

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.

1. Respondent Universe

The target population of the Survey of E-Verify Employers includes all firms that signed an MOU before September 30, 2012 in which they agreed that all or part of the firm would participate in E-Verify with the following exceptions:

- **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 April 1, 2012 and September 30, 2012. Excluding companies with no recent involvement in E-Verify helps to ensure that company representatives can accurately recall aspects of their participation, and also avoids the possibility of companies responding to a program characteristic that has recently changed.
- **E-Verify Employer Agents (EEAs) and their clients** will be excluded because this population was studied in the 2010 case study. In addition, this qualitative work revealed an extensive amount of variation among the characteristics, operations, and experiences of EEAs and their clients that would be difficult to capture in an online survey. We plan to include these employers in the 2014 employer survey which will allow time to develop and pretest questions that will adequately cover the unique experiences of these populations.
- **Employers in Puerto Rico, Guam and other U.S. territories** will be 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:

- **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;

- **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;
- **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:

- Contact information from the Point of Contact file will be appended to the appropriate employer records;
- 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 parent company), information will be obtained through Web searches and/or telephone inquiries to determine the appropriate firm level information¹;
- 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 2,800 completed surveys based on our experiences with the FY2010 and FY2008 surveys. This estimate assumes 8 percent of sampled cases are excluded because the intended respondent is ineligible for inclusion, and a survey response rate of approximately 80% of eligible respondents.

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

¹ The Employer Database contains a mix of establishments and firms, because of the E-Verify enrollment procedures.

- **Active Employers with TNCs:** employers that had transmitted one or more cases receiving a TNC between July 1 and September 30, 2012 *and* had not formally terminated participation in the program before October 1, 2012. 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 July 1 and September 30, 2012, had no cases receiving a TNC, and had not formally terminated participation in the program before October 1, 2012. These employers may well have different experiences from those active employers with TNCs, and these differences may impact both their satisfaction and compliance with E-Verify procedures. Thus, 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 January 1 and March 31, 2012 but had not transmitted any cases between April 1, and September 30, 2012, or had formally terminated participation in the program between April 1, and September 30, 2012. The insights from these employers 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:

- **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.
- **Industries** (e.g., hospitality services and food processing) known to have relatively large percentages of undocumented workers.² 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.
- **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.

Cross-classifying E-Verify companies by E-Verify status and industry provides a total of nine strata. These strata, their estimated population sizes (including ineligible cases that cannot be identified prior to sample selection), and proposed sample sizes are included in Table B-1. This

² 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 is based on information as of September 30, 2012, and represents the most recent transaction data that we have.

Table B-1. Estimated population and sample sizes of companies for the Survey of E-Verify Employers

	Industry sector	Population	Sample draw	Expected response rate ³	Expected # of completes	Percent sampled
Active (TNCs)	Staffing Agencies	480	400	0.80	320	83.3
	High percent undocumented	3,235	675	0.80	540	20.9
	Other	5,927	751	0.80	601	12.7
Active (no TNCs)	Staffing Agencies	378	302	0.80	242	80.0
	High percent undocumented	14,699	425	0.80	340	2.9
	Other	36,045	425	0.80	340	1.2
Inactive	Staffing Agencies	45	45	0.62	28	100.0
	High percent undocumented	4,449	350	0.62	217	7.9
	Other	11,570	350	0.62	217	3.0
	<i>Total</i>	76,828	3,723	0.76	2,845	4.8

Notes: Data included in the table are based on information for E-Verify employers as of September 30, 2012. 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).⁴ Classification of industries by percent undocumented is based on the most recent industry data available at the time this table was created: 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>). If available, updated data will be used when drawing the survey sample.

Within strata, we propose using probability proportional to size (PPS) sampling, where the square root of the number of transactions submitted between July 1 and September 30, 2012 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. s. 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 provides a good compromise between the two objectives.

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.

³ The response rate for the 2010 Web Survey of Employers was 83 percent.

⁴ Twenty percent of employers were ineligible in the 2010 data collection because they were no longer in business, were duplicate listings of a company, or were EEAs, or clients of EEAs (two groups of employers that were excluded from the study).

Power Analysis

As in the 2010 survey, the 2013 survey is designed to collect data for both categorical and continuous variables. This power analysis focuses on categorical variables because they are more important to address the key study questions.

From the results of the 2010 survey, we estimated the design effect (DEFF). For 8 key categorical variables that characterize the whole survey population, the estimated DEFF's range is from 2.3 to 5.7, and their average is 4.6. Considering that the 2013 survey is similar in key aspects of the design (in particular, the composition of the population and the definition of the measure of size) to the 2010 survey, the same DEFF for the 2013 survey is assumed in this power analysis.

The power is calculated under the following set-up:

1. The 2010 survey results are compared with the results from the 2013 survey using a two-sided normal test. While the t-test is usually used for these comparisons, degrees of freedom are large enough to use the normal test.
2. The significance level of the test is set at the usual level of 5 percent.
3. The base proportions (i.e., 2010 population proportions) denoted as P_0 are assumed to be 30 and 50 percent for two binomial variables.
4. The corresponding proportions in 2013 (denoted as P_1) are assumed to be larger than the 2010 counterparts (expecting an increasing trend).
5. The respondent sample size for the 2010 survey was 2,928, and that for the 2013 survey is expected to be 2,845.
6. The DEFF for both surveys is assumed to be 4.6. Thus, the effective sample size is 637 and 618, respectively, for the 2010 and 2013 surveys.

The following table provides the powers for various increments of the 2013 proportions from the 2010 proportions under the set-up described above. The table shows, for example, if a 2010 proportion of 30 percent has increased to 37 percent in 2013, such change can be detected by the normal test with 83.8 percent chance (i.e., power). However, if the 2010 proportion was 50 percent, the power to detect 7 percent change in 2013 is 66.6 percent.

Increment (%) <i>Δ</i>	Power (%)	
	$P_0 = 30\%$	$P_0 = 50\%$
5.0	56.5	39.8
5.5	64.4	46.5
6.0	71.8	53.4
6.5	78.3	60.1
7.0	83.8	66.6
7.5	88.3	72.6
8.0	91.8	78.0
8.5	94.5	82.8
9.0	96.4	86.9
9.5	97.7	90.2
10.0	98.6	92.9

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.

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 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 2010 survey of employers which obtained an overall unweighted survey response rate of 83 percent and a weighted response rate of 83 percent. These employers have signed MOUs 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 2013 questionnaire is based directly on the 2010 Survey of E-Verify Employers, and thus has been tested through both pretests and full data collections. Knowledge obtained from conducting the previous surveys was used to modify the 2010 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 2010 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 2010 survey. In addition, in January and February of 2012 we conducted focus groups and telephone interviews to examine the 2013 questionnaire, focusing particularly on items that are new or that have been modified from 2010.

(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 privacy; 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 individuals identified as primary contact persons in USCIS materials. This letter will be from USCIS (Attachment E) and will state that this is part of the evaluation effort the employers 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 persons reiterating the importance of the study and providing information on how to log on to the Web site (Attachment E).
- o At the same time that the emails from Westat are sent out, a “news” 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 who received a survey request to complete the survey.

- 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.
- 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.
- 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 (e.g., the employer now uses an E-Verify Employer Agent to perform all verifications).
- 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.
- Approximately two weeks after the reminder email, a second reminder e-mail will be sent to non-respondents.
- 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.
- 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 on the study who will be in contact with potential respondents by phone or email will be trained in ways to optimize response without placing undue 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.

(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 previous 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).

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 series of focus groups.

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:

Basmat Parsad
Westat
1600 Research Blvd.
Rockville, MD 20850
301-294-3946
basmatparsad@westat.com

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