger expected item sample size (1,443), the expected standard error for the estimated proportion receiving pay for their longest leave is 1.778.

The accuracy of subgroup estimates is also of interest. In the FMLA Employee Survey, the three main subgroups are leave-needers, leave-takers, and other employees. One question asked of all three groups in both the 2000 and 2011 questionnaires is, “Have you ever heard about the federal Family and Medical Leave Act?” As mentioned above, the sample sizes in 2011 are expected to be larger, but the design effect is also expected to be slightly higher. The observed standard errors for each subgroup in 2000 are reported in the top half of Exhibit 4. The expected standard errors for 2011 are reported in the bottom half. Even with our oversampling of leave-needers and leave-takers, we expect to survey many more leave-takers and employed only workers than leave-needers, and so the expected standard errors for those subgroups are much smaller (1.8983, 1.7151, and 3.8940, respectively).

Exhibit 4. Precision of Subgroup Estimates in 2000 and 2011 for Percent Reporting That They Have Ever Heard about the Family Medical Leave Act

	Leave-taker	Leave-needer	Employed only
2000 Observed Values*			
% Yes, have heard of FMLA	63.1	53.1	58.4
n	1,221	201	1,118
DEFF	2.0831	1.3600	1.4422
se(p) assuming SRS	1.3809	3.5199	1.5080
se(p) actual	1.9936	4.1150	1.8110
2011 Expected Values			
DEFF	2.2331	1.5100	1.5922
n	1,443	248	1,315
se(p) assuming SRS	1.2703	3.1689	1.3592
se(p) actual	1.8983	3.8940	1.7151

* Source 2000 FMLA Employee Survey

We will release the sample for interviewing in replicates, which are small random sub-samples of the larger sample. This technique enables us to compute preliminary estimates of the incidence rates for the three worker groups, part-way through the field period. For example, adjustments in the proportion of respondents interviewed who are employed only (but neither a leave-taker nor a leave-needer) may be made, if the preliminary estimates suggest that sample eligibility and response rate assumptions are inaccurate. It will be possible to monitor these

assumptions on a daily basis using Computer Assisted Telephone Interviewing (CATI) which allows the instantaneous transmission of finalized cases.

Employer Survey:

We adopt the stratification by establishment size approach to address the heavily skewed distribution of the survey population on this variable. Large employment size establishments have a great influence on the estimates of employment size related variables (e.g., total number of employees covered by the Act). By over-sampling larger establishments, we optimize the survey for estimates based on employees (which historically have been the focus of the analysis) as opposed to estimates based on employers. In this context, optimization means maximizing the precision of the estimates (i.e., minimizing the standard errors). The over-sampling is done using the size stratification and higher sampling fractions for the larger size strata than smaller size strata. This procedure also ensures adequate sample sizes for analysis of the survey data by size category.

Industry stratification was used in the 2000 survey, albeit with previously used SIC codes rather than the currently used NAICS codes. In 2000, the stratification featured five broad SIC groups, but in the analysis the results were collapsed to four groups (Manufacturing, Retail, Services, and Other). In the 2011 Employer Survey, the stratification will be based on these four broad groups that were used in the previous analysis. This NAICS stratification will ensure reasonably large sample sizes for these industry groups and facilitate data analysis by industry group.

In order to obtain valid survey estimates, estimation will be done using properly weighted survey data. Weighting starts with the calculation of sampling base weight as the inverse of the selection probability based on the sample design. There will inevitably be some non-respondents to the survey and weighting adjustments will be used to compensate for them. Even among respondents, some will fail to provide all the survey items in the questionnaire and some item missing values will result. As was done in 2000, we will impute item missing values. Weighting adjustment for unit non-responses will be carried out within each weighting class appropriately created to minimize the possible bias due to non-response. The final weight is then obtained by multiplying the post-stratification adjustment factor calculated to reflect more reliable information on employment size available from the U.S. County Business Patterns

published by Census Bureau. Such adjustment helps to reduce coverage error due to deficiency in the sampling frame and improve the precision of estimates.

The survey will be used primarily to compute descriptive statistics such as estimated percentages and totals. Those computations will be done using the weighted data. Standard errors of the estimates will also be calculated by a method appropriate for complex survey data. For more detail about the weighting and estimation procedure see Attachment D.

The sample sizes in 2000 and 2011 (1,839 and 1,800) are quite similar as is the expected design effect. The 2000 and 2011 Employer Surveys use the same sampling frame and a quite similar disproportionate, stratified sample based on industry code and establishment size. In evaluating the expected accuracy of estimates in the 2011 Employer Survey we can draw on the results from the 2000 survey.

The continuous variables collected in both the 2000 and 2011 questionnaire are used not as stand-alone outcomes, but rather as numerators for percentages. For example, both surveys ask, “How many employees at this location have been denied leave because they used their entire time allotment covered by FMLA?” In the analysis, this value would be divided by the total number of relevant employees at the location. Given this, we consider the accuracy of estimated proportions in the Employer Survey, rather than the accuracy of continuous variables.

Exhibit 5 reports the precision of subgroup estimates observed in the 2000 Employer Survey for the estimate proportion covered by the FMLA. The subgroups are the major SIC classification codes. In 2011, we will instead use major NAICS codes, but the nature of the analysis is the same. The estimated proportion of establishments covered by the FMLA are 25.9% (s.e.= 6.75) for Manufacturing, 11.4% (s.e.=1.84) for Retail, 9.1% (s.e.=1.72) for Service, and 9.9% (s.e.=2.38) for All other industries. Given the similarity between the 2000 and 2011 sample design, there is no reason to expect that the design effect and/or standard errors in 2011 will be higher than in 2000. The results from 2000 shown in Exhibit 5 are the best available approximations for the precision levels expected in 2011.

Exhibit 5. Precision of Subgroup Estimates in 2000 for Percent of Establishments Covered by Family Medical Leave Act

	Manufacturing	Retail	Service	All Other
2000 Observed Values*				
% Covered by FMLA	25.9	11.4	9.1	9.9
n	328	346	654	473
DEFF	1.7610	0.6295	2.0813	4.2541
se(p) assuming SRS	2.4189	1.7085	1.1246	1.3732
se(p) actual	6.7466	1.8401	1.7178	2.3757

* Source 2000 FMLA Employer Survey

Preliminary Analysis:

DOL/WHD has requested “early” results for preliminary regulatory filings; these results will be updated with the “full period” results when data collection is complete. These early results have been requested for July 31, 2011 from the contractor. The current plan is for the contractor to deliver a first draft of the early preliminary results trial results by June 15, 2011; give comments; and then have them revise with the final preliminary analysis of the “early” results. With the revision, we will also update the analysis file with another (approximately) two weeks of data. These revised “early” results will be used in DOL preliminary regulatory filings.

Planned methods are expected to be identical for the “early” and “full period” data; that is, the same weighting and estimation methods. More collapsing of post-strata may be required for the early period relative to the later period due to smaller sample sizes, but this is not likely to have a large effect on the estimates (see Attachment D for collapsing rules). The weighting and estimation methods proposed are the best approach to non-response and the possibility of differential non-response. In particular, changes to the screener and the reference period allow better control totals (from the March CPS). The contractors consider this to be a crucial improvement of non-response adjustment relative to the 2000 surveys (see below for more information on non-response).

The accuracy of the data will depend on the date of receipt of OMB approval; in particular, 30 day vs. 60 day review. If 30 day review is required, we project that we will have approximately three-fifths of the data for the early results. If 60 day review is required, we project that we will have approximately two-fifths of the data for the early results. Standard power formulae imply that with two-fifths and three-fifths of the data, our standard errors will

be 58 percent larger (square root of 2.5) and 29 percent larger (square root of 1.67), respectively. These figures assume a common rate of completed interviews for the entire field period. In fact, we would expect slightly more data as of these dates since the last month will be focused on non-response refusal and similar issues. During that period, we expect fewer completed interviews than during earlier periods. On the other hand, non-response conversion cases may be different from cases completed earlier.

In discussions with DOL/WHd about “early” data, the contractor noted the concerns about smaller samples (and therefore larger variances) and also concerns about differential response of non-respondents. However, DOL/WHd indicated that even such “early” data would be useful for their regulatory process. The contractor has stated that all written and oral deliverables based on these “early” data, will include explicit and strong caveats; i.e., that full results will differ from “early” results, both because of more data and because of the possibility of differential response.

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.

Employee Survey:

Several recruitment strategies will be used to increase the response rate.

1. Interviewers will make 10 attempts to contact landline cases. More calls will be attempted if contact is made with an eligible household but the interviewer is asked to call back later.
2. Interviews will be conducted during various times of the day and seven days a week to increase the likelihood of finding the respondent at home.
3. Respondents will be provided with the option of scheduling the interview at the time that is convenient for them.

4. For soft-refusals, “interview converters” who have extensive training in telephone interviewing and converting non-responders will be used to increase the response rate.

This recruitment design may prove to be overly intrusive for prospective cellular frame respondents; therefore we will use an 8-call design for the cellular sample (whether or not a request for callback is made). We will make one conversion attempt on all soft refusals on all landline sample cases. In adherence to the recommendations of the American Association for Public Opinion Research (AAPOR) Cell Phone Task Force, we will not attempt refusal conversion for soft refusals on cell cases:

“Logic and anecdotal evidence to date suggest that refusal conversion attempts to cell phone respondents should be of a limited nature so as to reduce the potential for further agitating [cell phone respondents]. This is in large part a result of likely reaching the same respondent who previously refused rather than reaching some other member of the sampling unit (household), as often is the case when trying to convert refusals in RDD landline surveys.” (AAPOR, 2010)

Non-response Follow-up Survey

We will evaluate non-response in the Employee Survey in four ways: a non-response follow-up survey (NRFU), a comparison of easy-to-reach versus harder-to-reach respondents, fitting response propensity models, and comparing survey estimates with external benchmarks. The NRFU will collect information on employees who fail to respond to the survey and provide insight into whether the nonrespondents differ from the respondents on the characteristics of interest (e.g., family and medical leave). Specifically, interviewers will call back a subsample (n=400) of households that declined the original survey. These interviewers will attempt to recruit an eligible employee to complete a shortened interview featuring a \$20 remuneration. Details on incentives are provided in Part A.

Incentives are a common feature in NRFU surveys because, by definition, the NRFU sample did not cooperate with the original survey, and so a major change in the recruitment protocol is required to elicit cooperation in the NRFU. Zimowski and colleagues (1997) noted in their report to the Federal Highway Administration (FHWA-PL-98-029) that large monetary incentives (e.g., \$20 to \$50) are a common element of NRFU designs for household surveys. For example, Peytchev et al. (2009) documented how a \$20 incentive was used in a successful

NRFU to the National Intimate Partner and Sexual Violence Survey for the Centers for Disease Control and Prevention.

In addition, all landline sample cases that can be matched to an address (though reverse lookup) will receive a letter encouraging them to cooperate with the interview. We expect to complete approximately 200 NRFU interviews. This will provide a sufficient case base for meaningful nonresponse analysis.

We will compare the employment and leave characteristics of Employee Survey respondents with the characteristics of NRFU respondents. This analysis will provide insights about the direction and magnitude of possible nonresponse bias. We will investigate whether any differences remain after controlling for major weighting cells (e.g., within race and education groupings). If weighting variables eliminates any differences, this suggests that the weighting adjustments discussed in Attachment D will reduce nonresponse bias in the final survey estimates. If, however, the differences persist after controlling for weighting variables, then this would be evidence that the weighting may be less effective in reducing non-response bias.

Comparison of Easy-to-Reach versus Harder-to-Reach Respondents

Second, we will conduct an analysis of the level of recruitment difficulty. This analysis will compare the leave-related characteristics of respondents who were easy to reach with respondents who were harder to reach. The level of difficulty in reaching a respondent will be defined in terms of the number of call attempts required to complete the interview and whether the case was a converted refusal. In some studies, this is described as an analysis of “early versus late” respondents, though we propose to also explicitly incorporate refusal behavior. If the employment and leave-related characteristics of the harder-to-reach cases are not significantly different from characteristics of the easy-to-reach cases, this would suggest that survey estimates may not be substantially undermined by non-response bias. The harder-to-reach cases serve as proxies for the non-respondents who never complete the interview. If the harder-to-reach respondents do not differ from the easy-to-reach ones, then presumably the sample members never reached would also not differ from those interviewed. Support for this “continuum of resistance” model is inconsistent (Lin and Schaeffer 1995; Montaquila et al. 2008), but it can still be a useful “straw man” framework for assessing the relationship between level of effort and non-response bias.

Estimating Response Propensity Models

Another technique that we will use to understand the risk of nonresponse bias is response propensity modeling (Little 1986; Groves and Couper 1998; Olson 2006). Response propensity is the theoretical probability that a sampled unit will be respond to the survey request. Many respondent characteristics can influence response propensity. Disentangling these effects requires multivariate modeling.

In order for a response propensity model to be informative, the researcher must know the values for respondent and non-respondents on one or more predictors of survey response. In RDD surveys, propensity models are often quite limited because little information is generally known for the non-respondents. For the Employee Survey, we propose to fit a response propensity model predicting the probability of completing the extended interview conditional on having completed the screener. This analysis will be based only on households for which we have a completed screener. By focusing on screened households, we can include richer independent variables in the model, including the selected respondent's age, gender, employment status, and leave status. In addition, the model will include an indicator for sampling frame, an indicator for whether or not the household ever refused the interview, and a log-transformed variable for the number of call attempts made to the household. Our preliminary plan is, thus, for the Employee Survey response propensity model to predict survey response using the following variables:

- Respondent age
- Respondent gender
- Respondent employment status
- Respondent FMLA leave status
- Household composition (count of leave-needers, leave takers, and/or other employed
- Census region
- Sampling frame (landline RDD or cellular RDD)
- Screener respondent is the selected extended interview respondent (yes/no)
- Household refused the interview once
- Number of call attempts made to the household

The estimated logistic regression model will be used to create summary “response propensity scores” (i.e., the predicted probability from the logistic regression model) that estimate how likely the selected respondent was to participate in the survey, regardless of the actual outcome. We will create five groups (response propensity classes) from the response propensity scores. In a well-specified model, respondents and nonrespondents will be equivalent on the characteristics of interest within each class, and likelihood of survey participation will vary across the classes.

The response propensity model will help us to identify the most powerful predictors of response when all available predictors are tested simultaneously. If employment-related or leave-related variables show a significant association with response to the extended interview (after controlling for other factors), this would be evidence of possible non-response bias. If, however, the employment and leave-related predictors do not have a significant effect, this suggests that the screener non-response adjustment described in Attachment D will be effective in reducing nonresponse bias. Similarly, comparisons of the respondent characteristics across the five response propensity classes will also provide insight on which types of screened respondents were most likely complete the extended interview and which types were less likely to do so.

Comparisons to External Benchmarks

One limitation of the aforementioned techniques is that they analyze only a subset of all non-respondents to the survey. The NRFU analysis relies on the NRFU participants as proxies for all non-respondents; the level of effort analysis relies on the “harder-to-reach” respondents as proxies for all non-respondents; the response propensity model captures only variation between the screened extended interview respondents and the screened extended interview non-respondents.

One approach for evaluating the total level of nonresponse bias in a survey is to compare the weighted survey estimates with external estimates based on a “gold standard” survey. The “gold standard” survey should feature a more rigorous protocol (e.g., area-probability sampling with in-person interviewing) and a higher response rate than the target survey (the 2011 Employee Survey). Critically, the gold standard survey and the target survey must feature one or more questions administered in a highly similar manner. Estimates based on these questions

can be compared. By virtue of its more rigorous design, the estimates from the gold standard survey are assumed to contain less non-response bias than those from target survey.

Differences in the question wording or mode of administration, however, may confound the comparison. Differences in population coverage between the gold standard and target survey may also confound the comparison. In light of these considerations, results from external comparisons must be interpreted with caution.

We propose to compare weighted Employee Survey estimates with those from the CPS. Examples of possible analytic variables administered in both surveys are: marital status, employment status, average hours worked per week, and labor union membership.

Non-contact versus Non-cooperation

Where possible, we will treat non-contact and non-cooperation as two distinct outcomes. Non-contact and non-cooperation are generally considered to reflect two different dimensions on which sample households can be placed. (Stinchcombe et al. 1981; Goyder 1987; Groves and Couper 1998; Lynn et al 2002). As noted by Stoop (2005), decomposing non-response into these two different dimensions can be analytically useful in several ways:

- When trying to enhance response rates different measures apply to improving contactability and improving cooperation;
- When comparing surveys over time or across countries different nonresponse rates and a different composition of the non-respondents (non-contacts and refusals) may be confounded with substantive differences;
- When estimating response bias or adjusting for nonresponse, knowledge about the underlying nonresponse mechanism (noncontact, refusal) should be available as contacting and obtaining cooperation are entirely different processes;
- When estimating response bias or adjusting for nonresponse, information on the difficulty of obtaining contact or cooperation is often used assuming that “difficult respondents are more like final refusers than easy respondents.

For the NRFU we will sample both non-contact nonresponding households and non-cooperative nonresponding households. This way, we can evaluate if employment and leave characteristics differ between these two groups and if they differ from the responding sample. In the easy-to-reach versus hard-to-reach analysis, we will define the easy/hard dimension in three ways: (1) in terms of ease of contactability as defined by the number of calls required to complete the interview; (2) in terms of amenability as defined by whether or not the case was a converted refusal; and (3) a in terms of both contactability and amenability as defined by a hybrid metric combining number of call attempts and converted refusal status. This analysis will provide some evidence as to which, if either, of these two mechanisms may be leading to nonresponse bias in survey estimates. For the response propensity modeling, we plan to condition on contacted households and model the probability of cooperating with the interview. This approach is based on the fact that the Employee Survey features an RDD sample design, which means that there is little information available on non-contacted households. Given this lack of data, models predicting the probability of contact would not be very informative. Finally, the benchmark comparison analysis is designed to compare survey estimates with external benchmark estimates. The outcomes of interest are NET differences between these two sets of estimates. In this analysis nonresponse must be treated in the aggregate. Decomposing non-contact and non-cooperation is not possible when evaluating estimates based on the responding sample.

Summary of Non-response Analyses

While these analyses rely on imperfect assumptions, all are standard techniques for assessing potential non-response error. No single nonresponse analysis for this study can be definitive because the true scores the non-respondents are not known. That said, by using several different methodologies (non-response follow-up analysis, easy-to-reach versus hard-to-reach comparisons, response propensity models, and comparisons of estimates to external benchmarks), we draw some meaningful conclusions about the level of risk to survey estimates from non-response bias. This information may also be helpful in modifying nonresponse weighting adjustments to reduce bias to the extent possible.

Employer Survey:

In order to achieve our estimated response rate, we anticipate the need for multiple modes of survey administration and will implement telephone and web-based modes. We have

developed a strategy that will maximize internet interviews, which in turn, will limit field data collection costs and minimize respondent burden. Survey administration will proceed as follows:

1. Prepare sample file and determine respondent location;
2. Establish telephone and internet identification (e.g., locating human resource personnel from a company directory that we are unable to contact via phone) of key respondent;
3. Mail advanced letter and information packet to identify key informant; includes personalized link to web-based survey; and
4. Send follow-up e-mail reminders and place telephone calls to non-responders

As mentioned in response to Question 1 of this submission, the employer survey includes two components, the identification of the key informant and the survey administration to that identified individual. Once identified, we will mail the key informant an advance package of materials providing background about the project. (Provided here as Attachment A). The package will include a letter from the Department of Labor explaining the importance of the survey and inviting the key informant to complete the survey either online (on a secure Web site) or by calling a toll free number to complete the survey over the phone with an interviewer. The letter will be printed on Department of Labor letterhead so that the recipient can clearly distinguish the survey materials from junk mail. We will send the package via priority mail. All non-responders to the pre-notification mailing will be contacted by interviewers to attempt to complete the survey over the phone. We anticipate that providing a web-based survey option will enhance the overall response rate in that it provides the respondent with additional flexibility in the time, location and pace of completing the survey. This is particularly important given the potential need for the respondent to consult with administrative records. The respondent can leave and re-enter the survey as frequently as they wish and at any time.

Non-response Analyses

Given that the anticipated response rate for the Employer Survey is under 70%, we will conduct an extensive nonresponse analysis. The main approaches that we will implement are comparisons of easy-to-reach versus hard-to-reach establishments, fitting response propensity models, and comparisons of survey estimates to external benchmarks. Discussion of each of

these approaches was presented above. The same experimental design theories and analytic steps apply in the context of the Employer Survey. The analysis for the Employer Survey will, of course, be customized for that survey.

For the Employer Survey response propensity models, we will estimate two separate outcomes: contact and cooperation conditional on contact. We plan to present two separate logistic regression models in this analysis because most readers find those results easiest to interpret. We will also estimate a multinomial logistic regression model for all three outcomes (non-contact, contact but non-cooperation, and cooperation). We will report whether the results of the multinomial model differ in any meaningful way from the two logistic regressions.

We plan to model contact in the Employer Survey using the following predictor variables:

- NAICS group
- Establishment size
- Census region

The contact model will be based on all establishments in the released replicates. We plan to model survey cooperation using the following predictor variables:

- NAICS group
- Establishment size
- Census region
- Establishment maintains records of FMLA leave (yes/no)
- FMLA requests processed internally versus outsourced

The cooperation model will be based only on establishments for which we have a completed screener. By focusing on screened establishments, we can include richer independent variables in the model. In a well-specified model, responding and non-responding establishments will be equivalent on the characteristics of interest within each response class, and likelihood of survey participation will vary across the classes. The cooperation propensity model will help us to identify the most powerful predictors of Employer Survey cooperation when all available predictors are tested simultaneously.

To explore the relationship between establishment characteristics and level of effort, we will compare the mean number of calls for establishments of different size, NAICS code, FMLA coverage status, workforce gender ratio, and workforce unionization. If establishment types are not shown to differ by the number of calls required to complete the interview, this suggests that non-response bias may be minimal. If, however, large differences are observed and cannot be addressed through weighting, then the risk of non-response bias is likely to be higher. Potential “gold standard” sources for the comparisons with external information are the Quarterly Census of Employment and Wages and Current Employment Statistics.

Non-contact versus Non-cooperation

As in the Employee Survey non-response analysis, we will treat non-contact and non-cooperation as two distinct outcomes where possible. We expect non-contact analysis in particular to be somewhat richer and more informative in the Employer Survey because of the availability of frame data that are related to constructs of interest in the survey (e.g., NAICS code and establishment size). In other words, this survey has useful data on both contacted and non-contacted sample units, which is not the case with the Employee Survey.

The easy-to-reach versus hard-to-reach analysis for the Employer Survey is somewhat different because we do not anticipate having “converted refusals” as we do in the Employee Survey. In a landline survey there is a reasonable chance that the interviewer may reach a different, more amenable household member on a subsequent call. In an establishment survey, by contrast, this chance is very small, and so we do not plan to attempt refusal conversion on establishments that have expressly declined the survey. As a result, the Employer Survey easy-to-reach versus hard-to-reach analysis will focus on whether or not leave-related practices are related to the number of calls required to complete the interview. For the response propensity modeling, we will model contact and cooperation separately as discussed above. The benchmark comparison analysis is designed to compare survey estimates with external benchmark estimates. The outcomes of interest are NET differences between these two sets of estimates. In this analysis nonresponse must be treated in the aggregate. Decomposing non-contact and non-cooperation is not possible when evaluating estimates based on the responding sample.

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

Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.

Employee Survey:

A total of 9 cognitive interviews have been conducted. This testing considered individuals' comprehension of survey items. Based on the results of the cognitive testing, some items were changed or eliminated. The survey included in this package reflects the findings from those interviews.

The current version of the employee survey is included as Attachment C.

Employer Survey:

A total of 9 cognitive interviews will be conducted; to date, 3 interviews have been conducted. This testing determines employers' comprehension of survey items. Based on the results of the cognitive testing, some items may be changed or eliminated. The survey included in this submission contains edits that resulted from this preliminary testing. Further testing on the internet version of the survey is scheduled once the survey has been fully programmed. No significant, substantive changes are expected. Changes to the formatting of survey for internet administration are expected. A final version of the survey instrument will be sent to OMB by February 2011.

The current version of the employer survey is included as Attachment C.

5. Disclosure Limitation Methods

Public use files (PUF) for both the Employee Survey and Employer Survey will be made available after completion of the data collection. We will implement a disclosure limitation protocol for each survey so that the PUF fully protects respondent confidentiality.

The risk of disclosure in either the Employee Survey or the Employer survey is extremely low for the following reasons:

- (1) **No sampling frame information, contact information, or other person or**

establishment identifying information will be included in the PUFs. It will not be possible to link the survey records to administrative data. Each record will have a unique case ID, but that value will be randomly assigned and will carry no information about the record.

- (2) **No geographic variables will be included in either PUF.** The surveys are designed for national-level analysis rather than sub-national analysis. Eliminating geographic detail is one of the most effective methods for limiting disclosure risk.
- (3) **The surveys are cross-sectional rather than longitudinal, and they do not feature clustering in the sample designs.**
- (4) **The sampling fractions in both surveys are extremely small.** In the Employee Survey cell RDD and landline RDD frames, the expected sampling fractions are 0.00010 and 0.00046, respectively. In the Employer Survey, the expected sampling fraction is 0.00020. Surveys with very small sampling fractions entail a lower risk of disclosure than surveys with larger sampling fractions.
- (5) **Sample design variables will not be released.** Replicate weights will be provided so that data users can account for the complex nature of the sample designs. When replicate weights are provided, it is not necessary to provide sample design variables, such as PSU or stratum.

According to guidelines published by the Federal Committee on Statistical Methodology Report on Statistical Disclosure Methodology (2005) and the National Center for Health Statistics Staff Manual on Confidentiality (2004) these properties of the Employee and Employer surveys reduce the risk of disclosure limitation.

Below we describe the specific additional steps that will be taken to ensure that the data released in the PUFs fully protect respondent confidentiality. We will employ variable suppression, rounding, top-coding, bottom-coding, and other data coarsening as needed so that no identifying values are released in the PUFs. We prefer these techniques over data swapping because for variables like respondent age, recoding has been shown to improve protection more than random data swapping (Reiter 2005).

Employee Survey

Basic demographic variables are often the most susceptible to matching. In order to make sure

that no identifying values are release, we will make the following manipulations to the Employee Survey dataset. These manipulations are in addition to the disclosure limitation procedures mentioned above.

- D1 We will collapse the cells for “GED” and “High school graduate.” Having a GED is a fairly rare characteristic.
- D4 The variables D4h and D4j will be suppressed (not included in the PUF). These variables detail relatively small income categories. The lowest income classification will, thus, be under \$20,000 and the highest will be \$100,000 or above. Specifically, we will bottom-code income. The top code (\$100,000 or above) is not a rare characteristic and will not be manipulated.
- D6 The “Native Hawaiian or Pacific Islander” cell will be collapsed with the cell for “Some other race.” The incidence of that group is very low (0.3% of the US population), meaning that it could potentially be an identifying variable if used in conjunction with other variables.
- D7 The number of children under 18 in the respondent’s care will be top coded at 4 or more children. Employees with 5 or more children in their care are relatively rare and potentially identifiable.
- D8 The number of people over age 65 in the respondent’s care will be top coded at 3 or more. Employees with 3 or more people over age 65 in their care are relatively rare and potentially identifiable.
- D12 The continuous variable for age of spouse/partner will be suppressed. We will instead provide a categorical variable with age values: 18 to 34, 35 to 49, 50 to 64, and 65 or over.
- D13 ZIP code (and all other geographic or personally identifying information) will not be released.
- END3 Name and address (and all other geographic or personally identifying information) will not be released.

Screener data (S1 through T6) collected for household members other than the selected respondent will not be included in the PUF. The main sections of the questionnaire contain a

series of questions asking about the start dates (month and year), stop dates (month and year), and reasons why respondents took leave from work. Given that this type of information may be known by numerous people in the respondent's life and some combinations of values may be quite rare, these variables pose a disclosure risk. We propose to suppress all variables containing the month/year of a leave beginning or ending. Instead, we will report the duration of the leave in a specially-constructed variable.

- A4 The number of total reasons the respondent took leave will be top coded so that larger values are not personal identifiable information.
- A8 The age of the care patient will be top coded so that larger values are not personal identifiable information.
- A13 Month/Year of leave start will be suppressed.
- A15 The number of separate blocks of time taken off work for the leave will be top coded so that larger values are not personal identifiable information.
- A16 Month/Year of leave start will be suppressed.
- A17 Month/Year of leave end will be suppressed.
- A34 Amount paid for medical certification will be coarsened into broad categories.
- A40 Amount paid for medical re-certification will be coarsened into broad categories.
- B4 Number of times leave was needed but not taken will be top coded so that larger values are not personal identifiable information.
- B5 Number of times leave was needed but not taken will be top coded so that larger values are not personal identifiable information.
- B9 The age of the care patient will be top coded so that larger values are not personal identifiable information.
- B14 Number of times leave was needed will be top coded so that larger values are not personal identifiable information.

In addition to these pre-identified data edits, we will review the final data for rare responses.

As necessary, we will recode so that no single response category or combination of closely related response categories has an unweighted frequency below five.

Employer Survey

No screener data (S1 through S21) will be included in the PUF. The following manipulations will be made in addition to the disclosure limitation procedures mentioned above.

- Q1 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q2 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q3 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q4 This variable will be suppressed. Its function is procedural not substantive.
- Q5 This variable will be suppressed. Values could potentially identify an establishment.
- Q6 Union participation will be reported only as a percentage.
- Q6a Union participation will be reported only as a percentage.
- Q7 Female work force will be reported only as a percentage.
- Q8 Employees working for at least one year will be reported only as a percentage.
- Q9 Employees working who worked at least 1,250 hours will be reported only as a percentage.
- Q16x2 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q16x4 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.

- Q16x5 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q19 This will be reported only as a percentage.
- Q20 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q21 This will be reported only as a percentage.
- Q24 This will be reported only as a percentage.
- Q26 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q27 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q29 This will be reported only as a percentage.
- Q31 This will be reported only as a percentage.
- Q33 This will be reported only as a percentage.
- Q46 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q58 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q59 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q60 Values will be coarsened and reported only as a categorical variable with no establishment identifying values.
- Q71 This variable will be suppressed.
- Q72 This variable will be suppressed.

Again, in addition to these pre-identified data edits, we will review the final data for rare

responses. As necessary, we will recode so that no single response category or combination of closely related response categories has an unweighted frequency below five.

6. Provide the name and telephone number of individuals consulted on statistical aspects of the design, and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Abt SRBI has been contracted to conduct both the Employee and Employer survey. The individuals at Abt SRBI assigned to this project include:

Jacob Klerman, Principal Associate, (617) 520-2613

Kelly Daly, PhD, Senior Analyst, (312) 529-9703

Alyssa Pozniak, PhD, Senior Analyst (617) 520-2455

Courtney Kennedy, PhD, Senior Methodologist, (734) 972-4283

In addition, the Project Officer for DOL is Jonathan Simonetta, (202) 693-5085

References

- American Association for Public Opinion Research (AAPOR). 2010. "New Considerations for Survey Researchers When Planning and Conducting RDD and Other Telephone Surveys in the U.S. With Respondents Reached via Cell Phone Numbers" (http://aapor.org/AM/Template.cfm?Section=Cell_Phone_Task_Force&Template=/CM/ContentDisplay.cfm&ContentID=2818)
- Curtin, R., S. Presser, and E. Singer. 2005. "Changes in Telephone Survey Nonresponse Over the Past Quarter Century." *Public Opinion Quarterly* 69:87-98.
- Federal Committee on Statistical Methodology Report on Statistical Disclosure Methodology. 2005. "Statistical Policy Working Paper 22. Report on Statistical Disclosure Limitation Methodology." Office of Management and Budget. Available at <http://www.fcsm.gov/working-papers/spwp22.html>.
- Goyder, J. 1987. *The Silent Minority*, Boulder, CO: Westview Press.
- Groves, R. and M. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Lin, I. and N. Shaeffer. 1995. "Using Survey Participants to Estimate the Impact of Nonparticipation." *Public Opinion Quarterly* 59: 236-258.
- Little, R. 1986. "Survey Nonresponse Adjustments for Estimates of Means." *International Statistical Review* 54:139-57.
- Lynn, P. and P. Clarke. 2002. "Separating Refusal Bias and Non-Contact Bias: Evidence from UK National Surveys." *Journal of the Royal Statistical Society Series D (The Statistician)*, 51(3): 319-333.
- Montaquila, J., J. Brick, M. Hagedorn, C. Kennedy, and S. Keeter. 2008. "Aspects of Nonresponse Bias in RDD Telephone Surveys." In *Telephone Survey Methodology*, edited by J. Lepkowski, C. Tucker, J. M. Brick, E. de Leeuw, L. Japac, P. Lavrakas, M. Link, R. Sangster. New York: Wiley.
- National Center for Health Statistics. 2004. "Staff Manual on Confidentiality." Available at <http://www.cdc.gov/nchs/data/misc/staffmanual2004.pdf>.
- Olson, K. 2006. "Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias." *Public Opinion Quarterly* 70: 737-758.
- Peytchev, A., R. Baxter, and L. Carley-Baxter. 2009. "Not All Survey Effort is Equal: Reduction of Nonresponse Bias and Nonresponse Error." *Public Opinion Quarterly* 73: 785-806.
- Reiter, J. 2005. "Estimating Risks of Identification Disclosure in Microdata." *Journal of the American Statistical Association* 100, 1103 - 1113.

Stinchcombe, A. L., C. Jones, and P. Sheatsley. 1981. "Nonresponse Bias for Attitude Questions." *Public Opinion Quarterly* 45: 359–379.

Stoop, I.A.L. 2005. *The Hunt for the Last Respondent*. The Hague: Social and Cultural Planning Office of the Netherlands.

Zimowski, M., T. Tourangeau, and R. Ghadialy. 1997. "Nonresponse in Household Travel Surveys (FHWA-PL-98-029)" Prepared for the Federal Highway Administration.

DRAFT