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

Part B: Statistical Methods for Baseline Data Collection

December 5, 2011

Revised March 9, 2012

CONTENTS

All revisions to the original document have been highlighted as yellow text.

B. Statistical Methods for Baseline Data Collection

B1. Respondent Universe and Sampling Methods

The respondent universe includes three groups for analysis: (1) households receiving SNAP benefits[1]; (2) low-income households not receiving SNAP benefits; and (3) higher-income hnouseholds (not receiving SNAP benefits). The non-SNAP households will be divided into three subgroups that will be sampled at different rates, so that, for sampling, four groups will be selected: (1) households receiving SNAP benefits; (2) households not receiving SNAP benefits whose income is below the Federal poverty guideline, (3) households not receiving SNAP benefits whose income is between the poverty guideline and 185 percent of that guideline and, (4) households not receiving SNAP whose income is greater than 185 percent of the poverty guideline.

These subgroups were defined to provide a sample that assures adequate representation of SNAP participants, low-income households who are potentially eligible for USDA food assistance programs, and higher-income households so that the results of the study can be generalized to all U.S. households.[2] We will use the 2012 poverty guidelines to define the groups. We expect a screener response rate of 80 percent calculated using the American Association for Public Opinion Research formula 4 (AAPOR 2008). We expect an overall survey response rate of 55 percent calculated using the AAPOR formula. (The response rate is addressed in more detail in section B.3.)

A field test was conducted (from February to May 2011) in two purposively selected sites in the mid-Atlantic region within proximity to Mathematica's survey operations center. Some results of the field test are discussed in section B.4. They have been used in this section to help derive estimates of the samples of addresses needed to conduct the survey. Nationally, the estimated number of households in each of the study population groups[3] are:

- 15.2 million SNAP households;

- 2.0 million non-SNAP households with income below the federal poverty guidelines (very low income);

---

[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 poverty level thresholds for the non-SNAP households are comparable to eligibility limits for the largest food assistance programs: 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 2010 estimates of families in poverty (U.S. Census Bureau, Current Population Reports, P60-239) and USDA estimates of participation rates (Leftin and Wolkowitz, 2009).

- 16.2 million non-SNAP households with income between 100 and 185 percent of the federal poverty guidelines (low income).

- 84.1 million non-SNAP households with income above 185 percent of the federal poverty guidelines (higher income).

The sample sizes for the main study (number of completed observations) are:

- 1,500 SNAP households (SNAP);

- 800 non-SNAP households with income less than poverty (very low income).

- 1,200 non-SNAP households with income between poverty and 185 percent of poverty (low income); and

- 1,500 non-SNAP households with income greater than 185 percent of poverty (higher income).

These sample sizes will allow adequate precision for estimates of food acquisition, both overall and for the SNAP and non-SNAP groups. Expected design effects and power for statistical tests are discussed in Section B2.

The full-scale survey uses 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 (approximately 16,000 to 22,000 overall) will be selected from a sampling frame constructed from multiple sources (described below) at the third stage of sampling within each SSU.

All sampled addresses will be randomly grouped into replicate subsamples and each replicate assigned to one of three waves. Beginning with the first wave, addresses will be released in batches every two weeks (each batch will comprise one or more replicates). Up to 10 batches will be released. The first batch will be released during the first month of data collection; and the final batch in the fifth month (data collection will last 6 months). Those released will be contacted to determine if they are occupied by households. Group quarters residences and seasonal units will be excluded. Households will be screened on income and receipt of SNAP so they can be correctly categorized. Targets will be set for each wave. Within each wave, once targets for any of the four groups are met, further batches of sample release (replicates) within the wave will treat that group as ineligible for data collection beyond the screener.

The full-scale survey will employ 50 PSUs selected with known probability; the target groups will include all households in the U.S. (excluding Alaska and Hawaii). The four target groups (SNAP and non-SNAP identified in three income groups) will be sampled at different rates. The group with the highest sampling rate are the very low-income non-SNAP households, and the lowest sampling rate will be for higher-income (non-SNAP) households. Higher-income households were not included in the field test of survey protocols because the task of reporting food acquisitions over a seven-day period is not considered a substantial cognitive burden for population groups with higher levels of education.

Before sampling, PSUs were defined as counties or groups of contiguous counties. In forming PSUs, metropolitan statistical area (MSA) boundaries were used (some MSAs were split into multiple PSUs, but in no case was part of one MSA joined to part of another MSA to form a PSU).

After the PSUs were formed, a stratified sample of 50 PSUs was selected using probability proportional to size (PPS) selection. The measure of size (MOS) for each PSU was a composite of four estimates derived from American Community Survey (ACS) Public Use Microdata Sample (PUMS) files: the number of SNAP households in the PSU, and numbers of very low-income, low-income, and higher-income non-SNAP households in the PSU. The composite measure reflects the numbers within each PSU of households in each of the sampling strata and the relative overall sampling rate of households within the PSU. Using the composite MOS to select PSUs (and SSUs) will help us to obtain samples of households that have nearly equal probabilities of selection within each study population group.

Within the sampled PSUs, secondary sampling units (SSUs) wereformed. Eight SSUs have been selected in each sampled PSU. Each SSU comprises a Census Block Group (CBG) or a group of contiguous block groups if any CBG does not meet minimum size requirements (that is, if it is expected to contain fewer than 50 survey-eligible households).[4]

SSUs were selected using probability proportional to size (PPS) sampling. The MOS for selecting the SSUs was a composite MOS constructed in the same way as the MOS for PSUs.

Within sampled SSUs, we will sample addresses for screening from a single sampling frame. Each frame will be constructed from one, two or three list sources:

- **SNAP Address list.** State SNAP agencies in the states in which the PSUs are located have been asked to provide addresses of all SNAP recipients in the PSUs (or for multi-county PSUs, those in counties with SSUs in the sample); these addresses will be flagged on the final sampling frame to more efficiently select households expected to be receiving SNAP.

- **Non-SNAP Address list.** A commercial list of addresses (in each SSU), known as an Address-Based Sampling (ABS) list, compiled from the United States Postal Service Delivery Sequence File, will be obtained for each SSU; we expect that a great majority of non-SNAP households will be identified at these addresses. Addresses of SNAP households from the SNAP list will be matched to the ABS list and the lists unduplicated so that each address appears only once on the final frame.

- **Alternative non-SNAP Address list.** If the ABS list 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 (along with any addresses from the SNAP list that can be identified[5] and flagged) will comprise the sampling frame in that SSU. If the SSU needing listing is large

---

[4] We hope to complete an average of 12.5 interviews per SSU, so having an estimated number of 50 survey eligible households will allow for inaccuracies in the estimates of survey eligible housheolds as well as for difficulties in an SSU that could lead to lower-than-expected response rates.

[5] In listed SSUs, address matching with SNAP data may be more difficult than in SSUs for which the ABS frame is used as the non-SNAP frame: addresses from the SNAP frame may be similar to those on the ABS frame (e.g., largely P.O. Boxes or Rural Delivery) or may be difficult to match with the addresses recorded by the field staff.

(more than 200 housing units), it will be divided into listing areas and two of these will be selected with PPS.

- **Joint SNAP/non-SNAP Address list**. In states where agencies are unable to provide addresses of all SNAP recipients in the PSUs, the non-SNAP lists noted above (ABS or field-listed) will be used to identify both SNAP and non-SNAP households. Additional screening effort is expected in these States relative to States providing SNAP lists.[6]

To obtain the planned 5,000 completes, the overall sample is estimated to require release of between 1,918 and 3,551 SNAP addresses and between 13,956 and 17,998 non-SNAP addresses. In addition to response rates at various steps of the overall data collection process, the number of addresses needed from each list will depend on the accuracy of the SNAP lists, the speed with which the information in the SNAP lists grows stale due to household moves and exits from SNAP, and the distribution of the three non-SNAP groups among addresses not occupied by SNAP households.[7]

The accuracy of the SNAP list refers to the accuracy of household address information in SNAP data systems. Address information which is accurate at the time of SNAP enrollment may no longer be accurate at the time address information is sent to the survey contractor. We expect there will be variation among states in the accuracy of their SNAP files, and we cannot predict exactly how accurate these files will be on average or how much variation there will be from state to state. Furthermore, as the survey effort progresses, some SNAP households will leave the program or move elsewhere. These events will cause the SNAP frame to become less efficient in identifying SNAP addresses as time passes.[8][9] We do not have information that would allow us to precisely estimate the timing of these events at the PSU or SSU level, which leads to the uncertainty of how many addresses from the SNAP list will need to be released.[10] The aging of SNAP information during the field period may require increased screening effort to locate SNAP households from the non-SNAP lists.

---

[6] The source of address data has no impact on field operations. The source of the address is not communicated to field staff.

[7] In calculating the sample sizes, we assume: (1) Mathematica will determine the residential status of all addresses released; (2) that, based on the field test, 83 percent of the non-SNAP and 88 percent of SNAP addresses will be occupied by households; (3) that 80 percent of attempted screeners will be completed; and (4) that 90 percent of households asked to provide food acquisition data will do so.

[8] It is possible to measure the percentage of addresses from the SNAP list that are not occupied by SNAP participants at the time of screening. It is not possible to decompose this measure into its component parts: (1) inaccuracy at the time the SNAP list is received, and (2) inaccuracy due to SNAP households moving during the field period.

[9] We considered asking State SNAP agencies to send additional extracts of up-to-date addresses of then-current SNAP households during the survey period, but we decided against this strategy because we did not want to increase the burden of our data requests on State agencies.

[10] Note also that the need to release addresses from the SNAP frame will be offset somewhat as some of the addresses in the initial non-SNAP frame become populated by SNAP households, either through a recent residential move or successful application for program benefits.

As with the SNAP list, it is difficult to estimate how many addresses from the non-SNAP lists will need to be released to achieve complete data collections with 800 very-low-income non-SNAP households, 1,200 low-income non-SNAP households, and 1,500 higher-income non-SNAP households. Predicting the distribution of these non-SNAP groups cannot be exact because of a number of factors: changing economic conditions in the recent past; use of data to select the PSUs and SSUs that, while the best available, is based on sample data and will be a few years out of date by the time the study is fielded; and the fact that the PSUs and SSUs comprise a sample of areas in the U.S. so any estimate of the distribution of the target groups in these areas is subject to sampling error.

The initial samples of SNAP and non-SNAP selected for screening for Wave 1,[11] will include approximately 1,200 SNAP addresses and 6,500 non-SNAP addresses.[12] The numbers ultimately released over all three waves will depend on the results of screening in the earlier waves and, of course, response rates within each group.

The sample for the first wave will allow for flexibility and the results will allow for more precise calibration of the sample distribution for the second and third waves. Although the total number of addresses is expected to be approximately the same for each SSU, the number of SNAP and non-SNAP addresses will vary depending on the sample allocation to each SSU of the target groups and the expected prevalence of each group in the SSU.

Addresses selected for the sample will be contacted and screened for the presence of eligible households.[13] When an eligible household is identified, we will attempt to collect data for the study. As discussed above, since data will be collected over a six-month period, separate samples of addresses will be selected for three waves to spread the interviewing over time. Target numbers of interviews for each of the four groups (SNAP, very-low income SNAP, low-income non-SNAP and higher income non-SNAP) will be established for each wave. The target for the first wave will be one-third of the overall target; targets for other waves will be adjusted based on the results of the first. The initial sample for each wave will be randomly sorted into "R" replicate subsamples and grouped for releases into batches of one or more replicates. Multiple batches will be released for each wave, after the target number of interviews for any of the four groups is attained, then that group will be ineligible for interviewing in subsequent batches within that wave. Further, if the SNAP target is met before all batches have been released, subsequent batches will not include any addresses from the SNAP frame.

To efficiently allow attainment of an adequately high response rate, we will employ sampling for non-response follow-up, known as two-phase sampling. This procedure has been used recently on the National Survey of Family Growth (DHHS 2010). The method and procedures for computing a response rate for such a design are discussed in AAPOR (2009, pp. 1-2). The initial sampling rate to

---

[11] As discussed above (page 2), there will be three waves of sample release.

[12] The balance will depend on how many States ultimately provide SNAP lists.

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

be used for sampling non-respondents[14] to the screener will be set two weeks after data collection starts for each wave. The sampling rate may be adjusted for subsequent batches within the wave. This procedure may also be employed to target refusal conversion for households that are screened and eligible for the study but refuse to participate.

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 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.[15] Our analysis will include: household nonresponse patterns; item nonresponse patterns; sample characteristics; distribution of FAH acquisitions by type (store type and nonpurchase types); distribution of FAFH acquisitions by type (type of eating place and nonpurchase types); and means, ranges, and distributions of key output variables (including total food expenditures, FAH and FAFH expenditures, food security, and food acquistions by food groups) by SNAP and income groups.

Degree of Accuracy. The sample is designed to provide adequate precision to estimate food expenditures and other outcome variables for the sample as a whole and to compare group averages between (1) SNAP and non-SNAP households; (2) SNAP and SNAP-eligible non-participating households; and (3) all low-income (SNAP and non-SNAP) and higher-income (above 185% FPL) households.

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 2.99 for measures with low intracluster correlations (ICCs). These estimates are based on Deffs reported in Cohen et al. 1999[16] and further analysis of the same data. They are also consistent with findings from the field test. 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 eight rows) and 0.05 (last eight rows). We anticipate that the value of ICC will be lower for more heterogeneous groups (e.g., All) than for more homogeneous groups. The value of Deff_w of 1.07 was derived from the same study. The value of Deff_w=1.5, shown for the "All" and "All Non-SNAP" groups, allows for the additional complexity of the current design (i.e., combining groups sampled at very different rates). The value of Deff_w=1.3 covers cases where groups with different sample rates are combined, but where the rates are not as different as in the cases of "All" and "All Non-SNAP."

---

[14] As noted below, our plan is to use two-phase sampling only for addresses where contact has not been established.

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

[16] 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
For measures with low ICC								
All	5000	50	99	0.01	1.99	1.50	2.99	1675.04
SNAP	1500	50	29	0.01	1.29	1.07	1.38	1086.72
All Non-SNAP	3500	50	69	0.01	1.69	1.50	2.54	1380.67
Non SNAP < 185% FPL	2000	50	39	0.01	1.39	1.07	1.49	1344.72
Eligible Non-SNAP*	1224	50	23.48	0.01	1.23	1.30	1.61	762.50
Eligible Non-SNAP**	1000	50	19	0.01	1.19	1.30	1.55	646.41
All Low (SNAP and non)	3500	50	69	0.01	1.69	1.50	2.54	1380.67
Non SNAP >185% FPL	1500	50	29	0.01	1.29	1.07	1.38	1086.72
For measures with high ICC								
All	5000	50	99	0.05	5.95	1.50	8.93	560.22
SNAP	1500	50	29	0.05	2.45	1.07	2.62	572.19
All Non-SNAP	3500	50	69	0.05	4.45	1.50	6.68	524.34
Non SNAP < 185% FPL	2000	50	39	0.05	2.95	1.07	3.16	633.61
Eligible Non-SNAP*	1224	50	23.48	0.05	2.17	1.30	2.83	433.09
Eligible Non-SNAP**	1000	50	19	0.05	1.95	1.30	2.54	394.48
All Low (SNAP and non)	3500	50	69	0.05	4.45	1.50	6.68	524.34
Non SNAP >185% FPL	1500	50	29	0.05	2.45	1.07	2.62	572.19

* Eligible non-SNAP include those with income <130% of FPL. This sample size estimate asssumes a uniform distribution of those between 100 and 185% of FPL, so that 35% (30/85) of the non-SNAP in the sample with incomes between 100 and 185% of FPL have incomes <130% of FPL

** Assumes that only 200 of the non-SNAP in the sample with incomes between 100 and 185% of FPL have incomes <130% of FPL

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 A2 in Appendix A identifies food item prices and quantities as among the highest priority data items for the National Food Study. When multiplied together, prices and quantities form expenditures. Estimating and comparing levels of weekly food expenditures and how those food expenditures are allocated between food alternatives (e.g., FAH vs. FAFH or fruits and vegetables vs. all oather foods) are critical outcomes for this study. As shown in Table B.2, which is based on the above expected design effects, the sample will produce confidence intervals, with low values of ICC, for estimates of percentages (e.g., percent of food acquisitions obtained away from home[17])

---

[17] Evaluated at an estimate of 50 percent, where the variance is highest. Estimates of other percentages would have narrower confidence intervals.

ranging from no greater than plus or minus 2.40 percentage points for the sample as a whole, to no greater than plus or minus 3.86 percentage points for eligible non-SNAP households. For the higher values of the ICC, the maximum confidence intervals would range from plus or minus 3.90 to 4.94 percentage points. Table B.2 summarizes the anticipated half-width confidence intervals. For continuous variables, such as the amount spent on food purchases, the projected confidence intervals for total weekly food expenditures range from plus or minus $5.93 to plus or minus $9.54 for the lower value of the ICC, and from plus or minus $9.64 to plus or minus $12.22 for the higher value of the ICC. The confidence intervals for FAH and FAFH purchases are narrower, as seen in Table B.2. ERS has determined that these confidence intervals are within acceptable limits for informing policies regarding agriculture and food assistance programs.

Table B.2. Half Width 95% Confidence Intervals

Group	Completed households	Effective Sample	P near			Weekly food expenditures***		
			10%	25%	50%	FAH	FAFH	Total
For measures with low ICC								
All	5000	1675.04	1.44	2.07	2.40	$5.11	$2.49	$5.93
SNAP	1500	1086.72	1.78	2.58	2.97	$6.34	$3.10	$7.36
All Non-SNAP	3500	1380.67	1.58	2.28	2.64	$5.63	$2.75	$6.53
Non SNAP < 185% FPL	2000	1344.72	1.60	2.32	2.67	$5.70	$2.78	$6.62
Eligible Non-SNAP*	1224	762.50	2.13	3.08	3.55	$7.57	$3.70	$8.79
Eligible Non-SNAP**	1000	646.41	2.31	3.34	3.86	$8.22	$4.02	$9.54
All Low (SNAP and non)	3500	1380.67	1.58	2.28	2.64	$5.63	$2.75	$6.53
Non SNAP >185% FPL	1500	1086.72	1.78	2.58	2.97	$6.34	$3.10	$7.36
For measures with high ICC								
All	5000	560.22	2.49	3.59	4.14	$8.83	$4.31	$10.25
SNAP	1500	572.19	2.46	3.55	4.10	$8.74	$4.27	$10.14
All Non-SNAP	3500	524.34	2.57	3.71	4.28	$9.13	$4.46	$10.60
Non SNAP < 185% FPL	1500	633.61	2.34	3.37	3.90	$8.31	$4.06	$9.64
Eligible Non-SNAP*	1224	433.09	2.83	4.08	4.71	$10.05	$4.91	$11.66
Eligible Non-SNAP**	1000	394.48	2.96	4.28	4.94	$10.53	$5.14	$12.22
All Low (SNAP and non)	3500	524.34	2.57	3.71	4.28	$9.13	$4.46	$10.60
Non SNAP >185% FPL	1500	572.19	2.46	3.55	4.10	$8.74	$4.27	$10.14

* Eligible non-SNAP include those with income <130% of FPL. This sample size asssumes a uniform distribution of those between 100 and 185% of FPL, so that 35% (30/85) of the non-SNAP in the sample with incomes between 100 and 185% of FPL have incomes <130% of FPL

** Assumes that only 200 of the non-SNAP in the sample with incomes between 100 and 185% of FPL have incomes <130% of FPL

Analysis of the data will also include comparing outcomes between groups of respondents. The comparisons of interest are: (1) SNAP with all non-SNAP with income less than 185% of FPL; (2) SNAP with eligible non-SNAP (income less than 130% of FPL); and (3) all low income (SNAP and non-SNAP) with the non-SNAP having income greater than 185% of FPL. Table B.3 presents

expected minimum detectable differences (MDD) for estimated perecentages and for estimated food expenditures.[18]

Table B.3. Minimum Detectable Differences (MDDs)[a]

Group 1	Group 2	Total households (groups 1&2)	Design effect for contrasts	P near 10%	P near 25%	P near 50%	Weekly food expenditures*** FAH	FAFH	Total
For measures with low ICC									
Non SNAP < 185% FPL	SNAP	3000	1.07	2.97	4.28	4.95	$10.55	$5.15	$12.25
Eligible Non-SNAP*	SNAP	2724	1.13	3.44	4.96	5.73	$12.22	$5.97	$14.18
Eligible Non-SNAP**	SNAP	2500	1.11	3.62	5.22	6.03	$12.85	$6.27	$14.91
All Low income (SNAP and non)	Non SNAP >185% FPL	5000	1.30	2.95	4.26	4.92	$10.49	$5.12	$12.17
For measures with high ICC									
Non SNAP < 185% FPL	SNAP	3000	1.87	4.20	6.06	7.00	$14.92	$7.28	$17.31
Eligible Non-SNAP*	SNAP	2724	2.05	4.64	6.69	7.73	$16.47	$8.05	$19.12
Eligible Non-SNAP**	SNAP	2500	1.93	4.76	6.88	7.94	$16.93	$8.27	$19.65
All Low income (SNAP and non)	Non SNAP >185% FPL	5000	2.88	4.40	6.35	7.33	$15.64	$7.64	$18.15

[a]   MDDs are based on 80% power, 95% confidence and a 2-tail test.

*   Eligible non-SNAP include those with income <130% of FPL. This sample size asssumes a uniform distribution of those between 100 and 185% of FPL, so that 35% (30/85) of the non-SNAP in the sample with incomes between 100 and 185% of FPL have incomes <130% of FPL

**   Assumes that only 200 of the non-SNAP in the sample with incomes between 100 and 185% of FPL have incomes <130% of FPL

For the lower values of the ICC, the maximum MDDs (for P near 50%) range from 4.92 to 6.03 perentage points, and MDDs for total expenditures range from $12.17 to $14.91. With the higher value of the ICC, the maximum MDDs for estimated percentages range from 7.00 to 7.94 percentage points and for estjmated total expenditures from $17.31 to $19.65. ERS has determined that these MDDs are within acceptable limits for informing policies regarding agriculture and food assistance programs.

---

[18] The design effect for contrasts is based on a weighted average of the design effects for the two groups being compared but the clustering component of the design effect has been reduced since the groups being compared are not selected independently but come from the same set of PSUs and SSUs. In the case of "two subclass means from clustered samples…the positive correlation between cluster influences… tends to reduce the variance of the difference…" (Kish, 1965, Survey Sampling, Wiley. page 582).

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. Data for the sampling frame of SNAP participants will be obtained from State SNAP agencies two months prior to beginning field efforts. Two-phase sampling of screener non-responses is proposed as a cost-efficient way to reduce the potential of non-response bias.

B3. Methods to Maximize Response Rates and Deal With Nonresponse

Ten days prior to each sample release, the contractor will mail a full-color postcard to each address with information about the survey and the incentive for participation (Appendix D). This postcard is designed to convince potential respondents of the survey's legitimacy, value and the importance of participation. In addition, we plan to inform local area organizations, such as senior centers, WIC agencies, SNAP offices, and law enforcement about the study being conducted in their area in case members of the public inquire about the legitimacy of the survey.

To maximize response to the household screener, all sampled addresses will be screened in person by trained, professional field staff. To encourage respondent cooperation in completing the screener, a prepaid $5.00 token of appreciation will be offered when field interviewers introduce the study and before conducting the screener (see Section A9). The incentive is provided unconditionally. Instead of converting refusals, it seeks to prevent them at the first point of contact. Households that are contacted and refuse to complete the screener will be contacted again by an interviewer who has received extra training in refusal conversion.

To control costs and maximize the representativeness of the sample, the National Food Study will use two-phase sampling at the screening stage, of addresses where a household member has not been contacted (as discussed in B1). We will implement this strategy with the first batch of sample released (in the first wave). Phase 1 for each batch will include all sampled addresses (within the replicates) assigned to that batch. Field interviewers will visit sampled addresses up to eight times on different days of the week and times of day to attempt contact with a household member.[19] After eight unsuccessful attempts at contact, sampled addresses will be retired from Phase 1 (unless an appointment was made on the 7th or 8th attempt, in which case the interviewer will make up to two more attempts to complete the screener). Phase 2 will include a random sample of addresses retired from Phase 1. This random sample of "hard to reach" cases will receive up to 10 additional field attempts at contact.[20] The size of the Phase 2 sample for each batch will be determined by estimating the maximum number of cases that could be worked to bring the two-phase screening response rate to 80 percent for the completed batch.. The size of the first Phase 2 sample will be based on our best judgment given the Phase 1 results. The sample sizes for subsequent Phase 2 releases will be

---

[19] The National Food Study uses in-person screening and recruitment to maximize response. It is expected that the refusal rate to a proposed one-week study introduced by telephone would be significantly higher than when introduced in person with study materials ready for demonstration.

[20] A random sample will be selected after sorting the retired cases by SSU and strata (SNAP and non-SNAP address).

adjusted based on the results of prior releases. It is not necessary that each Phase 2 sample be of the same size or that they be sampled at the same rate. The Phase 2 sample will be selected after every two batches are completed in the field. For instance, we will sample for Phase 2 from cases retired from Phase 1 of batch 1 and 2 while Phase 1 of batch 3 is in the field. Phase 2 for batches 1 and 2 will be re-released with Phase 1 of batch 4, and so on. We will follow this pattern through the first two waves and make adjustments as needed for the third wave.

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 75 percent of households that complete the screener and are found eligible for the survey. This estimate is based on a 65% participation rate in the field test among the high incentive group, in two PSUs purposively selected to provide challenging survey conditions.

To maximize response throughout the data collection week, respondents will be assigned to a single field interviewer. The field interviewer will conduct the initial screening, train the respondent on survey protocols, conduct the initial household interview at the start of the week, and pick up materials and conduct the final household interview at the end of the week. To the extent possible, a single telephone interviewer will be assigned to conduct the food-away-from-home interviews on days two, five, and seven. During the field 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 primary respondent is asked to complete two household interviews, receive training on food reporting protocols, track all food acquisitions for a week, and report food acquisitions by telephone three times during the week. The survey will provide a base incentive of $100 to primary respondents to encourage participation in the survey. The primary respondent may also receive up to three $10 telephone bonuses for initiating telephone calls to report food acquisitions on days two, five, and seven of the data collection week (averaging 15 minutes each).[21] 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. The base incentive will be provided to the main respondent as a check at the end of the survey week; telephone bonuses will be provided as gift cards at the end of the survey week.

The National Food Study will provide individual food books to all household members age 11 and older so that they may report their own food acquisitions (the Youth Food Book is for children age 11-14; the Adult Food Book is for persons age 15 and older; the Primary Food Book is for the primary respondent). It is very important to maximize response of all household members in order to get a complete account of *household* food acquisitions. To encourage participation by household members other than the primary respondent, the study provides a small incentive for additional members (other than the primary respondent) who report food acquisitions for the week. Findings

---

[21] The survey does not provide separate reimbursement for costs incurred by respondents in placing calls to the survey's toll-free number.

from the cognitive tests indicate that teenagers, in particular, might be reluctant to participate without a targeted incentive. Additional household members who complete food books will receive an incentive of $10 (age 11-14) or $20 (ages 15 and older). Incentives for additional household members will be provided as gift cards at the end of the survey week.

The total incentive available to households depends on household composition, as shown in Table B.4 (same as Table A.1).

Table B.4. Incentive Levels for the National Food Study

| Type of Household | Percent of Population | Average Household Size | Average Number of Additional HH Members Eligible For Incentive | | | | Average Incentive |
			Age <11	Youth, 11-14	Teens, 15-18	Adults	
Single Adult Households							
1 One person household	43.5	1.0	0	0	0	0	$130
2 No youth or teens	18.0	2.8	1.8	0	0	0	$130
3 Youth only	7.0	3.3	1.1	1.2	0	0	$142
4 Teens only	5.3	2.8	0.6	0	1.2	0	$153
5 Youth and teens	3.4	4.3	0.8	1.3	1.2	0	$166
Multiple Adult Households							
6 Adults, no youth or teens	14.6	3.3	1.3	0	0	1.1	$151
7 Adults and youth	3.2	4.8	1.5	1.3	0	1.1	$162
8 Adults and teens	2.7	4.3	0.8	0	1.3	1.2	$181
9 Adults, youth, and teens	2.1	5.8	1.1	1.3	1.3	1.2	$197
Average							**$139**

Note: The average incentive by type of household assumes eligibility for three telephone bonuses.

[a]The distribution of types of households is based on the distribution of the SNAP caseload in 2008 (Source: USDA, FNS. Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2008.

The incentives were tested in the field test conducted from February to May 2011. The field test included two alternative base incentive levels ($50 or $100) randomly assigned to households. The telephone bonus and additional member incentives were identical to the specification for the full-scale survey. The overall response rate for the field test was nine percentage points higher for the high incentive group ($100 base incentive) compared with the low incentive group ($50 base incentive). The nonresponse bias analysis (Appendix V) found that the likelihood of agreeing to participate in the study (among eligible households) and the likelihood of completing the data collection week (among those that started the week) was twice as high for those in the high incentive group. Thus, the higher incentive was adopted for the full-scale survey. The field test was unable to evaluate the effectiveness of the additional household member incentive because there was no variation in incentives within sample.

The first half of the National Food Study field period will include an experiment with households randomly assigned to one of two alternate incentive levels for respondents other than the primary respondent:

1. $10 gift cards to persons age 11 to 14 and $20 gift cards to persons age 15 and older
2. $15 gift cards to persons age 11 to 14 and $25 gift cards to persons age 15 and older

Assignment of households to incentive level will occur when addresses are sampled for the first half of the field period. The sample release for the first half of the field period will be sufficient to obtain 2,500 completes (approximately 12,500 sampled addresses will be released).

Additional methods for maximizing response rates include user-friendly survey materials, including a handheld scanner and colorful, well-designed food books. Field staff will train respondents to use the handheld scanner and food book 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 book. 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. The field test included a Respondent Feedback Form and 70 percent of respondents reported that the survey was easy or very easy; eight percent reported that it was difficult or very difficult; and the remainder reported that it was neither easy nor difficult. 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.[22]

The primary respondent in each sampled household will be asked to report FAFH 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 books, 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 books 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" included 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.

---

[22] During the field test, 90 percent of households reported at least one FAH and one FAFH acquisition. (Appendix B: " FoodAPS Field Test Findings", Memorandum to TWG, page 6).

Maximing Overall Response Rates

The 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.

Nonresponse Bias

The expected response rates for this ABS survey are less than 80 percent. We expect a final completion rate of 55 percent. This response rate is derived by multiplying the expected screener completion rate (80 percent), recruitment rate (75 percent), and food reporting completion rate (90 percent). The expected screener completion rate is based on the results of the field test and the improved protocols planned for the full study (two-phase sampling, $5 token of appreciation at screening). The expected recruitment and final completion rates are based on response to the field test for the high incentive group, and accounting for the characteristics of the field test PSUs.[23]

We examined nonresponse and the potential of bias due to nonresponse to the field test (Appendix V). We compared respondents and nonrespondents at the following stages of response:

1. Screener contact

2. Screener cooperation (among addresses contacted)

3. Agreement to participate (among those screened and eligible)

4. Starting the data collection week (among those agreeing to participate)

5. Completing the data collection week (among those starting)

At the first two stages of response, we compared respondents and nonrespondents by sample characteristics and timing of contact. There was significant variation in response by geography and interviewing team. A pattern of lower cooperation rates in areas with high contact rates (and vice versa) suggests that a higher contact rate throughout all areas would not result in higher response. The planned two-phase sampling for the full-scale study is designed to address the problem of nonresponse at the initial stage by targeting resources (including the most experienced interviewers) to obtain a representative sample of "hard-to-reach" cases in phase 2.

Findings from analysis of response at stages 3 – 5 indicate procedural problems that have been addressed in revised procedures for the full study. Receipt of the study's advance letter was a significant determination of response at each stage. For the field test, these letters were mailed to most of the sample in January but many households were not successfully contacted until March or later. The full study will release sample in smaller batches so that households are contacted soon after the postcard is mailed. The higher incentive was associated with greater likelihood of agreement and completion, and the higher incentive was adopted for the full study.

---

[23] The field test obtained recruitment and final completion rates of 65 and 93 percent, respectively, among the high incentive group.

Examination of nonresponse bias is limited by the nature of the address-based sample: we do not know the characteristics of nonrespondents at the early stages of response.[24] Among those agreeing to participate, the elderly and less educated were less likely to complete the data collection week and this is likely due to the perceived complexity of the data collection for these households. We will modify field procedures for the full study so that field interviewers call these vulnerable households in mid-week and provide technical assistance as needed.[25]

We also assessed the impact of the higher incentive on nonresponse bias. This evaluation is limited because: (1) the incentive level for the main data collection was not known to the respondent until they completed the screener and were determined eligible for the study, thus the incentive level could only affect response among those who completed the screener and were eligible for the main data collection; (2) household and respondent characteristics that are useful for non-response bias analysis were not measured until the first household interview, after the respondent agreed to participate in the study. As recommended by OMB, we use multivariate modeling of response using respondent and nonrespondent frame variables to determine if nonresponse bias exists and varies by incentive level (Guideline 3.2.9); and examination of item nonresponse to determine if it is random, or varies by incentive level (Guideline 3.2.10). The full results are presented in Appendix V. We found that the higher incentive moderates nonresponse bias; household characteristics that have a statistically significant relationship with response, at various stages, are either not significant among the high incentive group, or significantly moderated by the higher incentive.

For the full study, 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. Five methods of nonresponse bias analysis will be used for the full study:

1. Compare the distributions of respondents and nonrespondents, at each stage of response, across subgroups using sample frame characteristics

2. Use multivariate analysis to identify characteristics of cases associated with nonresponse at each stage

---

[24] The screener included a "short form for refusals" to collect information about shopping patterns and household members by age group. These questions were completed by half of those who declined to participate at the time of screening. The questions were not asked of respondents who initially agreed to participate and later declined at the first household interview. Revised procedures for the full survey will ask households who initially agree to participate but then decline to complete the "short form for refusals."

[25] Household are asked to call our telephone center on Days 2, 5, and 7. During these food reporting calls, telephone interviewers assess whether the respondent needs additional assistance or is overwhelmed and in danger of dropping out. Telephone interviewers have the capability to send email direct to the field interviewer for the case, and request a follow-up visit to the household.

3. Compare respondents to known national population characteristics from the American Community Survey[26]

4. Compare the characteristics of easy/early completed cases with the characteristics of later/difficult completed cases (this assumes that nonrespondents are similar to "hard to reach" respondents)

5. Examine the relationship between characteristics associated with nonresponse and study outcomes of interest.

Factors that are associated both with non-response and with study outcomes will be considered of higher priority to use in making non-response adjustments to the weights.

To assess the potential for nonresponse bias in the survey estimates of the SNAP population, we will compare respondents with national estimates of SNAP household characteristics (household size; benefit level; age, race, and gender of the head of household). We will also compare the characteristics of the respondents and nonrespondents on these measures using State administrative data.[27]

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.

B4. Tests of Procedures or Methods to be Undertaken

As part of the field test, the contractor evaluated several components of the sampling design, specifically: the utility of the ABS sampling frame as the non-SNAP frame; the accuracy of the SNAP frame; and the ability of the composite MOS to increase the prevalence of the low-income and very-low-income non-SNAP groups within the sample. The results indicate that the commercial ABS frame was useful: the addresses provided were within the CBGs selected as SSUs; some buildings had to be field listed because no individually identified apartments were on the ABS, but these could be identified from the ABS frame itself; fewer than 3 percent of addresses provided were not residential addresses (no building or non-residential). The SNAP frame was not as accurate as hoped: only 88 percent of SNAP addresses that could be verified were occupied dwellings (versus the anticipated 95 percent) and only 67 percent of households (compared to the 95 percent estimated) identified on the SNAP frame identified themselves as receiving SNAP benefits. On the other hand the composite MOS increased the prevalence of the very-low and low-income non-SNAP households in the sample. The very-low-income non-SNAP group comprises an estimated

---

[26] We did not use this method of analysis for the field test because the field test was not designed to be nationally representative and site-specific benchmark estimates have less precision than national estimates.

[27] This latter analysis will be limited to households providing informed consent for the match with administrative data.

1.0 to 2.0 perecnt of all U.S. households[28], but almost 18 percent of households sampled from the ABS frame and screened were identified as being in that group. This large of an increase is unlikely to be repeated in the main study because the composite MOS for the field test did not include the higher income non-SNAP group.

Prior to the field test, the contractor conducted a cognitive test of the food instruments, including the handheld scanner and food books. Three versions of the food books 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 book; 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).[29] 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.

A full pretest of all survey instruments was conducted with a total of six additional households from the cognitive test convenience sample from June 21 to July 19, 2010.[30] 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 three household

---

[28] Approximately 12 percent of US households were in poverty in 2009 (estimate derived from ACS income and household size estimates), but their SNAP participation rate is estimated to be 89 percent (Leftin, Andrew Gothro, and Esa Eslami, 2010).

[29] 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.

[30] 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.

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 were reported in the Supporting Statement for OMB clearance for the field test.[31]

A field test was conducted from February through May 2011 and with 400 households in two Mid-Atlantic counties. Findings from this field test are described in Appendix B. As noted earlier, the field test obtained lower-than-expected response rates which have led to substantial changes in procedures for the full-scale survey: additional field interview training, redesigned advance mailing (postcard rather than letter), two-phase sampling, a small token of appreciation for potential respondents at screening, and revised refusal conversion procedures for eligible households that decline to participate in the study.[32] The field test found that most households considered the survey procedures easy. Challenges were encountered in matching data reported by different modes (telephone, food book, scanner, and receipt) and system changes will be implemented to better match these data. Findings from the field test also indicated that respondents underreported household income on the screener and were unable to provide complete income data for all household members in the telephone interview. The screener has been revised to ask respondents about sources of income and then remind them of these sources when asking about total household income. The mid-week telephone interview (which asked about income, assets, and non-food expenditures) has been shortened and combined with the final in-person interview for the full survey. In addition, to obtain more complete and accurate income data, respondents will be given an Income and Expenses Worksheet to complete at their convenience during the data collection week so that they can they can consult other household members or documents as necessary.

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

The National Food Study will be administered by ERS's contracting organization, Mathematica Policy Research. Individuals whom ERS consulted or whom we expect to consult on the collection and/or analysis of the data include those listed below.

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

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

Barbara Devaney
Mathematica Policy Research

Diane Herz
Vice President & Director, Washington DC

---

[31] Supporting Justification for OMB Clearance of the Field Test for the National Household Food Acquisition and Purchase Survey (OMB 0536-0066).

[32] Mathematica will also implement system changes to provide better management of the screening effort.

Vice President and Director of Human
Services Research
P.O. Box 2393
Princeton, NJ 08543
(609) 275-2389
bdevaney@mathematica-mpr.com

Survey Research
Mathematica Policy Research
600 Maryland Avenue, Suite 550
Washington, DC 20024-2512
(202) 250-3529
dherz@mathematica-mpr.com

Mary Kay Fox
Mathematica Policy Research, Inc.
955 Massachusetts Avenue, Suite 801
Cambridge, MA 02139
(617) 301-8993

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

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

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

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

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.