

SUPPLEMENTAL SUPPORTING STATEMENT B E-VERIFY DATA COLLECTIONS

B. Collection of Information Employing Statistical Methods

Introduction

We are proposing to use statistical methods for our Web survey of E-Verify employers, as described in Sections B.1-B.3. We do not, however, intend to use statistical methods in conducting our case studies of Designated Agents (DAs) and Users of Designated Agents (UDAs). The remainder of this introduction explains why we are not proposing statistical methods in the DA/UDA study.

The primary goal of the exploratory case studies of DAs and UDAs is to obtain a more in-depth understanding of how well the DA version of E-Verify is working in terms of employer satisfaction and compliance with E-Verify procedures. We plan to use semi-structured telephone interviews to provide us with a better understanding of this rapidly growing component of the E-Verify Program. This qualitative study will also provide input into structured questions to ask on future Web surveys of DAs and UDAs.

Twenty DAs will be selected for the case study and 3 clients (UDAs) will be selected for each selected DA. Selection of DAs will be purposive and will be designed to obtain information from DAs that differs in terms of number of clients, mix of the industries represented by their clients, geographic location, the number of transmissions to E-Verify, and the number of tentative nonconfirmations resulting from the transmissions to E-Verify. Similarly, UDAs will be purposively selected to ensure that different types (based on industry, number of transmissions, and number of tentative nonconfirmations) of UDAs are included in the study. To apply the proposed methodology to a statistically representative sample of DAs and UDAs would greatly increase the cost of this component of the evaluation. We do not believe that the additional accuracy obtained would be warranted by the additional cost.

1. Respondent Universe

The target population of the Survey of E-Verify Employers includes all firms that signed an MOU before April 1, 2010 in which they agreed that all or part of the firm would participate in E-Verify with the following exceptions:

- o **Companies with no recent involvement in E-Verify.** Recent involvement is defined as having taken an action (signing an MOU, submitting cases to E-Verify, or formally terminating participation in the Program) within the six months between October 1, 2009 and March 31, 2010. This exclusion helps to ensure that the company representative can accurately recall aspects of their participation, and also avoids the possibility of companies responding to a program characteristic that has recently changed.
- o **Employers that participated in the recent study of Arizona employers** are excluded to avoid undue burden on this small group of employers. The only Arizona users that were excluded were the 126 firms that participated in the Arizona study. The sampling frame after the exclusions contained 4,585 active Arizona firms, 495 of these firms are in the final sample.

- o **DAs and UDAs** will be excluded, because this population is being studied in the exploratory case study described in the Introduction to this Supplementary Supporting Statement. A total of 2,762 firms identified themselves as DAs or UDAs, which constitutes 9.4 percent of the original database of 29,406 firms.¹
- o **Employers in Puerto Rico, Guam and other U.S. territories** are excluded primarily for pragmatic reasons (e.g., different time zones require telephone interviewers to follow-up at impractical times; some language barriers). Since this is a very small segment of the employer population, their exclusion should not result in a significant coverage problem.

Sampling Frame

The sampling frame will be developed from three databases provided by the contractor responsible for E-Verify operations:

- o **Employer Database** containing information provided by employers at the time they registered for E-Verify and any subsequent modifications the employers may have made to the information. The records contain the following fields needed for sample selection: employer name; North American Industry Classification System (NAICS) code; the company's "parent company"; the number of sites covered by the MOU, the date the employer signed the MOU and, where relevant, the date the employer terminated participation in E-Verify;
- o **Point of Contact database** which contains contact information associated with employers that have enrolled in E-Verify linked to the Employer Database through unique employer IDs;
- o **Transaction database:** This file contains information on case submissions to E-Verify, including date of case initiation, ID of the employer submitting the case, and the dates and types of subsequent case actions.

To produce the data file to be used for sample selection:

- o Contact information from the Point of Contact file will be appended to the appropriate employer records;
- o Unique records will be compiled at the firm level based on the information in the Employer Databases, since the sampling units will be single location companies (a business establishment with no branches or subsidiaries reporting to it) and the headquarters of the companies that have multiple branches. Most of these will be identified as companies without "parent companies" in the employer database. In cases in which it appears that employers may be branch companies of larger companies without a specified parent company (e.g., when there are large numbers of employers with the same name that have not specified a

¹ Both DAs and UDAs were included in the last user survey. They were excluded from the frame for the current year because they are being studied separately and in greater depth through telephone interviews. Given the interview component, it seemed a duplication of effort and unnecessary burden to use both interviews and surveys. DAs and UDAs have a very different role with E-Verify than other users: for example, they conduct the verification activities to meet another company's employment needs that are different from their own, the verifications may be conducted within the context of other services (such as payroll services) that they also provide to their clients, and their clients may request special approaches to the verification process that the DAs/UDAs would not use on their own workers. Therefore, the in-depth study of DAs and their users will enable us to identify the similarities and differences in how DAs operate with their clients, which can be used to develop questions for a national study. To make valid comparisons of users other than DAs/UDAs in the 2008 and 2010 survey, we will exclude DAs/UDAs from the 2008 analysis database. We anticipate using information for DAs/UDAs in the 2008 study in preparing the DA/UDA report.

parent company), information will be obtained through Web searches and/or telephone inquiries to determine the appropriate firm level information²;

- The Transaction Database will be purged of duplicate records and records that the employer indicates were “submitted in error” (typically records with typographical errors detected by employers after submission);
- An outcome variable will be calculated from the Transaction Database information to indicate the final outcome of the case such as immediately found work authorized, found work authorized after a tentative nonconfirmation (TNC), etc.
- Information on the number of transactions and the number of cases receiving a TNC in this database will be aggregated by employer and appended to the records in the employer file.

Ineligible cases (as defined above) will be excluded from the sampling process.

Sample Design and Sample Size

We plan to select a total sample of approximately 3,700 E-Verify employers for the survey—a number that should provide approximately 2,400 completed surveys based on our experiences with the FY2008 surveys. This estimate assumes 35 percent of sampled cases are excluded either because the intended respondent is ineligible for inclusion (approximately 20% of sampled respondents) or because the respondent does not respond (approximately 20% of eligible respondents).

We propose to stratify the sample based on employer’s E-Verify status as follows, i.e.:

- **Active Employers with TNCs:** employers that had transmitted one or more cases receiving a TNC between January 1 and March 31, 2010 *and* had not formally terminated participation in the program before April 1, 2010. These employers are of great interest to policy makers concerned with discrimination and compliance, since many of the E-Verify procedures, including many procedures designed to prevent discrimination, are only relevant for employers that have TNC cases. Since experiencing TNCs is relatively rare for companies, creating a separate stratum will help to ensure that a sufficient number of such companies are in the sample to provide reliable estimates.
- **Active Employers without TNCs:** employers that had transmitted one or more cases between January 1 and March 31, 2010, had no cases receiving a TNC, and had not formally terminated participation in the program before April 1, 2010. These employers may well have different experiences from those active employers with TNCs; these differences may impact both their satisfaction and compliance with E-Verify procedures, so it is not possible to assume that the employers with TNCs are representative of those without TNCs, making the creation of a separate stratum helpful in ensuring that a sufficient number of such companies are in the sample to provide reliable estimates.
- **Inactive Employers:** employers that had signed an MOU to participate in the E-Verify program between July 1 and September 30, 2009 but had not transmitted any cases between before October 1, 2009 and March 31, 2010 or had formally terminated participation in the program between October 1, 2009 and March 31, 2010. The insights from these employers

² The Employer Database contains a mix of establishments and firms, because of the E-Verify enrollment procedures.

are vital to understanding why some employers do not find E-Verify beneficial. Because they are asked a number of questions not asked of active employers, it is important that we have a sufficiently large sample of inactive employers to permit meaningful analyses of the unique questions asked of them.

We further propose dividing each of the above E-Verify status groups into three groups based on the following employer industry classification:

- o **Staffing agencies**, i.e., employment agencies and temporary help services. These employers have some experiences and needs that are known to differ from many other employers, because of their need to satisfy their clients.
- o **Industries** (like hospitality services and food processing) known to have relatively large percentages of undocumented workers.³ The experiences of these employers differ significantly from other employers, because participation in E-Verify is relatively likely to affect their ability to attract low salaried workers.
- o **Other industries**. The remaining industries represent the largest industrial subgroup of employers and needs to be a large enough to ensure that overall statistics are reliable.

A Cross-classifying E-Verify company by E-Verify status and industry provides a total of nine strata. These strata, their estimated population sizes (including ineligible cases that can't be identified prior to sample selection), and proposed sample sizes are included in Table B-1. This table has been updated from the original OMB submission to reflect information that became available as of March 31, 2010.

The sample size shown in Table B-1 is based on the assumption that 20 percent of those who were sampled will be found ineligible. This percentage is a rough estimate, and may be a little high, but it is better to have too many respondents than too few. In the 2008 user survey, 351 of 3,203 sampled employers (11 percent) were ineligible, either because they were no longer in business or there were duplicate listings. In the 2009 non-user survey, 1,072 of 3,819 companies (28 percent) were ineligible, most commonly because they had closed, but also because they were E-Verify users or they had no employees. This survey is much more like the 2008 survey than the 2009 survey in terms of the sampling approach and available data, and we therefore expect the percentage of ineligibles to be more similar to 2008 than to 2009. However, based on the high number of closed companies that we found in 2009 (presumably due to the recession), we expect that the percentage of ineligible companies will be higher in 2010 than in 2008.

³ Information on the percent of undocumented workers is based on information in Jeffrey S. Passel, Senior Demographer, Pew Hispanic Center, and D'Vera Cohn, Senior Writer, Pew Research Center, *A Portrait of Unauthorized Immigrants in the United States*, 4.14.2009 (<http://pewhispanic.org/reports/report.php?ReportID=107>).

Table B-1. Population and sample sizes of companies for the Survey of E-Verify Employers by sampling strata

Employer E-verify status	Industry sector	Population in Jan-Mar10	Sample size	Expected number of completes	Sampling rate
Active (TNCs)	Staffing Agencies	158	100	64	
	1-199 transactions	67	28	18	42%
	200+ transactions	91	72	46	79%
	High percent undocumented	1,316	850	544	
	1-19 transactions	621	301	193	48%
	20+ transactions	695	549	351	79%
	Other	1,936	870	557	
	1-49 transactions	1,087	267	171	25%
	50-149 transactions	396	245	157	62%
	150+ transactions	453	358	229	79%
Active (No TNCs)	Staffing Agencies	193	150	96	78%
	High percent undocumented	3,991	424	271	
	1-4 transactions	1,885	96	61	5%
	5-9 transactions	733	65	42	9%
	10-19 transactions	626	79	51	13%
	20-49 transactions	483	90	58	19%
	50+ transactions	264	94	60	36%
	Other	12,013	425	272	
	1-19 transactions	8,871	182	116	2%
	20-49 transactions	1,733	90	58	5%
50-99 transactions	766	60	38	8%	
100+ transactions	643	93	60	14%	
Inactive	Staffing Agencies	34	27	17	79%
	High percent undocumented	2,013	435	278	
	1 to 4 employees	589	88	56	15%
	5 to 19 employees	747	127	81	17%
	20 to 99 employees	583	146	93	25%
	100+ employees	94	74	47	79%
	Other	4,990	446	285	
	1 to 4 employees	1,753	88	56	5%
	5 to 19 employees	1,646	99	63	6%
	20 to 99 employees	1,120	112	72	10%
100 to 999 employees	416	104	67	25%	
1000+ employees	55	43	28	78%	
Total		26,644	3,727	2,385	

Notes: Data included in the table are based on information for E-Verify employers as of March 31, 2010. The table assumes that 65 percent of selected cases are eligible employers that respond to the survey (i.e., 20 percent of employers are ineligible and the response rate is 80 percent among eligible respondents). Classification of industries by percent undocumented is based on Jeffrey S. Passel, Senior Demographer, Pew Hispanic Center, and D'Vera Cohn, Senior Writer, Pew Research Center, *A Portrait of Unauthorized Immigrants in the United States*, 4.14.2009 (<http://pewhispanic.org/reports/report.php?ReportID=107>).

Within strata, we propose using probability proportional to size (PPS) sampling, where the square root of the number of transactions submitted between January and March 2010 is the measure of size (MOS) for active E-Verify employers without TNCs and the cube root of the number of employees is the measure of size for inactive employers. In our original OMB submission, we planned to select certainty samples of all active employers with TNCs, but the increase in the number of such active employers and the desire to avoid overlap with the customer satisfaction survey that is expected to be in the field at the same time as this study led to a change in that strategy. Instead, we plan to use PPS sampling for these employers with a MOS equal to the square root of the number of TNCs. PPS sampling to transactions/number of employees/TNCs would be very efficient for making estimates at the transaction/worker/TNC level, while equal probability would be very efficient for making inferences about the characteristics of companies. The proposed PPS sampling provide a good compromise between the two objectives. In a few cases, Table B-1 shows relatively high sampling rates (such as 79 percent) to better support analyses comparing small subgroups.

The sample design will generate a national probability sample of employers that have enrolled in E-Verify. The survey will utilize a stratified random sample design.

Table B-1 assumes that all active employers with TNCs will be selected with certainty; if the number of active employers with TNCs is significantly larger than indicated in Table B-1 when the sample is selected, sampling rates for these three strata may need to be reduced.⁴ For the remaining strata, a target sample size is set, so that the total sample will be 3,700, assuming the size of each stratum is set as the smaller of the target sample size (approximately 425 in the example) and the population size of the stratum.

Precision of the estimates

The domains of employer population of interest are defined by cross-classifying the three employer E-verify status classes with the three industry sectors. Table B-2 shows the expected precision levels for various percentage statistics for the resulting nine domains. The first two columns of Table B-2 show the nine domains. Column 3 shows the expected number of completed surveys in each domain. The remaining columns show the expected percent sampling errors for various levels of percent statistics. (The expected percent errors are calculated by multiplying the standard errors by 2.) For example, for a 50 percent statistic for active staffing agencies (with TNCs), the percent error will be around ± 10.4 percentage points, with a 95 percent confidence. As can be seen from Table B-2, the percent error is the largest for a 50 percent population proportion and decreases as proportion moves away from 50 percent/50 percent split. For a 20 percent/80 percent split, the percent error decreases to ± 8.4 percent for active staffing agencies (with TNCs). Thus, for active staffing agencies (with

⁴ The number of active employers with TNCs in January through March 2010 may be higher than the number observed in April through June 2009, since enrollment in E-Verify has been increasing. However, the percent of cases with TNCs has been declining, making it difficult to project the strata sizes.

TNCs), the sampling error will not exceed ± 10.4 percentage points for a percentage statistics of any magnitude.

Note that these precision estimates are provided for estimates of employer characteristics in the form of percentages. Since we oversampled the employers with a larger number of transactions (or number of employees), the precision of estimates correlated with the number of transactions (or number of employees) is expected to be higher than shown in Table B-2 in the response to question 15.

Table B-2. Expected number of completed surveys and percent error¹ for various magnitudes of population percentages by employer E-verify status and industry sector domains

Employer E-verify status	Industry sector	Expected number of completes	Percentages		
			50/50	30/70	20/80
Active (TNCs)	Staffing Agencies	64	± 10.4	± 9.6	± 8.4
	High percent undocumented	544	± 3.4	± 3.2	± 2.8
	Other	557	± 4.3	± 3.9	± 3.4
Active (No TNCs)	Staffing Agencies	96	± 7.2	± 6.6	± 5.8
	High percent undocumented	271	± 7.1	± 6.5	± 5.7
	Other	272	± 7.2	± 6.6	± 5.7
Inactive	Staffing Agencies	17	± 16.9	± 15.5	± 13.5
	High percent undocumented	278	± 6.1	± 5.5	± 4.8
	Other	285	± 6.7	± 6.2	± 5.4

¹ Percent errors are obtained by multiplying the expected standard errors by 2.

2. Procedures for the Collection of Information

The following data collection approach will be used for the Web survey to collect self-reported data on employers' experiences with E-Verify. It will ask employers about their verification procedures, labor force characteristics, and opinions on employment verification and possible improvements to E-Verify. The statistical methodology for stratification and sample selection of employers was described in Section B-1.

Development and testing of questionnaires

Louise Hanson took the lead role in designing the paper version of the questionnaire, working in consultation with Denise Glover, Carolyn Shettle, Brad Chaney, Joan Michie, Lisa Roney (as a consultant to Westat), and USCIS staff. Louise Hanson also took the lead role in coordinating with the programming staff to develop the web questionnaire. Ed Mann led the programming effort, assisted by Mei Dong, and with programming activities coordinated by John Brown. The instrument was tested by Westat's usability testing group, led by Jennifer Crafts. The instrument was also tested by roughly ten of Westat's E-Verify staff, including Louise Hanson, John Porton, Mike Walewski, and Roberta Pike.

Estimation Procedures

The sampling strategy used will result in unequal selection probabilities for the companies. We therefore will create statistical weights based on the selection probabilities in order to produce nationally representative statistics. In addition, we will examine the strata for differential rates of response, and will statistically adjust for nonresponse as needed to provide nationally representative statistics. The analyses will use the final weights adjusted for nonresponse.

Standard statistical software will not produce correct variance estimates when complex sampling schemes are used. We will add replicate weights to the analysis file and use WesVarPC to produce appropriate variance estimates.

3. Methods to Maximize Response Rates and Deal with Issues of Non-Response

To minimize nonresponse, the USCIS contractor will devote considerable resources to developing and implementing approaches likely to achieve good respondent cooperation with the Survey of E-Verify employers.

We expect high levels of cooperation with the evaluation among employers that have enrolled in E-Verify, based on the completion rates for the 2008 survey of users which obtained an overall un-weighted survey response rate of 81 percent and a weighted response rate of 84 percent. These response rates are calculated by dividing the number who completed the survey by the number of eligible firms in the sample; ineligible firms are excluded from the calculations. These employers as well as the Designated Agents and Users of Designated Agents have signed an MOU with the DHS and have agreed to respond to DHS and SSA designees' inquiries about E-Verify. Specifically, the MOU states the employer's responsibilities as follows:

“The Employer agrees to cooperate with DHS and SSA in their compliance monitoring and evaluation of E-Verify, including by permitting DHS and SSA, upon reasonable notice, to review Forms I-9 and other employment records and to interview it and its employees regarding the Employer's use of E-Verify, and to respond in a timely and accurate manner to DHS requests for information relating to their participation in E-Verify.”

The techniques that will be used to ensure high response rates are:

(1) Pre-testing. Much of the 2010 questionnaire is based directly on the 2008 Survey of E-Verify Employers, and thus has been tested through both pretests and full data collection.

Knowledge obtained from conducting the previous surveys was used to modify the 2008 survey, including a review of data issues from the previous survey, a review of responses to open-ended items, and a review of the frequencies of the 2008 survey (to identify items lacking sufficient variation or a sufficient number of responses to be useful). In addition, questions were added based on an update of new program features provided by USCIS. These items replaced some of the questions examining the implementation of features that were new at the time of the 2008 survey.

In January and February of 2010 we conducted focus groups and telephone interviews to examine the 2010 questionnaire, focusing particularly on items that are new or that have been modified from 2008. The focus group pretest was very informative. The questionnaire worked well, but the pretest identified a few places where the wording was unclear and should be revised. In an attachment we provide a revised questionnaire that addresses all of the problems that were identified; all revisions from the focus group sessions are highlighted in yellow.

There were no ‘new’ questions added to the questionnaire in response to the pretest. However, one question was split in two, and the revised numbering may create the appearance that a question was added. The original question D15 was in table format, and one of the items in that table asked about the use of the E-Verify Photo Tool. It became clear during the focus groups that this item needed to be pulled out of the table and made into an individual question so that a skip pattern could be added for those firms that had never used Photo Tool. Therefore, the original D15 has been split into questions D15 and D16.

Following are the results for the three questions in which OMB expressed special interest.

- a. **A5 – Which description best fits your company? Is the terminology used in the response options widely understood by respondents? Did more than an expected number of respondents choose “Don’t know”?**

We found that all of the focus group respondents understood the terms used in A5 without any additional wording needed. No respondents chose “Don’t know”.

- b. **C2 – estimated total direct expenditures associated with setting up e-Verify – How do respondents get these numbers? Are they direct from records? Did some struggle to locate historical information?**

Question C2 has been used in previous waves of the E-Verify User Survey and so, for purposes of longitudinal analysis, we were reluctant to make any revisions to the wording of the question. We found that some of the focus group respondents answered with a ‘best guess estimate’ and others indicated that they looked up actual records in order to respond. None of the respondents expressed any problems in answering the question.

- c. **C5 – estimated total annual direct expenditures associated with maintaining e-Verify – please see item b above.**

Question C5 has also been used in previous waves of the E-Verify User Survey and so, for purposes of longitudinal analysis, we again were reluctant to make any revisions to the wording of the question. We found that the focus group respondents indicated no problem in answering the question.

The focus groups also led to a change in current question D22. This was a new question in the 2008 survey, and in 2008 it was worded:

“Since the start of the Photo Tool, have you noticed any decreases in the use of immigration documents provided by employees during the verification process?”

For 2010 we initially revised the time reference to “Since the start of the Photo Tool in September 2007 ...”. During the focus group tests we found that respondents who had never used Photo Tool were having difficulty with that phrase. Therefore we ended up revising it to “In the past few years ...”. The focus group respondents indicated they had no difficulty with the rest of the wording of the question. They also indicated that they were answering from their own knowledge about the company’s point of view.

Current question D23 has been used in several previous waves of the E-Verify User Survey and so, for purposes of longitudinal analysis, we were reluctant to make any revisions to the wording of the question. We also found that none of the focus group respondents expressed any problems in answering the question. In response to OMB’s request, we have added the confidentiality reminder that appears in G5 to this question.

(2) Motivational material. Information about the E-Verify data collection will be placed on a Web site to be accessed by employers that wish to obtain additional information about the evaluation. Continued care will be taken in the final production of survey materials to:

- o Create a professional image for the study;
- o Emphasize the importance of participation towards shaping future directions in Federal immigration policy;
- o Emphasize the steps that will be taken to ensure respondent confidentiality; and
- o Use language appropriate for the target population.

(3) Aggressive follow-up. One of the major factors that increases study response rates is the use of aggressive follow-up procedures to gain cooperation with the study. The Web Survey of Employers, therefore, includes multiple contacts with selected respondents. More specifically, the data collection procedures consist of the following steps:

- o A personalized pre-notice letter will be sent to all primary contact people identified in USCIS materials. This letter will be from USCIS **(Attachment D)** and will state that this is part of the evaluation effort they authorized when they signed the MOU. The letter will stress both the importance of participation to future employment verification efforts and the fact that USCIS will only use the information for research purposes.

- o A personalized email will be sent to all contact people selected by Westat, re-iterating the importance of the study and providing information on how to log on to the Web site (Attachment D).
- o At the same time that the emails from Westat are sent out, a banner message will be placed on the Web site that employers use when verifying employees through E-Verify. It will indicate that the evaluation has started, provide a link to a Web page with additional information about the study, and ask employees that received a survey request to complete the survey.
- o If the email results in a response indicating the email address is no longer valid, an email will be sent to the alternate contact person, if any.
- o If no email is provided for the primary contact person or if there is no alternate contact person for a non-valid email address, phone interviewers will research the company's contact information online to identify a staff person in Human Resources or a similar office who might be knowledgeable about the appropriate contact person.
- o If calling the Human Resources Department does not result in identifying the correct contact person, a phone interviewer will call the main number of the company to determine who is the correct contact person and, if possible, obtain the name and contact information for an alternate person who will be responsible for the study, if the primary contact person is not available. The phone interviewers will also collect information on important changes in status among the contacted companies to determine if they are out-of-scope for the study.
- o A reminder email will be sent to contact persons approximately one week after the initial contact and a second banner message will be placed on the verification Web site at that time.
- o Approximately two weeks after the reminder email, a second reminder e-mail will be sent to non-respondents.
- o Approximately two weeks following the second reminder e-mail, phone interviewers will contact non-respondents. Reasons for nonresponse will be requested and participation will be encouraged. If necessary, reluctant respondents will be reminded of the MOU in which the employer had agreed to participate in the evaluation. Information on how to access the Web survey site will be provided, if necessary.
- o A second phone reminder will be made approximately two weeks after the first phone reminder. At that time, the interviewer will offer to send a hard copy survey if the respondent prefers to answer in this fashion. Again, nonrespondents will be reminded to complete the web survey and will receive specific log-in instructions, if necessary.

If necessary, a final contact will be made approximately two weeks after the second phone reminder. The non-respondents will be sent, via Federal Express, another cover letter, log-in information to access the survey Web site, and a hard copy survey. In addition to the above contacts, a thank you email will be sent to respondents that complete the survey.

While the survey data collection is in process, Westat will maintain a help desk (using a toll free telephone number) that companies may call to ask questions about both the mechanics of the survey (such as how to access the survey and enter responses) and the survey content (e.g., if employers are uncertain of the meaning of a particular question).

(4) Training. All individuals working who will be in contact with potential respondents by phone or email will be trained in ways to optimize response without placing undo pressure on

potential respondents. In addition to general survey procedures, they will be trained to respond to specific questions that are likely to be raised in this study. This training will include help desk personnel as well as telephone interviewers.

Westat staff on the E-Verify qualitative team and who are intimately familiar with the Program will conduct the majority of the 80 telephone interviews with DAs and UDAs. A few more experienced telephone interviewers will also be trained on the E-Verify Program, and the general and specific protocol procedures and questions. (The recruitment of DAs and UDA employers will be conducted by the same Westat staff who successfully recruited employers and workers for the Arizona case studies conducted in 2009 and the national case studies conducted in 2008.)

(5) Nonresponse conversion. Experienced interviewers who are particularly skilled in nonresponse conversion will re-contact initial refusals. The major exception to this rule is for hard refusals (i.e., sampled companies who have requested not to be called again).

(6) Unit nonresponse adjustments. Weights will be used to adjust for nonresponse within cells identified by key variables known prior to sample selection (industry, location, and number of verifications).

(7) Editing and data cleaning. A number of editing features will be built into the Web survey. For example, if the respondent attempts to provide multiple answers to a question requiring a single response, the respondent will be asked to select only one response. Additional editing checks will be done subsequent to survey completion to check for completeness, and inter-item consistency, extraneous remarks, and, for respondents completing a mail survey, proper adherence to any skip instructions.

(8) Item nonresponse adjustments. Although our procedures are designed to maximize item response rates, the analysis will need to confront the issue of missing data. Experience with similar surveys indicates that some respondents will omit responses to some specific items (e.g., sensitive items), although they may have provided most of the data required. By employing good survey data collection practices, we expect to minimize the amount of missing data on any single variable to a very low level. However, if item nonresponse is unexpectedly high for any of the key analytic variables, hot deck imputation techniques will be used to estimate missing-item values.

For analyses involving just one or two variables that have not been subject to imputation, we will handle the problem by omitting the cases with missing data; or, in the case of categorical response variables, we will use an explicit “missing” or “unknown” category. When multivariate techniques involving several variables are used, analytic techniques for missing values will be used (such as using the variable mean or adding a dummy variable to reflect how the nonrespondents differ from the other companies).

Based on our prior experience in which we used incentives and extensive follow-up procedures, we do not believe that it is feasible to obtain a sufficiently high response rate to permit inferences from the sample to the entire population. In the 2008 evaluation, we achieved an unweighted 37 percent response rate for employees due to the inability to locate the sampled employees.

Employee contact information either was missing or incorrect and accurate updated information was unavailable from the employer, the tracing service, or neighbors. In a few cases, interviewers were fairly certain that the person they were trying to interview was the sampled employee, but the person denied that the identification was correct. Finally, a few workers refused to participate because they were afraid of employer retribution (i.e., they would be fired if their employer discovered they participated in the interview).

The purpose of the case studies is to examine in depth the procedures that employers and workers follow in the verification process, not to produce representative statistics. We are using sampling to ensure that a variety of employer/employee situations are examined, but do not require the statistics to be generalized in order to identify problems and potential solutions in the verification process.

Nonresponse bias studies

As part of the weighting process before analyzing the survey data, we plan on incorporating a nonresponse adjustment in the final weight. The final weight thus will be the inverse of the original probability of selection, multiplied by the inverse of the percentage who responded within that category of firms (e.g., if 70 percent responded within a particular stratum, we would multiply the base weight by $1/0.7$, or 1.43). The nonresponse adjustment will be based on comparing the rates of response in the various sampling strata; also, in general we expect to combine contiguous strata that have similar response rates. Our experience when conducting formal nonresponse bias studies is that this approach is effective in adjusting for nonresponse bias.

The above approach provides a way of comparing the characteristics of respondents and nonrespondents, and also statistically adjusts for those differences that are identified. Of course, it is possible that respondents and nonrespondents differ systematically in ways that are unrelated to the sampling strata. Depending on the nature of those differences, we may or may not have data on nonrespondents available in the TDB that could also be used in the nonresponse adjustments. As part of the data collection process, we will watch for potential systematic factors that affect response rates; to the extent that we identify such biases, we will examine the TDB for measures that might be correlated with such factors. Additionally, depending on the level of nonresponse, we will consider comparing key items among the questionnaire responses of early responders with those of late responders as a way of making additional data available on response patterns, and thus of making inferences about the characteristics of nonrespondents.

The interview data are not intended to provide nationally representative statistics, and thus will not be weighted. However, we will examine whether the respondents differed in systematic ways from the nonrespondents, and if so, discuss the implications for interpreting the data.

4. Tests of Procedures for Refining Data Collections

The employer survey instrument submitted in this request for clearance was well pre-tested during prior evaluations. Some changes have been made to accommodate the differences in programs and scope compared to the previous employer data collection activities. New questions will be explored with employers during a focus group.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The following statisticians were consulted on the statistical aspects of the design and analysis of the current study:

Carolyn Shettle
Westat
1600 Research Blvd.
Rockville, MD 20850
301-251-4324
carolynshettle@westat.com

Huseyin Goksel
Westat
1600 Research Blvd.
Rockville, MD 20850
301-251-4395
huseyingoksel@westat.com

The following individuals will collect and/or analyze data for the current study:

Brad Chaney
Westat
1600 Research Blvd.
Rockville, MD 20850
301-294-3946
bradchaney@westat.com

Denise Glover
Westat
1600 Research Blvd.
Rockville, MD 20850
301-251-2269
deniseglover@westat.com