SUPPORTING STATEMENT FOR

THE NATIONAL BENEFICIARY SURVEY—

GENERAL WAVES AND SEMI-STRUCTURED INTERVIEWS

Revision Request

OMB No. 0960-0800

Supporting Statement: Part B

CONTENTS

B. Collections of Information Employing Statistical Methods

B.1. Statistical Methodology

B.1.1. Respondent Universe and Sampling Methods

In this information collection request (ICR), The Office of Retirement and Disability Policy within the Social Security Administration (SSA) seeks a three year extension for the National Beneficiary Survey (NBS)-General Waves (OMB no. 0960-0800; expiration date September 30, 2017).

Originally, SSA intended the new NBS-General Waves survey design to include a national sample of SSA disability beneficiaries, and a sample of beneficiaries whose benefits are in suspense due to successful work in all three rounds of data collection. We plan to complete approximately 4,000 interviews with active beneficiaries in each of the three rounds while the sample sizes for those whose benefits SSA placed in suspense would vary across rounds.[1] In addition, we will also take into account in rounds 2 and 3 some beneficiaries we identified as individuals who were in suspense status at the time of the round 1 interview.

However, due to difficulties associated with developing a sample design that provides sufficient numbers of beneficiaries who earned enough to have their benefits suspended in the recent past, SSA postponed the start of the NBS-General Waves survey to 2015, and did not include a sample of successful workers as part of round 1. In lieu of including a sample of successful workers in the round 1 survey data collection, we conducted 90 semi-structured interviews with this group. This gave us additional time to settle on an adequate design for the successful worker sample so that they could be in rounds 2 and 3, while still allowing us to collect important information about factors that aid or inhibit beneficiaries in their efforts to obtain and retain employment on the general sample. Mathematica Policy Research (Mathematica), the data collection contractor for all prior rounds, modified the sample design and revised the questionnaire in preparation for rounds 2 and 3 of the NBS-General Waves (see Attachment A). Mathematica will conduct the two remaining rounds of data collection and will prepare data files and documentation.

The primary purpose of the NBS-General Waves is to assess beneficiary well-being and interest in work; learn about their work experiences (successful and unsuccessful); and identify how factors such as health; living arrangements; family structure; pre-disability occupation; use of non-SSA programs (e.g., Supplemental Nutrition Assistance Program or SNAP); knowledge of Social Security Disability Insurance (SSDI) and Supplementary Security Income

[1] At round 1, we planned to interview approximately 4,500 suspended beneficiaries due to their work. In rounds 2 and 3, we plan to complete approximately 3,000 interviews with suspended beneficiaries selected for the cross-sectional samples. In addition, we will follow 2,500 suspended beneficiaries from round 1 longitudinally in rounds 2 and 3.

(SSI) work incentive programs; obstacles to work; and how beneficiary interest to return to work promote or restrict long-term work success. The NBS-General Waves uses a sample design similar to that used for prior NBS (conducted by SSA in 2004, 2005, 2006, and 2010). Under the current OMB approval, SSA completed the first wave of the NBS-General Waves in 2015. We will conduct the subsequent two rounds in 2017 (round 2) and 2019 (round 3). We collect survey data not available from SSA administrative data or other sources. The NBS's design is as a dual-mode survey to: collect data primarily using CATI with CAPI available 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. We will seek a proxy respondent only if the sample persons are unable to complete either a telephone or in-person interview as a result of their disability.

B.1.2. Universe and Sample

The target population or "universe" for the NBS-General Waves includes all SSI or SSDI beneficiaries who meet the following criteria:

- Between the ages of 18 and full retirement age (FRA)-18 to 65 if receiving SSI, and 18 to 66 if receiving SSDI.
- In active benefit status[2] as of June 30 of the sampling year (2016 for round 2 and 2018 for round 3) in either the SSI or SSDI program
- Are not nondisabled dependents of SSDI beneficiaries

To maintain consistency and support trend analyses, we apply essentially the same sample selection criteria for round 1 of the NBS-General Waves that we used to prepare the national samples in the prior NBS. In round 1, we selected only a nationally representative sample of active SSI recipients, SSDI and concurrent beneficiaries (the Representative Beneficiary Sample, or RBS). In rounds 2 and 3, we will apply the same criteria for the RBS, but will also independently select a sample of SSI recipients and SSDI beneficiaries whom we identified (using SSA administrative data) as having had high earnings from work in the year prior to the interview. We refer to this sample as the 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 exceeds 14 million. To ensure a sufficient number of persons seeking work, the contractor stratifies the active beneficiary population into four age categories: 18 to 29; 30 to 39; 40 to 49; and 50 years old or over, as seen in Table B.1. We

[2] Active status includes beneficiaries currently receive cash benefits, as well as those whose benefits SSA temporarily suspended for work or other reasons. It does not include beneficiaries in terminated status.

will select persons in the younger age categories at a higher rate than those in the oldest age category.

Successful Workers - In rounds 2 and 3, we plan to add SWS to the NBS. The target population for the SWS is a subset of the RBS target population, only including those SSI or SSDI beneficiaries who meet the criteria for successful work. Moreover, 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.[3] We will draw the sample of successful workers 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 Substantial Gainful Activity (SGA)[4] threshold, based on monthly earnings data from the SSA Disability Control File (DCF) administrative data[5]
- No older than 62 years of age as of June 30th of the sampling year (to ensure age eligibility for the next round)

Based on data from 2011 and 2013, we anticipate that the size of the SWS target population we can identify at the time of sample selection based on DCF earnings data is about 65,000.

For the SWS, to identify a sufficient number of respondents whose successful work began recently, we will need to 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, either with SSDI or not) 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.

In round 3, we plan to re-interview successful workers who completed an interview at round 2.[6] We anticipate that approximately one-half of successful

[3] The sample of successful workers will therefore not be representative of all SSI/SSDI beneficiaries who met the definition of successful work, since the earnings for some of these beneficiaries are not available in SSA administrative data in time for sample selection.

[4] To be eligible for disability benefits, a person is unable to engage in substantial gainful activity (SGA). A person who earns more than a certain monthly amount is ordinarily considered engaging in SGA. The monthly SGA amount for statutorily blind individuals for 2016 is $1820. For non-blind individuals, the monthly SGA amount for 2016 is $1130. (From https://www.socialsecurity.gov/oact/cola/sga.html. Accessed 2/24/2016.)

[5] The DCF is a centralized, electronic database that stores, supports, and controls data on post-entitlement disability-related actions and determinations.

[6] In the questionnaire, we ask successful workers if they are currently working or worked the past 6 months to ensure that all those interviewed are still employed, or recently employed,

workers at the time of round 2 sample selections will continue to work at the time of the round 2 interview and respond to the round 3 survey.[7]

Table B.1. NBS–General Waves Sample Sizes by Strata

Sampling Strata	Sample Size	Target # of Completed Interviews
Round 2		
Active beneficiaries	5,000	4,000
Age range in years		
18 to 29	1,389	1,111
30 to 39	1,389	1,111
40 to 49	1,389	1,111
50 to FRA	833	667
Successful Workers	5,625	4,500
SSI (both with and without SSDI)	2,813	2,250
SSDI only	2,812	2,250
Round 3		
Active beneficiaries	5,000	4,000
Age range in years		
18 to 29	1,389	1,111
30 to 39	1,389	1,111
40 to 49	1,389	1,111
50 to FRA	833	667
Successful Workers	3,750	3,000
SSI (both with and without SSDI)	1,875	1,500
SSDI only	1,875	1,500
Longitudinal successful worker sample	2,812	2,250

The sampling design will include the selection of 80 primary sampling units (PSUs), along with selection of zip-code-based secondary sampling units

---

at the time of interview. Those that are currently working at round 2 will be included in the longitudinal follow up. We will use administrative data to track their work status between the round 2 and round 3 interviews.

[7] We assume that at least 62 percent of the 4,500 successful workers identified at round 2 will be currently employed at the time of the round 2 interview, resulting in a sample size of approximately 2,812 for the longitudinal follow-up. Mathematica is in the process of using SSA administrative data to verify these estimates. Mathematica is continuing work to verify this assumption.

(SSUs), within certainty PSUs. Specifically, we will select PSUs using a four-level composite size measure, incorporating the four age-based strata of the active national beneficiary sample. 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.

B.1.3. Response Rates

Our target response rate for both samples is 80 percent, however, based on the contractor's experience with the round 1 General Waves NBS, we expect this will be difficult to obtain. We recognize that it is becoming increasingly challenging to locate sample members (especially with electronic payments) and to 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 non-response bias analysis and take the results into account during weighting procedures.

B.2. Procedures for Collecting the Information

B.2.1. Statistical Methodology for Stratification and Sample Selection

We use the same multi-stage clustered design developed for the prior NBS 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 developed in the prior NBS, we used data from SSA on the counts of eligible SSDI beneficiaries and SSI recipients in each county to form 1,330 PSUs consisting of one or more counties. From this list, we selected a stratified national sample of 80 PSUs in the first round of the current NBS. As in the prior NBS, two of these 80 PSUs corresponded to Los Angeles and Cook (Chicago) counties, which we selected with certainty because of the number of SSA beneficiaries in these locations. Because of the size of these two counties (in both the beneficiary population and geographic size), we formed SSUs using beneficiaries' zip codes. Using the same set of SSUs created for the prior NBS, we selected four and two SSUs from the Los Angeles and Cook (Chicago) counties, respectively. In round 1, we selected 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 rounds 2 and 3, we will use the same set of PSUs and SSUs selected from round 1.

B.2.2. 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) 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-General Waves 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*, $_{hi}$, is $_{hi}$ = 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.

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

$$\pi_{hij} = n_{hi} \frac{\text{MOS}_{hij}}{\text{MOS}_{hi}}.$$

   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 nonresponse in the survey. 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 the logistic regression models are variables available for both respondents and nonrespondents, and are chosen 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 logistic propensity models will include the strata used in sampling. It is important that each level of the model covariates has a sufficient number of sample members to ensure a stable adjustment. As with prior rounds, the contractor 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 for most of the survey populations. 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 view propensity modeling as the extension of the standard weighting class procedure. We will use propensity modeling instead of the standard weighting class procedure because it allows us to use more factors (including both continuous and discrete 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 step-wise 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 we may 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 in suspense due to work versus people in current pay status).

After making adjustments for non-response, we will further adjust the weights so that some weighted sample statistics match known population values. For example, if the weights for recipients of SSI only, SSDI only, or both 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 most likely used are beneficiary type; state; age; sex; months since award; and primary impairment. We will investigate the feasibility of including earnings information in the weighting process.

In computing final weights, some individuals may end up with large weights. Variability in sampling weights can severely impact standard errors, particularly in the extreme case where one observation has a sampling weight that is orders of magnitude higher than other respondents. We will use “weight trimming” to alleviate this problem. In this procedure, the value of very large weights is simply reduced in magnitude, with the amount “trimmed” being distributed among other individuals 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. Since we will use propensity modeling instead of weighting classes to do the nonresponse adjustments, we will define trimming classes using the most important variables in the propensity models.

B.2.3. Standard Errors

For the NBS-General Waves, the sampling variance estimate is a function of the sampling 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 sampling 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 developed 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 such as 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. We will need sample design information (such as stratum and analysis weight) for each sample unit, to calculate variance.

B.2.4. Degree of Accuracy Needed

Active Beneficiaries - In Table B.2 (below) the minimal detectable difference for the active beneficiary strata is a measure of the smallest difference between subgroups that 4,000 completes will be able to detect with 80 percent power and 90 percent confidence. For example, for a proportion of 0.10, a minimal detectable difference equal to 6.7 percentage points indicates that if 10 percent of the beneficiaries are employed who never attended college, and at least 16.7 percent of the beneficiaries are employed who attended at least some college, the analysis will detect a significant difference between those who never attended college and those who attended at least some college. The table presents minimum detectable differences where we compare one half of the sample to the other half of the sample, and minimum detectable differences where we compare 70 percent of the sample to 30 percent.

Table B.2. Projected Minimal Detectable Differences Between Groups In Representative Beneficiary Sample

	Half the Sample Compared to Other Half (2,000 vs. 2,000)			70% of Sample Compared to 30% (2,800 vs. 1,200)		
	Mean of Binomial Distribution			Mean of Binomial Distribution		
Stratum	10%	30%	50%	10%	30%	50%
Overall (100 Percent)	5.9%	9.1%	9.9%	6.5%	9.9%	10.8%

Stratum	Half the Sample Compared to Other Half (2,000 vs. 2,000)			70% of Sample Compared to 30% (2,800 vs. 1,200)		
Age 18 to 29	4.8%	7.3%	8.0%	5.2%	8.0%	8.7%
Age 30 to 39	4.8%	7.3%	8.0%	5.2%	8.0%	8.7%
Age 40 to 49	4.8%	7.3%	8.0%	5.2%	8.0%	8.7%
Age 50 to 64	8.0%	12.1%	13.3%	8.7%	13.3%	14.5%

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

$$Var(\hat{p}) = P(1-P)*(1/n_{eff\,1} + 1/n_{eff\,2}),$$

where $n_{eff\,1}$ and $n_{eff\,2}$ are the effective sample sizes of the two populations being compared and $n_{eff\,1} = n_1 / deff_1$, $n_{eff\,2} = n_2 / deff_2$. We compute the design effect using the design effect due to unequal weighting and the design effect due to clustering, assuming 80 PSUs and an intracluster correlation of 0.02. The minimum detectable differences (using alpha = 0.10 and 80 percent power) are 2.49 square root (Var($\hat{p}$)).

Successful Workers - Of the 4,500 completed successful worker cases in round 2; 2,250 will be among SSI recipients who may or may not be concurrent beneficiaries of SSDI, who were working successfully as of the date six months prior to sample selection; and 2,250 will be among SSDI-only beneficiaries who were successful workers in that period. Because SSA is interested in differences between successful workers and beneficiaries who are back on the rolls as of data collection and those who are not, a comparison of interest might be a comparison between these groups. Given that approximately 30 percent of successful workers could be back on the rolls at the time of data collection, such a comparison might involve a comparison between 70 percent of the sample (3,150 successful workers) and 30 percent of the sample (1,350 successful workers back on benefits). A comparison within title (SSI, SSDI, or concurrent successful workers) would involve (for example) a comparison between 1,575 and 675 successful workers who were SSDI beneficiaries. Table B.3 (below) presents minimum detectable differences between the successful worker groups (sustained work vs. back on benefits), both for a comparison between two halves, and a comparison between 70 percent and 30 percent. We also present these comparisons within title (SSI, SSDI, and concurrent). These figures assume 60 percent of the SSI stratum is SSI only, and 40 percent is concurrent.

Table B.3. Projected Minimal Detectable Differences Between Successful Worker Groups

	Half the Sample Compared to Other Half (2,250 vs. 2,250)			70% of Sample Compared to 30% (3,150 vs. 1,350)		
	Mean of Binomial Distribution			Mean of Binomial Distribution		
Title	10%	30%	50%	10%	30%	50%
Overall (100 Percent)	3.5%	5.3%	5.8%	3.8%	5.8%	6.4%
SSI only	4.7%	7.1%	7.8%	5.1%	7.8%	8.5%
SSDI only	3.8%	5.8%	6.3%	4.1%	6.3%	6.9%
Concurrent	5.6%	8.5%	9.3%	6.1%	9.3%	10.2%

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

$$Var(\hat{p}) = P(1-P) \bullet (1/n_{eff\,1} + 1/n_{eff\,2}),$$

where $n_{eff\,1}$ and $n_{eff\,2}$ are the effective sample sizes of the two populations being compared and $n_{eff\,1} = n_1 / deff_1$ $n_{eff\,2} = n_2 / deff_2$. The design effect is computed using the design effect due to unequal weighting and the design effect due to clustering, assuming 80 PSUs and an intracluster correlation of 0.02. The minimum detectable differences (using alpha = 0.10 and 80 percent power) are 2.49 square root (Var($\hat{p}$)).

B.2.5. Unusual Problems Requiring Specialized Sampling Procedures

For the successful workers strata, there is a probability 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. If this occurs, we plan to use a hybrid design, which combines an unclustered stratified random sample with the clustered sample design. While both the unclustered and clustered samples are nationally representative, the data collection in the unclustered component is limited 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-frame” strategy, as we have in the prior NBS.

The result will be lower response rates for the non-PSU participants, 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 expect the telephone response rate will be higher for participants than for all beneficiaries, because we know that these are individuals who are not being prevented from at least attempting to work by their physical or mental conditions, or by other personal circumstances.

B.2.6. Periodic Cycles to Reduce Burden

We will administer the remaining rounds of the NBS-General Waves in 2017, and 2019. Beneficiaries in the RBS will complete the survey one time only (cross-sectional), with a new sample drawn before each round. Thus, there is no cyclic burden for these respondents. However, we will follow a subset of beneficiaries in the SWS longitudinally. We will follow successful workers who complete an interview at round 2, and in round 3, so we can better understand factors that positively or negatively affect the ability to sustain employment over time. To minimize burden, we will administer follow-up surveys biennially rather than annually, and will skip some items because they had previously been answered and are not prone to change. We will also use data from the previously completed surveys to serve as question fills to reduce respondents' cognitive burden of recalling previously reported employers and service providers.

B.3. Methods to Maximize Response Rates

B.3.1. Maximizing Response Rates

Locating sampled beneficiaries and participants is our first challenge to obtaining a high response rate. While SSA has contact information for all potential respondents, we know from past experience that it will often not lead directly to the SSDI beneficiary or SSI recipient. Telephone numbers could be particularly problematic because there is no administrative reason to keep them updated in SSA records. Addresses could be more reliable because they are sometimes used for mailing correspondence. These might, however, be a post office box, address of a guardian, financial institution, or other types of addresses that may make it difficult for us to locate the beneficiary or recipient. Since SSA now requires direct deposit of payment checks, we minimized the importance of keeping address information current.

To improve contact information, we will mail an advance letter written on SSA letterhead and a study brochure to each sampled person prior to the survey, using the address of record (either from SSA administrative data or provider record). This letter describes the survey and indicates that we will soon contact the SSDI beneficiary or SSI recipient. In round 3, we will tailor the advance letter for longitudinal cases, as appropriate. We will begin locating with letters returned to the contractor as undelivered. When an address is available without a phone number, we will conduct a directory search to obtain a number. When direct 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 payees). In round 1 of the NBS-General Waves, we located approximately 84 percent of SSDI beneficiaries and SSI recipients.

If a phone number is available or obtained, we will attempt to call the respondent to conduct the interview. The contractor 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.

In the first three months of data collection, we will send locating letters; reminder letters; reminder postcards; and refusal conversion letters, as appropriate (see Attachment B, Respondent Correspondence). In round 3, we will tailor correspondence for longitudinal cases, as appropriate. After two to three months of CATI interviewing, we will begin to transfer cases to field staff for locating. Delaying the start of field locating and interviewing allows an adequate number of cases to accumulate so field staff will have sufficient work and travel could be more cost-effective. Prior to deploying field staff, we will send all cases assigned to the field with a valid mailing address a prefilled letter, informing them that a representative from the contractor will be visiting their home (Attachment B). 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 our efforts and sets clear parameters for when a case should cease because of lack of response. We will train locators not to reveal any private information about the participant 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 the contractor’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.

The impairments and health of some SSDI beneficiaries and SSI recipients will make responding problematic, especially by telephone. To facilitate responses to the CATI interview, we will offer the use of several assistive devices (amplifier, Telecommunications Relay Service, instant messaging, and sign interpreters for in-person interviews) 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 necessary. Despite these efforts, we know that some respondents will be unable to complete the interview by telephone; 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 beneficiaries or recipients 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 mental 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 built into the prior NBS instrument. The test provides interviewers a tool for determining when to seek a proxy rather than leaving the decision to interviewer discretion or a gatekeeper. We will also develop and administer a Spanish-language version of the instrument administered by Spanish-speaking interviewers to Spanish-speaking subjects. We will use translation interpretation services for other non-English speakers.

SSA believes that some compensation is important to engender a positive attitude about the study and reduce attrition in follow-up interviews. If we are able to allocate additional funds to our incentive budget, we will provide respondents with a $30 gift card to compensate them for their time[8]. As we did in round 1, we may provide a pre-paid gift card of $5 in the final months of the data collection period to encourage call-ins from non-respondents with the promise of a $25 gift card after completion[9]. These steps should also encourage sampled individuals to cooperate with the interviewer once contacted.

Between rounds 2 and 3, the contractor will track the address and telephone number of those respondents that the contractor will attempt to re-contact in round 3 of the survey. Tracking methods may include database searches, interim mailings, and interim phone calls to alternate contacts. We also added two questions to the revised instrument to aid with longitudinal locating—one was about the likelihood of moving within the next two years and one was about home ownership. Sample members who move (or have a higher propensity to relocate because they rent) will require more intensive tracking.

---

[8] If we are unable to allocate additional funds to our incentive budget, we will provide a $20 gift card to respondents as we did in Round 1.

[9] SSA will provide a $15 gift card to respondents once they complete the survey, for a total incentive payment of $20, if we are unable to allocate additional funds to our incentive budget.

To minimize burden for longitudinal respondents in round 3, we will administer follow-up surveys biennially rather than annually, and will skip some items because they had previously been answered and are not prone to change. We will also use data from the previously completed survey to serve as question fills so that respondents' cognitive burden of recalling previously reported employers and service providers is reduced.

B.3.2. 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. We expect such comparisons will be especially important to assess attrition bias in analyses of the follow-up survey for the longitudinal samples. We will also be able to use the administrative data to assess the extent to which nonresponse in the follow-up survey is due to mortality.

B.4. Tests of Procedures

We developed the original NBS survey and initially pre-tested as part of a separate contract held by Westat approximately ten years ago. For NBS–General Waves, we removed two sections that were specific to the TTW program. We made only minor modifications for round 1.

A technical working group informed the round 2 instrument content through a review of the literature, a review of existing instruments, and, by the results of 90 semi-structured interviews conducted in 2015 with beneficiaries who had periods of successful work where employment was sustained or not sustained over time. We submitted the instrument to two rounds of testing. We conducted cognitive interviews with a total of 18 beneficiaries; six each in three groups (successful workers in current benefit suspense because of earnings; successful workers with recent benefit suspense because of earnings, but back on benefits; and beneficiaries selected from the general NBS population). We pre-tested three paper-and-pencil protocols with a total of 36 beneficiaries sampled from the same three groups (10 in Spanish and 26 in English). The primary purposes of the cognitive and pretest interviews was to ensure that the new questions were clear and understandable to respondents; that they worked well in conjunction with the existing questions; and to assess the time needed to complete the questionnaire. We timed the pretests, and used them to figure our burden estimates, provided in Part A. We do not anticipate making substantive changes to the instrument between rounds 2 and 3.

B.5. Statistical Agency Contact for Statistical Information

Table B.4 below shows the individuals consulted on technical and statistical issues related to the data collection.

Table B.4. Individuals Consulted on Technical and Statistical Issues

Name	Affiliation/Address	Telephone Number
Elaine Gilby Paul O'Leary	Social Security Administration, Office of Research, Evaluation, and Statistics Washington, DC	(202) 358-6449 (202) 358-6227
Eric Grau	Mathematica Policy Research Washington, DC 20002 (Mathematica)	(609) 945-3330
Gina Livermore	Mathematica	(202) 264-3462
Frank Potter	Mathematica	(609) 936-2799
David Stapleton	Mathematica	(202) 484-4224
Debra Wright	Mathematica	(202) 554-7576
Kirsten Barrett	Mathematica	(202) 554-7564