# Appendix C

## Estimation of the Work-Authorization Status of Unresolved Cases

# APPENDIX C. ESTIMATION OF THE WORK-AUTHORIZATION STATUS OF UNRESOLVED CASES

## A. BACKGROUND

Several of the pilots' key goals, as articulated in the Illegal Immigration Reform and Immigrant Responsibility Act (IIRIRA) and by stakeholders, required the evaluation team to estimate work-authorization rates for various groups of newly hired employees in establishments participating in the Machine-Readable Document Pilot (MRDP). Examination of the transaction database provided only limited information of use in meeting these goals. The work-authorization rate could not be accurately estimated, since only 0.07 percent of all cases were determined by the MRDP system to be unauthorized, while 7.16 percent of the cases were final nonconfirmation cases. In other words, on the basis of this information alone, the estimated percentage of screened employees who were not work-authorized was between 0.07 and 7.23 percent. This range is too broad to provide a meaningful estimate.

Information from employer and Federal interviews indicated that the final nonconfirmation cases included a mix of work-authorized and non-work-authorized employees. However, this information was not specific enough to provide precise estimates of the percentage of the cases in each category. The evaluation team therefore developed a model for estimating the work-authorization status of employees who had information transmitted to the MRDP. Information about the model is provided in this appendix.

## B. MODEL-BASED ESTIMATION FOR SSA CASES

For employees whose records were never sent to INS, the model used information on the observed relationship between the initial findings of the Social Security Administration (SSA) database match and final case resolution (i.e., authorized, not authorized, or final nonconfirmation) to estimate the percentage of unauthorized employees. The model also included assumptions that have not been empirically tested.

Exhibit C-1 provides the basic model for cases in which SSA issues a tentative nonconfirmation. The bold letters in parentheses on the exhibit are for reference purposes.

The purpose of this model is to estimate how many employees who received final nonconfirmation outcomes from SSA would have been found to be work-authorized, given what is known about the cases and a set of "reasonable assumptions." For each SSA tentative nonconfirmation case, the transaction database provides the following information:

- Final case disposition (i.e., authorized or final nonconfirmation)
- Initial SSA finding (i.e., why the computer match failed)

This information is insufficient for estimating how many work-authorized employees are in the final nonconfirmation category. Therefore, the evaluation team had to use additional reasonable assumptions to estimate the number of employees with final nonconfirmations who would have been determined to be work-authorized by the MRDP system if all cases had been resolved. The following assumptions were used:

1. The probability that an employee receiving a final nonconfirmation from SSA is actually work-authorized[1] depends on the initial reason for the case not being matched on the SSA database. For example, it is reasonable to believe that there are more work-authorized individuals among those non-matched cases for which the employee's name did not match the SSA database than among those for whom both the name and date of birth disagreed. This assumption is consistent with data on the percentage of employees in each category who contested tentative nonconfirmations, assuming that employees in categories with high concentrations of authorized employees are more likely to contest than those in categories with few authorized employees. Employees with employer-input names that disagreed with SSA names were more likely to contest than were employees with a date of birth that did not match the SSA database (18 percent compared to 1 percent) (Exhibit C-2).

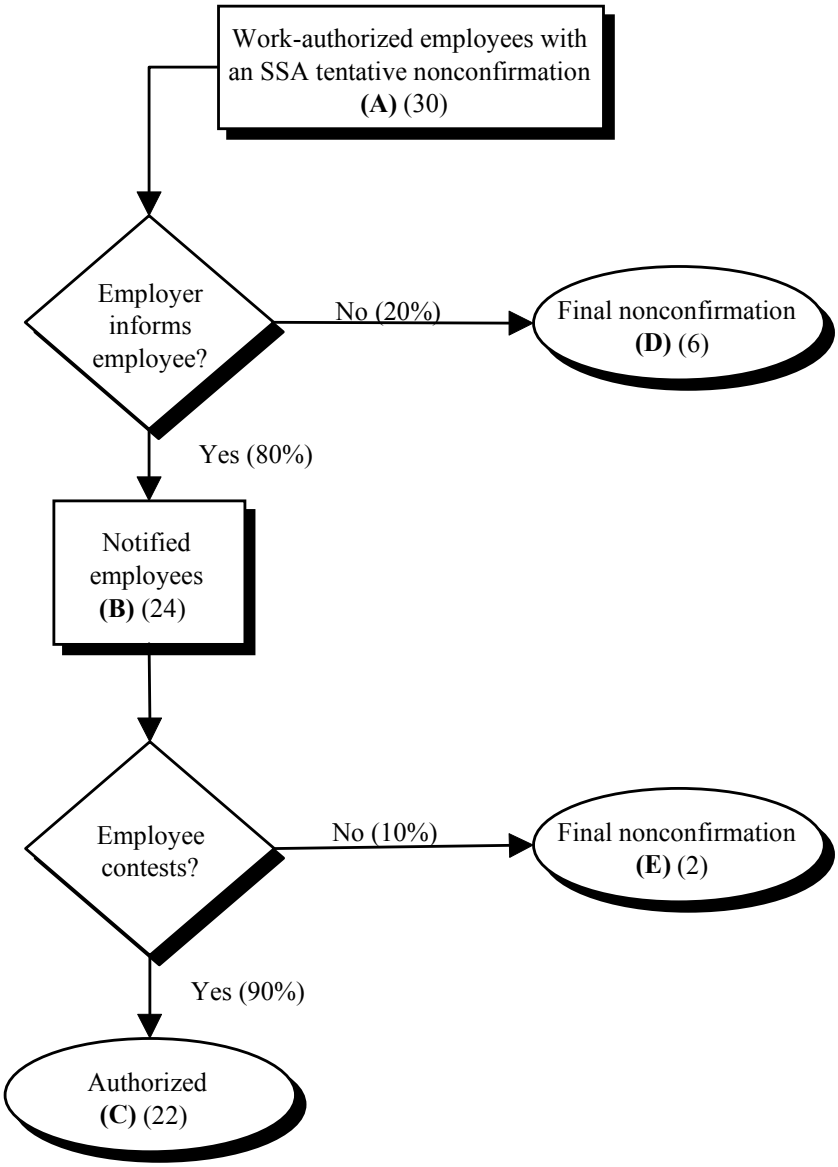
2. The percentage of employees informed by their employers of a tentative nonconfirmation from SSA does not depend on the reason for issuing the tentative nonconfirmation, since the employer does not know the reason for the tentative nonconfirmation finding. For example, employees not matched because of an invalid Social Security number were no more or less likely to have been informed of nonconfirmation than were employees whose names did not match the SSA database. In the model, the user estimates this percentage, so alternative scenarios can be tested. Exhibit C-1 models the probability that the employee will move from **(A)** to **(B)** (from tentative nonconfirmation to notification). The illustration assumes that the user has set the percentage of notified employees equal to 80 percent.

3. The percentage of *work-authorized* employees contesting SSA tentative nonconfirmations does not depend on the reason for issuing the tentative nonconfirmation. For example, work-authorized employees not matched because of an invalid Social Security number are no more or less likely to contest than are employees who did not match on date of birth. In the model, the user estimates this percentage, so alternative scenarios can be tested. Exhibit C-1 models the probability that the employee will move from **(B)** to **(C)** (from notification to authorization). The illustration assumes that the user has set the percentage of employees who contest tentative nonconfirmations equal to 90 percent.

---

[1] To simplify the explanation of the model, employees who were or would have been authorized by the system are referred to as work-authorized. In reality, as discussed in the report, some employees determined to be work-authorized were not actually work-authorized.

**Exhibit C-1: MRDP Process Between SSA Tentative Nonconfirmation and Final Determination for Work-Authorized Employees***




* The numbers refer to employees with an initial SSA determination of "name disagrees with SSA" and assume that 80 percent of employees are informed of the tentative nonconfirmation and that 90 percent of work-authorized employees contest the tentative nonconfirmation.

SOURCE: The number authorized **(C)** is from the transaction database. The remaining numbers are estimated.

**Exhibit C-2: Percentage of Employees Receiving Tentative Nonconfirmations from SSA Whose Cases Were Resolved, by Initial SSA Finding***

| Initial SSA Finding | Percent of Tentative Nonconfirmation Cases Resolved | No. of Tentative Nonconfirmations |
|---|---|---|
| Invalid Social Security number | 0.0 | 170 |
| Date of birth disagrees with SSA database | 1.3 | 310 |
| Name disagrees with SSA database | 17.7 | 124 |
| Name and date of birth disagree with SSA database | 0.6 | 640 |
| Missing initial Social Security Code or Social Security number belongs to deceased person | 0.0 | 70 |
| Unlawful permanent resident – INS must confirm work authorization | 32.1 | 209 |
| **Total** | **6.4** | **1,523** |

* The table includes only cases that were verified only by SSA.

SOURCE: Transaction Database

The number of employees determined to be work-authorized for a given SSA initial finding is known. For example, 22 employees with "name disagrees with SSA database" were authorized **(C)**.

Mathematically, the number of authorized cases **(C)** is equal to the number of contested cases **(B)** multiplied by the probability that a work-authorized employee will contest the finding. The probability that a work-authorized employee will contest a finding is a user-driven input. For illustrative purposes, assume that the model user has estimated that 90 percent of work-authorized employees will contest an SSA tentative nonconfirmation. In this case, **(C)** = 0.9 * **(B)**. Mathematically, this is equivalent to **(B)** = **(C)**/0.9. In the example, **(B)** = 22/0.9 = 24.

The number of employees who were informed of a tentative nonconfirmation can be estimated in a similar fashion, given the user-input assumption of the percentage of work-authorized employees with tentative nonconfirmations who are informed of their status. Continuing with the example and assuming that 80 percent of employees with SSA tentative nonconfirmations are notified, the number of work-authorized employees can be estimated as **(A)** = **(B)**/0.8 = 30. Thus, the model estimates that the MRDP would have found 30 employees with employer-input names that did not match their SSA names initially to be authorized, if all employees had been properly notified and had chosen to contest their cases.

## C. MODEL-BASED ESTIMATES FOR INS

The basic model for cases in which INS issues a tentative nonconfirmation is the same as the model for the SSA cases except that the estimates of final case outcomes are based on the relationship between the initial ISV finding and the final case finding.

## 1. RANGE ESTIMATION

To obtain a preliminary estimate of the range of possible values for the percentage of unauthorized employees, two scenarios were tested. In the first, all of the user-input parameters were set to 100 percent. This scenario assumes that all work-authorized individuals have already been identified by the system. As expected, this results in an estimate of the percentage unauthorized of 7.2 percent. This is the maximum value.

To obtain a reasonable minimum value, the evaluation team assumed that the product of the percentage of tentative nonconfirmation cases who are informed and the percentage of informed tentative nonconfirmation cases who contest is set equal to the minimum value consistent with the observed rate of employees who actually contested. This resulted in an estimate of 5.9 percent. Thus, the range of estimated values is from 5.9 to 7.2 percent.

## 2. MATHEMATICAL FORMULATION OF THE MODEL

The tentative nonconfirmation cases were assigned to *k* strata, based on the reason they were classified as tentative nonconfirmations (see Exhibit C-2).

Assuming that a work-authorized finding occurs only when work-authorized employees are told that they have tentative nonconfirmations and when the employees contest their cases, the observed percentage of work-authorized cases in stratum *k* can be defined as described below:

The total number of work-authorized employees is equal to

$$N_{\bullet}^{WT} = \sum_{k} N_{k}^{WV} + N_{\bullet}^{WI}, \tag{1}$$

where

$N_{\bullet}^{WT}$ = the total number of work-authorized employees in the population;

$N_{k}^{WV}$ = the number of work-authorized employees who received tentative nonconfirmations in stratum *k;* and

$N_{\bullet}^{WI}$ = the total number of work-authorized employees in the population who were resolved as work-authorized without a tentative nonconfirmation.

Since $N_{\bullet}^{WI}$ is observed, $N_{\bullet}^{WT}$ can be estimated from an estimate of $\sum_{k} N_{k}^{WV}$.

A formula for estimating $N_k^{WV}$ can be derived by solving the following equation for $N_k^{WV}$:

$$N_k^{WR} = N_k^{WV} \bullet t_k \bullet c_k , \tag{2}$$

where

$N_k^{WR}$ = the number of tentative nonconfirmations in stratum *k* that were resolved as work-authorized;

$t_k$ = the percentage of work-authorized employees with tentative nonconfirmations who were told of the tentative nonconfirmation; and

$c_k$ = the percentage of work-authorized employees told of a tentative nonconfirmation who contested and resolved the finding.

Equation (2) can be solved for $N_k^{WV}$ as follows:

$$N_k^{WV} = N_k^{WR} / t_k \bullet c_k . \tag{3}$$

The maximum value of $N_k^{WV}$ is the total number of tentative nonconfirmations in stratum *k* ($N_k^{TV}$). This maximum occurs when $t_k \bullet c_k = N_k^{WR} / N_k^{TV}$:

$$N_k^{WV} = N_k^{WR} / ( N_k^{WR} / N_k^{TV} ),$$

which is equivalent to

$$N_k^{WV} = N_k^{WR} \bullet N_k^{TV} / N_k^{WR} .$$

Further, the value of ($t_k \bullet c_k$) cannot be greater than 1, since both $t_k$ and $c_k$ are proportions.

Thus, the range of $t_k \bullet c_k$ is given as

$$( N_k^{WR} / N_k^{TV} ) \geq t_k \bullet c_k \geq 1. \tag{4}$$

To simplify the model, assume that the initial reason for receiving the tentative nonconfirmation affects neither the probability that the employer will tell the work-authorized employee of the tentative nonconfirmation nor the probability that the employee will contest the finding after being informed. In this situation, $t_k$ and $c_k$ are constant across strata (i.e., $t_k \bullet c_k = t_\bullet c_\bullet$ for all *k*).

Given this assumption, equation (3) becomes

$$N_k^{WV} = N_k^{WR} / t_\bullet c_\bullet , \tag{5}$$

and summing both sides of the equation leads to

$$\sum_k N_k^{WV} = \sum_k N_k^{WR} / t_\bullet c_\bullet \text{ and} \tag{6}$$

$$\sum_k N_k^{WV} = (1/t_\bullet c_\bullet) \bullet \sum_k N_k^{WR}. \tag{7}$$

$\sum_k N_k^{WR}$ is known and is equal to 22 in the example used in Exhibit C-2.