Part B: Statistical Methods

Part B of the Justification for this information collection activity, the *Food and Nutrition Service Evaluation of the Fresh Fruit and Vegetable Program (FFVP)*, addresses the five points outlined in Part B of the OMB guidelines.

In designing this study, we have dual goals. First, the FFVP authorizing legislation calls for an impact evaluation. By an impact evaluation we mean an estimate of outcomes with FFVP relative to what outcomes would have been without FFVP (for the same units, in the same time period). Second, in addition to the impact study required by statute, we are interested in an implementation analysis to understand how the FFVP is implemented nationally.

As we discuss in detail below, no national list of schools participating in FFVP exists. Instead, to generate samples of schools for either the impact or the implementation analysis, we must contact the States, secure their lists of participating schools, and process that data. This process has moderate cost per State and evaluation resources are limited. We therefore chose to use a common sample of States for both the impact and implementation analysis. After computations discussed in detail below, we decided to use 16 States—stratified by the four Census regions. We choose States using probability proportional to size (PPS) sampling, and as a result, the included 16 States account for approximately 72 percent of the total number of public elementary school students in schools where at least 50 percent of students are approved for free/reduced price meals.

Our primary concern for the impact analysis is internal validity. Schools that apply for FFVP funds are likely to be systematically different from schools that do not apply (e.g., in their initiative, creativity, and openness to new ideas). Among schools that apply, poor schools (defined as schools where at least 50 percent of students are approved for free/reduced price lunches) are more likely to be selected. Free/reduced price percentage is the primary selection criterion according to the statute and to FNS guidance. It seems plausible that both of these factors would be correlated with nutritional patterns, even in the absence of FFVP. Thus, finding a control group sufficient to yield convincing estimates of causal impact is challenging.

Random assignment was not an option for the evaluation since the statute requires awarding FFVP funds to the poorest schools. We considered difference-of-difference (i.e., pre/post with an unaffected control group) and regression discontinuity (RD) designs. Given the project time line, difference-of-differences was not feasible. Furthermore, RD is widely believed to have better internal validity. Both of those considerations suggested that RD was the most appropriate design for this evaluation.

The basic idea of RD is that schools near the cutoff (i.e., those who were just barely selected to participate in the program and those who were just barely not selected) are nearly identical. In most States, the FFVP selection rule is based on the percentage of students eligible to receive free/reduced price meals (FRSL). If we were to compare outcomes in *all* schools above the FRSL cutoff for participation in FFVP to those in *all* schools below the FRSL cutoff who were not selected to participate, we would expect the resulting impact estimates to be biased, since nutrition outcomes are likely to be worse on average in schools with high percentages of FRSL students. However, we would not expect that correlation to be very strong in the narrow range on either side of the FRSL cutoff covered by our impact sample; for example, one would not expect *a priori* that nutrition outcomes would differ appreciably in a school with 57 percent of students eligible to receive FRSL as compared to a school with 56 percent of students eligible to receive FRSL. To the extent that each school's percentage of FRSL-eligible students is measured imprecisely, the case for RD is even stronger; for example, schools are selected based on FRSL in the prior year. FRSL varies enough from year-to-year that some schools chosen (i.e., above the cutoff based on prior year FRSL) would actually be below the cutoff based on the current year FRSL (and vice versa). This chattering around the cutoff adds to the plausibility of treating schools just on either side of the line as essentially identical (except for FFVP participation). To address any remaining concern, for the impact analysis, we sample the schools that are as close as possible to the cutoff point— and therefore as similar as possible on the selection rule variable *while still differing in their FFVP participation status*. Larger States will have more schools within a given absolute range of the FRSL cutoff point. We therefore allocate impact sample of schools to the States in proportion to their size as

measured by the total number of public elementary school students in schools where at least 50 percent of students are eligible for free or reduced price meals.

We acknowledge that this focus on internal validity has a cost. RD only estimates the impact at the point of discontinuity. Impacts away from the point of discontinuity will have no effect on our estimate of the program. There is no evidence of impact heterogeneity for FFVP (but, of course, no evidence against impact heterogeneity) and it is our reading of the social science literature that evidence for program impact heterogeneity globally is weak.

On the other hand, there is overwhelming evidence of selection bias that would threaten internal validity. We have therefore focused the design of our impact analysis to minimize threats to internal validity.
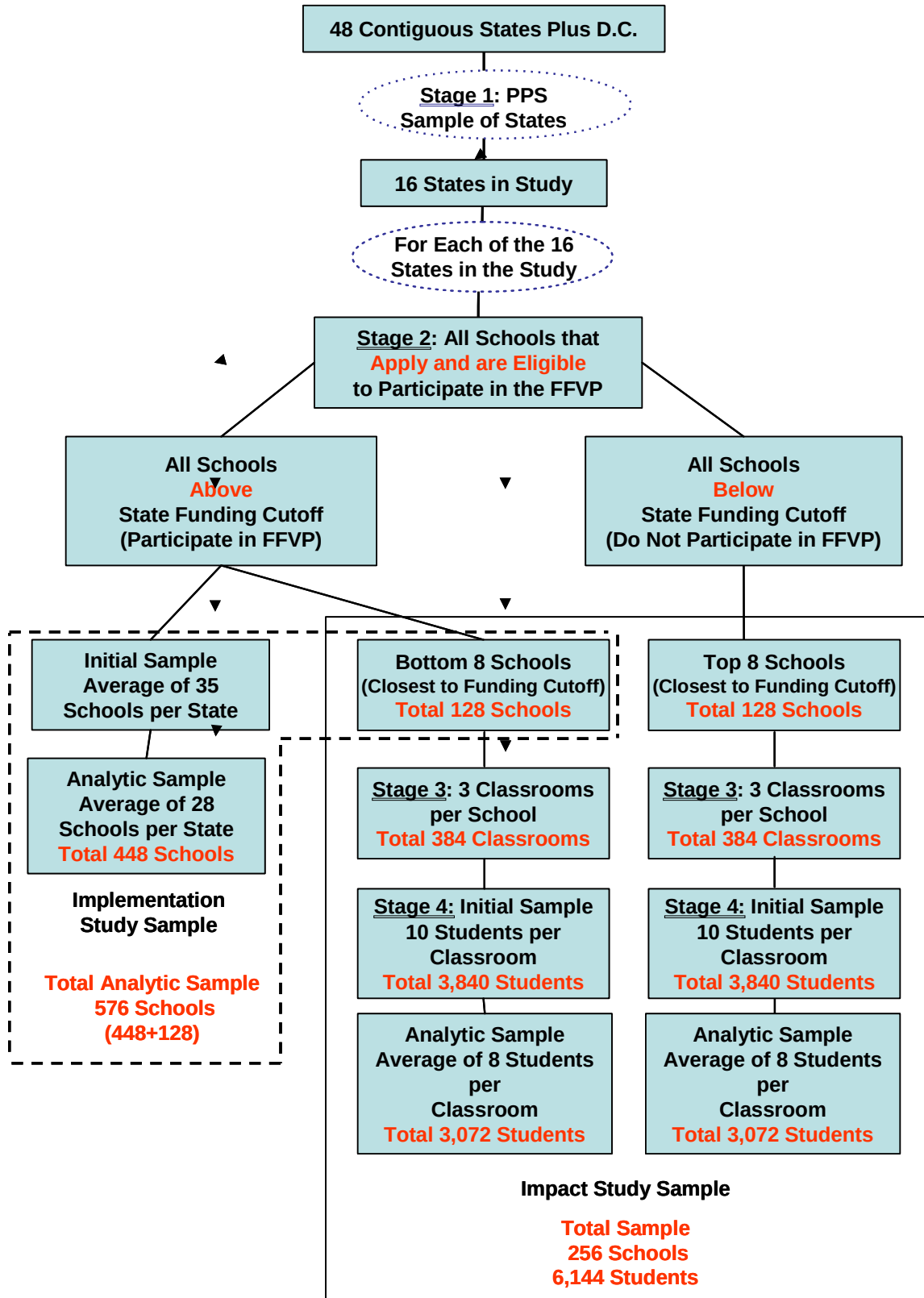
To conduct the impact evaluation, we need to collect information nearly equivalent to the information collected in the implementation study. Once we have gained cooperation from the schools for impact analysis data collection, it will be relatively inexpensive to collect the implementation data from all of the impact sample schools. As discussed in more detail below, we form a second stratum of participating schools not in the impact analysis sample. From that second stratum, we randomly select other participating schools. We then combine the two strata using standard methods.

B.1 RESPONDENT UNIVERSE AND SAMPLING METHODS

The sampling plan for the study will produce analytic samples for each of the two study components: the Impact Study and the Implementation Study. Exhibit 3 provides an overview of our proposed sampling plan. The Impact Study requires samples of students in elementary schools that participate in the FFVP and *eligible* schools that do not participate in the program. The Impact Study will use a regression discontinuity (RD) design, and appropriate statistical methods, which will produce unbiased estimates of the impact of the program. Expected response rate for the data collection as a whole is estimated at 89 percent.

The RD design is particularly well suited to the evaluation of the FFVP. The FFVP legislation requires that to "the maximum extent practicable," States give FFVP funding to the poorest schools, as

Exhibit 3: Four-Stage Sample Design—Impact Study and Implementation Study

48 Contiguous States Plus D.C.

Stage 1: PPS Sample of States

16 States in Study

For Each of the 16 States in the Study

Stage 2: All Schools that Apply and are Eligible to Participate in the FFVP

All Schools Above State Funding Cutoff (Participate in FFVP)

All Schools Below State Funding Cutoff (Do Not Participate in FFVP)

Initial Sample Average of 35 Schools per State

Bottom 8 Schools (Closest to Funding Cutoff) Total 128 Schools

Top 8 Schools (Closest to Funding Cutoff) Total 128 Schools

Analytic Sample Average of 28 Schools per State Total 448 Schools

Implementation Study Sample

Total Analytic Sample 576 Schools (448+128)

Stage 3: 3 Classrooms per School Total 384 Classrooms

Stage 3: 3 Classrooms per School Total 384 Classrooms

Stage 4: Initial Sample 10 Students per Classroom Total 3,840 Students

Stage 4: Initial Sample 10 Students per Classroom Total 3,840 Students

Analytic Sample Average of 8 Students per Classroom Total 3,072 Students

Analytic Sample Average of 8 Students per Classroom Total 3,072 Students

Impact Study Sample

Total Sample 256 Schools 6,144 Students

measured by the percent of students eligible for free and reduced price school meals. In selecting the sample for the RD design, all eligible schools are arrayed from highest to lowest along the selection criterion dimension (e.g., percent students eligible for free/reduced price meals). The RD sample includes schools around the State funding cutoff so that schools just above the funding cutoff are participating in the FFVP and schools just below the cutoff, while eligible for the program, are not participating due to funding limitations. RD estimates the causal impact of the FFVP by comparing eligible schools directly above and below the cut-off for funding. The last school to get the FFVP and the first school not to get the FFVP differ only very slightly along the selection criterion dimension (e.g., percent of students eligible for free and reduced price school meals). Within this narrow range we can view which schools got the FFVP as approximately random. Recent developments in methods for estimating causal impacts suggest that such RD designs are the strongest possible designs, when random assignment is not feasible (see: Imbens, G.W., and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics, 142*(2), 615-635).

However, the nature of the RD design is such that the schools (and students) included in the sample are not nationally representative. By contrast, the Implementation Study is intended to provide detailed information on how the FFVP is implemented in participating schools across the country, which requires a national probability sample of participating schools. For the Implementation Study, the sample of *participating* schools used in the Impact Study will be supplemented with a larger randomly selected sample of participating schools that were not included in the Impact study. A probability proportional to size sample of 16 States will be drawn that will serve as the first stage sample for both the Impact and Implementation Studies.

Impact Study Sample

Within each of the 16 States, *eligible* elementary schools will be sampled in order to end up with an average of 16 schools per State (8 above the State cutoff and 8 below the State cutoff), for a total of 256 eligible schools (128 participating in the FFVP and 128 not participating in the FFVP). Elementary schools in the U.S. have different grade ranges (e.g., K-4, K-5, K-6, etc.) The impact study is targeted to

older elementary school children because of the need to focus on the fruit and vegetable consumption of children as they transition into higher grade levels and because some aspects of the data collection are not suited to younger children. Thus an eligible elementary school must have at least one eligible grade present (i.e., K-3 elementary schools are not eligible).

From each of the 256 eligible schools we will sample one classroom from each of grades 4, 5 and 6 for those schools that have all three grades present. For K-5 and K-4 schools we will sample from the eligible grades that are present since our primary interest is in having a sample of older elementary school children defined as children in grades 4 to 6. This will yield a total sample of 768 classrooms (384 participating and 384 nonparticipating). At the final stage of sampling we will select a random sample of approximately 10 students per classroom. This will result in an initial sample size of 30 students per school, and a total of 7,680 students (3,840 participating and 3,840 nonparticipating) from the 256 schools. We expect that the initial sample of approximately 10 students from each classroom will allow us to obtain around 8 students who participate in the Impact Study. Thus, the total analytic sample size for the Impact Study is expected to be around 24 per school or around 6,144 (3,072 participating and 3,072 nonparticipating).

Implementation Study Sample

As previously noted, the regression discontinuity sample for the Impact Study does not yield a random sample of all FFVP schools. Therefore, within each of the 16 States we plan to create two strata of schools participating in the FFVP: 1) the schools sampled for the regression discontinuity portion of the study that are above the State-specific cut-off; and 2) all remaining schools in the State that participate in the FFVP. Across the 16 States the first stratum contains 128 participating schools. For the second stratum we plan to select an initial probability proportional to size (PPS) sample of around 35 schools per State for a total of 560 participating schools. Allowing for refusals, we expect to end up with a mean of around 28 participating schools per State for a total of 448 that respond to the school survey. These 448 schools will not participate in any of the other aspects of the study. Thus, from the 16 States we should

have a total of 576 FFVP participating schools with school survey data that will allow us to make national estimates regarding FFVP implementation.

B.2   PROCEDURES FOR THE COLLECTION OF INFORMATION

Procedures for the collection of information addressed below include:

- 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.

Statistical Methodology for Stratification and Sample Selection

The evaluation design involves a four-stage sample design which reflects the structure and administration of the FFVP. Stages of sampling include: States, schools, classrooms and students.

State selection: Because States vary considerably in size we will use probability proportional to size (PPS) sampling to select States. In PPS sampling larger States are given a higher probability of selection than the more numerous smaller States. From the National Center for Health Statistics Common Core of Data Public Elementary/Secondary School Universe Survey: School Year 2006-07 we formed a measure of size for each State equal to the number of public elementary school students in schools where at least 50 percent of students receive free or reduced price meals. The sum of the State measures of size equals 11,603,861 elementary school students from a total of 26,257 eligible elementary schools. The South Census Region accounts for 44.2% of the 26,257 schools while the Northeast Region only accounts for 11.1%. The median percentage of students approved for free or reduced price meals in these schools is 72.7% (range: 50% to 100%). The median student enrollment is 415 students (range: 4 to 5,944). Median Hispanic student enrollment is 12.5% but ranges from 0% to 100%.

Based on the sampling of 16 States we identified three very large States that would enter a PPS sample with certainty: California, Texas, and Florida. These three States will be included in the first stage sample with a probability of one. The remaining 46 noncertainty States will be grouped into Census

region strata to ensure a good geographic spread of the sample. Within each Census region the States will be ordered by the percentage of students in the State that are non-Hispanic White. Ordering States within Census Region on this variable will ensure that the sample of States gives good representation to states with a substantial minority student enrollment. The Census Region State sample sizes are 2 in the Northeast, 3 in the Midwest, 6 in the South, and 2 in the West. The sample of 13 noncertainty States will be drawn using PPS sequential sampling. This PPS sampling method takes advantage of the ordering of the States within each Census Region on the race/ethnicity variable to help ensure a good distribution of the noncertainty States in terms of the percentage of elementary school students who are non-Hispanic white. It is possible that a sample noncertainty State will use an FFVP school application and selection process that is not compatible with the regression discontinuity technique. If that occurs we will draw a different State from the same stratum using PPS sampling. A first stage sample of 16 states was determined to be appropriate for this study based on three criteria: 1) there is a moderate cost and time associated with obtaining a list of participating schools from a state, 2) we expect the intracluster correlation for schools within states to be fairly low, around 0.01, because actual implementation of the program is determined primarily at the school district and school level and much less at the state level, and 3) the 16 sample states will account for around 72% of the total number of public elementary school students in the U.S. in schools where at least 50 percent of students are approved for free or reduced price meals.

School selection: The sample of schools will consist of 128 schools above the State-specific cut-offs and 128 schools below the State-specific cut-offs. The number of sample schools above and below the cut-off in a given State is m_h where h references State. The mean value of m_h is 8. For our design, we are not selecting schools with probabilities proportional to size. Rather, the m_h schools above and below the cut-off are selected with a probability of one. For this type of school sample design, it is preferable to allocate the sample of 256 schools to the 13 noncertainty States proportional to the selection probability of the State. This is the best sample allocation for the regression discontinuity design. For example, the State of California which accounts for 16.3 percent of the total measure of size would be allocated 41.8

schools (20.9 above and 20.9 below the cutoff) = 16.3 percent of the 256 schools. Small noncertainty States entering the sample with a low probability of selection will receive a smaller allocation of sample schools. We will set the minimum State allocation to a sample size of a total of 2 schools above the cut-off and two schools below the cut-off.

Selecting classrooms within each school: As data will be collected from elementary school students in grades 4 through 6, schools will be asked to provide us with a list of classrooms in these grades. The impact study is targeted to older elementary school children because of the need to focus on the fruit and vegetable consumption of children as they transition into higher grade levels and because some aspects of the data collection are not suited to younger children.

One classroom in each of these three grades (4, 5, and 6) will then be randomly selected, and students in the selected classrooms will be rostered. Eligible elementary schools in the U.S. do however have different grade ranges (e.g., K-4, K-5, K-6, etc.) In schools that do not include one or two of the three eligible grades, we will select three classrooms from among the eligible (i.e., 4 to 6) grades present.

Selecting students within classrooms: An initial sample of about 30 students will be selected from each school in the evaluation. This initial sample assumes approximately 20 percent of the students will either not provide parental consent or be absent on the day of data collection in each school. This will provide an analytic sample of approximately 24 students per school.

The initial sample of students within each school will be a stratified cluster sample of students in grades 4-6. That is, the student sample will be stratified by grade level, and to maximize the efficiency of data collection will be clustered within classroom. Within each of the three sampled classrooms we will randomly select an initial sample of about 10 students.

Supplementary Sample of Schools

To describe implementation of the FFVP, we will conduct a web survey of schools participating in the FFVP. For this survey, we will retain all 128 schools in our initial sample of schools that are above the State-specific cut-off (the impact treatment group), and we will draw a supplementary probability proportional to size sample of 560 schools which have FFVP funding and are not in our regression

discontinuity sample. The measure of size for the PPS selection of around 35 participating schools per state is the total the number of students in the school. We project an 80 percent response rate for this stratum, so this will yield a final sample of 448 schools in the random sample, plus the 128 treatment schools in the regression discontinuity sample. Thus 78% of the implementation sample comes from the supplementary sample. Since every State had a positive probability of inclusion in the sample of States and every FFVP participating school has a positive probability of being included in this web survey sample, the resulting sample is a national probability sample of all FFVP schools in the U.S. in SY 2009-2010.

Estimation Procedures

Impact study: For measuring the impacts of FFVP on student outcomes, we will use econometric models that incorporate our regression discontinuity design. One consistent estimate of the impact of FFVP will be to compare mean outcomes for students in treatment and comparison schools. As with random assignment studies, however, we can increase our power and decrease our standard errors by modeling how the outcomes vary with observed covariates at the student level and school level. Covariates available for the students in the impact sample will include student gender, grade, race/ethnicity, and school meals eligibility status (Paid, Reduced, Free), and impact sample school Census Region, urban status, grade range, racial/ethnic composition, and percent free/reduced price school meals. We will account for the clustering of students in schools and other features of the sample structure by using SAS PROC SURVEYREG or SUDAAN for computation of appropriate standard errors, confidence intervals, and test statistics.

Implementation study: Sampling weights for the Implementation Study schools will be calculated as the reciprocal of the product of the selection probability of the State and the selection probability of the school within the State. Within each State the FFVP participating schools will be sampled from two strata, as described above. We will examine patterns of nonresponse at the school level and determine if any sampling frame variables should be used to form nonresponse adjustment cells.

It is possible that larger schools will implement FFVP differently than smaller schools. We will therefore develop two school-level weights corresponding respectively to FFVP schools, and FFVP schools times the number of students attending these schools. These will enable us to make statements about the prevalence of particular approaches to FFVP implementation (a) in terms of the percentage of FFVP schools that use an approach, and more importantly (b) in terms of the percentage of students in schools that use the approach.

Degree of Accuracy

Impact Study: For assessing statistical power, we take as our focal outcomes the fractions of children satisfying the MyPyramid 1-day recommendations for fruits and vegetables (1 ½ cups of fruit for boys and girls 9-13 years old; 2 cups vegetables for girls and 2 ½ cups vegetables for boys aged 9-13). Our power calculations incorporate assumptions about the levels of these outcomes taken from estimates in the peer-reviewed literature; the most currently-available estimates are for children aged 6-11, which we assume will be fairly similar to levels in our target population aged 9-13 years. 24% of children 6-11 met the MyPyramid 1-day recommendation for fruit and 16% met the recommendation for vegetables. The mean cup equivalents of fruits for children age 6-11 is 0.99 with a standard deviation of 2.14. For vegetables the mean cup equivalents is 0.98 with a standard deviation of 1.29.[1]

In terms of the percentage of children meeting the MyPyramid 1-day recommendations, our samples will provide 80 percent power to detect impacts of 4.1 to 5.9 percentage points for fruits, and 3.6 to 5.2 percentage points for vegetables. The sample sizes also allow for the detection of a difference of 0.20 to 0.29 cup equivalents for fruits, assuming a mean of 0.99 in the comparison sample, and a difference of 0.12 to 0.17 cup equivalents for vegetables, assuming a mean of 0.98 in the comparison sample. We consider this to be more than sufficient power. In fact, it allows some power against large subgroups. For example, the Minimum Detectable Effect (MDE), with respect to the percentage of children meeting the MyPyramid 1-day recommendations, for a subgroup that is about a third of the

---

[1]   B. A. Lorson, H. R. Melgar-Quinonez, and C. Taylor, Correlates of fruit and vegetable intake in U.S. children, *Journal of the American Dietetic Association*, 109(3), 474-478, 2009.

population (e.g., whites vs. non-Hispanic blacks) would be about 6.4 percentage points for fruits, 5.6 percentage points for vegetables. Also, the 33% subgroup sample sizes allow for the detection, with 80% power, of a difference of 0.31 cup equivalents for fruits, assuming a mean of 0.99 in the comparison sample, and a difference of 0.19 cup equivalents for vegetables, assuming a mean of 0.98 in the comparison sample.

Implementation Study: The primary purpose of the Implementation Study is to provide national estimates of program implementation procedures by FFVP participating schools. Most of these estimates will be descriptive in nature, consisting primarily of proportions. The expected sizes of the 95-percent confidence interval half-widths are shown in the table below. We anticipate a design effect for many school estimates around 1.35 given that we expect the intracluster correlation to be around 0.01 for many variables.

| | Percentage: | | | | |
	10% or 90%	20% or 80%	30% or 70%	40% or 60%	50%
95-percent CI half-width (±)	2.8%	3.8%	4.3%	4.6%	4.7%

Unusual Problems Requiring Specialized Sampling Procedures

No specialized sampling procedures are involved other than the regression discontinuity design is used to define the two strata of schools within each sample State.

Use of Periodic Data Collection Cycles to Reduce Burden

This is a one-time survey data collection effort.

B.3   METHODS TO MAXIMIZE RESPONSE RATES AND DEAL WITH NONRESPONSE

It is expected that a minimum of 80 percent of the sample schools will participate in the Impact Study and/or Implementation Study. This is a very reasonable estimate given the 93 percent response rate for SNDA-III schools in the student-level data collection.[2] For the sample of students it is expected that

---

[2]   U.S. Department of Agriculture, Food and Nutrition Service, Office of Research, Nutrition and Analysis, *School Nutrition Dietary Assessment Study-III: Vol. II: Student Participation and Dietary Intakes,* by Anne Gordon, et al. Project Officer: Patricia McKinney. Alexandria, VA: 2007.

the intensive efforts to gain participation will yield an 80 percent response rate. The response rate among students sampled for dietary recalls reported in the SNDA-III data collection was 63 percent. Although this is somewhat lower than our estimate, we believe that differences in the methodology used for the FFVP evaluation (e.g., not requiring a follow-up meeting with parents within 48 hours to complete the dietary recall, and more rigorous procedures for ensuring that consent forms are returned) will mitigate some of the response issues encountered in SNDA-III. In addition, in a study of low-income 9- and 10-year old girls, using the same diary-assisted recall methodology for three days, we achieved an 86 percent response rate among black girls and a 95 percent response rate among white girls.[3]

If the response rate for the school sample or the student sample falls below 80 percent, a nonresponse bias study will be undertaken to assess the potential for nonresponse bias given the lower than expected response rate. The nonresponse bias study will use available school sampling frame information on the characteristics of participating on nonparticipating sample schools, and available characteristics of the participating and nonparticipating sample students to determine the magnitude of the difference between participants and nonparticipants on a given characteristic variable, and to then determine among the participants whether that variable is correlated with the key study outcome measures. The findings may point to the use of specific characteristic variables to form unit nonresponse adjustment weighting cells. For the samples of respondent and nonrespondent schools we will have a wide range of school characteristics from the National Center for Health Statistics Common Core of Data Public Elementary/Secondary School Universe Survey. For sample students for whom consent is not obtained, we will have little information on individual student characteristics other than grade and possibly gender. We will however have school level characteristics which will allow us to examine the relationship between school level consent rates and schools characteristics.

---

[3]   E. Obarzanek, G.B. Schreiber, P.B. Crawford, S.R. Goldman, P.M. Barrier, M.M. Frederick, and E. Lakatos. Energy intake and physical activity in relation to indexes of body fat: The National Heart, Lung, and Blood Institute Growth and Health Study, *American Journal of Clinical Nutrition*, 60, 1522, 1994.

The procedures to be used to ensure a high rate of response for the study are largely not statistical in nature and focus on methods to ensure the cooperation of State Child Nutrition Agencies, School Food Authorities (SFAs), and school staff.

In eliciting cooperation from States, School Food Authorities, school food service managers, and school principals, we have found that the following guidelines prove successful:

- Use senior-level staff for recruitment and refusal conversion;

- Provide sufficient information about the study purposes, objectives, and methodology so that potential participants have an informed basis for their decision;

- Provide a realistic appraisal of what contributions in time, information, space and human resources the participants will be expected to invest in the study effort and a statement of anticipated benefits to them;

- Demonstrate knowledge and understanding of local FFVP procedures and a sensitivity to the problems facing SFA and school staff in trying to complete their day-to-day activities; and

- Obtain the endorsement and support of State agencies for the objectives of the study.

The data collection will include a web survey of 54 State Agencies and web surveys of SFA directors and school principals representing the 256 schools in the primary sample and 560 schools in the supplementary sample. In order to ensure a high response rate, several key elements will be employed, including development of an accessible and intuitive survey system; contacting respondents and enlisting their cooperation in order to ensure a high response rate; supporting respondents during data collection to assure complete and accurate data; monitoring responses and following up with nonrespondents; and protecting the confidentiality of information collected. Respondents will first be contacted by both letter and e-mail to ensure that they receive information about the study and that we have accurate contact information. At the start of the web survey, respondents will receive an e-mail with instructions about completing the web survey. Nonresponders will receive 2-3 reminder e-mails; phone follow-up will be implemented as needed.

For the on-site data collection a study liaison will be appointed in each sampled school to facilitate the data collection by visiting classrooms of selected students to distribute the packets, describe the study, encourage students to participate, motivate students to return the parent consent form the next day, and monitor the number of parents that return the consent form. Reminder letters for parents will be supplied to class teachers to send home to parents whose consent forms are not received back the following day. A copy of the parent consent form and the reminder letter for parents are in Appendix B. The liaison will be asked to identify two suitable locations for dietary interviewing when the data collectors arrive in the following week, and to assist the data collectors on the interview day by accompanying the selected students to and from their classrooms to interviews. In addition, use of highly trained data collectors helps minimize item nonresponse. These steps have proven to yield an honest, collaborative relationship between the research team and participants in the study.

B.4   TESTS OF PROCEDURES OR METHODS TO BE UNDERTAKEN

The FNS contractor, Abt Associates and it subcontractor, the Center for Weight and Health (CWH) at the University of California, Berkeley conducted pretests of all survey instruments for which OMB clearance is being requested. Since each survey instrument requires feedback from a distinctive target population of respondents, pretesting enables us to begin data collection, analyze the data, and submit a report in time to meet the congressionally mandated final report deadline of September 2011. For this reason, each of the nine instruments were pretested on 9 or fewer respondents, consistent with our understanding of OMB guidelines regarding collection of information prior to receipt of OMB clearance. Field procedures and testing of on-site school instruments were conducted by CWH staff at two elementary schools in northern California in late October. The instruments tested included the Food Diary, the Student Questionnaire, the Parent Questionnaire, the Teacher Questionnaire, the School Food Service Manager Survey and the School Food Environment Assessment instrument. Abt Associates conducted paper-version pretests of the web-based Surveys of State CN Agencies, School Food Authorities, and School Principals in late October and early November.

The primary objectives of the pretest were to evaluate the:

- Ability of respondents to understand and respond to questions;

- Methods of administering the survey instruments;

- Appropriateness of response categories;

- Assumptions regarding availability of certain data items; and

- Length of time required to administer the survey instruments.

Web-Based Surveys

Surveys were sent to respondents by Federal Express with pre-addressed, pre-paid materials for returning surveys to Abt Associates, Respondents were also provided with instructions for recording the time they spent completing the surveys and making note of questions, instructions that were unclear and response options that were inadequate. After receiving completed surveys, Abt staff followed up with each respondent to discuss their experience filling out the survey and noted any comments received on the draft instrument. SFA directors and school principals were given $30 as a gift for pretesting the survey and participating in a debriefing call. The instruments were revised based on feedback obtained from respondents, including clarification of instructions, changing some of the questions, and modifications to question wording. In addition some questions were deleted or simplified to reduce the overall survey burden.

The *Survey of State Child Nutrition Agencies* was sent to respondents at State agencies in four States: California, Tennessee, Indiana and Massachusetts. All four States willingly completed the survey and provided extensive comments in the debriefing. Most questions were seen as clear, appropriate, and not burdensome. None was seen as inappropriate, and there were no suggestions of additional questions to be asked. The most difficult questions were those that required quantitative data. Ease or difficulty of response, level of detail available, and burden depended on how this information was collected and organized at the State level, and whether data files were available, or required manipulation to be able to provide the information in the way it was requested. Thus, the time to complete the surveys all four respondents was longer than planned. Adjustments were made to address the burden issue, including

eliminating questions on the number and enrollment of eligible schools, which we will collect from CCD data, and rewording or reordering questions or response categories.

The ***Survey of School Food Authorities*** was sent to nine respondents in SFAs in the same four States. Six surveys were completed and interviews were conducted with each respondent. Overall the survey took longer than expected to complete, partly because some of the school-specific questions could not be answered by SFA level staff and follow-up was needed was long for many directors, and some questions in a matrix format were problematic for some respondents. The survey was edited to eliminate some questions and restructure others. Questions on popularity of specific fruits and vegetables were removed because SFAs often had to ask their staff at the schools for responses. The number of opinion-type questions were also reduced or consolidated.

The ***Survey of School Principals*** was sent to nine principals of schools within the four pretest States and SFAs. Debriefing interviews were conducted with the seven principals who completed and returned their surveys. The group of respondents included two from non-FFVP schools. Overall, most principals found the survey reasonable and easy. The survey took somewhat longer than expected for some principals, mostly due to the information they had to look up—such as questions dealing with enrollment or changes in competitive food sales. Adjustments were made to the survey by streamlining and simplifying some of the questions that were in matrix format, some reordering of the questions and some of the opinion-type questions were eliminated or consolidated.

A second round of testing will occur once the on-line versions of the surveys are developed, which will occur while the paper instruments are under OMB review. The purpose of the second round of testing is to ensure that the presentation of the survey questions is clear on screen and the data entry and editing procedures are user-friendly and minimize error. The second pilot testing of the instrument will be conducted by up to nine respondents in each respondent group. This test will be followed by a similar telephone interview to get feedback about using the on-line instrument (ease of navigation through the instrument, clarity of on-line instructions, layout and structure of the screens, etc.).

In-School Data Collection Instruments

The procedures and instruments to be used in the FFVP evaluation are similar to those that were developed, tested, and administered for other studies of school-based meals programs by CWH. A total of nine students from one northern California school participated in the pretest of the diary-assisted recall and the Student Questionnaire. Five parents completed the Parent Survey, two teachers completed the Teacher Survey and three school food service managers participated in the School Foodservice Manager Interview.

Interviewers made note of questions that needed clarification, questions that required adjustments, and those that needed to be reworded. The time required for each interview was also recorded. The pilot test wrapped up with a short session asking the respondents their opinions of the survey/interview, including what should be changed and what would make it better. The results of the pilot test were used to revise the instruments and the data collection procedures.

All students who completed the *1-Day Food Diary* returned a completed, useable food diary and successfully completed a multiple pass 24 hr recall based on the record. All received a gift card and reported that the gift card was a good incentive. Two main issues with the instrument were identified: students recorded more than one food per line, a few students did not understand am/pm; and the 8.5 x 11" format of the diary was thought to be a little cumbersome for carrying around, and students mostly reported that a notebook, or smaller size record booklet would be easier. To address these and other issues, the diary has been modified to expand the example (one food per line, whether fruits and vegetables are fresh, frozen or canned, start with lunch) to cover many of the issues which arose in question time. Also, a practice page where students can practice recording their breakfast of that day so they know not to include it in the record was also added. The format of the diaries for the study will be changed to a folded booklet format.

After completion of the 24 hr interview, students were given the *Self-Administered Student Questionnaire* to complete. All students completed the questionnaire. They took 8-12 minutes to complete it. Debriefing with students revealed that additional prompts on the fruits and vegetables, such

as suggested by reviewers would not have changed the answers they gave. Only minor edits were made to this instrument resulting from the pretest.

Parents completed the ***Parent Questionnaire*** in the cafeteria during the school day and were interviewed for their reaction to the questionnaire. The questionnaire took approximately 3 minutes to complete. Overall, it was easy to complete. Few comments or revisions were suggested. Several parents indicated concern about educational time taken by eating FFVP in classroom and wished that teachers used snack time to incorporate more learning about fruits and vegetables. Questions about this concern were added to both teacher and parent questionnaires.

Two teachers were selected by the Principal at one of the schools to complete the ***Teacher Questionnaire*** individually in the Principal's office. The questionnaire was easily completed in less than 5 minutes and the questions were generally clear. Teachers are more familiar with the FFVP in schools where the snack is distributed and consumed in the classroom. Additional questions were thus considered relevant for teachers to capture burden of FFVP and teacher observed waste and student preferences. Questions from the food service manager survey were added so that both teachers and foodservice staff were asked similar questions. Questions on Fruit and Vegetable availability were deleted as teachers are not generally aware of what occurs in food service. Questions throughout were shortened so that respondent burden not changed despite addition of several questions.

The ***Foodservice Manager Interview*** took about 40 minutes to complete, and although the FSMs answered most questions well, the time required is considered too large a burden. Thus, questions were reviewed after the pre-test to pare back the time required to approximately 20 minutes. We retained only questions absolutely required for analysis and essential to clarifying student food records (e.g. menus, satisfaction/perceptions). A page was added for afterschool snack as food service may provide this and foods may be in students' 24 hr recall, and some minor edits to wording were made to the instrument. Questions regarding satisfaction/perception were reformatted to be self-administered (***School Foodservice Manger Survey***) to save time, and it appeared to be easier for FSMs to answer as a survey rather than in interview format. Additionally, questions were made more consistent with the teacher

survey in order to collect comparable data depending on FFVP method of distribution. Other methods for collecting the necessary details about foods served in school meals and snacks are also being considered to further reduce burden.

The *School Food Environment Assessment* instrument was pretested in two schools by two to three observers who observed FFVP distribution and school lunch at both schools and recorded their observations. The instrument worked well and few modifications were needed. Layout improvements were made to reduce time in recording observations. Because distribution methods for FFVP differ between schools, modifications were required to address the options. Minor modifications included a question added to determine who serves FFVP to students and a column to report an estimate of portion size for the fruit or vegetable served.

B.5 INDIVIDUALS CONSULTED ON STATISTICAL ASPECTS AND INDIVIDUALS COLLECTING AND/OR ANALYZING DATA

The design for the study was developed by the contractor, Abt Associates Inc., under the direction of: Dr. Susan Bartlett, Project Director; Michael Battaglia, Sampling Statistician; and Jacob Klerman, Director of Analysis. Dr. Bartlett may be reached at (617) 349-2799 or susan_bartlett@abtassoc.com; Mr. Battaglia may be reached at (617) 349-2425 or mike_battaglia@abtassoc.com; Mr. Klerman may be reached at (617) 520-2613 or jacob_klerman@abtassoc.com.

In addition, Ms. Tracy Palmer and Dr. Ted Macaluso of FNS' Office of Research and Analysis have reviewed the study design and instruments. Ms. Palmer can be reached at (703) 305-2126 or tracy.palmer@fns.usda.gov. Dr. Macaluso can be reached at (703) 305-2121 or ted.macaluso@fns.usda.gov. Abt Associates Inc. is responsible for all data collection and analysis for this study.