

**MEMORANDUM**

 **Mathematica Policy Research**

 1100 1st Street, NE, 12th Floor

Washington, DC 20002-4221

 Telephone (202) 484-9220

 Fax (202) 863-1763

 www.mathematica-mpr.com

**TO:** Elaine Gilby and Paul O’Leary

**FROM:** Eric Grau **DATE:** 5/11/2015

 NBS - 21

**SUBJECT**: Review of Potential Bias in Earnings in the Prior NBS—Round 3 and Round 4

This memorandum summarizes the results of our investigation into the potential bias in earnings (as reported in SSA’s administrative records) between the released and unreleased sample members for rounds 3 and 4 of the prior National Beneficiary Survey (NBS), with a focus on round 4 where the greatest differences are apparent. It summarizes differences in average (mean and median) earnings between released and unreleased sample members noted above, particularly in round 4, and assesses differences in other related variables. In summary, we could not find a completely satisfactory explanation for the differences in earnings between released and unreleased cases, especially in round 4. Some differences could be due to differing release rates by Primary Sampling Unit (PSU), which is discussed in Section II.C, though we show that this does not come close to completely explaining the differences in earnings.

The need for this analysis was identified by an earlier SSA analysis of mean and median earnings using data from round 3 and round 4. That analysis showed that, in round 4, the mean and median earnings of unreleased sample members in the augmented sample were significantly larger than those of the released sample members. The mean and median earnings among unreleased sample members in round 3 of the prior NBS were also larger than those of the released sample members, though most of these differences were not statistically significant. These differences were perplexing, as Mathematica controlled for variables assumed to be correlated with earnings when sample cases were released in both rounds. Mean earnings among sample members with completed interviews, using nonresponse-adjusted weights, were also smaller than for the population as a whole. Any differences in earnings between these groups should have been random and not statistically significant.

I. Description of Analysis

For this review, we ran three sets of analyses. Because we do not have access to IRS earnings, we made requests to Emily Roessel at SSA, who conducted the computer processing and analyses involving earnings for us. The three analyses included: 1) a comparison of the average earnings of cases in the released sample to those in the unreleased sample; 2) a comparison of other variables in the released sample to those in the unreleased sample; and 3) an evaluation of differences in average earnings by release rate (percentage of the sample released). The latter, (#3) was accomplished in two ways: 3.1) comparing average earnings in the first release to those in the unreleased sample, and 3.2) evaluating the release rate by PSU. Each of these analyses used data from the augmented sample, where the results are weighted using augmented sampling weights.[[1]](#footnote-1) In the analyses involving earnings (#1 and #3), we viewed the data in three different ways:

1. the average earnings of all beneficiaries, including non-workers;
2. the average earnings of those working in the year closest to the creation of the frame (2009 for round 4, and 2005 for round 3), and

the average earnings of those working in the year defined by the earnings in that year.

Some analyses of the earnings did not include the third item mentioned above, as the conclusions drawn between the second and third items did not differ appreciably. The findings are summarized below.

II. Findings and summary

### A. Comparing mean earnings in the released sample to that of the unreleased sample

When comparing released and unreleased cases, it was necessary to create an estimate of the standard error of the difference. For the purposes of this test, the survey data analysis procedures in SAS were used because of its availability and ease of access and use at the SSA data processing center. The standard errors were estimated using PROC SURVEYMEANS, where clusters were identified as the selected PSUs. The standard error of the differences was estimated by taking the square root of the weighted average of the squared standard errors. Although not fully correct, this estimate is close to the correct standard error.[[2]](#footnote-2) In addition, although the distribution of the earnings of individuals was not close to normal, the analyses were conducted using means calculated from such large samples that the t-test is nonetheless valid.[[3]](#footnote-3) The tables that follow include all beneficiaries (Tables 1 and 4), those beneficiaries who were working in the year of available data closest to the frame creation (2005 for round 3 in Table 2 and 2009 for round 4 in Table 5), and those beneficiaries who were working in the year associated with the earnings variable in question (Tables 3 and 6). The tables include means and standard errors for earnings of all cases in the corresponding samples, plus the percentage working (that is, with positive earnings) during the year. Note that the means and standard errors in Tables 1 and 4 are much smaller than those in the other tables because these tables include all beneficiaries, and most beneficiaries have zero earnings (they do not work for pay). Tables 1 to 3 below use data from round 3. Significant t-values are shown in red.

Table 1. Weighted earnings of all round 3 beneficiaries, comparison of released and unreleased casesa

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released (n= 3,382) | Unreleased (n = 11,618) |  |
| Year | Mean | Standarderror | PercentWorking | Mean | StandardError | Percent Working |  t |
| 2004 | 1,295.78 | 84.86 | 21.9% | 1,483.15 | 94.82 | 18.4% | 2.02 |
| 2005 | 1,063.93 | 79.78 | 18.5% | 1,104.98 | 69.78 | 16.3% | 0.57 |
| 2006 | 1,270.51 | 93.04 | 18.7% | 1,236.03 | 89.01 | 16.2% | -0.38 |
| 2007 | 1,510.63 | 113.06 | 18.4% | 1,421.35 | 101.30 | 16.5% | -0.86 |

a15,000 cases were in the augmented sample, of which 3,382 cases were released and 11,618 cases were not released in round 3.

Table 2. Weighted earnings of round 3 beneficiaries working in 2005, comparison of released and unreleased casesa

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released, Working in 2005 (n= 819) | Unreleased, Working in 2005 (n = 2,684) |  |
| Year | Mean | Standard error | PercentWorking | Mean | Standarderror | PercentWorking |   t |
| 2004 | 4,939.33 | 360.21 | 79.2% | 6,652.79 | 436.59 | 78.7% | 4.41 |
| 2005 | 5,748.56 | 421.52 | 100% | 6,362.44 | 347.65 | 100% | 1.54 |
| 2006 | 5,920.91 | 420.15 | 78.4% | 6,227.23 | 408.35 | 77.5% | 0.74 |
| 2007 | 6,281.47 | 501.89 | 72.0% | 6,852.43 | 540.44 | 70.2% | 1.11 |

a819 of 3,382 released cases and 2,684 of 11,618 unreleased cases were working in 2005.

Table 3. Weighted earnings of round 3 beneficiaries working in the year of earnings, comparison of released and unreleased casesa

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released  | Unreleased |  |
| Year | Mean | Standarderror | Median | n | Mean | Standarderror | Median | n |  t |
| 2004 | 5919.25 | 357.92 | 2900.94 | 883 | 7406.55 | 397.52 | 3251.00 | 2863 | 3.93 |
| 2005 | 5748.56 | 421.52 | 2671.86 | 819 | 6362.44 | 230.88 | 2886.00 | 2684 | 1.59 |
| 2006 | 6811.92 | 469.23 | 3356.78 | 838 | 7237.47 | 276.39 | 3135.49 | 2733 | 0.91 |
| 2007 | 8209.07 | 605.39 | 4802.65 | 836 | 8289.99 | 325.71 | 3659.15 | 2738 | 0.14 |

aAmong the 3,382 released cases, the number of working beneficiaries varied by year (from 819 to 883). Among the 11,618 unreleased cases, the number of working beneficiaries varied by year (from 2,684 to 2,863).

In all three tables, the earnings among unreleased cases appear to be significantly greater than those of released cases in 2004, but not in the year when the population of beneficiaries is defined (2005), or in following years. Though standard errors for the medians are not available, the median earnings among those with positive earnings do not appear to differ substantively between the released and unreleased cases in 2004 or 2005.[[4]](#footnote-4)

Looking at round 4 data, we see a stronger pattern of differences. In all three tables, earnings among unreleased cases are significantly greater than those of the released cases. Again, significant results are shown in red.

Table 4. Weighted earnings of all round 4 beneficiaries, comparison of released and unreleased casesa

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released (n=3,683) | Unreleased (n= 8,316) |  |
| Year | Mean | Standarderror | %Working | Mean | StandardError | %Working |  t |
| 2008 | 1285.11 | 113.64 | 19.6% | 1719.54 | 127.13 | 21.2% | 3.52 |
| 2009 | 919.59 | 86.34 | 14.8% | 1181.12 | 75.51 | 16.5% | 3.31 |
| 2010 | 907.26 | 91.06 | 14.0% | 1125.29 | 77.47 | 15.5% | 2.66 |
| 2011 | 924.54 | 85.78 | 13.4% | 1189.04 | 88.61 | 14.7% | 3.01 |

a11,999 cases were in the augmented sample, of which 3,683 cases were released and 8,316 cases were not released in round 4.

Table 5. Weighted earnings of round 4 beneficiaries working in 2009, comparison of released and unreleased cases.a

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released, Working in 2009 (n=774) | Unreleased, Working in 2009 (n=1,778) |  |
| Year | Mean | Standarderror | %Working | Mean | Standarderror | %Working |   t |
| 2008 | 6125.18 | 553.68 | 83.8% | 8386.33 | 596.79 | 86.8% | 3.87 |
| 2009 | 6222.71 | 471.29 | 100% | 7139.30 | 376.72 | 100% | 2.25 |
| 2010 | 5695.02 | 506.59 | 77.4% | 6192.86 | 412.90 | 75.1% | 1.12 |
| 2011 | 5060.52 | 485.39 | 67.6% | 6045.55 | 464.25 | 66.3% | 2.09 |

a774 of 3,683 released cases and 1,778 of 8,316 unreleased cases were working in 2009.

Table 6. Weighted earnings of round 4 beneficiaries working in the year of earnings, comparison of released and unreleased cases.a

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released | Unreleased |  |
| Year | Mean | Standarderror | Median | n | Mean | Standarderror | Median | n |  t |
| 2008 | 6570.75 | 466.32 | 3440.79 | 927 | 8108.45 | 466.28 | 4334.67 | 2215 | 3.30 |
| 2009 | 6222.71 | 336.94 | 3069.65 | 774 | 7139.30 | 365.12 | 3919.39 | 1778 | 2.51 |
| 2010 | 6471.08 | 527.86 | 3194.07 | 741 | 7285.19 | 434.22 | 3746.40 | 1714 | 1.87 |
| 2011 | 6917.92 | 558.37 | 3719.77 | 733 | 8086.14 | 505.39 | 4483.67 | 1669 | 2.24 |

aAmong the 3,683 released cases, the number of working beneficiaries varied by year (from 733 to 927). Among the 8,316 unreleased cases, the number of working beneficiaries varied by year (from 1,669 to 2,215).

B. Comparison of other variables in the released sample to that of the unreleased sample

In the prior analysis, we saw that average earnings in the released sample were significantly less than for the unreleased sample, particularly for cases in round 4. What is puzzling about this is that, when the augmented sampling weights were created, the weights were poststratified so that the total number of beneficiaries in each of the four age categories matched the population. Marginal totals for other variables were also tracked closely. In this analysis, we compare the marginal proportions for several administrative variables, including variables that were available at the time of sample selection (title, race, and ethnicity [Hispanic/Non-Hispanic]), and variables that were extracted more recently. We limit our attention to round 4, since that is where the greatest differences were observed. Table 7 looks at the variables that were readily available at the time of sampling. Standard errors used to calculate the t-statistics were calculated in the same manner as in the earnings tables, where the same caveats apply.

Table 7. Weighted comparison of age, title, ethnicity, race, gender, and broad disability categories among released and unreleased cases for augmented sample of round 4 beneficiariesa

| Variable | Frame | Augmented Sample | Released Sample (n=3,683) | Unreleased Sample (n=8,316) |  |
| --- | --- | --- | --- | --- | --- |
|  | Propor-tion | Proportion | Proportion | Standard error | Proportion | Standard error | t |
| **Age** |  |  |  |  |  |  |  |
| 18-29 | 10.72% | 10.77% | 10.93% | 0.27% | 10.70% | 0.21% | 0.97 |
| 30-39 | 10.87% | 10.86% | 11.06% | 0.29% | 10.77% | 0.17% | 1.32 |
| 40-49 | 20.85% | 20.88% | 20.98% | 0.44% | 20.84% | 0.27% | 0.44 |
| 50-FRA | 57.56% | 57.49% | 57.03% | 0.73% | 57.69% | 0.43% | -1.23 |
| **Title** |  |  |  |  |  |  |  |
| SSI only | 31.22% | 32.43% | 32.72% | 1.38% | 32.30% | 1.20% | 0.34 |
| SSDI only | 52.10% | 51.02% | 49.61% | 1.61% | 51.63% | 1.26% | -1.46 |
| Concurrent | 16.68% | 16.56% | 17.67% | 1.00% | 16.07% | 0.64% | 2.06 |
| **Race/Ethnicity** |  |  |  |  |  |  |  |
| White | 62.69% | 62.60% | 61.60% | 2.46% | 63.04% | 2.20% | -0.63 |
| Black | 21.12% | 20.57% | 21.62% | 2.11% | 20.11% | 2.06% | 0.73 |
| Hispanic | 4.33% | 4.68% | 5.16% | 0.96% | 4.47% | 0.88% | 0.76 |
| Other | 4.84% | 4.96% | 5.56% | 0.87% | 4.69% | 0.63% | 1.21 |
| Unknown | 7.02% | 7.19% | 6.06% | 0.59% | 7.69% | 0.38% | -3.61 |
| **Gender** |  |  |  |  |  |  |  |
| Male | 50.44% | 50.23% | 50.92% | 1.09% | 49.93% | 0.62% | 1.25 |
| Female | 49.56% | 49.77% | 49.08% | 1.09% | 50.07% | 0.62% | 1.25 |
| **Region** |  |  |  |  |  |  |  |
| Midwest | 21.32% | 23.47% | 23.17% | 4.69% | 23.60% | 4.72% | -0.09 |
| West | 18.88% | 18.40% | 19.94% | 4.24% | 17.73% | 3.94% | 0.55 |
| Northeast | 18.94% | 15.53% | 15.08% | 3.89% | 15.73% | 4.00% | -0.16 |
| South | 40.86% | 42.61% | 41.82% | 5.62% | 42.95% | 5.62% | -0.20 |
| **General disability** |  |  |  |  |  |  |
| Physical disability | 54.09% | 53.70% | 53.73% | 1.25% | 53.69% | 0.82% | 0.04 |
| Mental disability | 41.01% | 41.39% | 41.61% | 1.18% | 41.29% | 0.79% | 0.34 |
| Hearing impairment | 0.84% | 0.95% | 0.79% | 0.18% | 1.03% | 0.15% | -1.36 |
| Unknown | 4.06% | 3.96% | 3.88% | 0.43% | 3.99% | 0.31% | -0.34 |

a11,999 cases were in the augmented sample, of which 3,683 cases were released and 8,316 cases were not released in round 4.

This analysis shows that there is very little difference between the proportions for the released and unreleased cases. In fact, nearly all proportions for the unreleased sample are within one sample standard error of the proportion for the released cases. There are some apparent differences that might be related to earnings, though only two of these differences seem to be statistically significant. In particular, 1) the released cases include fewer cases with unknown race than the unreleased cases, and 2) the released cases include more concurrent and fewer SSDI only cases than the unreleased cases. None of these comparison account for simultaneous multiple comparisons, and are unlikely to explain the large differences in earnings. Any differences that exist are likely related to the different release rates in PSUs, which we explore in more detail later.

In Table 8, we review other variables that were not readily available at the time of sampling but can now be obtained from SSA administrative records. We looked at Trial Work Period (TWP) status (number of months that were TWP months), 1619b status (number of months that were 1619b months), average months on the rolls, Suspended or Terminated due to Work (STW) status (number of months that were STW months), and 25 categories of primary impairment. For TWP status, we tallied the proportion for whom at least 4 months in 2009 were TWP months, the proportion for whom 1 to 3 months in 2009 were TWP months, and the proportion for whom none of the months in 2009 were TWP months. For 1619b status and STW status, we tallied the proportion for whom at least 7 months in 2009 were 1619b or STW months, respectively, the proportion for whom 1 to 6 months in 2009 were 1619b or STW months, respectively, and the proportion for whom no months in 2009 were 1619b or STW months, respectively. Again, the standard errors were not calculated for this analysis.

Table 8. Weighted comparison of TWPa status, 1619b status, average months on the rolls, STWb status, and 25 categories of primary impairment among released and unreleased cases for the augmented sample of round 4 beneficiariesc

|  | Released Sample (n=3,683) | Unreleased Sample (n=8,316) |
| --- | --- | --- |
| **TWP status** |  |  |
| Four to nine months in 2009 | 1.00% | 0.90% |
| One to three months in 2009 | 0.98% | 0.95% |
| No months in 2009 | 98.02% | 98.15% |
| **1619b status** |  |  |
| Seven to twelve months in 2009 | 1.17% | 1.18% |
| One to six months in 2009 | 1.11% | 1.48% |
| No months in 2009 | 97.72% | 97.34% |
| **STW status** |  |  |
| Seven to twelve months in 2009 | 1.22% | 1.39% |
| One to six months in 2009 | 0.90% | 1.23% |
| No months in 2009 | 97.88% | 97.38% |
| **Average months on rolls** |  |  |
| Mean | 157.20 | 158.66 |
| Standard error | 1.71 | 1.16 |
| **25 categories of physical impairment** |  |  |
| Major Affective Disorders | 14.74% | 15.51% |
| Schizophrenia & Psychoses | 7.90% | 8.25% |
| Anxiety & Neurotic Disorders | 3.88% | 3.66% |
| Other Mental Disorders | 8.72% | 8.57% |
| Intellectual Disability | 19.11% | 19.20% |
| Back Disorders | 5.62% | 5.64% |
| Musculoskeletal system | 4.10% | 4.49% |
| Infections & Parasitic Diseases | 0.22% | 0.19% |
| HIV/AIDS | 1.06% | 0.84% |
| Neoplasms | 1.60% | 1.83% |
| Endocrine/Nutritional Disorders | 1.98% | 2.06% |
| Blood/Blood-forming Diseases | 0.49% | 0.49% |
| Severe Visual Impairment | 2.17% | 1.94% |
| Severe Hearing Impairment | 1.25% | 1.08% |
| Severe Speech Impairment | 0.08% | 0.06% |
| Nervous System Disorders | 5.86% | 6.53% |
| Circulatory System Disorders | 3.48% | 3.20% |
| Respiratory System Disorders | 1.49% | 1.30% |
| Digestive System Disorders | 0.98% | 0.84% |
| Genitourinary System Disorders | 1.38% | 1.41% |
| Skin/Subcutaneous Tissue Disorders | 0.19% | 0.10% |
| Congenital Anomalies | 0.98% | 0.78% |
| Injuries | 2.96% | 3.17% |
| Other Disorders | 1.66% | 1.53% |
| Missing | 8.09% | 7.34% |

aTrial Work Period

bSuspended or Terminated Due to Work

c11,999 cases were in the augmented sample, of which 3,683 cases were released and 8,316 cases were not released in round 4

In all cases, the differences do not appear to be large enough to account for the large differences in average earnings. The proportions of cases that were in TWP status, STW status, and 1619b status are very small, much smaller than the portion of beneficiaries who had earnings.[[5]](#footnote-5) Even though some differences are apparent (the proportion in STW or 1619b status for 1 to 6 months is higher in the unreleased sample), the proportions themselves are very small, and the overall proportion who were not in either status for any period in 2009 is nearly the same between the released and unreleased samples. For round 3, the differences observed were of the same order as those shown here.

C. Comparisons of earnings between different releases and PSUs with differing release rates

In the prior analyses, we observed that even though released and unreleased cases were different in average earnings, they did not differ in a variety of auxiliary variables. The only conclusion that can be drawn is that earnings are somehow related to whether a case was released into the sample, but unrelated to any of the auxiliary variables shown here. For both NBS round 3 and 4, the proportion of the augmented sample that was released in each PSU varied according to the response rate in that PSU.[[6]](#footnote-6) We accomplished this in two ways: (1) in the first release, we used information from previous rounds to anticipate in which PSUs we would expect to have more difficulty obtaining responses, and released more cases in those PSUs, and (2) after the first release, we monitored data collection, and in those PSUs in which we continued to have difficulty obtaining responses, we released more cases in those PSUs in the second and third release. In round 4 the proportion of the augmented sample that was released in the first release in each PSU ranged from 20.8 to 24.0 percent, whereas for the final released sample the range was from 25.0 to 41.0 percent of the augmented sample. This suggests that perhaps differences in average earnings are somehow related to characteristics of the beneficiates in the PSU; perhaps the beneficiates in the PSUs with lower response rates (and a higher release rate) had lower earnings, leading to lower earnings overall for all released cases. This hypothesis suggests that, if the release rate had been constant across PSUs in the prior NBS, then we would not have observed significant differences in average earnings between released and unreleased cases. Indeed, in the current NBS, the release rates were in fact constant across PSUs, and we have shown that the average earnings did not differ significantly between released and unreleased cases.

We investigate this hypothesis in two ways. First, we compare the average earnings among cases in each release separately against the average earnings in the unreleased cases, looking at all beneficiaries and those who were working in 2009. The results are presented in Tables 9 and 10.[[7]](#footnote-7)

Table 9. Weighted earnings of all round 4 beneficiaries, comparison of released and unreleased cases by releasea

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released | Unreleased (n=8,316) |  |
| Year | Mean | Standarderror | n | PercentWorking | Mean | StandardError | n | PercentWorking |  t |
| First Release |
| 2008 | 1303.75 | 146.63 | 2685 | 19.9% | 1719.54 | 127.13 | 8316 | 21.2% | 3.15 |
| 2009 | 867.47 | 108.56 | 2685 | 14.7% | 1181.12 | 75.51 | 8316 | 16.5% | 3.70 |
| 2010 | 828.37 | 101.55 | 2685 | 14.0% | 1125.29 | 77.47 | 8316 | 15.5% | 3.54 |
| 2011 | 893.91 | 102.90 | 2685 | 13.3% | 1189.04 | 88.61 | 8316 | 14.7% | 3.20 |
| Second Release |
| 2008 | 1145.36 | 202.23 | 491 | 18.4% | 1719.54 | 127.13 | 8316 | 21.2% | 4.33 |
| 2009 | 981.75 | 214.30 | 491 | 13.0% | 1181.12 | 75.51 | 8316 | 16.5% | 2.24 |
| 2010 | 1073.10 | 200.07 | 491 | 12.0% | 1125.29 | 77.47 | 8316 | 15.5% | 0.59 |
| 2011 | 1062.74 | 206.39 | 491 | 11.9% | 1189.04 | 88.61 | 8316 | 14.7% | 1.28 |
| Third Release |
| 2008 | 1299.88 | 198.74 | 507 | 18.9% | 1719.54 | 127.13 | 8316 | 21.2% | 3.17 |
| 2009 | 1128.00 | 166.26 | 507 | 16.8% | 1181.12 | 75.51 | 8316 | 16.5% | 0.64 |
| 2010 | 1168.22 | 255.84 | 507 | 15.5% | 1125.29 | 77.47 | 8316 | 15.5% | -0.44 |
| 2011 | 969.78 | 215.14 | 507 | 14.8% | 1189.04 | 88.61 | 8316 | 14.7% | 2.19 |

a11,999 cases were in the augmented sample, of which 3,683 cases were released (including 2,685 in the first release, 491 in the second release, and 507 in the third release) and 8,316 cases were not released in round 4.

Table 10. Weighted earnings of round 4 beneficiaries working in 2009, comparison of released and unreleased casesa

|  | Released | Unreleased (n = 1,778) |  |
| --- | --- | --- | --- |
| Year | Mean | StandardError | n | PercentWorking | Mean | StandardError | n | PercentWorking | t |
| First Release |
| 2008 | 6168.09 | 762.84 | 550 | 84.2% | 8386.33 | 596.79 | 1778 | 86.8% | 3.47 |
| 2009 | 5919.92 | 615.69 | 550 | 100% | 7139.30 | 376.72 | 1778 | 100% | 2.74 |
| 2010 | 5263.67 | 603.20 | 550 | 78.1% | 6192.86 | 412.90 | 1778 | 75.1% | 2.00 |
| 2011 | 4916.97 | 597.38 | 550 | 69.9% | 6045.55 | 464.25 | 1778 | 66.3% | 2.26 |
| Second Release |
| 2008 | 5249.09 | 865.48 | 108 | 77.1% | 8386.33 | 596.79 | 1778 | 86.8% | 5.10 |
| 2009 | 7580.35 | 1188.05 | 108 | 100% | 7139.30 | 376.72 | 1778 | 100% | -0.95 |
| 2010 | 7395.83 | 1194.33 | 108 | 72.2% | 6192.86 | 412.90 | 1778 | 75.1% | -2.44 |
| 2011 | 6595.56 | 1288.69 | 108 | 60.1% | 6045.55 | 464.25 | 1778 | 66.3% | -1.01 |
| Third Release |
| 2008 | 6455.27 | 878.24 | 116 | 86.3% | 8386.33 | 596.79 | 1778 | 86.8% | 3.13 |
| 2009 | 6723.60 | 711.14 | 116 | 100% | 7139.30 | 376.72 | 1778 | 100% | 1.03 |
| 2010 | 6545.39 | 1328.78 | 116 | 77.5% | 6192.86 | 412.90 | 1778 | 75.1% | -0.68 |
| 2011 | 4775.03 | 1158.53 | 116 | 61.9% | 6045.55 | 464.25 | 1778 | 66.3% | 2.38 |

a774 of 3,683 released cases and 1,778 of 8,316 unreleased cases were working in 2009. The 774 cases included 550 in the first, 108 in the second, and 116 in the third release. All unreleased cases working in 2009 are included in the unreleased columns.

Tables 9 and 10 show that the same patterns evident overall are also more or less apparent in each of the releases. There is no consistent pattern which would indicate that the overall earnings were brought down by the addition of cases in the second and third release.

Another way to evaluate this question is to compare earnings in the PSUs with the highest release rate (where it was more difficult to get a response) with earnings in the PSUs with the lowest release rate (where it was easier to get a response), regardless of which release the observations came from. Table 11 shows the ten PSUs with the highest release rates and the ten PSUs with the lowest release rates. In general, the counties with the highest release rates (with lower response rates) tend to be urban or inner suburban; the counties with the lowest release rates (with higher response rates) tend to be rural or outer suburban areas.

Table 11. PSUs with highest and lowest release rates.

| PSU ID | Description | Release Rate |
| --- | --- | --- |
| Ten highest release rates |
| 36035 | New York County, NY (Manhattan) | 41.0% |
| 12025 | Orange County, FL (Orlando) | 37.5% |
| 51019 | Suffolk City, VA (Virginia Beach-Norfolk-Newport News MSA) | 36.9% |
| 4010 | Maricopa County, AZ (Phoenix) | 36.8% |
| 53003 | Whatcom County, WA (Bellingham) | 36.6% |
| 12027 | Volusia County, FL (Daytona Beach) | 36.4% |
| 36040 | Nassau County, NY (Long Island) | 36.4% |
| 10002 | Kent County, DE (Dover) | 36.3% |
| 31009 | Sarpy & Saunders Counties, NE (Omaha-Council Bluffs MSA) | 35.6% |
| 27002 | St. Louis, Cook, & Lake Counties, MN (Duluth & rural NE MN) | 35.5% |
| Ten lowest release rates |
| 55018 | Chippewa, Rush, & Taylor Counties, WI (Eau Claire MSA & rural NW WI) | 25.0% |
| 40024 | Adair & Cherokee Counties, OK (rural NE OK) | 25.9% |
| 13015 | Fayette & Spalding Counties, GA (Atlanta-Sandy Springs-Roswell MSA) | 26.0% |
| 5034 | Lonoke, Monroe, Prairie, & Woodruff Counties, AR (Little Rock-N. Little Rock-Conway MSA & rural EC AR) | 26.3% |
| 37003 | Buncombe & Madison Counties, NC (Asheville) | 26.6% |
| 21040 | Barren & Metcalfe Counties, KY (rural SC KY) | 26.8% |
| 42015 | Somerset County, PA (rural SW PA) | 27.0% |
| 17042 | Coles, Douglas, Moultrie, & Piatt Counties, IL (Champaign-Urbana MSA & rural EC IL) | 27.2% |
| 26011 | Ingham County, MI (Lansing-East Lansing MSA) | 27.3% |
| 42012 | Fayette County, PA (Pittsburgh MSA) | 27.3% |

The combined earnings for each group are presented in Tables 12 and 13, where Table 12 is focused on all beneficiaries, and Table 13 is focused on working beneficiaries only. These tables include both released and unreleased cases.

Table 12. Weighted earnings of all round 4 beneficiaries, comparison by PSU groupa

|  |  |  |  |
| --- | --- | --- | --- |
|  | High Release Rate PSUs (n=1,475) | Low Release Rate PSUs (n=1,535) | All Other PSUs (n=8,989) |
| Year | Mean | Std.error | PercentWorking | Mean | Std.Error | PercentWorking | Mean | Std.Error | PercentWorking |
| 2008 | 1675.9 | 280.5 | 23.5% | 1722.0 | 467.9 | 19.1% | 1549.5 | 102.9 | 20.5% |
| 2009 | 1228.1 | 163.9 | 19.3% | 1077.7 | 195.7 | 14.4% | 1084.7 | 71.8 | 15.7% |
| 2010 | 1095.6 | 170.7 | 17.7% | 1034.5 | 212.0 | 13.6% | 1057.1 | 68.1 | 14.8% |
| 2011 | 1232.4 | 232.0 | 16.2% | 1057.1 | 222.8 | 12.7% | 1097.0 | 69.0 | 14.3% |

a11,999 cases were in the augmented sample, of which 1,475 resided in the 10 PSUs with the highest release rate, 1,535 resided in the 10 PSUs with the lowest release rate, and 8,989 resided in the remaining PSUs.

Table 13. Weighted earnings of all round 4 beneficiaries working in 2009, comparison by PSU groupa

|  |  |  |  |
| --- | --- | --- | --- |
|  | High Release Rate PSUs (n=364) | Low Release Rate PSUs (n=308) | All Other PSUs (n=1,880) |
| Year | Mean | Std.error | Percent Working | Mean | Std.Error | Percent Working | Mean | Std.Error | Percent Working |
| 2008 | 6776.3 | 627.4 | 86.2% | 9501.9 | 2665.0 | 87.4% | 7672.8 | 440.8 | 85.7% |
| 2009 | 6360.0 | 435.2 | 100% | 7472.9 | 1216.8 | 100% | 6894.5 | 338.7 | 100% |
| 2010 | 5018.5 | 535.2 | 71.2% | 6740.9 | 1401.3 | 75.4% | 6154.5 | 345.2 | 76.5% |
| 2011 | 4928.6 | 609.6 | 65.4% | 6557.7 | 1532.9 | 71.7% | 5815.0 | 360.3 | 66.1% |

aOf the 11,999 cases in the augmented sample, 2,552 were working in 2009. Of these, 364 resided in the 10 PSUs with the highest release rate, 308 resided in the 10 PSUs with the lowest release rate, and 1,880 resided in the remaining PSUs.

If the hypothesis was correct, we would expect two things to occur: (1) the earnings in the 10 PSUs with a high release rate would be much lower than that of the 10 PSUs with a low release rate, dampening the earnings of released cases overall and (2) when removing these 20 PSUs from the analysis, the differences in average earnings between released and unreleased cases would be reduced. The first item is evaluated in Tables 12 and 13, and the second evaluated in Tables 14 and 15.

For the first item, the conclusion is mixed. As Table 12 indicates, there is not much difference in earnings between the three groups of PSUs when non-working beneficiaries are included, which does not support the hypothesis. However, in Table 13, we do see a large difference, particularly for 2008 earnings, between the PSUs with a high release rate and a low release rate (the standard errors of these means are large, so none of these differences are likely to be statistically significant). The number of working beneficiaries is small in the sample, but, upon review of Table 12, there does appear to be a substantial difference in the percentage of working beneficiaries between PSUs with high and low release rates. The high percentage of working beneficiaries in the PSUs with a high release rate might in part explain the resulting low response rates in these PSUs. The large differences in the mean earnings may be at least partially explained by the differences in the percentage of working beneficiaries. The direction of earnings differences in Table 13 would also seem to support the hypothesis.

To evaluate the second item, we removed the observations in these 20 PSUs from the analysis. As is apparent from Tables 14 and 15, the earnings in PSUs with release rates that are closer to the average also have large differences in earnings, nearly as large as those observed when the 20 PSUs with high release rates and low release rates are included. Clearly, this does not support the hypothesis. Something we have not observed and cannot account for is causing much of the differences in average earnings.

Table 14. Weighted earnings of round 4 beneficiaries, released and unreleased cases, comparison of all beneficiaries with beneficiaries where 20 PSUs were removed, with highest and lowest release ratesa

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released (n=2,733) | Unreleased (n=6,256) |  |
| Year b | Mean | Standarderror | PercentWorking | Mean | StandardError | PercentWorking |  t |
| 2008 all PSUs | 1285.11 | 113.64 | 19.6% | 1719.54 | 127.13 | 21.2% | 3.52 |
| 2008 63 PSUs | 1253.72 | 135.56 | 19.6% | 1676.47 | 123.27 | 20.9% | 3.33 |
| 2009 all PSUs | 919.59 | 86.34 | 14.8% | 1181.12 | 75.51 | 16.5% | 3.31 |
| 2009 63 PSUs | 869.87 | 99.52 | 14.7% | 1176.96 | 85.49 | 16.2% | 3.41 |
| 2010 all PSUs | 907.26 | 91.06 | 14.0% | 1125.29 | 77.47 | 15.5% | 2.66 |
| 2010 63 PSUs | 907.40 | 110.83 | 13.7% | 1121.37 | 83.64 | 15.3% | 2.31 |
| 2011 all PSUs | 924.54 | 85.78 | 13.4% | 1189.04 | 88.61 | 14.7% | 3.01 |
| 2011 63 PSUs | 933.13 | 101.56 | 13.2% | 1167.29 | 89.13 | 14.7% | 2.52 |

a11,999 cases were in the augmented sample, of which 8,989 resided in PSUs that were not among the 10 highest or 10 lowest release rates. Of these, 2,733 were released in round 4, and 6,256 were not released in round 4.

bThe term “PSU” in these labels refers to both PSUs and Secondary Sampling Units (SSUs). The rows labeled “all PSUs” correspond to the rows in Table 4.

Table 15. Weighted earnings of round 4 beneficiaries working in 2009, released and unreleased cases, comparison of all working beneficiaries with working beneficiaries where 20 PSUs were removed, with highest and lowest release ratesa

|  |  |  |  |
| --- | --- | --- | --- |
|  | Released (n=574) | Unreleased (n=1,306) |  |
| Year b | Mean | Standarderror | PercentWorking | Mean | Standarderror | PercentWorking |   T |
| 2008 all PSUs | 6125.18 | 553.68 | 83.8% | 8386.33 | 596.79 | 86.8% | 3.87 |
| 2008 63 PSUs | 6187.65 | 726.14 | 83.8% | 8251.52 | 540.01 | 86.4% | 3.42 |
| 2009 all PSUs | 6222.71 | 471.29 | 100% | 7139.30 | 376.72 | 100% | 2.25 |
| 2009 63 PSUs | 5919.75 | 566.63 | 100% | 7274.40 | 430.95 | 100% | 2.84 |
| 2010 all PSUs | 5695.02 | 506.59 | 77.4% | 6192.86 | 412.90 | 75.1% | 1.12 |
| 2010 63 PSUs | 5703.65 | 620.98 | 76.7% | 6330.22 | 455.30 | 76.4% | 1.22 |
| 2011 all PSUs | 5060.52 | 485.39 | 67.6% | 6045.55 | 464.25 | 66.3% | 2.09 |
| 2011 63 PSUs | 5097.30 | 578.60 | 65.2% | 6094.67 | 510.85 | 66.5% | 1.87 |

a11,999 cases were in the augmented sample, of which 2,552 were working in 2009. Of these, 1,880 resided in PSUs that were not among the 10 highest or 10 lowest release rates. Of these, 574 were released in round 4, and 1,306 were not released in round 4.

bThe term “PSU” in these labels refers to both PSUs and Secondary Sampling Units (SSUs). The rows labeled “all PSUs” correspond to the rows in Table 5.

III. Summary

It is still not entirely clear why we see differences between released and unreleased cases in average earnings, particularly in round 4. There do not appear to be significant differences between any of the auxiliary variables that we have at our disposal that might be related to earnings. In addition, the differences that do exist are not large enough to explain the large differences in average earnings. When we examined differences by PSU, we did find that the 10 PSUs with the highest release rate appeared to have lower earnings than the 10 PSUs with the lowest release rate. However, when these 20 PSUs are removed, the observed differences in mean earnings do not change that much.[[8]](#footnote-8) Thus, we are left with the conclusion that the observed differences in average earnings for round 4 are the result of randomness that is mostly unrelated to any of the variables we have can observe and investigate.

IV. Recommendations for Subsequent Rounds

Although it is unlikely that the samples in future rounds will show the differences in earnings that were apparent in round 4 of the prior NBS, the risk of a recurrence of this problem is non-trivial, particularly given that (1) we have no way of controlling the sample for earnings since we don’t have access to this data and (2) the priority of maintaining relatively equal weights (to minimize the variance) makes it necessary to release more cases in PSUs with low response rates, which may be related to earnings, as we have seen. Because there is a lag in the determination of accurate earnings data, our recommendation is to do nothing different at the time of sample selection. However, it may be useful to create a new set of weights two or three years after the initial sample selection that accounts for different earnings levels through another post stratification. Though it is not possible for Mathematica to create these weights themselves, it may be possible for Mathematica statisticians to coach staff members at SSA in the creation of these poststratified weights.

cc: Debra Wright, Gina Livermore, David Stapleton, Frank Potter, Emily Roessel

1. Other analyses were also conducted that are not included in this memo. For both rounds, we verified that the mean earnings for sample cases do not differ significantly from the earnings on the sampling frame when using the augmented sample weights. In those instances where the mean earnings in the released sample significantly differed from that of the unreleased sample, the mean earnings in the released sample also differed from that of the sampling frame when using the release-adjusted weights. In addition, we compared the mean earnings from the released sample only to the sampling frame using the release-adjusted sampling weights. None of the results of these analyses are presented here for the sake of brevity. [↑](#footnote-ref-1)
2. Because we did not have direct access to the data, calculation of the correct standard error was not feasible. Combining the standard errors in this way assumes the groups are independent, which they are not. However, the correlation between the groups is likely to be small. This is verified by an initial the analysis in which we compared released cases to the frame and unreleased cases to the frame. In both cases, the differences were not significant. We present the present the comparison of released vs. unreleased cases for simplicity. [↑](#footnote-ref-2)
3. The estimated means in Table 1 and Table 4 include the approximately around 80 percent of all cases with a valid zero value—the the beneficiary was not working. Therefore, the population distribution has around 80 percent of the observed data at zero and the remaining 20 percent with a value greater than zero—obviously not normal. [↑](#footnote-ref-3)
4. The large differences between means and medians were due to the skewed distribution of earnings. There did not appear to be an issue with outliers. We analyzed the data with a few of the top cases removed, and the conclusions did not change. [↑](#footnote-ref-4)
5. If we also consider TWP completion before the year, we might have larger proportions that could be different. [↑](#footnote-ref-5)
6. This variation in release rates likely explains the small differences in auxiliary variables that we did observe. [↑](#footnote-ref-6)
7. The standard errors of the differences in these tables that are used in the t-tests are calculated in the same way as in Tables 1-6. The same caveats with the calculation of the standard errors and t-tests in those tables also apply here. [↑](#footnote-ref-7)
8. Another way to verify how much impact differing release rates across PSUs impacted earnings between released and unreleased cases is to run a regression of earnings on all of the auxiliary variables, including PSU release rate, and show that the means *after adjusting for these variables* are still different for the two samples. This would require more time, however, and is unlikely to provide a different conclusion. [↑](#footnote-ref-8)