

PART B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

B.1. 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.

B.1.1. Respondent Universe

The respondent universe for the proposed study will include the following:

- 1) **SNAP authorized and redeem benefits.** SNAP authorized farmers markets and direct marketing farmers that redeemed benefits during the last 12 months;
- 2) **SNAP authorized but don't redeem benefits.** SNAP authorized farmers markets and direct marketing farmers that did not redeem benefits during the last 12 months;
- 3) **SNAP authorized in the past.** Farmers markets and direct marketing farmers that were SNAP-authorized sometime during the last 5 years, but not authorized during the last 12 months; and
- 4) **Never SNAP authorized.** Farmers markets that have never been authorized to redeem SNAP benefits.
- 5) **SNAP recipients.** SNAP recipients (individuals or households).

Details on the respondent universe for each of the above groups are provided below in the context of sampling methods.

B.1.2. Sampling Methods

National Survey of Farmers Markets (FM) Managers and Direct Marketing (DM) Farmers

The sampling frame for the survey will be farmers market and direct marketing farmer contacts identified from two sources: the FNS Store Tracking and Redemption System II (STARS II) and the latest Agricultural Marketing Service (AMS) Farmers Market Directory. The unit of analysis will be multi-stall markets and direct marketing farmers included in these data files.

The STARS II data file captures detailed information on all farmers markets that have ever been authorized to participate in the Supplemental Nutrition Assistance Program (SNAP) and includes names and contact information for each market. “Farmers markets” on STARS II are defined as multi-stall markets, direct marketing farmers, and umbrella organizations. Umbrella organizations represent multi-stall markets in more than one location; these organizations will be identified and disaggregated to their individual market locations to allow for an equal probability of the individual market locations being sampled. The first three groups (current or ever SNAP authorized) will be randomly selected from the STARS II sampling frame. An

extract of the STARS database indicated that about 60% of the list includes farmers markets and the remaining 40% are direct marketing farmers.

The AMS directory includes names and contact information for farmers markets (SNAP and non-SNAP) located in the United States, and provides a limited set of additional information regarding the market. It does not contain information on direct marketing farmers. The **Never SNAP authorized** group will be randomly selected from the AMS directory. We will compare the STARS II and AMS data files using an auto match program based on market name and address, and will remove all duplicates from the AMS directory to identify those markets that have *never been SNAP authorized*. Market information included in the AMS directory is voluntary and self-reported to AMS from market managers, state market representatives, state associations, and consumers, and as such, is known to be incomplete. The combination of SNAP and AMS files are *believed* to account for about 85 percent of all farmers markets, but the true degree of undercoverage is not known with any certainty. All contacts identified on the STARS II and AMS directory that represent farmers markets and direct marketing farmers are eligible to participate in the study.

Substudy to Examine Undercoverage of Farmers Markets

This substudy involves the identification and survey of a sample of markets that are not to be found on the AMS data file, specifically those that have never been SNAP authorized. There will be one, or possibly two stages

of sampling, depending on how many such markets are identified. The frame for the first stage of sampling will be a complete list of the 3,143 U.S. counties. County is a convenient geographic unit for the type of data gathering involved in the substudy. We plan to sample about 20% (n=630) of the counties across the U.S. and employ standard tracing procedures to identify markets that are missing from the AMS file. We will identify new markets by (1) contacting farmers markets on the AMS directory to ask if they know of any other markets in their county; (2) reviewing the list with the Farmers Market Coalition; and (3) contacting the county extension service and other state and local agricultural agencies. The resulting list of farmers markets will be compared to the AMS directory in order to generate a list of markets that are not on the current AMS directory. This list will then further be reduced by consulting the STARS database and eliminating those markets that have ever been SNAP authorized. At that point we will take stock of how many newly identified markets remain in the list overall. It is not possible to predict with any real confidence the number of markets that will be identified during the substudy frame building activities. Our plan is if 125 markets or fewer are identified, all will be sent the farmers market manager survey using the same procedures as for the national survey. If the number exceeds 125 we will subsample down to 125.

It is of note that the AMS directory does not include DM farmers. We will not identify the DM farmer's undercoverage in this sub-study – for two reasons – first, there is no public-use data file of DM farmers that can be

used as a sampling frame. Second, the high level of effort involved in screening information on all farms (there are over 2 million farms in the US) to identify those that market directly to consumers.

Focus Group with SNAP Recipients

We will conduct four in-person focus group sessions with English-speaking SNAP recipients. We will recruit 40 SNAP participants (10 per group) with the expectation that about 32 will participate (8 per group) in the discussions (the remaining may be no-shows or cancellations). **Appendix F1-F8** presents the procedures for recruiting participants and conducting the focus groups.

Recruitment Procedures

We will select two markets for the focus group study: the 32nd Street/Waverly Farmers' Market in Baltimore and the Ward 8 Farmers' Market in Washington, DC. These markets were chosen because they are located in high-poverty neighborhoods and both have EBT programs in place. In addition, we will reserve meeting rooms for the focus groups in community locations at or near each market. In Baltimore, the groups will be held at the Waverly Branch Public Library (approximately one block from the Market); and in Washington, we will reserve the community meeting room at THEARC, a community theater and activity center who parking lot actually hosts the Ward 8 market. Both locations can be reached easily via several public transportation routes, and both are wheelchair accessible. At each location,

we will reserve space for one weeknight group and one Saturday midday group.

Focus group participants will be recruited via flyers (Appendix F7) posted in an array of community locations within 2-3 miles of each market. In addition to posting flyers at the markets themselves, posting locations will include bus stop kiosks, local convenience stores, service stations, and the facilities where the focus groups will be conducted. Flyers are worded generally (“focus groups about food-shopping”) so that we attract individuals who may shop in the area, but do not make purchases at the farmers’ market. Interested individuals can call a toll-free number at Westat, where our recruiter will screen for persons who have shopped at the market (“shoppers”) and those who are at least aware of the market, but who have not made any purchases there (“non-shoppers”) (see attached screening call script, beginning on page 3t). We intend to recruit a diverse set of participants for each group, that is, individuals who represent both ethnic and racial diversity, both men and women, as well as participants who have varying levels of education. We will recruit up to 12 individuals for each focus group, recognizing that there will be some attrition among the invitees. Our aim is to have 8-10 participants in each discussion. In an effort to reduce attrition, a confirmation letter (Appendix F8) will be mailed to individuals within one day of their agreeing to participate in a discussion. Our recruiter will also contact each participant by phone one day before the scheduled focus group.

B.1.3. Response Rates and Non-Response Analysis

Response Rates for the National Survey of Farmers Market Managers and Direct Marketing Farmers

Table B.1 below presents sample sizes for the four strata included in the FM Manager and DM Farmer survey. Sample sizes have been set so that after nonresponse, 95 percent confidence intervals for estimated proportions do not extend beyond plus or minus three percentage points for each of the four strata. The initial sample sizes assume an 80% response rate. Note that the requirement of equal precision across strata means that markets in smaller strata are sampled at a higher rate. Major subgroups within strata will have somewhat lower precision. For example, in strata 1 and 2 for the FM Manager survey we would expect the 95 percent confidence intervals for estimated proportions to be within plus or minus 4 percent; for the DM Farmer survey, within plus or minus 5 percent.

Table B.1. Sample Sizes for Precision Estimates

Sampling stratum (cohort)	Sample frame	Initial sample size	Responding farmers markets
1*	1208	709	567
2*	454	398	319
3**	415	374	299
4**	4,188	1,063	851
Total	6,265	2,544	2,036

* Includes FM (for-profit and not-for-profit) and DM farmers (for profit and not for profit)

** Includes FM (for-profit and not-for-profit), does not include DM farmers

To best ensure a representative sample the frame will be sorted within each stratum prior to sampling. In strata 1 and 2 the highest level of sort variable will be survey type; i.e. FM manager vs. DM farmer. In all four

strata, we will also sort by geography (region), type of locale using the USDA Rural-Urban Continuum Code that classifies a county based on population,²¹ and market size (e.g., SNAP transactions, number of vendors). Systematic equal probability sampling with a random start will be employed. Sample weights will be developed for responding farmers markets. These weights will permit inference to a population of farmers markets listed in the STARS and AMS databases. In the case of strata 1, 2, and 3, coverage bias will not be an issue as the STARS database contains the entire population of markets ever authorized to redeem SNAP benefits. The coverage for stratum 4 will only be as good as the AMS database's, which is known to be less than 100 percent. However, as discussed above, we plan to conduct a sub-study to examine the undercoverage of farmers markets on the AMS database.

Weights will be computed in such a way as to compensate for variable probabilities of selection and differential response rates. The weighting procedure will consist of the calculation of the "base weight" and a nonresponse adjustment procedure. The base weight is the reciprocal of the probability of selecting a unit for the sample. With a stratified sample of farmers markets in which markets are selected at constant rates within strata, and each farmers market manager has one and only one chance of selection, the selection probability for the i-th sample farmers market in sampling stratum h is simply:

$$p_h = \frac{n_h}{N_h}$$

²¹<http://www.ers.usda.gov/Data/RuralUrbanContinuumCodes>

Where N_h is the total number of farmers markets in the frame in stratum h ; and n_h is the corresponding number of sample farmers markets.

New Farmers Markets. We have assumed a 75 percent response rate for the farmers market managers at the new farmers markets that are identified through the undercoverage substudy. We plan to use standard procedures to achieve this response rate, similar to those used in the national survey.

Analysis for Non-response

To compensate for unit nonresponse, a standard approach is to calculate adjustment factors within selected weighting classes, and then to use these factors to inflate the base weights. The weighting classes may be the same as the original sampling strata, or they may be defined on the basis of other relevant characteristics that are available for both responding and nonresponding units. The sum of the weights of the respondents after the adjustment will equal the sum of the weights of the respondents and nonrespondents before the adjustment.

Replication will be used to calculate sampling errors of survey estimates and to conduct statistical significance tests of survey findings. To implement a replication method, a series of weights are attached to each record in the data file: a full sample weight and a corresponding set of

replicate-specific weights. Replication methods work by dividing the sample into subsample replicates that mirror the design of the survey. To obtain the replicate weights, the complete weighting process applied to the full sample is also applied to each replicate. Each recalculation generates a different set of weight adjustment factors, and these in turn are used to produce the required replicate-specific weights. The survey estimate is calculated for each replicate and variation among the subsample replicates is then used to estimate the variance for the full sample estimate.

B.2. 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

As noted above, we will stratify the sample of farmers markets. The use of explicit stratification will allow us to fix sample sizes at levels designed to achieve the desired precision for each of the four groups of interest. Within each stratum we will implicitly stratify by census region, type of locale and

market size (for strata where a size measure is available). This will be accomplished by sorting the frame by these variables and selecting the samples using systematic equal-probability sampling with a random start. Implicit stratification will ensure good representation of markets across the sort variables and is likely to improve the precision of survey estimates somewhat.

B.2.2. Estimation Procedures

When a survey is conducted using a complex sample design, the design must be taken explicitly into account to produce unbiased estimates and standard errors for these estimates. This is accomplished by dividing the complete sample into a number of subsamples known as replicates so that each replicate sample, when properly weighted, will provide appropriate estimates of population characteristics of interest. In general, replicate samples are formed to mirror the original sampling of primary sampling units.

Replication will be used in this study to calculate sampling errors of survey estimates and to conduct statistical significance tests of survey findings. WesVar Software for Complex Survey Analysis will be used with replicate weights to take the sample design into account when calculating point estimates, correlation and regression coefficients and their associated standard errors. WesVar uses a set of orthogonal Hadamard matrices to

create balanced half-samples for use with balanced repeated replication (BRR) procedures (Morganstein, 1998).

B.2.3. Degree of Accuracy Needed for the Purpose Described in the Justification

The sample has been designed to estimate response proportions in each of the four sample strata within a three-percentage point margin of error. Here we address the minimally detectable difference considered to be statistically significant when response proportions reported from two different strata are compared.

The respondent universe is represented by the sampling frame from which each of the four strata of farmers markets have been selected. Sampling fractions for sample stratum h yield selection probabilities $p_h = n_h / N_h$, where n_h represents the number of farmers markets selected for interviews and N_h represents the total number of such units in the sampling frame. Sampling proportions to be used in the present study are large, implying that a finite population correcting is needed when standard errors are estimated. In this case, the correction for stratum h is $fpc_h = 1 - p_h = 1 - (n_h / N_h)$.

Minimum Detectable Differences (MDDs)

MDD for between stratum comparisons of proportions are reported in Table B.2.3.1, together with their respective stratum sample sizes n_h ,

selection probabilities p_h and finite population corrections fpc_h . Above the diagonal at right, MDD are reported for between stratum comparisons involving a response proportion in the neighborhood of $p_0 = 0.2$. The actual values of these MDD are somewhat different depending on the strata in question, but these are all equal to 0.048 to three decimal places. Below the diagonal at left, MDD are reported for response proportions around $p_0 = 0.5$. Although there are also differences in these MDD values, all values are close to 0.061.

Table B.2.3.1. Sample Sizes, Selection Probabilities and Finite Population Corrections for Minimally Detectable Differences at $p_0 = 0.2$ and $p_0 = 0.5$

Stratum m	n_h	p_h	fpc_h	Stratum			
				1	2	3	4
1	567	0.469	0.531	--	0.048	0.048	0.048
2	319	0.703	0.297	0.061	--	0.048	0.048
3	299	0.720	0.280	0.061	0.061	--	0.048
4	851	0.203	0.797	0.061	0.061	0.061	--

MDD_h at $p_0 = 0.2$ above the diagonal. MDD_h at $p_0 = 0.5$ below.

These values show for example that for $p_0 = 0.2$ observed in one stratum, a value of $p_1 = 0.248$ or larger observed in a second stratum would be considered statistically significant. Alternatively, for a value of $p_0 = 0.5$ in one stratum, a value of 0.561 or larger a second stratum would be considered statistically significant. This provides a range for MDD that are likely to be encountered in this study. These values have been calculated with the usual assumptions that power = 0.8 and $\alpha = 0.05$.

Estimation and Calculation of Sampling Errors

A series of jackknife replicate weights will be created and attached to each data record for variance estimation purposes. In order to have stable variance estimates for subgroup analyses, 100 jackknife replicates will be formed by deleting selected cases from the full sample and adjusting the base weights of the retained cases accordingly. The weighting process developed for the full sample will be applied separately to each jackknife replicate resulting in a series of 100 replicate weights. These replicate weights will then be imported into WesVar to calculate appropriate standard errors for survey-based estimates.

B.2.4. Unusual Problems Requiring Specialized Sampling Procedures

No specialized sampling procedures are involved.

B.2.5. Any use of Periodic (less frequent than annual) Data Collection Cycles to Reduce Burden

The study design requires a one-time data collection from respondents. All data collection activities will occur with 6 month period.

B.3. 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.

Overall response rate projections were presented earlier. By carefully and convincingly explaining the importance and potential usefulness of the study findings in the introductory letters from FNS, and by implementing a series of follow-up reminders with a final attempt to complete the survey by telephone, we expect to achieve an overall survey response rate of 80% on the national survey and 75% response rate for the undercoverage substudy. Specific procedures to maximize response rates include:

- A cover letter from USDA/FNS (**Appendix A1**).
- \$5 prepaid incentive included with introductory letter and survey
- Two reminder mail and email to respondents who have not completed the survey after one week and two weeks of the original survey mailing (**Appendix D1-D2**).
- Two data collection modes (mail or web) for participants' convenience
- Telephone follow-up interview for non-responders (begins 4 weeks after receiving survey) (**Appendix E1**).
- Make up to 9 unsuccessful call attempts to a number without reaching someone before considering whether to treat the case as “unable to contact.”

- Implement refusal conversion efforts for first-time refusals and use interviewers who are skilled at refusal conversion and will not unduly pressure the respondent (**Appendix E2**).
- Provide a toll-free number for respondents to call to verify the study's legitimacy or to ask other questions about the study.
- Implement standardized training for telephone data collectors. The interviewer training will focus on basic skills of telephone interviewing, use of CATI platforms for interviews.

B.4. Describe any Test 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.

In-depth interviews were conducted with nine farmers market managers and site visits were conducted at five farmers markets to inform the development of the survey instrument. Interview findings identified the most important constructs and close-ended response categories for developing a quantitative survey. The resulting quantitative survey instrument was cognitively tested with 8 FM managers and 1 DM farmer to ensure participant understanding of survey questions as intended, further refine question wording, and identify additional response categories. The results of cognitive testing indicated that a separate survey should be developed for

DM farmers. Two new survey instruments were created from the original instrument – one for FM manager and a second for DM farmers. The new FM instrument was cognitively tested with 2 FM managers and the new DM instrument was tested with 3 DM farmers. The final instruments were revised to reflect the comments from the cognitive testing.

B.5. 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.

Name	Affiliation	Telephone Number	e-mail
John Burke	Senior Study Director, Westat	301-294-2057	JohnBurke@westat.com
Susie McNutt	Project Director, Westat	301-251-3554	SusieMcNutt@westat.com
Sujata Dixit-Joshi	Senior Study Director, Westat	508-293-4029	SujataDixit-Joshi@westat.com
Cynthia Robins	Senior Study Director, Westat	301-738-5424	CynthiaRobins@westat.com