

## **PART B.COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS**

### **B.1. Describe (including a numerical estimate) the Potential Respondent Universe and any Sampling or Other Respondent Selection Method to be used.**

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

#### **B.1.1. Respondent Universe**

The Healthy Incentives Pilot (HIP) will be implemented by the Massachusetts Department of Transitional Assistance (DTA) in Hampden County, MA beginning in November 2011 and ending 14 months later in early 2013. Using tested algorithms, we will randomly assign 7,500 of the 53,000 recipients to the treatment group and the remainder to the control group. We will then stratify each group by person-level characteristics such as age, race/ethnicity, gender to ensure that they are balanced. We will randomly select 2,535 individuals (referred to as SNAP participants) from each group

(7500 HIP participants and the remainder of the SNAP caseload not participating in HIP) for survey follow-up<sup>23</sup>. We will then estimate the impact of HIP by comparing regression adjusted outcomes in the treatment group to regression adjusted outcomes in the control group. For this evaluation, the primary and confirmatory outcome is Modified Targeted Fruit and Vegetable Intake (MTFV). We provide more detail on sample sizes, blocking, stratification, variable definition, and analysis below.

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

- SNAP participants;
- SNAP retailers;
- State and local SNAP agency staff;
- State and local partners (including community groups); and
- EBT vendors and third-party transaction processors.

More detail on the respondent universe for each of the above is given below in the context of sampling methods.

## **B.1.2. Sampling Methods**

### ***SNAP Participants***

There are three details of eligibility for the pilot program and the survey that warrant discussion. First, only SNAP households who do their own

---

<sup>23</sup> We expect 20.3% attrition from SNAP between sampling and the start of data collection, yielding 2,020 respondents on SNAP in each group (a total of 4,040 participants, as shown in Exhibit A12.1.). We will attempt to contact these 4,040 respondents to complete the baseline survey.

shopping are eligible for the pilot. For example, SNAP participants who sign over their benefits to a residential or treatment facility are not eligible for the study and will be excluded from the sampling frame (based on identification in the SNAP casefiles). Homeless participants who do not turn over their benefits are eligible for HIP and will remain in the sampling frame.

Second, participants who move to other counties within Massachusetts will remain in the study. Even those that move out of the county will be able to utilize the HIP incentive at large retailers that operate stores outside Hampden County. Participants who move outside the state will be dropped from the study. They are no longer eligible for SNAP in Massachusetts and will no longer appear in the Massachusetts DTA SNAP casefiles.

Third, HIP eligibility is by SNAP case (i.e., a household), but the food intake measure is at the individual level. We will therefore sample adults for the evaluation from the households in the pilot program sample. At Round 2 and Round 3, we will attempt to interview the adult as long as he/she remains in the original SNAP case (or becomes part of another HIP evaluation case).<sup>24</sup> Since SNAP cases are defined by the “head of household,” this means that if the head of household is sampled, we continue to sample him/her as long as he/she remains on SNAP. However, if some other adult is sampled, if he/she goes to another SNAP case, he/she is no longer sampled

---

<sup>24</sup>The SNAP case is tied to the Head of Household (HoH); therefore the HIP incentives will also be tied to the HoH. If the original HoH leaves the SNAP household, by DTA rule that SNAP case closes. Other household members may form a new case, but that new case will not earn HIP incentives. Similarly, if a member of a HIP household other than the original HoH leaves the household, that person will not be eligible to earn HIP incentives; but the household with the original HoH will retain the HIP incentives.

(unless he/she joins a HIP evaluation household). This rule follows the rule for HIP benefits; i.e., they follow the head-of-household (not any individual in the household at random assignment).

**Respondent Universe.** Information on food intake, and in particular the primary outcome MTFV, will be collected from SNAP participants. The Massachusetts Department of Transitional Assistance (DTA) will provide a casefile with all SNAP cases in Hampden County that have an adult, not living in group quarters (the evaluation considers impact on adults, we discuss the details of program and sample inclusion below). That SNAP casefile is expected to have approximately 53,000 SNAP cases. All households in that casefile will be considered to be in the HIP universe for the duration of the pilot.

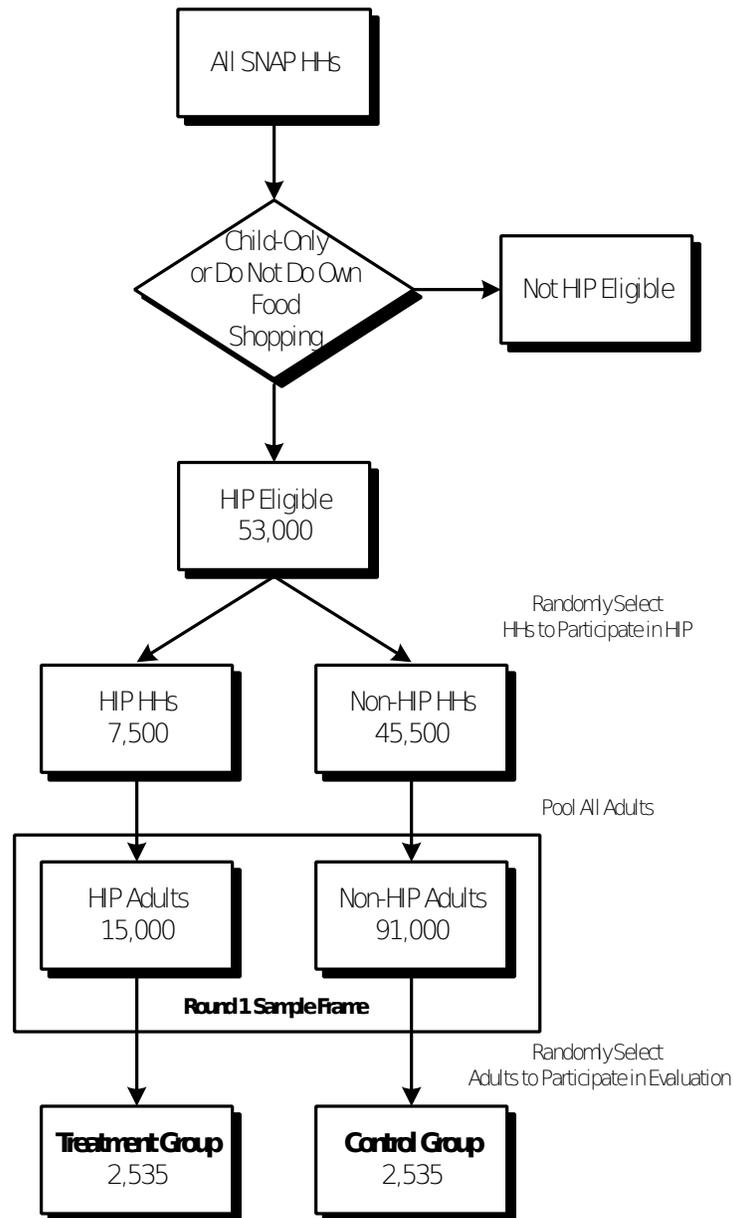
**Selecting the Sample.** We will use the Hampden County SNAP casefile to select the Treatment and Control Groups for the study. The sample will be selected in two phases (see Exhibit B1.1). First, using the SNAP casefile we will randomly assign 7,500 cases to the Treatment Group (i.e., to receive the HIP benefit) and the balance of cases (approximately 45,500) to the Control Group. These groups represent the universe for the evaluation sample.

Second, within these two groups, we will pool all of the adults in the households (SNAP casefiles will provide a list of individuals age 16 and over in each household) to form the sample frame for selecting the adults that will constitute the Treatment and Control Groups for the evaluation. From the

sampling frame for Treatment and Control groups, we will randomly select 2,535 adults for each group. Thus, the Treatment and Control samples will both be random samples from the common population of SNAP households in Hampden County, making them appropriate for comparison as a way of measuring the demonstration's impact without risk of selection bias.

While we will use random assignment to select households to receive HIP and to select adults to participate in the evaluation, we will not use simple random sampling to assign cases to treatment or control. We propose to give each case an equal probability of random assignment; however to assure that the sample of households receiving HIP are as similar as possible to the households not receiving HIP, we will block the list of SNAP households on key household characteristics. That is, we will sort the case records by up to four key characteristics including geography, household size, benefit as a percentage of the maximum benefit, and race/ethnicity. We will use systematic random sampling, selecting every 7<sup>th</sup> household ( $\approx 53,000/7,500$ ) on the blocked list to select the 7,500 that will receive HIP benefits. This procedure ensures that the 7,500 selected households are similar to the remaining 45,500 SNAP cases.

### Exhibit B1.1. Round 1 Sampling Plan



Similarly, we will not use simple random sampling to select the samples of 2,535 adults in each group to participate in the evaluation. Instead, we will use the same systematic random sampling procedure to select the 2,535 adults in each group, selecting every 6<sup>th</sup> adult ( $\approx 15,000/2,535$ ) from the blocked list of 15,000 HIP adults and every 36<sup>th</sup>

adult ( $\approx 91,000/2,535$ ) from the blocked list of 91,000 non-HIP SNAP adults. While this procedure might result in sampling more than one person in the same household, the likelihood of this occurring is minimal because of the way the frame will be sorted prior to the sample selection and it will not affect the generalizability of the sample.

Our decision rule is to interview sample cases only if they are on SNAP at the time of the interview.<sup>25</sup> We expect that some members of the initial sample will exit SNAP between the time the sample is selected and the time of their scheduled interview. Those not on SNAP at the time of their scheduled Round 1 interview will be totally dropped from the sample. Similarly, we expect that some cases that complete the Round 1 interview (baseline) will exit SNAP before their scheduled Round 2 interview, and some that complete the Round 2 interview will exit SNAP before their scheduled Round 3 interview. In addition, some cases that are not on SNAP in Round 2 will be back on SNAP for the Round 3 interview. Thus, the possible response patterns are: completed Rounds 1, 2 and 3; completed Rounds 1 and 2; completed Rounds 1 and 3.

---

<sup>25</sup>The primary impact of SNAP will be through a price effect (the rebate implicitly lowers the price of fruits and vegetables) and an income effect (any rebate earned, even if the absence of a behavioral response). Neither the price effect, nor the income effect, will be operative for those off SNAP. It is possible that short-term participation in SNAP will have effects on tastes for fruit and vegetable intake even after individuals leave SNAP, but any such impacts seem likely to be second order (i.e., quite small). Given that SNAP attrition rates are not low, interviewing those no longer on SNAP will either result in much lower precision (for a given sample size) or much higher costs (for larger samples, in order to maintain a given level of precision). Instead, we will only interview those on SNAP as of each interview. We acknowledge that if HIP causes differential exit from SNAP, failure to interview those who leave SNAP will violate the experimental design. We judge that any differential exit from SNAP is likely to be small, therefore (as confirmed by our Technical Work Group) following those who remain on SNAP is the appropriate design decision. To control for any differential exit from SNAP, we will reweight the survey sample to assure treatment/control balance on observable characteristics. Furthermore, we will model exit rates from SNAP and test for differential exit. We note that SNAP exit is available for the entire evaluation sample (not merely for the smaller survey sample), so these tests will have considerable statistical power.

Our target is to have 750 completes for the Treatment Group and 750 completes for the Control Group in both Round 2 and Round 3. These sample sizes will provide sufficient statistical power to detect meaningful differences in the consumption of MTFVs for the population as a whole, as well as important sub-groups of the population (discussed below). Exhibit B1.2 summarizes the expected sample sizes for the Treatment Group and Control Group under the proposed decision rule for each of the three rounds of surveys based on published national SNAP exit rates (Cody et al., 2007).<sup>26</sup> To minimize data collection costs, we plan to only conduct 750 interviews for each group in Rounds 2 and 3.<sup>27</sup> An estimated SNAP retention rate of 91.6 percent (i.e., an attrition rate of 8.4 percent) and a response rate of 70 percent imply that an initial baseline sample of 2,535 is sufficient to yield 750 completed interviews at Round 2 and Round 3. Specifically, we plan to interview a random sample of the Round 1 sample at Round 2— yielding 750 completes. We will attempt to re-interview all of these Round 2 completes (who are still on SNAP) at Round 3. However, SNAP attrition and Round 3 non-response (among Round 2 respondents) imply that the 750 completes at Round 2 will not be sufficient to yield 750 Round 3 completes. We will therefore draw an additional sample of Round 1 completes to meet the target 750 Round 3 completes. This additional sample will include all of the

---

<sup>26</sup>Massachusetts statewide exit rates provided by DTA are broadly similar to the national rates. According to analysis done by DTA, exit rates in Hampden County are somewhat less than Massachusetts statewide exit rates. This suggests that we will be able to achieve the desired number of completed interviews in Rounds 2 and 3 with the baseline sample size.

<sup>27</sup>A somewhat simpler, but more expensive approach would be to “funnel down” to 750 completes in Round 3. Allowing for SNAP exit rates and response rates this would require at least 949 completed interviews in Round 2.

Round 2 non-interviews, plus a sample of those not selected for Round 2 interviews.

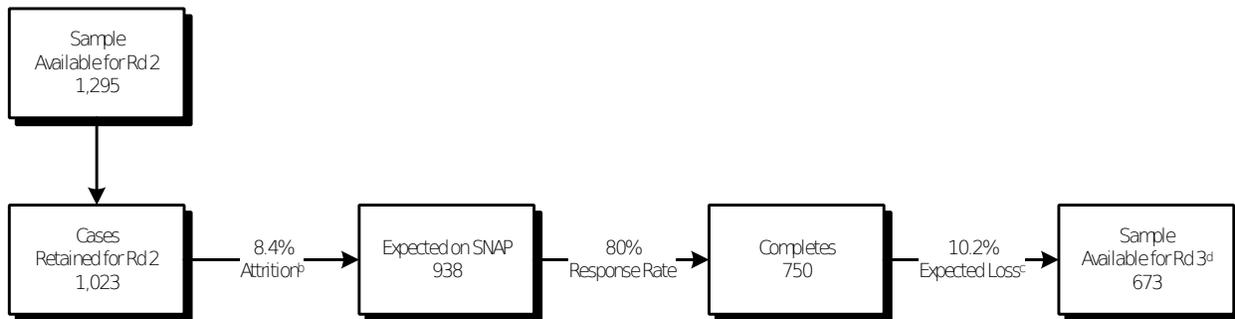
In addition, we will conduct six in-person focus group sessions with HIP participants who are not part of the survey sample. These focus groups will coincide with the Round 2 and Round 3 surveys (three groups at each Round). We will recruit 66 participants (11 per group) with the expectation that about 60 will participate (10 per group) in the discussions (the remaining may be no-shows or cancellations). Appendix D1-D5 presents the procedures for recruiting participants and conducting the focus groups.

## Exhibit B1.2. Expected sample sizes of each group (Treatment and Control)

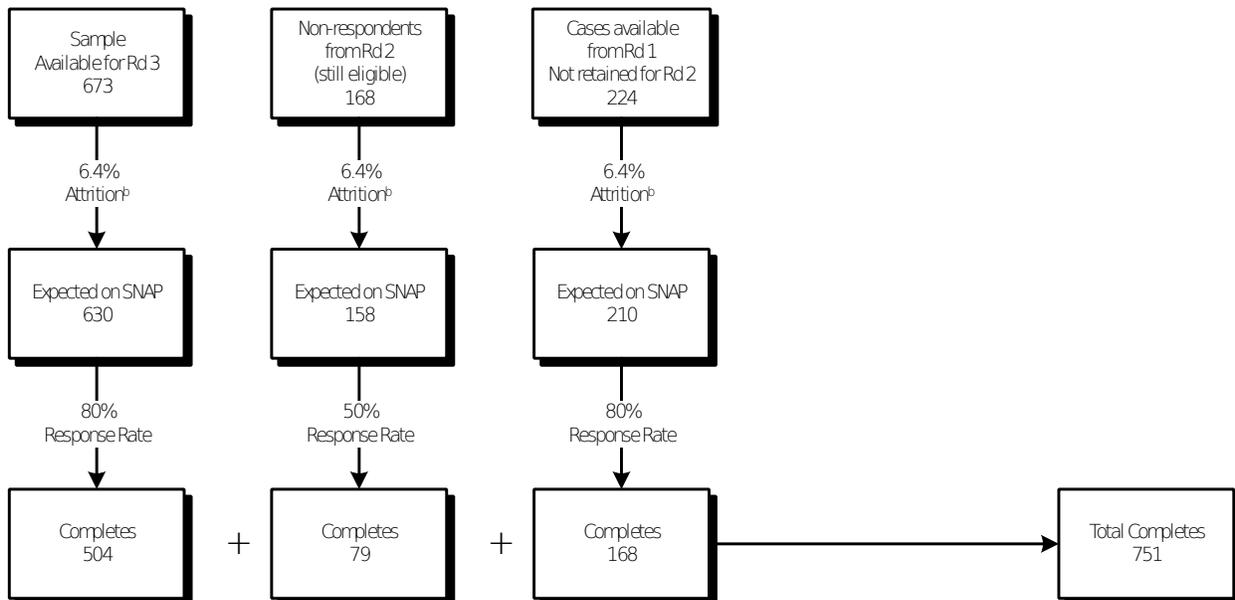
### ROUND 1



### ROUND 2



### ROUND 3



<sup>a</sup> For Round 1, the number of participants in sample is the number of persons to be sampled from SNAP administrative files. By the time data collection for Round 1 starts (approximately 1 month later), an estimated 5.5 percent of the original sample will no longer be in SNAP. Any sampled participants who are determined to be no longer in SNAP will not be fielded in Round 1.

<sup>b</sup> SNAP exit rates derived from hazard rates in Table II.14 of Cody et al., (2007).

<sup>c</sup> At the beginning of the data collection round, we assume that everyone is in SNAP. The exit rates in this column reflect the expected cumulative losses by the end of the data collection period.

<sup>d</sup> Losses determined by matching survey respondents against SNAP administrative files just prior to fielding next round.

## ***SNAP Retailers***

We will also survey SNAP retailers. The samples for the retailer surveys will be drawn from two discrete populations of retailers: those that choose to participate in HIP and those that are eligible to participate but do not. Hampden County has approximately 455 SNAP retailers eligible to participate in HIP. The recruiting process is in progress, so the final number of participating retailers is unknown.

We will stratify stores participating in HIP by store type. FNS has official store types, which we will combine into superstores, supermarkets, small and medium grocery stores, other stores, and farmers markets. Exhibit B1.3 provides the distribution of SNAP retailers in Hampden County and the expected sample size by store type. As indicated, we project that 150 retailers will participate. A survey of all participating retailers would be excessively burdensome. In addition, we expect that responses from stores with the same corporation (superstores, supermarkets, and convenience stores) will not be independent. For these reasons, we will select a sample of 75 participating retailers (one-half of all participating retailers). This sample will be proportionately allocated across strata, and retailers will be randomly selected within strata. With an expected response rate of 80 percent, this sample will yield 60 completed surveys, with at least 10 per stratum (except for farmers' markets, where we will survey all participating markets). While the sample sizes within strata will be too small for statistical inference, they will be sufficient for qualitative analysis. For supermarkets and superstores,

we will select one store from each participating corporation; if necessary, additional stores will be selected in these strata to obtain the desired sample.

Exhibit B1.3. Hampden County SNAP Retailers: Population and Expected Sample Sizes

<b>Store type</b>	<b>Population</b>	<b>Pct. of Population</b>	<b>Expected Participants</b>	<b>Pct. of Partic. Retailers</b>	<b>Partic. Sample*</b>	<b>Non-part. Total</b>	<b>Non-partic. Sample**</b>
Superstore	35	8%	22	15%	11	13	4
Supermarket	19	4%	19	13%	10	0	4
Other grocery	88	19%	44	29%	22	44	4
Convenience/ other store	306	67%	61	41%	30	245	4
Farmer's market	7	2%	4	3%	2	3	3
<b>TOTAL</b>	<b>455</b>		<b>150</b>		<b>75</b>	<b>305</b>	<b>19</b>

\* Participating retailer sample allocated proportionately by strata.

\*\* Non-participating retailer sample allocated equally by strata.

FNS and DTA will provide the evaluation contractor with the official list of retailers participating in SNAP and whether they are participating in HIP. FNS/DTA will provide the retailers files monthly; files from August 2011 and August 2012 will be used to draw the retailer samples. At each round, we will randomly draw a representative sample of 75 retailers participating in HIP. Given the size of the samples relative to the universe, there will be some overlap between the samples at the two rounds. However, we will not attempt a panel design, because we expect that there will be sufficient turnover among smaller stores, and we would therefore need a much larger baseline sample to assure the targeted sample at follow-up. Instead, we have designed the sample to be representative of the participating retailer population at each wave. In addition, this approach allows us to include retailers that do not participate at the start of the demonstration but later choose to participate (“drop-ins”).

The primary focus of the evaluation is on the participating retailers, so a smaller sample of 19 non-participating retailers will be selected in each wave. In the first wave, these will be retailers who decline the initial offer to participate. In the second wave, they will be retailers who dropped out. This total sample size allows for selecting 4 retailers per stratum, except in the farmers' market stratum where the sample will include all non-participating retailers (estimated to be 3).

### ***State and Local Agency SNAP Staff***

Massachusetts DTA staff will assist the evaluation contractor in identifying State-level and county-level staff who will be able to provide needed information on the HIP implementation process and on-going operations. Approximately 19 respondents will be purposively selected to participate in the in-person interviews. The counts of respondents are based on the number of organizational units that are known to be involved in the demonstration.

### ***State and Local Partners (Including Community Organizations)***

DTA is working with a number of community-based organizations to implement and operate HIP. We will work with DTA staff to identify all organizations involved in HIP and, in consultation with DTA, purposively select up to 6 to interview.

## ***EBT Vendor and Third Party Transaction Processors***

ACS, the Massachusetts SNAP EBT vendor has a central role in HIP. We will therefore interview their staff to understand the work involved in implementing and operating HIP. In addition to ACS, four third-party processors are involved in processing SNAP EBT transactions. We will purposively select approximately 9 respondents from across these organizations (ACS and the four third-party processors) and interview them.

### **B.1.3. Response Rates and Non-Response Analysis**

#### ***Survey Response Rates***

**SNAP Participants.** We expect the following response rates from SNAP participants:

- Round 1: 70 percent
- Round 2: 80 percent
- Round 3: a) Round 2 respondents, 80 percent; b) Round 2 non-respondents 50 percent; and c) Round 1 respondents not selected for Round 2, 80 percent.

We will use the American Association of Public Opinion Research (AAPOR) Response Rate 4<sup>28</sup> to calculate the response rates for each round of

---

<sup>28</sup>The AAPOR Response Rate 4 includes both full and partial completes in the numerator (partial completes are instruments where critical items are completed, though the respondent may have broken off before the very end). The AAPOR Response Rate 4 denominator includes full and partial completes, refusals, non-contacts, and an estimate of the proportion of cases of unknown eligibility that are eligible. We anticipate that the number of cases of unknown eligibility will be fairly small, due to the use of monthly case extracts to verify both SNAP receipt and household status. Thus, we will assume conservatively that all unknown eligibility cases are eligible. The impact on response rates will be small given the small number of cases involved. We will exclude ineligible households from the calculation of response rates.

the participant survey. The response rate assumptions reflect our experience on studies that include a diverse samples, including low income households, such as the National Health and Nutrition Examination Survey (NHANES) (response rate of 75% or higher since 1999<sup>29</sup>), NCI Food Attitudes and Behavior Survey (FAB)<sup>30</sup> and the longitudinal Fragile Families and Child Wellbeing study<sup>31</sup> (response rate of about 80% with mothers and 50% with fathers over a four year period; internal communication with Westat project staff). In particular, the NCI Food FAB survey is similar to the HIP survey along a number of dimensions. In the NCI FAB survey, more than 25 percent of the sample was African-American, 55 percent was below \$50,000 and more than 13 percent below \$17,500. The survey included three rounds of telephone data collection to conduct a 24-hour dietary recall; participants were paid \$5 at each round for their time. Twenty-five percent of the sample had bad phone numbers, of the remaining, 64 percent were recruited into the study. Given that we plan to use in-person recruitment to follow-up bad phone numbers and non-response households, and will reimburse participants \$20 for their time, we believe 70 percent is a reasonable response rate for Round 1. The response rate estimates for Rounds 2 and 3 are also based on the FAB data, where 92 percent of the sample completed the second interview and about 82 percent completed the third interview.

---

29<http://www.cdc.gov/nchs/data/bsc/NHANESReviewPanelReportrapril09.pdf>

30Food Attitudes and Behavior (FAB) Study Final Report, Pilot Studies 1 and 2. National Cancer Institute. Westat, May 23, 2007

31<http://www.fragilefamilies.princeton.edu/>

As in those studies, to achieve these response rates we will use standard operational procedures, including in-person recruitment for households with no phone contact information, advance letters explaining the study, and call scheduling to vary the contact times. A detailed list of field procedures to maximize response rates is provided in Section B3.

**SNAP Retailers.** We have assumed an 80 percent response rate for the retailer surveys. This is consistent with response rates achieved on retailer surveys in studies of EBT implementation and operation. We plan to use standard procedures to achieve this response rate, including advance calls and extensive follow-up with respondents. See Section B3 for details of our procedures to maximize response rates.

**Stakeholders (State and Local SNAP Agency Staff, State and Local Partners, EBT Vendors and Third Party Transaction Processors).** We have assumed response rates exceeding 80 percent for all stakeholder interviews. These response rates for key informant interviews are consistent with those attained on other studies examining the implementation and operation of pilot studies.

### ***Analysis for Non-response***

Even with a response rate goal of 80 percent, unit non-response comprises 20 percent of the sample. As OMB has noted (in a presentation made by OMB's Bridget Dooling and Brian Harris-Kojetin on May 27, 2010 at Abt Associates), there may not be a strong relationship between the

response rate in a survey and the magnitude of non-response bias. For a survey that achieves a 60 percent or 70 percent response rate, some substantive variables may be subject to a small degree of non-response bias, while other variables in the same survey may be subject to more substantial non-response bias. The same statement may be made for a survey that achieves a higher response rate of 80 percent.

The real issue relates to how potential non-response bias is being addressed in a study that is seeking OMB clearance. More specifically:

- What information is available from the sampling frame?
- What analyses are being planned and conducted to examine potential non-response bias?
- What adjustments can be made to attempt to reduce bias?
- What special studies are planned to examine potential non-response bias?

For this study, we can assess non-response bias and make adjustments for the participant Round 1 (baseline) survey and the Round 2 and 3 surveys. Our sampling frame of households includes the blocking variables: geography, household size, benefit as a percentage of the maximum benefit, and race/ethnicity. These variables will be available for both the respondent and non-respondent baseline sample households in both the Treatment and Control Groups, and we can compare respondents and non-respondents on these variables. This analysis may identify auxiliary variables to use in post-stratification weighting adjustments to reduce non-response bias.

For the Round 2 and 3 surveys, will have the baseline survey data for the Round 2 and 3 samples, which will allow us to examine the use of response propensity modeling to assess and potentially adjust for unit non-response in these samples. This non-response bias analysis should identify statistically significant predictor variables in a logistic regression model of response in the Round 2 and 3 samples. The predicted probabilities from the model can then be used to form response propensity weighting cells.

## **B.2. Describe the Procedures for the Collection of Information including:**

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

### **B.2.1. Statistical Methodology for Stratification and Sample Selection**

As noted above, we will stratify our sample of SNAP participants to improve precision. Such stratification has some small implications for standard errors. Stratifiers include geography, household size, benefit as a percentage of the maximum benefit, and race/ethnicity. In addition we

expect to stratify the Round 2 and 3 samples based on information obtained in the baseline interviews.

## **B.2.2. Estimation Procedures**

The main objective of the HIP evaluation is to estimate the causal impact of HIP on fruit and vegetable consumption. As discussed in more detail below, the primary measure of fruit and vegetable consumption will be modified target fruit and vegetable (MTFV) intake, averaged over two rounds of the participant survey. Impact will be measured as the HIP/non-HIP difference in this measure, with regression adjustment for selected control variables. Appendix J presents detailed analysis plans for the evaluation.

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

### ***Minimum Detectable Differences (MDDs)***

For this evaluation, FNS' primary outcome of interest is impact on "Targeted Fruits and Vegetables" (TFV); i.e., foods eligible for the incentive. Furthermore, limitations on what can be measured using the 24-hour food intake recall instrument lead us to use Modified Targeted Fruits and Vegetables (MTFV) that can be measured using the 24-hour recall instrument. We define TFV and MTFV as follows:

- TFVs eligible for the financial incentive are the same foods that are allowed by Federal regulations for the WIC fruit and vegetable voucher. These foods include fresh, canned, frozen, and dried fruits

and vegetables without added sugars, fats, oils or salt. Fruit juices, mature legumes, and white potatoes are excluded, but yams and sweet potatoes are included. As with SNAP, the class of foods eligible for HIP also excludes food-away-from-home and hot food served ready to eat.

- MTFV is identical to TFV except that it does not incorporate the restriction against added sugars, fats, oils, and salt. We make this modification because the 24-hour recall records cannot always identify whether such ingredients were included in a purchased product or added later as part of a recipe.

For the entire sample, we estimate that our design has a MDD with respect to MTFV of 0.164 cups targeted fruits and vegetables at either follow-up interview. This power calculation assumes 1,500 completed interviews (750 in the Treatment Group and 750 in the Control Group),  $R^2=12\%$  (a conservative value based on unpublished Westat studies of the predictive power of the Fruit and Vegetable Screener),  $d_{eff}=1.05$  (for non-response), and conventional power parameters  $1-\alpha=80\%$ ;  $\beta=5\%$ , and a one-sided test. A one-sided test is appropriate because we would treat a negative impact as equivalent to no impact.

Furthermore, this is the MDD for a single Round of 24-hour recall of food intake. Our analysis strategy treats the *average impact across Rounds 2 and 3* as the focal outcome. (In the language of multiple hypothesis testing, this outcome will be treated as “confirmatory”; all other outcomes will be treated as “exploratory”.) The MDD for this focal outcome—average impact across the two waves—should be lower (which is better) than the MDD for any single round which we reported above (i.e., 0.164).

How much smaller will depend on the correlation between MTFV intake across the two rounds. If the correlation is zero, the MDE would drop from 0.164 to 0.116. The correlation will be zero for that part of the sample which completes Round 2, but not Round 3, or vice versa. We estimate that group to be about a third of the sample. Furthermore, even for people interviewed at both rounds, given the high day-to-day variability in TFV consumption, it seems likely that the correlation will be low. It therefore seems plausible that the MDD for the average impact across the two rounds will be in the range of 0.125 cups; i.e., an eighth of a cup.

Even the MDD for a single round is well below a plausible estimate of the impact of HIP. Estimates from the 1999-2002 National Health and Nutrition Examination Survey (NHANES) suggest that among near poor households (less than 130 percent of poverty line), average fruit and vegetable purchase is 2.0 cups.<sup>32</sup> Mean estimates of the elasticity of fruit and vegetable consumption with respect to price are 0.70 and 0.58, respectively. Together, these figures suggest an impact of HIP on TFV of about 0.384 cups. Thus, our study's MDD for a single round of 0.164 cups will be able to detect the most likely impacts (0.384 cups) and even impacts that are moderately smaller than this plausible estimate.

For subgroup analyses, the MDDs will be larger because of the smaller sample sizes involved. For example, assuming an R-squared of 12%,

---

<sup>32</sup>Computed from Dong, Diansheng, and Biing-Hwang Lin. *Fruit and Vegetable Consumption by Low-Income Americans: Would a Price Reduction Make a Difference?* 2009. Economic Research Report No. 70, U.S. Department of Agriculture, Economic Research Service, January 2009. <http://www.ers.usda.gov/publications/err70/err70.pdf>. Table 1: 2.0 cups = 1.43 Total Vegetables + 0.96 Total Fruits - 0.39 White Potatoes. This estimate is close, but not ideal. MTFV also excludes fruit juices, consumption away from home, and some fruits and vegetables consumed in combination with other foods.

an impact of  $\frac{1}{4}$  cup of TFV (or smaller) can be detected for several important subgroups: adults in households with children (53% of participants), adults in households without children (47%), and adults age 16 to 59 (85% of participants). For smaller subgroups (e.g., subgroups defined by various levels of baseline fruit and vegetable consumption or older age groups), the MDDs under these assumptions are: slightly above  $\frac{1}{4}$  cup MDD for each of three levels (low/moderate/high) of baseline fruit and vegetable consumption (0.28 cup), and well above the  $\frac{1}{4}$  cup MDD for persons 60 and older (0.43 cup).

### ***Estimation and Calculation of Sampling Errors***

Analysis will proceed using sampling weights. Specifically, we will compute multiple sampling weights, as follows: For each round and for complete cases (i.e., all three rounds), for households and for the sampled SNAP participant. In general, weights are needed in analysis to compensate for differential probabilities of selection and non-response.

**Person Weights.** We will compute person weights as follows. First, we will compute the base weight for a sampled person as the reciprocal of the probability of selecting that person from sampling frames derived from SNAP casefiles. This base weight will differ between treatment and control, but will be common within the treatment and control groups.

To compensate for nonresponse, the base weights will be adjusted within classes determined by the non-response bias analysis. We will conduct a non-response bias analysis to determine characteristics to be used in the weighting (i.e., that are known for respondents and for non-respondents and that are correlated with non-response). For the Round 1 survey, the non-response adjustment cells will be defined using both household-level and person-level characteristics that are available from SNAP casefiles. Within these cells, we will compute a weighted response rate and apply it to the person base weights to obtain the corresponding baseline non-response-adjusted weights. These weights will then be adjusted to account for subsampling prior to fielding in Round 2, and carried over as the “base weights” for Round 2.

To construct appropriate non-response adjustment classes for Round 2 non-response, in addition to the SNAP casefile data used at Round 1, we will use data from the Round 1 survey. Finally, to form the Round 3 adjustment classes, we will carry over the non-response-adjusted weights for Round 2 as the base weights for Round 3, and adjust them for non-response in Round 3 using data collected in Round 2.

**Household Weights.** At each of the three rounds of data collection, the evaluation will collect household-level data from a household member knowledgeable about food purchases and other household characteristics. Because large households (defined by the number of persons 16 years of age or older in the household) will have higher probabilities of

selection than small households, the household sample arising from the proposed sample design is not self-weighting. As described above for the person-level weights, household weights will be produced in a series of steps. The first step will be to create a base weight that reflects the probability of selecting the sampled household. Since household members are sampled at the same rate, the corresponding household weight would be inversely proportional to the number of persons 16 years or older in the household. Next, the household base weights will be adjusted for non-response within adjustment classes that are internally homogeneous with respect to response propensity. For the Round 1 survey, the non-response adjustment classes will be defined using household-level characteristics that are available from SNAP casefiles. For Round 2 non-response, data from the Round 1 survey in addition to administrative data will be used to construct the non-response adjustment classes. Finally, data from Round 2 will be used to form adjustment classes for Round 3 non-response.

**Longitudinal Person Weights.** We project that of the 750 people interviewed at Round 2 and at Round 3, approximately 500 sample persons will complete food intake interviews at both Rounds 2 and 3. (Recall that cases are only interviewed at Round 2 or Round 3 if they complete the Round 1 interview.) To estimate changes between Round 2 and Round 3 among those persons who responded in both Round 2 and Round 3, a separate set of (person-level) “longitudinal” weights will be constructed. These weights will include an adjustment to compensate for the loss of

persons who responded in Round 2, and were still eligible (receiving SNAP), but did not respond in Round 3.

**Replicate Weights for Variance Estimation.** In addition to the full sample weights described above, a series of jackknife replicate weights will be created and attached to each data record for variance estimation purposes. Replication methods provide a relatively simple and robust approach to estimating sampling variances for complex survey data.<sup>33</sup> Under the proposed replication approach, 100 jackknife replicates will be formed by deleting selected cases from the full sample and adjusting the base weights of the retained cases accordingly. The entire weighting process developed for the full sample will then be applied separately to each jackknife replicate resulting in a series of replicate weights. The replicate weights can be imported into variance estimation software (e.g., SAS, SUDAAN, WESVAR) to calculate standard errors of the survey-based estimates. In addition to the replicate weights, stratum and unit codes will also be provided in the data files to permit calculation of standard errors using Taylor series approximations if desired.

**SNAP Retailers.** All SNAP retailers that sell TFVs will have the opportunity to participate in HIP. Participating retailers will represent a broad range of store types, including supermarkets, warehouses/"big box" stores, local groceries, convenience stores, and farmers markets. The objective of the retailer sample is to represent the range of participating retailers. We will

---

<sup>33</sup>Rust, KF, & Rao JNK (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research* 5, 283-310.

stratify stores participating in HIP by store type and sample stores randomly within strata, subject to the constraint of a minimum of one retailer per store type. However, because of the limited sample sizes for the retailer sample, the analysis will largely be qualitative and not require any weighting.

**Stakeholders (State and Local Agency SNAP Staff, State and Local Partners, EBT Vendor/Third Party Processors).** We will purposively select the State and local agency SNAP staff, State and local partners (including community groups), EBT vendor/third party processor staff to participate in in-person interviews. Qualitative data analysis will be undertaken for the data that will be collected from a small sample of stakeholders.

#### **B.2.4. Unusual Problems Requiring Specialized Sampling Procedures**

No specialized sampling procedures are involved.

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

All data collection activities will occur with an 18 month period. The evaluation design requires that respondents be surveyed at multiple times, as described in Section B.1.

### **B.3. Describe Methods to Maximize Response Rates and to Deal with Issues of Non-Response.**

**The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield “reliable” data that can be generalized to the universe studied.**

#### ***SNAP Participants***

Overall response projections were presented earlier. Achieving these response rates involves locating the sample members and securing participation. As discussed above, we estimate that about 70 percent of the HIP participants will complete the Round 1 survey and 80 percent of those who complete Round 1 will complete Round 2 and 3 surveys. Below we describe procedures to be followed to maximize the number of HIP participants who complete the survey:

- Launch a 2-part data collection process that involves contacting individuals by phone and then making in-person field contacts for those who cannot be reached by phone (up to 50 percent of the sample at Baseline, and 10 percent at Rounds 1 and 2). The telephone script is provided in Appendix B7.
- Carefully develop invitation letters for participants that emphasize the importance of this study and how the information will help the Food and Nutrition Service (FNS) to better understand and address current policy issues.
- Use call scheduling procedures that are designed to call numbers at different times of the day (between 8 am and 9 pm) and week

(Sunday through Saturday), to improve the chances of finding a respondent at home.

- Make every reasonable effort to obtain an interview at the initial contact, but allow respondents flexibility in scheduling appointments to be interviewed.
- Conduct silent monitoring of interviews to identify and promptly correct behaviors that could be inviting refusals or otherwise contributing to low cooperation rates.
- Leave a message on voice mail in order to let the respondent know the call was for a research study. The voice mail message script is provided in Appendix B8.
- Provide a toll-free number for respondents to call to verify the study's legitimacy or to ask other questions about the study.
- Require up to 9 unsuccessful call attempts to a number without reaching someone before considering whether to treat the case as "unable to contact."
- Implement refusal conversion efforts for first-time refusals and use interviewers who are skilled at refusal conversion and will not unduly pressure the respondent (Appendix B9).
- Implement standardized training for all data collectors. The interviewer training will focus on basic skills of telephone interviewing, use of CATI platforms for interviews. The training will also include self-paced training materials to familiarize the data collectors with the study backgrounds and questionnaires. It will focus on gaining participant cooperation, questionnaire delivery, accurate coding, effective neutral probing, and appropriate contact procedures. Data collectors will also conduct live interviews with each other; these interviews will be monitored by experienced supervisors; data collectors may receive further coaching and evaluation or will be replaced if they are not comfortable with the data collection instrument and/or procedures.

## ***SNAP Retailers***

By carefully and convincingly explaining the importance and potential usefulness of the study findings in the introductory letters from FNS

and DTA, and by implementing a series of follow-up reminders (Appendix E9) and offers to complete the survey by telephone, we expect to achieve an overall survey response rate of 80 percent for participating retailers. Specific procedures to maximize response rates include:

- Initial notification of selected retailers by telephone to describe the study and solicit participation on behalf of FNS and DTA (Appendix E10).
- Survey mailing that includes letters from FNS and DTA (with HIP logo for recognition).
- Availability of technical assistance through a toll-free telephone number for respondents.
- Follow-up procedures that will monitor return rates and make reminder calls to non-responders and collect data by phone, if necessary.

Prior to participating in the HIP, retailers will be required to sign a letter or memorandum of understanding (MOU) with the DTA, agreeing to comply with the terms of the MOU and adhere to the procedures specified by FNS. DTA is developing a three-party memorandum of understanding among DTA, ACS and retailers. The MOU will include language about evaluation participation requirements, which will contribute to maximizing response rates.

Securing the cooperation of the non-participating retailers and those who dropped out of the pilot may be more challenging than for participating retailers. We will encourage non-participating retailers or those who dropped out of the pilot to cooperate by noting that if a HIP-like program was to be implemented nationwide, retailers might see an increase in sales

of fruits and vegetables that earned SNAP participants an incentive. We would also appeal to their interest in improving the health and nutrition of Americans participating in SNAP. In addition, we will also appeal to their desire to be heard about the barriers to participating in HIP. Finally, we will offer a modest \$40 monetary incentive to the non-participating retailers for completing a survey.

***Stakeholders (State and Local Agency SNAP Staff, State and Local Partners, EBT Vendor/Third Party Processors***

We will work with DTA to identify the most appropriate individuals to participate in the stakeholder interviews. DTA will also assist in scheduling interviews. These efforts by DTA will ensure that we are able to achieve projected response rates of 83-89 percent.

#### **B.4. Describe any Test of Procedures or Methods to be Undertaken.**

**Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.**

#### ***SNAP Participants***

Two methodological tests were conducted simultaneously to cognitively evaluate two instruments – the English Participant/Primary Shopper survey and the Spanish Participant/Primary Shopper survey. The English instrument was cognitively tested to identify problems with question wording and the flow of the interview. The Spanish instrument was cognitively tested to evaluate whether Spanish-speaking respondents understood the translated questions the way they were intended in English. No more than 9 respondents were asked the same question. Both tests involved 2 rounds; the English test included 9 participants in Round 1 and 4 in Round 2, while the Spanish test included 9 in Round 1 and 5 in Round 2. In Round 1, all newly developed questions in the English and Spanish surveys were cognitively tested; in Round 2, we tested the complete interview flow, as well as new questions that were developed as a result of findings from Round 1. We also derived the total interview administration time in Round

2. The final instruments were revised based on the findings from the two rounds of cognitive testing; Appendix B1-B6 include the revised Round-specific English and Spanish Participant surveys that will be used in the evaluation.

## ***SNAP Retailers***

With the assistance of DTA and FNS, the evaluation contractor recruited 9 stores to pretest items from the retailer survey that do not refer specifically to HIP. These stores included three supermarket chains, two convenience store chains, and four small/medium grocery stores. Abt sent the pretest version of the retailer survey to the 9 stores, with a cover letter explaining the pretest, letters from DTA and FNS explaining HIP and the importance of the survey, a form for comments on the survey, and a prepaid FedEx form and envelope to return the survey. Store managers/owners provided the start and end times for each section, circled confusing language or formatting, and provided feedback on the comments form after completion of the survey. Telephone debriefings were conducted with the retailers to determine any unclear or difficult questions, missing questions, or other recommended changes.

The observation form was pretested on-site in 3 stores – one supermarket, one chain and one independent store. For survey and store observation questions referring to HIP, and thus not included in the pretest, we reviewed the content and wording with DTA, retailer corporate contacts

and retailer association representatives. Instruments were revised to reflect the comments from the pretest and other reviews.

***Stakeholders (State and Local Agency SNAP Staff, State and Local Partners, EBT Vendor/Third Party Processors***

DTA reviewed all stakeholder interview guides to ensure that they were consistent with plans for HIP implementation and operation, and with terminology used by DTA and its partners.

**B.5. Provide the Name and Telephone Number of Individuals Consulted on Statistical Aspects of the Design and the Name of the Agency, Unit, Contractor(s), Grantee(s), or Other Person(s) Who Will Actually Collect and/or Analyze the Information for the Agency.**

<b>Name</b>	<b>Affiliation</b>	<b>Telephone Number</b>	<b>e-mail</b>
Susan Bartlett	Project Director, Abt Associates Inc.	617-349-2799	Susan_bartlett@abtassoc.com
Parke Wilde	Director of Design, Abt Associates Inc.	339-368-2975	Parke_wilde@abtassoc.com
Jacob Klerman	Director of Analysis, Abt Associates Inc.	617-520-2613	Jacob_klerman@abtassoc.com
Susie McNutt	Project Director, Westat	301-251-3554	SusieMcNutt@westat.com
Adam Chu	Associate Director, Statistics, Westat	301-251-4326	CHUA1@westat.com
Janet Tooze	Wake Forest University	336-716-3833	jtooze@wfubmc.edu
Frances Thompson	Applied Research Program, National Cancer Institute	301-435-4410	thompsonf@mail.nih.gov