

**Supporting Statement Part B For  
National Beneficiary Survey (Round 8)  
OMB NO. 0960-NEW**

**B. Collections of Information Employing Statistical Methods**

The primary purpose of the NBS is to assess beneficiary well-being and interest in work, learn about their work experiences, and identify how factors such as health; living arrangements; family structure; pre-disability occupations; use of non-SSA programs; obstacles to work; beneficiary interest to return to work; and knowledge of SSDI and SSI work incentive programs promote or hamper long-term work success. Round 8 of the NBS uses a sample design similar to that used for the seven previous rounds of the NBS (conducted by SSA in 2004, 2005, 2006, 2010, 2015, 2017, and 2019). The NBS design uses a dual-mode survey to collect data primarily using computer-assisted telephone interviewing (CATI) and computer-assisted personal interviewing (CAPI) for those who request or require an in-person interview to facilitate their participation in the survey. The survey instrument will be identical in each mode. In all cases, we will attempt to interview the sample person. SSA will seek a proxy respondent only if the sample people are unable to complete a telephone or in-person interview because of their disability. We also propose to conduct a web and paper survey experiment for a sample of 1,000 nonrespondents.

**1. Statistical Methodology**

The target population or “universe” for the Round 8 NBS includes all SSI or SSDI beneficiaries who meet the following criteria:

- Are age 18 to full retirement age (FRA) (age 18 to 65 if receiving SSI, and age 18 to 66 if receiving SSDI)
- Are in active pay status<sup>1</sup> as of June 30 of the sampling year (2022 for Round 8) in either the SSI or SSDI program
- Are not nondisabled dependents of SSDI beneficiaries

To maintain consistency and support trend analyses, we will apply essentially the same sample selection criteria for Round 8 of the NBS that we used to prepare the national samples in the prior NBS rounds. In Round 8 of the NBS, we will select a nationally representative sample of active SSI recipients and SSDI beneficiaries (the Representative Beneficiary Sample, or RBS). We will also independently select a sample of SSI recipients and SSDI beneficiaries whom we identified (using SSA administrative data) as a subset of beneficiaries who maintained a minimum level of earnings for a sustained period (a “successful worker” sample or SWS).

**Representative Beneficiary Sample.** For the RBS, the target population includes SSI recipients and SSDI beneficiaries in all 50 states and the District of Columbia. The estimated size of the target population for the RBS is approximately 13 million. To ensure enough sample members seeking work, Mathematica stratifies the active beneficiary population into

---

<sup>1</sup>Active status includes beneficiaries who currently receive cash benefits, as well as those whose benefits SSA temporarily suspended for work or other reasons.

four categories: (1) beneficiary had positive earnings in the year prior to the sampling year (2021 for Round 8); (2) age 18 to 39, no earnings in the year prior to the sampling year; (3) age 40 to 49, no earnings in the year prior to the sampling year; and (4) age 50 to FRA, no earnings in the year prior to the sampling year (Table B.1).<sup>2</sup> We will select people in the first category (those most likely to be working at interview) at a higher rate than other categories; we will select people in the fourth category (those least likely to be working at interview) at the lowest rate.

**Successful Worker Sample.** The target population for the SWS is a subset of the RBS target population that includes SSI or SSDI beneficiaries who maintained a minimum level of earnings for a sustained period. Because of the lag in identifying earnings for some successful workers, the SWS target population is further limited to successful workers SSA identified using administrative data at the time of sample selection. The weights from this selected sample are therefore provisional. Approximately three years after the end of data collection, we will review updated administrative data from the same period and identify the beneficiaries who maintained a minimum level of earnings for a sustained period during that period, which will include many of the successful workers that could not be included in the initial frame.<sup>3</sup> Using these updated data, we will post-stratify the weights so that the final sample more closely matches the full successful worker population.<sup>4</sup> We will draw the SWS from the same frame as the RBS, with the same criteria for inclusion listed above for the RBS. However, SSA will include additional criteria for the SWS:

- Identified as having three consecutive months of earnings above the nonblind substantial gainful activity (SGA) level,<sup>5</sup> based on earnings data reported in SSA’s Disability Control File (DCF)<sup>6</sup>
- No older than age 62 as of June 30, 2022 (the sampling month)<sup>7</sup>

---

<sup>2</sup> This stratification is a departure from the stratification used in Rounds 1 through 7, which was based solely on age categories. It takes advantage of the relationship between earnings in the year prior to the sampling year and whether the respondent was working at the time of the interview. Earnings information was not readily available for stratification and sampling prior to Round 6.

<sup>3</sup> This updated information will contain about three times the number of successful workers that will be in the provisional frame from which the sample is selected. A small number of cases who were initially identified as successful workers in the provisional frame may subsequently be found to not meet the successful work criteria.

<sup>4</sup> The SWS will therefore not be representative of all SSI/SSDI beneficiaries who met the definition of successful work because the earnings for some of these beneficiaries are not available in SSA administrative data in time for sample selection. Applying the final post-stratified weights will alleviate some of this bias, but we have no opportunity to interview successful workers whose earnings were not included in the administrative data in time to be included in the provisional frame.

<sup>5</sup> To be eligible for disability benefits, a person must be unable to engage in SGA. In 2022, the monthly nonblind SGA threshold is \$1,350. (From <https://www.socialsecurity.gov/oact/cola/sga.html>. Accessed 2/13/2022.)

<sup>6</sup> The DCF is a centralized, electronic database that stores, supports, and controls data on post-entitlement disability-related actions and determinations.

<sup>7</sup> This constraint on the age of the successful worker was applied in Round 6 to ensure successful workers who indicated that they were working at the time of the Round 6 interview (and therefore eligible for the Round 7 longitudinal sample) would still be age-eligible for disability benefits at Round 7. Although no sample members will be followed longitudinally in Round 8, we will apply the same constraint to maintain consistency across rounds.

Based on data from prior NBS rounds, we expect the size of the SWS target population we will be able to identify at sample selection will be about 100,000.<sup>8</sup> Note that because of the COVID-19 pandemic, the size of the actual SWS target population for Round 8 might differ. To identify enough respondents whose successful work began recently for the SWS, we must create seven subset frames at six-week intervals during the data collection period. Each of the samples from these subset frames will act as strata, with independent samples of approximately equal size drawn from each frame. Beneficiary type, or “Title” (SSDI only and SSI [with or without SSDI]) will act as a substratum within each subset frame. In Table B.1, we collapsed the seven samples into a single sample, broken out by beneficiary title.

For the NBS Round 8, we plan to complete 5,000 interviews with the RBS sample and 3,000 interviews with the SWS sample.

**Table B.1. NBS Round 8 Sample Sizes by Strata**

<b>Sampling Strata</b>	<b>Sample Size</b>	<b>Target # of Completed Interviews</b>
RBS	6,250	5,000
Positive earnings in 2021	1,500	1,200
Zero earnings in 2021, age 18 to 39	2,000	1,600
Zero earnings in 2021, age 40 to 49	1,375	1,100
Zero earnings in 2021, age 50 to FRA	1,375	1,100
SWS	3,750	3,000
SSI (with and without SSDI)	1,875	1,500
SSDI only	1,875	1,500

The sample design will include the selection of 80 primary sampling units (PSUs), along with selection of zip-code-based secondary sampling units (SSUs), within certainty PSUs. Specifically, we will select PSUs using a four-level composite size measure, incorporating the four age-based strata of the RBS. The RBS and SWS will use the same PSUs; SSUs will only be used in the RBS. Subsequent rounds for both the RBS and SWS will include the same set of PSUs and, for the RBS only, SSUs.

To administer the web and paper experiment, we will sample 1,000 SWS nonrespondents from multiple sample extracts after those extracts on the main NBS effort. We will randomly assign the 1,000 non-respondents and offer 500 a web option and 500 a paper questionnaire.

The target response rate for the RBS and SWS is 80 percent; however, based on Mathematica’s experience with prior rounds of the NBS, we expect this will be difficult to obtain. We recognize that it is becoming increasingly challenging to locate sample members and gain their cooperation with the survey process. To achieve the target number of completed interviews, we will release as many sample cases as needed (releasing additional sample cases in waves after the initial release as necessary). If the response rate for both samples is less than 80 percent, we will conduct a nonresponse bias analysis and take the results into account during weighting procedures.

<sup>8</sup> We define the SWS target population here in terms of the population of successful workers that is available for sampling. In the final post-stratification three years after data collection, the SWS target population will be expanded to include most successful workers.

The target number of responses rate for the web experiment is 125 (25 percent) and paper experiment is 100 (mode 20 percent). These completed experimental surveys are in addition to the 3,000 completed surveys we expect for the SWS sample for the NBS.

## 2. Procedures for Collecting the Information

SSA will use a multi-stage clustered design similar to the design used in prior rounds to facilitate in-person interviews of beneficiaries selected for the NBS, but who cannot be reached by telephone or who cannot be interviewed by telephone because of their disability or impairment. For the multi-stage design, we used recent test frame information to create PSUs based on the counts of eligible SSDI beneficiaries and SSI recipients in each county, as well as the counts of successful workers, to form PSUs consisting of one or more counties. The list of PSUs consists of 1,326 units.<sup>9</sup> From this list, we will select a stratified national sample of 80 PSUs. As in the previous NBS rounds, we anticipate that we will select two of these 80 PSUs, corresponding to Los Angeles and Cook (Chicago) counties, with certainty because of the number of disability beneficiaries in these locations, though it is possible that the set of certainty PSUs will differ from these two. Because of the size of the certainty counties, we will form SSUs using beneficiaries' zip codes and select the SSUs as an intermediate stage of selection before selecting beneficiaries. We will select the PSUs and SSUs with probability proportional to size, where we defined size as a composite size measure that accounts for the number of active beneficiaries and recipients in each age group. In prior rounds, we selected four and two SSUs from the Los Angeles and Cook (Chicago) counties, respectively; the number of SSUs that we select in each certainty county will depend upon the relative size of those counties.

**Estimation Procedure:** The analysis involves computation of descriptive statistics (means and percentages) for the entire sample or specified subsamples. We will use multivariate models (primarily multiple regression and probit or logit models) in some instances.

The analysis of survey data from such complex sample designs requires the use of weights to compensate for various probabilities of selection and special methods to compute standard errors. We compute from the inverse of the selection probability the base weight associated with a sampled SSDI beneficiary or SSI recipient for the NBS survey. The probability of selection is the product of the selection probability at each sampling stage—the PSU (as needed), and the individual. Therefore, the initial sampling weight will be the inverse of the full selection probability for each case. The following component probabilities are the basis to calculate the probability of selection:

1. The probability of selecting PSU  $i$  within PSU stratum  $h$ ,  $\pi_{hi}$ , is  $\pi_{hi} = 1$  for certainty PSUs; for noncertainty PSUs, the selection probability is given by

$$\pi_{hi} = n_h \frac{\text{MOS}_{hi}}{\text{MOS}_h},$$

where  $n_h$  is the sample size for stratum  $h$ . Typically,  $n_h = 1$  or  $2$ .

---

<sup>9</sup> This is close to the number of PSUs that were used in Rounds 1 through 7 (1,330), which was based on the number and distribution of beneficiaries in Round 1.

2. If secondary units are selected within the  $hi$ -th PSU, the probability of selecting secondary unit  $j$  is given by

$$\pi_{hij} = n_{hi} \frac{MOS_{hij}}{MOS_{hi}} .$$

where  $n_{hi}$  is the sample size for secondary units in PSU  $hi$ ,  $MOS_{hij}$  is the measure of size of the secondary unit, and  $MOS_{hi}$  is the total measure of size for all secondary units in PSU  $hi$ .

3. When subareas are used, the probability of selecting a given beneficiary within stratum  $s$  of secondary unit  $j$  in the  $hi$ -th PSU is given by

$$\pi_{hijsk} = \frac{n_{hijsk}}{N_{hijsk}} ,$$

where  $n_{hijsk}$  and  $N_{hijsk}$  are the sample and population size, respectively, for the  $hijsk$ -th stratum within secondary unit  $j$  of PSU  $hi$ , assuming subareas are used. When subareas are not used,  $j$  drops out of the subscripts.

Finally, the overall selection probability is given by the following:

$$\text{Overall selection probability} = \pi_{hi} \pi_{hij} \pi_{hijsk} .$$

The initial sampling weight is calculated as

$$\text{Base weight} = W_{hijsk} = \frac{1}{\pi_{hi} \pi_{hij} \pi_{hijsk}} .$$

The subscript  $j$  is dropped from the last two formulas for PSUs in which subareas are not sampled.

The use of base weights will yield unbiased estimates if there is adequate coverage and no survey nonresponse. Unit nonresponse (that is, whole questionnaire nonresponse) occurs when an eligible sampled beneficiary fails to respond to the survey. We will adjust the base weights with propensity scores to reduce the potential for bias due to unit nonresponse, created using logistic regression models. Covariates in these models are variables available for both respondents and nonrespondents and we choose them because of their relation to the likelihood of poor survey response and an assumed relationship to the data outcomes. At a minimum, candidates for covariates used in the models will include the strata used in sampling. It is important that each level of the model covariates has enough sample members to ensure a stable adjustment. As with prior rounds, Mathematica develops two logistic propensity models: one for locating a person and another for response among located individuals. We will develop the models using data in the SSA database available on all sample members, which is extensive. The location and response logistic models provide estimated propensity scores for each respondent accounting for individuals with similar

characteristics who we cannot locate or who did not respond. We will use the inverse of the propensity score as the adjustment factor. The adjusted weight for each sample case will be the product of the initial sampling weight and the adjustment factor.

We will use propensity modeling instead of the standard weighting class procedure because it allows us to use more factors and complex interactions among factors to explain the differential propensity located or to respond. In addition, we will use available standard statistical tests to evaluate the selection of variables for the model. To identify the factors for inclusion in the models, we will use bivariate cross-tabulations and multivariate procedures, such as interaction detection procedures (for example, Chi-squared Automatic Interaction Detection, or CHAID, software). To evaluate the candidate factors and interactions, we will use a weighted stepwise procedure. We will then check the final model using survey data analysis software to obtain design-based precision estimates for assessing the final set of factors. We expect to require separate models for some survey populations because the factors explaining the ability to locate a person or response could be unique to these populations (for example, people whose benefits are suspended due to work versus people in current pay status).

After adjusting for nonresponse, we will further adjust the weights so that some weighted sample statistics match known population values, using the raking method for post-stratification. For example, if the weights for beneficiaries by program (SSDI only and SSI) do not correspond to population values, we will adjust the weights in a proportional fashion, so the weighted sample and population values correspond. Potentially, we can control population statistics for any variable observed in SSA administrative data. The variables we will most likely use are program; age; sex; earnings level; census division, and primary impairment. For the SWS, we will conduct a second, final post-stratification three years after the end of data collection. This final post-stratification will account for successful workers whose earnings were not recorded in the DCF in time to be included in the provisional SWS frame used for sampling.

In computing the final weights, some individuals might have large weights. Variability in sampling weights can severely affect standard errors, particularly in the extreme case where one observation has a sampling weight that is substantially larger than others. We will use “weight trimming” to alleviate this problem. In this procedure, the value of very large weights is simply reduced, with the amount “trimmed” being distributed among other sample members in some way. Reducing the weight can create biased estimates, but when one or two individuals have extremely large weights, the contribution to variance reduction outweighs the bias that might be created by trimming.

One way to protect against bias is to redistribute the “trimmed” amount over a group of individuals who share some common characteristic with those whose weights were trimmed. We will define these “trimming classes” using variables selected in the same manner we use to select variables for the nonresponse adjustments. Because we will use propensity modeling instead of weighting classes to make the nonresponse adjustments, we will define trimming classes using the most important variables in the propensity models.

**Standard Errors:** For the NBS, the sampling variance estimate is a function of the sample design and the population parameter we are estimating; this is the design-based sampling variance. The design-based variance assumes the use of “fully adjusted” sampling weights, which derive from the sample design with adjustments to compensate for locating a person; individual nonresponse; and ratio-adjusting the sampling totals to external totals. We will follow the same method used in the prior NBS, developing a single fully adjusted sampling weight and information on analysis parameters (that is, analysis stratification and analysis clusters) necessary to estimate the sampling variance for a statistic, using the Taylor series linearization approach.

The Taylor series procedure is the most appropriate sampling variance estimation technique for complex sample designs like the NBS. The Taylor series procedure is based on a classic statistical method in which one can approximate a nonlinear statistic by a linear combination of the components within the statistic. The accuracy of the approximation is dependent on the sample size and the complexity of the statistic. For most commonly used nonlinear statistics (such as ratios, means, proportions, and regression coefficients), the linearized form is already developed and has good statistical properties. Once a linearized form of an estimate is developed, one can use the explicit equations for linear estimates to estimate the sampling variance. Because one can use the explicit equations, one can estimate the sampling variance using many features of the sampling design (for example, finite population corrections; stratification; multiple stages of selection; and unequal selection rates within strata). This is the basic variance estimation procedure used in SUDAAN, the survey procedures in SAS, STATA, and other software packages to accommodate simple and complex sampling designs. To calculate variance, we will need sample design information (such as stratum and analysis weight) for each sample unit.

**RBS.** Table B.2 shows the minimum detectable differences for age subgroups, a measure of the smallest difference between subgroups that we will be able to detect in 5,000 completed surveys with 80 percent power and 95 percent confidence. For example, for a proportion of 0.10, a minimum detectable difference equal to 5.0 percentage points indicates that if 10 percent of the beneficiaries are employed and never attended college, and at least 15.0 percent of beneficiaries are employed and attended at least some college, the analysis will detect a significant difference between these groups. The table presents minimum detectable differences where we compare one half of the sample to the other half and 70 percent of the sample to 30 percent.

**Table B.2. Estimated Minimum Detectable Differences Between RBS Age Subgroups**

Subgroup	Half the Sample Compared to Other Half (2,500 vs. 2,500 for full sample)			70% of Sample Compared to 30% (3,500 vs. 1,500 for full sample)		
	Mean of Binomial Distribution			Mean of Binomial Distribution		
	10%	30%	50%	10%	30%	50%
All	3.7%	5.7%	6.2%	4.1%	6.2%	6.8%
Age 18 to 29	7.3%	11.1%	12.2%	8.0%	12.2%	13.3%
Age 30 to 39	7.5%	11.4%	12.4%	8.1%	12.4%	13.6%

Subgroup	Half the Sample Compared to Other Half (2,500 vs. 2,500 for full sample)			70% of Sample Compared to 30% (3,500 vs. 1,500 for full sample)		
	10%	30%	50%	10%	30%	50%
Age 40 to 49	7.1%	10.8%	11.7%	7.7%	11.8%	12.8%
FRA	5.9%	8.9%	9.8%	6.4%	9.8%	10.6%

Note: This table assumes a comparison of binary outcomes, with a 5 percent significance level for a two-sided test and 80 percent power. Design effects are estimates obtained from the Round 7 and Round 8 test NBS.

The minimum detectable difference between two populations of an estimated percentage,  $\hat{p}$ , can be approximated by the following formula:

$$Var(\hat{p}) = \frac{p(1-p)}{n_{eff1}} + \frac{p(1-p)}{n_{eff2}}$$

where  $n_{eff1}$  and  $n_{eff2}$  are the effective sample sizes of the two populations being compared and  $n_{eff1} = n_1 / deff_1$ ,  $n_{eff2} = n_2 / deff_2$ . We estimated the design effect using data from the Round 7 NBS.

**SWS.** Of the 3,000 completed SWS interviews in Round 8, 1,500 will be among SSI recipients who may or may not concurrently receive SSDI and 1,500 will be SSDI-only beneficiaries. Because SSA is interested in differences between SWS beneficiaries who are receiving or returning to benefits and those who are not, we may conduct a comparison of interest between these groups. Because about half of the SWS is likely to be receiving or returning to benefits at the time of data collection, such a comparison might involve comparing the 50 percent of the sample not receiving benefits (N=1,500) and the 50 percent receiving or returning to benefits (N=1,500). Table B.3 presents minimum detectable differences for a comparison between two halves of the SWS and a comparison between 70 percent and 30 percent of the SWS. We also show the minimum detectable differences for comparisons by title (SSI, SSDI, and concurrently receiving both). Those estimates assume that 57 percent of the SSI stratum is SSI only and 43 percent is concurrently receiving SSDI.

**Table B.3. Estimated Minimum Detectable Differences Between SWS Subgroups**

Group	Half the Sample Compared to Other Half (1,500 vs 1,500 for full sample)			70% of Sample Compared to 30% (900 vs. 2,100 for full sample)		
	Mean of Binomial Distribution			Mean of Binomial Distribution		
	10%	30%	50%	10%	30%	50%
All	3.4%	5.3%	5.7%	3.8%	5.7%	6.3%
SSI only	6.4%	9.8%	10.7%	7.0%	10.7%	11.7%
SSDI only	4.8%	7.4%	8.1%	5.3%	8.1%	8.9%
Concurrent	7.5%	11.5%	12.5%	8.2%	12.5%	13.7%

Note: This table assumes a comparison of binary outcomes, with a 5 percent significance level for a two-sided test and 80 percent power. Design effects are estimates obtained from the Round 7 NBS.

The following formula approximates the minimum detectable difference between two populations of an estimated percentage,  $\hat{p}$  :



$$\text{Var}(\hat{p}) = \frac{n_{eff1} - 1}{n_{eff1}} \frac{p_1(1-p_1)}{n_{eff1}} + \frac{n_{eff2} - 1}{n_{eff2}} \frac{p_2(1-p_2)}{n_{eff2}},$$

where  $n_{eff1}$  and  $n_{eff2}$  are the effective sample sizes of the two populations being compared and  $n_{eff1} = n_1 / deff_1$ ,  $n_{eff2} = n_2 / deff_2$ . The design effects are obtained using data from the Round 7 NBS, with both clustered and unclustered components in a dual sample design.

**Unusual Problems Requiring Specialized Sampling Procedures:** For the SWS strata, it is possible that the number of beneficiaries with successful work within the PSU areas, and within the three beneficiary types, will be too small to meet sample size requirements. Even though we have data from Round 7 to help us evaluate the number of successful workers in PSUs, we do not know whether the Round 7 experience is representative of what we will see in Round 8. We plan to use a hybrid design, referred to as a dual sample design, which combines an unclustered stratified random sample with the clustered sample design. Although we will ensure both the unclustered and clustered samples will be nationally representative, we will limit the data collection in the unclustered component to CATI-only (no in-field follow-up or interviewing) due to the high cost that would be associated with field follow-up for the unclustered cases. For national estimates, we will compute sampling weights to account for this “dual-sample” strategy, as we have in the prior NBS.

The result will be lower response rates for the non-PSU sample members, and potential bias in the estimates. To address the bias issue, we will compare the responses of the within-PSU phone interview sample to those of the within-PSU in-person interview sample.

We will administer Round 8 of the NBS once in 2023. Beneficiaries will complete the survey one time only. Thus, there is no cyclic burden for these respondents.

### 3. Methods to Maximize Response Rates

Locating sample members is the first challenge to obtaining a high response rate. Although SSA has contact information for all potential respondents, we know from past experience that it often does not lead directly to the sample member. Telephone numbers are particularly problematic because there is no administrative reason to keep them updated in SSA records. Addresses can be more reliable because SSA sometimes uses them for mailing correspondences. These might, however, be a post office box, address of a guardian, financial institution, or other types of addresses that can make it difficult to locate the sample member. Because SSA now requires direct deposit of payment checks, it has minimized the importance of keeping address information current. To improve contact information, we will mail an advance letter on SSA letterhead and a study brochure to each sample member before attempting to survey them using the address of record. The letter will describe the survey and indicate that we will soon contact the sample member. We will begin locating with letters returned to Mathematica as undelivered. When an address is available without a phone number, we will conduct a directory search to obtain a phone number. When directory searches are unproductive, we will submit searches to Accurint, a comprehensive database compiled from multiple sources, and use locating letters and telephone tracing (calling former neighbors or representative payees). In Round 7 of the NBS, we located approximately 86 percent of the sample in this way.

If we find or obtain an available phone number, we will attempt to call the respondent to conduct the interview. Mathematica will use a protocol that calls for repeat efforts, including attempts on different days and different times. If we make successful contact and the beneficiary or recipient consents to the interview, the caller will conduct the interview using CATI technology.

For sample members who do not respond to our telephone calls or refuse to participate in the survey, we will mail locating letters, reminder letters, reminder postcards, and refusal conversion letters, as appropriate. After two to three months of CATI interviewing, we will begin to transfer cases to field (in-person) data collection staff for locating. Delaying the start of field locating and interviewing allows an adequate number of cases to accumulate so field staff will have enough work for travel to be cost-effective. Prior to deploying field staff, we will send all cases assigned to the field with a valid mailing address a pre-field letter, informing them that a representative from Mathematica will be visiting their home. Once in the field, staff will have several other tools at their disposal to support field locating efforts, including a “Sorry I Missed You” card, appointment card, post-office letter, study brochure, interviewer field letter, and locating checklist. The locator checklist identifies steps a field interviewer should take when locating a respondent, with the steps listed hierarchically from most to least likely to be effective. The checklist helps prevent duplication of efforts and sets clear parameters for when a case should cease because of nonresponse. We will train locators to not reveal private information about a sample member to any informants, including the study’s name or unique details about the study. We will equip all field staff with cellular telephones so sample members, once found, can call into Mathematica’s telephone center to complete the interview. We will monitor respondent characteristics throughout the data collection effort to detect the potential for bias so that resources could be allocated as needed to target specific sub-groups.

To promote response among Spanish speakers, we will develop a Spanish-language version of the CATI instrument that Spanish-speaking interviewers will administer. In addition, we will mail respondent letters and notification materials in Spanish to sample members who indicated a preference for Spanish language communications according to SSA’s administrative data. We will also provide Spanish versions of the advance letters and other communications to all sample members who live in NBS sampling areas with a large proportion of beneficiaries with a Spanish-language preference.

The impairments of some sample members will make responding to the survey problematic, especially by telephone. To facilitate responses to the CATI interview, we will offer the use of several assistive devices (for example, amplifiers, Telecommunications Relay Service, instant messaging,) and will instruct interviewers to remain patient; repeat answers for clarification; and identify signs of respondent fatigue so we can complete the interview in one session if possible. Despite these efforts, we know that some respondents will be unable to complete the interview by telephone and others will be unable to complete the interview at all.

To increase opportunities for self-response, we will permit assisted interviews, which differ from proxy interviews in that the sample members answer most questions themselves. The assistant, typically a family member, provides encouragement, interpretation, and verifies answers as needed. These interviews minimize item nonresponse; improve response

accuracy; and help with some limiting conditions such as hearing difficulties and language barriers.

As a last resort, we will rely on proxy respondents to complete the survey on behalf of sample members who are unable to do so (even with assistance) either by telephone or in person. This includes individuals with severe communication impairments or physical disabilities that preclude participation in any mode, and those with cognitive impairments that might compromise data quality. The use of proxies can minimize the risk of nonresponse bias that would result from the exclusion of individuals with severe physical or cognitive impairments. To identify the need for proxy respondents, we will administer a mini-cognitive test used in prior rounds of the NBS. The test provides interviewers a tool for determining when to seek a proxy rather than leaving the decision to interviewer discretion or a gatekeeper.

SSA will mail sample members a \$30 gift card upon completion of the Round 8 NBS interview. In addition to the post-paid incentive, we propose incorporating a \$2 cash pre-paid incentive experiment into the Round 8 incentive structure. To operationalize the experiment, we will randomly assign all of the approximately 10,000 RBS and SWS members in the first sample release to a treatment or control group. Treatment group members will receive \$2 cash inside of their advance letter, and a \$30 gift card upon completion of the survey. Control group members will receive a \$30 gift card upon survey completion.

**Dealing with Issues of Nonresponse:** We will adjust the base weights for survey nonresponse using the procedures described above and to control distributions for some variables to known totals from the administrative data. We can assess the extent of remaining bias by comparing weighted outcomes for the survey sample that we can observe in administrative data (for example, annual earnings and SSI and SSDI payments) to outcomes for the population that the weighted sample is intended to represent.

#### 4. Tests of Procedures

We based the Round 8 survey on the previous surveys, with the exception of the following new questions:

- Question B2a: This new question captures a self-reported limitation scale (0 to 100) about the respondent's main physical or mental health condition.
- Questions C5c and C\_BP5c: These new questions ask how respondents reported their work status to SSA.
- Questions CP9a, C\_BP9a, and DP4: These new questions ask if most other workers at the place of employment have disabilities and are part of the additions to the instrument, to determine if the job is competitive integrated employment or some other employment program. They are similar to questions from the prior round that were removed: CP10 and C\_BP10.
- Questions CP9b, C\_BP9b, and DP5: These new questions ask if anyone who has a disability or does not have a disability could have taken the respondent's job. They are part of the additions to the instrument to determine if the job is competitive integrated employment or some other employment program.

- Questions C39b.d2, C\_B39b.d2, and D25a.d2: These new questions ask if the respondent worked fewer hours or earned less money, because they want to keep other benefits like food stamps, housing assistance, or workers' compensation.
- Questions C39\_7, C39\_8, C\_B39\_7, C\_B39\_8, D31, and D32: For respondents who were asked to repay SSA because they were working while receiving benefits, these new questions ask if the respondents tried to appeal or challenge the request to repay benefits and, if so, if they ended up repaying some, all, or none of the benefit amount.
- Questions EP2a and EP2b: These new items ask about respondents' awareness of rules that allow work while on disability benefits and rules to report work status to SSA.
- Question E22: This new item asks respondents to provide a usefulness rating for the Ticket to Work program.
- Questions G59 and G59a: These new questions collect information about respondents' experiences about getting information about services they need.

SSA tested the new aspects of the Round 8 survey by conducting pretests of the NBS instruments in June 2022. The purpose of the pretests was to test the order, flow, clarity, and timing of the instruments, especially for new or modified questions. We also tested the new web and paper versions of the NBS that will be incorporated in the Round 8 causal experiment. In addition to the pretests, we conducted an abbreviated cognitive interviewing session at the conclusion of each pretest to better understand how respondents interpreted and answered any new or modified questions. The pretest sample was drawn from contact information that SSA pulled to match a test SWS extract. Following the pretest, we made minor updates to the instruments to improve the clarity and flow of the instruments.

## 5. Statistical Agency Contact for Statistical Information

John Travis Jones, Contracting Officer Representative  
 Social Security Administration Office of Research, Demonstrations and Employment Support  
 Telephone: 443-834-8092  
 Email: [John.Travis.Jones@ssa.gov](mailto:John.Travis.Jones@ssa.gov)

Eleanor Stinnett, Contracting Officer Representative  
 Social Security Administration  
 Telephone:  
 Email: [Eleanor.Stinnett@ssa.gov](mailto:Eleanor.Stinnett@ssa.gov)

Eric Grau, Statistician,  
 Mathematica  
 Telephone: 609.945.3330  
 Email: [EGrau@mathematica-mpr.com](mailto:EGrau@mathematica-mpr.com)

Gina Livermore, Senior Fellow  
 Mathematica  
 Telephone: 202-642-3462  
 Email: [GLivermore@mathematica-mpr.com](mailto:GLivermore@mathematica-mpr.com)

Mustafa Karakus, Associate Director  
Westat  
Telephone: 240-370-4907  
Email: [MustafaKarakus@westat.com](mailto:MustafaKarakus@westat.com)