National Beneficiary Survey-General Waves Round 7: Nonresponse Bias Analysis

Final Report

October 20, 2021

Eric Grau and Hanzhi Zhou

This page has been left blank for double-sided copying

CONTENTS

TABLES

ACRONYMS

AHRF	Area Health Resource File
CAPI	Computer-assisted personal interviewing
CATI	Computer-assisted telephone interviewing
CHAID	Chi-Squared Automatic Interaction Detector
DCF	Disability Control File
FRA	Full retirement age
IRS	Internal Revenue Service
MCAR	Missing completely at random
NBS	National Beneficiary Survey
PSU	Primary sampling unit
RBS	Representative beneficiary sample
SGA	Substantial gainful activity
SPSS	Statistical Package for the Social Sciences (SPSS is a registered trademark of SPSS Inc., of Chicago, Illinois)
SSA	Social Security Administration
SSDI	Social Security Disability Insurance (Title II of the Social Security Act)
SSI	Supplemental Security Income (Title XVI of the Social Security Act)
SSU	Secondary sampling unit
SWS	Successful worker sample
TTW	Ticket to Work and Self-Sufficiency

This page has been left blank for double-sided copying.

NBS DATA DOCUMENTATION REPORTS

The following publicly available reports are available from SSA on its website (https://www.ssa.gov/disabilityresearch/nbs_round_7.html):

- **User's Guide for Restricted- and Public-Use Data Files** (Callahan et al. 2021). This report provides users with information about the restricted-use and public-use data files, including construction of the files; weight specification and variance estimation; masking procedures employed in the creation of the public-use file; and a detailed overview of the questionnaire design, sampling, and data collection for the National Beneficiary Survey (NBS)–General Waves. The report provides information covered in the Editing, Coding, Imputation, and Weighting Report and the Cleaning and Identification of Data Problems Report (described below)—including, procedures for data editing, coding of open-ended responses, and variable construction—as well as a description of the imputation and weighting procedures and development of standard errors for the survey. In addition, this report contains an appendix addressing total survey error and the NBS.

- **NBS Public-Use File Codebook** (McDonald et al. 2021). This codebook provides extensive documentation for each variable in the file, including variable name, label, position, variable type and format, question universe, question text, number of cases eligible to receive each item, constructed variable specifications, and user notes for variables on the public-use file. The codebook also includes frequency distributions and means as appropriate.

- **NBS–General Waves Questionnaire** (Callahan et al. 2021). This document contains all items on Round 7 of the NBS–General Waves and includes documentation of skip patterns, question universe specifications, text fills, interviewer directives, and checks for consistency and range.

- **Editing, Coding, Imputation, and Weighting Report** (Grau et al. 2021). This report summarizes the editing, coding, imputation, and weighting procedures as well as the development of standard errors for Round 7 of the NBS–General Waves. It includes an overview of the variable naming, coding, and construction conventions used in the data files and accompanying codebooks; describes how the sampling weights were computed to the final post-stratified analysis weights for the representative beneficiary sample; outlines the procedures used to impute missing responses; and discusses procedures that should be used to estimate sampling variances for the NBS.

- **Cleaning and Identification of Data Problems Report** (McDonald et al. 2021). This report describes the data processing procedures performed for Round 7 of the NBS–General Waves. It outlines the data coding and cleaning procedures and describes data problems, their origins, and the corrections implemented to create the final data file. The report describes data issues by sections of the interview and concludes with a summary of the types of problems encountered and general recommendations.

- **NBS Nonresponse Bias Analysis** (current report). This report discusses whether the nonresponse adjustments applied to the sampling weights of Round 7 of the NBS–General Waves appropriately accounted for differences between respondents and nonrespondents or whether the potential for nonresponse bias still exists.

The following restricted-use report is available from SSA through a formal data sharing agreement:

- **NBS Restricted-Access Codebook** (McDonald et al. 2021) This codebook provides extensive documentation for each variable in the file, including variable name, label, position, variable type and format, question universe, question text, number of cases eligible to receive each item, constructed variable specifications, and user notes for variables on the restricted-access file. The codebook also includes frequency distributions and means as appropriate.

I. INTRODUCTION

In all studies, final survey estimates are based solely on the survey's respondents. Errors may arise in the estimates resulting from unit nonresponse if there are systematic differences between individuals who respond to a survey and those who do not. Nonresponse-adjusted weights attempt to mitigate these differences by identifying respondents and nonrespondents who are similar on characteristics available for both and by adjusting the weights of the respondents to compensate for the nonrespondents. In studies where these adjustments are able to account for differences between nonrespondents and respondents, the survey estimates would have minimal potential for nonresponse bias.

The purpose of this report is to determine whether the nonresponse and raking[1] adjustments applied to the sampling weights of Round 7 of the National Beneficiary Survey (NBS)–General Waves appear to have effectively accounted for differences between respondents and nonrespondents, or whether the potential for nonresponse bias has been minimized. The study team cannot directly measure nonresponse bias without knowing how nonrespondents would have answered survey items; however, we can examine variables available for both respondents and nonrespondents that we believe are correlated with responses to survey items. For Round 7 of the NBS, we met the goals of the study through three samples: (1) a cross-sectional sample of all beneficiaries (the representative beneficiary sample, or RBS); (2) a cross-sectional sample of a subset of beneficiaries who maintained a minimum level of earnings for a sustained period (a successful worker sample, or SWS); and (3) a subset of SWS cases from Round 6, which were followed longitudinally in Round 7. The survey was administered to all three of these samples simultaneously. Mathematica collected data by using computer-assisted telephone interviewing (CATI). We also deployed in-person field locators to follow up with some CATI nonrespondents.[2]

While more detail is provided in this report, the main findings of our analysis indicate that the nonresponse adjustment alleviated all differences observed between respondents and nonrespondents in all three samples for the variables that we had at our disposal. We did find that, after weighting, the estimate of the proportion of the "all others" race category was significantly less than in the frame in the cross-sectional SWS, though this was primarily due to issues other than nonresponse. Any conclusions involving race should be viewed with caution due to the high levels of missing data in that variable. There were other sources of potential bias

---

[1] Raking, also called "iterative proportional fitting," is a method of adjusting weights in an iterative, sequential manner so that weighted marginal totals on key variables of interest match those of the population one variable at a time. It is considered a type of post-stratification and is the method we used to post-stratify weights in the NBS. For the remainder of this report, we use the terms "raking" and "post-stratification" interchangeably, even though "post-stratification" is a more general term than "raking."

[2] For a portion of the RBS, we did not employ field follow-up. Instead, we randomly selected telephone nonrespondents for a second phase of data collection involving field follow-up, described later in Section A.1.a. We also did not employ field follow-up for a portion of the SWS. This portion, referred to as the "unclustered" sample, is also described later in Section A.1.b.

representing some small populations where the weighted estimates from the sample differed from the frame, because these populations could not be controlled for when creating the base weights. This was because the variables representing these populations were (1) not important enough to be considered for the variables used in raking and (2) not included as covariates in the final nonresponse models, most likely because the sample counts were too small. As a result, these differences could not be rectified when adjusting these weights for nonresponse or post-stratifying them to marginal population totals.

A. Study overview

Sponsored by the Office of Retirement and Disability Policy at the Social Security Administration (SSA), the NBS–General Waves collects data on the employment-related activities of working-age beneficiaries of Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI). In 2019, Mathematica conducted the seventh round of data collection since the NBS began in 2004. The first four rounds of the survey—in 2004, 2005, 2006, and 2010—helped glean information about beneficiary impairments; health; living arrangements; family structure; occupation before disability; and use of non-SSA programs (for example, the Supplemental Nutrition Assistance Program). Rounds 1 to 4 also evaluated the Ticket to Work and Self-Sufficiency (TTW) program. In Rounds 5 (2015), 6 (2017), and 7 (2019), we sought to uncover important information about the factors that promote beneficiaries' self-sufficiency and, conversely, the factors that impede beneficiaries' efforts to maintain employment.[3]

The NBS–General Waves collects important beneficiary data that are not available from SSA administrative data or other sources, including more detailed information about beneficiaries' disabilities (other than their general disability classification) as well as disability payment information, interest in work, use of services, and employment.

The survey addresses five major questions:

- What are the work-related goals and activities of SSI and SSDI beneficiaries, particularly as they relate to long-term employment?

- What are the short-term and long-term employment outcomes for SSI and SSDI beneficiaries who work?

- What supports help SSA beneficiaries with disabilities find and keep jobs and what barriers to work do they encounter?

- What are the characteristics and experiences of beneficiaries who work?

---

[3] In this report, the NBS rounds conducted in 2004, 2005, 2006, 2010, 2015, 2017, and 2019 are referred to as Round 1, Round 2, Round 3, Round 4, Round 5, Round 6, and Round 7, respectively.

- What health-related factors, job-related factors, and personal circumstances hinder or promote employment and self-sufficiency?

SSA will combine data from Round 7 of the NBS–General Waves with SSA administrative data to provide critical information on access to jobs and employment outcomes for beneficiaries. As a result, SSA and external researchers who are interested in disability and employment issues may use estimates from the survey for policymaking and program planning efforts.

In this report, we assess the potential for nonresponse bias separately for the three samples (RBS, cross-sectional SWS, and longitudinal SWS). We first describe the sample design and the population that each sample is supposed to represent, followed by a description of the nonresponse adjustments to the sampling weights. We then provide the unweighted and weighted response rates for each sample and its substrata. In Tables 4, 7, and 10 for the RBS, we assess (1) how well the sample represents the data from the sampling frame; (2) how ineligible sample cases differ from the rest of the sample and how this would affect comparisons with the frame; (3) how sample respondents differ from nonrespondents; and (4) how well nonresponse adjusted weights account for these differences, using the unadjusted base weights and weights adjusted for nonresponse. We follow this with similar comparisons for the cross-sectional SWS in Tables 5, 8, and 11. Finally, we make three of these four comparisons with the longitudinal SWS in Tables 6, 9, and 12. (We do not compare the sample estimates to those of the frame because the frame values for the longitudinal SWS are unknown.) The comparison between the estimates with adjusted and unadjusted weights allows us to (1) see the potential for nonresponse bias after removing nonrespondents and making no nonresponse adjustments to the weights and (2) assess the effectiveness of nonresponse adjustment procedures on the potential for nonresponse bias.

1.   Sample design and target populations

For all survey rounds, the NBS has used a multistage sampling design for both the RBS and cross-sectional SWS, with an independently drawn, supplemental, single-stage sample for some successful worker populations.[4] In Round 7, we drew the cross-sectional SWS and RBS independently from separate frames, although the SWS frame was a subset of the RBS frame. This means that some sample members could have been selected for both the RBS and the cross-sectional SWS—which occurred for 90 individuals (of which, 30 responded[5]). Because most

---

[4] The RBS and the main sample of the SWS involved selecting individuals within selected clusters of geographic areas; thus, they are referred to as clustered samples. The supplemental sample (for the SWS only) was selected across the entire population of successful workers and was therefore not limited to those residing in selected clusters. It is therefore referred to as an unclustered sample. This is discussed in detail later in the report.

[5] Of the 30 who responded, 28 were considered completes for both the cross-sectional SWS and RBS. Of the remaining 2 respondents, 1 was completed in the field for the SWS but was not selected for field operations in the second phase of the RBS, and thus was not an RBS complete. The other was an RBS complete but was considered ineligible for the cross-sectional SWS because the person had not been working in the past six months. Therefore, there were 29 total RBS completes and 29 total cross-sectional SWS completes.

analyses do not require combining the samples, we did not adjust the RBS and cross-sectional SWS weights for these duplicates. However, in case an analysis would require combining the samples, we also created composite weights that accounted for duplicates (individuals who were selected for both samples). These composite weights also accounted for those in the RBS that were not part of the cross-sectional SWS but could have been potentially sampled for the cross-sectional SWS because they were part of the SWS frame.

The longitudinal SWS was composed of all cases that (1) completed a Round 6 SWS interview and (2) reported currently working at the time of the Round 6 survey.[6] Table 1 summarizes the actual sample sizes and number of completed interviews for the RBS, cross-sectional SWS, and longitudinal SWS. Note that longitudinal SWS cases carried over from Round 6 also had a chance of being selected, if eligible, for the independently selected Round 7 RBS or the Round 7 cross-sectional SWS.[7]

In Rounds 1 through 4, we used data from SSA on the counts of eligible beneficiaries in each county in 2003 to form 1,330 primary sampling units (PSUs), each of which consisted of one or more counties. We selected a new sample of PSUs in Round 5 from the same group of 1,330 PSUs that were formed prior to Round 1 (in 2003). We used the same PSUs in Rounds 6 and 7 (for both the RBS and the SWS main sample) that we had selected in Round 5. We classified two PSUs as certainty selections (Los Angeles County and Cook County).[8] These counties were certainty selections based on the selection frequencies for the PSUs computed using a composite size measure.[9] Within these two counties, we formed secondary sampling units (SSUs) composed of one or more five-digit zip codes within each county—using counts of beneficiaries for the SSUs in each age stratum for the composite size measure—and selected a sample of SSUs within the certainty PSUs for the RBS. In the SWS, we did not use the SSUs; individuals were selected within the entire certainty PSUs and were not limited to the selected SSUs. Details on the sample design of the NBS–General Waves, including the selection of PSUs and SSUs, are available in the Editing, Coding, Imputation, and Weighting Report (Grau et al. 2021).

a. RBS

For the RBS, we fielded a nationally representative sample of 11,299 SSA disability beneficiaries. The sample design for the RBS in Round 7 was similar to the design of the RBS in

---

[6] We did not create composite weights that combined sample cases from the longitudinal SWS with any other sample. Longitudinal SWS sample members were selected based on their work activity at Round 6; therefore, they cannot be meaningfully combined with any of the other Round 7 samples.

[7] In general, the only way a longitudinal SWS case would be sent for field follow-up in Round 7 was if it was also selected for one of these other samples. It would be sent to the field under those samples' protocols.

[8] Los Angeles County includes the city of Los Angeles; Cook County includes the city of Chicago.

[9] The composite size measure was the sum of the products of the sampling fraction for each age category stratum in the RBS and the estimated count of beneficiaries in that stratum and PSU (Folsom et al. 1987).

prior rounds, though there were two important changes: (1) we stratified the sample of PSUs differently in Rounds 1 through 4 than we did in Rounds 5 through 7,[10] and (2) all telephone nonrespondents were followed up in the field in Rounds 1 through 6, but only a random sample of telephone nonrespondents were followed up in the field in Round 7, as described in more detail below.

The target population for the RBS consisted of SSI recipients and SSDI beneficiaries between the ages of 18 and full retirement age (FRA) who resided in all 50 states and the District of Columbia (excluding outlying territories) and who were in an active pay status[11] as of June 30, 2018. As of that date, the target population consisted of approximately 13.7 million beneficiaries. We stratified the cross-sectional RBS by four age-based strata within the PSUs: (1) age 18 to 29, (2) age 30 to 39, (3) age 40 to 49, and (4) age 50 and older. To ensure a sufficient number of persons seeking work, we oversampled beneficiaries in the first three cohorts (age 18 to 49). The target number of completed interviews for Round 7 was 1,111 beneficiaries in each of the three younger age groups. For those age 50 and older, the target number of completed interviews was 667 beneficiaries.

To reduce data collection costs, we implemented a two-phase sample design for the RBS in Round 7. Our goal was to achieve the same number of completed interviews (4,000) as in past rounds, but with a greater proportion completed by phone instead of in the field. In Phase 1, we reserved a minimum of 12 weeks for cases to work their way through the prespecified phone interview protocol for each sample release. Next, in Phase 2, we randomly subsampled telephone nonrespondents for field follow-up instead of fielding all of these cases. This approach necessitated increasing the sample size for the RBS compared with prior rounds. Note that, when weighted for the two-phase design, the weighted response rate was the same regardless of what proportion of Phase 1 nonrespondents was subsampled for Phase 2.

b.   Cross-sectional SWS

The cross-sectional SWS target population was limited to SSI and SSDI beneficiaries who were eligible for the RBS but were considered successful workers because their earnings for a sustained period were sufficiently high. In particular, the SSI and SSDI beneficiaries were required to (1) have earnings above SSA's non-blind substantial gainful activity (SGA) monthly earnings level ($1,180 in 2018 and $1,220 in 2019) for a minimum of three consecutive calendar

---

[10] As noted earlier, the sample design for Rounds 1 through 4 included two samples: one for all beneficiaries (the RBS) and one for the ticket participants (the TPS). To accommodate the rollout of the TTW program, the PSUs were sampled within strata defined by the three phases of the rollout. The design for Round 5 included one sample only: a sample of all beneficiaries. The PSUs were not drawn within strata, except those defined by the two certainty PSUs. The Round 6 and Round 7 samples used the same PSUs as those sampled in Round 5.

[11] Active status includes beneficiaries who are currently receiving cash benefits as well as those whose benefits have been temporarily suspended for work or other reasons. It does not include beneficiaries whose benefits have been terminated.

months at any time between August 1, 2018, and July 31, 2019, and (2) be younger than age 62 on June 30, 2018.[12] The successful work must have occurred within a time frame so that the vast majority would be interviewed within six months of the end of their successful work (if they were not currently working) and their earnings had to have been revealed in the Disability Control File (DCF) at the time of data extraction—removing from the population any successful workers who had a long delay in having their earnings recorded on the DCF.[13] Using these constraints to define the target population, we identified a population of 101,698 successful workers.[14] From this frame, we fielded a nationally representative sample of 8,590[15] successful workers. We included one screening question as an additional constraint: the sampled successful workers had to indicate that they had been working in the past six months.[16]

To ensure a large enough number of successful workers for sampling, we formed seven successive frames of successful workers over time. Each one was revealed by comparing the full sampling frame to updated earnings information and identifying all successful workers at that time, then removing them from subsequent frames to make the frames mutually exclusive. The SWS sampling frames were all subsets of the same sampling frame used for the Round 7 RBS sample, and are therefore referred to as extracts from the larger frame. Within each of the seven extracts, we stratified the SWS into two strata defined by beneficiary type (SSDI only and SSI, which included both SSI only and concurrent beneficiaries) and selected a probability sample from each extract. The targeted number of completed interviews for the two strata was 1,500 interviews apiece across all extracts. We did not know the size of each extract before sample selection; the first sample size allocation to the samples in each extract was based on historical data. After the release of each extract, the allocation of sample sizes to the samples from the

---

[12] We used an age limit of 62 in Round 6 to ensure that longitudinal cases would still be younger than age 65 at the time of the Round 7 interview. Although we did not plan to follow the Round 7 cross-sectional successful workers longitudinally, we maintained the age limit of 62 in the Round 7 cross-sectional sample for the sake of consistency with Round 6.

[13] Some SSI and SSDI beneficiaries would be considered successful workers because their earnings and age met the threshold, but they had to be excluded from the target population for the sampling effort due to a delay in recording their earnings on the DCF. For these individuals, a lag of up to three years existed between the time that they received their earnings and the time that the earnings data were recorded in the DCF. There was no way they could be identified in time for the data extraction. Two years after the completion of this document, the DCF earnings data will be revisited and the weights will be post-stratified to account for the new information that the updated DCF earnings data will provide.

[14] This count does not include all beneficiaries who had three consecutive months of earnings above non-blind SGA. It only includes those who met that condition and an additional condition: their earnings were recorded in the DCF at the time of the extraction.

[15] The cross-sectional SWS includes 152 duplicates across two sample components (discussed later in this section). As a result, 8,438 unique cases were sampled.

[16] This screening question was included to account for situations where a long period of time had elapsed between the date when the case was released for data collection and the interview date. Few cases were actually removed from the sample due to this screening question, especially in later extracts.

remaining extracts was adjusted to make the allocation as proportional as possible to the population of successful workers over time, within each of the two beneficiary-type strata (SSDI only and SSI). We did not complete sample selection until after the release of the last extract.

Because of the concerns about the number of successful workers within strata and their distribution across PSUs within each extract, we decided to supplement the main SWS (within the PSUs) with a second, independent sample of successful workers. This supplemental sample was divided into two geographic strata (successful workers residing in a sampled PSU and successful workers not residing in any of the sampled PSUs), in addition to the strata based upon the extract and beneficiary title (SSI versus SSDI only).[17] We refer to the multistage sample design as the clustered sample and to the second independent sample as the unclustered sample.[18] We call the combination of data from the clustered and unclustered samples to calculate estimates a dual sample design. The clustered sample included in-person follow-up for sample members who could not be located or otherwise did not respond by phone; the unclustered sample did not have in-person follow-up.

After the completion of the sample selection for all seven extracts, we created a single set of cross-sectional SWS composite weights that combined information from the clustered and unclustered cross-sectional SWS, which appropriately accounted for the different follow-up rules between the two samples.[19]

Table 1 includes selected and completed sample sizes for the RBS, cross-sectional SWS, and longitudinal SWS. It includes the total across the clustered and unclustered samples in both the cross-sectional SWS and longitudinal SWS. It does not break out the counts between the clustered and unclustered samples in either case and also does not account for duplicates.[20] Because of the availability of administrative data for all SSI and SSDI beneficiaries, we were able to ascertain many of the true properties of each target population, which provided us with the tools we needed for the processing of this analysis.

---

[17] Given that the target population for the NBS did not include Puerto Rico or other outlying territories, we excluded from the frame all beneficiaries and successful workers who resided in those areas.

[18] Because of the small populations of successful workers, Mathematica often selected successful workers who resided in both the selected PSUs for the clustered and in-PSU strata of the unclustered samples. Hence, we had to account for these duplicate cases in the weighting process.

[19] These composite weights, which combined weights from the clustered and unclustered samples in the SWS, should not be confused with the composite weights that combined the RBS sampling weights and the cross-sectional SWS sampling weights that we briefly alluded to in the introductory paragraphs.

[20] Duplicates occur between the clustered and unclustered samples within both the cross-sectional SWS and the longitudinal SWS. Duplicates also occur between the RBS, cross-sectional SWS, and longitudinal SWS.

c. Longitudinal SWS

The Round 7 longitudinal sample consisted of Round 6 cross-sectional SWS respondents who were working at the time of the Round 6 interview. In the Round 6 survey, we defined successful workers as SSI or SSDI beneficiaries who (1) were active or in suspense status[21] on June 30, 2016; (2) had earnings above SSA's nonblind SGA earnings level[22] for at least three consecutive calendar months at any time from August 1, 2016, through July 31, 2017; and (3) were younger than 62 on June 30, 2016. We used the same definition for successful workers in Round 7, except for the dates and SGA earnings levels. We used an age limit of 62 to ensure that the longitudinal sample cases would be younger than 65 on the date of the Round 7 interview. Of the 4,587 respondents in the Round 6 SWS, 3,712 were eligible for and included in the Round 7 longitudinal SWS.

Table 1. NBS–General Waves Round 7 actual sample sizes, target completed interviews, and completes

Sampling strata	Selected sample size	Original target completed interviewsa	Actual completed interviews
RBS			
Total	11,299	4,000	4,008
18- to 29-year-olds	3,237	1,111	1,127
30- to 39-year-olds	3,291	1,111	1,059
40- to 49-year-olds	3,060	1,111	1,118
50-year-olds or older	1,711	667	704
Cross-sectional SWS			
Total	8,590	3,000	3,017
SSDI only	4,221	1,500	1,493
SSI (SSI only + concurrent)	4,369	1,500	1,524
December 2018 extract	1,757	516	714
SSDI only	833	218	328
SSI (SSI only + concurrent)	924	298	386
January 2019 extract	1,438	456	592
SSDI only	747	222	305
SSI (SSI only + concurrent)	691	234	287
March 2019 extract	1,327	559	446
SSDI only	609	266	207
SSI (SSI only + concurrent)	718	293	239
April 2019 extract	1,043	394	339
SSDI only	545	215	175
SSI (SSI only + concurrent)	498	179	164

[21] Suspense status refers to the beneficiaries whose benefits have been temporarily suspended because of work or for other reasons.

[22] This threshold was $1,180 in 2017 and $1,220 in 2018.

Table 1 (*continued*)

Sampling strata	Selected sample size	Original target completed interviews[a]	Actual completed interviews
June 2019 extract	1,450	444	429
SSDI only	732	230	216
SSI (SSI only + concurrent)	718	214	213
July 2019 extract	998	348	319
SSDI only	468	193	161
SSI (SSI only + concurrent)	530	155	158
September 2019 extract	577	283	178
SSDI only	287	156	101
SSI (SSI only + concurrent)	290	127	77
Longitudinal SWS			
Total	3,712	2,040	2,078
SSDI only	1,863	1,019	1,080
SSI (SSI only + concurrent)	1,849	1,021	998

Source: NBS Round 7.

[a]The target completed interviews for the SWS shown here were calculated prior to receiving the first extract, using historical data from simulated successful worker populations in 2011–2012, 2013–2014, 2015–2016 and from Round 6 of the NBS. In fact, there were actually seven allocations, with a new sample allocation calculated after the population sizes for each previous extract were revealed. This explains the sometimes large deviation between the target allocation and the actual number of completed interviews.

2.   Calculation of nonresponse adjustments

Each observation had a base weight that accounted for the sample design.[23] Because the proportion of ineligible cases was small and there were few nonrespondents confirmed to be eligible, we treated ineligible cases as respondents when calculating nonresponse adjustments, then dropped those cases after adjusting the weights. For each sample, we calculated two adjustments to the weights to account for sample members who did not complete the questionnaire: a location adjustment to compensate for unlocated sample members and, among located cases, a response adjustment to compensate for those who refused to respond. The product of these adjustments, which constituted a nonresponse adjustment to the base weight, were intended to reduce the potential for bias attributable to differential ability to locate or respond, across levels of a set of auxiliary variables. In this report, we assess whether the adjustments successfully decreased the potential for bias or whether a potential for significant nonresponse bias still exists.

In the absence of information about how nonrespondents would have answered survey questions, we used data from three sources for this analysis: (1) administrative data from the sampling

---

[23] For the RBS, the base weight was the product of the inverse of the probability of selection, which was adjusted to account for telephone nonrespondents who were sampled for the second (field) phase. For the cross-sectional SWS, the base weight was the composite weight that combined the clustered and unclustered sample weights. For the longitudinal SWS, the base weight consisted of the final analysis weight from Round 6 adjusted for the different follow-up rules for the clustered and unclustered samples in Round 7.

frame provided by SSA; (2) earnings data from the DCF, also provided by SSA; and (3) data from the Area Health Resource File (AHRF), which contains demographic, health, and economic data for every county in the United States (Area Health Resource File 2018–2019). The administrative data included demographic characteristics about each beneficiary; whether the beneficiary received SSI, SSDI, or both; and the beneficiary's general disability classification and disability payment status, including why and to whom the payments were provided. The DCF earnings data included monthly earnings for each beneficiary for 2017 and 2018, though much of the earnings data, particularly from 2018, were not complete.[24] The AHRF data were used to classify the county where each beneficiary lived and included urbanicity and metropolitan status and information about the county's economic and racial and ethnic characteristics.

We used selected levels of a small number of these variables to calculate the nonresponse adjustments. In this analysis, however, we looked across all the levels for the variables of greatest interest, assuming correlation between them and key survey outcomes. We believe that these data provide an effective assessment of the potential for bias in this sample.

3. NBS–General Waves Round 7 data collection effort

Mathematica completed 9,103 interviews by the end of the Round 7 data collection. Of these, 4,008 were completed from the RBS; 3,017 from the cross-sectional SWS; and 2,078 from the longitudinal SWS. An additional 261 beneficiaries from the RBS, 310 successful workers from the cross-sectional SWS, and 46 longitudinal SWS cases were deemed ineligible for the survey.[25] Because of the independence of the sample selections for the RBS and the cross-sectional SWS, the clustered and unclustered samples within the cross-sectional SWS, and the Round 6 SWS (the source for the Round 7 longitudinal SWS), individuals could be selected for more than one sample. After accounting for 279 cases actually selected for more than one sample, the number of unique completed interviews was 8,824.[26] Mathematica completed all of these interviews by telephone.

---

[24] It would generally take approximately three years after the beneficiaries received the earnings for all monthly earnings data to be recorded in the DCF file. By 2020, when this analysis was conducted, the 2017 earnings data were mostly complete, but the 2018 earnings data were not.

[25] Ineligible sample members include those who were deceased, incarcerated, in active military, or no longer living in the continental United States and those whose benefit status was pending at the time of the interview. For the cross-sectional SWS, ineligibles also included sample members who had not worked in the past six months at the time of the interview.

[26] Among sample cases that were completed interviews only, there were 23 duplicates (46 sample cases total) between the RBS and cross-sectional SWS and 76 duplicates (152 sample cases total) between the clustered and unclustered samples within the cross-sectional SWS. Duplicates and triplicates also occurred with the longitudinal SWS.

The weighted response rates for Round 7 of the NBS were 54.7 percent for the RBS, 41.0 percent for the cross-sectional SWS, and 54.5 percent for the longitudinal SWS (see Table 2).[27] These response rates, along with the response rates broken out by the substrata for the RBS (age category); the cross-sectional SWS (extract and beneficiary title, with concurrent and SSI only collapsed); and longitudinal SWS (beneficiary title as defined in Round 6, with concurrent and SSI only collapsed) are provided in Table 2.

Table 2. Sample sizes and response rates, overall and by strata

| | Sample (unweighted counts) | | | | | Response rate (percent)[a] |
| | | | Nonrespondents | | | |
	Total sample	Respondents	Located	Unlocated	Ineligibles	Weighted
RBS	**11,299**	**4,008**	**5,177**	**1,853**	**261**	**54.7**
Age 18–29	3,237	1,127	1,470	576	64	51.9
Age 30–39	3,291	1,059	1,557	605	70	51.0
Age 40–49	3,060	1,118	1,430	442	70	50.7
Age 50–65	1,711	704	720	230	57	56.6
Cross-sectional SWS	**8,590**	**3,017**	**3,158**	**787**	**311**	**41.0**
Dec. 2018	1,757	714	595	92	82	48.9
Jan. 2019	1,438	592	510	100	56	47.6
Mar. 2019	1,327	446	555	173	37	38.2
Apr. 2019	1,043	339	330	73	42	39.2
June 2019	1,450	429	582	178	44	33.7
July 2019	998	319	361	106	32	38.3
Sept. 2019	577	178	225	65	18	37.1
SSDI only	4,221	1,524	1,548	452	151	40.5
SSI	4,369	1,493	1,610	335	160	41.7
Longitudinal SWS	**3,712**	**2,078**	**363**	**1,204**	**46**	**54.5**
SSDI only	1,863	1,080	189	563	18	56.7
SSI	1,849	998	174	641	28	50.7

Source: NBS Round 7.

[a]Response rates are calculated by taking the weighted number of respondents and ineligibles as the numerator and dividing by the total number of sample members. Because the eligibility of very few nonrespondents is known, the response rate calculation is close to a more commonly used response rate calculation: numerator = number of respondents and denominator = number of respondents + number of eligible nonrespondents + eligibility rate * number of nonrespondents with unknown eligibility. In subpopulations where a dual sample design was used, we did not include some sample cases in the denominator. Details are beyond the scope of this report but may be found in the Editing, Coding, Imputation, and Weighting Report (Grau et al. 2021).

4. Rationale for nonresponse bias analysis

The purpose of the nonresponse bias analysis was to determine if there were systematic differences between respondents and nonrespondents that could result in the potential for

---

[27] We expected response rates for the SWS to be very low, particularly for later extracts (as shown in Table 2), given the short amount of time available for data collection.

nonresponse bias for either sample, and to determine whether the nonresponse adjustments to the weights addressed those differences.

B. Response rates

As indicated previously, the beneficiary population includes all SSI or SSDI beneficiaries age 18 to FRA in active pay status as of June 30, 2018. The total number of beneficiaries in the target population (excluding U.S. territories) was 13,670,658, with some data missing for items of interest. In Table 2, we present the total number of beneficiaries sampled and the number of respondents, nonrespondents, and sample members who were ineligible due to death, incarceration, or other reasons, by stratum. In addition, we present the weighted response rates by using the base weight. Note that the base weight used in the RBS response rate calculation was already adjusted for the second phase sample selection. We only present the weighted response rates because (1) with two-phase sampling, the unweighted rates are not meaningful;[28] (2) the sampling rates—and thus the sampling weights—vary substantially across the sampling strata (as seen in Table III.2); and (3) the weighted rates better reflect the potential for nonresponse bias. The weighted response rates ranged from a low of 50.7 percent for 40- to 49-year-olds to a high of 56.6 percent for those who were 50 years old or older.

The successful worker population included successful workers within the beneficiary frame described above, provided they could be identified. For the cross-sectional SWS, the total number of successful workers in the target population was 101,698. The weighted response rates ranged from a low of 33.7 percent in the fifth extract (June 2019) to a high of 48.9 percent in the first extract (December 2018). Weighted response rates in the later extracts were much lower than in earlier extracts due to the shortened data collection period, but they did not differ much between the SSDI-only stratum and the SSI stratum. Table 2 also presents the total number of successful workers within beneficiary title strata.

The longitudinal successful workers shown in Table 2 constituted the Round 7 longitudinal sample. These sample members included successful workers identified in Round 6 who had responded affirmatively to Question B24 in Round 6, "Are you currently working at a job or business for pay or profit?" This restriction removed people who had been working within six months of the Round 6 interview but were not working at the time of the Round 6 interview. We do not know what proportion of the 89,636 successful workers in Round 6 were working at the time of the Round 6 interview, which would provide us with the size of the longitudinal SWS population. However, we have an estimate based on weighted totals from our responding longitudinal sample (65,871), of which 64,225 were eligible. However, after we processed an updated extract from Round 6, we found that there was a total of 288,576 successful workers, of

---

[28] If we included the second-phase-eligible cases that were not selected for the second phase, the unweighted response rate would be too low and it would not reflect the fact that the cases' base weights were transferred to other sample members. If we excluded these cases, the unweighted response rate would be too high and it would not reflect the unsuccessful effort to get a response from these cases in the first phase.

which 265,514 were eligible. We poststratified the Round 6 weights to this new total; however, we still need to recalculate the longitudinal weights to determine an estimated size of the eligible longitudinal population.[29]

We did not present a breakdown by extract in Table 2 because the numbers within each extract were small and the variable was of limited analytic utility. The weighted response rates ranged from 50.7 percent among successful workers who were beneficiaries of SSI to 56.7 percent among successful workers who were beneficiaries of SSDI only.

C. Methodology

The nonresponse bias analysis used data on individual members of the sampling frame (for the RBS and cross-sectional SWS) and sample (for all three samples). The following were the variables that we used in the analysis (all categorical):

1. Age category

    -   18 to 29

    -   30 to 39

    -   40 to 49

    -   50 to FRA

2. Gender

    -   Male

    -   Female

3. Beneficiary type

    -   SSI only

    -   SSDI only

    -   Both SSI and SSDI

4. Race and ethnicity

    -   Non-Hispanic white

    -   Non-Hispanic black

    -   Non-Hispanic Asian

    -   Non-Hispanic American Indian

---

[29] After we conducted a final extract of Round 6 earnings data in November 2020, we determined that the estimated number of eligible successful workers in Round 6 was actually 265,514; the discrepancy was due to a lag in recording earnings for many successful workers. We will need to redo the longitudinal weights to account for this new total and obtain a new estimate of successful workers who were eligible for the longitudinal population.

- Non-Hispanic other

- Hispanic

5. Constructed disability status

    - Hearing disability

    - Cognitive disability

    - Mental illness

    - Physical disability

    - Disability not given

6. Racial and ethnic characteristics of beneficiary's county

    - County with plurality or majority non-Hispanic black population

    - County with plurality or majority Hispanic population

    - County with racially and ethnically mixed population, no majority group

    - County with majority but less than 90 percent non-Hispanic white population

    - County with at least 90 percent non-Hispanic white population

7. Economic characteristics of county (10 overlapping levels, each listed as binary variables)

    - Government-dependent economy county[30]

    - Recreation-dependent economy county[31]

    - Nonspecialized-dependent economy county[32]

    - Manufacturing-dependent economy county[33]

    - Counties with high levels of children living in persistent poverty[34]

---

[30] Fifteen percent or more of average annual labor and proprietors' earnings derived from federal and state government from 2010 to 2012.

[31] This indicator is based on three sources: (1) percentage of wage and salary employment in entertainment and recreation, accommodations, eating and drinking places, and real estate as a percentage of all employment reported by the Bureau of Economic Analysis; (2) percentage of total personal income reported for these same categories by the Bureau of Economic Analysis; and (3) percentage of vacant housing units intended for seasonal or occasional use reported in the 2010 census. The AHRF documentation does not provide the specific percentage from these three sources.

[32] County did not meet the dependence threshold for service, government, farming, mining, or manufacturing.

[33] Twenty-three percent or more of the county's average annual labor and proprietors' earnings were derived from manufacturing, or 16 percent or more of jobs were in manufacturing.

[34] Twenty percent or more of the county's children under age 18 were persistently poor, based on census data from 1980, 1990, and 2000, and recent data from the American Community Survey.

- Counties with high levels of poverty[35]

- Counties with persistent levels of poverty[36]

- County with low education[37]

- Population-loss county[38]

- Retirement-destination county[39]

8. Metropolitan status of county

- Metropolitan area of 1 million population or more

- Metropolitan area of 250,000 to 999,999 population

- Metropolitan area of fewer than 250,000 population

- Nonmetropolitan area adjacent to large metropolitan area

- Nonmetropolitan area adjacent to medium or small metropolitan area

- Nonmetropolitan area not adjacent to metropolitan area

9. Geographic region (U.S. census region) of beneficiary's residence

- West

- South

- Northeast

- Midwest

10. Geographic region (U.S. census division) of beneficiary's residence

- East North Central

- West North Central

- New England

- Middle Atlantic

- South Atlantic

---

[35] Twenty percent or more of residents were poor, based on recent data from the American Community Survey.

[36] Twenty percent or more of residents were persistently poor, based on census data from 1980, 1990, and 2000, and recent data from the American Community Survey.

[37] Twenty-five percent or more of residents age 25 to 64 had neither a high school diploma nor GED, as determined by the American Community Survey, 5-year average data for 2008 to 2012.

[38] Number of residents declined both between the 1990 and 2000 censuses and between the 2000 and 2010 censuses.

[39] Number of residents 60 and older grew by 15 percent or more between 2000 and 2010 due to net in-migration.

- East South Central

- West South Central

- Mountain

- Pacific

11. Earnings category

    a. Earnings category levels used for RBS [40]

        o Three consecutive months of earnings above SGA[41] at some point in 2017 or 2018

        o At least one month of earnings above $7,000 in 2017 or 2018

        o At least one month of earnings above $2,000 in 2017 or 2018

        o At least one month of earnings above $0 in 2017 or 2018

        o No monthly earnings in 2017 or 2018

    b. Earnings category (levels used for cross-sectional SWS)[42]

        o At least one month of earnings above $30,000 in 2017 or 2018

        o At least one month of earnings above $20,000 in 2017 or 2018

        o At least one month of earnings above $15,000 in 2017 or 2018

        o At least one month of earnings above $7,000 in 2017 or 2018

        o Did not meet the earnings thresholds given above

    c. Earnings category (levels used for longitudinal SWS)[43]

        o At least one month of earnings above $30,000 in 2015 or 2016

        o At least one month of earnings above $20,000 in 2015 or 2016

---

[40] We arrived at the five categories used for the earnings variable in the RBS after a lengthy investigation in Round 5 using both (annual) Internal Revenue Service (IRS) and (monthly) DCF earnings. Using data from the 2014 sampling frame, we calculated the percentage with positive IRS earnings in 2014 (considered "working"), as well as the mean and median IRS 2014 earnings, both overall and among those who were working. We compared these values to several sets of post-stratified weights, where the raking was based on a variety of earnings-related categorical variables—each with different cut points; some with IRS earnings and some with DCF earnings. We determined that, although the IRS earnings are more accurate than DCF earnings, IRS earnings are only available annually, which raised timing issues and diluted the advantage of accuracy. It was also more difficult to use IRS earnings because only SSA staff could access them. We arrived at the cut points given above because using them resulted in estimated annual earnings that were closest to the IRS values. We used the same cut points for the earnings category variable used for the RBS in Round 7.

[41] The monthly non-blind SGA earnings level was $1,170 in 2017 and $1,180 in 2018.

[42] We arrived at the five categories used for the earnings category variable in the SWS by looking at the distribution of earnings across all successful workers and creating cut points that resulted in five roughly equal categories.

[43] We arrived at the five categories used for the earnings category variable in the SWS by looking at the distribution of earnings across all successful workers and creating cut points that resulted in five roughly equal categories.

- o At least one month of earnings above $15,000 in 2015 or 2016

- o At least one month of earnings above $7,000 in 2015 or 2016

- o Did not meet the earnings thresholds given above

12. Extract (for cross-sectional and longitudinal SWS only)

For the RBS and cross-sectional SWS, these variables were defined according to their frame values from Round 7. However, for the longitudinal SWS, most variables were defined according to their Round 6 values, except for age, which was recalculated using the sample members' dates of birth, and the geographic variables, which were determined based upon their residence in Round 7, if available.

After examining the extent of missing data for the above variables in the sampling frame (Table 3), we used the base weights to compare the distributions of the variables across the frame; the total sample; and the sample split into two parts, the ineligibles and the remainder of the sample with ineligibles removed (Tables 4 and 5 for the RBS and cross-sectional SWS, respectively). We made similar comparisons for the longitudinal SWS (Table 6), though we did not include frame values because no frame was available for the longitudinal SWS. In Tables 7 (RBS), 8 (cross-sectional SWS), and 9 (longitudinal SWS), we compared the distributions of variables between the respondents (with ineligibles) and nonrespondents. We then compared the distributions among respondents with ineligibles by using nonresponse-adjusted weights against the original sample with the original sample weights (Table 10 for the RBS, Table 11 for the cross-sectional SWS, and Table 12 for the longitudinal SWS).

In each table, we used SUDAAN to calculate standard errors in order to accommodate the sample design.[44] The sample statistics consisted of proportions with an attribute (presented as percentages). We conducted comparisons for all beneficiaries. Several variables had missing values in the sample frame. In particular, in the beneficiary frame, race and ethnicity and disability type had missing values. In each case, the proportions with each attribute that were used in the following analyses were calculated among cases without missing data.

Table 3. Percentage of missing values for variables of interest

| Variable | Weighted percent missing[a] | | | | |
	In frame	In entire sample	Among respondents	Among nonrespondents	Among ineligibles
RBS					
Race/ethnicity	12.6	12.8	12.3	13.1	15.9
Disability status	2.0	2.1	1.9	1.8	7.3
Cross-sectional SWS					

[44] SUDAAN is a statistical package that was developed specifically for survey data by Research Triangle Institute, now called RTI International. A hard copy manual is available for Version 9.0 (Research Triangle Institute 2004). An online version is available for Version 11.0 (see www.rti.org/sudaan).

Variable	Weighted percent missing[a]				
	In frame	In entire sample	Among respondents	Among nonrespondents	Among ineligibles
Race/ethnicity	22.6	23.4	23.2	23.5	24.0
Disability status	1.0	0.9	0.8	0.9	1.7
Longitudinal SWS					
Race/ethnicity	Not available	20.0	17.2	23.1	25.6
Disability status	Not available	1.1	1.0	1.2	1.5

Source: NBS Round 7.

[a]The weights in the table are the base weights. None of the other variables have missing values in the sampling frame.

As is apparent from Table 3, the level of missingness for race and ethnicity was high, with approximately 13 percent of the frame missing this variable among all beneficiaries and approximately 22 percent of the frame among Round 7 (cross-sectional) successful workers. The estimate of the proportion missing race/ethnicity in the longitudinal SWS was 20 percent. Any conclusions drawn from race and ethnicity therefore must be viewed with caution.

D. Results

In Tables 4, 5, and 6, we compared sample statistics of the variables for all sampled cases in the RBS, cross-sectional SWS, and longitudinal SWS, respectively. The values are percentages for each level of the categorical variables, with the associated standard errors (se) in parentheses. The frame values (for the RBS and cross-sectional SWS) do not have a standard error because they represent the original population and are without sampling error (no frame values are available for the longitudinal SWS). Unknown categories were not included in the levels for these variables; proportions were calculated for the cases without missing data.[45] In all three tables, we applied base weights to sample values for all columns except the frame percentage for Tables 4 and 5, for which no weights were required (percentages calculated by using the entire population).

We compared two types of variables. Greater emphasis was placed on the variables that were likely to be correlated with important outcome variables: beneficiary type, disability type, demographic variables, and the categorical earnings variable. Other variables were less likely to be highly correlated with outcome variables and thus received less emphasis: geographic and economic characteristics associated with the beneficiary's county.

---

[45] Values are assumed to be missing completely at random (MCAR). Although MCAR is normally a strong assumption, the level of missingness was so small for all but race and ethnicity that deviations from this assumption would not significantly change the conclusions.

Table 4. Percentages with various attributes (categorical variables) in the RBS, using base weights

Variable[a]	Frame percent	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
		Percent	(se)	Percent	(se)	Percent	(se)
Beneficiary type							
SSI only	29.9	30.0	(1.3)	29.1	(1.3)	50.7	(7.0)*
SSDI only	56.6	55.8	(1.4)	56.5	(1.5)	40.7	(6.7)*
Both SSI and SSDI	13.5	14.2	(0.8)	14.4	(0.8)	8.6	(2.4)*
Constructed disability status							
Hearing	0.4	0.4	(0.1)	0.3	(0.1)	0.7	(0.6)
Cognitive	11.9	11.0	(0.7)	11.2	(0.8)	4.8	(1.4)*
Mental	29.6	29.9	(1.5)	30.2	(1.5)	22.0	(4.6)
Physical	58.1	58.8	(1.6)	58.3	(1.6)	72.5	(4.9)*
Sex							
Female	49.1	49.2	(1.3)	49.6	(1.4)	39.8	(7.0)
Male	50.9	50.8	(1.3)	50.4	(1.4)	60.2	(7.0)
Beneficiary's age							
18–29 years	9.9	9.9	(0.3)	9.9	(0.4)	7.4	(1.4)
30–39 years	10.7	10.7	(0.4)	10.7	(0.4)	8.9	(1.9)
40–49 years	15.2	15.2	(0.4)	15.4	(0.5)	11.2	(2.2)
50 years–FRA[b]	64.2	64.2	(0.8)	63.9	(0.9)	72.4	(4.4)
Race/ethnicity							
White	65.9	65.4	(2.7)	65.5	(2.8)	64.9	(7.8)
Black	22.7	24.5	(2.5)	24.4	(2.5)	26.9	(7.7)
Hispanic	3.6	3.3	(0.8)	3.3	(0.8)	3.5	(1.7)
All others	7.8	6.7	(0.9)	6.8	(1.0)	4.7	(2.2)
County racial/ethnic profile							
County with plurality or majority non-Hispanic black population	4.7	3.9	(2.2)	3.9	(2.2)	4.0	(3.3)
County with plurality or majority Hispanic population	9.5	7.4	(2.5)	7.4	(2.4)	7.9	(3.5)

Table 4 (*continued*)

Variable[a]	Frame percent	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
		Percent	(se)	Percent	(se)	Percent	(se)
County with majority but less than 90% non-Hispanic white population	37.9	46.0	(5.2)	46.0	(5.3)	44.0	(8.2)
County with racially/ethnically mixed population, no majority group	34.4	34.2	(5.1)	34.1	(5.2)	37.4	(8.0)
County with at least 90% non-Hispanic white population	12.9	8.5	(2.7)	8.6	(2.7)	6.8	(3.3)
Economic characteristics of county[c]							
Government-dependent economy county	11.9	10.5	(3.2)	10.6	(3.3)	9.5	(4.2)
Manufacturing-dependent economy county	9.8	8.9	(2.7)	8.8	(2.7)	9.6	(5.1)
Nonspecialized-dependent economy county	65.3	67.7	(4.8)	67.9	(4.8)	63.5	(8.6)
Recreation-dependent economy county	8.2	9.1	(3.1)	8.8	(3.1)	14.9	(7.7)
County with high levels of poverty	18.5	12.2	(3.5)	12.4	(3.5)	5.7	(2.5)*
County with high levels of persistent poverty	8.3	4.4	(2.1)	4.5	(2.1)	2.2	(1.3)*
County with high levels of persistent child poverty	19.1	15.0	(3.9)	14.8	(3.8)	20.8	(7.2)
County with low education	13.3	11.9	(3.1)	11.7	(3.0)	17.8	(7.7)
Population-loss county	8.7	3.6	(1.8)*	3.7	(1.9)*	0.3	(0.2)
Retirement-destination county	14.1	14.9	(3.6)	14.6	(3.5)	23.5	(8.9)
Metropolitan status of county							
Metropolitan area of 1 million population or more	46.5	43.9	(5.4)	43.9	(5.4)	42.9	(8.7)
Metropolitan area of 250,000 to 999,999 population	22.6	27.2	(4.9)	27.2	(4.9)	26.8	(6.7)
Metropolitan area of fewer than 250,000 population	10.7	13.0	(3.7)	13.0	(3.7)	14.0	(7.5)
Nonmetropolitan area adjacent to large metropolitan area	4.4	3.8	(1.9)	3.9	(1.9)	2.9	(1.9)
Nonmetropolitan area adjacent to medium or small metropolitan area	9.0	8.8	(2.9)	8.8	(2.9)	8.8	(5.6)
Nonmetropolitan area not adjacent to metropolitan area	6.9	3.2	(1.7)*	3.2	(1.7)*	4.5	(2.7)
Census region							
West	18.5	17.3	(4.0)	17.5	(4.1)	13.8	(4.8)
South	41.4	42.4	(5.6)	41.8	(5.6)	56.5	(8.4)
Northeast	18.2	18.8	(4.4)	19.0	(4.5)	12.1	(4.7)
Midwest	21.9	21.5	(4.5)	21.7	(4.5)	17.7	(5.5)

Table 4 (*continued*)

Variable[a]	Frame percent	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
		Percent	(se)	Percent	(se)	Percent	(se)
Census division							
East North Central	15.7	15.3	(3.9)	15.4	(3.9)	12.5	(4.6)
West North Central	6.2	6.2	(2.7)	6.2	(2.7)	5.2	(3.0)
New England	5.0	5.1	(2.5)	5.1	(2.5)	5.9	(3.4)
Middle Atlantic	13.2	13.7	(3.9)	14.0	(4.0)	6.1	(3.2)*
South Atlantic	20.5	20.1	(4.7)	20.0	(4.7)	22.1	(8.0)
East South Central	9.2	10.2	(3.5)	10.3	(3.5)	8.1	(4.0)
West South Central	11.8	12.0	(3.6)	11.4	(3.5)	26.3	(9.4)
Mountain	5.7	5.9	(2.6)	6.0	(2.7)	4.1	(2.2)
Pacific	12.7	11.4	(3.3)	11.5	(3.3)	9.7	(4.2)
Earnings categories for 2017–2018 time period							
Three consecutive months of monthly earnings above SGA	3.2	4.2	(0.6)	4.3	(0.6)	1.3	(1.1)
Monthly earnings above $7,000 in at least one month in '17 or '18	1.4	1.8	(0.4)	1.9	(0.4)	0.1	(0.1)*
Monthly earnings above $2,000 in at least one month in '17 or '18	2.4	3.0	(0.4)	2.8	(0.3)	8.2	(6.6)
Monthly earnings above zero in at least one month in '17 or '18	2.6	3.0	(0.4)	3.1	(0.4)	0.9	(0.9)
Monthly earnings always zero in 2017 and 2018	90.4	87.9	(0.9)*	87.8	(0.9)*	89.5	(6.5)

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

*Denotes a difference between the sample and frame value of more than two standard errors.

Table 5. Percentages with various attributes (categorical variables) in the cross-sectional SWS, using base weights

Variable[a]	Frame Percent	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
		Percent	(se)	Percent	(se)	Percent	(se)
Beneficiary type							
SSI only	26.0	25.7	(0.9)	25.7	(0.9)	23.7	(2.8)
SSDI only	52.5	52.5	(1.1)	52.6	(1.1)	50.9	(3.6)
Both SSI and SSDI	21.5	21.8	(0.6)	21.7	(0.6)	25.3	(2.9)
Constructed disability status							
Hearing	2.0	1.7	(0.2)	1.7	(0.2)	3.4	(1.3)
Cognitive	14.2	14.5	(0.6)	14.5	(0.6)	15.5	(2.2)
Mental	36.2	36.1	(1.0)	36.3	(1.0)	30.7	(3.0)
Physical	47.6	47.6	(1.1)	47.5	(1.1)	50.4	(3.2)
Sex							
Female	45.9	44.8	(0.7)	44.6	(0.7)	49.6	(3.4)
Male	54.1	55.2	(0.7)	55.4	(0.7)	50.4	(3.4)
Beneficiary's age							
18–29 years	22.7	23.0	(0.7)	23.3	(0.8)	16.2	(2.4)*
30–39 years	23.9	24.5	(0.7)	24.8	(0.7)	18.6	(2.3)*
40–49 years	22.4	21.7	(0.6)	21.6	(0.6)	23.8	(2.5)
50 years–FRA[b]	30.9	30.7	(0.8)	30.3	(0.8)	41.4	(3.3)*
Race/ethnicity							
White	54.9	55.8	(2.3)	56.2	(2.3)	46.1	(3.9)*
Black	31.4	32.0	(2.2)	31.5	(2.2)	43.2	(3.8)*
Hispanic	4.9	4.5	(0.5)	4.6	(0.5)	2.5	(0.9)*
All others	8.8	7.8	(0.7)	7.7	(0.7)	8.2	(2.1)
County racial/ethnic profile							
County with plurality or majority non-Hispanic black population	4.6	3.4	(1.6)	3.3	(1.7)	4.5	(1.7)
County with plurality or majority Hispanic population	9.7	8.4	(2.1)	8.4	(2.1)	8.1	(2.6)

Table 5 (*continued*)

Variable[a]	Frame Percent	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
		Percent	(se)	Percent	(se)	Percent	(se)
County with majority but less than 90% non-Hispanic white population	38.7	45.5	(4.8)	45.7	(4.8)	39.1	(5.2)
County with racially/ethnically mixed population, no majority group	36.5	34.3	(4.6)	34.0	(4.6)	41.9	(5.4)
County with at least 90% non-Hispanic white population	10.2	8.4	(2.4)	8.5	(2.5)	5.9	(2.0)*
Economic characteristics of county[c]							
Government-dependent economy county	12.4	11.9	(3.2)	11.9	(3.2)	11.8	(3.3)
Manufacturing-dependent economy county	8.2	8.9	(2.8)	8.9	(2.8)	8.6	(3.5)
Nonspecialized-dependent economy county	67.3	66.5	(4.5)	66.4	(4.6)	69.0	(4.8)
Recreation-dependent economy county	8.3	9.3	(2.9)	9.5	(2.9)	5.1	(1.7)
County with high levels of poverty	14.4	10.6	(2.8)	10.5	(2.8)	14.1	(3.5)
County with high levels of persistent poverty	6.1	3.9	(1.6)	3.8	(1.6)	5.8	(2.6)
County with high levels of persistent child poverty	15.2	12.3	(3.0)	12.2	(3)	14.9	(3.8)
County with low education	11.6	12.2	(2.8)	12.0	(2.8)	16.4	(4.4)
Population-loss county	8.1	4.2	(1.8)*	4.0	(1.8)*	8.6	(3.1)
Retirement-destination county	12.4	13.1	(3.0)	13.1	(3.1)	12.4	(3.6)
Metropolitan status of county							
Metropolitan area of 1 million population or more	54.0	52.9	(4.8)	52.7	(4.8)	57.2	(5.3)
Metropolitan area of 250,000 to 999,999 population	21.9	24.3	(4.0)	24.5	(4.0)	19.8	(4.3)
Metropolitan area of fewer than 250,000 population	9.8	11.5	(3.1)	11.6	(3.1)	10.4	(3.8)
Nonmetropolitan area adjacent to large metropolitan area	3.3	3.7	(1.5)	3.7	(1.5)	3.5	(1.7)
Nonmetropolitan area adjacent to medium or small metropolitan area	6.2	4.8	(1.4)	4.8	(1.4)	4.3	(2.2)
Nonmetropolitan area not adjacent to metropolitan area	4.8	2.8	(1.0)*	2.7	(1.0)*	4.9	(2.3)
Census region							
West	22.3	22.9	(4.2)	23.1	(4.3)	18.5	(4.0)
South	33.5	33.2	(4.5)	32.9	(4.5)	42.5	(5.5)
Northeast	21.3	21.3	(4.2)	21.4	(4.2)	19.2	(4.0)
Midwest	23.0	22.6	(4.1)	22.7	(4.1)	19.8	(4.1)
Census division							
East North Central	16.0	15.2	(3.4)	15.2	(3.4)	15.6	(3.7)

Table 5 (*continued*)

Variable[a]	Frame Percent	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
		Percent	(se)	Percent	(se)	Percent	(se)
West North Central	7.0	7.3	(2.7)	7.5	(2.8)	4.3	(2.0)
New England	6.7	6.2	(2.6)	6.2	(2.6)	4.8	(2.3)
Middle Atlantic	14.5	15.1	(3.7)	15.2	(3.7)	14.4	(3.5)
South Atlantic	16.3	14.6	(3.2)	14.6	(3.2)	16.7	(4.1)
East South Central	6.0	7.2	(2.4)	7.2	(2.4)	6.6	(2.4)
West South Central	11.1	11.4	(3.1)	11.1	(3.0)	19.1	(5.0)
Mountain	6.0	6.7	(2.5)	6.7	(2.5)	6.3	(2.5)
Pacific	16.3	16.2	(3.8)	16.4	(3.8)	12.2	(3.3)
Earnings categories for 2017–2018 time period							
Monthly earnings above $30,000 in at least one month in '17 or '18	22.3	22.7	(0.6)	23.0	(0.7)	16.7	(2.3)*
Monthly earnings above $20,000 in at least one month in '17 or '18	23.5	23.7	(0.5)	23.7	(0.6)	24.1	(2.5)
Monthly earnings above $15,000 in at least one month in '17 or '18	19.3	19.5	(0.5)	19.4	(0.5)	19.6	(2.3)
Monthly earnings above $7,000 in at least one month in '17 or '18	21.4	21.2	(0.5)	21.1	(0.5)	22.5	(2.6)
Monthly earnings did not meet the thresholds above	13.5	12.9	(0.5)	12.7	(0.5)	17.1	(2.3)
Extract							
December 2018 extract	19.5	19.5	(0.6)	19.3	(0.5)	24.7	(3.1)
January 2019 extract	15.9	15.9	(0.4)	15.8	(0.4)	17.7	(2.3)
March 2019 extract	16.2	16.2	(0.4)	16.4	(0.4)	12.0	(2.3)
April 2019 extract	12.0	12.0	(0.4)	11.9	(0.4)	14.2	(2.1)
June 2019 extract	17.0	17.0	(0.5)	17.1	(0.5)	14.8	(2.5)
July 2019 extract	10.9	10.9	(0.3)	11.0	(0.4)	9.5	(1.8)
September 2019 extract	8.4	8.4	(0.3)	8.5	(0.4)	7.1	(1.9)

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

*Denotes a difference between the sample and frame value of more than two standard errors.

Table 6. Percentages with various attributes (categorical variables) in the longitudinal SWS, using base weights

Variable[a]	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
	Percent	(se)	Percent	(se)	Percent	(se)
Round 6 beneficiary type						
SSI only	27.8	(0.9)	27.8	(0.9)	30.5	(7.4)
SSDI only	51.1	(1.0)	51.2	(1.0)	42.3	(8.1)
Both SSI and SSDI	21.1	(0.8)	21.0	(0.8)	27.2	(7.1)
Round 6 constructed disability status						
Hearing	3.1	(0.4)	3.1	(0.4)	4.4	(4.3)
Cognitive	12.8	(0.7)	12.9	(0.7)	5.7	(3.6)
Mental	36.6	(1.0)	36.6	(1.0)	33.3	(8.8)
Physical	47.5	(1.0)	47.4	(1.0)	56.6	(9.0)
Sex						
Female	46.4	(1.1)	45.4	(1.1)	44.8	(8.8)
Male	54.6	(1.1)	54.6	(1.1)	55.2	(8.8)
Beneficiary's age						
18–29 years	24.8	(0.9)	24.9	(0.9)	18.4	(6.1)
30–39 years	23.0	(0.8)	23.0	(0.8)	19.5	(6.6)
40–49 years	21.6	(0.7)	21.5	(0.7)	25.1	(7.7)
50 years–FRA[b]	30.7	(0.9)	30.6	(0.9)	37.0	(8.5)
Race/ethnicity						
White	57.6	(2.0)	57.7	(2.0)	45.4	(9.8)
Black	29.7	(1.8)	29.6	(1.8)	34.3	(9.4)
Hispanic	5.4	(0.6)	5.3	(0.6)	14.7	(6.5)
All others	7.4	(0.7)	7.4	(0.7)	5.6	(4.1)
County racial/ethnic profile						
County with plurality or majority non-Hispanic black population	3.3	(1.1)	3.3	(1.1)	1.5	(1.5)
County with plurality or majority Hispanic population	8.2	(1.4)	8.2	(1.5)	7.7	(3.5)
County with majority but less than 90% non-Hispanic white population	43.7	(3.7)	43.5	(3.7)	54.3	(8.6)

Table 6 (*continued*)

Variable[a]	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
	Percent	(se)	Percent	(se)	Percent	(se)
County with racially/ethnically mixed population, no majority group	35.5	(3.6)	35.6	(3.6)	33.4	(7.7)
County with at least 90% non-Hispanic white population	9.1	(1.9)	9.2	(1.9)	3.1	(3.0)
Economic characteristics of county[c]						
Government-dependent economy county	12.7	(2.8)	12.7	(2.8)	15.3	(7.5)
Manufacturing-dependent economy county	8.1	(1.9)	8.2	(1.9)	3.2	(3.1)
Nonspecialized-dependent economy county	66.8	(3.5)	66.9	(3.5)	62.6	(9.4)
Recreation-dependent economy county	9.3	(2.0)	9.1	(2.0)	19.0	(8.5)
County with high levels of poverty	11.1	(1.9)	11.0	(1.9)	14.3	(5.1)
County with high levels of persistent poverty	4.7	(1.4)	4.7	(1.4)	6.9	(3.5)
County with high levels of persistent child poverty	12.4	(2.2)	12.3	(2.2)	18.1	(5.7)
County with low education	11.3	(2.0)	11.4	(2)	7.0	(3.5)
Population-loss county	4.9	(1.2)	4.9	(1.2)	4.3	(3.0)
Retirement-destination county	12.4	(2.3)	12.2	(2.3)	24.1	(8.4)
Metropolitan status of county						
Metropolitan area of 1 million population or more	52.3	(3.7)	52.3	(3.7)	56.0	(9.1)
Metropolitan area of 250,000 to 999,999 population	24.5	(3.4)	24.6	(3.4)	22.2	(8.4)
Metropolitan area of fewer than 250,000 population	10.9	(2.2)	10.9	(2.3)	11.9	(5.5)
Nonmetropolitan area adjacent to large metropolitan area	2.9	(0.7)	2.8	(0.7)	8.0	(6.5)
Nonmetropolitan area adjacent to medium or small metropolitan area	5.8	(1.4)	5.9	(1.4)	0.0	(0.0)*
Nonmetropolitan area not adjacent to metropolitan area	3.6	(0.6)	3.6	(0.6)	1.9	(1.9)
Census region						
West	22.8	(3.2)	22.6	(3.2)	42.0	(9.2)*
South	31.1	(3.3)	31.2	(3.3)	24.2	(7.2)
Northeast	23.3	(3.2)	23.4	(3.3)	12.5	(4.6)*
Midwest	22.7	(3.1)	22.7	(3.1)	21.4	(6.8)
Census division						
Pacific	15.8	(2.7)	15.8	(2.7)	13.8	(5.7)
East North Central	15.0	(2.5)	15.0	(2.5)	15.9	(5.9)

Table 6 (*continued*)

Variable[a]	Entire sample		Sample with known ineligibles removed		Sampled ineligibles	
	Percent	(se)	Percent	(se)	Percent	(se)
Middle Atlantic	16.2	(2.8)	16.3	(2.9)	10.0	(4.2)
South Atlantic	15.7	(2.5)	15.6	(2.5)	19.2	(6.2)
West South Central	9.2	(1.9)	9.2	(1.9)	4.3	(4.2)
East South Central	6.3	(2.0)	6.4	(2.0)	0.7	(0.7)*
Mountain	7.1	(2.1)	6.8	(2.1)	28.2	(9.4)*
New England	7.1	(2.0)	7.1	(2.0)	2.4	(1.7)*
West North Central	7.7	(2.2)	7.7	(2.2)	5.5	(3.8)
Earnings categories for 2015–2016 time period						
Monthly earnings above $30,000 in at least one month in '15 or '16	20.9	(0.8)	20.7	(0.8)	36.4	(8.5)
Monthly earnings above $20,000 in at least one month in '15 or '16	23.0	(1.0)	23.0	(0.9)	23.4	(6.9)
Monthly earnings above $15,000 in at least one month in '15 or '16	19.2	(0.8)	19.2	(0.8)	18.7	(6.4)
Monthly earnings above $7,000 in at least one month in '15 or '16	24.0	(0.9)	24.1	(0.9)	17.2	(6.0)
Monthly earnings did not meet the thresholds above	12.9	(0.7)	13.0	(0.7)	4.3	(3.7)*
Extract						
December 2016 extract	19.7	(0.7)	19.6	(0.7)	28.9	(8.4)
January 2017 extract	14.7	(0.6)	14.6	(0.6)	23.1	(7.0)
March 2017 extract	19.0	(0.7)	19.0	(0.8)	17.2	(7.4)
April 2017 extract	12.2	(0.5)	12.3	(0.5)	4.9	(2.9)*
June 2017 extract	14.7	(0.6)	14.7	(0.6)	10.1	(4.7)
July 2017 extract	11.3	(0.5)	11.3	(0.5)	10.5	(5.7)
September 2017 extract	8.5	(0.3)	8.5	(0.3)	5.3	(2.8)

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

*Denotes a difference between the sample and frame value of more than two standard errors.

For each variable, approximate 95 percent confidence intervals were created by adding and subtracting two standard errors from each point estimate among the sample values. We did not account for the fact that these confidence intervals were considered simultaneously, which would increase the Type I error (the probability that the confidence interval did not include the true value, or the probability of rejecting the null hypothesis when it was true). Hence, one must consider this when significant results are observed.

1. Comparison of entire sample with frame

Before conducting a nonresponse analysis, we must determine if the sample distribution adequately matches the frame distribution on important variables. This is necessary to ascertain whether the estimates using the sampling weights produce estimates that are consistent with population values. As shown in Tables 4 and 5, the statistics estimated from the entire sample (using the base weight) among all beneficiaries were generally close to those computed with the full frame, although a few estimates—especially among the county-level variables defined from the AHRF—deviated from the frame value by more than two standard errors. (Those varying by more than two standard errors are denoted by an asterisk.) Because no frame statistics were available for the longitudinal SWS in Table 6, no such comparison was possible for this sample.

a. Representative beneficiary sample

Within PSUs, the RBS was selected within explicit strata defined by age category and implicit strata defined by disability status; beneficiary title (SSI, SSDI, or both); race and ethnicity; gender; and zip code, in that priority order.[46] We would expect the distribution of all these variables to resemble the frame. This is especially true for age category and the higher priority implicit stratification factors. Looking at Table 4, we see that this is the case, as none of the listed covariates differed significantly from the frame. Among other non-geography-based variables, the weighted sample percentages of beneficiaries with zero earnings in 2017 and 2018 were significantly less than the frame. Larger differences were found with geographic county-level variables, including some levels of the racial and ethnic profile, economic characteristics, and metropolitan status of the sample member's county of residence. In particular, the estimates of the proportions of beneficiaries residing in population-loss counties and those residing in nonmetropolitan counties were significantly less than the frame values. The estimates of the proportions of beneficiaries living in counties with high levels of poverty or high levels of persistent poverty were also a lot less than that of the frame. Although the differences did not meet the 5 percent significance threshold, the *p*-value would not exceed 0.05 by very much and

---

[46] With explicit stratification, the population is subdivided into subpopulations (strata) defined by the levels of the explicit stratification variables. Independent samples are drawn from each stratum, where the sampling fraction may or may not differ between strata. With this type of stratification, the size of the sample of each stratum is controlled. With implicit stratification, population members within each explicit stratum are sorted in priority order by the implicit stratification and the sample is selected using a sequential selection procedure. This imposes some control on the distribution of these variables in the sample.

would fit the pattern of a sample that underestimates, using base weights, regions that are poor, very rural, and in decline. Other variables for which the weighted sample estimates differed from the frame by more than one but less than two standard errors included (1) two of the five county racial and ethnic profile variables and (2) counties with high levels of persistent child poverty.[47]

b. Cross-sectional SWS

Within PSUs, the clustered sample of the SWS was selected within explicit strata defined by extract and beneficiary title, with SSI only and concurrent beneficiaries included in the same strata. We defined implicit strata for the SWS by using the same variables that we used for the RBS: disability status, beneficiary title, race and ethnicity, gender, and zip code, in that priority order. For the unclustered sample, the explicit strata were the same as those used as secondary strata in the clustered sample: extract and beneficiary title, with all SSI-only and concurrent cases combined into a single stratum. Whether the sample case was in a PSU or not was also used for explicit stratification. The implicit strata were identical to those used in the clustered sample. As with the RBS, we would expect the distribution of all these variables to resemble the frame. Looking at Table 5, we see that this is the case, as none of the listed covariates differed significantly from the frame. As with the RBS, larger differences were found with geographic county-level variables. The estimated proportions of successful workers residing in population-loss counties and those residing in nonmetropolitan counties that were not adjacent to metropolitan counties were significantly less than the frame values using the base weights. This is a similar story to the RBS. Other variables for which the weighted sample estimates differed from the frame by more than one but less than two standard errors included (1) hearing disability, (2) one of the five county racial and ethnic profile variables, (3) counties with high levels of poverty, and (4) counties with high levels of persistent poverty.[48]

c. Longitudinal SWS

We do not know which of the 89,936 persons who were identified as successful workers in Round 6 were actually working at the time of the Round 6 interview. We only know this for the sample members who were sampled and responded in Round 6. Therefore, it is not possible to compare sample values to the frame.

2. Removal of ineligible cases from the sample

If there were systematic differences between the estimates for the sampled eligible and ineligible cases, then this could point to a problem in the frame—where the sample frame covers a different population than the target population. For example, if the sample frame consisted of a large number of individuals with a particular disability who were found to be deceased, the target

---

[47] We only present statistically significant findings in this text to show patterns. Some results may be significant simply because we are conducting multiple tests without correcting for multiple comparisons.

[48] As with the RBS, statistical significance should be read with caution because we did not correct for multiple comparisons.

population (as measured by the eligible sample) could have a smaller proportion with that disability than the sample frame. We estimate that approximately 3.9 percent of the population of SSI and SSDI beneficiaries on June 30, 2018, was ineligible at data collection. Similarly, approximately 3.6 percent of the SWS was found to be ineligible at data collection, representing about 3.8 percent of the population of successful workers. With these small percentages, it is unlikely that the populations that include ineligible cases will differ significantly from the populations that do not. Nevertheless, it was instructive to investigate whether the population represented by eligible sample cases differed from the sample frame. There was some imprecision in this exercise because the eligibility for the majority of nonrespondents was unknown. Therefore, some of the cases included in the column for sample cases with known ineligibles removed will in fact be ineligible because they were nonrespondents with unknown eligibility. Because we observed a small percentage of ineligible sample members among our respondents, we assumed that the number of cases like this would be small. In Tables 4, 5, and 6, we placed asterisks by the estimates from the sample with ineligible cases removed (using base weights) that differed from the frame by more than two standard errors. There were instances where ineligible cases differed from the sample and from the frame, but when the ineligible cases were removed from the sample, the proportions did not change very much due to the small number of ineligible cases. For these samples, it appears that the eligible samples did not differ markedly from the initial samples; the patterns of deviation from the frame that were observed in the initial sample were also observed with eligible cases.[49] Any differences between the sample and the frame did not change markedly whether ineligible cases were included or not.

3. **Assessment of differences between respondents and nonrespondents before nonresponse adjustment**

To avoid the issue of unknown ineligibles among nonrespondents, we looked at the comparison between respondents and nonrespondents by including ineligibles among the respondents. These comparisons are shown in Tables 7, 8, and 9. We calculated the t-statistic by calculating the differences between the proportions within the levels of each covariate and creating an estimate of the variance of the difference by combining the standard error estimates obtained from SUDAAN. Looking at general tendencies for the non-geographic variables in Table 7, RBS respondents and ineligibles were more likely than nonrespondents to be age 50 or older and less likely to be age 40 to 49. They were also less likely to be SSI only beneficiaries (particularly as opposed to concurrent beneficiaries) and male, but these differences were not statistically significant. No significant differences were apparent for the other variables between respondents and nonrespondents.

For the SWS (Table 8), respondents and ineligibles were more likely than nonrespondents to be age 50 or older, while they were less likely to be age 18 to 39, deaf, or male. We observed significant differences between respondents and nonrespondents for the earnings category

---

[49] In other words, the pattern of asterisks between the columns for entire sample percentage and eligible sample percentage are nearly identical.

variable that was used for the SWS. Finally, response rates in later extracts were far lower than those observed in the first two extracts.

Table 7. Percentages with various attributes (categorical variables) in the RBS, using base weights among respondents (with ineligibles) and nonrespondents

Variable[a]	Frame percent	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
		Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
Beneficiary type									
SSI only	29.9	30.0	(1.3)	27.6	(1.3)	32.8	(2.4)	-5.2	-1.9
SSDI only	56.6	55.8	(1.4)	57.7	(1.6)	53.6	(2.7)	4.1	1.3
Both SSI and SSDI	13.5	14.2	(0.8)	14.7	(1.2)	13.6	(1.6)	1.0	0.5
Constructed disability status									
Hearing	0.4	0.4	(0.1)	0.4	(0.1)	0.3	(0.1)	0.1	0.9
Cognitive	11.9	11.0	(0.7)	11.1	(1.0)	10.9	(1.2)	0.2	0.1
Mental	29.6	29.9	(1.5)	28.7	(1.6)	31.3	(2.4)	-2.6	-0.9
Physical	58.1	58.8	(1.6)	59.8	(1.8)	57.5	(2.5)	2.3	0.7
Sex									
Female	49.1	49.2	(1.3)	51.3	(1.5)	46.7	(2.4)	4.7	1.6
Male	50.9	50.8	(1.3)	48.7	(1.5)	53.3	(2.4)	-4.7	-1.6
Beneficiary's age									
18–29 years	9.9	9.9	(0.3)	9.3	(0.4)	10.5	(0.7)	-1.1	-1.4
30–39 years	10.7	10.7	(0.4)	9.9	(0.5)	11.5	(0.7)	-1.6	-1.9
40–49 years	15.2	15.2	(0.4)	14.1	(0.6)	16.6	(1)	-2.4	-2.2*
50 years–FRA[b]	64.2	64.2	(0.8)	66.6	(1.1)	61.4	(1.6)	5.1	2.6*
Race/ethnicity									
White	65.9	65.4	(2.7)	65.3	(3.1)	65.6	(3.1)	-0.3	-0.1
Black	22.7	24.5	(2.5)	24.7	(2.8)	24.3	(2.7)	0.5	0.1
Hispanic	3.6	3.3	(0.8)	3.9	(1.0)	2.7	(0.7)	1.2	1.0
All others	7.8	6.7	(0.9)	6.1	(1.0)	7.4	(1.4)	-1.4	-0.8
County racial/ethnic profile									

Table 7 (*continued*)

Variable[a]	Frame percent	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
		Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
County with plurality or majority non-Hispanic black population	4.7	3.9	(2.2)	4.3	(2.5)	3.5	(2.2)	0.8	0.2
County with plurality or majority Hispanic population	9.5	7.4	(2.5)	6.8	(2.6)	8.1	(2.3)	-1.3	-0.4
County with majority but less than 90% non-Hispanic white population	37.9	46	(5.2)	45.0	(5.3)	47.2	(5.5)	-2.2	-0.3
County with racially/ethnically mixed population, no majority group	34.4	34.2	(5.1)	35.3	(5.4)	32.9	(5.3)	2.4	0.3
County with at least 90% non-Hispanic white population	12.9	8.5	(2.7)	8.6	(2.9)	8.4	(2.7)	0.3	0.1
Economic characteristics of county[c]									
Government-dependent economy county	11.9	10.5	(3.2)	10.2	(3.2)	10.9	(3.5)	-0.8	-0.2
Manufacturing-dependent economy county	9.8	8.9	(2.7)	9.4	(2.9)	8.2	(2.7)	1.1	0.3
Nonspecialized-dependent economy county	65.3	67.7	(4.8)	66.8	(5.0)	68.8	(4.9)	-2.1	-0.3
Recreation-dependent economy county	8.2	9.1	(3.1)	9.8	(3.4)	8.3	(3.0)	1.5	0.3
County with high levels of poverty	18.5	12.2	(3.5)	12.7	(3.7)	11.6	(3.4)	1.1	0.2
County with high levels of persistent poverty	8.3	4.4	(2.1)	4.3	(2.0)	4.6	(2.2)	-0.2	-0.1
County with high levels of persistent child poverty	19.1	15.0	(3.9)	17.3	(4.4)	12.2	(3.4)	5.1	0.9
County with low education	13.3	11.9	(3.1)	12.4	(3.5)	11.3	(2.8)	1.1	0.2
Population-loss county	8.7	3.6	(1.8)	4.1	(2.0)	2.9	(1.6)	1.2	0.5
Retirement-destination county	14.1	14.9	(3.6)	15.8	(3.9)	13.9	(3.4)	1.9	0.4
Metropolitan status of county									
Metropolitan area of 1 million population or more	46.5	43.9	(5.4)	41.8	(5.5)	46.4	(5.7)	-4.6	-0.6
Metropolitan area of 250,000 to 999,999 population	22.6	27.2	(4.9)	26.2	(4.9)	28.3	(5.2)	-2.1	-0.3
Metropolitan area of fewer than 250,000 population	10.7	13.0	(3.7)	13.8	(4.0)	12.2	(3.6)	1.6	0.3
Nonmetropolitan area adjacent to large metropolitan area	4.4	3.8	(1.9)	4.9	(2.4)	2.5	(1.3)	2.3	0.9
Nonmetropolitan area adjacent to medium or small metropolitan area	9.0	8.8	(2.9)	10.0	(3.4)	7.4	(2.5)	2.7	0.6
Nonmetropolitan area not adjacent to metropolitan area	6.9	3.2	(1.7)	3.3	(1.7)	3.2	(1.7)	0.1	0.0
Census region									

Table 7 (*continued*)

Variable[a]	Frame percent	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
		Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
West	18.5	17.3	(4.0)	15.5	(3.9)	19.5	(4.4)	-4.0	-0.7
South	41.4	42.4	(5.6)	44.8	(5.8)	39.4	(5.7)	5.5	0.7
Northeast	18.2	18.8	(4.4)	17.6	(4.3)	20.2	(4.8)	-2.6	-0.4
Midwest	21.9	21.5	(4.5)	22.0	(4.8)	20.9	(4.5)	1.1	0.2
Census division									
East North Central	15.7	15.3	(3.9)	16.1	(4.2)	14.4	(3.6)	1.6	0.3
West North Central	6.2	6.2	(2.7)	6.0	(2.7)	6.5	(2.9)	-0.5	-0.1
New England	5.0	5.1	(2.5)	5.1	(2.5)	5.2	(2.7)	-0.1	0.0
Middle Atlantic	13.2	13.7	(3.9)	12.5	(3.7)	15.0	(4.3)	-2.5	-0.4
South Atlantic	20.5	20.1	(4.7)	20.4	(4.8)	19.8	(4.9)	0.6	0.1
East South Central	9.2	10.2	(3.5)	11.5	(3.9)	8.7	(3.5)	2.8	0.5
West South Central	11.8	12.0	(3.6)	13.0	(3.9)	10.8	(3.5)	2.1	0.4
Mountain	5.7	5.9	(2.6)	5.9	(2.6)	5.9	(2.7)	0.1	0.0
Pacific	12.7	11.4	(3.3)	9.6	(3.1)	13.7	(3.7)	-4.1	-0.8
Earnings categories for 2017–2018 time period									
Three consecutive months of monthly earnings above SGA	3.2	4.2	(0.6)	3.2	(0.6)	5.5	(1.1)	-2.3	-1.8
Monthly earnings above $7,000 in at least one month in '17 or '18	1.4	1.8	(0.4)	1.8	(0.5)	1.9	(0.6)	-0.1	-0.1
Monthly earnings above $2,000 in at least one month in '17 or '18	2.4	3.0	(0.4)	3.5	(0.6)	2.5	(0.5)	1.0	1.2
Monthly earnings above zero in at least one month in '17 or '18	2.6	3.0	(0.4)	3.3	(0.5)	2.7	(0.5)	0.7	0.9
Monthly earnings always zero in 2017 and 2018	90.4	87.9	(0.9)	88.2	(1.0)	87.5	(1.4)	0.7	0.4

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

*Denotes a difference between the sample and frame value of more than two standard errors.

Table 8. Percentages with various attributes (categorical variables) in the SWS, using base weights among respondents (with ineligibles) and nonrespondents

Variable[a]	Frame Percent	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
		Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
Beneficiary type									
SSI only	26.0	25.7	(0.9)	26.4	(0.9)	25.1	(1.1)	1.3	0.9
SSDI only	52.5	52.5	(1.1)	51.8	(1.2)	53.0	(1.4)	-1.2	-0.6
Both SSI and SSDI	21.5	21.8	(0.6)	21.7	(0.9)	21.8	(0.8)	-0.1	-0.1
Constructed disability status									
Hearing	2.0	1.7	(0.2)	1.3	(0.2)	2.0	(0.3)	-0.8	-2.2*
Cognitive	14.2	14.5	(0.6)	13.6	(1.0)	15.2	(0.8)	-1.6	-1.3
Mental	36.2	36.1	(1.0)	35.7	(1.3)	36.4	(1.1)	-0.8	-0.4
Physical	47.6	47.6	(1.1)	49.5	(1.4)	46.3	(1.1)	3.2	1.8
Sex									
Female	45.9	44.8	(0.7)	47.5	(1.0)	43.0	(1.0)	4.5	3.2*
Male	54.1	55.2	(0.7)	52.5	(1.0)	57.0	(1.0)	-4.5	-3.2*
Beneficiary's age									
18–29 years	22.7	23.0	(0.7)	20.3	(0.8)	24.9	(1.0)	-4.7	-3.6*
30–39 years	23.9	24.5	(0.7)	23.1	(0.8)	25.5	(0.9)	-2.4	-2.0*
40–49 years	22.4	21.7	(0.6)	21.7	(0.7)	21.7	(0.9)	0.0	0.0
50 years–FRA[b]	30.9	30.7	(0.8)	34.9	(1.1)	27.8	(1.0)	7.0	4.7*
Race/ethnicity									
White	54.9	55.8	(2.3)	54.1	(2.3)	57.0	(2.5)	-2.8	-0.8
Black	31.4	32.0	(2.2)	34.7	(2.2)	30.1	(2.4)	4.6	1.5
Hispanic	4.9	4.5	(0.5)	3.7	(0.5)	5.0	(0.6)	-1.4	-1.6
All others	8.8	7.8	(0.7)	7.5	(0.8)	7.9	(0.8)	-0.5	-0.4
County racial/ethnic profile									

Table 8 (*continued*)

Variable[a]	Frame Percent	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
		Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
County with plurality or majority non-Hispanic black population	4.6	3.4	(1.6)	3.7	(1.6)	3.1	(1.7)	0.7	0.3
County with plurality or majority Hispanic population	9.7	8.4	(2.1)	7.9	(1.8)	8.7	(2.4)	-0.8	-0.3
County with majority but less than 90% non-Hispanic white population	38.7	45.5	(4.8)	43.8	(4.3)	46.6	(5.2)	-2.8	-0.4
County with racially/ethnically mixed population, no majority group	36.5	34.3	(4.6)	36.0	(4.2)	33.2	(5.0)	2.8	0.4
County with at least 90% non-Hispanic white population	10.2	8.4	(2.4)	8.5	(2.2)	8.3	(2.7)	0.1	0.0
Economic characteristics of county[c]									
Government-dependent economy county	12.4	11.9	(3.2)	12.4	(3.0)	11.5	(3.4)	0.9	0.2
Manufacturing-dependent economy county	8.2	8.9	(2.8)	9.1	(2.7)	8.8	(2.9)	0.3	0.1
Nonspecialized-dependent economy county	67.3	66.5	(4.5)	67.7	(4.1)	65.7	(5.0)	2.0	0.3
Recreation-dependent economy county	8.3	9.3	(2.9)	7.6	(2.2)	10.4	(3.4)	-2.8	-0.7
County with high levels of poverty	14.4	10.6	(2.8)	12.1	(2.7)	9.6	(2.9)	2.4	0.6
County with high levels of persistent poverty	6.1	3.9	(1.6)	4.5	(1.6)	3.4	(1.6)	1.1	0.5
County with high levels of persistent child poverty	15.2	12.3	(3.0)	13.6	(2.9)	11.4	(3.2)	2.2	0.5
County with low education	11.6	12.2	(2.8)	11.8	(2.6)	12.4	(3.0)	-0.6	-0.2
Population-loss county	8.1	4.2	(1.8)	5.2	(2.0)	3.5	(1.7)	1.6	0.6
Retirement-destination county	12.4	13.1	(3.0)	12.7	(2.8)	13.4	(3.3)	-0.6	-0.1
Metropolitan status of county									
Metropolitan area of 1 million population or more	54.0	52.9	(4.8)	52.3	(4.3)	53.3	(5.2)	-1.0	-0.1
Metropolitan area of 250,000 to 999,999 population	21.9	24.3	(4.0)	24.7	(3.8)	24.1	(4.3)	0.6	0.1
Metropolitan area of fewer than 250,000 population	9.8	11.5	(3.1)	11.9	(3.0)	11.3	(3.3)	0.6	0.1
Nonmetropolitan area adjacent to large metropolitan area	3.3	3.7	(1.5)	3.7	(1.6)	3.7	(1.6)	0.0	0.0
Nonmetropolitan area adjacent to medium or small metropolitan area	6.2	4.8	(1.4)	5.0	(1.3)	4.6	(1.5)	0.4	0.2
Nonmetropolitan area not adjacent to metropolitan area	4.8	2.8	(1.0)	2.4	(0.8)	3.0	(1.2)	-0.6	-0.4
Census region									

Table 8 (*continued*)

Variable[a]	Frame Percent	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
		Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
West	22.3	22.9	(4.2)	20.6	(3.6)	24.5	(4.8)	-3.9	-0.7
South	33.5	33.2	(4.5)	35.0	(4.2)	32.0	(4.8)	3.0	0.5
Northeast	21.3	21.3	(4.2)	20.3	(3.6)	22.0	(4.7)	-1.6	-0.3
Midwest	23.0	22.6	(4.1)	24.1	(3.9)	21.5	(4.3)	2.6	0.4
Census division									
East North Central	16.0	15.2	(3.4)	16.5	(3.3)	14.3	(3.5)	2.2	0.5
West North Central	7.0	7.3	(2.7)	7.5	(2.5)	7.2	(2.9)	0.4	0.1
New England	6.7	6.2	(2.6)	6.2	(2.4)	6.2	(2.8)	0.0	0.0
Middle Atlantic	14.5	15.1	(3.7)	14.2	(3.0)	15.8	(4.1)	-1.7	-0.3
South Atlantic	16.3	14.6	(3.2)	15.1	(2.9)	14.3	(3.5)	0.8	0.2
East South Central	6.0	7.2	(2.4)	7.9	(2.5)	6.7	(2.3)	1.2	0.4
West South Central	11.1	11.4	(3.1)	12.0	(2.9)	11.0	(3.3)	1.0	0.2
Mountain	6.0	6.7	(2.5)	6.2	(2.1)	7.0	(2.9)	-0.8	-0.2
Pacific	16.3	16.2	(3.8)	14.4	(3.1)	17.5	(4.3)	-3.1	-0.6
Earnings categories for 2017–2018 time period									
Monthly earnings above $30,000 in at least one month in '17 or '18	22.3	22.7	(0.6)	19.7	(0.8)	24.8	(1.0)	-5.1	-4.0*
Monthly earnings above $20,000 in at least one month in '17 or '18	23.5	23.7	(0.5)	23.6	(0.8)	23.9	(0.7)	-0.3	-0.3
Monthly earnings above $15,000 in at least one month in '17 or '18	19.3	19.5	(0.5)	20.8	(0.7)	18.5	(0.8)	2.2	2.2*
Monthly earnings above $7,000 in at least one month in '17 or '18	21.4	21.2	(0.5)	22.4	(0.8)	20.3	(0.7)	2.1	1.9
Monthly earnings did not meet the thresholds above	13.5	12.9	(0.5)	13.5	(0.7)	12.4	(0.6)	1.1	1.2
Extract									
December 2018 extract	19.5	19.5	(0.6)	23.3	(0.8)	16.9	(0.7)	6.4	6.0*
January 2019 extract	15.9	15.9	(0.4)	18.4	(0.7)	14.2	(0.5)	4.3	5.0*
March 2019 extract	16.2	16.2	(0.4)	15.1	(0.7)	17.0	(0.6)	-1.9	-2.2*

Table 8 (*continued*)

Variable[a]	Frame Percent	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
		Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
April 2019 extract	12.0	12.0	(0.4)	11.5	(0.5)	12.4	(0.6)	-0.9	-1.1
June 2019 extract	17.0	17.0	(0.5)	13.9	(0.6)	19.1	(0.7)	-5.2	-5.3*
July 2019 extract	10.9	10.9	(0.3)	10.2	(0.6)	11.4	(0.5)	-1.3	-1.7
September 2019 extract	8.4	8.4	(0.3)	7.6	(0.5)	9.0	(0.6)	-1.4	-1.8

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

*Denotes a statistically significant difference between nonrespondents and respondents (which include ineligible cases).

Table 9. Percentages with various attributes (categorical variables) in the longitudinal SWS, using base weights among respondents (with ineligibles) and nonrespondents

Variable[a]	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
	Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
Round 6 Beneficiary type								
SSI only	27.8	(0.9)	27.6	(1.2)	28.1	(1.4)	-0.5	-0.3
SSDI only	51.1	(1.0)	53.2	(1.2)	48.6	(1.6)	4.6	2.3*
Both SSI and SSDI	21.1	(0.8)	19.2	(0.9)	23.3	(1.4)	-4.1	-2.4*
Round 6 Constructed disability status								
Hearing	3.1	(0.4)	2.1	(0.5)	4.4	(0.7)	-2.3	-2.8*
Cognitive	12.8	(0.7)	11.8	(0.9)	13.9	(1.2)	-2.1	-1.4
Mental	36.6	(1.0)	36.9	(1.3)	36.2	(1.7)	0.8	0.4
Physical	47.5	(1.0)	49.2	(1.3)	45.5	(1.6)	3.6	1.7
Sex								
Female	45.4	(1.1)	46.9	(1.4)	43.6	(1.6)	3.3	1.6
Male	54.6	(1.1)	53.1	(1.4)	56.4	(1.6)	-3.3	-1.6
Beneficiary's age								
18–29 years	24.8	(0.9)	22.4	(1.2)	27.6	(1.3)	-5.2	-3.0*
30–39 years	23.0	(0.8)	20.3	(1.0)	26.2	(1.3)	-5.9	-3.7*
40–49 years	21.6	(0.7)	22.3	(1.1)	20.7	(1.1)	1.6	1.0
50 years–FRA[b]	30.7	(0.9)	35.1	(1.2)	25.4	(1.2)	9.6	5.5*
Race/ethnicity								
White	57.6	(2.0)	54.7	(2.1)	61.2	(2.6)	-6.5	-2.0*
Black	29.7	(1.8)	32.0	(1.9)	26.7	(2.3)	5.2	1.7
Hispanic	5.4	(0.6)	5.4	(0.7)	5.4	(0.8)	0.0	0.0
All others	7.4	(0.7)	7.9	(0.8)	6.6	(0.8)	1.3	1.1
County racial/ethnic profile								
County with plurality or majority non-Hispanic black population	3.3	(1.1)	3.3	(1.2)	3.2	(1.1)	0.1	0.0

Table 9 (*continued*)

Variable[a]	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
	Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
County with plurality or majority Hispanic population	8.2	(1.4)	8.4	(1.3)	7.9	(1.7)	0.5	0.2
County with majority but less than 90% non-Hispanic white population	43.7	(3.7)	42.8	(3.5)	44.7	(4.1)	-1.9	-0.4
County with racially/ethnically mixed population, no majority group	35.5	(3.6)	36.4	(3.4)	34.5	(4.0)	2.0	0.4
County with at least 90% non-Hispanic white population	9.1	(1.9)	8.7	(1.7)	9.5	(2.3)	-0.8	-0.3
Economic characteristics of county[c]								
Government-dependent economy county	12.7	(2.8)	12.5	(2.7)	12.9	(3.2)	-0.4	-0.1
Manufacturing-dependent economy county	8.1	(1.9)	8.0	(1.9)	8.3	(2.2)	-0.3	-0.1
Nonspecialized-dependent economy county	66.8	(3.5)	66.6	(3.4)	67.2	(3.9)	-0.6	-0.1
Recreation-dependent economy county	9.3	(2.0)	9.4	(2.0)	9.1	(2.2)	0.3	0.1
County with high levels of poverty	11.1	(1.9)	11.8	(2.0)	10.3	(1.9)	1.5	0.5
County with high levels of persistent poverty	4.7	(1.4)	4.8	(1.4)	4.6	(1.4)	0.2	0.1
County with high levels of persistent child poverty	12.4	(2.2)	12.7	(2.2)	12.1	(2.2)	0.6	0.2
County with low education	11.3	(2.0)	11.5	(2.0)	11.1	(2.3)	0.5	0.2
Population-loss county	4.9	(1.2)	5.2	(1.2)	4.6	(1.2)	0.6	0.4
Retirement-destination county	12.4	(2.3)	10.8	(2.0)	14.3	(2.8)	-3.5	-1.0
Metropolitan status of county								
Metropolitan area of 1 million population or more	52.3	(3.7)	53.5	(3.4)	51.0	(4.2)	2.5	0.5
Metropolitan area of 250,000 to 999,999 population	24.5	(3.4)	24.9	(3.2)	24.1	(3.8)	0.8	0.2
Metropolitan area of fewer than 250,000 population	10.9	(2.2)	9.7	(1.9)	12.3	(2.9)	-2.6	-0.7
Nonmetropolitan area adjacent to large metropolitan area	2.9	(0.7)	2.6	(0.7)	3.2	(0.8)	-0.7	-0.6
Nonmetropolitan area adjacent to medium or small metropolitan area	5.8	(1.4)	5.9	(1.4)	5.8	(1.6)	0.1	0.1
Nonmetropolitan area not adjacent to metropolitan area	3.6	(0.6)	3.5	(0.6)	3.7	(0.9)	-0.2	-0.2
Census region								
West	22.8	(3.2)	23.7	(3.2)	21.8	(3.5)	1.9	0.4

Table 9 (*continued*)

Variable[a]	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
	Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
South	31.1	(3.3)	30.1	(3.1)	32.4	(3.8)	-2.4	-0.5
Northeast	23.3	(3.2)	22.6	(3.0)	24.1	(3.7)	-1.6	-0.3
Midwest	22.7	(3.1)	23.7	(2.9)	21.6	(3.5)	2.1	0.4
Census division								
Pacific	15.8	(2.7)	16.7	(2.7)	14.7	(2.9)	2.0	0.5
East North Central	15.0	(2.5)	16.4	(2.4)	13.4	(2.7)	3.0	0.8
Middle Atlantic	16.2	(2.8)	15.7	(2.7)	16.9	(3.2)	-1.2	-0.3
South Atlantic	15.7	(2.5)	14.9	(2.3)	16.7	(2.9)	-1.8	-0.5
West South Central	9.2	(1.9)	9.2	(1.9)	9.2	(2.3)	0.0	0.0
East South Central	6.3	(2.0)	6.1	(1.9)	6.6	(2.2)	-0.5	-0.2
Mountain	7.1	(2.1)	7.0	(2.1)	7.2	(2.3)	-0.1	0.0
New England	7.1	(2.0)	6.9	(1.8)	7.2	(2.3)	-0.3	-0.1
West North Central	7.7	(2.2)	7.3	(1.9)	8.2	(2.7)	-1.0	-0.3
Earnings categories for 2015–2016 time period								
Monthly earnings above $30,000 in at least one month in '15 or '16	20.9	(0.8)	19.2	(1.0)	22.9	(1.4)	-3.7	-2.2*
Monthly earnings above $20,000 in at least one month in '15 or '16	23.0	(1.0)	22.9	(1.2)	23.1	(1.3)	-0.2	-0.1
Monthly earnings above $15,000 in at least one month in '15 or '16	19.2	(0.8)	19.6	(1.0)	18.8	(1.2)	0.8	0.5
Monthly earnings above $7,000 in at least one month in '15 or '16	24	(0.9)	24.5	(1.0)	23.4	(1.4)	1.1	0.6
Monthly earnings did not meet the thresholds above	12.9	(0.7)	13.8	(0.8)	11.8	(0.9)	2.0	1.6
Extract								
December 2015 extract	19.7	(0.7)	20.7	(0.9)	18.5	(1.2)	2.3	1.5
January 2016 extract	14.7	(0.6)	14.7	(0.7)	14.9	(1.1)	-0.2	-0.1
March 2016 extract	19.0	(0.7)	16.4	(0.8)	22.0	(1.3)	-5.6	-3.6*
April 2016 extract	12.2	(0.5)	11.1	(0.7)	13.5	(1.0)	-2.4	-2.1*

Table 9 (*continued*)

Variable[a]	Entire sample		Respondents and ineligibles		Nonrespondents		Comparison	
	Percent	(se)	Percent	(se)	Percent	(se)	Difference	t-statistic
June 2016 extract	14.7	(0.6)	15.2	(0.8)	14.0	(1.1)	1.2	0.9
July 2016 extract	11.3	(0.5)	12.0	(0.7)	10.5	(1.0)	1.5	1.3
September 2016 extract	8.5	(0.3)	9.9	(0.5)	6.7	(0.6)	3.2	4.2*

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

*Denotes a statistically significant difference between nonrespondents and respondents (which include ineligible cases).

Differences between respondents plus ineligibles and nonrespondents for the longitudinal SWS are provided in Table 9. There were some similarities with the other two samples, though we observed more differences between respondents and nonrespondents in this sample. In particular, respondents and ineligibles were more likely than nonrespondents to be age 50 or older, beneficiaries of SSDI only, and from the seventh extract.[50] Nonrespondents were more likely than respondents plus ineligibles to be (1) age 18 to 39, (2) beneficiaries of both SSDI and SSI, (3) deaf, (4) white, and (5) from the third or fourth extract. Nonrespondents were also more likely than respondents plus ineligibles to have earnings of $30,000 or more in at least one month.

4.   Nonresponse adjustment

Nonresponse adjustments made to base weights seek to reduce the potential for bias that might result from differential nonresponse on the basis of the variables used in the nonresponse adjustment. These adjustments were calculated separately for each sample. In addition, for the longitudinal SWS, we also separated the sample into two pieces, depending upon whether the sample member remained an SSI or SSDI beneficiary on June 30, 2018 (the date of the Round 7 frame). We conducted the nonresponse adjustments separately for these two pieces.

For the RBS, cross-sectional SWS, and longitudinal SWS in the Round 7 frame, we calculated two separate nonresponse adjustments by using a logistic propensity model for location and another logistic propensity model for cooperation. Known ineligibles were considered to be located and cooperating for these models. The predicted value from each model was the probability that a sample member was located or responded to the survey. We used a Chi-square Automatic Interaction Detector (CHAID) analysis in SPSS to identify possible statistically significant interactions.[51] If an interaction was included in a candidate model, then the main effects associated with that interaction were also included. At a particular level of a given covariate or interaction, if all respondents either were located or unlocated (for the location models), complete or not complete (for the cooperation models), or the total number of sample

---

[50] Extracts in the longitudinal SWS were based upon when they were defined as successful workers in Round 6. We did not have the same data collection constraints on later extracts for the longitudinal SWS as we did in the cross-sectional SWS. The only difference between extracts was when their successful work was defined.

[51] CHAID normally is attributed to Kass (1980) and Biggs and colleagues (1991). Its application in SPSS is described in Magidson (1993). The CHAID procedure iteratively segments a data set into mutually exclusive subgroups that share similar characteristics based on their effect on nominal or ordinal dependent variables. It automatically checks all variables in the data set and creates a hierarchy that shows all statistically significant subgroups. The procedure generates a tree that identifies the set of variables and interactions among the variables that have an association with the ability to locate a sample member (and the propensity of a located sample member to respond or be ineligible). We first ran CHAID with all covariates and then re-ran it a few times with the top variable in the tree removed, in order to ensure that all potentially important interactions were retained for further consideration.

members at that level was fewer than 20, we combined levels if doing so was possible or logical. If combining levels was not possible, we excluded the covariate or interaction from the pool.

We used forward and backward stepwise selection logistic regression procedures with normalized weights to reduce the pool of covariates, which included both main effects and the interactions from CHAID.[52] Next, we carefully evaluated a series of models by comparing the following measures of predictive ability and goodness of fit: the Generalized Coefficient of Determination (also known as the Generalized R-squared statistic),[53] percentage of concordant and discordant pairs,[54] and the Hosmer-Lemeshow goodness-of-fit test.[55] The selection of the final model involved evaluating these measures in concert, choosing a parsimonious model that was among the best in all or most of these measures using SUDAAN. Model fitting also involved a review of the statistical significance of the coefficients of the covariates in the model and avoidance of any unusually large adjustment factors. In addition, we manipulated the set of variables to avoid data warnings in SUDAAN.[56] Once we finalized the model, we calculated the location and cooperation adjustments as the inverse of the propensity scores. We multiplied the base weight by the two adjustments to form the nonresponse-adjusted weights. We then trimmed the nonresponse-adjusted weights (if necessary) to reduce the variance attributable to outlier weights.[57] Finally, we raked the RBS weights so that the weighted totals for beneficiary type, age category, gender, and RBS earnings category added up to frame totals for the RBS. In addition,

---

[52] The stepwise logistic regression procedure does not take into account the sampling design when computing standard errors, so the variances are usually underestimated. The final model is developed by using SUDAAN to incorporate the sample design features of stratification and clustering as well as weighting. For the locatability stepwise regression model, we normalized the base weight so that it summed to the sample size. For the cooperation stepwise regression model, we normalized the location-adjusted weight so that it summed to the sample size.

[53] The Generalized Coefficient of Determination (Cox and Snell 1989) is a measure of the adequacy of the model, where higher numbers indicate a greater difference between the likelihood of the model in question and the likelihood of the null model. The Max rescaled R-Square scales this value to have a maximum of 1.

[54] A pair of observations is concordant if a responding subject has a higher predicted value than the nonresponding subject, discordant if not, and tied if both members of the pair are either respondents, nonrespondents, or have the same predicted values. The predicted value is the probability of location or response from the logistic propensity model. It is desirable to have as many concordant and as few discordant pairs as possible among all possible pairs of observations (Agresti 1990).

[55] The Hosmer-Lemeshow goodness-of-fit test is a test for goodness of fit of logistic regression models. Unlike the Pearson and deviance goodness-of-fit tests, it may be used to test goodness of fit even when some covariates are continuous (Hosmer and Lemeshow 1989).

[56] SUDAAN data warnings usually included one or more of the following: (1) an indication of a response cell with a zero count, (2) one or more parameters approaching infinity (which may not be readily observable with the parameter estimates themselves), and (3) degrees of freedom for overall contrast less than the maximum number of estimable parameters. We tried to avoid all such warnings, although avoiding the first two was the highest priority. The warnings almost always were caused by a response cell with a count that was too small, which required dropping the covariate or combining categories of a covariate.

[57] Trimming is a process whereby outlier weights are trimmed to be closer to the rest of the weights in the distribution. The trimmed amount is reallocated to the rest of the weights in the same trimming class. The decision about how much to trim is a subjective one, and is based on the balance between reducing the variance in the weights, and minimizing any increase in bias that might result from trimming.

for the 45 cases in the RBS who were also in the SWS frame, we ensured that their weights added up to the total in the SWS frame. We raked the cross-sectional SWS weights so that the weighted totals by beneficiary type, age category, SWS earnings category, and extract added up to the SWS frame totals.

For the longitudinal SWS not in the Round 7 frame, the number of cases involved was small. As a result, we did not create the adjustments using logistic regression models. Instead, we calculated the adjustments using simple weighting classes. We created the final longitudinal SWS weights by trimming and raking to marginal totals together across the two groups (the Round 7 frame and non-Round 7 frame cases). We raked the longitudinal SWS weights so that the weighted totals matched the estimated population counts of beneficiary type (in Round 6), age category (as of June 30, 2018), SWS earnings category (in Round 6), and extract (in Round 6). We determined these estimated population counts by summing the base weights for all sample cases in the longitudinal sample.

5. **Comparison of respondents and ineligibles to the sampling frame after nonresponse adjustment for the RBS and SWS**

The purpose of nonresponse adjustments is to account for differences between respondents and nonrespondents, in order to make respondents look like the original sample as much as possible. In this section, we evaluate how well the nonresponse adjustments accounted for those differences.

In Tables 10, 11, and 12, we included percentages from the sample frame; estimates from the entire sample (using base weights); and nonresponse-adjusted weighted estimates among respondents and ineligibles (again, including ineligibles because the number of ineligibles among nonrespondents was unknown) for the RBS, cross-sectional SWS, and longitudinal SWS, respectively. We made comparisons between the estimates using nonresponse-adjusted weights and the sample frames, but there were differences between the selected sample and the frame that the nonresponse adjustments could not rectify. We post-stratified the counts to match the frame for the beneficiary type, age, gender, and earnings categories in the RBS; for the beneficiary type, age, and earnings categories and the extract for the cross-sectional SWS; and for Round 6 beneficiary type, age, Round 6 earnings categories, and Round 6 extract for the longitudinal SWS.

a. RBS

As Table 10 indicates, the nonresponse and raking adjustments to the sampling weights alleviated all of the differences observed between respondents and nonrespondents, as the proportions observed for age, gender, three of the four race categories, three of the four disability categories, beneficiary type, and the RBS earnings categories (using nonresponse-adjusted weights) were all within one standard error of the sampling frame and no new differences were generated by these adjustments. In fact, for the RBS earnings categories, the raking procedure rectified the deviation between the sample proportions and the frame proportions. The difference

between the non-hearing physical disability proportion in the frame and final sample estimate exceeded one standard error, but the proportion estimated matched the original sample estimate (using base weights) very closely. However, the post data collection adjustments did not rectify differences that were observed in some of the geographic variables between the selected sample and the frame. In particular, the estimated proportions of beneficiaries residing in nonmetropolitan counties not adjacent to metropolitan counties and those residing in population-loss counties were significantly underestimated when compared to the frame (using nonresponse-adjusted weights). The proportion of beneficiaries residing in high poverty counties was also quite a bit less than that observed in the frame, though the difference was not statistically significant at the 5 percent level.

b.   Cross-sectional SWS

As with the RBS, the nonresponse and raking adjustments to the sampling weights in the cross-sectional SWS alleviated all of the differences observed between respondents and nonrespondents. In the case of the cross-sectional SWS, as Table 11 shows, the estimated proportions observed for age, gender, disability, four of the five county racial and ethnic composition variables, three of the four race categories, SWS earnings categories, and the extracts (using nonresponse-adjusted weights) were all within one standard error of the sampling frame values. The only variable for which differences between the original sample and the frame were exacerbated by post–data collection adjustments was the "all others" race category, which had a significantly lower estimate using nonresponse adjusted weights than was observed in the frame. Among non-geographic variables, other observed estimates using nonresponse adjusted weights that differed from the frame by more than one standard error were very close to those of the selected sample and were within two standard errors of the proportion in the frame. Among the geographic variables, differences between the nonresponse-adjusted estimates and the sampling frame exceeded two standard errors, but these were also due to the differences between the selected sample estimates, based on the base weights and the frame values that were not rectified by the adjustments. In particular, the estimated proportions of beneficiaries residing in nonmetropolitan counties not adjacent to metropolitan counties were significantly less than that observed in the sampling frame. The county-based poverty variable proportions also differed substantially from the frame, though these differences did not meet the threshold of two standard errors.

c.   Longitudinal SWS

As with the other samples, the nonresponse and raking adjustments to the sampling weights in the longitudinal SWS mitigated all of the differences observed between respondents and nonrespondents. We do not have frame values to compare to, but the estimated sample proportions using the post–data collection adjustments were within one standard error of the original sample estimates (using base weights) for all of the variables except two levels of the disability category, which were within two standard errors of the original sample estimates.

Table 10. Percentages with various attributes (categorical variables) in the RBS, comparing frame percent with final weighted estimate (using nonresponse-adjusted weights)

Variable[a]	Frame percent	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
		Percent	(se)		Percent	(se)
Beneficiary type						
SSI only	29.9	30.0	(1.3)	1,533	29.9	(1.3)
SSDI only	56.6	55.8	(1.4)	1,714	56.6	(1.4)
Both SSI and SSDI	13.5	14.2	(0.8)	761	13.5	(1.0)
Constructed disability status						
Hearing	0.4	0.4	(0.1)	25	0.4	(0.1)
Cognitive	11.9	11.0	(0.7)	788	11.0	(0.8)
Mental	29.6	29.9	(1.5)	1,447	29.7	(1.5)
Physical	58.1	58.8	(1.6)	1,696	58.9	(1.6)
Sex						
Female	49.1	49.2	(1.3)	2,045	49.1	(1.4)
Male	50.9	50.8	(1.3)	1,963	50.9	(1.4)
Beneficiary's age						
18–29 years	9.9	9.9	(0.3)	1,127	9.9	(0.4)
30–39 years	10.7	10.7	(0.4)	1,059	10.7	(0.5)
40–49 years	15.2	15.2	(0.4)	1,118	15.2	(0.6)
50 years–FRA[b]	64.2	64.2	(0.8)	704	64.2	(1.1)
Race/ethnicity						
White	65.9	65.4	(2.7)	2,067	66.7	(2.9)
Black	22.7	24.5	(2.5)	742	23.8	(2.5)
Hispanic	3.6	3.3	(0.8)	145	3.3	(0.8)
All others	7.8	6.7	(0.9)	212	6.2	(1.0)
County racial/ethnic profile						
County with plurality or majority non-Hispanic black population	4.7	3.9	(2.2)	131	3.7	(2.2)
County with plurality or majority Hispanic population	9.5	7.4	(2.5)	318	7.1	(2.7)

Table 10 (*continued*)

Variable[a]	Frame percent	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
		Percent	(se)		Percent	(se)
County with majority but less than 90% non-Hispanic white population	37.9	46.0	(5.2)	1,878	44.8	(5.3)
County with racially/ethnically mixed population, no majority group	34.4	34.2	(5.1)	1,310	35.9	(5.4)
County with at least 90% non-Hispanic white population	12.9	8.5	(2.7)	371	8.5	(2.8)
Economic characteristics of county[c]						
Government-dependent economy county	11.9	10.5	(3.2)	426	10.7	(3.3)
Manufacturing-dependent economy county	9.8	8.9	(2.7)	351	9.1	(2.8)
Nonspecialized-dependent economy county	65.3	67.7	(4.8)	2,706	67.8	(4.9)
Recreation-dependent economy county	8.2	9.1	(3.1)	346	8.5	(3.0)
County with high levels of poverty	18.5	12.2	(3.5)	485	12.0	(3.5)
County with high levels of persistent poverty	8.3	4.4	(2.1)	203	4.4	(2.1)
County with high levels of persistent child poverty	19.1	15.0	(3.9)	612	15.2	(4.0)
County with low education	13.3	11.9	(3.1)	490	12.4	(3.5)
Population-loss county	8.7	3.6	(1.8)*	150	3.8	(1.9)*
Retirement-destination county	14.1	14.9	(3.6)	570	15.1	(3.7)
Metropolitan status of county						
Metropolitan area of 1 million population or more	46.5	43.9	(5.4)	1,735	43.7	(5.5)
Metropolitan area of 250,000 to 999,999 population	22.6	27.2	(4.9)	1,119	27.5	(5.0)
Metropolitan area of fewer than 250,000 population	10.7	13.0	(3.7)	493	13.0	(3.8)
Nonmetropolitan area adjacent to large metropolitan area	4.4	3.8	(1.9)	168	4.3	(2.1)
Nonmetropolitan area adjacent to medium or small metropolitan area	9.0	8.8	(2.9)	352	8.5	(2.9)
Nonmetropolitan area not adjacent to metropolitan area	6.9	3.2	(1.7)*	141	2.9	(1.6)*
Census region						
West	18.5	17.3	(4.0)	734	16.3	(4.0)
South	41.4	42.4	(5.6)	1,659	42.7	(5.7)
Northeast	18.2	18.8	(4.4)	718	18.6	(4.4)

Table 10 (*continued*)

Variable[a]	Frame percent	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
		Percent	(se)		Percent	(se)
Midwest	21.9	21.5	(4.5)	897	22.3	(4.8)
Census division						
East North Central	15.7	15.3	(3.9)	634	16.5	(4.3)
West North Central	6.2	6.2	(2.7)	263	5.8	(2.6)
New England	5.0	5.1	(2.5)	205	5.2	(2.5)
Middle Atlantic	13.2	13.7	(3.9)	513	13.5	(3.9)
South Atlantic	20.5	20.1	(4.7)	767	19.7	(4.6)
East South Central	9.2	10.2	(3.5)	384	10.5	(3.6)
West South Central	11.8	12.0	(3.6)	508	12.5	(3.8)
Mountain	5.7	5.9	(2.6)	269	6.0	(2.6)
Pacific	12.7	11.4	(3.3)	465	10.3	(3.2)
Earnings categories for 2017–2018 time period						
Three consecutive months of monthly earnings above SGA	3.2	4.2	(0.6)	192	3.2	(0.4)
Monthly earnings above $7,000 in at least one month in '17 or '18	1.4	1.8	(0.4)	75	1.2	(0.3)
Monthly earnings above $2,000 in at least one month in '17 or '18	2.4	3.0	(0.4)	216	2.6	(0.4)
Monthly earnings above zero in at least one month in '17 or '18	2.6	3.0	(0.4)	202	2.6	(0.4)
Monthly earnings always zero in 2017 and 2018	90.4	87.9	(0.9)	3,323	90.4	(0.7)

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

*Denotes a difference of more than two standard errors between the sample estimate (either from the entire sample or using the nonresponse-adjusted weight) and the frame value.

Table 11. Percentages with various attributes (categorical variables) in the SWS, comparing frame percent with final weighted estimate (using nonresponse-adjusted weights)

Variable[a]	Frame percent	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
		Percent	(se)		Percent	(se)
Beneficiary type						
SSI only	26.0	25.7	(0.9)	860	26.0	(0.9)
SSDI only	52.5	52.5	(1.1)	1,493	52.5	(1.2)
Both SSI and SSDI	21.5	21.8	(0.6)	664	21.5	(0.9)
Constructed disability status						
Hearing	2.0	1.7	(0.2)	39	1.7	(0.3)
Cognitive	14.2	14.5	(0.6)	379	13.6	(1.0)
Mental	36.2	36.1	(1.0)	1,093	36.1	(1.3)
Physical	47.6	47.6	(1.1)	1,480	48.5	(1.4)
Sex						
Female	45.9	44.8	(0.7)	1,426	46.8	(1.1)
Male	54.1	55.2	(0.7)	1,591	53.2	(1.1)
Beneficiary's age						
18–29 years	22.7	23.0	(0.7)	648	22.7	(0.9)
30–39 years	23.9	24.5	(0.7)	689	23.9	(0.8)
40–49 years	22.4	21.7	(0.6)	643	22.4	(0.7)
50 years–FRA[b]	30.9	30.7	(0.8)	1,037	30.9	(1.0)
Race/ethnicity						
White	54.9	55.8	(2.3)	1,192	54.6	(2.3)
Black	31.4	32.0	(2.2)	810	33.5	(2.1)
Hispanic	4.9	4.5	(0.5)	101	4.6	(0.6)
All others	8.8	7.8	(0.7)	189	7.3	(0.7)*
County racial/ethnic profile						
County with plurality or majority non-Hispanic black population	4.6	3.4	(1.6)	130	3.5	(1.6)
County with plurality or majority Hispanic population	9.7	8.4	(2.1)	274	8.1	(1.8)

Table 11 (*continued*)

Variable[a]	Frame percent	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
		Percent	(se)		Percent	(se)
County with majority but less than 90% non-Hispanic white population	38.7	45.5	(4.8)	1,237	44.3	(4.3)
County with racially/ethnically mixed population, no majority group	36.5	34.3	(4.6)	1,185	35.6	(4.2)
County with at least 90% non-Hispanic white population	10.2	8.4	(2.4)	189	8.4	(2.2)
Economic characteristics of county[c]						
Government-dependent economy county	12.4	11.9	(3.2)	364	12.1	(2.9)
Manufacturing-dependent economy county	8.2	8.9	(2.8)	228	9.2	(2.7)
Nonspecialized-dependent economy county	67.3	66.5	(4.5)	2,141	67.4	(4.1)
Recreation-dependent economy county	8.3	9.3	(2.9)	207	8.1	(2.4)
County with high levels of poverty	14.4	10.6	(2.8)	354	10.6	(2.4)
County with high levels of persistent poverty	6.1	3.9	(1.6)	140	4.0	(1.4)
County with high levels of persistent child poverty	15.2	12.3	(3.0)	434	12.9	(2.8)
County with low education	11.6	12.2	(2.8)	380	11.8	(2.6)
Population-loss county	8.1	4.2	(1.8)*	130	4.8	(1.8)
Retirement-destination county	12.4	13.1	(3.0)	365	13.3	(3.0)
Metropolitan status of county						
Metropolitan area of 1 million population or more	54.0	52.9	(4.8)	1,780	52.9	(4.3)
Metropolitan area of 250,000 to 999,999 population	21.9	24.3	(4.0)	752	24.3	(3.7)
Metropolitan area of fewer than 250,000 population	9.8	11.5	(3.1)	256	12.3	(3.1)
Nonmetropolitan area adjacent to large metropolitan area	3.3	3.7	(1.5)	78	3.5	(1.5)
Nonmetropolitan area adjacent to medium or small metropolitan area	6.2	4.8	(1.4)	106	4.7	(1.3)
Nonmetropolitan area not adjacent to metropolitan area	4.8	2.8	(1.0)*	45	2.3	(0.8)*
Census region						
West	22.3	22.9	(4.2)	672	21.0	(3.6)
South	33.5	33.2	(4.5)	967	35.2	(4.2)
Northeast	21.3	21.3	(4.2)	682	20.0	(3.6)

Table 11 (*continued*)

Variable[a]	Frame percent	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
		Percent	(se)		Percent	(se)
Midwest	23.0	22.6	(4.1)	696	23.8	(3.8)
Census division						
East North Central	16.0	15.2	(3.4)	490	16.0	(3.2)
West North Central	7.0	7.3	(2.7)	206	7.7	(2.5)
New England	6.7	6.2	(2.6)	223	6.0	(2.3)
Middle Atlantic	14.5	15.1	(3.7)	459	14.0	(3.0)
South Atlantic	16.3	14.6	(3.2)	455	15.3	(2.9)
East South Central	6.0	7.2	(2.4)	210	7.6	(2.4)
West South Central	11.1	11.4	(3.1)	302	12.3	(3.1)
Mountain	6.0	6.7	(2.5)	163	6.2	(2.2)
Pacific	16.3	16.2	(3.8)	509	14.8	(3.1)
Earnings categories for 2017–2018 time period						
Monthly earnings above $30,000 in at least one month in '17 or '18	22.3	22.7	(0.6)	621	22.3	(0.9)
Monthly earnings above $20,000 in at least one month in '17 or '18	23.5	23.7	(0.5)	695	23.5	(0.8)
Monthly earnings above $15,000 in at least one month in '17 or '18	19.3	19.5	(0.5)	635	19.3	(0.7)
Monthly earnings above $7,000 in at least one month in '17 or '18	21.4	21.2	(0.5)	681	21.4	(0.8)
Monthly earnings did not meet the thresholds above	13.5	12.9	(0.5)	385	13.5	(0.7)
Extract						
December 2018 extract	19.5	19.5	(0.6)	714	19.5	(0.7)
January 2019 extract	15.9	15.9	(0.4)	592	15.9	(0.6)
March 2019 extract	16.2	16.2	(0.4)	446	16.2	(0.7)
April 2019 extract	12.0	12.0	(0.4)	339	12.0	(0.5)
June 2019 extract	17.0	17.0	(0.5)	429	17.0	(0.7)
July 2019 extract	10.9	10.9	(0.3)	319	10.9	(0.6)
September 2019 extract	8.4	8.4	(0.3)	178	8.4	(0.4)

Table 11 (*continued*)

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

* Denotes a difference of more than two standard errors between the sample estimate (either from the entire sample or using the nonresponse adjusted weight) and the frame value.

Table 12. Percentages with various attributes (categorical variables) in the longitudinal SWS, comparing frame percent with final weighted estimate (using nonresponse-adjusted weights)

Variable[a]	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
	Percent	(se)		Percent	(se)
Round 6 Beneficiary type					
SSI only	27.8	(0.9)	606	27.9	(1.2)
SSDI only	51.1	(1.0)	1,074	51.1	(1.1)
Both SSI and SSDI	21.1	(0.8)	388	21.0	(1.0)
Round 6 Constructed disability status					
Hearing	3.1	(0.4)	29	2.4	(0.6)
Cognitive	12.8	(0.7)	224	12.3	(0.9)
Mental	36.6	(1.0)	713	38.2	(1.3)
Physical	47.5	(1.0)	1,074	47.1	(1.3)
Sex					
Female	45.4	(1.1)	1,032	46.3	(1.5)
Male	54.6	(1.1)	1,036	53.7	(1.5)
Beneficiary's age					
18–29 years	24.8	(0.9)	403	24.6	(1.3)
30–39 years	23.0	(0.8)	384	23.0	(1.1)
40–49 years	21.6	(0.7)	450	21.8	(1.1)
50 years–FRA[b]	30.7	(0.9)	831	30.7	(1.1)
Race/ethnicity					
White	57.6	(2.0)	865	56.8	(2.1)
Black	29.7	(1.8)	594	29.4	(1.8)
Hispanic	5.4	(0.6)	94	5.8	(0.8)
All others	7.4	(0.7)	151	8.0	(0.8)
County racial/ethnic profile					
County with plurality or majority non-Hispanic black population	3.3	(1.1)	95	3.1	(1.1)
County with plurality or majority Hispanic population	8.2	(1.4)	199	8.4	(1.3)
County with majority but less than 90% non-Hispanic white population	43.7	(3.7)	763	44.1	(3.6)

Table 12 (*continued*)

Variable[a]	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
	Percent	(se)		Percent	(se)
County with racially/ethnically mixed population, no majority group	35.5	(3.6)	866	35.5	(3.4)
County with at least 90% non-Hispanic white population	9.1	(1.9)	138	8.7	(1.7)
Economic characteristics of county[c]					
Government-dependent economy county	12.7	(2.8)	234	12.6	(2.8)
Manufacturing-dependent economy county	8.1	(1.9)	127	7.6	(1.7)
Nonspecialized-dependent economy county	66.8	(3.5)	1,505	66.7	(3.4)
Recreation-dependent economy county	9.3	(2.0)	135	9.9	(2.1)
County with high levels of poverty	11.1	(1.9)	258	11.3	(2.0)
County with high levels of persistent poverty	4.7	(1.4)	115	4.5	(1.4)
County with high levels of persistent child poverty	12.4	(2.2)	310	12.2	(2.2)
County with low education	11.3	(2.0)	279	11.3	(1.9)
Population-loss county	4.9	(1.2)	94	5.0	(1.2)
Retirement-destination county	12.4	(2.3)	194	12.3	(2.3)
Metropolitan status of county					
Metropolitan area of 1 million population or more	52.3	(3.7)	1,275	52.6	(3.5)
Metropolitan area of 250,000 to 999,999 population	24.5	(3.4)	486	25.6	(3.3)
Metropolitan area of fewer than 250,000 population	10.9	(2.2)	131	10.3	(2.1)
Nonmetropolitan area adjacent to large metropolitan area	2.9	(0.7)	39	2.5	(0.6)
Nonmetropolitan area adjacent to medium or small metropolitan area	5.8	(1.4)	81	5.4	(1.3)
Nonmetropolitan area not adjacent to metropolitan area	3.6	(0.6)	56	3.6	(0.6)
Census region					
West	22.8	(3.2)	478	23.5	(3.1)
South	31.1	(3.3)	605	30.5	(3.2)
Northeast	23.3	(3.2)	527	22.8	(3.1)
Midwest	22.7	(3.1)	458	23.2	(2.9)
Census division					
Pacific	15.8	(2.7)	376	16.6	(2.6)
East North Central	15.0	(2.5)	331	15.3	(2.3)

Table 12 (*continued*)

Variable[a]	Entire sample percent using base weights		Respondents/ ineligibles with attribute	Respondents/ineligibles weighted percent using adjusted weights	
	Percent	(se)		Percent	(se)
Middle Atlantic	16.2	(2.8)	360	15.7	(2.8)
South Atlantic	15.7	(2.5)	309	15.4	(2.4)
West South Central	9.2	(1.9)	182	9.1	(1.9)
East South Central	6.3	(2.0)	114	6.0	(1.9)
Mountain	7.1	(2.1)	102	6.9	(2.1)
New England	7.1	(2.0)	167	7.1	(1.8)
West North Central	7.7	(2.2)	127	7.9	(2.1)
Earnings categories for 2015–2016 time period					
Monthly earnings above $30,000 in at least one month in '15 or '16	20.9	(0.8)	354	20.9	(1.1)
Monthly earnings above $20,000 in at least one month in '15 or '16	23.0	(1.0)	448	23.1	(1.3)
Monthly earnings above $15,000 in at least one month in '15 or '16	19.2	(0.8)	429	19.4	(1.0)
Monthly earnings above $7,000 in at least one month in '15 or '16	24.0	(0.9)	500	23.6	(1.0)
Monthly earnings did not meet the thresholds above	12.9	(0.7)	337	13.0	(0.8)
Extract					
December 2015 extract	19.7	(0.7)	472	19.8	(0.9)
January 2016 extract	14.7	(0.6)	323	14.6	(0.7)
March 2016 extract	19.0	(0.7)	296	18.8	(0.9)
April 2016 extract	12.2	(0.5)	248	12.2	(0.7)
June 2016 extract	14.7	(0.6)	270	14.9	(0.7)
July 2016 extract	11.3	(0.5)	201	11.1	(0.6)
September 2016 extract	8.5	(0.3)	258	8.6	(0.4)

Source: NBS Round 7.

[a]The percentages for race and disability category were calculated among non-missing values. No other variables had any missing data.

[b]FRA = Full retirement age.

[c]The listed categories for "economic characteristics of county" are not mutually exclusive and the percentages do not add up to 100.

* Denotes a difference of more than two standard errors between the sample estimate (either from the entire sample or using the nonresponse adjusted weight) and the frame value.

E. Summary and implications for analyses

Our analysis has shown that, despite a few minor differences between the sample frame and the weighted estimates from the sample using base weights, the selected samples for both the RBS and cross-sectional SWS were representative of their populations of interest among variables used for either explicit or implicit stratification. Because we did not achieve an 80 percent response rate, the main purpose of this nonresponse bias analysis was to determine if systematic differences between respondents and nonrespondents were alleviated by nonresponse adjustments to the weights, or if the potential for nonresponse bias was still likely in the weighted estimates.

We found that the nonresponse and raking adjustments alleviated all differences observed between respondents and nonrespondents in all three samples. In addition, for the RBS and longitudinal SWS, it did not appear that the nonresponse and raking adjustments created new biases. However, we did see evidence that the slight underrepresentation of those in the "all others" race category in the cross-sectional SWS may have been exacerbated by the post–data collection adjustments. That said, any conclusions involving race/ethnicity should be viewed with caution due to the large amount of missing data with this variable. Although we did not find evidence that the potential for nonresponse bias exists, we found bias in the base weights representing some small populations that could not be controlled for when selecting the sample and creating the base weights. Due to their small sample counts, it was also not possible to create adjustments to the sampling weights that would accommodate their small populations. The selected sample in the RBS underestimated (using the base weights) the proportion of beneficiaries residing in population-loss counties and those residing in nonmetropolitan counties not adjacent to metropolitan counties. Further, the nonresponse and raking adjustments to the weights did not rectify this. Similarly, the selected sample in the SWS underestimated (using the base weights) the proportions of successful workers residing in nonmetropolitan areas not adjacent to metropolitan areas. The nonresponse and raking adjustments to the weights did not rectify this. Although we did not have a longitudinal SWS frame to compare the longitudinal SWS estimates to, we found that the sample estimates using the final adjusted weights were very close to those of the original sample (using the base weights).

REFERENCES

Agresti, A. *Categorical Data Analysis.* New York: John Wiley and Sons, 1990.

Biggs, D., B. deVille, and E. Suen. "A Method of Choosing Multiway Partitions for Classification and Decision Trees." *Journal of Applied Statistics*, vol. 18, 1991, pp. 49-62.

Callahan, R., E. Grau, A. Wec, K. McDonald, B. Mory, L. Pranschke, and J. Markesich. "The National Beneficiary Survey-General Waves Round 7 (Volume 3 of 3): User's Guide for Restricted and Public Use Data Files." Washington, DC: Mathematica, 2021.

Callahan, R., K. McDonald, J. Markesich and G. Livermore. "The National Beneficiary Survey-General Waves Round 7 Questionnaire." Washington, DC: Mathematica, 2021.

Cox, D.R., and E.J. Snell. *The Analysis of Binary Data, Second Edition.* London: Chapman and Hall, 1989.

Folsom, R., F. Potter, and S. Williams. "Notes on a Composite Size Measure for Self to Weighting Samples in Multiple Domains." *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 1987, pp. 792-796

Grau, E., Y. Zheng, B. Smit, B. Mory, K. McDonald, R. Callahan, H. Zhou, and J. Markesich "National Beneficiary Survey—General Waves Round 7 (volume 1 of 3): Editing, Coding, Imputation, and Weighting Procedures." Washington, DC: Mathematica, 2021.

Hosmer, D.W., Jr., and S. Lemeshow. "Goodness-of-Fit Tests for the Multiple Logistic Regression Model. *Communications in Statistics, Theory and Methods*, vol. A9, no. 10, 1980, pp. 1043-1069.

Kass, G.V. "An Exploratory Technique for Investigating Large Quantities of Categorical Data." *Applied Statistics*, vol. 29, 1980, pp. 119-127.

McDonald, K., A. Wec, R. Callahan, J. Markesich, B. Mory, and E. Grau. "The National Beneficiary Survey—General Waves Round 7: Public-Use File Codebook." Washington, DC: Mathematica, 2021.

McDonald, K., B. Mory, R. Callahan, A. Wec, and J. Markesich. "The National Beneficiary Survey—General Waves Round 7: Restricted-Use File Codebook." Washington, DC: Mathematica, 2021.

McDonald, K., R. Callahan, A. Wec, B. Mory, L. Pranschke, E. Grau, and J. Markesich. "National Beneficiary Survey—General Waves Round 7 (volume 2 of 3): Data Cleaning and Identification of Data Problems." Washington, DC: Mathematica, 2021.

Magidson, J. SPSS for Windows CHAID Release 6.0. Belmont, MA: Statistical Innovations, Inc., 1993.

This page has been left blank for double-sided copying.