

Appendix A  
Methodological Notes



## APPENDIX A. METHODOLOGICAL NOTES

These methodological notes provide additional details on the methods used in the evaluation, including information about weighting, variance estimation, and statistical testing. This section supplements the information outlined in Chapter II.

### A. BACKGROUND

The sample selection and estimation used for data collection for the pre- and post-test employer mail surveys, employer on-site interview, Form I-9 sample, and employee on-site interview are based on probability sampling methods. Probability samples allow the evaluation team to compute sampling weights, estimate the precision of sample estimates, and test the statistical significance of study hypotheses. Nonresponse adjustment through weighting implies that, within adjustment cells, nonrespondents are similar to respondents in the characteristics or behavior reported in the surveys or interviews. To the extent that this assumption does not hold, bias in the estimates may result.

### B. EMPLOYER WEIGHTING

#### 1. FIRST EMPLOYER MAIL SURVEY

Data from a sample survey typically need to be weighted to achieve an unbiased estimator of the population characteristics. The completed interviews formed the basis of an analytic data set, which required weighting to produce unbiased estimates of the eligible target population of establishments. The following subsections describe the calculation of the weights, including nonresponse adjustment and the calculation of initial weights.

##### *a. INITIAL WEIGHTS*

Let  $p_i^{M1}$  be the selection probability of the  $i$ th establishment for the first mail survey. The initial weight of the  $i$ th establishment,  $W_{1,i}^{M1}$ , for the first mail survey is equal to the inverse of the selection probability, that is,

$$W_{1,i}^{M1} = 1/p_i^{M1}.$$

Since all 96 establishments were selected to be included in the survey,  $p_i^{M1}$  is 1 and the initial weight is equal to 1.

##### *b. NONRESPONSE ADJUSTMENT*

In the presence of nonresponse, statistics calculated from a survey may be biased as estimators of the corresponding population characteristics if the nonrespondents and respondents have different characteristics. Therefore, for each sample member on the file,

a weight is produced that adjusts for establishment nonresponse. The nonresponse-adjusted weights were based on weighting-cell adjustment methods,<sup>1</sup> which are described below.

The nonresponse adjustment cells were constructed by classifying the sample establishments by size.<sup>2</sup> Each eligible establishment was classified into one of several mutually exclusive adjustment cells based on its size. Within each adjustment cell, if there were at least 20 establishments and the response rate was not less than two-thirds of the overall response rate, the nonresponse adjustment factor was calculated. When a given cell did not meet these criteria, a new set of adjustment cells was constructed.

Let  $I_i^{M1} = 1$  if the  $i$ th employer responded to the mail survey; otherwise,  $I_i^{M1} = 0$ . The cell-specific response rate ( $R_{W.C}^{M1}$ ) in the  $c$ th cell is

$$R_{W.C}^{M1} = \frac{\sum_{i=1} (W_{1.ci}^{M1} \cdot I_{ci}^{M1})}{\sum_{i=1} W_{1.ci}^{M1}},$$

where the summation is over the eligible units. Exhibit A-1 shows the cell-specific response rates and the number of eligible establishments in each cell.

**Exhibit A-1: Response Rates by Establishment Size in the MRDP First Mail Survey of Employers**

<b>Establishment Size (employees)</b>	<b>Response Rate (%)</b>	<b>Number of Eligible Employers</b>
1-99	84.5	33
100-249	91.7	36
250 and over	81.0	21
<b>Total</b>	<b>86.7</b>	<b>90</b>

The nonresponse-adjusted weight,  $W_{2.i}^{M1}$ , for the  $i$ th responding establishment is

$$W_{2.i}^{M1} = \frac{W_{1.ci}^{M1}}{R_{W.C}^{M1}}.$$

The nonresponse-adjusted weights should be used for unbiased estimation of statistics for the first mail survey.

<sup>1</sup> Little and Rubin, 2002.

<sup>2</sup> The establishment size was also used for nonresponse adjustment for Basic Pilot surveys.

## 2. SECOND EMPLOYER MAIL SURVEY

All employees using the pilot and responding to the first mail survey were included in the second mail survey. All 29 eligible establishments responded to the survey. The nonresponse-adjusted weight,  $W_{2,i}^{M2}$ , for the  $i$ th establishment is

$$W_{2,i}^{M2} = W_{2,i}^{M1}.$$

The nonresponse-adjusted weight was used for unbiased estimation of statistics for the second mail survey.

## 3. TELEPHONE SURVEY OF EMPLOYER NON-USERS

Employers that responded to the first mail survey but had never used the pilot were included in the telephone survey. All 46 eligible establishments responded to the survey. The nonresponse-adjusted weight,  $W_{2,i}^P$ , for the  $i$ th establishment is

$$W_{2,i}^P = W_{2,i}^{M1}.$$

The nonresponse-adjusted weights were used for unbiased estimation of statistics for the telephone survey.

## 4. EMPLOYER ON-SITE INTERVIEW

All 29 quasi-experimental employers responding to the second mail survey were visited for the on-site interview. An additional 12 non-experimental employers were also included in the on-site interview. All 41 employers participated in the on-site interview. The nonresponse-adjusted weight,  $W_{2,i}^O$ , for the  $i$ th establishment is

$$W_{2,i}^O = W_{2,i}^{M1} \text{ for the 29 respondents to the second mail survey and}$$

$$W_{2,i}^O = 1 \text{ for the additional 12 establishments.}$$

The nonresponse-adjusted weights were used for unbiased estimation of statistics for the on-site interview.

## 5. LIMITATION OF WEIGHTING

The nonresponse-adjusted mail and on-site weights were constructed by weighting the respondents to known population totals in the given size categories. To the extent that there are differences in the survey items or variables between the respondents and nonrespondents in each size category, however, some nonresponse bias will remain.

## C. ANALYSIS OF I-9 FORMS

The first step in selecting I-9 forms was for the interviewer to locate I-9 forms for all employees receiving tentative nonconfirmations at the establishment visited for an on-site

interview, using a list compiled for that purpose from the transaction database. These cases were not eligible for selection during the second and third steps in the process.

The second step consisted of sampling I-9 forms for the pre-test sample of employees. The number of forms to be selected was based on the estimated number of employees hired in the 6 months before the date that the employer received MRDP materials. The number of I-9 forms collected was based on an estimate reported by pilot employers. If the employer reported that 190 or fewer employees had been hired in the 6-month period prior to the start of the MRDP, the interviewer was to photocopy all I-9 forms for these employees. If more than 190 employees were hired during this period, the interviewer selected a systematic random sample from among the I-9 forms.

The third step consisted of sampling I-9 forms for the post-test sample of employees. The same general procedures were used to select these forms that were used for the pre-test sample except for the period for which cases were sampled. For quasi-experimental employers, the post-test included employees hired from the start of the MRDP to the date of the interview. For the 12 non-experimental employers, I-9 forms were collected for employees hired during the 6 months preceding the interview.

Weights for the I-9 forms collected during steps two and three were set equal to the product of the on-site employer survey weight and the inverse of the Form I-9 sampling rate used by the interviewer in selecting the forms for the specified time period. For I-9 forms collected during the first step, the case was given a Form I-9 weight equal to the transaction database weight, if the case was in scope for the Form I-9 sample. Being in scope meant that an I-9 form was found for the employee during step one and that the employee was hired during the time period used for selecting I-9 forms for that employer. If the employee was out of scope, the Form I-9 weight was set to 0.

#### **D. ANALYSES OF THE TRANSACTION DATABASE**

INS provided the evaluation team with a data file consisting of 32,663 records from the contractor responsible for capturing information input by employers and Immigration Status Verifiers, as well as the results of the automated verification processes. The file contains data from the inception of the MRDP program until March 18, 2002. The evaluation team merged multiple records for the same hiring event [e.g., a Social Security Administration (SSA) finding and an INS finding] and eliminated duplicate records. The resulting file, referred to in this report as the transaction database, contains data for 21,422 hiring situations (defined as a unique employer/employee combination).

Because basic information is available for all cases transmitted to the transaction database, the resulting sample can be viewed as constituting a census of all transmitted cases and is therefore not subject to sampling error. However, nonsampling errors cannot be completely eliminated. One source of nonsampling error resulted from inaccuracies that occurred during the resolution of duplicate cases. Identification of duplicate cases was not always straightforward. When the employee's name and Social Security number did not match exactly, data coders had to scan the cases visually to determine whether they were duplicates. Thus, the unduplication process was subject to classification errors.

## 1. EMPLOYEE WEIGHTING

The MRDP evaluation also included in-person interviews with employees of those employers that used the system. Employees were selected from two separate but overlapping sampling frames: the employee transaction database sample and the Form I-9 sample.

### *a. EMPLOYEE TRANSACTION DATABASE SAMPLE*

#### **Sampling Frame**

The first sampling frame consisted of all employees having records on the transaction database as of January 2002<sup>3</sup> who met one or both of the following criteria:

1. All employees electronically verified through the MRDP on or after June 2001. This sample was restricted by time rather than taking a random sample because older cases are harder to locate than more recent cases (approximately 3,600 cases).
2. All employees who received tentative nonconfirmations from the program (i.e., those who were not immediately verified by the MRDP as work-authorized). These cases are of greatest interest to the evaluation (approximately 600 additional cases).

The transaction database contains the information electronically transmitted by employers during the automated verification process (the database contains limited information such as Social Security number, name, and verification results). This database also includes information from the INS contractor that manages the database of employer queries for new employees' work authorizations.

#### **Base Weights**

The base weight for a transaction database employee sample record is computed as the inverse of the selection probability.<sup>4</sup> The base weight for the  $i$ th employee in the transaction database sample is given by

$$w_i = \frac{1}{p_i},$$

where  $p_i$  is the overall probability of including the  $i$ th employee in the sample.

---

<sup>3</sup> The evaluation team excluded 13,184 employees of employers under investigation by INS. These employees constitute 62 percent of employees screened by the MRDP.

<sup>4</sup> The sample was not selected on a probability basis, but is treated as a probability sample for weighting purposes.

## Nonresponse Adjustment

The total number of cases selected from the transaction database was 4,402. However, only 2,044 cases were assigned for interview. The remaining cases had no addresses or addresses outside of the interview area (Iowa and locations close to Iowa in neighboring States). The weights assigned to the 532 respondents in the employee sample are based on information about all 4,402 cases. To conduct nonresponse adjustment for the transaction database sample, a response status was assigned to every sample record based on the final field disposition documented in the receipt control file. Since the eligibility of some nonrespondents was not known, one of the following four response status groups was assigned to each sample record:

**Group 1: Responding records.** This group consists of sample records for all eligible employees who provided substantially complete and usable survey data (n=532).

**Group 2A: Eligible nonresponding records.** This group consists of sample records for all employees who were eligible but did not provide substantially complete and usable survey data, such as employees who refused, were not at home, or lived in Iowa but could not be located (n=88).

**Group 2B: Other nonresponding records.** This group consists of sample records for all nonresponding employees whose eligibility could not be ascertained (n=1,415).

**Group 3: Ineligible records.** This group consists of all sample records that are ineligible for the study (n=2,366).

**Stratification by outcome.** The first step in the weighting process was to divide employee cases into strata on the basis of case outcomes. Since 487 of the 532 interviewed employees had received immediate clearance from SSA, only two strata were used: SSA first-stage employment authorized and all other findings. Since there were only 45 cases in the second stratum, further division of this stratum was not feasible. The following weighting steps were then performed for each of the two strata separately.

**First-stage adjustment.** The record-level nonresponse adjustment was made in two stages. The first stage distributed the weights of Group 2B to the remaining three groups, for which eligibility had been determined. That is, the first-stage nonresponse adjustment factor within an adjustment class  $c$  is defined as the following ratio of sums:

$$F_c^{(1)} = \frac{\sum_{i=1}^{n_1} w_i + \sum_{i=1}^{n_{2A}} w_i + \sum_{i=1}^{n_{2B}} w_i + \sum_{i=1}^{n_3} w_i}{\sum_{i=1}^{n_1} w_i + \sum_{i=1}^{n_{2A}} w_i + \sum_{i=1}^{n_3} w_i},$$

where the sums extending over  $n_1, n_{2A}, n_{2B}$ , and  $n_3$  correspond to each of the four groups.

The adjustment factor  $F_c^{(1)}$  was then applied to the base weight of the  $n_1 + n_{2A} + n_3$  sample records for which eligibility could be determined; that is, the first-stage nonresponse-adjusted weight,  $w_i^{A1}$ , is calculated as follows:

$$w_i^{A1} = F_c^{(1)} w_i \text{ for records with eligibility status determined (Groups 1, 2A, and 3)}$$

and

$$w_i^{A1} = 0 \text{ for "other nonresponding" records (Group 2B).}$$

**Second-stage adjustment.** At the second stage, the previously adjusted weight was distributed to the responding records in the sample (Group 1). The second-stage nonresponse adjustment factor within an adjustment class  $c$  is defined as the following ratio of sums:

$$F_c^{(2)} = \frac{\sum_{i=1}^{n_1} w_i^{A1} + \sum_{i=1}^{n_{2A}} w_i^{A1}}{\sum_{i=1}^{n_1} w_i^{A1}},$$

where the sums extending over  $n_1$  and  $n_{2A}$  correspond to the responding and eligible nonresponding groups.

**Final weights.** Analysts applied the second-stage adjustment factor to the first-stage nonresponse-adjusted weight of the  $n_1$  responding records (Group 1) in the sample. That is, the final nonresponse-adjusted weight,  $w_i^{A2}$ , for the transaction database is calculated as follows:

$$w_i^{A2} = F_c^{(2)} w_i^{A1} \text{ for responding record (Group 1);}$$

$$w_i^{A2} = 0 \text{ for an eligible nonresponding record (Group 2A); and}$$

$$w_i^{A2} = w_i^{A1} \text{ for an ineligible record (Group 3).}$$

#### **b. THE FORM I-9 SAMPLE**

The second sampling frame consisted of I-9 forms (which are supposed to be completed by all new employees and only new employees, whether or not they are participating in a pilot program) collected from the 41 employers interviewed on-site. The interviewers were guided by the sampling forms in selecting the I-9 forms. For the 29 quasi-experimental employers, two samples were selected:

- The Form I-9 pre-test sample, which consisted of employees hired during the 6 months prior to the receipt of the MRDP materials
- The Form I-9 post-test sample, which covered the time from the receipt of the MRDP materials to the time of the interview.

For the 12 non-experimental employers, I-9 forms were selected for the 6 months preceding the interview date.

### **Initial Weights**

The base weight for a Form I-9 employee sample record is computed as the inverse of the selection probability. The base weight for the  $i$ th employee in the Form I-9 sample is given by

$$w_i = \frac{1}{p_i},$$

where  $p_i$  is the probability of selecting the I-9 form during sample selection.

### **Final Weights**

The final weights for the Form I-9 cases were calculated by multiplying the base weight for the employee by the employer's on-site weight.

### **Combined Weights**

In some cases, statistics are based on information from all respondents in either the Form I-9 or transaction database sampling frame. To prevent undue weight from being given to employee cases in both the Form I-9 sample and the transaction database sample, cases included in both sampling frames were given a combined weight equal to their transaction database weight. For the Form I-9 cases that were not on the transaction database, the Form I-9 weight was used as the combined weight.

## **E. VARIANCE ESTIMATION AND TESTS OF SIGNIFICANCE**

In sample surveys, the variance is used to estimate the precision of the survey estimates. This is especially important when the researcher wants to know whether any differences observed can be explained by chance. Tests of significance indicate the likelihood that observed differences occurred by chance.

Statistical software packages such as SPSS and SAS permit variance estimation for equal probability samples when there is no nonresponse weighting adjustment. These variance estimates are not appropriate for use with the sampling and nonresponse adjustment methods used for the employer surveys and the employee interviews. Instead, in computing test statistics the evaluation team used WesVarPC for variance adjustment among respondent subgroups.

The variance estimates generated by WesVarPC were used for statistical tests of significance. The significance level for the statistical tests was set at 0.05 ( $\alpha=5\%$ ), a commonly used significance level.