National Beneficiary Survey–General Waves Round 7: (Volume 2 of 3) Data Cleaning and Identification of Data Problems

Final Report

October 20, 2021

Kim McDonald, Ryan Callahan, Aleksandra Wec, Bevin Mory, Leah Pranschke, Eric Grau and Jason Markesich

Submitted to:

Social Security Administration
Office of Retirement and Disability Policy
ITC Building
500 E Street, SW, 9th Floor
Washington, DC 20254
Project Officer: Mark Trapani
Contract Number: SS00-12-60094

Submitted by:

Mathematica
1100 1st Street, NE, 12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Jason Markesich
Reference Number: 40160.324

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

CONTENTS

TABLES

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

ACRONYMS

ADL	Activities of daily living
CAPI	Computer-assisted personal interviewing
CATI	Computer-assisted telephone interviewing
DCF	Disability Control File
IADL	Instrumental activities of daily living
ICD-9	International Classification of Diseases—9th revision
NAICS	North American Industry Classification System
NBS	National Beneficiary Survey
PSU	Primary sampling unit
RBS	Representative Beneficiary Sample
SGA	Substantial gainful activity
SOC	Standard Occupational Classification
SSA	Social Security Administration
SSDI	Social Security Disability Insurance (Title II of the Social Security Act)
SNAP	Supplemental Nutrition Assistance Program
SSI	Supplemental Security Income (Title XVI of the Social Security Act)
SVRA	State Vocational Rehabilitation Agency (also called VR)
SWS	Successful Worker Sample
TRS	Telecommunications relay service
TTW	Ticket to Work and Self-Sufficiency Program

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

NBS DATA DOCUMENTATION REPORTS

The following publicly available reports are available from SSA through 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 NBS–General Waves data collection. The report provides information covered in the Editing, Coding, Imputation and Weighting Report and the Cleaning and Identification of Data Problems Report—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 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** (current report). 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 types of problems encountered and general recommendations.
- **NBS Nonresponse Bias Analysis** (Grau et al. 2021). 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 existed.

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

The National Beneficiary Survey-General Waves (NBS-General Waves), sponsored by the Social Security Administration's (SSA's) Office of Retirement and Disability Policy, 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, or SNAP). Rounds 1–4 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.

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 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 deployed in-person field locators to follow-up with some CATI non-respondents[1] and we offered computer-assisted personal interviewing (CAPI) to sample members who preferred or needed an in-person interview to accommodate their disabilities.[2]

Mathematica conducted an extensive review of the NBS data in order to identify data problems before analysis. In the following discussion, we describe the data processing procedures that we performed for Round 7 of the NBS–General Waves. In the remaining sections of Chapter I, we provide an overview of the NBS-General Waves, including the objectives of the study, sample design, survey procedures, and final response rates. In Chapter II, we summarize the NBS data collection instrument while in Chapter III, we describe the data coding and cleaning procedures and highlight the main data issues that we encountered. We present a section-by-section summary of findings in Chapter IV, and provide concluding comments in Chapter V.

A. NBS–General Waves Objectives

The NBS–General Waves collects important beneficiary data that are not available from SSA administrative data or other sources. The survey addresses five major questions:

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

[1] 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 this chapter, in Section B.1. 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 this chapter, in Section B.2.

[2] In Round 7, none of the NBS respondents requested a CAPI interview.

2. What are the short-term and long-term employment outcomes for SSI and SSDI beneficiaries who work?
3. What supports help SSA beneficiaries with disabilities find and keep jobs and what barriers to work do they encounter?
4. What are the characteristics and experiences of beneficiaries who work?
5. What health-related factors, job-related factors, and personal circumstances hinder or promote employment and self-sufficiency?

SSA will combine data from 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 the survey data for policymaking and program planning efforts.

We addressed the core research questions in Rounds 1 through 4 through two surveys, one of all beneficiaries (the RBS) and one of participants in the TTW program (the Ticket Participant Sample, or TPS). The NBS–General Waves (Rounds 5 through 7) no longer focuses on TTW. The survey design for Rounds 5 through 7 initially called for three national cross-sectional surveys of SSI and SSDI beneficiaries (the RBS)—one each in 2014, 2016, and 2018. It also called for cross-sectional surveys, in the same years, of beneficiaries whose benefits were suspended or terminated due to work (with a subset followed longitudinally across rounds). However, due to difficulties in identifying beneficiaries experiencing benefit suspense in SSA's administrative data, we subsequently revised the design to focus on beneficiaries with successful work attempts (the SWS). We delayed the start of NBS–General Waves by one year (from 2014, 2016, and 2018, to 2015, 2017, and 2019) to allow time to redesign the successful worker portion of the survey and sample, and ultimately opted not to administer the SWS in Round 5. In lieu of the Round 5 SWS survey, we conducted in-depth qualitative interviews with 91 successful workers about their experience with benefits and their attempts to find and keep a job (O'Day et al. 2016). In Round 6, we conducted the second cross-sectional survey for the RBS in the NBS–General Waves, using the same primary sampling units (PSUs) that were selected in Round 5, while simultaneously conducting the first cross-sectional survey for the SWS. In Round 7, we conducted the third cross-sectional survey for the RBS in the NBS–General Waves,[3] the second cross-sectional survey for the SWS, and a longitudinal follow-up survey for a subset of SWS cases from Round 6. Table I.1 shows the samples that were processed in Rounds 1 through 7.

[3] Although this is the third RBS in the NBS–General Waves, it is the seventh RBS over the entire history of the NBS project.

Table I.1. Summary of Samples Processed in Rounds 1 through 7 [a]

Round	Year	Study	RBS	TPS	SWS	Longitudinal SWS
1	2004	NBS-TTW	√	√		
2	2005	NBS-TTW	√	√		
3	2006	NBS-TTW	√	√		
4	2010	NBS-TTW	√	√		
5	2015	NBS-General Waves	√			
6	2017	NBS-General Waves	√		√	
7	2019	NBS-General Waves	√		√	√

[a]Qualitative interviews were also conducted in Round 5 of the NBS-General Waves, in 2015.

Source: NBS Round 7.

B. NBS–General Waves Sample Design Overview

For all survey rounds, the NBS 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. The longitudinal SWS was composed of all sample cases that (1) completed a Round 6 SWS interview and (2) reported to be currently working at the time of the Round 6 survey. [5] Table I.2 summarizes the actual sample sizes and number of completed interviews for the RBS, cross-sectional SWS, and longitudinal SWS.

1. RBS

For the RBS in Round 7, 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 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, 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 who resided in all 50 states and the District of Columbia, excluding outlying territories, and who were in an active pay status as of June 30, 2018.[6] As of that date, the target population consisted of approximately 13.7 million beneficiaries. As in prior rounds, we stratified the

[4] The RBS and the main sample of the SWS involved selecting individuals within selected clusters of geographic areas, and they are therefore referred to as "clustered samples." The supplemental sample (for the SWS only) was sampled 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.

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

[6] 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. Active status does not include beneficiaries whose benefits have been terminated.

cross-sectional RBS by four age-based strata within the PSUs: (1) 18- to 29-year-olds, (2) 30- to 39-year-olds, (3) 40- to 49-year-olds, and (4) 50-year-olds and older. To ensure a sufficient number of persons seeking work, we oversampled beneficiaries in the first three cohorts (18- to 49-year-olds). 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 during CATI data collection, instead of fielding all of these cases. Because the length of the Phase 1 protocol varied on a case-by-case basis, not all cases were ready for Phase 2 after 12 weeks. By week 15, 27 percent of the total cases that would be selected for Phase 2 were active in Phase 2. By week 25, this increased to 75 percent. By week 35, nearly all Phase 2 selected cases were active in Phase 2. The two-phase 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 is the same regardless of what proportion of Phase 1 nonrespondents is subsampled for Phase 2.

2. Cross-sectional SWS

The cross-sectional SWS 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/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 months at any time between August 1, 2018, and July 31, 2019, and (2) be less than 62 years old on June 30, 2018.[7] 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.[8] Using these constraints to define the target population, we identified a population of 101,698 successful workers.[9] From this frame, we fielded

[7] We used a 62-year age limit in Round 6 to ensure that longitudinal cases would still be under 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 62-year age limit in the Round 7 cross-sectional sample for the sake of consistency with Round 6.

[8] Some SSI/SSDI beneficiaries would be considered successful workers because their earnings and age met the threshold, but 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 exists 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.

[9] 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.

a nationally representative sample of 8,590[10] successful workers. We included a screening question as an additional constraint: the sampled successful workers had to indicate that they had been working in the past six months.[11] To ensure a large enough number of successful workers for sampling, we selected seven samples of successful workers from seven successive mutually exclusive frames. 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).

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 PSU, and successful workers not residing in any of the PSUs).[12] We refer to the initial sample design as the "clustered" sample, and to the second independent sample as the "unclustered" sample.[13] We call the combination of data from the clustered and unclustered samples to calculate estimates a "paired" 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. We created a single set of composite weights that combined information from the clustered and unclustered SWS, appropriately accounting for the different follow-up rules between the two samples.[14] Table I.2 includes the total across the two samples in the SWS, and does not break out the counts between clustered and unclustered samples; the 152 duplicate cases that were selected for both the clustered and unclustered samples are counted twice in this table.

3. Longitudinal SWS

The Round 7 longitudinal sample consists of Round 6 cross-sectional SWS respondents who indicated that they 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 due to work [15] on June 30, 2016, (2) had earnings above SSA's nonblind SGA earnings level[16] for at least three

[10] For reasons explained later in this chapter, the cross-sectional SWS includes 152 duplicates in the sample. As a result, 8,438 unique cases were sampled.

[11] 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.

[12] 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 these areas.

[13] 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 (discussed later).

[14] These composite weights, combining 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.

[15] "Suspense status due to work" refers to the beneficiaries whose benefits have been temporarily suspended because of work. Those in suspense status for other reasons were not eligible for the sample.

[16] This threshold was $1,090 in 2015 and $1,130 in 2016.

consecutive calendar months at any time from August 1, 2016, through July 31, 2017; and (3) were younger than 62 on June 30, 2016. (This is the same definition for successful workers that we used 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 I.2. NBS–General Waves Round 7 Actual Sample Sizes, Target Completes, and Completes

Sampling strata	Selected sample size	Original target completed interviews[a]	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
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.

Table I.2 (*continued*)

[a]The target completed interviews for the SWS shown here were calculated prior to receiving the first extract, using data from simulated successful worker populations in 2011-12, 2013-14, 2015-16, and Round 6 of the NBS. In fact, there were actually seven allocations, with a new sample allocation calculated after the population sizes for each extract were revealed. This explains the sometimes large deviation between the target allocation and the actual number of completed interviews.

C. Round 7 Survey Overview

The NBS was designed and implemented to maximize both response and data quality. Table I.3 describes the most significant sources of potential error identified at the outset of the NBS and how we attempted to minimize the impact of them. A more detailed discussion of our approach to minimizing total survey error can be found in Appendix A of the Round 7 User's Guide (Callahan et al. 2021).

Table I.3. Sources and Descriptions of Potential Error, Description, and Methods to Minimize Impact

Source of Error	Description	Method to Minimize Impact
Sampling	Error that results when characteristics of the selected sample deviates from the characteristics of the population.	Select a large sample size; select PSUs with probability proportional to size, basing the measure of size for each PSU on the counts of beneficiaries in the study population; use stratified sampling by age categories to create units within each stratum that are as similar as possible.
Specification	An error that results when the concept intended to be measured by the question is not the same as the concept the respondent ascribes to the question.	Cognitive interviewing during survey development[a] and pre-testing; use of proxy if sample member unable to respond due to cognitive disability
Unit Nonresponse	An error that occurs when a selected sample member is unwilling or unable to participate (failure to interview). This can result in increased variance and potential for bias in estimates if nonresponders have different characteristics than responders.	Interviewer training; intensive locating, including field locating; in-person data collection; refusal conversion; incentives; nonresponse adjustment to weights
Item Nonresponse	An error that results when items are left blank or the respondent reports that he or she does not know the answer or refuses to provide an answer (failure to obtain and record data for all items). This can result in increased variance and potential bias in estimates if nonresponders have different characteristics than responders.	Use of probes; allowing for variations in reporting units; assurance of confidentiality; assistance during interview; use of proxy if sample member unable to respond due to cognitive disability; imputation on key variables
Measurement	An error that occurs as a result of the respondent or interviewer providing incorrect information (either intentionally or unintentionally). This may result from inherent differences in interview mode.	Use of same instrument in both interview modes; use of probes; adaptive equipment; interviewer training, validation of field interviews; assistance during interview; use of proxy if sample member is unable to respond due to cognitive disability
Data Processing	An error in data entry, coding, weighting, or analyses.	Coder training; monitoring and quality control checks of coders; quality assurance review of all weighting and imputation procedures

Source: NBS Round 7.

[a] Conducted during survey development phase under a separate contract held by Westat.

We did not expect item nonresponse to be a large source of error because the survey contained few obviously sensitive items. In fact, item nonresponse was greater than 6 percent only for select items asking for wages and household income. Unit nonresponse was the greater concern given the population; thus, the survey was designed with a dual-mode approach. Mathematica made all initial attempts to interview beneficiaries using CATI. If a sample member could not participate in the survey because of an intellectual disability, even with help from a friend or family member, Mathematica sought a proxy respondent. To promote response among Hispanic sample members whose primary language is Spanish, we translated the questionnaire into Spanish. For languages other than English or Spanish, interpreters, if available in the sample person's home, helped to conduct the interviews. If no one in the household was available to interpret for the respondent, then we flagged the case as a "language barrier." We elected not to use an interpreter service to help contact and complete these cases in Round 7, as we did not have a sufficient number in any language that made it cost effective to do so. At the conclusion of Round 7 data collection, we dispositioned 117 cases as language barriers. We made a number of additional accommodations for those with hearing or speech impairments, including using a telecommunications relay service (TRS) and amplifiers.

If Mathematica could not locate and contact a sample member by telephone and the case was selected for field follow-up, we deployed a field locator to make contact in person. After locating the sample member, the field locator attempted to facilitate an interview with him or her via CATI, using a cell phone (or the sample member's own phone, if preferred) to call into the data collection center. If a sample member could not complete the interview by telephone in this manner due to his or her disability, trained field staff were available to conduct the interview in person using CAPI. In Round 7, none of the NBS respondents requested a CAPI interview.

We began the Round 7 CATI data collection effort in February 2019. In May 2019, Mathematica began in-person locating, which continued concurrently with CATI interviewing through November 2019.

1. Completed Interviews and Response Rates

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.[17] Because of the independence of the RBS and cross-sectional SWS sample selections, the independence of the clustered and unclustered sample selections within the cross-sectional SWS, and the independence of 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[18] Mathematica completed all of these interviews by telephone.

[17] 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.

[18] 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

During Round 7, we completed proxy interviews for 1,113 sample members in the RBS, for 293 sample members in the SWS, and for 151 sample members in the longitudinal SWS, for a total of 1,557 proxy interviews across the three sample groups. In addition, we completed a total of 242 interviews in Spanish--104 in the RBS, 74 in the SWS, and 64 in the longitudinal SWS.

The weighted response rates for Round 7 of the NBS are 54.7 percent for the RBS, 41.0 percent for the cross-sectional SWS, and 54.5 percent for the longitudinal SWS. Please see the Round 7 User's Guide (Callahan et al. 2021) for detailed information on the final survey dispositions.

2. Nonresponse Bias

Because the weighted response rates were less than 80 percent for both samples, we conducted a nonresponse bias analysis at the end of data collection. We examined all 11,299 selected sample cases in the RBS, all 18,590 selected sample cases in the cross-sectional SWS, and all 3,712 cases in the longitudinal SWS to determine if there were systematic differences between respondents and nonrespondents for a variety of covariates. Our analysis revealed differences between respondents and nonrespondents for some variables, but the nonresponse adjustments to the weights appear to have eliminated all such differences in both samples. 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 potential sources of bias for some small populations representing county-level economic indicators, but this was unrelated to nonresponse. In these cases, the weighted estimates of the small populations differed from those in the frame because we did not control for those populations when we created the initial sampling weights. This was because the variables representing these populations (1) were not considered important enough to be used as variables for either sample stratification or post-survey raking, and (2) were not included as covariates in the final nonresponse models, generally because the samples were too small. We therefore could not reconcile these differences when adjusting these weights for nonresponse or when post-stratifying them to marginal population totals.

The full nonresponse bias analysis can be obtained from SSA (https://www.ssa.gov/disabilityresearch/nbs_round_7.html).

SWS. The counts of ineligible cases included 16 duplicates; the number of unique ineligible cases across both samples was 601.

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

II. DESCRIPTION OF THE NBS–GENERAL WAVES INSTRUMENT

The NBS collects data on a wide range of topics—including, employment, disability, experience with SSA programs, employment services used in the past year, health and functional status, health insurance, income and other assistance, and sociodemographic information. Under a separate contract, Westat developed and initially pre-tested the survey items. Mathematica subsequently made revisions to the survey items to prepare the instrument for CATI/CAPI programming and made minor wording changes in response to pre-testing results. For Round 7, we added 11 new questions to the instrument to capture information on longitudinal sample members' previous employment. We also added probes to several questions that longitudinal sample members answered in Round 6, and new response options to several "other/specify" questions. Finally, we revised a few questions in order to accommodate changes in reference periods and changes in federal programs. A detailed description of the changes between Round 6 and the Round 7 questionnaire is included in Appendix A. Prior to the launch of the Round 7 data collection effort, we pretested the survey instrument to confirm the order, flow, and clarity of the revisions. The instrument is available from SSA (https://www.ssa.gov/disabilityresearch/nbs_round_7.html).

To promote responses among Hispanic populations, Mathematica translated the questionnaire into Spanish. Certified bilingual interviewers administered the Spanish interviews. If a Spanish speaker was more familiar with a word or term in English than in Spanish, we provided the term in both languages—allowing interviewers to reinforce the question by using the second language as a probe, if necessary.[19] We treated measurements in a similar way. Questions that mentioned a particular weight also mentioned the kilogram equivalent.[20] For languages other than English or Spanish, interpreters, if available in the sample person's home, helped conduct the interviews. If no one in the household was available to interpret for the respondent, then we flagged the case as a "language barrier." We elected not to use an interpreter service to help contact and complete these cases, as we did not have a sufficient number in any language that made it cost effective to do so.

A. Summary of Modules

The questionnaire is divided into 13 sections, labeled A through M:

- Section A—Introduction and Screener
- Section B—Disability and Current Work Status
- Section C—Current Employment
- Section C_B—Employment in Past 6 Months
- Section D—Jobs/Other Jobs During 2018
- Section SC—Benefit Suspense

[19] For example, on Item K11: Did {you/NAME} receive any food stamps last month? Spanish: Recibió {usted/NAME} food stamps o cupones de alimentos el mes pasado?

[20] For example, on Item I35: {Do you/Does NAME} have any difficulty lifting and carrying something as heavy as 10 pounds, such as a full bag of groceries? Spanish: Tiene {usted/NAME} cualquier dificultad en levantar y cargar algo que pesa hasta unas 10 libras {5 kilos}, tal como una bolsa llena con compras del mercado?

- Section E—Awareness of SSA Work Incentive Programs
- Section G—Employment-Related Services and Supports Used in 2018
- Section I—Health and Functional Status
- Section J—Health Insurance
- Section K—Income and Other Assistance
- Section L—Sociodemographic Information
- Section M—Closing Information and Observations

Descriptions of each section follow.

1. Section A—Introduction and Screener

This section confirms that the interviewer has contacted the correct sample person and verifies that the sample person is still eligible for the survey. Ineligible respondents are deceased, incarcerated, not living in the continental United States, are active duty military, or have not received any SSA disability benefits in the last five years.[21] Additionally, cross-sectional SWS respondents who are not currently working, and did not work in the last six months, are ineligible for the survey. The screener allows interviewers to do the following:

- **Identify any barriers to participation** and, if needed, identify a proxy respondent. The sample member is offered every opportunity to complete the interview himself or herself; a proxy responds only if necessary.
- **Identify the need for an interpreter** for a respondent who speaks a language other than English or Spanish.
- **Administer a cognitive assessment** to ensure that the respondent is capable of completing a complex survey. This assessment includes elements of informed consent for participation, as it provides respondents with an overview of the survey and informs them of the voluntary nature of the interview.

We present three statements in the screener: (1) a brief description of what it means that the survey is confidential, (2) what it means that the survey is voluntary, and (3) an overview of the study topics. Then we ask respondents to reiterate the concepts in their own words. If a respondent cannot restate a concept, the question is read a second time. If the respondent still cannot restate a concept, we ask if someone else (such as a friend, parent, caseworker, or payee) can answer questions about the respondent's health, daily activities, and jobs. We then pursue an interview with the proxy respondent, if available. Proxy respondents are administered the same cognitive assessment to ensure that they are capable of completing the complex survey on the sample member's behalf and also understand the voluntary nature of the survey. To minimize bias in reporting, we do not ask the proxy respondent to provide subjective assessments on behalf of the sample person with respect to, for example, satisfaction with jobs or programs. The constructed variable C_Rtype indicates whether the sample person or a proxy completed most of the interview.

[21]The screening of respondents who had not received any SSA benefits in the last five years occurs in Section B of the instrument.

2. Section B—Disability and Current Work Status

This section collects information on the beneficiary's limiting physical or mental conditions and current employment status. We ask about the disability status of the beneficiary by identifying the health condition or conditions that affect the beneficiary's work or daily activities and the age at which the condition first began limiting the beneficiary's activities. Then, we ask questions about the beneficiary's current work status. If a beneficiary is not currently employed, we explore their reasons for not working, the reasons that health prevents work among those so indicating, and the reasons why they have no expectation of working in the near future or no expectation of receiving benefits. For all respondents who became limited as an adult (after the age of 18), we ask about the respondent's ability to perform the same job they performed before they started to receive disability benefits. If a respondent reports that they have not received SSA disability benefits in the last five years, they are determined to be ineligible for the survey. If a cross-sectional SWS respondent reports that they are not currently working and have not worked in the last six months, they are ineligible for the survey. For those longitudinal SWS cases that are not currently working, we ask why they left their last job. We also ask questions to determine the job characteristics that are important to beneficiaries and collect information about work-related goals and expectations.

3. Section C—Current Employment

In this section, we collect detailed information about the beneficiary's current job. We ask beneficiaries for information about their job, such as job title, the type of work performed, type of employer, hours worked, benefits offered, how they found their job, and wages earned. These questions are asked for each job that the beneficiary currently holds. We also ask questions about the beneficiary's primary job (if they have more than one job), including questions about work-related accommodations—those received as well as those needed but not received. We ask additional questions to determine if the beneficiary's employer made changes to the workplaces to help the beneficiary work. We solicit information about job satisfaction. We ask respondents about their motivation for working, the formal and informal supports they use to find or keep a job, the features of their current job that allow them to work with a disability, and the various challenges they face in their current job. We also ask questions that address disability disclosure in the workplace, whether other people with disabilities are employed at the respondent's place of work, and whether a benefit overpayment affected employment.

4. Section C_B—Employment in Past 6 Months

Questions in this section collect information about employment in the last six months, if the respondent is not currently working. We ask beneficiaries for information about all of the jobs they have worked in the last six months, including the type of employer; hours worked; benefits offered; how they found their job; wages earned; and the reasons for leaving employment, if applicable. We also ask whether beneficiaries worked or earned less than they could have (and, if so, why) and collect information about their experiences with adjustments to social security benefits due to work. We ask beneficiaries about their motivation for working in the last six months, the formal and informal supports they used to find or keep their main job, the features of their former main job that allowed them to work with a disability, and the various challenges they faced in their former job. We ask questions that address disability disclosure in the workplace, whether other people with disabilities were employed at the respondent's main place of work, and whether a benefit overpayment affected employment.

5. Section D—Jobs/Other Jobs During 2018

Questions in this section collect information about employment during the 2018 calendar year, excluding jobs noted in Section C or Section C_B. For example, we ask beneficiaries questions about the type of employer; hours worked; wages earned; and the reasons for leaving employment, if applicable. In other questions, we ask whether beneficiaries worked or earned less than they could have (and, if so, why) and collect information about their experiences with adjustments to social security benefits due to work (including if their work activity was affected by a disability overpayment).

6. Section SC—Benefit Suspense

This section is asked only of beneficiaries who are currently employed, or who have been employed within the last six months or in 2018. It asks beneficiaries how their work experiences have affected their social security disability benefits. Questions in this section differentiate between three types of beneficiaries: 1) beneficiaries who have not received a suspension of benefits because of employment in the past year, 2) beneficiaries who are no longer receiving social security benefits due to recent employment, and 3) beneficiaries who received a suspension of social security benefits because of employment in the past year, but are now receiving benefits again. If beneficiaries are currently experiencing a suspension of benefits, or did so in the last year, we ask them for more information about the factors that affected their benefit receipt, specifically factors related to health, employment, and personal circumstances.

7. Section E—Awareness of SSA Work Incentive Programs

In this section, we ask questions to assess whether the beneficiary is aware of or is participating in SSA work incentive programs and services, including where they obtain information about SSA programs. We inquire if beneficiaries are aware that their SSDI cash benefits cease if their earnings exceed the substantial gainful activity threshold after completing the trial work period. We also ask a question to measure whether sample members are aware that most people who start working and lose their disability benefits are able to keep their health insurance.

8. Section G—Employment-Related Services and Supports Used in 2018

Questions in this section ask beneficiaries about their use of employment-related services and supports in calendar year 2018, including employment, job training, medical, therapy or counseling, and educational services. We also ask sample members about their reasons for, and satisfaction with, services and the nature of any services needed but not received.

9. Section I—Health and Functional Status

In this section, we ask about the beneficiary's health status and daily functioning, including the need for special equipment or assistive devices. We ask for information about general health status (via the SF-8^{TM} scale), unmet health needs, informal supports, difficulties with activities of daily living (ADLs) and instrumental activities of daily living (IADLs), functional limitations, substance abuse or dependence, and

treatment for mental health conditions.[22] In addition, we ask about episodic poor health, number of days confined to a bed, informal supports for daily needs, and transportation usage.

10. Section J—Health Insurance

Questions in this section collect information about the beneficiary's sources of health insurance, both at the time of interview and during calendar year 2018.

11. Section K—Income and Other Assistance

In this section, we ask about sources of income, including income received from earnings, social security, workers' compensation, and other government programs and sources. Additionally, we ask sample members about their perception of their financial situation and ability to save for an emergency or crisis.

12. Section L—Sociodemographic Information

This section collects basic demographic information about the beneficiary, such as race, ethnicity, education, parental education, veteran status, height and weight, marital status, living arrangements including homeownership and possible plans to relocate, and household income.

13. Section M—Closing Information and Observations

In this section, we collect address information for the sample person so that the $30 incentive gift card may be mailed. The interviewer also records the reasons that a proxy or other assistance was required, if appropriate, and documents special circumstances.

B. Instrument Pathing and Preloaded Data

Round 7 of the NBS–General Waves required 65 minutes to administer on average. The interview length ranged from 30 minutes to 197 minutes, excluding TRS interviews.

Interviewers asked all respondents questions from Sections A, B, E, G, I, J, K, L, and M. Only respondents who reported that they were currently working answered the questions in Section C. Similarly, only respondents who reported working in the last six months received Section C_B, and only respondents who reported working in 2018 answered the questions in Section D. RBS and SWS respondents who reported working currently, in the last six months, or in 2018, received Section SC. In Round 7, longitudinal SWS respondents also received Section SC, even if they reported not working currently or in the last six months or in 2018. Table II.1 provides a summary description of the main questionnaire pathing.

[22] SF-8™ is a trademark of QualityMetric, Inc.

Table II.1. NBS–General Waves Instrument Sections

Section	Title of Section	Respondents Receiving the Section
A	Introduction and Screener	All respondents
B	Disability and Current Work Status	All respondents
C	Current Employment	Respondents who are currently working (B24 = YES) Question B24: Are you currently working at a job or business for pay or profit?
C_B	Employment in the Last 6 Months	Respondents who are not currently working, but who worked in the last 6 months {B24 = NO and B24b = YES) Question B24b: Did you work for pay or profit at any time during the last 6 months?
D	Jobs/Other Jobs During 2018	Respondents who worked in 2018 (B30 = YES) Question B30: Did you work at a job or business for pay or profit any time in 2018?
SC	Benefit Suspense	All longitudinal SWS respondents and RBS or cross-sectional SWS respondents who are currently working, have worked in the past 6 months, or worked in 2018 (LONGSAMP = YES or B24 = YES or B24b = YES or B30 = YES)
E	Awareness of SSA Work Incentive Programs	All respondents
G	Employment-Related Services and Supports Used in 2018	All respondents
I	Health and Functional Status	All respondents
J	Health Insurance	All respondents
K	Income and Other Assistance	All respondents
L	Sociodemographic Information	All respondents
M	Closing Information and Observations	All respondents

Source: NBS Round 7.

The NBS–General Waves instrument, which Mathematica programmed in Blaise, is complex and involves several integrated skips within and across sections. The use of preloaded SSA administrative data and allowances for proxy participation introduce further complexities into the questionnaire pathing. Preloaded data on respondents' disability benefits status (SSI, SSDI, or both) and age at which respondents first received SSI benefits determine pathing for certain survey items. A longitudinal SWS indicator (LongSamp) was used to determine pathing for longitudinal SWS respondents. Longitudinal SWS respondents who reported that they were not currently working, and had not worked in the past six months or in 2018 were asked a new series of questions about the reasons they left their last job. Other administrative variables serve as fills for particular items to provide respondents with names of local programs or to prompt recognition of program participation. Table II.2 provides a list and description of the preloaded variables.

Table II.2. Survey Preloads

Variable	Definition	Purpose
Bstatus	SSA benefit type (SSI only, SSDI only, or SSI and SSDI) received by sample member	Used to determine pathing for awareness of SSA work incentive items. Only respondents who received SSDI benefits were asked items E3 through E12. Only respondents who received SSI were asked items E15 and E17.
DOB	Sample member date of birth	Reported date of birth (or age) matched with administrative data to verify that the correct person was contacted in the screener portion of the survey.
SSIage	Age at which sample member first received SSI benefits	Used to determine pathing at item E12. Only respondents who received SSI before age 22 (and were 25-years-old or younger) were asked this item.
StateMed	State name for Medicaid, based on state of residence reported at time of survey	Used at item J2 to identify, by name, the Medicaid program in the respondent's state.
VRname	State name for State Vocational Rehabilitation Agency, based on state of residence reported at time of survey	Used at items B29 and to identify, by name, the State Vocational Rehabilitation Agency in the respondent's state.
SampGrp	Sample group (RBS or SWS)	Used to screen cross-sectional SWS respondents who have not worked within the last six months (A73b and B24c) and to collect information on future moves for SWS respondents (M2c).
LongSamp	Successful Worker Sample, Longitudinal case	Used to determine question pathing for longitudinal sample members. Longitudinal respondents were not screened out of the survey because of their current work status. Instead, longitudinal SWS members who were not working at the time of the interview, did not work in the past six months, and did not work in 2018, were asked questions about the reasons they left their last job (B36c-B36f). Longitudinal SWS respondents were also asked about benefit suspense (Section SC), even if they were not currently working, did not work in the past six months, and did not work in 2018.

Source: NBS Round 7.

Given that proxies are needed when the sample member's disability precludes participation, we programmed the instrument to fill in the proper pronoun or name in the question text after the interviewer indicated that the survey respondent would be either a sample member or a proxy. In addition, the instrument was programmed to skip attitudinal and opinion items for proxy respondents to minimize bias in reporting. (See Table II.3 for a complete list of items not asked of proxy respondents.) As mentioned previously, interviewers completed 1,557 proxy interviews.

Table II.3. Items Skipped for Proxy Respondents

Survey Item	Question Text
B29_3a	You said that one of the reasons you did not accept a job you were offered was because it did not pay enough. What is the lowest wage or salary you would have accepted for this job?
B29_3b	If you did get a job offer that matched your current needs and abilities, what is the lowest wage or salary you would be willing to accept for such a job?
B29_8a	You said that one of the reasons you are unable to find a job is that the jobs that are available do not pay enough. What is the lowest wage or salary you would accept for a job that matched your current needs and abilities?
B29_8b	If you did get a job offer that matched your needs and abilities, what is the lowest wage or salary you would be willing to accept for such a job?
B29_8c	How many hours per week would you expect to work for this amount of pay?
B29_8d	Would you expect to work full-time or part-time?
B29_12a	If you did get a job offer that matched your current needs and abilities, what is the lowest wage or salary you would be willing to accept for such a job?
B29_12b	How many hours per week would you expect to work for this amount of pay?
B29_12c	Would you expect to work full-time or part-time?
C18	Taking all things into account, how satisfied are you with your {main/current} job? Would you say very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied?
C39a – C39h	Again, thinking about your {main/current} job, how much do you agree with each of the following statements? Would you say you strongly agree, agree, disagree, or strongly disagree?
C39a	You had a chance to develop your abilities.
C39b	You had recognition or respect from others.
C39c	You could work on your own in your job if you wanted to.
C39d	You could work with others in a group or team if you wanted to.
C39e	Your work was interesting or enjoyable.
C39f	Your work gave you a feeling of accomplishment or contribution.
C39g	Your supervisor was supportive.
C39h	Your co-workers were friendly and supportive.
C_B18	Taking all things into account, how satisfied are you with your {main/current} job? Would you say very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied?
C_B39a – C_B39h	Again, thinking about the {main} job {you/NAME} had within the past six months, how much do you agree or disagree with each of the following statements? Would you say you strongly agree, agree, disagree, or strongly disagree?
C39a	You had a chance to develop your abilities.
C39b	You had recognition or respect from others.
C39c	You could work on your own in your job if you wanted to.
C39d	You could work with others in a group or team if you wanted to.
C39e	Your work was interesting or enjoyable.
C39f	Your work gave you a feeling of accomplishment or contribution.
C39g	Your supervisor was supportive.
C39h	Your co-workers were friendly and supportive.

Source: NBS Round 7.

During the study introduction, we informed respondents that we could stop the interview and resume it at a later date and/or time if they began to tire or otherwise felt that they could not continue with the interview. We also trained interviewers to periodically ask respondents about their level of fatigue during

the course of the interview. If an interviewer sensed that a respondent was tiring, they asked the respondent if it was okay to continue with the interview or if they needed to complete it in another call. In Round 7, 2,928 sample members, or about 12 percent of the total sample, broke off the interview after they had completed the cognitive screener. Of these cases, we completed 1,888 interviews, or about 64 percent of the total number of break off cases. Table II.4 provides additional detail on break-offs by sample group.

Table II.4. Round 7 Break-Off Cases by Sample Group

	Selected sample size	Number of break-off cases	Percentage of sample that broke off	Number of break-off cases completed	Percentage of break-off cases completed
Representative Beneficiary Sample	11,299	1,260	11%	741	59%
Successful Worker Sample	8,590	1,029	12%	669	65%
Longitudinal Successful Worker Sample	3,712	639	17%	478	74%
Total	23,601	2,928	12%	1,888	64%

Source: NBS Round 7.

C. Changes Made to Survey Instrument in Round 7

Mathematica modified the survey instrument prior to administration in Round 7. In Section 1 below, we describe the questions that we added to the Round 7 instrument. These items pertain to the longitudinal SWS members who were interviewed in Round 6. In Section 2, we discuss the modifications that we made to some of the questions from Round 6. No questions were removed between Round 6 and Round 7. A detailed description of the changes between the Round 6 and the Round 7 questionnaire is included in Appendix A.

1. New Questions Applicable to Longitudinal Sample Members

For Round 7 survey administration, we added a few questions to account for longitudinal sample members. Table II.5 summarizes the new questions that we added to the Round 7 instrument.

Table II.5. New Questions in Round 7

Question Number	Topic Addressed
Disability and Work Status (Section B)	
B36c, B26c_1, B36c_1_oth, B36d, B36d_1, B36d_1_oth, B36e, B36e_1, B36e_1_oth, B36f, B36f_oth	Reasons longitudinal sample members are no longer working

Source: NBS Round 7.

2. Other Modifications to the Round 7 Survey Instrument

Mathematica made several minor modifications to the Round 6 NBS instrument for administration in Round 7 of the NBS, including (1) changing reference periods from 2016 to 2018, (2) updating items to reflect changes in SSA programs or policies, (3) improving question wording and adding response categories, and (4) modifying skip logic.

Changes to the Reference Period. The NBS Round 7 was administered in 2019. As a result, we updated year references for questions and response categories. For example, in Section D (Jobs/Other Jobs in 2018), we changed the reference year from 2016 to 2018. Similarly, in Section G (Employment-Related Services and Supports in 2018), we changed the reference year from 2016 to 2018. Further, on items asking about the year in which services were last received, we changed the response options from "in 2016" or "before 2016" to "in 2018," or "before 2018," respectively.

The change in the reference period also necessitated changes to the upper bound of soft and hard edit checks for certain numeric items. For example, in Section C (Current Employment), we changed the upper bound for the year in which the respondent started his or her current job from 2017 to 2019 because Round 7 was fielded in that year.

Changes to Reflect Changes in SSA Programs or Policies. In some instances, we updated items to reflect the 2019 dollar amounts for some SSA work support provisions (e.g., trial work period).

Changes to Question Wording and Response Categories. For a few items, we revised the question wording slightly, added interviewer probes, and/or adjusted response categories. We made these changes as a result of (1) the need to incorporate the longitudinal SWS into the Round 7 instrument and (2) lessons learned during Round 6 data processing. We added interviewer probes for the longitudinal SWS respondents that acknowledged that we spoke to them in 2017, that we would like to conduct another interview, and that some questions may sound similar to their 2017 interview. We also modified some of the skip logic to ensure that longitudinal cases were not deemed ineligible for the survey if they had not worked recently. After completing the Round 6 data processing and back-coding tasks, we added new response options to a number of questions in the instrument. Chapter III provides more detail on this process and the new response options are included in Appendix A.

Modifying Skip Logic. During the Round 6 data cleaning effort, we discovered an error in the skip logic affecting item L10 (Do you have a long-term partner who lives in the same household?), which is one of the source variables for the cohabitation status constructed variable, C_COHAB. We corrected the skip logic in the Round 7 survey instrument to ensure that unmarried sample members who are living with their partner (L8=6 and L9=1) were not asked L10, and divorcees (L8=3) were asked L10.

III. ROUND 7 DATA PROCESSING

A. Coding of Open-Ended and Verbatim Responses

The NBS questionnaire includes several questions designed to elicit open-ended responses. To make it easier to analyze the data connected with these responses, we grouped the responses and assigned them numeric codes when possible. The methodology used to code each variable depended upon the variable's content. Three types of questions did not have designated response categories; rather, responses to these questions were recorded verbatim:

1. **Open-ended questions** have no response options specified. For example, item G61 asks, "Why {were you/was NAME} unable to get these services?" For such items, interviewers recorded the verbatim response. Using common responses, we developed categories and reviewed them with analysts. Coders then attempted to code the verbatim response into an established category. If the response did not fit into one of the categories, coders coded it as "other."
2. **Other/specify** is a response option for questions with a finite number of possible answers that may not necessarily capture all possible responses. For example, item B29 asks, "Did you do anything else to look for work in the last four weeks that I didn't mention?" For such questions, respondents are asked to specify an answer to the question "anything else?" or "anyone else?"
3. **Field-coded responses** are answers coded by interviewers into a predefined response category without reading the categories aloud to the respondent. If none of the response options seems to apply, interviewers select an "other/specify" category and type in the response. For example, item G53 asks "Thinking only about the services {you/NAME} used in 2018, what are the main reasons {you/he/she} decided to use these services?" Interviewers then coded the verbatim response into seven established categories. If the response did not fit into one of the categories, interviewers selected "other."

As part of data processing and based on an initial review of data, we examined verbatim responses in an attempt to uncover dominant themes for each question. To ensure high quality coding, we employed the same coding procedures in Round 7 that we used in prior rounds of data collection. For example, in Round 6, we added supplemental response categories to some field-coded and other/specify response options if a sufficient number of similar responses could not be back-coded into pre-existing categories. In general, we added a new response category to the Round 6 data file if it was provided by 10 percent or more of the respondents who offered a verbatim response to the question. In an effort to minimize back-coding during the Round 7 data cleaning effort, we added many of these response categories to the Round 7 instrument. We re-used the supplemental response categories that we identified during Round 6 coding, but did not add to the Round 7 instrument, during Round 7 coding. After reviewing verbatim responses to the Round 7 open-ended items, we determined that we did not need to add any other response categories to the data file.

In Chapter IV, we indicate which items in each instrument section required coding and list all additional response categories created during coding. Thus, we categorized verbatim responses for quantitative analyses by coding responses that clustered together (for open-ended and "other/specify" responses) or by back-coding responses into existing response options if appropriate (for field-coded and "other/specify" items).

If the need for changes to the coding scheme became apparent during coding (for example clarification of coding decisions), we discussed and documented new decision rules. Coders used the Ascribe coding software to apply codes to verbatim responses. The Ascribe program allowed coders to sort and filter verbatim responses in several ways to facilitate the coding effort. We sorted verbatim responses alphabetically by item for coders. Records could also be filtered to show responses that had been reviewed by a supervisor, or to show cases with comments for a coder. When it was impossible to code a response, when a response was invalid, or when a response could not be coded into a given category, we assigned a two-digit supplemental code to the response (Table III.1). The data files do not include the verbatim responses, although we note in Chapter IV which items in each instrument section required coding and list any response categories that we added during coding.

Table III.1. Supplemental Codes for "Other/Specify" Coding

Code	Label	Description
94	Invalid Response	Indicates that response should not be counted as an "other" response and should be deleted.
95	Refused	Used only if verbatim response indicates that respondent refused to answer question.
96	Duplicate Response	Indicates that verbatim response has already been selected in a "code all that apply" item.
98	Don't Know	Used only if verbatim response indicates that respondent does not know answer.
99	Not Codeable	Indicates that a code may not be assigned based on verbatim response.

Source: NBS Round 7.

B. Data Cleaning

Once we incorporated coded data into a preliminary data file, we conducted a systematic review of the frequency counts of the individual questionnaire items. We reviewed the counts by each questionnaire path to identify possible errors in skip patterns.

The CATI software sets data entry ranges to limit the entry of improbable responses. However, the ranges intentionally encompass a wide spectrum of values to permit the interview to continue in most instances. Several consistency checks embedded throughout the NBS instrument also flag potential problems during the interview. To minimize respondent burden, however, interviewers are able to suppress all consistency checks encountered during the interview. Interviewers are trained to probe and reconcile such inconsistencies before moving to the next question, but this is not always possible as it may be too time consuming to resolve the inconsistencies with the respondent. We reviewed the data for these potential inconsistencies, such as duplicate conditions listed under the main limiting condition at B2, jobs ending prior to the start date, take home job pay that is greater than gross pay, or very high or very low hours per week on the job. Many of these same consistency and range checks were used in prior rounds of NBS data processing. We reviewed the output of flagged cases and set data to "don't know" (.D) if an error likely occurred. As in earlier rounds, we took the general approach of editing only those cases with an obvious data entry or respondent error. As a result, although substantial time went into meticulously reviewing individual responses, some suspect values remain in the file.

During data processing, we created several constructed variables to combine data across items. For these items, we reviewed the specifications and all data values for the constructed variables based on composite variable responses and frequencies. Many of these constructed variables are the same as in prior rounds of the NBS.

For open-ended items assigned numeric codes, we examined frequencies to ensure the assignment of valid values.

C. Identification of Data Problems

The data problems we identified during the course of checking the data file may be characterized as either measurement error or processing error. Measurement error is the difference between the observed value of a variable and the true, but unobserved, value of that variable. Sources of measurement error may include the questionnaire itself (including, design, format, and content); data collection mode; the interviewer; and the respondent. As discussed below, the questionnaire, interviewer, and respondent likely all contributed to data problems identified in the NBS. Processing errors discussed in this report consist of incorrect specification or implementation of a complicated skip pattern or edit. We discuss programming errors that resulted in incorrect skip patterns throughout the NBS–General Waves and focus on measurement errors and processing errors at the individual-item level.

The identification of data problems in the NBS–General Waves file occurred at several points during the data cleaning and data preparation processes. Specifically, errors may have surfaced during the systematic review of frequency counts of individual questionnaire items or during the identification of cases flagged by the inconsistency and range checks. Other problems surfaced during development of the constructed variables and implementation of the imputation procedures. In Chapter IV, we describe the results of the review by instrument section.

Table III.2 summarizes the problems we encountered when we cleaned the data. These problems are described in further detail in Chapter IV.

Table III.2. Main Problems Encountered

Item	Question Universe Count	Cases Affected (n)	Cases Affected (% of question universe)	Description of problem and resolution
A68, A68a, and A68b (self-reported date of birth)	8,824	2	0.0%	A programming issue caused two cases to skip A68, A68a, and A68b. We restored the data for these cases by setting the date of birth as equal to the date of birth from SSA records.
B29_11a (contacted someone to find out how benefits affected if went to work)	635	4	0.6%	A programming issue caused four respondents to skip B29_11a even though these respondents met the entry criteria for the question. We set B29_11a to .M for these cases.
B29_12a (lowest wage/salary willing to accept job)	2,191	5	0.2%	Five respondents should have skipped out of B29_12a because they did not meet the entry criteria. We left all responses as collected for this question, even though the respondents should not have been asked this question.
C_B18 (satisfaction with main job)	765	1	0.1%	A programming error caused one respondent to skip C_B18. We set C_B18 to .M for this case.
C_B skip logic				Section C_B is asked of respondents who are not currently working, but who worked in the last 6 months (B24=0 and B24b=1). For these three cases, the response to question B24b was changed after Section B was complete. There was an issue related to the skip logic in the instrument as a result of the interviewer moving backward and forward in the instrument to revise the answers to earlier questions. These respondents met the entry criteria for the respective questions. We set these questions to .M as applicable for the three cases.
	814	3	0.4%	Due to this issue, three respondents skipped C_B39a2, C_BP13c, C_B39_1, C_B39_3_a through C_B39_3_h, C_B39_4, and C_B39_5.
	761	3	0.4%	These three respondents were also skipped out of C_B39_a through C_B39_h.
	814	2	0.2%	Due to the same programming error, two of these cases also skipped C_BP13a and C_BP13b.

Item	Question Universe Count	Cases Affected (n)	Cases Affected (% of question universe)	Description of problem and resolution
	774	1	0.1%	One of these cases also skipped C_BP2_1, C_BP3_a through C_BP3_j, C_BP3l, C_BP7, C_BP10, C_B16, C_B20_a through C_B20_i, and C_B33_a through C_B33_f, due to the same programming error.
	294	1	0.3%	The same case also skipped C_BP5_1
	814	1	0.1%	The same case also skipped C_BP4 and C_B34
CP4 (anyone helped find or keep work), C18 (satisfaction with current/main job)	4,354 3,782	3 3	0.1% 0.1%	Three respondents who should have been asked CP4 and C18 were skipped because of a programming error. We fixed the programming error after conducting a data review during the early stages of data collection. We set CP4 and C18 to .M for these cases.
CP7 (anyone at main/current job know about disability), CP8 (how comfortable discussing disability with others at main/current job), CP10 (other people with disabilities work at job), C16 (received promotion in last 12 months)	4,149	15	0.4%	Items CP7, CP8, CP10, and C16 are not meant to be asked of respondents who are self-employed. These items were asked of 15 self-employed respondents before we corrected the programming error during the early stages of data collection. Subsequent self-employed respondents correctly skipped these items. We left all responses as collected for these questions for the 15 self-employed respondents.
G53 (reasons decided to use the employment, job training, medical, or therapy services)	7,442	2	0.0%	Two respondents who should have been asked G53 were skipped because of a programming error. We resolved this issue after conducting a data review in the early stages of data collection. We set G53 to .M for these cases.

Source: NBS Round 7.

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

IV. SECTION-BY-SECTION DESCRIPTION OF FINDINGS

A. Section A—Introduction and Screener

We designed the NBS screener to identify and gain the cooperation of the respondent in addition to verifying that the sample person was still eligible for the survey. We also determined if the sample member was capable of completing the interview and if the sample member required special accommodations, such as TRS or an in-person interview.

1. Date of Birth

We used sample member name and date of birth from SSA records to verify that the correct person had been contacted. If two of the three date of birth elements provided by SSA matched self-reported information (for example, month and year), we continued the interview. If none or only one of the elements matched, we terminated the interview and the case was sent to locating. If the respondent could not provide date of birth, we requested the age of the sample member. If the age fell within two years (plus or minus) of the age in the SSA records, we continued the interview.

Of the successfully screened respondents, date of birth provided by SSA differed from the self-reported date of birth by one field (day, month, or year) in 320 cases. For 143 cases, the discrepancy was in the day of birth, but the month and year matched what was provided by SSA. For 72 cases, the month of birth was different, but the day and year matched. We did not edit these discrepancies and they remain in the file. The remaining 105 cases had a discrepancy in the year of birth, compared to SSA records. Of the 105 cases in which year of birth did not match:

- 22 cases differed by one year, and 22 cases differed by two to nine years. We did not edit these discrepancies and they remain in the file.
- 61 cases differed by 10 or more years. For these cases, we set the year of birth as equal to the year of birth from SSA records because the discrepancy appeared to be the result of a data entry error.

In addition, for cases in which age was provided in lieu of date of birth (11 cases), we used date of birth from SSA records to populate the self-reported date of birth (items A68, A68a, and A68b). As reported above, self-reported date of birth information is missing for two cases due to an instrument programming error. However, we can tell from internal audit trails that respondents provided answers to these questions and passed the identity verification. For these two cases, we used date of birth from SSA records to populate the items on self-reported date of birth (items A68, A68a, and A68b). For one case, the discrepancy was in the month and day of birth, but the year of birth matched. We did not edit this discrepancy and it remains in the file.

2. Discrepancies in Respondent Type

We used three screener items to determine if the sample member was cognitively able to participate in the survey. The items addressed important elements of informed consent—the study topics, the voluntary nature of participation, and confidentiality. If the sample member did not pass any of the three items (within two attempts), the interviewer sought a proxy respondent. For the proxy to complete the survey on the sample member's behalf, the proxy also had to pass the cognitive screener. In addition, interviewers could complete the interview with a proxy if a knowledgeable informant indicated that the sample person would not be able to participate even with an accommodation or if it became clear during the course of

the interview that the sample person was not capable of responding. Whenever possible, we preferred the beneficiary to participate instead of a proxy because sample members generally provide more complete and more accurate information than proxy respondents.

At the end of Section A, before the interview begins, the interviewer had to indicate whether the survey respondent was a sample member or a proxy. This determines which questions the respondent receives during the interview, as some questions are asked only of sample members and not asked of proxies. At the end of the interview (in Section M), the interviewer records whether the sample member or proxy completed the majority of the survey. For 111 cases, there was a switch in respondents after the cognitive screener. We were not surprised by a switch in respondents and expected that some sample members would pass the cognitive screener but then be unable to recall or report information for the vast number of survey questions. We considered the data collected by the interviewer at the end of the survey as the most accurate data for purposes of creating the constructed variable C_Rtype (respondent type). That is, if the sample member began the interview but the interviewer indicated that the proxy completed most of it, we recoded respondent type to proxy.

In terms of the survey questions, we asked perception and attitudinal questions only of sample members. These questions focused on overall job satisfaction and satisfaction with various work characteristics, awareness of SSA programs, and satisfaction with state vocational rehabilitation services. We skipped specific questions (see Table II.3) for proxy respondents. Based on our review of the data, we can confirm that the proxies did not skip any items that they should have been asked, nor receive any questions that they should not have been asked.

B. Section B—Disability and Current Work Status

In Section B, we asked about the sample member's limiting physical or mental conditions and employment status. We also included questions designed to determine what job characteristics were important to sample members and collected information about work-related goals and expectations.

1. Current Work Status

At item B24, we asked respondents if they are currently working. Respondents who reported they were not currently working (B24=0) were asked if they worked in the past six months (B24b). In order to be eligible for the Round 7 survey, all cross-sectional SWS respondents had to respond affirmatively to either B24 (currently working), or B24b (worked in the past six months).

All Round 7 longitudinal SWS members reported currently working at the time of the Round 6 interview (B24=1). These sample members were not required to be working at the time of the Round 7 interview, nor to have worked in the last six months, in order to be eligible for the Round 7 interview.

For longitudinal sample members who were not working at the time of the Round 7 interview, did not work in the past six months, and did not work in 2018, we added 11 questions about the reasons they left their last job to Section B.

2. Health Condition Coding

In Section B of the questionnaire, we asked each respondent to cite the primary and secondary physical or mental conditions that limit the type or amount of work or daily activities that he or she performs. Respondents could report main conditions in one of four questions: B2 (primary reason limited), B6 (primary reason eligible for benefits), B12 (primary reason formerly eligible for benefits if not currently eligible), and B15 (primary reason limited when first receiving disability benefits). The majority of

respondents (81 percent) reported a primary limiting condition in item B2. The main purpose of the other questions (B6, B12, and B15) was to collect information on a health condition from people who reported no limiting conditions in item B2. For example, if respondents reported no limiting conditions, they were asked if they were currently receiving social security benefits (item B5). If they answered "yes," they were asked about the main reason that made them eligible for benefits (item B6). If respondents said that they were not currently receiving benefits, we asked whether they had received disability benefits in the last five years (item B9). If they answered "yes," we asked for the condition that made them eligible for social security benefits (item B12) or the reason that first made them eligible if they no longer had that condition (item B15). If respondents said that they had not received disability benefits in the last five years, we screened them out of the survey and coded them as ineligible. We assigned a value for the three health condition constructs to each response for items B2, B6, B12, and B15. Although respondents were asked to cite one "main" condition in question B2, B6, B12, or B15, many listed more than one. We maintained the additional responses under the primary condition variable and coded in the order in which they were recorded.

For each item on a main condition, we asked respondents to list any other, or secondary, conditions. For example, in item B4, we asked respondents who had reported a main condition in item B2 to list other conditions that limited the type or amount of work or daily activities that they could perform. In item B8, we asked respondents who had reported the main reason for their eligibility for disability benefits in item B6 to list other conditions that made them eligible. For respondents who reported that they were not currently receiving benefits but who reported a main condition in item B12 (the condition that made them eligible to receive disability benefits in the last five years), we asked in item B14 for other reasons that made them eligible for benefits. Those who reported that their current main condition was not the condition that made them eligible for benefits and who were asked for the main reason for their initial limitation were also asked if any other conditions limited them when they started receiving benefits (item B17).

During Rounds 1 through 5 of data collection, we coded respondents' verbatim responses by using the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9) five-digit coding scheme. The ICD-9 is a classification of morbidity and mortality information developed in 1950 to index hospital records by disease for data storage and retrieval. A newer version of the coding scheme (ICD-10) was released prior to Round 6 of data collection. Rather than switching to the ICD-10, which included a new layout of the codes and more complex mapping, SSA agreed that we should use a broader, three-digit coding scheme derived from the ICD-9 categories for Rounds 6 and 7. The list of 21 codes used for Rounds 6 and 7 is included in Table IV.1. The coders, many of whom had medical coding experience, attended a four-hour training session before they started coding; they also attended weekly check-in meetings with coding supervisors throughout the coding effort.

For cases in which the respondent reported several distinct conditions, all conditions were coded (for example, three distinct conditions would be recorded and coded as B2_1, B2_2, and B2_3). Each code was applied a maximum of one time per question, even in instances in which the same medical code could be applied to more than one condition reported within a question. For instance, "bipolar" and "schizophrenia" are distinct conditions that fall under the same medical code (050 – mental disorders). If both conditions were reported within the same response, "bipolar" and "schizophrenia" would receive code 050 one time. If each condition was reported in a separate question (for instance, if the respondent reported "bipolar" at item B2 and "schizophrenia" at item B4), both conditions were coded.

Following the health condition coding, we created a series of three constructed variables based on item B2 in order to collapse the codes into three classes of broad disease groups:

1. Main Condition Primary Diagnosis Groups (C_MainConDiagGrpNew), 16 levels
2. Main Condition Primary Diagnosis Groups Collapsed (C_MainConColDiagGrp), 5 levels
3. Main Condition Body Groups (C_MainConBodyGroup), 18 levels

Each of these constructed variables are created for every condition listed at B2 (C_MainConDiagGrpNew_1, C_MainConDiagGrpNew_2, etc.).

Table IV.1. Round 7 Coding Scheme

Code	Label	Description of ICD-9 Codes	Corresponding ICD-9 Codes
010	Infectious and parasitic diseases	Borne by a bacterium or parasite and viruses that can be passed from one human to another or from an animal/insect to a human, including tuberculosis, HIV, other viral diseases, and venereal diseases (excluding other and unspecified infectious and parasitic diseases)	001.0–135, 137.0–139.8
020	Neoplasms	New abnormal growth of tissue (i.e., tumors and cancer), including malignant neoplasms, carcinoma in situ, and neoplasm of uncertain behavior	140.0–239.9
030	Endocrine/nutritional disorders	Thyroid disorders, diabetes, abnormal growth disorders, nutritional disorders, and other metabolic and immune disorders	240.0–279.9
040	Blood/blood-forming diseases	Diseases of blood cells and spleen	280.0–289.9
050	Mental disorders	Psychoses, neurotic and personality disorders, and other non-psychotic mental disorders. EXCLUDES Intellectual disability (formerly termed mental retardation)	290.0–302.9, 305.00–314.9, 315–316
051	Intellectual disability	Intellectual disability	317.0-319.9
060	Diseases of nervous system	Disorders of brain, spinal cord, central nervous system, peripheral nervous system, and senses, including paralytic syndromes	320.0–359.9
061	Diseases and disorders of the eye and ear	Disorders of eye and ear	360.0–389.9
070	Diseases of circulatory system	Heart disease; disorders of circulation; and diseases of arteries, veins, and capillaries	390-459.9
080	Diseases of respiratory system	Disorders of the nasal, sinus, upper respiratory tract, and lungs, including chronic obstructive pulmonary disease	460-519.9
090	Diseases of digestive system	Diseases of the oral cavity, stomach, esophagus, and duodenum	520.0-579.9
100	Diseases of genitourinary system	Diseases of the kidneys, urinary system, genital organs, and breasts	580.0-629.9
110	Complications of pregnancy, child birth, and puerperium	Complications related to pregnancy or delivery and complications of puerperium	630-677
120	Diseases of skin/ subcutaneous tissue	Infections of the skin, inflammatory conditions, and other skin diseases	680.0-709.9
130	Diseases of musculoskeletal system	Muscle, bone, and joint problems, including arthropathies, rheumatism, osteopathies, and acquired musculoskeletal deformities	710-719, 725-739

Code	Label	Description of ICD-9 Codes	Corresponding ICD-9 Codes
131	Diseases of the musculoskeletal system: back disorders.	intervertebral disc disorders, other disorders of cervical region, and other and unspecified disorders of the back	720-724
140	Congenital anomalies	Problems arising from abnormal fetal development, including birth defects and genetic abnormalities	740.0-759.9
150	Conditions in the perinatal period	Conditions that have origins in birth period, even if disorder emerges later	760.0-779.9
160	Symptoms, signs, and ill-defined conditions	Ill-defined conditions and symptoms; used when no more specific diagnosis can be made	780.01-799.9
170	Injury and poisoning	Problems that result from accidents and injuries, including fractures, brain injury, and burns (excluding complications of medical care not elsewhere classified)	800.00–998.9
180	Physical problem, not elsewhere classified	The condition is physical, but no more specific code can be assigned	No ICD-9 codes
95	Refused	Verbatim indicates that respondent refused to answer the question	No ICD-9 codes
96	Duplicate condition reported	The condition has already been coded for the respondent	No ICD-9 codes
97	No condition reported	The verbatim does not contain condition or symptom to code	No ICD-9 codes
98	Don't know	The respondent reports that he or she does not know the condition	No ICD-9 codes
99	Uncodeable	A code cannot be assigned based on the verbatim response	No ICD-9 codes

Source: NBS Rounds 6 and 7.

We created a set of separate constructs that use the same three methods to collapse responses provided in items B4 (other limiting conditions) and B6 (primary reason eligible for benefits) for those currently receiving benefits. The B4 constructs include the prefix "C_SecCon," as B4 is the secondary condition reported (C_SecConDiagGrpNew, C_SecConColDiagGrp, and C_SecConBodyGroup). The B6 constructs include the prefix with "C_MainReasElig" for the main reason the respondent became eligible for disability benefits (C_MainReasEligDiagGrpNew, C_MainReasEligDiagGrp, and C_MainReasEligBodyGroup).

Lastly, we created another set of three constructs to summarize responses provided across B6, B12, and B15 collectively to determine the main reason for becoming eligible for disability benefits, regardless of current status. These variables are prefixed with "C_ReasBecElig" for reason became eligible. They clarify the eligibility of sample members who indicated in item B2 that they did not have a disabling condition.

a. Several Primary Conditions

Health condition coding of respondent-provided data is complex. Often, respondents either do not know the name of a condition or describe it in vague terms (for example, "he is slow," or "she has trouble breathing"). As previously mentioned, although respondents were asked to provide one "main" condition in item B2, B6, B12, or B15, many listed more than one. Despite the emphasis in interviewer training on collecting one main condition, 16 percent of respondents reported more than one condition in item B2.

Rather than attempting to discern which condition was the main condition among more than one condition listed, we coded conditions in the order provided by the respondent and named on the file as _1, _2, and so on.

b. Duplicate Conditions

In 121 cases, respondents reported a secondary condition already reported as a main condition. During the process of coding such responses, coders identified any duplicate conditions by assigning the code 96. As mentioned previously, each code was applied only one time per question, even in instances in which the same medical code could be applied to more than one condition reported within a question. The same medical code was applied to both primary and secondary conditions when the secondary condition differed from the primary condition, but fell under the same medical category as the primary condition (for example, if a sample member reported "asthma" at item B2, and "Chronic Obstructive Pulmonary Disease" at item B4, both items would be coded as 080 – diseases of respiratory system). If the sample member reported the same condition at item B2 and item B4, the secondary condition was assigned the code 96 – duplicate condition. In the event that the only condition reported as a secondary condition in item B4, B8, B14, or B17 was a duplicate of the main condition reported in B2, B6, B12, or B15, the secondary code received a code of 96 indicating that it was a duplicate condition.

c. Uncodeable Conditions

We expected that not all verbatim responses would contain enough information to allow coders to assign a specific medical code. We could not code approximately 8 percent of the medical condition verbatim responses. In such cases, we coded responses as "don't know" (if the respondent indicated he or she did not know the answer; code 98), "refused" (if the respondent indicated he or she refused to answer; code 95), "uncodeable" (a response was provided but not enough information was given to code; code 99), or "no condition reported" (no condition is reported in the verbatim; code 97).

3. Back-Coding Responses to "Other/Specify" Items

At item B25 we asked respondents if any one of a series of items (item B25_a through item B25_p) was a reason that they were not currently working. In addition, at item B26 we asked respondents if there were any other reasons not already mentioned that explain why they were not working. If they answered "yes," we collected a verbatim response at B27. Before coding, verbatim responses to item B27 were reviewed to determine whether they could be back-coded into items B25_a through B25_p. Table IV.2 shows the response categories that we added during Round 6 coding and reused during Round 7. We back-coded responses whenever possible into one of the existing categories, and retained responses that could not be coded as "other." If all responses could be coded, we recoded item B26 to "no" to indicate that there were no additional reasons that explain why the sample member was not working. If a verbatim response could not be coded into any of the B25 categories, we kept the code for item B26 as "yes."

Table IV.2. Additional Response Categories Used in Section B During Coding

Item	Question Text	Categories Used
B25	Are you not working because…	q = Can't find a job r = Lack skills
B29_6	What benefits are you most worried about losing?	14 = Health Insurance Unspecified
B29_10	What benefits are you most worried about losing?	14 = Health Insurance Unspecified
B29_11b	What benefits were you most worried about losing?	14 = Health Insurance Unspecified

Source: NBS Round 7.

At items B29_6, B29_10, and B29_11b, we asked respondents which benefits they were most worried about losing if they took a job. Responses coded as "other" by the interviewer were reviewed by coders and back-coded into existing response options when possible. We added the category "Health Insurance Unspecified" during coding in a previous round to capture responses that could not be classified into the more specific insurance categories.

We also included "other/specify" responses at items B29 (what did you do to look for work), B29_2 (reasons did not accept a job), B29_7 (reasons have not found a job), B39 (main person you discuss work goals with), BP1b (reasons you are no longer able to do the same kind of work), BP3 (reason health prevents you from working), B36c_1-B36e_1 (health-, job-, or personal circumstance-related problems that might prevent you from working), BP4a1-BP4a3 (health-, job-, or personal circumstance-related problems that made you leave this job), and BP4b1-BP4b3 (health-, job-, or personal circumstance-related problems that might prevent you from working enough to leave benefits), which were examined by coders and back-coded when possible.

C. Section C—Current Employment

In Section C, we collected information about the respondent's current job or jobs by asking respondents about the type of work performed, type of business, hours worked, benefits offered, and wages earned. We also asked about the receipt of work-related accommodations and those needed but not received. We also gathered information on job satisfaction.

We collected job-specific information (items C2 through C13) separately for each current job held. We represent these items in the data file with an _n indicating the job to which the data refer (for example, C4mth_1 indicating month started first job, C4mth_2 indicating month started second job, and so on). Respondents reported first on their main job (that is, the job at which they worked the most hours) and then on other jobs currently held. For purposes of the constructed variables based on data collected in Section C, we named constructs pertaining to the "main" job based on responses provided in the first job slots (_1).

1. Occupation and Industry Coding

In item C2, we asked respondents to describe the type of work they performed at each of their current jobs (occupation). For Rounds 6 and Round 7, we used the 2010 Standard Occupational Classification (SOC) scheme for coding.[23] The SOC classifies all occupations in the economy, including private, public, and military occupations in which work is performed for pay or profit. Occupations are classified on the basis of work performed, skills, education, training, and credentials. The sample member's occupation was assigned an occupation code. The first two digits of the SOC codes classify the occupation to a major group and the third digit to a minor group. For the NBS–General Waves, we assigned three-digit SOC codes to describe the major group that the occupation belonged to and the minor groups within that classification (using the 23 major groups and 96 minor groups). Rounds 6 and 7 codes applied using the 2010 SOC are comparable to earlier rounds that used the 2000 SOC, as all major and minor group codes remained consistent across both coding schemes.

[23] See Standard Occupational Classification Manual (2010) available at http://www.bls.gov/soc. In Rounds 5 and earlier, we used the Bureau of Labor Statistics 2000 Standard Occupational Classification to code verbatim responses to occupation items.

At item C3, we collected information about the type of business employing the sample person (industry). In previous rounds of data collection, we coded verbatim responses to the industry items according to the 2002 North American Industry Classification System (NAICS). For Rounds 6 and 7, we used the 2017 North American Industry Classification System (NAICS).[24] The NAICS is an industry classification system that groups establishments into industrial categories based on the activities in which those establishments are primarily engaged. It uses a hierarchical coding system to classify all economic activity into 20 industry sectors. For the NBS—General Waves, we coded NAICS industries to three digits, with the first two numbers specifying the industry sector and the third specifying the subsector. Rounds 6 and 7 codes applied using the 2017 NAICS are comparable to earlier rounds that used the 2002 NAICS, as all industry sector and subsector codes remained consistent across both coding schemes. Most federal surveys use both the SOC and NAICS coding schemes, thus providing uniformity and comparability across data sources. Although both classification systems allow coding to high levels of specificity, SSA and the analysts decided based on research needs to limit coding to three digits. More information on coding responses to the occupation and industry items is available in "The National Beneficiary Survey–General Waves: Round 7 Editing, Coding, Imputing, and Weighting Procedures" report (Grau et al. 2021).

The verbatim responses to items C2 and C3 do not appear in either the restricted- or public- use version of the file. Rather, we provided the coded responses to item C2 in the constructed variables C_MainCurJobSOC, C_CurJob2SOC, and so on; the coded responses to item C3 were provided in C_MainCurJobNAICS, C_CurJob2NAICS, and so on.

a. Uncodeable Occupation and Industry Verbatim Responses

We expected that some verbatim responses would lack sufficient detail to permit coding at the three-digit level. We provided coders with supplemental two-digit codes to allow assignment of a general code in such cases (Table IV.3).

Table IV.3. Two-digit Supplemental Codes for Occupation and Industry Coding

Code	Label	Description
94	Sheltered workshop	Code used if occupation is in sheltered workshop and a specific occupation cannot be coded from the verbatim response. All industry responses for sheltered workshop are coded as 624.
95	Refused	The respondent refused to give his or her occupation or type of business.
98	Don't know	The respondent reports that he/she does not know the occupation or industry.
99	Uncodeable	A code cannot be assigned based on the verbatim response.

Source: NBS Round 7.

If a respondent did not provide a codeable occupation but indicated either in the verbatim response or in item C7 (job part of sheltered workshop) that the occupation was a sheltered workshop position, we assigned code 94 only if the position could not be assigned an SOC code. If a position in a sheltered workshop was described by an accompanying codeable occupation, we coded the occupation with the SOC classification. When respondents indicated in item C7 that their current job was a sheltered workshop position, we coded the industry as 624 (social assistance), which encompasses service for people with disabilities. If the occupation was uncodeable with no indication that the position was a sheltered workshop position, we assigned code 99 (uncodeable) to the occupation. In all, approximately 3

[24] For more information, see North American Industry Classification System, 2017, or https://www.census.gov/eos/www/naics/index.html

percent of both the current occupation verbatim responses (item C2) and the industry verbatim responses (item C3) for each job were uncodeable.

2. Hours Worked

In item C8, we asked respondents to provide the number of hours per week that they usually worked at their current job. We incorporated a soft edit check into the Blaise instrument to prompt interviewers to verify that the response was correct for any response over 60 hours per week (244 cases for job 1). We reviewed all responses under 5 hours per week (120 cases for job 1) during data cleaning. After a review of other job-related information, including occupation and industry verbatim responses, wage rates, self-employment, and sheltered workshop indicators, we did not recode any cases. In general, if the respondent was working in a sheltered employment setting, we determined that low values for hours worked were not unreasonable and should be retained. Similarly, if the respondent's occupation was consistent with a high number of hours worked per week (for example, truck driver), we retained the values. Although some values were suspect, our general approach was to recode only those cases that appeared to be obvious data entry or respondent errors.

3. Weeks per Year

Item C9 asked respondents how many weeks per year they usually worked at their current job. We reviewed responses of fewer than 20 weeks during data cleaning (134 cases for job 1), along with other job-related information, in order to determine if the values were reasonable. In general, if the occupation verbatim and other job-related information was consistent with the possibility of minimal weeks worked per year, we retained the original values. In some cases, the respondents apparently interpreted the question as asking how many weeks they had worked if they had just started their job—despite the inclusion of the probe, "If you have worked less than a year, please answer for the number of weeks you expect to work." Because it was not possible based on other information to determine whether such values were errors, we retained them in the data file.

4. Pay

At item C11 or C12amt, we asked respondents to report their pretax earnings for each current job. In item C13amt, we asked for respondents' take-home pay. We created three constructed variables: one designed to combine pretax responses into an hourly wage (C_MainCurJobHrPay, C_CurJob2HrPay, and so on); one into a monthly wage (C_MainCurJobMnthPay, C_CurJob2MnthPay, and so on) regardless of where the initial reporting occurred; and one for monthly take-home pay (C_MainCurJobMnthPayTH, C_CurJob2MnthPayTH, and so on). In addition, we constructed a total monthly pay variable to sum across all jobs (C_TotCurMnthPay). Given that the earning constructs are subject to imputation, we were concerned that outliers might become imputation donors and exacerbate the outlier problem. Thus, we performed a detailed review of high and low values for both the source variables and constructs. The donor pool for imputation excluded cases with very high and very low values.

We included a soft edit check included in the Blaise instrument to prompt interviewers to verify any response that was more than $25 per hour in item C11. The check could be suppressed, however, leading to 146 cases reporting hourly rates over $25. Because other job-related information, including the verbatim occupation response, indicated that the entries could be valid, we retained all in the file. We also examined all hourly wage values of $3 and below. In such cases (11 cases), we retained the value because the verbatim job descriptions indicated that the low value for hourly wages was not unreasonable.

We built soft edit checks into the instrument to flag high entries for each of the various reporting units in items C12amt and C13amt. We examined values that were suppressed or that were at the high and low ends of the range. Because other job-related information, including the verbatim occupation and industry descriptions indicated that the values could be valid, we retained all high-value entries in the file. If the respondents were working in a sheltered employment setting or the verbatim job descriptions indicated that the low values for wages were not unreasonable, we retained the values.

During post-processing, we also compared take-home and pretax values. There were seven cases that reported a monthly take-home pay that exceeded the monthly gross pay. For five of these cases, we set C13amt_1 and/or C13hop_1 to "don't know" (D). For one case, we set both C12amt_1/C12hop_1 and C13_amt1/C13hop_1 to "don't know" (D) after reviewing other job-related information. For the remaining case, we adjusted the value of C12hop_1 to be consistent with the value of C13hop_1.

5. Back-Coding Responses to "Other/Specify" Items

In items C33_a through C33_e, we asked whether the sample member's employer made a series of accommodations. If the respondent indicated that other accommodations were made (C33_f = 1), we collected a verbatim response. We reviewed and back-coded the responses into questions C33_a through C33_e when possible.

We also included the "other/specify" option at item CP3 (things you did to find or keep working), item CP5 (who helped you find job or keep working), item CP7 (coworkers who know about disability), item C39b (reasons work fewer hours or earn less money), and item C39_3 (supports needed to work or earn more).We did not add any new categories during Round 7 data cleaning. Table IV.4 shows the response categories that we added during Round 6 coding and re-used during Round 7.

6. Back-Coding Field-Coded Responses

Items C1a (main reasons decided to work), CP2 (how did you find your job), CP3_k (special equipment or devices used to keep working), CP6 (type of help received), CP12a (supports provided to work with disability), CP13a-CP13c (health-, personal circumstances-, or job-related reasons for working less), CP14 (things that helped you keep working), and C39_2 (benefits reduced or ended as a result of job) were all open-ended items that interviewers attempted to code into one of several predefined response categories during the interview. We reviewed responses coded as "other" by the interviewer and back-coded them into existing response options when possible. We added the category "Health Insurance Unspecified" during a prior round of coding to item C39_2 to capture responses that could not be classified into the more specific insurance categories (Table IV.4). We kept verbatim responses that could not be recoded into one of these categories as "other."

Table IV.4. Additional Response Categories Used in Section C During Coding

Item	Question Text	Categories Used
CP13b1	What was it about [your/NAME's] [main/current] job that might have caused [you/NAME] to have to work less or stop working?	20= Found another job 22= Work schedule 23= Did not like/get along with co-workers 24= Did not like/get along with manager, supervisor, or boss 25= Did not like/get along with other staff responsible for hiring or providing accommodations (such as Human Resources)
CP13c1	What was it about [your/NAME's] personal circumstances that might have caused {you/NAME} to have to work less or stop working?	19= Moved to another area 21= Loss or potential loss of government benefits
C39b	{Do you/Does NAME} work fewer hours or earn less money than {you/he/she} could because {you/he/she}…	g = Poor health/health concerns
C39_2	What benefits have been reduced or ended as a result of your main/current job?	14 = Health insurance unspecified

Source: NBS Round 7.

7. Coding Open-Ended Responses

We asked respondents whether any changes were needed but not made to the sample member's workplace (item C34). If the answer was yes, we collected a verbatim response at item C35 on the specific changes needed. We used a five-level coding scheme developed in Round 6 to categorize the Round 7 responses. We retained responses that could not be coded into one of the five categories as "other."

D. Section C_B—Employment in the Last 6 Months

In Section C_B, we collected information about the respondent's job or jobs held in the last six months by asking them about the type of work performed, type of business, hours worked, benefits offered, and wages earned. We also asked about the receipt of work-related accommodations and those needed but not received. We also gathered information on job satisfaction.

We collected job-specific information (items C_B2 through C_B13) separately for each job held in the last six months. We represent these items in the data file with an _n indicating the job to which the data refer (for example, C_B4mth_1 indicating month started first job, C_B4mth_2 indicating month started second job, and so on). Respondents reported first on their main job (that is, the job at which they worked the most hours) and then on other jobs held in the last six months.

1. Uncertainty about Current Work Status

Section C_B is asked of respondents who report they are not currently working (B24 = 0) and report that they have worked in the last six months (B24b = 1). During data processing, we found three cases that were not sure, or refused to report, if they were currently working, but reported they had worked in the last six months. Given the ambiguity around current work status, these cases did not receive Section C or Section C_B; however, they reported working in 2018 and subsequently received Section D. For these three cases, we coded the first question in Section C_B as missing (".M").

2. Job End Dates

During data processing, we found 66 cases in which a respondent reported working at a job during the last six months in Section C_B, but provided a job end date indicating they had not worked at that job in the last six months. We built soft edit checks into the instrument to flag issues like this during the interview, but these checks could be suppressed by the interviewers. For these 66 cases, we coded the job end date to "don't know" (.D).

3. Occupation and Industry Coding

In item C_B2, we asked respondents to describe the type of work they performed at each of their jobs held in the last six months (occupation). At item C_B3, we collected information about the type of business employing the sample person (industry). As mentioned previously, we used the Bureau of Labor Statistic's 2010 Standard Occupational Classification (SOC) to code verbatim responses to the occupation items, and the 2017 North American Industry Classification System (NAICS) to code industry questions. The verbatim responses to items C_B2 and C_B3 do not appear in either the restricted- or public- use versions of the data file. Rather, we provided the coded responses to item C_B2 in the constructed variables C_Main6MoJobSOC, C_6MoJob2SOC, and so on; the coded responses to item C_B3 were provided in C_Main6MoJobNAICS, C_6MoJob2NAICS, and so on.

a. Uncodeable Occupation and Industry Verbatim Responses

As previously noted, we expected that some verbatim responses would lack sufficient detail to permit coding at the three-digit level. We provided coders with the same supplemental two-digit codes used in Section C to allow assignment of a general code in such cases (Table IV.3).

If a respondent did not provide a codeable occupation but indicated either in the verbatim response or in item C_B7 (job part of sheltered workshop) that the occupation was a sheltered workshop position, we assigned code 94 (sheltered workshop) only if the position could not be assigned an SOC code. If a position in a sheltered workshop was described by an accompanying codeable occupation, we coded the occupation with the SOC classification. When respondents indicated in item C_B7 that their recent job was a sheltered workshop position, we coded the industry as 624 (social assistance), which encompasses services for people with disabilities. If the occupation was uncodeable with no indication that the position was a sheltered workshop position, we assigned code 99 (uncodeable) to the occupation. In all, approximately 4 percent of both the recent occupation verbatim responses and the industry verbatim responses for each job were uncodeable.

4. Hours Worked

In item C_B8, we asked respondents to provide the number of hours per week that they usually worked at their job. We incorporated a soft edit check into the Blaise instrument to prompt interviewers to verify that any response over 60 hours per week (13 cases for job 1) was correct. We reviewed all responses under 5 hours per week (16 cases for job 1) during data cleaning. After a review of other job-related information, including occupation and industry verbatim responses, wage rates, self-employment, and sheltered workshop indicators, we did not recode any cases. In general, if the respondent was working in a sheltered employment setting, we determined that low values for hours worked were not unreasonable and should be retained. Similarly, if the respondent's occupation was consistent with a high number of hours worked per week (for example, truck driver), we retained the values. Although some values were suspect, our general approach was to recode only those cases that appeared to be obvious data entry or respondent errors.

5. Weeks per Year

Item C_B9 asked sample members how many weeks per year they usually worked at their job. We reviewed responses of less than 20 weeks during data cleaning (221 cases for job 1), along with other job-related information, in order to determine if the values were reasonable. As noted previously, if the occupation verbatim response and other job-related information was consistent with the possibility of minimal weeks worked per year, we retained the original values.

6. Pay

At item C_B11 or C_B12amt, we asked respondents to report their pretax earnings for each job worked during the last six months. In item C_B13amt, we asked for respondents' take-home pay. We created three constructed variables: one designed to combine pretax responses into an hourly wage (C_Main6MoJobHrPay, C_6MoJob2HrPay, and so on); one into a monthly wage (C_Main6MoJobMnthPay, C_6MoJob2MnthPay, and so on) regardless of where the initial reporting occurred; and one for monthly take-home pay (C_Main6MoJobMnthPayTH, C_6MoJob2MnthPayTH, and so on). In addition, we constructed a total monthly pay variable to sum across all jobs (C_Tot6MoMnthPay).

We included a soft edit check in the Blaise instrument to prompt interviewers to verify any response that was greater than $25 per hour in item C_B11. The check could be suppressed, however, resulting in 20 cases that reported hourly rates greater than $25 per hour. Because other job-related information, including the verbatim occupation response, indicated that the entries could be valid, we retained all responses in the file. We also reviewed the data for hourly wage values of $3 and below. There were two such cases in Section C_B.

We built soft edit checks into the instrument to flag high entries for each of the various reporting units in items C_B12amt and C_B13amt. We examined values that were suppressed or that were at the high and low ends of the range. Because other job-related information, including the verbatim occupation and industry descriptions indicated that the values could be valid, we retained all responses in the file. If the respondents were working in a sheltered employment setting or the verbatim job descriptions indicated that the low values for wages were not unreasonable, we retained the values.

7. Back-Coding Responses to "Other/Specify" Items

In items C_B33_a through C_B33_e, we asked whether the sample member's employer made a series of accommodations. If the respondent indicated that other accommodations were made (C_B33_f = 1), we collected a verbatim response. We reviewed and back-coded the responses into questions C_B33_a through C_B33_e when possible.

We also included the "other/specify" option at, item C_BP3 (things you did to find or keep working), item C_BP5 (who helped you find job or keep working), item C_BP7 (coworkers who know about disability), item C_B39b (reasons work fewer hours or earn less money), and item C_B39_3 (supports needed to work or earn more). We did not add any new coding categories during Round 7 data cleaning. Table IV.5 shows the categories added during Round 6 coding and re-used in Round 7.

8. Back-Coding Field-Coded Responses

Items C_B1a (main reasons decided to work), C_BP2 (how did you find your job), C_BP3_k (special equipment or devices used to keep working), item C_BP6 (type of help received), item C_BP12a (supports provided to work with disability), C_BP13a-C_BP13c (health-, personal circumstances-, or job-

related reasons for working less), item C_BP14 (things that helped you keep working), and C_B39_2 (benefits reduced or ended as a result of job) were all open-ended items that interviewers attempted to code into one of several predefined response categories during the interview. We reviewed responses coded as "other" by the interviewer and back-coded them into existing response options whenever possible. We added the category "Health Insurance Unspecified" during a prior round of coding to item C39_2 to capture responses that could not be classified into the more specific insurance categories. We incorporated the same additional response into C_B39_2 this round. We kept verbatim responses that could not be recoded into one of these categories as "other."

9. Coding Open-Ended Responses

We asked respondents whether any changes were needed but not made to the sample member's workplace (item C_B34). If the answer was yes, we collected a verbatim response at item C_B35 on the specific changes needed. We reviewed the verbatim responses before coding and used five categories from earlier rounds to summarize them. We retained responses that could not be coded into one of the five categories as "other."

Table IV.5. Additional Response Categories Used in Section C_B During Coding

Item	Question Text	Categories Used
C_BP13b1	What was it about [your/NAME's] [main/current] job that might have caused [you/NAME] to have to work less or stop working?	20= Found another job 22= Work schedule 23= Did not like/get along with co-workers 24= Did not like/get along with manager, supervisor, or boss 25= Did not like/get along with other staff responsible for hiring or providing accommodations (such as Human Resources)
C_BP13c1	What was it about [your/NAME's] personal circumstances that might have caused {you/NAME} to have to work less or stop working?	19= Moved to another area 21= Loss or potential loss of government benefits
C_B39b	{Do you/Does NAME} work fewer hours or earn less money than {you/he/she} could because {you/he/she}…	g = Poor health/health concerns
C_B39_2	What benefits have been reduced or ended as a result of your main/current job?	14 = Health insurance unspecified

Source: NBS Round 7.

E. Section D—Job/Other Jobs During 2018

In Section D, we collected information about employment during the 2018 calendar year, including types of employers, hours worked, wages earned, and reasons for leaving employment, if applicable. We also asked if respondents worked or earned less than they could have (and, if so, why) and collected information about experiences related to any adjustments made in social security benefits due to work.

As in Section C and C_B, we collected job-specific information (items D2 through D23) for each job held in 2018. We represent data for each job in the data file with an _n indicating the job to which the data refer (for example, D6mth_1 indicating month started first job, D6mth_2 indicating month started second job, and so on). Respondents reported first on their main job—that is, the job at which they worked the most hours—and then reported on other jobs held. To reduce respondent burden, respondents did not have to report on any jobs held during 2018 that were mentioned in Section C as current employment or

Section C_B as employment in the last six months. Rather, we copied employment data from Section C and C_B into Section D during data processing for all current jobs also held during 2018. Table IV.6 lists all job-specific items that were filled in with Section C or Section C_B data. Items in Section D that had no equivalent in Section C or Section C_B (C_B4bmth, C_B4byr, D8mnth, D8yr, C_BP13a-C_BP13c, DP1a-DP1c) were coded as .L (indicating logical skip).

Table IV.6. Job Variables in Sections C, C_B, and D

Variable in Section C	Variable in Section C_B	Variable in Section D	Variable Description
C2	C_B2	D4	Occupation
C3	C_B3	D5	Industry
C4mth, C4yr	C_B4amth, C_B4ayr	D6mth, D6yr	Start month and year of job
No equivalent item	C_B4bmth, C_B5byr	D8mth, D8yr	Stop month and year of job
C6	C_B6	D14	Self-employed status
C7	C_B7	D15	Sheltered workshop status
C8	C_B8	D16	Hours usually worked per week
C9	C_B9	D17	Weeks usually worked per year
C10	C_B10	D18	Paid by the hour
C11	C_B11	D19	Hourly pay
C12amt, C12hop	C_B12amt, C_B12hop	D20amt, D20hop	Amount of pretax pay
C13amt, C13hop	C_B13amt, C_B12hop	D21amt, D21hop	Amount of after tax pay
No equivalent item	C_BP13a-C_BP13c	DP1a-DP1c	Reasons for stopping work

Source: NBS Round 7.

1. Including Current Jobs Held in 2018 and Jobs in Last Six Months Held in 2018 in Section D

We included jobs mentioned in Section C and Section C_B as held in 2018 if item C4yr (year started current job) or item C_B4yr was earlier than or equal to 2018. We copied each applicable job from Section C and Section C_B into the first blank job slot in Section D (for example into D6mth_2 if D6mth_1 already contained data and into D6mth_3 if both D6mth_1 and D6mth_2 already contained data). We included the variables C_Job_From_SecC_1 through C_Job_From_SecC_5, and C_Job_From_SecC_B_1 through C_Job_From_SecC_B_5 in the data file to indicate which jobs from Section C and Section C_B (by job number) were copied into specific Section D job slots.

2. Determining Main Job Held in 2018

In addition to copying job data from Section C and C_B into the Section D items, we had to determine which job held in 2018 was the main job. Before including the jobs from Section C and C_B, we stored the main jobs held in 2018 as job 1. Given that it was possible that a job reported in Section C or C_B was the respondent's main job in 2018, we compared hours worked in 2018 on each job with the first job mentioned in Section D after incorporating the jobs from Section C and C_B. We deemed the job with the greatest number of hours per year (number of hours per week multiplied by the number of weeks per year) the main 2018 job.[25] We used the variable C_Main_Job_grid_num, which identifies the job number

[25] If hours per year could not be calculated because of missing data on either number of hours per week or number of weeks per year, we coded the hours as missing. If hours per year were missing for all 2018 Section C or C_B jobs, job 1 in Section D was counted as the main job in 2018. If no jobs were listed in Section D and hours per year

of the main job held in 2018 based on number of hours worked, to create a series of variables ending with _m to represent each job-specific item listed in Table IV.6 for the main job held in 2018 (for example D6mth_m and D6yr_m). We did not delete information related to the main job from the job_1 through job_5 variables. For example, for a case with three jobs listed in Section D (after copying relevant jobs from Section C or C_B) that had the second job deemed the main job, information related to hours worked on the second job is available in both items C8_m and C8_2. Therefore, _m jobs should not be counted as additional jobs. The public-use version of the file provides only the main job variables (_m) for jobs held in 2018.

For purposes of the constructed variables created in Section D, we created separate constructs for each job mentioned (job 1, job 2, and so on) as well as additional constructs for the "main" job (C_MainJob2018SOC, C_MainJob2018NAICS, C_MainJobHrPay2018, C_MainJobMnthPay2018, C_MainJobMnthPayTH2018, and C_MnthsMain2018Job) as identified by the variable C_Main_Job_grid_num. As stated, information in the main job constructs is replicated in one of the other job slots in the restricted-use file and does not represent an additional job.

During data processing, we found 129 cases in which a respondent answered "don't know," "refused," or reported in item B30 that he or she did not work in 2018 (B30 = D, R, or 0), but the respondent also had a reported a job start date in Section C or C_B that indicated that the individual had held a job in 2018. We recoded such cases to B30 = 1 (indicating that the respondent did work in 2018). It is important to note that we did not recode item D3 ("Other than the current jobs you just told me about, how many other jobs did you hold for at least one month in 2018?") to reflect the number of jobs held in 2018 after including jobs from Section C or C_B. To determine the total number of jobs held in 2018, the data user should sum item D3 and C_Totjobcopied, a construct that indicates the number of jobs copied from Section C and C_B to Section D.

3. Occupation and Industry Coding

In item D4, we asked respondents to describe the type of work they performed on each job held in 2018 (occupation). In item D5, we asked respondents to describe the corresponding type of business (industry). As for equivalent items in Section C and C_B, we coded the verbatim responses to these items by using the SOC and NAICS classification systems described in Section IV.C.1. We did not include the verbatim responses to items D4 and D5 in the restricted- or public-use version of the data file. Rather, the coded responses to item D4 are in the construct C_MainJob2018SOC, C_Job12018SOC, and so on. The coded responses to item D5 are in C_MainJob2018NAICS, C_Job12018NAICS, and so on.

4. Uncodeable Occupation and Industry Verbatim Responses

Coders used the same codes described in IV.C.1 during the coding process. In all, we deemed uncodeable approximately 4 percent of the 2018 occupation verbatim responses and industry verbatim responses for each job in Section D for any given item.

---

were missing for all 2018 jobs in Section C, the first job listed in Section C or C_B that was a 2018 job was counted as the main job in 2018. If hours per year were missing for job 1 in Section D, the Section C or C_B job with most hours per year was counted as the main 2018 job. If there was no 2018 job from Section C or C_B or hours per year were missing for all Section C and C_B 2018 jobs, job 1 in Section D was counted as the main 2018 job. If hours per year were missing for all 2018 Section C or C_B jobs and for job 1 in Section D, job 1 in Section D was counted as the main job in 2018.

5. Dates Worked at 2018 Job

In items D6mth, D6yr, D8mth, and D8yr, we collected start and stop dates for each job held in 2018. We built soft edit checks into the Blaise instrument to verify that stop dates were later than start dates and that each job was held for at least one month in 2018. If the interviewer verified that the job ended before 2018 or was held for less than one month in 2018, we skipped items collecting job-specific information in items D14 through D21hop. We retained occupation and industry data as well as start and stop dates for these jobs in the data file because respondents answered other items in Section D (why they stopped working at the job in item D23 and general questions about working in 2018 in items D25 through D30).

6. Hours Worked

In item D16, we asked respondents for the number of hours per week usually worked in their 2018 job. As in Section C and C_B, we incorporated a soft edit check into the Blaise instrument to prompt interviewers to verify that the response was correct for any response greater than 60 hours per week. We examined responses over 60 hours per week (57 cases for job 1, for example) and under 5 hours per week (89 cases on job 1) during data cleaning. After a review of other job-related information, we retained all data. In general, if the respondent was working in a sheltered employment setting, we determined that low values for hours worked were not unreasonable and should be retained. Similarly, if the respondent's occupation was consistent with a high number of hours worked per week, we retained the values.

7. Weeks per Year

In item D17, we asked respondents how many weeks per year they usually worked in their 2018 job. We reviewed responses indicating fewer than 20 weeks during data cleaning (337 cases for job 1). In general, if the occupation verbatim response and other job-related information were consistent with the possibility of few weeks worked per year, we retained the original values. It was not possible, based on other information, to determine whether the values were errors; therefore, we retained them in the file.

8. Pay

Respondents reported their pretax earnings for each 2018 job in item D19 (if reported as an hourly wage) or in item D20amt (if reported in another unit, such as daily, weekly, monthly, or annually) and their take-home pay in item D21amt. We combined pretax responses from three constructed variables into an hourly wage (C_MainJobHrPay2018, C_Job1HrPay2018, and so on), a monthly wage (C_MainJobMnthPay2018, C_Job1MnthPay2018, and so on) regardless of where the initial reporting occurred, and monthly take-home pay (C_MainJobMnthPayTH2018, C_Job1MnthPayTH2018, and so on). In addition, we created a constructed total monthly pay variable to sum pay across all jobs (C_Tot2018Pay). We examined source variables and constructed variables for extremely high and low values. There was 1 case where the monthly take-home pay reported exceeded the monthly gross pay. For this case, we set D21_1 and D21hop_1 to "don't know" (D).

We incorporated a soft edit check into the Blaise instrument to prompt interviewers to verify any response over $25 an hour in item D19. We examined responses over $25 an hour (174 cases for job 1). We retained all of these entries in the file because other job-related information, including the verbatim occupation response, indicated that the entries could be valid. We also examined hourly wage values of $3 and below (13 cases). In all cases, respondents were working in a sheltered employment setting or the verbatim job description indicated that the low values for hourly wages were not unreasonable; thus, we retained the values.

We built soft edit checks into the instrument and flagged high entries for each of the various reporting units in items D20amt and D21amt. As for hourly wages, we examined values that were suppressed or that were at the high and low ends of the range by looking at other job-related information. In most cases, the verbatim occupation and industry descriptions indicated that the values could be valid; thus, we retained the values in the file. Generally, if the respondent was working in a sheltered employment setting or the verbatim job description indicated that the low values for wages were not unreasonable, we retained the values. Recoding of data occurred only in cases of an obvious data entry error or when the respondent's job characteristics were inconsistent with reported earnings or pay.

9. Back-Coding Responses to "Other/Specify" Items

We asked items D25_a through D25_f if respondents indicated several issues were the reasons that they had worked fewer hours than they might have worked. In items D26a through D26h, we asked if several issues were the reasons that the sample member did not work or earn more. We reviewed responses coded as "other" during data processing. For both items D25 and D26, we added categories during a prior round of coding to allow further categorization of responses (Table IV.7).

10. Back-Coding Field-Coded Responses

Where possible, interviewers attempted to code verbatim responses to items DP1a-DP1c (why the sample person quit working at the job held in 2018) and D25_2 (benefits reduced or ended as a result of 2018 job) into a series of predetermined categories. We reviewed responses coded as "other reason" to determine if they could be back-coded into an existing category. We retained responses that could not be coded as "other."

Table IV.7. Additional Response Categories Used in Section D During Coding

Item	Question Text	Categories Used
DP1b_1	What was it about [your/NAME's] job that made [you/him/her] leave it?	20= Found another job 22= Work schedule 23= Seasonal/Temporary job
DP1c_1	What was it about [your/NAME's] personal circumstances that made [you/him/her] leave the job?	19= Moved to another area 21= Loss or potential loss of government benefits
D25	Did you work fewer hours or earn less money than you could have because [you/he/she] you…	g= Had medical problems/complications
D25_2	What benefits were reduced or ended as a result of [your/NAME's] job in 2018?	14= Health insurance unspecified
D26_h	In 2018, do you think [you/NAME] could have worked or earned more if [you/he/she] had:	i= Better health/treatment j= More supportive/helpful employer and/or coworker

Source: NBS Round 7.

F. Section SC – Benefit Suspense

In Section SC, we ask beneficiaries who are currently employed (B24=1), have been employed within the last six months (B24b=1), or were employed in 2018 (B30=1) how their work experiences affected their social security disability benefits.

1. Skip Logic from Other Sections

Items SC1a and SC1b both ask if the respondent ever completely stopped receiving cash disability benefits in the past year because of employment. Item SC1a is asked if the respondent reports that his or her Social Security disability benefits were reduced or ended because of a recent job (C39_2=06 or CB39_2=06 or D25_2=06), while item SC1b is worded slightly differently for those who do not report reduced benefits. After back-coding "Other" responses at C39, CB39, and D25, we flagged several cases that should have received SC1a instead of SC1b. We did not recode these two variables during data cleaning because SC1a and SC1b are essentially the same question. SC1a just acknowledges the earlier reference to reduced benefit receipt.

2. Back-Coding Responses to "Other/Specify" Items

Several questions in Section SC (items SS2a_1, SS2b_1, SS2c_1, SS3a, SB1a_1, SB1b_1, SB1c_1, SB2a_other, SB3a, SB4a, and SB4b) included an "other" option that prompted a verbatim response. During data processing, we reviewed the verbatim responses to determine whether they could be clustered into additional categories. In Table IV.8, we provide the response categories that we added during the coding process during Round 6. We did not add any new categories during Round 7 and coded responses to these same categories. We retained responses that could not be coded as "other."

Table IV.8. Additional Response Categories Used in Section SC During Coding

Item	Question Text	Categories Used
SS2b_1	What was it about [your/NAME's] job that makes [you/NAME] think [you/he/she] might go back on benefits?	20= Found another job 22= Work schedule 23= Did not like/get along with co-workers 24= Did not like/get along with manager, supervisor, or boss 25= Did not like/get along with other staff responsible for hiring or providing accommodations (such as Human Resources)
SS2c_1	What was it about [your/NAME's] personal circumstances that makes [you/NAME] think [you/he/she] might go back on benefits?	19= Moved to another area 21= Loss or potential loss of government benefits
SB1b_1	What was it about [your/NAME's] job that made [you/NAME] have to go back on benefits?	20= Found another job 22= Work schedule 23= Did not like/get along with co-workers 24= Did not like/get along with manager, supervisor, or boss 25= Did not like/get along with other staff responsible for hiring or providing accommodations (such as Human Resources)
SB1c_1	What was it about [your/NAME's] personal circumstances that made [you/NAME] have to go back on benefits?	19= Moved to another area 21= Loss or potential loss of government benefits

Source: NBS Round 7.

G. Section E—Awareness of SSA Work Incentive Programs

In Section E, we assessed whether the beneficiary was aware of or participating in specific SSA work incentive programs and services.

1. Student Earned-Income Exclusion

Item E12 was intended to be asked of all SSI beneficiaries 25 years old or younger, and who received SSI benefits before they were 22 years old. During data review, we found six cases that met these criteria, but the instrument incorrectly routed these cases to skip item E12. To correct for this routing error, we recoded E12 to be missing (.M) instead of logically skipped (.L) for these six cases.

H. Section G—Employment-Related Services and Supports Used in 2018

In Section G, we collected information from respondents about their use of employment-related services in 2018, including types of services received, types of providers used, how long services were received, payments for services, and reasons for and satisfaction with service utilization. We asked questions about the sources of information for services and the nature of any services needed but not received.

1. Back-Coding Responses to "Other/Specify" Items

Each of the questions on type of service received and provider type in Section G (items G2, G7, G11, G13, G16, G18, G20, and G22) included an "other" option that prompted a verbatim response. During data processing in Round 6, we reviewed the verbatim responses to determine whether they could be clustered into additional categories. We did not add any new categories during Round 7 data cleaning. In Table IV.9 we provide the response categories added during coding in Round 6, which we also utilized in Round 7. Responses were then back-coded whenever possible into one of the existing or newly created categories. We retained responses that could not be coded as "other."

2. Back-Coding Field-Coded Responses

Items G28 (type of degree working toward) and G53 (reasons for service use) were open-ended items that required interviewers to attempt to code respondents' verbatim responses into predetermined categories. Coders reviewed responses coded as "other" by the interviewer and back-coded them into existing response options when possible.

3. Coding Open-Ended Items

Item G61 (reasons unable to get needed services) was an open-ended question with no response options. We reviewed responses during Round 6 and developed seven categories based on common responses (Table IV.9). Round 7 coders re-used these codes to code the verbatim responses. We retained the response "other" if it did not fit into one of the categories.

Table IV.9. Additional Response Categories Used in Section G During Coding

Item	Question Text	Categories Used
G13	Where did {you/NAME} go to get this training? Please think about all of the places {you/NAME} went in 2016.	9= On the job training (unspecified)
G18	Where did {you/NAME} go to receive these medical services? Please think about all of the places {you/NAME} went in 2016. Did {you/NAME} go to:	05= A school 06= A nursing home/group home 07= A government agency 08= In home care 09= A medical equipment store 10= A rehabilitation/counseling center 11= Physical therapy center
G22	Where did {you/NAME} receive this mental health therapy or counseling? Please think about all of the places {you/NAME} went in 2016. Did {you/NAME} go to:	06= Residential treatment program/facility 07= Rehab center/counseling center/day program 08= Church or religious institution
G61	Why [were you/was NAME] unable to get these services?	01= Not eligible/request refused 02= Lack information on how to get services/didn't know about services 03= Could not afford/insurance would not cover 04= Did not try to get services 05= Too difficult/too confusing to get services 06= Problems with the service or agency 07= Other

Source: NBS Round 7.

I. Section I—Health and Functional Status

In Section I, we collect information about the respondent's general health status and daily functioning, including the need for special equipment or assistive devices. We also collect information on difficulties with ADLs and IADLs, functional limitations, substance abuse or dependence, and treatment for mental health conditions.

1. Back-Coding Responses to "Other/Specify" Items

Items I20 (equipment used for seeing), I24 (equipment used for hearing), I28 (equipment used for speaking), and I32 (equipment used for walking) were all open-ended items that required interviewers to attempt to code respondents' verbatim responses into predetermined categories. Coders reviewed responses coded as "other" by the interviewer and back-coded them into existing response options when possible.

J. Section J—Health Insurance

In Section J, we collected information about the sources of the beneficiary's health insurance coverage both at the time of the interview and during calendar year 2018.

1. Back-Coding Responses to "Other/Specify" Items

Items J6 (type of private insurance), J9 (type of health coverage), and J11 (type of health coverage in 2018) were all open-ended items that required interviewers to attempt to code respondents' verbatim responses into predetermined categories. Coders reviewed responses coded as "other" by the interviewer and back-coded them into existing response options when possible.

K. Section K—Income and Other Assistance

In Section K, we asked about sources of income, including income received from earnings, social security, workers' compensation, and other government programs and sources.

1. Earnings Last Month

In item K3, we asked respondents how much they earned last month before taxes and deductions. In item K3a, we then asked how much remained after taxes and deductions. We built soft edit checks into the instrument and flagged high and low values for both items, although the checks were set to accept a wide range of responses. According to the distribution of responses, we examined extremely low (less than $50 per month) and high (over $5,000 per month) values for both pretax and take-home pay. In most cases, we were able to evaluate the values in the context of the job-specific information in Section C by reviewing the number of jobs currently held by the sample person, the number of hours worked, the sample person's occupation, and whether the sample person was in a supported employment setting. Of the sample members who reported earnings less than $50 a month in either pretax or take-home pay (n=62), about half worked in a sheltered workshop or were involved in a self-employment activity that could explain low monthly wages (n=30). In most cases that reported $0 income, sample persons were employed in seasonal work or sporadic work such as substitute teaching and photography.

2. Income from Other Sources

We built soft edit checks into the instrument to flag high and low values for income received from each source specified (items K7_a through K7_h). We examined values for cases in which the edit check had been suppressed (over $1,000 per month) and cases at the high and low ends of the distribution. In general, as in prior rounds, although some values exceeded the maximum benefit amounts for 2018, we decided to retain the values on the original items. However, for purposes of creating the imputed variable, we did not use values above the limits when calculating the median from which the imputed values were derived. In addition, we did not use values over $8,000 per month for item K7_g (other regular sources) when calculating the median for the imputed variables. Similarly, we reviewed values associated with questions K7_h (other non-regular sources) and K14 (other government assistance), but we did not edit them because we could not clearly identify them as data entry errors. In general, we did not recode values of "0" for amounts received from other sources.

3. SNAP Dollar Value

In question K12, we asked respondents who had reported receiving SNAP last month to report the dollar value of the SNAP. We instructed respondents to include SNAP benefits received only by the sample person, not by other family members. Despite reports of some high values (72 values of $500 or greater), we retained the values in the file.

4. Irregular SSI Income

Item K7_h (amount of income received from other sources not on a regular basis) inquired about irregular SSI payments as non-regular income. For cases that reported the receipt of income from other sources on an irregular basis, we reviewed verbatim responses at item K6_h regarding the source of the income to determine if any SSA or SSI benefits were included. None of the responses suggested that SSA or SSI benefits were the source. A constructed variable, C_AmtOthNonReg contains the amount captured in this survey question.

5. Back-Coding Responses to "Other/Specify" Items

If respondents indicated receipt of income from other sources on either a regular (item K6_g) or non-regular (item K6_h) basis, they were asked to specify the source. Although we could have created additional categories during coding to cluster responses to the query about income sources, such categories would have necessitated the development of additional amount variables in item K7 in order have appropriate coding of the amount of income received from each source. Cases reporting more than one source would not lend themselves to disaggregation of amounts. Therefore, we did not back-code "other" responses for these items into any newly created categories. We back-coded the "other" responses to K6_a through K6_f when applicable (17 cases). This involved also moving the income reported at K7 to the correct response option for K7_a-K7_f.

6. Coding Open-Ended Items

Item K14 (type of assistance received from other government program) was an open-ended question with no response options. Following a review of the responses in Round 6, we developed categories based on common responses (Table IV.10) and coders used these same categories in Round 7 to code each verbatim response into an established category. If the response did not fit into one of the categories, we kept it as "other."

Table IV.10. Additional Response Categories Used in Section K as During Coding

Item	Question Text	Categories Used
K14	What other assistance did {you/NAME} receive last month?	1 = Housing assistance 2 = Energy assistance 3 = Food assistance 4 = Other

Source: NBS Round 7.

L. Section L—Sociodemographic Information

In Section L, we collected basic demographic information about the beneficiary, such as race, ethnicity, education, parental education, marital status, living arrangements, and household income.

1. Living Situation

In item L11, we asked respondents to indicate whether they live alone; live with parents, guardians, a spouse or partner, or other relative; live with friends or roommates; live in a group setting; or live in some other arrangement. Then, in item L12, we asked respondents to describe the place where they live. We built a soft edit check into the instrument to prompt interviewers to clarify answers in which the respondent indicated that he or she lived alone at item L11, but at item L12 reported living in a group setting, such as a supervised apartment, group home, halfway house, personal care or board-and-care home, assisted living facility, nursing or convalescent home, center for independent living, or some other type of supervised group residence or facility. In some cases, the interviewer suppressed the edit check (61 cases) and the inconsistency remained. For these cases, we recoded question L11 to 4 (live in another group setting).

2. Number of Children

In item L17, we asked how many children under age 18 lived in the sample person's household. We then asked respondents who reported the presence of children how many of the children were their own (item

L19). In 53 cases, the number of the respondent's own children living in the household (item L19) was greater than the number of children living in the household (item L17). For these cases, we set the value of question L19 to don't know (.D).

3. Reporting of Household Income

In item L23Aamt, we asked respondents to provide either their total income in 2018 or the total combined income of their household, before taxes and other deductions. Respondents who experienced difficulty in calculating an annual amount could report their income in monthly, twice-a-month, weekly, biweekly, or daily units (recorded in item L23Ahop). The level of item nonresponse for L23Aamt was 36 percent. We asked those answering "don't know" or "refused" to indicate which of a series of ranges described their income (item L24). Of the 3,115 respondents who did not respond to item L23Aamt, 57 percent (1,765 cases) provided income data in item L24.

We created the construct C_HhInc2018 to combine the responses expressed in various units into an annual income amount. We first examined high and low values of item L23Aamt by unit reported (item L23Ahop) and then examined high and low values on C_HhInc2018 to determine if any appeared to be invalid. Twenty-nine cases reported an annual income of less than $100. After reviewing work status, household size, and other sources of income, we did not change any values of L23Aamt. We created and imputed a more general construct C_FedPovertyLevel (2018 Federal Poverty Level) based on reported income, household size, and respondent's age.

4. Back-Coding Responses to "Other/Specify" Items

As mentioned, we asked respondents to indicate which of a series of items best described their living situation in item L11. Coders reviewed responses of "some other living situation" and back-coded them when possible. We also reviewed "other" responses to item L23ahop (how often paid), and recoded three cases appropriately.

5. Back-Coding Field-Coded Responses

Item L12 (type of place respondent lives) was an open-ended item that required interviewers to attempt to code the respondent's verbatim response into a predetermined category. Coders reviewed responses coded as "other" by the interviewer and back-coded them into existing response options when possible. If a respondent indicated they reside in a townhouse or condo in the "other" field, coders back-coded this response as a single family home.

Several questions in Section L are contingent upon the response to item L12 (type of place respondent lives). Item L15 asks if the home you live in is primarily a place for people with disabilities, and is asked only of sample members living in a group setting. L16 and L17 which collect information on household size, is only asked of respondents who do not live in a group setting. As a result of back-coding "Other" responses at item L12, several cases were recoded into pre-existing responses that would have resulted in different instrument pathing. Specifically, during data cleaning we assigned a value of ".M" to 52 cases at items L16 and L17 that were determined to live in a group setting based on backcoded responses. Additionally, at L11, we recoded ten cases to different responses that necessitated instrument pathing changes. Based on this recoding, ten cases were no longer eligible to receive L16, L17, and L19. We assigned a value of ".L" to these ten cases. Following a review of responses, we added a new response category to question L12 to the Round 7 instrument (Table IV.11). Interviewers did not read this response category to the respondent during the interview, but selected it if it was mentioned by the respondent.

Table IV.11. Additional Response Categories Used in Section L During Coding

Item	Question Text	Categories Used
L12	The next question is about the place where you live. Was this place a…	13 = Homeless

Source: NBS Round 7.

M. Section M—Closing Information and Observations

In Section M, we updated the sample member's contact information so that the incentive check could be mailed. The interviewer recorded the reasons that a proxy or assistance was required, if appropriate, and documented special circumstances.

1. Back-Coding Field-Coded Responses

In items M2a_oth1, M2_oth2, and M13, we asked interviewers to indicate the relationship of the proxy respondent to the sample person. We reviewed responses coded as "other relative" or "other not related" and back-coded them when possible. At item M14 (why assistant/proxy needed), we required interviewers to attempt to code respondents' verbatim responses into predetermined categories. We reviewed responses coded as "other" by the interviewer and back-coded them into existing response options when possible.

2. Respondent Type

At item M11, we asked the interviewer to code whether the interview was conducted with the sample member or with a proxy. We compared this item to the responses asked of sample members only or proxies only. We used the value of M11 and information from these items to construct C_RType. We did not make any edits to the value of M11.

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

V. CONCLUSIONS

In this report, we highlighted data quality issues identified during the NBS–General Waves Round 7 data editing and cleaning process. In an effort to improve data quality in Round 7, we updated the survey instrument to address programming errors identified during the Round 6 data cleaning process. In addition, we conducted more systematic data reviews during the Round 7 data collection effort. As a result of these data reviews, we identified and corrected several minor programming and instrumentation issues.

In general, although survey data processing could have been made more efficient by introducing stricter range checks for unusually high or low values, we were hesitant to apply checks that could have overwhelmed and frustrated respondents by rejecting survey responses during the interview. Any addition of checks must balance the complications associated with survey instrument programming to account for known data complexities against the need to address data complexities after survey completion.

We continued to strengthen interviewer training to emphasize areas of the questionnaire where data problems surfaced during Round 7. Such areas include the importance of correct data entry for job-specific items, probing for sufficient information on open-ended items, and avoiding the suppression of edit checks. The improved training aimed to sensitize interviewers to areas of the questionnaire that were particularly error-prone or to survey concepts that were particularly difficult.

In conclusion, the NBS Round 7 data file provides a rich array of data. As a result of data cleaning and editing, we identified and reported some instances where micro-level errors were obvious.

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

REFERENCES

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 Policy Research, 2021.

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

Grau, E., and H. Zhou. "National Beneficiary Survey—General Waves Round 7: Nonresponse Bias Analysis." Washington, DC: Mathematica, 2021.

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.

Livermore, G., Barrett, K., Wright, D., Stange, M. "National Beneficiary Survey General Waves Design Report." Washington, DC: Mathematica Policy Research, November 2015.

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

O'Day, B., Hannah Burak, K. Feeney, E. Kelley, F. Martin, G. Freeman, G. Lim, and K. Morrison. "Employment and Experiences of Young Adults and High Earners Who Receive Social Security Disability Benefits: Findings from Semi-structured Interviews." Washington, DC: Mathematica, March 2016.

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

APPENDIX A

CHANGES IN QUESTIONNAIRE CONTENT BETWEEN NBS ROUND 6 AND NBS ROUND 7

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

Table A.1. Changes in Questionnaire Content Between Round 6 and Round 7 NBS

Item	Change	Reason
Cross-cutting sections		
Sections A, B, C, C_B, and D	Items modified	Throughout the Round 7 instrument, we added probes for SWS longitudinal sample members, acknowledging that we spoke to them in 2017, that we would like to conduct another interview, and that some questions may sound similar to 2017.
Sections B, C, C_B, D, and L	Items modified	In order to improve respondent and interviewer comprehension, we added clarifying text to the "bi-weekly" response option in questions that ask how often the sample member is paid. This response option now reads "Bi-weekly (every other week)" throughout the survey instrument.
Section A		
A1. Hello, my name is [NAME]. I'm calling on behalf of the Social Security Administration. May I please speak with {IF RTYPE=PROXY AND WE DON'T HAVE PROXY NAME: someone who can answer questions about {NAME's} health, daily activities, and any jobs {he/she} might have?}, else: {FIRST NAME} {LAST NAME}?	Item modified	We added a description of a proxy for the interviewer to read when the name of a proxy was unavailable.
Section B		
B36c. I'm going to ask you about reasons {you/NAME} might have left {your/his/her} last job. Did {you/NAME} leave this job because of {your/NAME's} health, for example, because of worsening illness or the need to go to medical appointments?	Item added	Question added for longitudinal SWS cases.
B36c_1. What was it about {your/NAME's} health that made {you/him/her} leave this job?	Item added	Question added for longitudinal SWS cases.
B36c_1_oth. INTERVIEWER: PLEASE SPECIFY	Item added	Question added SWS for longitudinal sample members.
B36d. Did {you/NAME} leave {your/his/her} last job because of {your/his/her} job, for example because of the need for accommodations or problems with {your/his/her} co-workers?	Item added	Question added for longitudinal SWS cases.
B36d_1. What was it about {your/NAME's} job that made {you/him/her} leave it?	Item added	Question added for longitudinal SWS cases.
B36d_1_oth. INTERVIEWER: PLEASE SPECIFY	Item added	Question added for longitudinal SWS cases.

Item	Change	Reason
B36e. Did {you/NAME} leave {your/his/her} last job because of {your/his/her} personal circumstances, for example because {you/he/she} needed childcare, didn't have reliable transportation, or worried about losing other benefits?	Item added	Question added for longitudinal SWS cases.
B36e_1. What was it about {your/NAME's} personal circumstances that made {you/him/her} leave {your/his/her} last job?	Item added	Question added for longitudinal SWS cases.
B36e_1_oth. INTERVIEWER: PLEASE SPECIFY	Item added	Question added for longitudinal SWS cases.
B36f. Are there any other reasons that we haven't talked about why {you/NAME} left {your/his/her} last job?	Item added	Question added for longitudinal SWS cases.
B36f_oth. What other things made {you/NAME} leave this job?	Item added	Question added for longitudinal SWS cases.
BP4a3_1. What was it about {your/NAME's} personal circumstances that may prevent {you/NAME} from working?	Item modified	Two response options added: 19= MOVED TO ANOTHER AREA 21= LOSS OR POTENTIAL LOSS OF GOVERNMENT BENEFITS
BP4b3_1. What is it about {your/NAME's} personal circumstances that may cause {you/NAME} to not work enough to leave benefits?	Item modified	Two response options added: 19= MOVED TO ANOTHER AREA 21= LOSS OR POTENTIAL LOSS OF GOVERNMENT BENEFITS
Section C		
CP13b1. What was it about [your/NAME's] [main/current] job that might have caused [you/NAME] to have to work less or stop working?	Item modified	Five response options added: 20= FOUND ANOTHER JOB 22= WORK SCHEDULE 23= DID NOT LIKE/GET ALONG WITH CO-WORKERS 24= DID NOT LIKE/GET ALONG WITH MANAGER, SUPERVISOR, OR BOSS 25= DID NOT LIKE/GET ALONG WITH OTHER STAFF RESPONSIBLE FOR HIRING OR PROVIDING ACCOMMODATIONS (SUCH AS HUMAN RESOURCES)
CP13c1. What was it about [your/NAME's] personal circumstances that might have caused {you/NAME} to have to work less or stop working?	Item modified	Two response options added: 19= MOVED TO ANOTHER AREA 21= LOSS OR POTENTIAL LOSS OF GOVERNMENT BENEFITS
C39b. [Do you/Does NAME] work fewer hours or earn less money than [you/he/she] could because [you/he/she]:	Item modified	One response option added: g= [Are/is] in poor health or [have/has] health concerns?

Item	Change	Reason
Section C_B		
C_BP13b1. What was it about [your/NAME's] [main/current] job that might have caused [you/NAME] to have to work less or stop working?	Item modified	Five response options added: 20= FOUND ANOTHER JOB 22= WORK SCHEDULE 23= DID NOT LIKE/GET ALONG WITH CO-WORKERS 24= DID NOT LIKE/GET ALONG WITH MANAGER, SUPERVISOR, OR BOSS 25= DID NOT LIKE/GET ALONG WITH OTHER STAFF RESPONSIBLE FOR HIRING OR PROVIDING ACCOMMODATIONS (SUCH AS HUMAN RESOURCES)
C_BP13c1. What was it about [your/NAME's] personal circumstances that might have caused {you/NAME} to have to work less or stop working?	Item modified	Two response options added: 19= MOVED TO ANOTHER AREA 21= LOSS OR POTENTIAL LOSS OF GOVERNMENT BENEFITS
C_B39b. [Do you/Does NAME] work fewer hours or earn less money than [you/he/she] could because [you/he/she]:	Item modified	One response option added: g= [Are/is] in poor health or [have/has] health concerns?
Section D		
DP1b_1. What was it about [your/NAME's] job that made [you/him/her] leave it?	Item modified	Three response options added: 20= FOUND ANOTHER JOB 22= WORK SCHEDULE 23= SEASONAL/TEMPORARY JOB
DP1c_1. What was it about [your/NAME's] personal circumstances that made [you/him/her] leave the job?	Item modified	Two response options added: 19= MOVED TO ANOTHER AREA 21= LOSS OR POTENTIAL LOSS OF GOVERNMENT BENEFITS
D25a. Did you work fewer hours or earn less money than you could have because [you/he/she] you…	Item modified	One response option added: g=Had medical problems/complications
Section E		
E12. {Have you/Has NAME} ever heard of the student earned-income exclusion? This is a Social Security incentive where if {you are/a beneficiary is} in school, up to $1,870 of earnings per month are not counted when Social Security figures {your/the} benefit.	Item modified	We updated the Student Earned-Income Exclusion rate with 2019 rate.
E15a. Most people receiving Social Security disability benefits will lose their cash benefits if they work and earn more than $1,220 in a month for more than nine months. Is this something {you/NAME} knew before today?	Item modified	We updated the Trial Work Period rate with 2019 rate.

Item	Change	Reason
E15. {Have you/Has NAME} ever heard of a Trial Work Period? This is a Social Security incentive that lets {you/beneficiaries} earn above $880 per month for nine months without losing {your/their} benefits.	Item modified	We updated the Substantial Gainful Activity (SGA) rate with 2019 rate.
Section SS		
SS2b_1. What was it about [your/NAME's] job that makes [you/NAME] think [you/he/she] might go back on benefits?	Item modified	Five response options added: 20= FOUND ANOTHER JOB 22= WORK SCHEDULE 23= DID NOT LIKE/GET ALONG WITH CO-WORKERS 24= DID NOT LIKE/GET ALONG WITH MANAGER, SUPERVISOR, OR BOSS 25= DID NOT LIKE/GET ALONG WITH OTHER STAFF RESPONSIBLE FOR HIRING OR PROVIDING ACCOMMODATIONS (SUCH AS HUMAN RESOURCES)
SS2c_1. What was it about [your/NAME's] personal circumstances that makes [you/NAME] think [you/he/she] might go back on benefits?	Item modified	Two response options added: 19= MOVED TO ANOTHER AREA 21= LOSS OR POTENTIAL LOSS OF GOVERNMENT BENEFITS
Section SB		
SB1b_1. What was it about [your/NAME's] job that made [you/NAME] have to go back on benefits?	Item modified	Five response options added: 20= FOUND ANOTHER JOB 22= WORK SCHEDULE 23= DID NOT LIKE/GET ALONG WITH CO-WORKERS 24= DID NOT LIKE/GET ALONG WITH MANAGER, SUPERVISOR, OR BOSS 25= DID NOT LIKE/GET ALONG WITH OTHER STAFF RESPONSIBLE FOR HIRING OR PROVIDING ACCOMMODATIONS (SUCH AS HUMAN RESOURCES)
SB1c_1. What was it about [your/NAME's] personal circumstances that made [you/NAME] have to go back on benefits?	Item modified	Two response options added: 19= MOVED TO ANOTHER AREA 21= LOSS OR POTENTIAL LOSS OF GOVERNMENT BENEFITS
SB4a. Why {don't you/doesn't NAME} think {you/he/she} will {go back to work / work and earn enough to stay off benefits in the future}?	Item modified	We altered the format of this question from "choose one" to "choose all that apply."

Item	Change	Reason
SB4b. Why {are you/is NAME} unsure about whether {you/he/she} will {go back to work / work and earn enough to stay off benefits in the future}?	Item modified	We altered the format of this question from “choose one” to “choose all that apply.”
Section G		
G13. Where did {you/NAME} go to get this training? Please think about all of the places {you/NAME} went in 2016.	Item modified	One response option added: 9= On the job training (unspecified)
G18. Where did {you/NAME} go to receive these medical services? Please think about all of the places {you/NAME} went in 2016. Did {you/NAME} go to:	Item modified	Two response options added: 10=A rehabilitation/counseling center 11=Physical therapy center
G22. Where did {you/NAME} receive this mental health therapy or counseling? Please think about all of the places {you/NAME} went in 2016. Did {you/NAME} go to:	Item modified	Three response options added: 06=Residential treatment program/facility 07=Rehab center/counseling center/day program 08=Church or religious institution
Section L		
L9. Do {you/NAME} and {your/his/her} {spouse/unmarried partner} live in the same household?	Item modified.	We added unmarried partner to this question text.
L10. {Do you/Does NAME} have a long-term partner who lives in the same household with {you/him/her} in a marriage-like relationship?	Item modified	We corrected the skip logic in the Round 7 survey instrument to ensure that unmarried sample members who are living with their partner (L8=6 and L9=1) are not asked L10, and divorcees (L8=3) are asked L10.
L12. The next question is about the place where you live. Was this place a…	Item modified	We added “homeless” as an unread answer category to this “other/specify” question.

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

Mathematica

Princeton, NJ • Ann Arbor, MI • Cambridge, MA
Chicago, IL • Oakland, CA • Seattle, WA
Tucson, AZ • Woodlawn, MD • Washington, DC

EDI Global, a Mathematica Company

Bukoba, Tanzania • High Wycombe, United Kingdom

mathematica.org