Survey of SNAP and Work

OMB Supporting Statement 0584-NEW

Paperwork Reduction Act of 1995

Part B. Justification

Collections of Information Employing Statistical Methods

April 29, 2020

TABLE OF CONTENTS

B.1   **Respondent Universe and Sampling Methods**

Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

Respondent Universe

There are two respondent universes. The first universe includes State SNAP agency staff from each of the 50 states and the District of Columbia. We will request SNAP administrative caseload files from State agencies to use in the building of a sampling frame of SNAP participants. One week after OMB approval, we will send SNAP State agencies an email requesting the SNAP administrative data along with a fact sheet about the study. Two weeks after OMB approval, we will send an email reminder to agencies that do not respond to the initial email. We will send a second reminder email to agencies that do not respond within four weeks of OMB approval. When we receive the data, we will send a thank you email to the State SNAP agencies.

The second respondent universe is nondisabled SNAP recipients, ages 18 to 69. This population includes work registrants and able-bodied adults without dependents (ABAWDs), as well as individuals not required to register for work or participate in SNAP work programs because they are already working 30 hours or more - or are exempt for other reasons, such as receiving

unemployment insurance benefits, caring for young children or disabled family members, or age 60 or older. [1]

Sampling Methods

For the State SNAP agencies, no sampling will be used. We will contact the universe of the 50 states and the District of Columbia. For the respondent universe of nondisabled SNAP recipients, ages 18 to 69, we will use a two-stage stratified sample design to produce State-level estimates. A two-stage design provides efficiency by clustering sampled SNAP participants in smaller geographic areas for a more efficient in-person data collection effort. In the first stage, we will select primary sampling units (PSUs) in each State and the District of Columbia (DC) with probability proportional to the number of SNAP participants as measured by the American Community Survey (ACS). The Small Area Income and Poverty Estimates Program provides single year estimates for all U.S. states and counties. Model input data includes state SNAP benefits data. State-level files contain the number of SNAP recipients by month, and the county level files contain the number of recipients in July of each year. To ensure representation of rural areas in the PSU sample, we will employ a measure of urbanicity to form PSU strata with similar population density levels. The measure of urbanicity will be the percentage of the population living in urban areas according to data from the 2010 Census.[2] This will result in the selection of rural PSUs from rural strata at the appropriate representative proportion for each State.

---

[1] FNS considers SNAP participants ages 18 to 59 as "adults" or "working age" and are generally subject to work requirements. In contrast, FNS considers SNAP participants ages 60 to 69 as elderly and are not subject to work requirements.

[2] The U.S. Census Bureau defines urban areas as those with populations of 50,000 or more and urban clusters as those with populations of at least 2,500. Rural includes all areas that are not urban.
https://www.census.gov/geo/reference/urban-rural.html

At the second stage, we will sample nondisabled SNAP participants ages 18 to 69 in each PSU. SNAP participants include all SNAP participants living in SNAP households and eligible to receive SNAP benefits. We will select the sample in each State from State SNAP agency data on participants during a calendar month, referred to as the "sample month." The State agency data will include lists of all SNAP participants residing in sampled PSUs, including names, addresses, and telephone numbers, as well as demographic variables about individuals/households. We will sample age groups (18-49, 50-59, 60-64, 65-69) to proportionally represent nondisabled SNAP participants in each State, thereby ensuring adequate representation of age groups at the National-level for more detailed analysis than will be possible at the State-level. We will seek 780 completed interviews per State, and 39,780 interviews overall.

B.2　　　Procedures for the Collection of Information

Describe the procedures for the collection of information including:
- **Statistical methodology for stratification and sample selection,**
- **Estimation procedure,**
- **Degree of accuracy needed for the purpose described in the justification,**
- **Unusual problems requiring specialized sampling procedures, and**
- **Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

B.2.1　　　Statistical Methodology for Stratification and Sample Selection

The sample design for the Survey of SNAP and Work will provide for State-level estimates of employment among nondisabled, adult SNAP participants between ages 18 and 69. The State-level estimates will have a plus or minus 5 percentage point precision level. The design will also provide State-level estimates for subgroups, with reasonable precision. Precise estimates for subgroups are important to allow us to calculate and compare the descriptive characteristics of

interest. The State-level samples will proportionately represent nondisabled SNAP participant in terms of urbanicity and age to support National-level estimates. The National-level estimates will have high precision (plus or minus 1.3 percent), permitting the comparison of estimates among even relatively small subgroups in the population (with precision of plus or minus 3.8 percent).

We will sample PSUs with probability proportion to size (PPS) using data on SNAP participation from the American Community Survey (ACS). The Small Area Income and Poverty Estimates Program provides single year estimates from the ACS for all U.S. states and counties. Model input data includes state SNAP benefits data. State-level files contain the number of SNAP recipients by month, and the county level files contain the number of recipients in July of each year. The measure of size (MOS) used for the PPS selection will be as closely associated with the inference populations as possible. We will select with certainty large PSUs with relatively higher counts of SNAP participants (i.e. large MOS) and form smaller PSUs into non-certainty strata from which to randomly select one PSU. We will identify critical stratification variables to define sufficiently homogeneous strata and will be able to ensure the PSU sample is representative of critical characteristics, such as, degree of urbanicity. To ensure the appropriate level of representation of rural areas in the PSU sample, we will also form rural PSU strata by employing a measure of urbanicity to form PSU strata with similar population density levels. This will result in the selection of rural PSUs from rural strata at the appropriate representative proportion for each state.[3] We will not oversample rural PSUs but will ensure representative sample proportional to the population. On average, about 10 to 12 PSUs in each state will be needed, but a few states with very few counties we will likely need to define most if not all the

---

[3] We do not plan to use a tribal area as a sampling domain but tribal areas will be included in the sample if they are in a sampled PSU.

counties as certainty counties in order to ensure sufficiently reliable survey estimates for the State.

Upon receiving SNAP caseload data from the States, we will organize the sampling frame. The first step will be to subset the SNAP participants to those that are eligible for the Survey of SNAP and Work inference population—that is, non-disabled adults, age 18 to 69. To create samples that are well controlled and proportionally representative with respect to demographic and household characteristics, we will sort the frame of SNAP participants prior to sample selection—sorting cases by age (18 to 49, 50 to 59, 60 to 64, and 65 to 69) to obtain implicit stratification of the sample. Other candidate sorting variables include household size, gender, and race/ethnicity.

A critical issue may be locating SNAP participants and the associated costs of doing so. Sampled SNAP participants that do not respond through the web or telephone will be eligible for field followup. However, the cost of completing cases out in the field are substantially greater than that for web or telephone. Thus, we will select a 25-percent subsample of non-respondents to the web and telephone in each State for field location. We will select a subsample because the expected response rate to in-person data collection is higher than web or telephone. If the initial non-in-person response rate is lower than expected, we can increase the subsampling rate, and it can vary between States and by PSU within States to ensure a sufficient overall yield. We will apply appropriate weighting adjustments to subsampled field followup cases to ensure unbiased estimates.

We anticipate that there will be differences between States that could influence the sample design. For example, there may be variation in the proportion above age 60 that may necessitate a differential sampling rate for this group. Additionally, the set of variables available from each State could vary requiring some level of modification in the frame creation, explicit stratification formation to inform oversampling, or sorting of SNAP participants within the explicit strata.

B.2.2   Estimation Procedures

We will weight outcome measures derived from the surveys to adjust for differential probabilities of selection, in-person field data collection subsampling, and nonresponse. Weighting the samples will involve three steps. First, assign a base weight equal to the reciprocal of the probability of selecting the SNAP participant for the study to each sampled person. Specifically, the base weight for person i in stratum h in PSU g will be computed as $w_{ghi}^{base} = ¿¿/$ $n_{gh}¿/\pi_g$, where N_{gh} is the number of eligible SNAP participants in stratum h within PSU g, n_{gh} is the corresponding number of SNAP participants selected for the sample, and π_g is the probability of selection of the sample PSU. Second, apply an adjustment factor to those cases selected for in-person field data collection. The subsampling rate and thus the factor could vary by PSU. Specifically, the intermediate weight for person i in PSU g will be computed as $w_{gi}^{\int ¿¿} = S_g w_{gi}^{base}$, where S_g is the inverse of the subsampling rate applied to cases eligible for in-person field data collection in PSU g. Note that S_g is equal to 1 for all initial respondents not subject to the subsampling process.

Finally, since not all of the sampled persons will agree to participate in the study, we will apply a nonresponse adjustment to the base weights to obtain the final weight for analysis. Thus, the final weight for responding person i in a specified weighting cell k will be computed as $w_{ki}^{final} = R_k w_{ki}^{\int \dot{c}\dot{c}}$, where R_k is the inverse of the weighted response rate in weighting cell k using a CHAID (Chi Square Automatic Interaction Detector) analysis or similar algorithm to determine the weighting cells to use in the nonresponse adjustment. The CHAID algorithm is effective in identifying cells within which we expect response propensities to be similar.

In addition to the full-sample analytic weights, we will create a corresponding series of jackknife replicate weights and attach them to each data record for variance estimation purposes. Replication methods provide a relatively simple and robust approach to estimating sampling variances for complex survey data.[4] Under this replication approach, form up to 100 jackknife replicates by deleting selected cases from the full sample and adjusting the weights of the retained cases accordingly. Then apply the entire weighting process developed for the full sample separately to each replicate resulting in a series of replicate weights. The replicate weights can be imported into variance estimation software (e.g., SAS, SUDAAN, WESVAR) to calculate standard errors of the survey-based estimates and perform statistical tests that take account of the complex sample design.

B.2.3   Degree of Accuracy Needed

We require that precision for State-level estimates be a ±5 percent margin of error with a 95 percent level of confidence around a point estimate (proportion). The design effect is a factor to

---

[4] Rust, K.F. and Rao, J.N.K. (1996). Variance estimation for complex surveys using replication techniques. Statistical. Methods in Medical. Research, 5, 283-310.

be incorporated into the power and precision expected. There will be some clustering effect due to the two-stage design and an expected design effect due to differential sampling rates by subsampling for field location.

Table B-1 provides estimates of the number of respondents needed to achieve 95-pecent confidence intervals that are less than ±5 percentage points at the State-level for different prevalence estimates and design effects.[5] The table shows that 780 respondents per State would provide the needed precision for an estimate in the middle of the distribution (50 percent) with a design effect of 2.03 (covering 10 of the 12 table cells). Our plan is to obtain 39,780 responses (39,780 = 780*51).

With 780 responses per State, we will achieve ±7 percentage points of precision for key subgroups for State-level estimates. We believe that having precise estimates among subgroups at the State-level is important. We expect to conduct many of the analyses on subgroups within each State, especially those defined by employment status. For example, the analysis of hourly wages or hours worked will be subset to employed participants, and the analysis of reasons for not working will be subset to unemployed participants. Reasonable precision among subgroups in each State will provide States with information that is more actionable.

---

[5] This assumes initial State-level design effects of 1.05, 1.10, 1.15 and 1.20 based on previous studies conducted with PSU clustering within States. Each design effect is multiplied by a factor to account for differential sampling rates for subsampling for field nonresponse to arrive at the estimated design effects.

Table B-1. Required number of responses for 95-percent confidence intervals less than plus or minus 5 percentage points at the State-level for selected prevalence and design effects

Proportion	State-level design effect			
	1.94	**2.03**	**2.13**	**2.22**
0.1	270	280	290	310
0.3	630	660	690	720
0.5	750	780	820	850

We will target 780 completed surveys in each State, and 39,780 completed surveys overall. Based on recent experiences of surveys of SNAP participants, we expect to achieve 60 percent response rate. Following American Association of Public Opinion Research (APPOR) guidelines, we will use a weighted response rate that "weights up" the subsampled group and sets the weight for the non-subsampled cases equal to zero because we did not complete the data collection effort for this group.[6] This is the traditional calculation of the response rate for surveys that include nonresponse subsampling, including the ACS. The weighted response rate provides a valid estimate of the percentage of responses where no response occur even after conducting all three modes.

Based on experience on similar studies, we expect a 20 percent response rate to web and 20 percent response rate to computer-assisted telephone interviewing. We will select a 25 percent subsample of non-respondents to the web and telephone for field location and expect a 33 percent response rate to this effort. To determine the number of sampled cases needed, it is important to define the yield rate, which takes into consideration the cases that are not

---

[6] American Association for Public Opinion Research (AAPOR). (2016). *Standard Definitions: Final Disposition Codes and Outcome rates for Surveys.* 9th edition. Lenexa, Kansas: AAPOR.

subsampled for non-response follow up. The yield rate is calculated as $YR = IRP + ((1\text{-}IRP) * SP * RR$, where IRP is the initial response proportion, or the proportion responding to web and telephone, SP is the subsampled proportion of nonrespondents, and RR is the response rate to the field location. The yield rate is estimated to be $.45 = (.40 + (1\text{-}.40)*.25*.33)$. We plan to sample about 88,400 (39,780/.45) SNAP participants. The weighted response rate is calculated as $RR = IRP + (SP * RR)$, or $0.40 + (.25*.33) = 0.6$.

B.2.4   Unusual Problems Requiring Specialized Sampling Procedures

There are no unusual problems that require specialized sampling procedures.

B.2.5   Any use of periodic (less frequent than annual) data collection cycles to reduce burden.

This is a one-time study; concern regarding the periodicity of data collection cycles is not applicable.

B.3   Methods to Maximize Response Rates and to Deal With Issues of Nonresponse

Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

B.3.1   Methods to Maximize Response Rates

The survey will use a sequential multi-mode data collection approach, to obtain the highest possible response rates while minimizing data collection costs. We will mail an invitation letter to study participants with valid mailing addresses and ask them to complete the web survey. This

web invitation will include a phone number for respondents to call into the Telephone Research Center (TRC) if, for example, they do not have easy access to a computer or the Internet. The TRC will contact those who do not complete the survey online (and those with bad addresses) by telephone to complete a Computer Assisted Telephone Interview (CATI). For the subsample of non-responders to these two modes, field data collectors will locate respondents and direct them to the CATI or web to complete the survey. Based on the expected response rates, we anticipate that among the 780 completed interviews per state, approximately 347 will be completed by web or inbound CATI, 347 will be completed by outbound CATI, and 86 will be completed by web or telephone following in-person recruitment efforts. We will assign a unique study identifier to all sample members to track them across survey modes to ensure that respondents complete the survey only once.

Using the mailing address contained in the SNAP case files, we will mail an invitation letter explaining the study and inviting the sampled SNAP participant to complete the survey. We request permission to include a $2 cash pre-incentive with the invitation letter. Previous studies, including the Farmers' Market Client Survey incentive experiment, have shown that a small non-contingent incentive has a positive effect on response rates.[7] The invitation letter will offer a $20 cash incentive for completing the survey and an additional $20 if completed within the first four weeks of the sample release. One week after the initial mailing, we will send a postcard and e-mail (when an email address is available) to sampled persons, reminding them to complete the survey. We will send a second reminder postcard after two weeks. See Table B-2.

---

[7] Mercer, A, Caporaso, A., Cantor, D. and Townsend R; How Much Gets You How Much? Monetary Incentives and Response Rates in Household Surveys. *Public Opinion Quarterly,* Spring 2015 vol. 79, no. 1 102-129.

Approximately four weeks after the initial mailing, we will contact those participants that have

not responded (and have a telephone number) to complete a CATI interview. Messages left on

answering machines will provide a toll-free 800 number so respondents can call if they want to

do the survey at their convenience.

Approximately eight weeks after the initial mailing, for a sample (25 percent) of those

participants that have not responded to either the web invitation or telephone follow-up, we will

send a data collector to their homes for in-person recruitment effort, directing them to a CATI

interview or complete the survey online.

Table B-2. Timeline for data collection events weeks after OMB approval

Time period	Activity
Week 1-2	Send invitation letter (1st sample release)
Week 2	Send first reminder post card
Week 3	Send second reminder post card
Week 4	Send invitation letter (2nd sample release)
Week 5	Begin computer-assisted telephone interviews
Week 5	Send first reminder post card
Week 6	Send second reminder post card
Week 8	Send invitation letter (3rd sample release)
Week 8	Begin field outreach to non-respondents
Week 9	Send first reminder post card
Week 10	Send second reminder post card
Week 14	Complete field work on 1st sample release

Time period	Activity
Week 18	Complete field work on 2nd sample release
Week 22	Complete field work on 3rd sample release

We anticipate that some sample members will move out of the primary sampling unit (PSU)**.** If our field location efforts discover that a sample member moved out of the PSU and into a PSU that is not in the sample, we will first attempt to interview the sample member by telephone. For sample members who moved out of the PSU and we cannot reach by telephone, we will set a maximum distance that the data collector has to travel to interview the sample member. In this case, we will count sample members who move more than the maximum distance outside of a sampled PSU as nonrespondents and we will adjust for this type of nonresponse for using nonresponse weighting.

Imputation is widely used to fill in item nonresponse in survey data sets because it makes analysis more consistent from one user to the next, it avoids large biases in estimates of totals, and it can reduce nonresponse bias in other types of estimates in some circumstances. Eliminating missing data by imputation also allows analysts to include all cases when conducting multivariate analyses, which is particularly helpful if the alternative is to drop all cases that have missing values for one or more of the variables involved in the analysis.

We will not impute missing data on outcome variables. However, we will impute key covariates if there are significant missing data. The general approach involved in imputation is to use the information available for a record to assign values for the record's missing items. We will flag

any imputed values so that the user of the data knows which values are original respondent values and which are imputed values.

B.3.2 Nonresponse Bias Analysis

Although we will make efforts to achieve as high a response rate as practicable with the available resources, nontrivial nonresponse losses are likely to occur. We will conduct a nonresponse bias analysis to assess the impact of nonresponse on the survey estimates and the effectiveness of the weight adjustments to dampen potential nonresponse biases. We will use variables that are available for both respondents and nonrespondents.

The types of analyses we will conduct to evaluate nonresponse will include:

- Evaluating differences found between survey respondents and nonrespondents using state SNAP participant data;

- Comparing weighted estimates of characteristics available for both respondents and nonrespondents using unadjusted (base) weights versus nonresponse-adjusted weights;

- Comparing characteristics of respondents providing completed data at different levels of data collection effort (e.g., cases completed with limited follow-up compared to those requiring field follow-up).

B.4 Test of Procedures or Methods to be Undertaken

Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or

more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.

The web and CATI versions of the survey instrument were pretested with SNAP participants in English and Spanish under FNS Generic Clearance for Pre-Testing, Pilot, and Field Test Studies, OMB Control Number 0584-0606 (expiration 03/31/2019). Pretesting began on March 25, 2019 and ended on June 6, 2019.

Specific pretest objectives included identifying problems related to communicating intent or meaning of questions and concepts; determining whether respondents could accurately provide the information requested; and assessing the adequacy of the range of responses. In total, we completed 30 interviews. We revised the draft instruments following the pretests, including:

- Simplifying and standardizing response options, and reducing the number of responses offered in long lists;

- Providing revised Spanish translations for greater clarity and understanding;

- Revising phrasing of questions and response options for greater consistency across questions on the same topics;

- Providing definitions for some key terms; and

- Including a few additional response options.

B.5   Individuals Consulted on Statistical Methods and Individuals Responsible for Collecting and/or Analyzing the Data

Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Table B-3 presents a summary of individuals consulted on statistical aspects of the design. Westat staff will be responsible for the collection and analysis of the study's data, in coordination with FNS.

Table B-3. Individuals Consulted on Data Collection or Analysis

Staff	Title	Contact Information (phone or email)
Contractor—Westat		
David Hubble	Associate director - senior statistician	301-610-8814
Jennifer Kali	Senior statistician	301-738-3588
Robert Fay	Senior statistician	240-314-2318
Lloyd Hicks	Senior statistician	301-610-4960
Joseph Gasper	Associate Director	240-314-2470
Frank Bennici	Senior Study Director	301-738-3608
Maeve Gearing	Senior Study Director	301-212-2168
Karen Stein	Senior Study Director	301-610-8869
Peer Advisory Panel		
Benjamin Reist	Assistant Center Chief for Research, Center for Adaptive Design in Research and Methodology Directorate at U.S. Census Bureau	Benjamin.M.Reist@census.gov
Dottie Rosenbaum	Senior Fellow, Center on Budget and Policy Priorities	rosenbaum@cbpp.org
Yvette Chocolaad	Policy Director, National Association of State Workforce Agencies	ychocolaad@naswa.org
Diane Whitmore Schanzenbach	Director, Institute for Policy Research at Northwestern University;0 research associate at the National Bureau of Economic Research	dws@northwestern.edu

Staff	Title	Contact Information (phone or email)
Laura Tiehen	Economist, Food Assistance Branch of the Economic Research Service's Food Economics Division	LTIEHEN@ers.usda.gov
FNS Staff		
Michael Burke	Social Science Research Analyst	michael.burke@fns.usda.gov
NASS Staff		
Lindsay Drunasky	Mathematical Statistician	Lindsay.drunasky@usda.gov