



U. S. Department of Labor

Office of the Chief Financial Officer
and
Employment and Training Administration

Unemployment Insurance
National Directory of New Hires Pilot Report

September 21, 2005

Table of Contents

<u>Section</u>	<u>Page Number</u>
I. Executive Summary.....	2
II. Purpose	5
III. Methodology	5
IV. Summary of Results	6
V. Conclusions	8
 <u>Appendices</u>	
A. Pilot Methodology	9
B. Detailed Results	14
C. List of Acronyms.....	20

I. Executive Summary

Congress enacted the Improper Payment Information Act (IPIA) in 2002 to require Federal departments and agencies to address the problem of erroneous payments. The act requires the head of each agency to review all programs and activities using a risk assessment to determine their susceptibility to improper payments. Last year, OMB identified the Unemployment Insurance (UI) program as one of the top seven government programs at risk of significant improper payments. The UI program, administered by the Department of Labor's (DOL) Employment and Training Administration (ETA), provides temporary unemployment benefits to eligible workers who have been unemployed through no fault of their own.

Since 1988, ETA has assessed the UI program's susceptibility to improper payments and estimated the annual amounts of improper payments through the Benefit Accuracy Measurement (BAM) program.¹ In FY 2004, the UI annual error rate was estimated to be 9.70 percent or roughly \$3.6 billion in total overpayments.² Several corrective action plans were developed to target the various root causes of overpayments and specifically targeted Benefit Year Earnings (BYE) violations because these are the largest cause of UI overpayments.³ BYE violations are defined as individuals who return to work and continue to claim and receive UI benefits. In FY 2004, the BYE violations, as detected through BAM, represented 27.2 percent of the \$3.6 billion in total improper payments for the UI program.⁴

The Office of the Chief Financial Officer (OCFO) and ETA initiated the Unemployment Insurance National Directory of New Hires pilot (pilot) to determine how a cross-match between the National Directory of New Hires (NDNH) database and state UI claimant data would help identify and reduce BYE violations. NDNH is a nationally consolidated database that contains employment and UI data on the nation's workforce and is maintained by the Department of Health and Human Services (HHS). Granting states access to NDNH should make available the full universe of newly employed persons, since NDNH consolidates all new hire information from every source. More specifically, NDNH not only includes all the new hire information from every State Directory of New Hires (SDNH) database, but also all Federal and military new hires, and the new hires from large multi-state employers (e.g., Wal-Mart) that report their new hires data to a single state.

The pilot involved the cooperation of HHS and the States of Texas, Utah, and Virginia. These three states volunteered for participation in the pilot and ultimately represented an adequate cross

¹ BAM is a quality control program designed to determine erroneous UI benefit payments by selecting weekly random samples of UI payments and denied claims. ETA investigates the samples to determine whether the claimant was properly paid or properly denied eligibility.

² DOL FY 2004 PAR, Improper Payments Information Act Reporting Details;
http://www.dol.gov/_sec/media/reports/annual2004/IPIA.htm

³ A BYE violation is just one of the several root causes of overpayments as identified through BAM. Other reasons for overpayments include separation issues where an employee was fired and still allowed to collect UI claims, "able and available" to work issues where an employee was paid benefits even though they were later determined to be able to work, and various other eligibility violation issues.

⁴ UI Benefit Accuracy Measurement Report for CY 2004,
<http://workforcesecurity.doleta.gov/unemploy/bam/2004/bam-cy2004.asp>

section of the nationwide UI program in terms of relative volume of claims and the level of cross-state employment activity.⁵

This pilot was executed through a cross-matching of the states' UI claimant data with the comparable data contained in NDNH. This cross-match involved comparing UI benefits data (defined as UI claimant data) sent to HHS by the pilot states to the W-4 new hire reports in NDNH database. This was done to detect individuals drawing UI benefits after returning to work. The results of this cross-match are considered to be BYE violations because the UI claimants were drawing UI benefits while earning wages.

Each pilot state was responsible for compiling and submitting their fourth quarter 2004 UI claimant data to HHS so it could cross-match the UI claimant information with the information contained in NDNH. A positive match was made when the same Social Security number was found on both sources of data for the cross match. A positive SDNH match represents those matches that would normally have been identified using the individual states' SDNH database (e.g., Texas UI claim matched with a W-4 new hire report from Texas). A NDNH match represents those matches that were discovered as a result of using the consolidated national data from NDNH (e.g., Texas UI claim matched with a W-4 new hire report from California). These matches would not normally be identified using individual states' SDNH databases.

The results for this cross-match were very encouraging for all pilot states. The potential benefit through use of NDNH is defined as the percentage increase in the additional BYE overpayments detected using NDNH as compared to the BYE overpayments that were detected through SDNH at that same time period. The pilot results indicated the following additional overpayments detected for fourth quarter 2004 claims using NDNH database compared to just using SDNH and the value of the additional overpayments that were identified.

TABLE 1

Additional Overpayments Identified in Fourth Quarter 2004 Using NDNH Compared to SDNH⁶		
	Additional Overpayments Identified	Value of Additional Overpayments Identified
Texas	114%	\$1.6 million
Utah	41%	\$0.4 million
Virginia	73%	\$0.8 million

There appears to be three main reasons why the additional overpayments were identified using NDNH. First, NDNH allows the states to identify claimants where UI claims are made in one state while wages are earned in a different state. Second, NDNH allows the states to identify

⁵These three states were the states that ultimately volunteered from a group of six pilot workgroup states. The other three states, which were part of the workgroup and did not participate, were Connecticut, Washington, and Florida.

⁶ The additional overpayments in Table 1 are for the identified 4th Quarter claims using the 2004 average BYE overpayment amount of \$482 per claim.

possible overpayments where a claimant is claiming benefits in one state while working in the same state for a multi-state employer that reports its W-4 data to a different state. Third, Federal and military employers are required to report only to NDNH database. This information is not available at the state level.

The pilot clearly demonstrated that the NDNH can be of significant benefit to states in identifying and reducing improper UI payments. The pilot also allowed states to identify possible changes for improving their own benefit payment quality control processes. For example, as part of the pilot workgroup meetings, the states and DOL discussed the possibility of making the submission of the Date of Hire and State of Hire fields in the W-4 reports mandatory; these fields are currently optional. This requirement would reduce the amount of potential overpayments that the states would have to sample and investigate for validity and would place them in a better position to determine sooner which claims represent potential overpayments.

As a result of these positive experiences and the overall success of the pilot, many states have expressed interest in implementing the NDNH cross-match. ETA is already moving ahead with the implementation of the NDNH cross-match with five states (Utah, Virginia, Texas, Connecticut, and Washington). An additional 24 states have expressed interest in the use of NDNH by end of FY 2006. The ultimate goal is for all states to use NDNH to detect and reduce the occurrence of BYE overpayments.

II. Purpose

The purpose of this pilot was to determine the extent to which a cross-match between the NDNH database and state UI data would help identify and reduce UI benefit overpayments. NDNH is a nationally consolidated database that contains employment and UI data on the nation's workforce and is maintained by the HHS. Granting states access to NDNH should make available the full universe of newly employed persons, since NDNH consolidates all new hire information from every source. More specifically, NDNH not only includes all the new hire information from every State Directory of New Hires (SDNH) database⁷, but also all Federal and military new hires and the new hires from large multi-state employers that report their new hires to a single state.

III. Methodology

The Department's OCFO and ETA initiated a pilot study to determine the added value of allowing states access to NDNH when administering their UI programs.⁸ The pilot involved the cooperation of HHS and the agencies for the States of Texas, Utah, and Virginia. These three states volunteered for participation in the pilot and ultimately represented an adequate cross section of the nationwide UI program in terms of relative volume of claims and the level of cross-state employment activity.⁹ HHS' involvement in the pilot was necessary as it maintains the NDNH database and uses it for various cross-agency data sharing agreements (e.g., SSA, IRS, Education, etc.). HHS was responsible for compiling the UI claimant information submitted electronically by each pilot state and matching it against the W-4 new hire report and UI claimant data contained in NDNH.

The focus of this pilot is on BYE violations as they are the leading cause of UI overpayments as identified by the BAM program.¹⁰ BYE violations are defined as individuals who return to work and continue to claim and receive UI benefits. Granting states access to the data contained in NDNH should help to curtail these and other violations.

The pilot states were responsible for sending the UI claimant data to HHS. The fourth quarter 2004 UI claimant data from the pilot states was matched by HHS to the fourth quarter 2004 W-4 data in NDNH. The purpose of this cross-match was to identify potential instances where individuals continued to claim and receive UI benefits after returning to work and earning income in another state. The total number and dollar amounts for these matches represented the additional UI overpayment detection associated with use of NDNH, compared with SDNH.

⁷ SDNH consolidates the employment and UI information for each state's workforce. It does not include new hire information on Federal and military new hires, new hires in other states, or new hires working for multi-state employers that are reported to other states.

⁸ OCFO provided guidance to the pilot states on sampling. The pilot states were responsible for conducting investigations and providing the results to ETA and OCFO. OCFO and ETA reviewed and analyzed the data from the pilot states.

⁹ These three states were the states that ultimately volunteered from a group of six pilot workgroup states. The other three states, which were part of the workgroup and did not participate, were Connecticut, Washington, and Florida.

¹⁰ BAM is a quality control program designed to determine erroneous UI benefit payments by selecting weekly random samples of UI payments and denied claims that they investigate further to determine whether the claimant was properly paid or properly denied eligibility.

HHS determined whether the match was a SDNH match or NDNH match. A positive SDNH match represents those matches that would normally have been identified using the individual states' SDNH databases (e.g., Texas UI claim matched with a W-4 new hire report from Texas). A NDNH match represents those matches that were discovered as a result of using the consolidated national data from NDNH (e.g., Texas UI claim matched with a W-4 new hire report from California). These matches would not normally be identified using individual states' SDNH databases. As part of the output, HHS provided a data output file for all unique matches to the pilot states with the associated NDNH or SDNH identifiers.

The SDNH and NDNH matches that were identified by HHS were then filtered to remove any invalid matches (e.g., a new hire match where the time period of employment did not overlap the time period of UI benefit payments). This post-filtering population represented the potential of all BYE overpayments. As the final step in the process, these potential BYE overpayments were sampled and investigated by the states to determine the valid number and amount of overpayments.

Although the results of the HHS cross-match were sampled and subsequently investigated, these results may not be statistically valid due to the variations in the standard operating procedures (SOP) across the pilot states. In addition to the impact of the employer non-response bias, where a non-response from an employer was considered to be a valid payment, there were several factors that may have impacted the results presented in this report. Further detail on some of these factors and a more detailed explanation of the methods used to analyze the data from the cross-match is provided in *Appendix A: Pilot Methodology*.

IV. Summary of Results

As part of the sampling and investigation process, the pilot states determined the number and total dollar amount of valid overpayments identified through the cross match. These results were then extrapolated using the sample size and population size of SDNH and NDNH matches in each state. However, because the identified overpayments represent multiple weeks of overpaid claims and are larger than what states should expect when implementing the cross-match on a regular basis, a proxy was used to calculate a more accurate extrapolation of the total dollar amount of overpayments.

OCFO used the average BYE overpayment amount per claimant for calendar year (CY) 2004, as identified through the states activities reported on the ETA 227 Report,¹¹ to estimate and extrapolate the amount of overpayment. This proxy represents what a typical BYE overpayment would be when overpayments are identified on a timely basis according to historical information from the ETA 227 Report.

The final results for the cross-match after extrapolating the sample results and using the average BYE overpayment amount are illustrated in the Table 2.

¹¹ ETA 227 UI Benefit Overpayments Report provides a summary of accounts receivables (overpayments) and recoveries. It is generated on a quarterly basis and is based on the data provided by the states.

TABLE 2

Fourth Quarter 2004 Extrapolation Results Using Average BYE Overpayment Amount		
Texas	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	3,005	\$1,448,540
NDNH Matches	3,415	\$1,645,837
NDNH as Percent of SDNH Matches	114%	
NDNH as Percent of Total Matches	53%	
Utah	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	1,821	\$877,905
NDNH Matches	747	\$359,837
NDNH as Percent of SDNH Matches	41%	
NDNH as Percent of Total Matches	29%	
Virginia	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	2,350	\$1,132,700
NDNH Matches	1,705	\$821,937
NDNH as Percent of SDNH Matches	73%	
NDNH as Percent of Total Matches	42%	

In the case of Texas, extrapolating the number of valid overpayments identified in the sample investigations resulted in 3,005 SDNH matches and 3,415 NDNH matches. These NDNH matches were never identified by Texas prior to the pilot because they did not have access to the national data contained in the NDNH database at that time. The total dollar amount of overpayment was derived by multiplying the number of overpayments by the average BYE overpayment amount of \$482.

The potential benefit of using NDNH is defined as the percentage increase in the BYE overpayment detection and is quantified as the total number of newly identified matches (NDNH) divided by the total number of previously identified SDNH matches. This benefit or lift of using NDNH is 114 percent for Texas, 73 percent for Virginia, and 41 percent for Utah. The biggest reason that this benefit is higher for Texas and Virginia is that these pilot states have higher claim volume with a lot more cross-border activity and multi-state employers. The final results, after extrapolation and using the average BYE overpayment amount, were approximately \$1.6 million for Texas, \$0.4 million for Utah, and \$0.8 million for Virginia. The results represent the overpayment amount each pilot state would have identified through the use of this match.

There appears to be three main reasons why the additional overpayments were identified using NDNH. First, NDNH allows the states to identify claimants where UI claims are made in one state while wages are earned in a different state. Second, NDNH allows the states to identify

possible overpayments where a claimant is claiming benefits in one state while working in the same state for a multi-state employer that reports its W-4 data to a different state. Third, Federal and military employers are required to report only to NDNH database. This information is not available at the state level.

Detailed step-by-step results are provided in *Appendix B: Detailed Results*.

V. Conclusion

The results of the pilot clearly indicate that use of the NDNH cross-match by states should help reduce the levels of BYE violations through earlier detection. The pilot focused on BYE overpayments because they comprise the majority of the UI program overpayments and reducing these overpayments should have the greatest impact on reducing the overall improper payment rate in the UI program.

As a result of these positive experiences and the overall success of the pilot, many states have expressed interest in implementing the NDNH cross-match. ETA is moving ahead with the implementation of the NDNH cross-match with five states (Utah, Virginia, Texas, Connecticut, and Washington). An additional 24 states have expressed interest in the use of NDNH by the end of FY 2006. The remaining states, the District of Columbia, Puerto Rico, and the Virgin Islands are interested in using NDNH in FY 2007 or later.

Appendix A: Pilot Methodology

The detailed descriptions of the activities performed for the pilot are described in this section. The detailed descriptions for the processes for the cross-match are provided in the following subsections:

- Matching Process
- Filtering Process
- Sampling Process
- Investigation Process

The 4th quarter 2004 UI claimant data from the pilot state was matched to the 4th quarter 2004 W-4 data in NDNH. The purpose of this cross-match was to identify instances where individuals were claiming UI benefits in one state while having W-4 new hire reports in other states. The total number and dollar amounts for these matches represent the additional UI overpayment detections associated with use of NDNH.

Matching Process

1. Each of the three pilot states compiled their fourth quarter 2004 UI claimant information and submitted it to HHS via IV-D Connect Direct (data transfer network). The UI claimant information consisted of a subset (SSN, name, record identifier) of their entire claimant data.
2. HHS compared fourth quarter 2004 UI claimant data vs. fourth quarter 2004 W-4 data in NDNH by matching SSNs of each claimant against the W-4 information in the NDNH database to determine the initial number of UI claimants with W-4 reports from other states. A match of the claimant's SSN (and name) with the W-4 data indicated the possibility that the individual claimed UI benefits after returning to work. This did not necessarily mean an overpayment occurred since the claim might have been made before the new hire report was filed. The match only cross-checked the names and SSNs; it did not look at the dates. Verification of employment data required follow-ups by the state agency.
3. As part of the cross match, HHS used the state or submitting agency code to identify whether the match was a NDNH or SDNH match.
 - a. Example of SDNH match: The Texas UI claimant SSN matched the W-4 data in NDNH and the submitting state or agency code was marked as Texas.
 - b. Example of NDNH match: The Texas UI claimant SSN matched the W-4 data in NDNH and the submitting state or agency code was marked as a state other than Texas.

4. HHS reviewed the results of the matches and filtered the data to remove duplicate matches. The total non-duplicate match numbers were provided for both the NDNH and SDNH matches.¹²
5. HHS then sent the matches back to the pilot state's mainframe servers for the states to start evaluating and investigating the matches for validity.

Filtering Process

1. The pilot states split the file received from HHS into two separate files - one for NDNH matches, the other for SDNH matches, using the "submitting state or agency code" field that was described in the previous section. The steps below were performed for each of the two match files.
2. The files were sorted by those with Date of Hire (DOH) vs. No Date of Hire. Approximately, 70 percent of the records had a DOH field. The purpose of using the DOH field as a filter is to remove those records where the DOH is greater than the last Benefit Week Ending (BWE) claimed. For those records where this is the case, it is not a true overpayment since the claimant stopped claiming benefits before he/she was hired. In about 30 percent of the cases, there was no DOH.
3. For those records with a DOH field, a time lag was computed, where the *Time Lag* = W-4 Processing Date – DOH (the date the new hire report was received by HHS).
4. All records were sorted and arrayed in ascending order by the *Time Lag* field.
5. The pilot states used pivot tables and/or other Excel tools to determine the frequency of observations that fall in each *Time Lag*.
6. The cumulative frequency for the *Time Lag* was determined until it reached 90 percent (90 percent of the *Time Lags* were less than or equal to this number). This *Time Lag* was used in subsequent steps to determine a proxy DOH for those records without a DOH, based on the assumption that the remaining 30 percent or so records without the DOH field will have the same *Time Lag* distribution as those with the DOH field.
7. For those records without a DOH field, the *Proxy DOH* was calculated, where the *Proxy DOH* = W-4 Processing Date – *Time Lag*. (Now all records have a DOH field).
8. The pilot states then subtracted the DOH (Original or Proxy) from the W-4 Processing Date and sorted the results in descending order.

12 Normally, most multi-state employers report their W-4 data to one state. One benefit of using this NDNH cross-match is the identification of possible overpayments where a claimant is claiming benefits in one state while working for a multi-state employer that does not report its W-4 data in that state. However, there are instances where the local branch of a multi-state employer might also report W-4 data to the local state. This would lead to a duplication of matches as they would be identified both to the local state and to the state that the multi-state employer normally reports its W-4 data.

9. All records that had a result that was zero or greater than zero (DOH was after the last BWE claimed) were removed. The remaining records where the result was less than zero (DOH was before the last BWE claimed) represent the true matches after this filtering process. These matches were sampled and further investigated.

Sampling Process

Due to the number of matches described above and because of time and resource constraints, it was not feasible to review all matches to determine whether they were associated with an overpayment. Consequently, DOL applied a sampling process to estimate the approximate number of matches that are associated with actual UI overpayments.

The process of selecting an appropriate sample size always involves judgment. In general, larger sample sizes lead to more precise estimates. Even so, there tend to be “diminishing returns” as sample size increases. That is, a sample that is twice as large will not yield a result that it is twice as precise. Unfortunately, it is not possible to determine exactly how precise a statistical estimate will be until the sampling procedure is completed. For this reason, DOL reviewed a statistical table (see Table 3 below) as a way to guide the judgment of what sample size should be selected for each match population.

The Department chose sample sizes for each match that afforded the pilot states enough time to conduct the match investigations and lower the rate of non-response that tend to occur in these investigations. Non-response is defined as the failure of an employer or UI claimant to respond to the initial inquiries of the UI Claims Investigator. If the rate of non-response is high, it can be more difficult to interpret the results of the sampling procedure because non response introduces a potential bias, which contributes to the uncertainty of the estimate. A smaller and more manageable representative sample will allow the states more time to follow up with employers and claimants that did not respond to the initial inquiry.

1. The sample sizes were based on attribute sampling with a confidence level of 90 percent with two-sided precision and an expected error rate (overpayments) based on a projected overpayment rate of 30 percent.

TABLE 3

Overpayment Rates for Expected Overpayment Rate of 30 Percent		
Sample Size	Lower Bound	Upper Bound
50	19.1%	40.9%
100	22.4%	38.5%
150	23.8%	36.8%
200	24.6%	35.8%
300	25.6%	34.7%
500	26.6%	33.6%

Note: The confidence interval will be different if the percentage of overpayments is not 30 percent, but this table can still serve to illustrate the “diminishing returns” that can be expected when increasing the sample size.

2. A total sample size (n) of 300 was selected for both the SDNH and NDNH matches. Although the total sample size was 300, this is actually a stratified sample of 200 for NDNH matches and 100 for SDNH matches. A higher sample size was selected for NDNH since the expected overpayment rate is more of an unknown as it has never been investigated through this process. With an expected 30 percent overpayment rate, the lower bound and upper bounds would be 24.6 percent and 34.7 percent for NDNH matches and 22.4 percent and 38.5 percent for SDNH matches.
3. The skip factor (k) was calculated for each match by dividing the sample size (n) into the total population of post-filtering matches (N).
4. The pilot states, using the Excel spreadsheet for each of the two matches, inserted a column at the beginning of the sheet. In this column, they generated a random number multiple (RAND function output * 50000) and re-sorted the filtered file by this random number multiple. *This step is performed to remove any bias from the population through the use of random numbers.*
5. Starting at the top of the re-sorted file, the pilot states highlighted the first record, and every kth record (rounding, where necessary). The output of steps 5 and 6 was two files: 1st with 200 random samples for NDNH matches; 2nd with 100 random samples for SDNH matches.

After completing the match investigations, each pilot state prepared and sent two files for both NDNH and SDNH matches to UI for further analysis. These files contained UI claimant-specific discrete data such as demographic codes, occupation codes, NAICS codes, BAM error codes, and aggregate results (i.e., total over/underpayment amounts and numbers and total costs for investigation) for both SDNH and NDNH matches.

Investigation Process

The pilot states investigated the matched cases that were sampled. These investigations were conducted to determine the actual number of fraud and non-fraud overpayments for both the SDNH and NDNH matches out of the potential matches identified as a result of the HHS cross match.

Each state used their existing standard operating procedures (SOPs) to investigate the cases. The method by which the investigations were conducted varied slightly among the pilot states although all three pilot states executed the same steps - contacting the claimant, contacting the employer, and adjudicating. The differences were mainly in the method of contact, mail vs. telephone vs. fax, and in the order in which the states executed the steps. Some states started by contacting the employer first via mail, fax, or phone, while other states contacted directly the claimant first via mail or phone. In all cases, the final step was the adjudication of results of the contact.

If the match involved an out-of-state employer, then the states usually submitted an official document to the employer requesting information on the claimant given the level of unfamiliarity on the part of the state to those employers. Once the employers or claimants were contacted, they were then asked to verify or provide information on the claimant's date of hire, amounts of benefits paid, earnings reported, and any other issues involved in the investigation. This information was then tabulated and sent off for adjudication within 24 hours of the investigation completion date.

During the adjudication process, the claimant was contacted to determine their reasons for continuing to draw benefit while earning wages. The adjudication unit then had 48 hours to issue a decision on whether or not the case was a fraud or non-fraud overpayment of UI benefits. Per the guidelines in the Computer Matching Act, the pilot states did not take any actions to set up and try to recover any of the overpayments identified through this pilot.

All three pilot states generally followed the steps in the approach detailed above. However, there were several factors that may have impacted the final results. These factors are noteworthy, but are not considered to have materially impacted the final results.

Protocol Exceptions

- After the receipt of the cross-match results from HHS, Texas conducted filtering which was over and above that prescribed for the pilot, but is part of Texas' normal process. In addition to the filtering prescribed for the pilot, Texas filtered out those records that had a prior new hire or fraud audit. The premise behind this additional filtering is that these prior new hire or fraud audits were identified as part of the agency's normal benefit payment control process; thus they should not be attributable to the use of NDNH.
- Utah expanded the filtering to eliminate those matches that had a disqualification that prevented the claimant from ever receiving benefit payments. The premise behind this filtering is that these are not overpayments because the benefits were never actually paid.
- Virginia's sample size varied. Virginia used a sample size of 190 (instead of 200) for NDNH matches and a sample size of 111 (instead of 100) for SDNH matches.
- Utah includes a 100 percent penalty as part of its fraud overpayments. As part of this pilot, we adjusted the data by dividing the Utah fraud overpayment amounts by two for both SDNH and NDNH matches.
- Each state used a modified adjudication process in order to meet the expedited deadlines for this pilot for determining if the overpayment was fraud or non-fraud. Texas determined fraud overpayments by reviewing the number of weeks overpaid, the overpayment amount, and input from the claimants themselves. For fraud to be determined there has to be six or more weeks of under or unreported earnings or the claimant has to make a statement such as, "I knew it was wrong, I just needed to pay my baby sitter." Non-fraud overpayments involve overpaid claims that do not exceed a certain dollar threshold and are incurred for six weeks

or less. However, they may be determined fraudulent if the claimant admits to fraud or had a prior fraud overpayment case. Utah conducted a “pseudo” adjudication process where they only looked at the evidence presented from the claimant (e.g., wages) and what was reported by the employer after contacting them. The investigator then made a decision as to whether the overpayment was fraud versus non-fraud based on this information. They determined that small dollar overpayments that were incurred for one week were considered non-fraud overpayments. Overpayments spanning multiple weeks were considered fraudulent overpayments.

Although sampling was done to investigate the matches identified through the HHS cross match, the results may not be statistically valid due to the aforementioned variations in the SOPs across the pilot states.

Appendix B: Detailed Results

The results of the pilot are organized in the following categories:

1. HHS Cross-match and Filtering Results
2. Sampling and Investigation Results
3. Extrapolation Results
4. Effect of Multi-State Employers

1. HHS Cross-match and Filtering Results

The results of the HHS cross-match between the fourth quarter 2004 UI claimant information and the NDNH W-4 data are presented below in Table 4. Further explanation about the fields in the table is provided below:

- Total input records: total number of UI claims sent from the pilot states to HHS.
- Total unique SSN matches for the State Directory of New Hires (SDNH): total non-redundant matches (e.g. matches that did not have duplicate SSNs) resulting from the cross-match between the UI claimant data and NDNH.
- Total unique SSN matches (NDNH): total non-redundant matches resulting from the cross-match between the UI claimant data and NDNH.
- Post Filtering population (SDNH): total actual matches resulting from the post-filtering process for the SDNH matches. These represent the actual number of matches that may indicate possible overpayment. These matches will be the basis of sampling and investigation for the pilot.
- Post Filtering population (NDNH): total actual matches resulting from the post-filtering process for the NDNH matches. These represent the actual number of matches that may indicate possible overpayment. These matches will be the basis of sampling and investigation for the pilot.

TABLE 4

Fourth Quarter 2004 HHS Cross-match and Post Filtering Results		
Texas - Total Input Records 295,951	Total Unique SSN Matches	Post Filtering Population
SDNH Matches	62,174	10,363
NDNH Matches	42,129	10,840
NDNH as Percent of SDNH Matches		105%
NDNH as Percent of Total Matches		51%
Utah - Total Input Records 32,591		
	Total Unique SSN Matches	Post Filtering Population
SDNH Matches	10,796	8,279
NDNH Matches	4,399	3,318
NDNH as Percent of SDNH Matches		40%
NDNH as Percent of Total Matches		29%
Virginia - Total Input Records 67,184		
	Total Unique SSN Matches	Post Filtering Population
SDNH Matches	11,192	8,695
NDNH Matches	7,963	6,000
NDNH as Percent of SDNH Matches		69%
NDNH as Percent of Total Matches		41%

The results were very encouraging for all pilot states as they each experienced an added benefit of using the NDNH cross match. One way of defining the benefit of using the NDNH cross-match would be resultant new BYE overpayments identified through this cross match. In this context, the benefit can be quantified as the total number of newly identified matches (NDNH) divided by the total number of SDNH matches. This benefit or lift of using NDNH is 105 percent for Texas, 69 percent for Virginia, and 40 percent for Utah. The biggest reason that this benefit is higher for Texas and Virginia is that these pilot states have higher claim volume with a lot more cross-border activity and multi-state employers. These newly identified matches represented new leads that the pilot states then sampled and investigated to determine the number and amounts of valid overpayments. More specifically, between an additional 3,318 (Utah) and 10,840 (Texas) potential overpayments would be sampled and investigated.

2. Sampling and Investigation Results

The post-sampling and investigation results are provided below in Table 5. Further explanation about the fields in the table is provided below:

- Total number of overpayment (SNDH): these are the total number of fraud and non-fraud overpayments that were identified as part of the investigation process for the sample size for this match for SDNH matches.

- Total number of overpayment (NDNH): these are the total number of fraud and non-fraud overpayments that were identified as part of the investigation process for the sample size for this match for NDNH matches.
- Total dollar amount of overpayment (SDNH): these are the total amount of fraud and non-fraud overpayments that were identified as part of the investigation process for the sample size for this match for SDNH matches.
- Total dollar amount of overpayment (NDNH): these are the total amount of fraud and non-fraud overpayments that were identified as part of the investigation process for the sample size for this match for NDNH matches.

TABLE 5

Fourth Quarter 2004 Sampling and Investigation Results		
Texas	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	29	\$76,188
NDNH Matches	63	\$178,226
Utah		
	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	22	\$14,318
NDNH Matches	45	\$52,007
Virginia		
	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	30	\$28,392
NDNH Matches	54	\$43,626

The investigation of the sampled NDNH matches (or new leads) resulted in the identification of significant amounts of overpayments for all pilot states. Texas had the highest amount of overpayments identified from the NDNH cross-match, with \$178,226; Utah had \$52,007; and Virginia had \$43,626. Though the corresponding SDNH overpayments seem much smaller, it should be noted that the sample sizes were different for the investigations for SDNH and NDNH matches.

DOH field was used as the primary filter to identify potential overpayment matches. The premise was that if the DOH is before the claim date, that claim is potentially an overpayment. Since the DOH field was not provided for all claims, the pilot states determined a proxy DOH for these claims through the process described in Approach section above.

Although all claims that had a DOH before the UI claim date have the potential to be overpayments, the investigations revealed only a sub-set of the sample to be overpayments. As

presented above in Table 5, not every investigation yielded a positive match; this is better illustrated below in Table 6:

TABLE 6

Fourth Quarter 2004 Percentage of NDNH Overpayments			
	Total NDNH Overpayments	Total Number Investigated	Percentage Overpayments
Texas	63	200	32%
Utah	45	200	23%
Virginia	54	190	28%
Fourth Quarter 2004 Percentage of SDNH Overpayments			
	Total SDNH Overpayments	Total Number Investigated	Percentage Overpayments
Texas	29	100	29%
Utah	22	100	22%
Virginia	30	111	27%

The results provided in the table above illustrate the true number of overpayments that may be identified through the use of this NDNH cross match. Possible reasons for the substantial amount of non-overpayments are listed below:

- The proxy DOH field was not entirely accurate when the actual hire date was later than the paid claims end date.
- The employer reported no wages earned during that period even if the person was hired and was identified through the cross match.
- Cases where employers did not respond to investigation inquiries were considered non overpayments by default.

3. *Extrapolation Results*

As the final step in the analysis of the data from the sampling and investigation, the results were extrapolated to the entire population of matches based on the final post-filtering number of matches and the sample size.

The overpayment amounts identified during the investigations cannot be used as a benchmark because they do not account for the fact that when the cross-match is in production the states will identify many of these overpayments earlier. As a result, the amounts of overpayments will be smaller. A better proxy for estimating and extrapolating the amount of overpayment may be the average BYE overpayment amount for 2004 as identified through the states activities reported on the ETA 227 Report.¹³ This proxy represents what a typical BYE overpayment would be when

¹³ ETA 227 UI Benefit Overpayments Report provides a summary of accounts receivables (overpayments) and recoveries. It is generated on a quarterly basis and is based on the data provided by the states.

overpayments are identified in a timely basis, according to historical information from the ETA 227 Report.

The extrapolation results are provided below in Table 7. The Further explanation about the fields in Table 7 is provided below:

- Total number of overpayment (SNDH): these are the total number of extrapolated fraud and non-fraud overpayments that were identified as part of the investigation process for the sample size for this match for SDNH matches.
- Total number of overpayment (NDNH): these are the total number of extrapolated fraud and non-fraud overpayments that were identified as part of the investigation process for the sample size for this match for NDNH matches.
- Total dollar amount of overpayment (SNDH): these are the total amount of extrapolated fraud and non-fraud overpayments, using the average BYE overpayment amount, that were identified as part of the investigation process for the sample size for this match for SDNH matches.
- Total dollar amount of overpayment (NDNH): these are the total amount of extrapolated fraud and non-fraud overpayments, using the average BYE overpayment amount, that were identified as part of the investigation process for the sample size for this match for NDNH matches.

TABLE 7

Fourth Quarter 2004 Extrapolation Results Using Average BYE Overpayment Amount		
Texas	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	3,005	\$1,448,540
NDNH Matches	3,415	\$1,645,837
NDNH as Percent of SDNH Matches	114%	
NDNH as Percent of Total Matches	53%	
Utah	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	1,821	\$877,905
NDNH Matches	747	\$359,837
NDNH as Percent of SDNH Matches	41%	
NDNH as Percent of Total Matches	29%	

TABLE 7 (continued)

Fourth Quarter 2004 Extrapolation Results Using Average BYE Overpayment Amount		
Virginia	Total Number of Overpayments	Total Dollar Amount of Overpayments
SDNH Matches	2,350	\$1,132,700
NDNH Matches	1,705	\$821,937
NDNH as Percent of SDNH Matches	73%	
NDNH as Percent of Total Matches	42%	

In the case of Texas, extrapolating the number of valid overpayments identified in the sample investigations resulted in 3,005 SDNH matches and 3,415 NDNH matches. These NDNH matches were never identified by Texas prior to the pilot because they did not have access to the national data contained in the NDNH database at that time. The total dollar amount of overpayment was derived by multiplying the number of overpayments by the average BYE overpayment amount of \$482, in the case of Texas this resulted in \$1,448,540 in SDNH overpayments and \$1,645,837 in NDNH overpayments.

The final results after extrapolation of the sample results and the use of the average BYE overpayment amount are illustrated in table above. The benefit is quantified as the total number of newly identified matches (NDNH) divided by the total number of SDNH matches. This benefit or lift of using NDNH is 114 percent for Texas, 73 percent for Virginia, and 41 percent for Utah. The final results, after extrapolation and using the average BYE overpayment amount, were approximately \$1.6 million for Texas, \$0.4 million for Utah, and \$0.8 million for Virginia. These results represent the overpayment amount each pilot state would have identified through the use of this match.

4. Effect of Multi-State Employers

OCFO and ETA were interested in the impact of multi-state employers on overpayments identified through NDNH since one major expected benefit of this cross-match is to detect those overpayments associated with multi-state employers. There were no specific data elements that would conclusively provide this information. The pilot states reviewed their NDNH matches and provided an estimate of the number of overpayments related to multi-state employers based on their knowledge of the state and its major employers. The overpayments identified in the table below are based on the investigation results on the sample size of 200 for NDNH cross-match, except for Virginia which had a sample size of 190. These results are provided below in Table 8:

TABLE 8

Fourth Quarter 2004 NDNH Overpayments for Multi-state Employers			
	Total Overpayments Identified	Overpayments related to Multi-state Employers	Percentage of Multi-state Employer Overpayments
Texas	63	36	57%
Utah	45	16	36%
Virginia	54	Not available	Not available

Texas had a much higher percentage of multi-state employer related overpayments because of the associated higher volume of claims with more multi-state employers than Utah. The remaining overpayments identified through the NDNH cross-match are due to claimants claiming benefits in one state (e.g., Texas) while earning wages and having W-4 wage reports in another state (e.g., Oklahoma).

Appendix C: List of Acronyms

Acronym	Description
BAM	Benefit Accuracy Measurement program
BWE	Benefit Week Ending
BYE	Benefit Year Earnings
DOH	Date of hire
DOL	The Department of Labor
ETA	Employment and Training Administration
HHS	The Department of Health and Human Services
IPIA	Improper Payment Information Act of 2002
NDNH	National Directory of New Hires
OCFO	Office of the Chief Financial Officer
SDNH	State Directory of New Hires
SOP	Standard operating procedures
UI	Unemployment Insurance program