

Contract Number:
AG-3K06-D-09-0212/GS-10F-0050L

Mathematica Reference Number:
06687.070

Submitted to:
USDA Economic Research Service
1800 M Street
Washington, DC 20036
Project Officer: Mark Denbaly

Submitted by:
Mathematica Policy Research
955 Massachusetts Avenue
Suite 801
Cambridge, MA 02139
Telephone: (617) 491-7900
Facsimile: (617) 491-8044
Project Director: Nancy Cole

**Supporting Justification for OMB
Clearance of the Field Test for
the National Household Food
Acquisition and Purchase Survey**

**Part B: Statistical Methods for
Baseline Data Collection**

September 23, 2010

CONTENTS

B. Statistical Methods for Baseline Data Collection	1
B1. Respondent Universe and Sampling Methods	1
B2. Procedures for Collection of Information	5
B3. Methods to Maximize Response Rates and Deal With Nonresponse	9
B4. Tests of Procedures or Methods to be Undertaken	12
B5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data	13

B. Statistical Methods for Baseline Data Collection

B1. Respondent Universe and Sampling Methods

The respondent universe for the field test includes two groups for analysis: (1) households¹ receiving SNAP benefits; and (2) households not receiving SNAP benefits, in which household income does not exceed 185 percent of the Federal poverty guidelines. Non-SNAP households will be selected from a stratified sampling frame so that, for sampling, three groups will be selected: (1) households receiving SNAP benefits; (2) households not receiving SNAP benefits whose income is below the Federal poverty guidelines, and (3) households not receiving SNAP benefits whose income is between the poverty guideline and 185 percent of that guideline. The three strata were defined to provide a sample of SNAP participants for analysis, and a sample of low-income households who are potentially eligible for USDA food assistance programs.² We will use the 2009 poverty guidelines unless they are updated before field test instruments are finalized; in that case, we will incorporate the 2010 poverty guidelines. We expect a response rate of 71 percent for SNAP households and 55 percent for other households calculated using the American Association for Public Opinion Research formula 4 (AAPOR 2008). (The response rate is addressed in more detail in section B.3.) The field test will be conducted in two purposively selected sites in the mid-Atlantic region within proximity to Mathematica's survey operations center. Because the sites have not yet been determined we do not know their population figures.

Nationally, the estimated number of households in each of the study population groups³ are

- 11.1 million SNAP households;
- 2.0 million non-SNAP households with income below the federal poverty guidelines; and
- 9.1 million non-SNAP households with income between 100 and 185 percent of the federal poverty guidelines.

The design for the full-scale survey, described in more detail later in this section, calls for a multistage design with 50 primary sampling units (PSUs) selected at the first stage. Eight secondary sampling units (SSUs) will be selected within each PSU at the second stage for a total of 400 SSUs. Addresses will be selected at the third stage of sampling within each SSU.

¹ The SNAP program definition of household: Everyone who lives together and purchases and prepares meals together is grouped together as one household. However, if a person is 60 years of age or older and he or she is unable to purchase and prepare meals separately because of a permanent disability, the person and the person's spouse may be a separate household if the others they live with do not have very much income. (More than 165 percent of the poverty level.) Some people who live together, such as husbands and wives and most children under age 22, are included in the same household, even if they purchase and prepare meals separately.) Source: www.fns.usda.gov/snap/applicant_recipients/Eligibility.htm

² The income eligibility limits for the largest food assistance programs are: 185 percent of poverty for the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC); 130 and 185 percent of poverty, respectively for free and reduced-price meals in the National School Lunch and School Breakfast Programs; and 130 percent of poverty in the Supplemental Nutrition Assistance Program (SNAP).

³ Based on 2008 current population estimates of families in poverty and USDA estimates of participation rates (Leftin and Wolkowitz, 2009).

The field test will purposively select two PSUs and 8 SSUs within each PSU, with a goal of obtaining low-income areas with a variety of retail food environments. Sampling of addresses at the third stage of selection will follow the same procedures planned for the full-scale survey.

In addition to its other purposes, the field test will provide information that will be used to evaluate and, if needed, revise the sampling procedures at the third stage of sampling. For the field test, our target sample size is 400 completed cases, divided among the groups of interest:

- 200 SNAP households;
- 120 non-SNAP households with income between poverty and 185 percent of poverty (low income); and
- 80 non-SNAP households with income less than poverty (very low income).

These sample sizes will allow adequate precision for estimates of response rates, household burden, and data quality, both overall and for the SNAP and non-SNAP groups. Expected design effects and power for statistical tests are discussed in Section B.2. For estimates of response rates, we expect 95 percent confidence intervals of between ± 5.7 and ± 6.2 percentage points for the sample as a whole (depending on the extent of the design effect) and of no more than ± 8.1 percentage points for each of the groups of SNAP and non-SNAP households. This will also enable us to evaluate sampling and data collection procedures for the two groups. The sample of 80 very low-income non-SNAP households will not provide precise estimates for this group but can inform us of large differences between this group and others.

The field test will also employ a randomized trial to evaluate response rates and data quality associated with two incentive levels and two versions of food booklets for collecting food acquisition data. This two-by-two test requires assignment of households to one of four treatment cells. Each case will be assigned to a treatment cell so that the distribution of cases within each cell reflects the distribution of the overall field test sample. The planned allocation will enable us to make precise estimates of overall treatment impacts and to see if there are large differences between SNAP and the non-SNAP households included in the field test.

For the field test, eight SSUs will be selected in each of two sites. Each site will be defined in the same way as we define PSUs for the main study (a county or group of counties). Each SSU will comprise Census Block Groups (CBGs) or a group of contiguous block groups (BGs) if any BG does not meet minimum size requirements (that is, if it is expected to contain fewer than 75 survey eligible households).⁴

SSUs will be selected using probability proportional to size (PPS) sampling. The measure of size (MOS) will be a composite of three measures: the number of SNAP households in the SSU (as calculated from SNAP administrative data), estimates of the numbers of low-income non-SNAP households in the SSU, and estimates of the numbers of very low-income non-SNAP households in

⁴ We hope to complete an average of 25 interviews per SSU, so having an estimated number of 75 survey eligible households will allow for inaccuracies in the estimates of survey eligible households as well as for difficulties in an SSU that could lead to lower-than-expected response rates.

the SSU. The composite measure, of numbers within each SSU of households in each of the sampling strata, reflects the relative overall sampling rate of households within the SSU. The composite MOS will enable us to obtain samples of households that have nearly equal probabilities of selection within each study population group.

Within sampled SSUs, we will sample addresses for screening from two or three sources (sampling frames):

- **SNAP frame.** The state or states in which the sites are located will be asked to provide addresses of all SNAP recipients in the sites; these addresses will comprise the frame for selecting households expected to be receiving SNAP.
- **Non-SNAP frame.** A commercial list of addresses, known as an Address-Based Sampling (ABS) frame, compiled from the United States Postal Service Delivery Sequence File will be the preferred frame for non-SNAP addresses.
- **Alternative non-SNAP frame.** If the ABS frame for an SSU contains a large number of addresses that are not useful for locating households (P.O. Boxes, Rural Delivery), the sampled SSU will be listed by Mathematica field staff, and the listed addresses will comprise the sampling frame (in that SSU) for non-SNAP addresses. If the SSU needing listing is large (more than 200 housing units), it will be divided into listing areas and two of these will be selected with PPS.

Within each SSU or listing area, addresses in the non-SNAP (ABS or field-listed) frame will be cross-checked against the SNAP frame and any that appear on both frames will be eliminated from the non-SNAP frame.

Initial samples of 400 SNAP and 4,000 non-SNAP addresses will be selected for screening.⁵ Although the number of SNAP addresses is expected to be approximately the same for each SSU, the number of non-SNAP addresses will vary depending on the allocation to the SSU of the two non-SNAP groups and the expected prevalence of each group in the SSU.

We will use two methods for sampling addresses from the non-SNAP frame. In 12 of the 16 SSUs we will select an equal probability sample of non-SNAP addresses. In the other 4 SSUs (2 per site), we will oversample addresses on the non-SNAP frame that are adjacent to addresses on the SNAP frame. The theory underlying this test is that nonparticipants who are eligible for SNAP, or

⁵ Mathematica expects to need 300 SNAP addresses, but will keep a random sub-sample of 100 in reserve. For SNAP households, we assume 95 percent of addresses contacted will be eligible. (The other 5 percent will either be invalid addresses, non-household housing units or will, at the time of the contact, be occupied by a household that is not eligible for the survey.) We also assume that 81 percent of these will provide complete data. In addition, we assume that 87 percent of the addresses provided will result in a contact. For other households, we expect the screening eligibility rate for the eligible non-SNAP group to be 25 percent in the SSUs selected. We expect to make contact with 90 percent of households and obtain a screener completion rate of 75 percent and complete data collection among 81 percent of eligibles. We expect 80 percent of the addresses on the ABS frame to be deliverable household addresses. If these assumptions are correct, we will need a sample of 1828 addresses after unduplicating with SNAP records, and a total of 2504 before unduplicating. To allow for inaccuracies in our assumptions we will select an initial sample of 4000 addresses.

close to the SNAP's income cutoff are more likely to live in close proximity to SNAP households. We will evaluate this procedure to see if it can reduce data collection cost, and also estimate the extent to which it may increase the design effect.

Addresses from each frame will be contacted and screened for the presence of eligible households.⁶ When an eligible household is identified, we will attempt to collect data for the study. The addresses in the initial samples will be randomly sorted into "r" replicate subsamples ($r = 1, 2, \dots, R$) and released in waves. After the target number of interviews for any of the three groups (SNAP, low-income non-SNAP, very low -income non-SNAP) is attained, then that group will be ineligible for interviewing in subsequent releases. Further, if the SNAP target is met before all replicates have been released, subsequent releases will not include any addresses from the SNAP frame.

The full-scale survey will use an approach similar to the field test, but PSUs will be selected with known probability rather than judgmentally and the target population will include a fourth group of non-SNAP households with income above 185 percent of poverty. Higher income households are included in the full-scale survey to provide a nationally representative sample of households. Higher-income households are not included in the field test of survey protocols because the cognitive task of reporting food acquisitions over a seven-day period is not considered a substantial burden for that population. Including higher-income households in the field test would also reduce the power to detect differences in response rates and data quality between subgroups defined by the two incentive levels and two survey protocols.

The preliminary design for the full-scale survey calls for a multistage sample. First, PSUs will be defined as counties or groups of contiguous counties. In forming PSUs, metropolitan statistical area (MSA) boundaries will be used (an MSA may be split, but part of one MSA will not be joined to part of another MSA to form a PSU).

After PSUs for the full-scale survey are formed, a stratified sample of 50 PSUs will be selected using PPS selection. The MOS for each PSU will be a composite that reflects the overall sample sizes and the estimated population in each PSU of each of the four groups for which separate sample sizes have been defined. These include households receiving SNAP benefits and three groups of non-SNAP households: (1) those with income at or less than the poverty level; (2) those with income at or above poverty but no more than 185 percent of that level; and (3) households with higher income.

Within the sampled PSUs, secondary sampling units (SSUs) will be formed. The preliminary plan is to define SSUs as CBGs or groups of CBGs. Based on the results of the field test, we may revise the SSU definition to include Census Tracts (CTs) or postal zip code areas rather than CBGs.

When SSUs are defined, we will assign each a composite MOS, defined in the same way as for PSUs, and select eight SSUs per PSU using PPS sampling.

The target number of completes for the main study are

⁶ We will include any eligible household (SNAP or eligible non-SNAP) identified on either frame, a procedure that will increase sample frame coverage.

- 1,500 SNAP households
- 800 non-SNAP households with income less than poverty
- 1,200 non-SNAP households with income between 100 and 185 percent of poverty.
- 1,500 non-SNAP households with income above 185 percent of poverty.

The process of sampling within SSUs will be the same as for the field test, except in cases in which the results of the field test suggest that changes should be made.

B2. Procedures for Collection of Information

Statistical Methodology. The proposed methods for stratification and sample selection for the field test are discussed in Section B.1.

Estimation procedures. Data from the field test will be analyzed using tabular analysis and statistical comparison of means and proportions. Before analysis, the data will be weighted to reflect differences in probability of selection (within site) and where appropriate, weighting adjustments will be made reflecting the difference in propensity to respond.⁷ Our analysis will include: nonresponse patterns, data quality (including item non-response), and the impact of clustering at the SSU level.

Degree of Accuracy. The field test sample is designed to provide adequate precision to evaluate data collection procedures and data quality and to provide adequate precision and adequate power to detect differences between the experimental treatments on cooperation rates.

Expected design effects for the full-scale survey are shown in Table B.1. We estimate design effects (Deffs) to range from 1.38 to 2.38 for measures with low intracluster correlations (ICCs). These estimates are based on Deffs reported in Cohen et al. 1999⁸ and further analysis of the same data. We expect values of the ICCs to be between 0.01 and 0.05, and the table shows the values of Deff at the ICC values of both 0.01 (first two rows) and 0.05 (last two rows). The value of Deff_w of 1.07 was derived from the same study. Expected design effects for the full-scale survey will be updated based on findings from the field test.

⁷ For example, household level non-response adjustments would be appropriate for analysis of consumption data, but not for computing response rates.

⁸ Cohen, Barbara, James Ohls, Margaret Andrews, Michael Ponza, Lorenzo Moreno, Amy Zambrowski, and Rhoda Cohen. "Food Stamp Participants' Food Security and Nutrient Availability: Final Report." Princeton, NJ: Mathematica Policy Research, July 1999.

Table B.1. Expected Design Effects for the Full-Scale National Food Study

Group	Completed Households	PSUs	b-1	ICC	Deff_c	Deff_w	Deff	Effective n
All	3,000	50	59	0.01	1.59	1.5	2.38	1,257
SNAP or Not	1,500	50	29	0.01	1.29	1.07	1.38	1,086
All	3,000	50	59	0.05	3.95	1.5	5.92	506
SNAP or Not	1,500	50	29	0.05	2.45	1.07	2.62	572.

Notes: $Deff = deff_c * deff_w$

deff_c is the design effect due to clustering

deff_w is the design effect due to unequal weights

$deff_c = 1 + ICC(b-1)$

ICC is the intracluster correlation

b is the number of cases per PSU

Table B.2 presents expected design effects for the field test. For the field test, SSUs will serve as PSUs. We estimate design effects (Deffs) for the SSUs to range from 1.3 to 2.6 for study outcomes (data from completed interviews). Measures of food acquisitions and population demographics are expected to have higher ICCs, and thus higher Deffs, while measures of response rates and burden are expected to have lower ICCs (0.005 to 0.01). The precision for estimates of response rates, burden, and data quality are based on the design effects due to weighting and clustering presented in Table B.2.

Statistical Power. The precision for the field test will vary for different groups. For measuring screener response, the sample will have a maximum 95 percent confidence interval of between ± 3.8 and ± 4.3 percentage points for the sample as a whole and from ± 4.9 to ± 5.3 percentage points for any treatment group or for either SNAP or non-SNAP households.

For field test estimates of the impact of either experimental treatment on cooperation rates, we focus on the ability of the sample to detect differences. The power to detect differences between treatment groups is shown in Table B.3 panels a through c.

Specialized Sampling Procedures. Area probability methods are proposed because personal visits to the households are required to train respondents and collect data. The use of ABS sampling frame enables the identification of non-SNAP households at a lower cost, compared to field listing of addresses. The use of two sampling frames (SNAP and ABS) is proposed because the SNAP frame is the most efficient way to sample SNAP households, and will yield few if any non-SNAP households. Data for the sampling frame of SNAP participants will be obtained from State SNAP agencies two months prior to beginning field efforts.

Table B.2. Expected Design Effects for the Field Test

Group	Sample Size	PSUS	b*-1	icc	Deff C	Deff W	Deff	Effective Sample Size	Expected CI
Low ICC Assumed									
All Households	1476	16	91.25	0.005	1.456	1.456	2.184	676	3.8
Per Treatment Group	738	16	45.125	0.005	1.226	1.226	1.839	401	4.9
All Eligible Households	494	16	29.875	0.005	1.149	1.5	1.724	287	5.8
Per Treatment Group	247	16	14.4375	0.005	1.072	1.5	1.608	154	7.9
Completed Interviews	400	16	24	0.01	1.240	1.18	1.463	273	5.9
Per Treatment Group	200	16	11.5	0.01	1.115	1.18	1.316	152	7.9
High ICC Assumed									
All Households	1476	16	91.25	0.01	1.913	1.5	2.870	514	4.3
Per Treatment Group	738	16	45.125	0.01	1.451	1.5	2.177	339	5.3
All Eligible Households	494	16	29.875	0.01	1.299	1.5	1.949	253	6.2
Per Treatment Group	247	16	14.4375	0.01	1.144	1.5	1.716	144	8.2
Completed Interviews	400	16	24	0.05	2.200	1.18	2.596	154	7.9
Per Treatment Group	200	16	11.5	0.05	1.575	1.18	1.859	108	9.4

Notes: Deff = deff_c * deff_w

deff_c is the design effect due to clustering

deff_w is the design effect due to unequal weights

deff_c = 1 + ICC(b-1)

ICC is the intracluster correlation

b* is the number of cases per SSU

Table B3 Power of Statistical Tests for the Field Test**A) Power to Detect Differences In Screener Completion**

Difference between treatment levels	Power, 2 tailed test	Power, 1 tailed test
0.3	>0.99	>0.99
0.25	>0.99	>0.99
0.2	>0.99	>0.99
0.15	0.989	>0.99
0.1	0.809	0.883
0.09	0.722	0.817
0.075	0.565	0.684
0.05	0.293	0.41

Total sample of 1476 screened households; design effect of 1.9; mid-range proportion of 0.30 to 0.70.

B) Power to Detect Differences In Interview Completion

Difference between treatment levels	Power, 2 tailed test	Power, 1 tailed test
0.3	>0.99	>0.99
0.25	>0.99	>0.99
0.2	0.939	0.969
0.175	0.866	0.923
0.15	0.748	0.838
0.125	0.591	0.707
0.1	0.418	0.543
0.05	0.139	0.221

Total sample of 494 interview-eligible households; design effect of 1.6; mid-range proportion of 0.30 to 0.70.

C) Power to Detect Differences In Data Among Completes

Difference between treatment levels	Power, 2 tailed test	Power, 1 tailed test
0.3	>0.99	>0.99
0.25	>0.99	>0.99
0.2	0.937	0.967
0.175	0.862	0.92
0.15	0.744	0.834
0.125	0.587	0.703
0.1	0.414	0.539
0.05	0.138	0.22

Total sample of 400 completed interviews; design effect of 1.3; mid-range proportion of 0.30 to 0.70.

B3. Methods to Maximize Response Rates and Deal With Nonresponse

After the sample of addresses is selected, the contractor will mail an attractively designed and easily understood survey brochure to each address, along with the advance letter signed by a USDA official. These materials are designed to convince potential respondents of the surveys' legitimacy, value and the importance of participation. In addition, we plan to inform local area organizations, such as senior centers, WIC agencies, and SNAP offices about the study being conducted in their area in case members of the public inquire about the legitimacy of the survey.

The field test will use all survey protocols planned for the full-scale survey, except with respect to the alternate incentive levels being tested and the alternate instruments tested for collecting food acquisition data. To maximize response to the household screener, all sampled addresses will be screened in person by trained, professional field staff. Immediately after screening, field interviewers will explain to eligible households the importance of the survey, its requirements and incentive, and recruit them to participate. We expect that in-person screening and recruitment, together with an incentive to offset the burden of participation, will result in participation of 86 percent of households that complete the screener and are found eligible for the survey. Testing this expectation is one of the purposes of the field test.

To maximize response throughout the data collection week, respondents will be assigned to a single field interviewer and a single telephone interviewer. The field interviewer will conduct the initial screening, train the respondent on survey protocols, and conduct the first and third household interviews. A single telephone interviewer will be assigned to conduct the food-away-from-home interviews on days two, five, and seven, as well as the second household interview conducted by telephone. During the pre-test, telephone interviewers were able to establish rapport with respondents during the week, and limiting the number of survey staff interacting with respondents reinforces the fact that the data obtained are confidential.

The National Food Survey will impose a large burden on participating households. The reporting burden increases with household size because (1) many questions in the household interviews are asked with regard to each household member and (2) the number of food acquisitions and number of food items per acquisition are expected to vary with household size. A base incentive will be offered to households for participation, along with a small additional incentive for each additional household member who reports food acquisitions for the week. This design is intended to encourage response from household members other than the primary respondent. Findings from the cognitive tests indicate that teenagers, in particular, might be reluctant to participate without a targeted incentive.

Two incentive levels will be tested in the field test, labeled "low incentive" and "high incentive" in Table B.4 (which repeats Table A.1 in Part A). The low and high incentives are tiered to reflect the additional burden for larger households, and to encourage participation of all household members. The base incentive will be provided to the main respondent as a check at the end of the survey week; the additional incentive per household member will be mailed as gift cards so that they can be easily distributed to other household members. The incentive scheme also includes a "telephone bonus" if the primary respondent initiates the telephone reporting of food acquisitions. This bonus is designed to reduce overall data collection costs for the survey. Interviews initiated by incoming calls from respondents are completed at significantly lower cost than outgoing calls with multiple callbacks to obtain these responses.

Table B.4. Incentive Levels Tested in the National Food Survey Field Test

Household Size	Percentage of Sample ^a	“Low” Incentive	Telephone Bonus	Total “Low” Incentive	“High” Incentive	Telephone Bonus	Total “High” Incentive
1	45.5	\$50	\$25	\$75	\$100	\$25	\$125
2	19.8	\$70	\$25	\$95	\$125	\$25	\$150
3	15.6	\$90	\$25	\$115	\$150	\$25	\$175
4	10.1	\$110	\$25	\$135	\$175	\$25	\$200
5	5.5	\$130	\$25	\$155	\$200	\$25	\$225
6	2.3	\$150	\$25	\$175	\$225	\$25	\$250
Average		\$72		\$97	\$128		\$152

^aCharacteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2008 (Table A-5).

Field staff will train respondents to use a handheld scanner, book of barcodes, and food booklets to track food acquisitions during the survey week. This training will include scanning practice food items from the field interviewer’s training kit and completing practice forms from the food booklet. The training is designed to put respondents at ease and leave them confident of their ability to use the survey tools to track their food acquisitions during the week. Cognitive tests were conducted with 16 households and respondents reported that the initial training to use the food instruments left them confident in their understanding of the tasks and ability to complete the survey protocols. We expect that 90 percent of households that agree to participate will track food obtained for home preparation and consumption using the scanner, and that 90 percent of food booklets will be completed.

The primary respondent in each sampled household will be asked to report food acquisitions by telephone on days two, five, and seven of the survey week. Response to these interviews is critical to obtaining accurate food acquisition data, because interviewers can ask clarifying questions to obtain precise food descriptions during the interview. In addition to collecting data from the food booklets, these interviews provide an opportunity to answer respondents’ questions, provide feedback on their tracking activities, and offer reminders about survey protocols (such as saving receipts and scanning food items). Respondents will receive multiple reminders of these interviews: the schedule will be printed on the food booklets and on a calendar magnet that they may place on their refrigerator. Respondents will be asked to call the survey center at a time that is convenient for them on the designated days. They will be offered a nominal incentive bonus for initiating these calls (the “telephone bonus” shown in Table B.4), thus offsetting the costs of outgoing calls and callbacks from the survey center.

The contractor will maintain a study website and toll-free number. Respondents can go to the website to get information about the study and browse the questions and answers regarding survey protocols. The toll-free number can be used to obtain specific help with survey procedures, to voice concerns, or to call in at a convenient time for reporting food acquisitions.

Maximizing Overall Response Rates

The full-scale National Food Survey will be conducted over a six-month period. Original plans for a four month field period were modified in response to recommendations from the Technical Work Group (TWG). The TWG recommended multiple sequential releases of sample and careful monitoring of rates of screening, eligibility, and response in order to maximize response rates for the

survey. The field test will be conducted over an eight week period with three releases of sample during that time period.

Nonresponse Bias

The expected response rates are less than 80 per cent: 71 per cent for SNAP households and 55 percent for other households. These response rates are derived by multiplying the expected screener completion rate, recruitment rate, and final completion rate. The difference in overall response rates for SNAP and non-SNAP households is entirely due to expected differences in screener completion rates (87 percent for SNAP and 67.5 percent for non-SNAP). The field test will provide information about the accuracy of these expected response rates and potential avenues for improving the screener completion rates for non-SNAP households.

We will take several steps to assess the potential of bias due to nonresponse and correct for it by weighting. Nonresponse can increase the potential for bias, and hence inaccurate estimates, when those in the sample who did not respond differ in important respects from those that did. Our nonresponse bias analysis will be coordinated with sample weighting. In evaluating the response patterns for the weighting activities, we will compute response rates for different subpopulations to look for variation among these rates and we will compare the characteristics of respondents and nonrespondents.

To assess the potential for nonresponse bias in the survey estimates of the SNAP population, we will use administrative data on size of household, household characteristics, and the amount of SNAP benefits and compare the characteristics of the respondents and nonrespondents on these measures. For the non-SNAP households we will have limited data, but can compare the respondents to estimated population characteristics of households in the SSUs and sites, such as race, income, and household size. As an additional step in the nonresponse bias analysis, we will compare initial respondents with initial refusals to provide an indication of potential bias that might exist between all respondents and the final refusal sample for the full sample and for subpopulations. We will also look at the characteristics of the nonrespondents by the reason for nonresponse and by data collection stage at which the nonresponse occurred to assess the potential for nonresponse bias. In addition, eligible households who complete the initial screener and refuse to participate will be asked a few questions about the number of adults and children in the household and the types of places at which members of the household have shopped in the past 30 days (Appendix E). This information will provide additional data to help assess nonresponse bias.

Steps in assessing the potential impact of nonresponse will include comparing the responding households with external data at several steps: before weighting, after weighting to account for differences in probability of selection, and after weighting for nonresponse. The first two steps will indicate the potential for bias and the third will indicate whether weighting has sufficiently reduced the potential for bias. As in most studies, we will not have external estimates of study outcomes (dependent variables), but we will use several measures of household characteristics that should be correlated with the dependent variables. For the SNAP households, we will use administrative data to tabulate caseload and respondent distributions by

- Household type (with children, without children, with elderly)
- Race of household head
- SNAP benefit amount

- Households with and without earnings
- Metro and nonmetro

Compared to the nonresponse analysis for the main study, the nonresponse analysis in the field test will be more difficult, particularly for the non-SNAP group. This is because accurate site-specific benchmark estimates are more difficult to obtain than national ones. For the non-SNAP households (or for the two groups aggregated) we will use American Community Survey (ACS) three year or five year (if available) data⁹ on a variety of economic and demographic characteristics, including type of household (family, nonfamily) household size, race of householder, home ownership, and earnings. We will use the external data to create estimates of these characteristics for the site population that has income less than 185 percent of poverty and compare these to the distribution of our sample.

B4. Tests of Procedures or Methods to be Undertaken

The contractor conducted a cognitive test of the food instruments, including the handheld scanner, book of barcodes, and food booklets. Three versions of the food booklets were tested with a total of 16 households from May 3 through May 24, 2010. These tests were designed to assess the clarity of the food instruments, identify possible modifications to content and/or sequence, and estimate respondent burden. Four local social services agencies were contacted and assisted in recruiting a convenience sample of low-income households of various sizes. Households were trained on how to use the scanner and record their food acquisitions in the food booklets; they were asked to track their food acquisitions for two days, after which time survey staff returned and completed an in-person debriefing with the primary respondent.

Findings from the cognitive test were used to modify data collection instruments and procedures. Verbal and written instructions for the handheld scanner were revised, and methods for associating scanned items with reported food acquisitions were modified. For example, the tested forms included a single checkbox for respondents to indicate that they did not acquire food, and this was removed so that respondents are not tempted to underreport. Some respondents reported that they guessed at quantities of food for school meals and non-packaged items, and instructions were revised to ask for the “size or amount if known” (if written on a package or menu).¹⁰ One of the tested versions of instruments was associated with reporting of foods consumed that were not acquired during the test period; this version was dropped. The cognitive test found that some respondents missed scanning a small number of items that were listed on store receipts; these data will not be lost because scanned data will be compared to store receipts as prices are data entered from receipts. The cognitive test also found that respondents did not save all receipts from food acquisitions. The training protocols have been revised to emphasize the importance of receipts, and the telephone interview scripts for days two, five, and seven were revised to include reminders about saving receipts.

⁹ If ACS data are not available we will use model-based estimates from a commercial vendor.

¹⁰ With this approach, the average reliability of reported data is improved and missing data can be imputed. This approach is also consistent with the Technical Work Group recommendation to sacrifice detailed data where it will improve accuracy and data quality.

A full pretest of all survey instruments was conducted with a total of six additional households from this convenience sample from June 21 to July 19, 2010.¹¹ Respondents were trained, asked to track food acquisitions for seven days, and to call the survey center to report food acquisitions on days two, five, and seven. Respondents were also administered the three household interviews. At the end of the data collection week, respondents participated in two brief debriefing sessions with contractor staff. As the food instruments had already been through cognitive testing, these debriefing sessions focused on the burden associated with completing the full data collection (including time spent reviewing instructions/training documents, and tracking and reporting food acquisitions); the appropriateness of the planned compensation; and potential improvements to procedures or materials. The results of the pretest and recommendations for adjustments for the field test are presented in Appendix P.

B5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The Field Test of the National Food Survey will be administered by ERS' contracting organization, Mathematica Policy Research. The same contractor will analyze field test data. Individuals whom ERS consulted or whom we expect to consult on the collection and/or analysis of the field test data include those listed below.

Nancy Cole
Mathematica Policy Research, Inc.
955 Massachusetts Avenue, Suite 801
Cambridge, MA 02139
(617) 674-8353

Mary Kay Fox
Mathematica Policy Research, Inc.
955 Massachusetts Avenue, Suite 801
Cambridge, MA 02139
(617) 301-8983

Laura Kalb
Mathematica Policy Research, Inc.
955 Massachusetts Avenue, Suite 801
Cambridge, MA 02139
(617) 301-8989

John Hall
Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543
(609) 275-2357

¹¹ Nine households were contacted for the pre-test from the group that expressed interest in the cognitive test when contacted in April. Three households declined to participate; two cases refused due to a significant change in circumstances, while the third had a mild language barrier and was deterred by the question about citizenship. The question about citizenship, which is needed to determine SNAP eligibility among non-SNAP households, has been moved to the last interview to minimize data loss due to respondent sensitivity to this question.

TECHNICAL WORK GROUP MEMBERS

Steven Heeringa
Institute for Social Research
University of Michigan
426 Thompson
Ann Arbor, MI 48104
(734) 647-4621
sheering@umich.edu

Sarah Nusser
Center for Survey Statistics & Methodology
Iowa State University
220 Snedecor
Ames, IA 50011-1210
(515) 294-9773
nusser@iastate.edu

Helen Jensen
Center for Agricultural and Rural
Development
Iowa State University
578E Heady Hall
Ames, IA 50011-1070
(515) 294-6253
hhjensen@iastate.edu

Roger Tourangeau
Director, Joint Program in Survey
Methodology
University of Maryland
1218 LeFrak Hall
College Park, MD 20742
(301) 314-7984
rtourang@umich.edu

Suzanne Murphy
Professor
Cancer Research Center of Hawaii
1236 Lauhala St., Suite 407
Honolulu, HI 96813
(808) 564-5861
Suzanne@crch.hawaii.edu

Parke Wilde
Friedman School of Nutrition Science
and Policy
Tufts University
150 Harrison Ave
Boston, MA 02111
(617) 636-3495
parke.wilde@tufts.edu

Inquiries regarding statistical aspects of the study design should be directed to:

Mark Denbaly
USDA Economic Research Service
1800 M Street
Washington, DC 20036
(202) 694-5390

Mr. Denbaly is the project officer.