IV. ECLS-K-NHANES III Design

This chapter describes the design selected for studying how participation in the SBP influences learning. The study design, which is nonexperimental, incorporates two analyses. The primary analysis is based on data from the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K), which tracks educational and related outcomes of students in the kindergarten class of 1998–1999 at various points through the fifth grade.

The outline of this analysis was sketched in the previous chapter and is described in detail in this one. The supplemental analysis is based on data from the third National Health and Nutrition Examination Survey (NHANES III), collected between 1988 and 1994. This NHANES dataset is *different* than the one considered in Chapter III, since it also contains cognitive and academic testing results, although it contains much of the same nutrition and health information.

NHANES III collected comprehensive information on school-aged children's school breakfast participation, dietary intakes, nutritional status, and health status (intermediate outcomes in the conceptual model shown in Figure II.1), and on their performance on standardized cognitive and academic tests (the main outcomes). A supplemental analysis of NHANES III data will provide a unique opportunity to fully examine the relationships between participation in the School Breakfast Program (SBP) and the intermediate outcomes that may ultimately be linked to the main outcomes of cognitive functioning and academic achievement in U.S. schoolchildren.

This chapter describes the full two-part design suggested for estimating the relationship between participation in the SBP and learning. Sections A through E present the proposed analysis of ECLS-K data introduced in Chapter III. Section F presents the proposed analysis of NHANES III data.

A. ECLS-K Nonexperimental Comparison Group Design

As described in Chapter III, the basic design we propose is a nonexperimental comparison group design. This approach involves comparing key outcomes in students who participate in the SBP with outcomes in students who do not eat a school breakfast. The design is primarily cross-sectional, with learning-related outcomes and SBP participation status measured at a given point in time. The ECLS-K is a longitudinal dataset, however, so the design calls for some longitudinal analysis.

We propose exploiting the longitudinal aspects of the ECLS-K data in at least two ways. First, rather than measuring students' level of cognitive functioning at a given point, the design calls for measuring the *growth* in students' cognitive functioning by contrasting their test scores at the most recent point of data collection with their scores at a previous point of data collection.[39]

We propose analyzing the growth in test scores because we recommend using a measure of SBP participation that refers only to the current school year, and we wish to measure only what has been

[39] As described in this chapter, rather than construct a "growth in test scores" measure and using it as a dependent variable, the analysis would use the current test score level as the dependent variable and would control for the previous test score level in a regression framework. In practice, this approach is similar to one that uses a "growth in test scores" dependent variable but is more flexible.

learned during that school year. A test score level reflects students' innate ability and learning throughout their lives, whereas the growth in test scores approaches our ideal of measuring what has been learned in the current school year.

Second, the longitudinal data include information on students' usual SBP participation status that covers more than one year. Thus, if students' level of cognitive functioning is used as a dependent variable, the definitions of SBP participants and nonparticipants could be based on participation (or nonparticipation) over multiple years prior to the measurement of the outcome. For example, when measuring growth in test scores between grades 1 and 3, students who participated in the SBP during both kindergarten and grade 1 could be compared with students who participated during only a single grade, or who did not participate at all.

B. Variables in the ECLS-K

The ECLS-K consists of a nationally representative sample of 16,906 students in kindergarten in 866 schools as of fall 1998. A wide variety of information about these children has been collected and included in this dataset. In this section, we describe in detail the key pieces of ECLS-K data to be used in the analysis.

1. Outcome variables

The ECLS-K includes a rich set of variables that measure three types of outcomes: (1) a student's cognitive development; (2) events and processes associated with learning, such as school attendance and tardiness; and (3) other aspects of a child's growth, including emotional, social, and physical growth.

a. Cognitive outcomes

Children's cognitive skills and knowledge are measured in three broad areas: (1) language and literacy skills (or reading), (2) mathematics, and (3) general knowledge. The ECLS-K uses a battery of assessments to measure these cognitive skills.[40]

The intent of the ECLS-K cognitive assessment battery is to measure children's common skills and knowledge. In particular, the battery assesses "typical and important elements of the curriculum with particular emphasis on content and process areas that are critical to growth and can be expected to reflect growth on the same scale over time" (West et al., 2000). Thus, the survey measures children's skill levels in a way that allows for comparisons over time.

Students' responses to the assessment battery in each of the three broad areas were converted into norm-referenced scores and criterion-referenced proficiency scores. The norm-referenced scores allow for the measurement of a child's performance against the norm of the performances of other children in the same cohort population. These scores were constructed as *t*-scores with a mean equal to 50 and a standard deviation of 10. The criterion-referenced scores evaluate children's performance in a given area on a specific set of skills thought to reflect cognitive development in that area.

The ECLS-K contains measures of children's basic literacy, such as recognizing printed words, identifying sounds, reading words, vocabulary, and reading comprehension. The criterion-referenced scores include the following five levels of reading proficiency (West et al., 2000):

[40] In addition to these direct tests, the ECLS-K includes a teacher survey in which teachers provide their assessments of sample members' cognitive skills and other outcomes.

- Level 1: Recognition of upper and lower case letters of the alphabet
- Levels 2 & 3: Phonologic sensitivity at the subword level (for example, knowledge of letter and sound relationships at the beginning and at the end of words)
- Level 4: Ability to read common words
- Level 5: Comprehension of written text

The battery of mathematical questions, as described by West et al. (2000), measures such skills as understanding the properties of numbers, mathematical operations, and problem solving; understanding patterns and relationships among numbers; formulating conjectures; and identifying solutions. The five proficiency levels by which the mathematics criterion-referenced scores measure students' skill levels are:

- Level 1: Ability to read numerals, recognize shapes, and count to 10
- Level 2: Ability to count beyond 10, sequence patterns, and use nonstandard units of length to compare objects
- Level 3: Ability to recognize simple number sequences, read two-digit numbers, identify the ordinal position of an object, and solve word problems
- Level 4: Ability to calculate sums up to 10 and to recognize relationships among numbers in sequence
- Level 5: Ability to solve problems using multiplication, division, and number patterns

The general knowledge that the ECLS-K questions assess includes factual information, such as information from history and the physical, earth, biological, and social sciences. The questions also test children on "the skills children need to establish relationships between and among objects, events, or people and to make inferences and to comprehend the implications of verbal and pictorial concepts" (West et al.,2000).

In a study of SBP participation on children's cognitive skills using the ECLS-K, the outcome measure would be the growth in children's test scores (or the change in test scores from one year to another). Because the survey would collect test scores for students included in kindergarten and grades 1, 3, and 5, and their scores would be normalized, we would be able to assess changes in children's cognitive skills relative to their peers from one year to another during the period of their early schooling. Measuring test score growth, rather than test score levels, reduces the variability of this outcome measure across students and allows for more precise estimation of the effect of SBP participation on cognitive skills.

b. School attendance and tardiness

Availability of the SBP may increase some students' school attendance. First, the offer of a school breakfast, especially a free one, may draw children to school for the day. Second, if eating school

breakfasts positively influences children's health, they may be less likely to miss school because of illness. Similarly, if eating school breakfast boosts cognitive skills, children may enjoy their classes more and feel less need to be absent from school.

School attendance rates can be constructed from the student transcript data schools provide for the ECLS-K. In measuring attendance, special care must be taken to ensure the comparability of the transcript data across schools. Often, such data as attendance rates are recorded by different schools in slightly different formats.

We believe the best measure of attendance is a measure of the number of days a student attends school as a percentage of the number of days the school has been open that year. In addition, to the extent that data on tardiness are available, the proportion of days on which children are late to school should be used as a measure potentially related to learning.

c. Emotional/social development

Although not the primary outcome of interest for this design, emotional/social development is important because it has direct implications for a child's well-being. It could easily be included in the analysis of outcomes. The ECLS-K assesses the emotional and social development of children at different grades. The assessments cover children's self-control, responsibility, ability to cooperate, and ability to avoid impulsive reactions and verbal and physical aggression. Parents' and teachers' responses are the primary sources of the assessments, at least during for children in kindergarten and first grade. To the extent that the measures are consistent across grades, it may be possible to estimate not only the impact of SBP participation on social competence at a particular grade, but also the impact of SBP participation on growth in social competence over time.

d. Physical development

Nutritious breakfasts may positively affect children's physical development. The ECLS-K measures this potentially important outcome in several ways. A key measure of children's physical development is their height and weight (and, hence, their body mass index, or BMI), which trained ECLS-K assessors measure at each data collection point. It will be important to estimate the effect of SBP participation on the change in children's height and weight (that is, their growth).

ECLS-K assessors also measure children's fine and gross motor skills. Fine motor skills include such activities as copying figures or manipulating blocks. Gross motor skills are measured by the ability to hop on one foot, walk backward in a line, and perform similar activities.

Children's general health and developmental difficulties may influence physical development. The ECLS-K collects information on general health status from the children's parents. Parents also report their children's activity levels and any developmental difficulties with vision, hearing, articulation, attention span, or coordination.

2. Participation in the SBP

The ECLS-K contains information on SBP participation collected at the student level and the school level. Most of the analysis we propose is based on child-level data, so we focus the following discussion on these data.

The number of days during the previous 5 school days that a student ate a school breakfast is a key piece of student-level participation information. However, there are limitations on the accuracy of the SBP weekly participation variable available in the ECLS-K. In particular, three types of errors have the potential to compromise the accuracy of this variable.

First, only parents who report that their child usually eats a school breakfast are asked the question on breakfast eating during the previous 5 school days. However, even though parents may have reported that their children do not "usually" eat a school breakfast, some of these children may have eaten one or more breakfasts during the previous week, and may do so regularly. The parents of these children would not be asked the question about the number of school breakfasts consumed during the previous week. Therefore, these children would be inaccurately coded as eating no school breakfasts during that 5-day period.

In the proposed design, the ECLS-K information on the number of days of breakfast eating during the previous 5 school days serves as a proxy for the number of days the child ate a school breakfast during a typical week of that school year. This can lead to a second type of proxy error. Having information on the "typical" number of days of SBP participation would be preferable, because we hypothesize that SBP participation over the full school year has a greater influence on student learning than does a single week of program participation.

One type of proxy error arises from using one week of participation as a proxy for participation in the usual week. For example, some children who usually eat a school breakfast every day may have eaten only a few school breakfasts during the previous 5 school days because of some unusual circumstance. Alternatively, children who usually eat school breakfasts only 2 or 3 days a week may have especially liked the food served during the previous week and eaten a school breakfast every day.

A third type of proxy error occurs because parents report their children's experiences.[41] Rather than eating breakfast at school, as the parents report in answer to the question, some children might play or do other things. Alternatively, parents may provide breakfast in the home because they do not realize that their children also receive a breakfast at school. Evidence from the 1994 to 1996 Continuing Survey of Food Intakes by Individuals (CSFII) suggests that parents overestimate their children's SBP participation.

In particular, a question in the CSFII about children's usual participation in the SBP obtained responses from parents suggesting a daily participation rate of 26 percent. According to administrative data from the Food and Nutrition Service (FNS), the actual SBP participation rate during 1994 to 1996 was 20 percent.

A definition of SBP participation that would minimize these errors would be a simple binary variable indicating whether the parent reports that a child usually eats a school breakfast. Although this variable would contain less information than would a variable on the number of days of participation per week, it probably would more accurately measure what it purports to measure. In addition, all parents whose children attend SBP schools would be asked the question.

Finally, although the variable would be subject to proxy error, we believe that parents are more likely to

[41] A different type of reporting error would have arisen if very young children had been asked the participation question. They might not have completely understood the concepts behind the question.

have general information on whether their children usually eat a school breakfast than they are to have more detailed information on the number of days of breakfast was eaten during the previous week.

An alternative definition of SBP participation would expand the scope of the binary variable. It would group parents' responses about participation into three categories: (1) children who do not usually eat breakfast, (2) children who eat breakfast somewhat regularly (one or two times the previous week), and (3) children who usually eat breakfast (three or more times the previous week).

This variable does not contain as much information as a variable giving the exact number of days of participation per week, and proxy error still might occur. Nevertheless, it would measure what it is supposed to measure. Furthermore, it makes it possible to differentiate between children who eat numerous breakfasts per week and those who eat only a few. Finally, having two alternative measures of SBP participation would provide an opportunity to test which measure is more accurate.

The more general SBP participation variable is likely to be more accurate than the more detailed variable, but it also contains less information. We place more importance on the accuracy of the participation data than on obtaining more information, however, so the ECLS-K design calls for the use of the binary participation variable throughout most of the analysis. Because it also would be important to test for the robustness of the results with respect to alternative definitions of SBP participation, the design also calls for estimating alternative versions of the key models that include the SBP participation variable.[42]

For any analysis of between-school differences in students' average learning, researchers would have to create a variable or variables measuring schoolwide SBP participation. Measuring schoolwide participation is somewhat difficult using the ECLS-K. The ECLS-K asks school principals how many students are eligible for free school breakfasts, and how many receive free breakfasts.

Principals are not asked how many students are eligible for reduced-price breakfasts, nor are they asked how many receive reduced-price breakfasts or pay the full price for breakfast.[43] As an alternative to this incomplete information on SBP participation at the school level, it is possible to construct school-level participation variables from student-level data.

There are likely to be approximately 20 students (and 4 or 5 SBP participants) per school in the ECLS-K, so it should be possible to base the school-level participation rate on the participation rate of the sampled students at that school. This estimate is inefficient relative to a percentage reported by the principal (which would be based on all students in the school), but it is unbiased because it includes all SBP participants (those paying a reduced price, those paying full price, and those receiving free breakfasts).

[42] Of course, any school breakfast participation variable will be affected by the degree of parental accuracy in reporting, which is unknown at this point.

[43] CFSII data indicate that 77 percent of school breakfasts were free, with the remaining 23 percent offered at reduced or full price (Gleason and Suitor, 2001).

3. Background characteristics of students and schools

The following text and tables describe the school-level and student-level background characteristics we believe are most relevant for the proposed analysis of learning outcomes. The discussion is based on the research described in Chapter III and on additional hypotheses about the factors potentially related to SBP participation.

a. Student-level background characteristics

The variation in student learning outcomes in a school depends on the characteristics of individual children and their families (and schools). Table IV.1 lists student background characteristics available in the ECLS-K that might affect learning outcomes, and that could be included in the multivariate analyses specified under this design. (Specific variables relating to these characteristics are included in Appendix B.) These characteristics include:

- Child's prior levels of learning (including prior test scores)
- Child's age, sex, and race/ethnicity
- Child's ESL (English as a second language) and disability status
- Number of siblings in the household and the child's relative birth order
- Whether the child is a member of a two-parent household
- Parents' education and employment
- Family income and food security
- Participation in federal nutrition programs (the Food Stamp Program [FSP], SBP, and National School Lunch Program [NSLP])

Table IV.I—Student background characteristics in the ECLS-K

VARIABLE	SOURCE	COMMENTS
Characteristics Affecting Learning		
Prior level of learning	Assessment data, teacher survey, transcripts	Captures starting point for growth in cognitive, social, and emotional skills
Age of child	Parent survey	Measured in months at time of assessment
Sex of child	Parent survey	
Race/ethnicity of child	Parent survey	
ESL status of child	Parent survey	

Disability status of child	Parent survey	
Number of siblings	Parent survey	
Indicator for youngest child	Parent survey	Captures birth-order effects
Indicator for middle child	Parent survey	Captures birth-order effects
Indicator for eldest child	Parent survey	Captures birth-order effects
Two-parent family	Parent survey	
Parental education	Parent survey	
Hours parent(s) work	Parent survey	
Family income	Parent survey	Measured relative to poverty level, including indicators of eligibility for free and reduced-price meals at school
Family food security	Parent survey	Based on a combination of questions
FSP participation	Parent survey	Assumed exogenous
SBP participation	Parent survey	Possibly endogenous
NSLP participation	Parent survey	Possibly exogenous; measured similarly to breakfast participation
Additional Characteristics Potentially Affecting SBP Participation		
Parents' knowledge of SBP	Parent survey, principal survey	Lack of awareness may reduce likelihood of SBP participation
Child's distance from school	Parent survey	Increased distance may reduce SBP participation

ECLS-K = Early Childhood Longitudinal Study, Kindergarten Cohort; ESL = English as a Second Language; FSP = Food Stamp Program; NSLP = National School Lunch Program; SBP = School Breakfast Program.

The ECLS-K also contains information on characteristics that may influence whether the child eats a school breakfast, but that do not influence learning-related outcomes. Variables representing these characteristics, also called instrumental variables or identifying variables, are highly important ones in estimating models of the influence of SBP participation on learning that account for selection bias (that is, bias resulting from unobserved differences between participants and nonparticipants).

The distance between the child's home and school, reported by parents, is a possible instrumental variable in the ECLS-K. Distance from school indicates the relative convenience of sending a child to school to receive a breakfast compared with having no breakfast or eating breakfast only at home. Although distance from school is likely to affect SBP participation, it is less likely to directly influence outcomes related to learning.

Another potential instrumental variable is parents' knowledge of the SBP (that is, awareness of the availability of school breakfasts), which is likely to affect SBP participation but not learning outcomes (after accounting for parents' education levels and the language spoken at home).[44] The quality of this

[44] Parents' knowledge of the SBP may be a weak instrumental variable because parents are informed at the beginning of the school year whether their child's school offers the SBP. Because most parents therefore are likely to know about the program, there will be little variation

instrumental variable depends on whether it is independent of unobserved factors affecting student learning, or whether it is correlated with them. With more than one potential instrumental variable, specification tests, such as those proposed by Wooldridge (1990) or Hausman (1983) could be used to examine whether questionable instrumental variables should be used for estimation.

Although additional information on parents' morning work schedules may be useful for predicting SBP participation, a variable representing the total hours that parents work is not a good candidate for an instrumental variable because it is likely to affect both SBP participation and learning. Furthermore, information on morning work schedules is likely to be a weak instrumental variable unless supplemented by additional information on the child's school schedule; that information is not included in the ECLS-K. In Section C, we suggest additional data that could be collected for use in identifying determinants of SBP participation.

b. School-level background characteristics

The ECLS-K includes a wealth of information on the characteristics of schools that could be used to estimate the contribution of SBP participation to learning outcomes at the school level. These characteristics include the average individual characteristics of sampled students in a given school (such as the average proportion eligible for free school breakfasts or lunches) as well as school characteristics obtained through surveys of principals and teachers, and through the inspection of school facilities.

Table IV.2 lists the following school characteristics that may affect the learning of students (specific variables relating to these characteristics are included in Appendix B):

- Type of school (regular public, magnet, charter, or private), location (urban, rural, or suburban), and region (Northeast, Midwest, South, or West)
- Number of kindergartners
- Number of full-time equivalent staff per student (distinguishes regular classroom teachers from teachers' aides and other staff)
- Teachers' salaries (minimum and maximum values) and background (education, work experience, and tenure)
- Principal's background (teaching experience, experience as a principal, and tenure)
- Allocation of the principal's time (with students, teachers, and parents)
- Use of standardized tests and grade retention
- Availability of special programs (ESL, bilingual education, and gifted/talented programs)

---

in this instrumental variable. Parents who report not knowing about SBP availability in their child's school may simply have responded incorrectly to the survey. Alternatively, their response may indicate that they do not carefully monitor information their child's school provides.

- Availability of full-day kindergarten
- Number of computers per pupil
- Physical characteristics of school library and classrooms

The ECLS-K also includes the following characteristics of schools that are likely to affect average levels of SBP participation without influencing learning outcomes:

- Availability of before-school care[45]
- Physical characteristics of the school cafeteria

Table IV.2—School background characteristics in the ECLS-K (in addition to average levels of student characteristics)

VARIABLE	SOURCE	COMMENTS
Characteristics affecting learning		
School type	Principal survey	Distinguishes regular public schools, magnet schools, charter schools, and private schools
School location	Principal survey	Distinguishes urban, rural, and suburban schools
Number of kindergartners	Principal survey	Average number of students can be calculated using the total school enrollment
Full-time equivalent staff per student	Principal survey	Distinguishes regular classroom teachers, teachers' aides, and other staff
Teachers' salaries	Principal survey	Minimum and maximum values
Teachers' background	Teacher survey	Education, work experience, tenure
Principal's background	Principal survey	Experience as teacher and as principal, tenure
Principal's time allocation	Principal survey	Time with students, teachers, and parents
Use of standardized tests	Principal survey	Proportion of grades assessed
Use of grade retention	Principal survey	Proportion of kindergarten students held back
Special programs	Principal survey	ESL, bilingual, and gifted/talented programs
Full-day kindergarten	Principal survey	
Computers per pupil	Principal survey	
School library and classroom characteristics	Facilities checklist	Includes observations on space, light, ventilation, physical condition, room temperature, noise, and handicap accessibility
Parents' involvement	Principal survey	

[45] Before-school care programs may directly influence learning if they include an educational component. This possibility should be more carefully examined before definitively concluding that this characteristic does not influence learning outcomes.

Additional characteristics affecting SBP participation		
Availability of before-school care	Principal survey	
School cafeteria characteristics	Facilities checklist	Includes observations on space, light, ventilation, physical condition, room temperature, noise level, and handicap accessibility

These two school characteristics could be used as instrumental variables in an analysis of the importance of SBP participation on schoolwide learning. Other school characteristics that the ECLS-K does not collect, such as school bus schedules, the time classes start, and the content of school breakfast menus, are also likely to affect SBP participation. The following section suggests questions that could be added to the ECLS-K survey of school administrators and parents to obtain this supplemental information.

C. ECLS-K Supplemental Data Collection

In a report by Jacobson and Briefel (2000), we discussed a broad range of supplemental data collection activities, including supplemental questions on the parent and school administrative surveys and an entirely new survey component—a brief interview with school food service personnel.[46] Given the high cost of adding survey items and survey components to a data collection effort already involving a broad range of data collection activities on more than 16,000 students in 850 schools, we have limited the number of supplemental data collection questions.

This section describes what we, along with ERS, consider to be the supplemental data collection activities that would have the greatest usefulness in an ECLS-K-based study of the impacts of SBP participation on learning. We suggest adding supplemental questions to the parent and school administrator surveys to facilitate the construction of instrumental variables. (The original questions in the ECLS-K survey that relate to these topics are included in Appendix C.) We have dropped our earlier suggestion (Jacobson and Briefel, 2000) to survey food service employees since the cost of collecting that data would be quite large and some of the data could be collected through the school administrator survey.

1. Parent survey

The ECLS-K parent survey asks parents about their child's school breakfast and lunch participation, and about factors that may influence SBP participation but not learning, such as parents' awareness of the SBP and the distance of the child's home from the school. Table IV.3 lists the proposed high-priority supplemental questions that we suggest be added to the parent survey. We also suggest that the ECLS-K questions on the number of school breakfasts and lunches received during the previous 5 school days be asked of all children, not just of those who usually receive these meals.

[46]Jacobson and Briefel (2000) discussed, but dismissed, the possibility of conducting dietary recall surveys of children.

Table IV.3—Supplemental survey questions for parents

Breakfast habits of the child and parent

1. During the last 5 days (CHILD) was in school, how many breakfasts did (he/she) eat that were NOT school breakfasts (for example, breakfasts eaten at home, at child care, at school but not part of a school breakfast, and so on)? Please count only one breakfast per day.
   None 1 2 3 4 5 Don't Know

(Skip to Question 3, if response is "None.")

2. Where did (CHILD) eat these breakfasts? (check all that apply)

————At home
————At a relative's or friend's home
————At child care location
————At school, but not part of school breakfast
————At a restaurant (this includes food taken-out from fast food restaurants)
————Other (specify)

3. During the last 5 days (CHILD) was in school, how many breakfasts did *you** eat? Please count only one breakfast per day.
   None 1 2 3 4 5 Don't Know
   *Assumes the question is addressed to the main caregiver.

Child's morning schedule

4 (a) How does (CHILD) usually get to school in the morning?

————School bus
————Parent drives {him/her}
————Carpool
————Walk
————Other

4 (b) How long does this take?

————Less than 15 minutes
————15-30 minutes
————More than 30 minutes

5. On school days, how much time does (CHILD) have between arriving at school and classes starting?

————Less than 10 minutes
————10-20 minutes
————More than 20 minutes

The first three questions in the table provide information on the breakfast habits of the child and parent. Question 1 asks the parent to provide the number of nonschool breakfasts the child ate during the previous

5 school days. ("Nonschool" breakfast is defined clearly in the question.)

Question 2 asks about where the nonschool breakfasts were eaten. The information from these two questions would present a more complete picture of children's overall breakfast habits. In particular, when these items are combined with the ECLS-K question on the number of school breakfasts received during the same period, it will be possible to determine how often children skip breakfast.

This information could be used to address such issues as whether the availability of the SBP in a school is related to the probability of skipping breakfast, and whether eating any breakfast influences school learning. It will also be possible to determine, amongst the children who do regularly eat a nonschool breakfast, what breakfast alternatives are favored.

The information on parents' breakfast habits, obtained from question 3, would serve as a proxy for parents' attitudes about breakfast eating at home. The more often a parent or other adult in the household eats breakfast at home, the greater the likelihood their child will eat at home with them, and the lower the likelihood that the child will eat a school breakfast.

By themselves, however, parents' breakfast habits could be exogenous with respect to the student learning outcomes in this study. In this case, parents' breakfast habits could be a useful instrumental variable for children's SBP participation status.

The information that can be drawn from questions 4 and 5, on a child's morning schedule, could be used to construct useful instrumental variables. The questions on children's morning schedules are designed to capture as much relevant data as possible to account for various possible scheduling arrangements, and to give researchers flexibility in constructing variables representing the morning schedule.

Children's morning schedules are likely to influence their SBP participation status but have no direct effect on learning. For example, a child who arrives at school via school bus well before the start of school would have time to eat a school breakfast. A child who must rise early because of a long school bus commute would have limited time at home for breakfast but might be more likely to eat a school breakfast, if time is available to do so.

A child who is driven to school by a parent might be more likely to eat at home with his or her parents (because parent and child would leave the house at the same time). Because variables derived from school bus schedules are likely to be independent of parental behavior and parental characteristics and are unlikely to directly affect learning outcomes, they are good candidates to be instrumental variables to predict school breakfast participation.

2. School administrator survey

Table IV.4 presents supplemental questions that we suggest be added to the school administrator survey. To derive the greatest benefit from the supplemental parent survey questions on children's morning schedules, we would have to add supplemental school administrator survey questions on the school's morning schedule and SBP participation.

Table IV.4—Supplemental survey questions for school administrators

Morning school schedule

1. What time does the FIRST bus usually arrive at school in the morning? _____ AM
2. What time does the LAST bus usually arrive at school in the morning? _____ AM
3. What time does school officially start in the morning? _____ AM

School-level breakfast and lunch eligibility and participation

4. Does your school participate in USDA's school breakfast program?
   Yes [skip TO Question 6] No
5. What are the reasons why your school does not participate in USDA's school breakfast program?
   _____ Too few eligible students
   _____ Program too costly
   _____ School starts too late to serve breakfast
   _____ School lacks facilities to serve breakfast
   _____ School lacks staff to serve breakfast
   _____ Other reason (SPECIFY)

   [skip to Question 10]

6. What time is breakfast served at the school? _____ start time _____ end time

7. Where is the breakfast typically served for 3rd graders?
   _____Cafeteria
   _____Classroom
   _____School bus (as a bag breakfast)
   _____In some other common area of school (as a bag breakfast)
   _____Other (SPECIFY)_____

8. Are children who are served breakfast in the cafeteria allowed to take it to the classroom?

   Yes No

9. How many children in your school were eligible for and participating in the school breakfast program as of October of the current (or most recent) school year? Write in numbers below.

	Eligible Children	Participating Children
a. Any school breakfast?	All Enrolled	_____
b. Free school breakfast?	_____	_____
c. Reduced-price breakfast?	_____	_____

10. How many children in your school were eligible for and participating in the school lunch program as of October of the current (or most recent) school year? Write in numbers below. If service is not provided, write "zero."

	Eligible Children	Participating Children
a. Any school lunch?	All Enrolled	_____
b. Free school lunch?	_____	_____
c. Reduced-price school lunch?	_____	_____

USDA = U.S. Department of Agriculture

Questions 1 and 2 would add information on the time that children arrive at the school by bus, and Question 3 would add information on the time that morning classes start. Question 4 checks whether a school even participates in the SBP and question 5 asks school administrators in schools that do not offer the SBP why the school does not participate in the program.

This information would be useful because a school's lack of participation in the SBP implies that all children of that school will be nonparticipants. The information could be used to develop strategies to make school breakfasts accessible to a broader range of students. In addition, because students in non-SBP schools are a potential comparison group to SBP participants, the information would help researchers determine the extent to which students at these schools are in fact comparable to students at SBP schools.

Question 6 would add information on the time that breakfast is served, and questions 7 and 8 would add information on where breakfast is served or eaten (to a typical class, such as third graders). The key feature of these questions, along with questions 1 through 3, is that they could be combined with the supplemental questions on children's morning schedules to construct more detailed measures of the scheduling and logistical constraints that might increase or decrease the likelihood that a particular child would participate in the SBP.

For example, combining the information from parents on the time that a particular child has between arriving at school and going to class with the information from the school administrator on the time and place that breakfast is served and on the time that classes start would enable us to determine whether the child has the opportunity to eat a school breakfast.[47]

The final supplemental questions, 9 and 10, add several options to a question in the ECLS-K school administrator survey that asks school administrators to provide the (certified) eligibility and participation data for free breakfasts, free lunches, and reduced-price lunches. To obtain complete data on school-level certification and participation, we suggest adding questions about eligibility for reduced-price breakfasts

[47] If the bus arrives at 8 a.m., breakfast is served from 8 a.m. to 8:30 a.m., and classes start at 8:30 a.m., the child could eat breakfast. Children who arrive by bus at 8:25 a.m. would not be able to eat breakfast, unless breakfast were served in the classroom or provided by the cafeteria in bags, to be eaten in the classroom.

and about participation among all students in the breakfast program, eligible students in the reduced-price breakfast program, and all students in the lunch program.[48]

This information would be especially helpful, as we would be able to increase the precision of the measures of school-level SBP participation, and to assess the quality of the parent-reported, child-level SBP participation data by comparing the implied participation rates based on this information with the implied participation rate based on the school-level data reported by principals. Finally, the ECLS-K already contains a question to obtain similar information, so it would be easy to add these extra items.

In isolation, the supplemental questions to the parent and school administrator survey would be less useful than when combined. In combination, the questions would enable researchers to construct a potentially powerful instruments that would address selection bias in estimations of the impact of SBP participation on learning. One simple example of such a variable would be a binary indicator of whether or not the child takes a bus to school in the morning and whether the bus arrives at least 20 minutes before school starts. We believe that these children would be much more likely to participate in the SBP.

D. ECLS-K Analytic Approach

We propose using the ECLS-K in a nonexperimental analysis. We must therefore rely on multivariate statistical methods to infer the differential effect of school breakfast participation (or, alternatively, attendance at a school offering the SBP) on the educational outcomes of otherwise identical students.

A strategy for estimating the impact of SBP participation on learning must account for the selection of SBP participants into the sample, conditional on students attending a school that offers the SBP. It also may be necessary to account for the selection of schools into the sample of schools offering the SBP.

In this section, we discuss the general framework for the analysis and the way we propose to address selection bias. We then discuss methods for identifying students with varying propensities of receiving school breakfasts, conditional on their attendance at a school offering the SBP. In the last two sections, we discuss, respectively, the minimum detectable differences (MDDs) that are likely to arise from our methodology and different procedures we might use to present estimated results.

The basic analytic framework was introduced in Chapter III. Because this model is based on multiple observations of children from the same school, it can be written in the form of two equations, one representing within-school variation in student participation and outcomes and the second representing between-school variation in mean participation and outcome levels.

(1) $(Y_{is} - Y_s) = \gamma (X_{is} - X_s) + \delta (P_{is} - P_s) + (e_{is} - e_s)$.
(2) $Y_s = \alpha A_s + \beta Z_s + \gamma X_s + \delta P_s + u_s$.

[48] Alternatively, each of these questions could be broken down into three parts: (1) the number of students certified for free or reduced-price meals as of October (when certifications are usually completed), (2) the number of free, reduced-price, and full-price meals served during a month or school year, and (3) the number of serving days covered in the period in the second part of the question. However, this step would increase the complexity of the survey. Furthermore, because the change would occur in the third or fourth year of the study, the longitudinal quality of the data would be affected adversely.

The variables in these equations are defined as in Chapter III.

Given the assumption that the independent variables are uncorrelated with the error terms (that is, that there is no selection bias), the models can be estimated through linear estimation methods, such as ordinary least squares (OLS) estimation. The assumption of no selection bias is a relatively strong one, however. It implies that, after controlling for observable characteristics, participants and nonparticipants are identical in all other ways related to the outcome of interest. The following section discusses methods for estimating the impact of SBP participation on learning if the selection bias assumption is relaxed.

1. Selection bias

In estimating the parameters of equations (1) and (2), we must recognize the possibility that selection into the sample of school breakfast participants is endogenous with respect to learning outcomes. That is, school breakfast participation levels (P_{is} and P_s) may be correlated with the unexplained portion of learning outcomes within schools (e_{is}) or between schools (u_s).

This correlation could occur because of unobserved characteristics of families or schools that both promote school breakfast participation and advance a child's learning. If such a correlation were present, the estimated relationship between school breakfast participation and learning (δ) would be biased, as would the other parameters of equations (1) and (2).

A standard correction for selection bias is to use an instrumental variables (IV) procedure to predict the endogenous explanatory variable, using a linear equation that has the following form for within-school participation:

(3) $(P_{is} - P_s) = \pi (X_{is} - X_s) + \theta (I_{is} - I_s) + (r_{is} - r_s)$,

and the following form for between-school participation:

(4) $P_s = \kappa Z_s + \phi X_s + \theta I_s + v_s$,

where P_{is} refers to each student's school breakfast participation; P_s refers to average school breakfast participation in each school; X_{is} refers to characteristics of each student; X_s refers to the average levels of these characteristics in each school; I_{is} refers to additional characteristics of each student affecting school breakfast participation but not student outcomes; I_s refers to the average levels of these characteristics in each school; Z_s refers to additional school characteristics affecting both school breakfast participation and student outcomes; r_{is}, r_s, and v_s are error terms (with $r_s = 0$); and κ, φ, and θ are parameters to be estimated (κ represents the estimated effect of Z_s on P_s, φ represents the estimated effect of X_s on P_s, and θ represents the estimated effect of I_s on P_s).[49] These equations would be estimated conditional on a school offering the SBP (that is, conditional on $A_s = 1$).

[49] When P_{is} is defined as a binary variable, equation (3) becomes equivalent to a linear probability model in which a heteroskedasticity correction is required to produce correct standard errors (see Maddala 1983). In contrast, P_s is an average participation rate and is assumed to be a continuous variable between 0 and 1 for the purpose of OLS estimation.

Researchers could use the estimates of the parameters in equations (3) and (4) to derive predicted values of school breakfast participation levels and could then substitute these values for the reported values in equations (1) and (2).

The resulting estimates of α, β, γ, and δ will be unbiased if the variables I_{is} and I_s (the instrumental variables) are validly excluded from equations (1) and (2). To qualify as a valid instrumental variable, a variable must be uncorrelated with the outcome variable but correlated with the endogenous explanatory variable.

In this case, I_{is} and I_s would be factors affecting school breakfast participation but *not* learning outcomes. Family income and household food security are likely to be important factors affecting SBP participation. However, they also are likely to have independent effects on learning, so they are not valid instrumental variables.

In this chapter, we discussed candidates for instrumental variables from the original ECLS-K survey and supplemental data collection from parents and schools that might be used in this analysis.

The strength of IV estimates depends on the availability of valid instruments as well as on their predictive power. As we note in Section B.3 of this chapter, without strong instruments, the ability of the analysis to detect small impacts of SBP participation on learning is seriously limited. An equally serious issue is the practical difficulty of identifying valid instrumental variables, even among those that appear to be theoretically sound. As illustrated by Bound et al. (1993), having instruments of questionable validity can lead to biased impact estimates.

A further complication of the analysis would arise if, in addition to the endogeneity of a student's SBP participation (P_{is}) with respect to learning outcomes, the student's attendance at a school offering the SBP (A_s) also were endogenous with respect to these outcomes. This situation would involve endogenous selection into the sample of students in SBP schools. For example, parents who expect their children to benefit academically from SBP participation might be more likely than other parents to place their children in schools offering the SBP. Alternatively, schools might adopt the SBP in the belief that their students are especially likely to benefit academically from SBP participation.

At least two approaches could be used to model the selection of schools into the SBP. First, linear instrumental variables could be used to predict a school's participation in the SBP, as well as the student's participation in the program. Second, a bivariate probit procedure and the proposed set of instrumental variables could jointly estimate the probability of the school offering the SBP and the probability of the student participating in the SBP where it is offered.

This procedure would not only make it possible to predict each student's probability of attending a school offering the SBP, it would also make it possible to predict each student's *joint* probability of attending an SBP school and participating in the program. The joint probabilities could be substituted for the reported SBP participation of the school and student in equation (2). The difficulties of identifying valid instruments in the previously mentioned IV models apply to an even greater extent to this model, which contains two possible sources of selection bias.

2. Propensity score methods

Estimating probabilities of SBP participation for students in SBP schools would make it possible to use propensity score methods to create subgroups of likely SBP participants in SBP schools and likely SBP participants in non-SBP schools. These methods effectively match participants and nonparticipants on the basis of their observable characteristics. However, because the two groups might differ in their unobservable characteristics, propensity score methods would not address the underlying selection bias problem, if it exists.

Under the propensity score methodology, the subclassification of participants and nonparticipants on the "propensity score" could be accomplished as follows. First, assume that the probability of SBP participation is estimated according to the following equation:

(5) $probability\ (P_{is} = 1 \mid A_s = 1) = F\ (\pi Z_s + \sigma X_{is} + \tau I_{is}),$

where P_{is} is each student's actual SBP participation; A_s is the availability of the SBP at the student's school; Z_s is a set of observed school-specific characteristics; X_{is} is a set of personal background characteristics for the student; I_{is} is a set of additional characteristics correlated with SBP participation; π, σ, and τ are parameters to be estimated; and $F(.)$ is the normal density function.[50]

Next, assume that equation (5) is used to assign all students, including those in non-SBP schools, to categories (Q_{is}) on the basis of their probability of SBP participation if they were in a school offering the SBP. At a minimum, Q_{is} would define quintiles of the distribution of students in non-SBP schools (or of students in SBP schools, depending on which population was of greater interest for policy purposes).[51] The following equation could be estimated separately for each subgroup k defined by Q_{is}:

(6) $Y_{is} = \alpha^k A_s + \beta Z_s + \gamma^{kk} X_{is} + w_{is},$ for each category k of Q_{is},

where Y_{is}, A_s, Z_s, and X_{is} would be defined as before; w_{is} would be an error term; and α^k, β^k, and γ^k would be parameters to be estimated. The key parameter of interest would be α^k, which would measure the impact of offering the SBP to students with a propensity to participate indicated by subgroup k. We expect students in lower-income families and in food-insecure households to be in subgroups with higher propensities of SBP participation. Averaging the values of α^k across all the subgroups would produce an estimate of the impact of expanding the SBP to students in all schools that do not offer the program.[52]

---

[50] If equation (5) is estimated jointly with the probability of the school offering the SBP, a bivariate normal density function also may be used.

[51] To study the effects of offering the SBP to students in non-SBP schools, the subclassification must be based on the distribution of estimated propensity scores for students in non-SBP schools. For some subclasses, relatively few observations may be available for comparison, suggesting that comparisons for these subclasses will be imprecise, at best, and impossible to make, at worst. This issue is relevant to an analysis of the SBP, because relatively few ECLS-K students in non-SBP schools might have high propensities of SBP participation.

[52] For additional information on propensity score methods, including methods for matching observations directly using the propensity score, we refer the reader to Rosenbaum and Rubin (1983, 1984, and 1985). For an interesting example of subclassification used to create subgroups for policy analysis, consult Kemple et al. (2000).

This propensity score approach is valuable for estimating the impact of *offering* the SBP to students, but it also could be used to develop an alternative set of estimates of the impact of SBP *participation* on learning. If, as appears likely, the estimated impact of offering the SBP is zero for nonparticipants, the impact of SBP participation on participants in each subgroup may be estimated by dividing the corresponding SBP participation rate in schools offering the SBP into the estimate of α^K. As before, averaging the different estimates for the various subgroups could be used to obtain an overall estimate for the entire population of interest. *Note that, to generate unbiased impact estimates, this analysis also would have to assume that attendance at a school offering the SBP is itself independent of unobserved variables influencing learning outcomes.*

3. Minimum detectable differences

The formula for calculating the MDD for the participation model (assuming no selection bias) is the following with 80 percent power:

$$MDD = 2.486\sqrt{\left(\frac{1-R_1^2}{1-R_2^2}\right)\left[(1-\rho)\sigma^2\left(\frac{1}{sn_t}+\frac{1}{sn_c}\right) + \rho\sigma^2\left(\frac{1}{s}+\frac{1}{s}\right)\right]}$$

where σ^2 is the variance of the outcome, ρ is the proportion of the total variance due to cross-school variance, R_1^2 is from the outcome equation, R_2^2 is from a supplemental regression of treatment status (participation) on the X variables from the outcome equation, s is the number of schools in the sample, n_t is the number of participants per school in the sample, and n_c is the number of nonparticipants per school in the sample.[53]

The main difference between this formula and the formula for an experimental evaluation, as described by Bloom (1995), is that the nonexperimental formula divides by ($1 - R_2^2$), which is the proportion of variation in participation status remaining after we control for the X variables. The greater the correlation between participation status and the X variables, the smaller this term will be and the larger the MDD will be. The formula is based on the formula for the standard error of a regression coefficient, taking into account multicollinearity.

In the case of selection bias, there is even less exogenous variation in participation status. All the techniques for controlling for selection bias effectively control for the unobservable determinants of participation status. Therefore, the only exogenous variation in participation comes from the portion of the variation explained by the identifying variables. The resulting formula is:

$$MDD = 2.486\sqrt{\left(\frac{1-R_1^2}{R_3^2-R_2^2}\right)\left[(1-\rho)\sigma^2\left(\frac{1}{sn_t}+\frac{1}{sn_c}\right) + \rho\sigma^2\left(\frac{1}{s}+\frac{1}{s}\right)\right]}$$

[53] This formula accounts for the clustering of observations in particular schools. In addition, the formula assumes that observations from students attending schools that do not offer the SBP are excluded, and that each school in the included sample has both SBP participants and nonparticipants.

where R_3^2 is from a regression of participation status on both the X variables from the outcome equation and the identifying variables. The term (R_3^2 - R_2^2) represents the increase in the R^2 that comes from adding the identifying variables to the participation equation. The better the identifying variables predict participation, the more they will add to the R^2 and the lower the resulting MDD will be.

We can use the two formulas shown here to calculate sample MDDs for an achievement test score outcome under different assumptions. In the following three sample scenarios, we assume that the analysis excludes schools that do not off the SBP. We also assume that, after accounting for sample attrition, there will be at least 750 schools in the ECLS-K that could be used in the analysis. We expect the schools to include a total of 15,000 students (20 students per school). An average of one-quarter of the students (3,750) will be SBP participants, and an average of three-quarters (11,250) will be nonparticipants.

Scenario I

$\rho = 0.05$
$\sigma^2 = 1$
$R_1^2 = 0.30$ (reasonable when estimating achievement growth as a function of prior achievement)
$R_2^2 = 0.25$ (0.23 was obtained using ECLS-K data and a reduced set of explanatory variables)
Participation rate = 25 percent
750 schools (that is, $s = 750$)
20 students per school
thus, $n_t = 5$
$n_c = 15$
no selection bias

MDD = 2.486 * 1 * sqrt {(.70)/(.75) * [(.95)(1/3,750 + 1/11,250) + (.05)(1/750 + 1/750)]} = 0.052 of a standard deviation.

(For comparison, the average distance between adjacent percentiles of the test score distribution could reasonably equal 0.025 of a standard deviation, so this magnitude of MDD should enable us to detect relatively small differences in outcomes, for example, a change of three percentiles.)

Scenario II

Same as Scenario I but with selection bias and
$R_3^2 = 0.35$ (under this assumption, the identifying variables add 0.10 to R^2)[54]

MDD = 2.486 * 1 * sqrt {(.70)/(.10) * [(.95)(1/3,750 + 1/11,2500) + (.05)(1/750 + 1/750) } = 0.143 of a standard deviation.

[54] We analyzed actual ECLS-K data for 1998-1999 and did not find any evidence of such an increase in predictive power from the addition of the identifying variables available in the ECLS-K to the participation equation. A dramatic increase would be possible only if supplemental data collection obtained an especially powerful set of identifying variables.

(This example indicates that selection bias can dramatically increase the MDD of the analysis, making it more difficult to detect small differences in outcomes between SBP participants and nonparticipants.)

Scenario III

Same as Scenario II but with identifying variables that are much less powerful predictors of selection:

$R_3^2 = 0.29$ (under this assumption, the identifying variables add only 0.04 to R^2)[55]

MDD = 2.486 * 1 * sqrt {(.70)/(.04) * [(.95)(1/3,750 + 1/11,250) + (.05)(1/750 + 1/750)} = 0.226 of a standard deviation.

(This example indicates how the absence of instrumental variables that are strong predictors of SBP participation can dramatically increase the size of the MDD when selection bias is present, making it highly unlikely that small differences in outcomes between SBP participants and SBP nonparticipants will be detected.)

In general, the large number of schools in the ECLS-K increases the power of ECLS-K-based analyses. If, instead of 15,000 students distributed across 750 schools, a research sample contained the same number of students distributed across 100 schools, the MDDs in scenarios I, II, and III would have equaled 0.088, 0.241 and 0.380, respectively, 69 percent larger than the MDDs estimated for the ECLS-K. Consequently, the MDDs for ECLS-K-based analyses are smaller primarily because of the large number of schools included in the study.[56]

E. Presentation of ECLS-K-Based Design Results

Table 5.5 indicates how estimated impacts of SBP participation (or attending an SBP school) might be presented for comparison purposes. Each proposed estimation method has certain limitations. A simple comparison of reported differences in outcomes between SBP participants and SBP nonparticipants would fail to take into account the observed differences between the two groups.

OLS, fixed-effects, and random-effects estimates account for observed differences between students and their schools, but not for the endogenous selection of students into the SBP (and, possibly, into schools offering the SBP). IV estimates may help correct for sample selection bias, but their ability to do so is depends on whether valid identifying variables are found. In addition, the precision of the IV estimates will depend on the predictive power of the instrumental variables used in the analysis.

[55] Preliminary regressions using ECLS-K data for 1998-1999 indicate that this increase in predictive power is obtained by adding parents' knowledge of the SBP and distance from the school as predictive variables.

[56] These calculations ignore the possibility of greater clustering of the schools in the ECLS-K resulting from sampling of multiple schools from the same school district. To the extent that clustering occurs, the MDDs understate the true MDDs obtainable through the evaluation.

Table IV.5—Presentation of estimated impacts of the SBP on a given outcome (for example, test scores measured in standard deviation units)

IMPACT/ESTIMATION	Estimated impact	Standard\ error	Statistical significance
Method			
Impact of SBP participation on SBP participants versus nonparticipants			
Reported values			
OLS			
Fixed effects for schools			
Random effects for schools			
IV with school fixed effects			
IV with school random effects			
IV with school SBP participation endogenous			
Propensity score approach (based on distribution in SBP schools)			
Propensity score approach (based on distribution in non-SBP schools)			
Propensity score approach (based on distribution in SBP schools)			
Students in top propensity quintile			
Students in second propensity quintile			
Students in third propensity quintile			
Students in fourth propensity quintile			
Students in bottom propensity quintile			
Propensity score approach (based on distribution in non-SBP schools)			
Students in top propensity quintile			
Students in second propensity quintile			
Students in third propensity quintile			
Students in fourth propensity quintile			
Students in bottom propensity quintile			

Propensity score methods can account for how attendance at an SBP school is associated with different outcomes for students with different probabilities of SBP participation. However, the ability of propensity score methods to identify the true impact of SBP participation on learning outcomes depends on whether attendance at an SBP school is independent of the unobserved factors influencing learning.

In other words, selection bias may influence the estimation of propensity score models as well. Propensity score methods cannot be used to correct for the endogenous selection of students into the sample of SBP participants, as propensity scores account only for differences in students' observable characteristics.

F. Supplemental NHANES III Analysis

The primary rationale for conducting the supplemental NHANES III analysis is that the ECLS-K-based analysis does not address the question of how SBP participation affects intermediate outcomes, such as nutrition and health status. Furthermore, NHANES III includes cognitive and academic performance tests as key indicators of learning.

By understanding the direct impact of participation on students' dietary intakes, nutritional status, and health status, it will be easier to interpret any impacts of participation on learning-related outcomes. For example, a positive impact of participation on learning could arise because students who eat school breakfasts have improved nutritional status or are more alert during morning classes. Alternatively, if participation improves students' overall health, participants may be absent from school because of sickness less often and thereby perform better academically.

Conducting the supplemental NHANES III analysis would improve the overall design for two additional reasons. First, the analysis would provide additional evidence on the impact of school breakfast participation on learning. Because NHANES III contains information on children's cognitive functioning and academic performance, the supplemental analysis may corroborate or fail to corroborate the primary findings from the ECLS-K.[57]

The issues of selection bias and statistical power are important in the ECLS-K-based design, so having additional evidence on the critical research question will be valuable. Second, the NHANES III data would give the design the ability to examine the relationship between SBP participation and various student outcomes among older children as well as elementary-school-aged children.

In particular, the analysis may suggest whether the effects of SBP participation change as children get older, although we know that SBP participation decreases as children age. Defining the variables that contribute to SBP participation in older children, especially teenaged girls who frequently skip breakfast, could help us understand the potential benefits of promoting the SBP to older age groups.

1. Basic Design Approach

In the NHANES III analysis, the intervention is participating in the SBP (that is, usually eating a school breakfast). To study learning outcomes, we are most interested in an intervention that covers a substantial

[57] The analysis of a main outcome, academic achievement, is more limited in NHANES III because it is measured at only one point in time. Therefore, no measure of change is available.

period—"usual participation" over a school year.[58] This information is collected in the NHANES as the number of times per week that the child selects a school breakfast. In particular, the intervention refers to participation in the SBP *as it was implemented in schools throughout the country from 1988 to 1994.*

The counterfactual against which the intervention is being assessed is nonparticipation in the SBP, or "not usually eating a school breakfast." The main counterfactual condition is attending a school that offers the SBP but "not usually eating a school breakfast". We also may want to examine the counterfactual condition of attending a school that does not offer the SBP. Finally, it may be useful to examine differences in outcomes among three primary groups: (1) SBP participants, (2) SBP nonparticipants, and (3) children who usually do not eat breakfast.

Elementary-school-aged children are of greatest interest among the populations the intervention targets. The design also calls for analysis of models that include older children (especially girls) because (1) the NHANES III data include information on older children and increase the available sample size for the proposed analysis; and (2) older children are less likely to report eating breakfast, which increases the potential role of the SBP.

Although the NHANES III includes a wealth of information pertinent to the research question of interest, the total sample size is smaller than the one available in the ECLS-K design, and the dataset is not longitudinal. Thus, the NHANES III data analysis will supplement the analysis of the ECLS-K data. The cross-sectional nature of the NHANES III dataset (and the resulting NHANES III analysis) means that key outcomes are measured at a given point in time during the school year.

The mean values of key outcomes are collected at the same time and cover the same time period as is covered by the information on SBP participation. One implication of this design is that it is not possible to measure changes in students' achievement levels, as in the ECLS-K-based analysis. This lessens the resulting statistical power of the estimated impact of SBP participation on learning.

2. Data and Measurement

a. The NHANES III Data

The NHANES provides data on a nationally representative sample of school-aged children, including SBP participants and nonparticipants. It measures much of the same family and school information that other national surveys such as the ECLS-K measure, but it also includes detailed diet, nutrition, and health information on younger and older children.

The NHANES sample is a representative, cross-sectional sample of the U.S. civilian, noninstitutionalized population. The sample design is a multistage, complex, stratified survey design of individuals living in households. The survey was designed as two 3-year national samples: (1) 1988 through 1991, and (2) 1991 through 1994. In addition, the entire six-year span constitutes a national sample.

The survey design includes participants of all ages and racial/ethnic groups and equivalent numbers of boys and girls in each age group. Blacks and Mexican Americans are oversampled to produce reliable estimates for these major racial/ethnic groups, and sampling weights are used to adjust for nonresponse

[58] Several possible definitions of "usual participation" could be used. Definitions and related measurement issues are fully described in Section B.2 of this chapter.

and to inflate the sample to produce nationally representative estimates (National Center for Health Statistics 1994). Though NHANES III data collection ended in 1994, when school breakfast participation rates were lower and school environments were different, the data's usefulness to supplement an ECLS-K based analysis outweighs these slight drawbacks.

The specific content in NHANES III that is pertinent to a study of SBP participation and learning is shown in Appendix D and includes a detailed list of recommended study variables. The table also presents the SAS variable names and corresponding NHANES III data file containing the variables. Details of the survey instrument are described in greater detail in other publications (National Center for Health Statistics, 1994) and are summarized here:

- ***SBP Participation.*** Participation in the SBP is a key variable for the study of learning and school breakfast, but the NHANES measure of SBP participation may be somewhat imprecise. The NHANES includes "usual" participation in the SBP during the school year (that is, the number of times per week the child usually receives school breakfast, as well as whether the breakfast is free or at a reduced price). This information is reported by the parents of all sampled children who are younger than age 17.[59]

- ***School Meals.*** Information is collected from the parent on whether the child's school offers the SBP or NSLP, how frequently the child participates in the SBP and NSLP, and whether the child receives free or reduced-price school meals.

- ***School Outcomes.*** Information is collected from the parent on the child's grade level, attendance, suspensions, expulsions, and skipped grades.
- ***Cognitive Tests.*** Standardized cognitive tests to evaluate short-term cognition and intelligence were administered in the mobile examination center during a private interview with 6- to 16-year-old children. Academic performance was assessed using subtests of the Wechsler Intelligence Scale for Children, Revised (WISC-R) and the reading and arithmetic subsections of the Wide Range Achievement Test, Revised (WRAT-R) (National Center for Health Statistics, 1996; and Kramer et al., 1995).

- ***Family Characteristics.*** Information on the family's income, food assistance program participation in the past year, food sufficiency, and socioeconomic status was collected, as was information on parents' education, health insurance coverage, and sources of medical care.

- ***Protective and Risk Behaviors.*** Assessed health behaviors for younger and older children include physical activity levels and time spent performing sedentary activities, such as watching television. Risk behaviors collected during private interviews with the child in the mobile examination center include smoking (for ages 8 years and older) and alcohol and drug use (for ages 12 years and older).

- ***Dietary Intake and Dietary Behavior.*** Dietary intake was assessed using 24-hour dietary recall methodology and additional interview questions about dietary habits. At least one 24-hour recall per

[59] Like with a variable constructed from the ECLS-K, any school breakfast participation variable constructed from the NHANES will be affected by the degree of parental accuracy in reporting, which is unknown at this point.

person was collected, with a second day's recall collected on a subsample. The 24-hour dietary recall provides information on whether breakfast was consumed, the time and source of breakfast(s), the foods and amounts consumed at breakfast, and the total day's intake. Total nutrient intake was estimated using information collected on dietary supplements use, discretionary salt use, and water intake. For 6- to 11-year-old children, a combination of self- and proxy-reporting was used for the 24-hour recall and dietary questions. For children 12 years or older, dietary intake and behavior were self-reported.

- ***Nutritional Status.*** The NHANES provides the most comprehensive picture of nutritional status available on a national sample of school children. Precise anthropometric measurements, such as height and weight, are used to assess growth and overweight in relation to the revised growth charts developed by the Centers for Disease Control and Prevention (CDC) (Kuczmarski et al., 2000). Blood and urinary measurements provide an assessment of vitamin and mineral status for a wide variety of nutrients, such as B vitamins and iron. Iron status is of particular interest, as iron deficiency is related to developmental and behavioral disturbances that may affect mental performance and learning in young children (Centers for Disease Control and Prevention, 1998).

- ***Health Status.*** General health status measures, such as blood pressure and presence of respiratory disease, provide an overall picture of the child's health and readiness to learn. Other health components in the NHANES related to a study of breakfast and learning include vision and hearing problems, which may affect classroom learning, and environmental exposures, such as to lead. Frequent health problems and illnesses may lead to more days absent from school and fewer opportunities to learn. Elevated levels of lead in the blood may be associated with iron deficiency anemia and are higher in low-income children (Centers for Disease Control and Prevention, 1998). Variables that relate to the child's prenatal environment, such as low birth weight or exposure to smoke, have been shown to relate to growth and development and are collected in the parent interview.

Table IV.6 presents the NHANES III sample sizes for school-aged children in three groups: (1) children who participated in SBP at least once per week, (2) children who did not participate in the SBP in schools that offered the program, and (3) children who did not have access to the SBP in the school they attended.

Between 1988 and 1994, about 18 percent of the NHANES III sample of school-aged children participated in the SBP, 32 percent did not participate, and 49 percent attended a school that did not offer the SBP. Oversampling produced approximately equal numbers of respondents who were participants, nonparticipants in SBP schools, and nonparticipants in non-SBP schools.[60] The sample contains equivalent numbers of boys and girls.

[60] For specific analytic purposes, nonparticipants can be further divided into subgroups that do or do not consume breakfast at home, or that do not consume breakfast at all. SBP participants may also consume a breakfast at home or at some other place, a factor to be considered when analyzing and interpreting the dietary data.

Table IV.6—Sample size of children aged 6 to 16 years, by SBP participation

SBP participation	6 to 11 years		12 to 16 years		6 to 16 years	
	N	(%)[a]	N	(%)[a]	N	(%)[a]
SBP participation	1,259	21.2	557	14.5	1,816	18.3
SBP nonparticipation	1,017	30.0	801	35.5	1,818	32.4
No SBP offered	1,174	48.8	742	50.0	1,916	49.3
Total[b]	3,450	100	2,100	100	5,550	100

Source: NHANES III, 1988-1994.
SBP = School Breakfast Program.
[a]Weighted to reflect the population based on the household interview weight.
[b]Excludes 186 children (3.2 percent) with missing information on SBP participation.

Table IV.7 presents the available sample sizes for 3 SBP participation groups by 2 income levels: (1) lower-income, defined as a family income at or below 185 percent of the poverty line and eligible for free or reduced-price school meals; and (2) higher income, defined as a family income above 185 percent of the poverty line. Weighted population data show that about 44 percent of the total population had a family income at or below 185 percent of the poverty line, and that 49 percent of the total population attended schools that did not offer the SBP. Only about 7 percent of those in the higher income group participated in the SBP.

Table IV.7—Sample sizes of children aged 6 to 16 years, by SBP participation, income level

SBP participation	Lower income[a]		Higher income[b]		Total	
	Sample size	Population percentage[c]	Sample size	Population percentage[c]	Sample size	Population percentage[c]
SBP participant	1,424	32.2	236	7.1	1,660	18.2
SBP nonparticipant	940	32.3	737	33.0	1,677	32.7
No SBP offered	797	35.5	933	59.9	1,730	49.1
Total (percent)[d]	**3,161**	**100 (44.5)**	**1,906**	**100 (55.5)**	**5,067**	**100**

Source: NHANES III, 1988-1994
SBP = School Breakfast Program.
[a]Lower income is defined as a family income at or below 185 percent of the poverty line.
[b]Higher income is defined as a family income above 185 percent of the poverty line.
[c]Weighted to reflect the population based on the household interview weight.
[d]The total sample size reflects a lower sample size than that shown in Table IV.6 due to missing information on family income.

Most children who participated in the household interview and had SBP information were also interviewed in the mobile examination center. Table IV.8 shows that about 91 percent of the total sample that had SBP information had complete, reliable 24-hour recall data; 89 percent had WISC-R and WRAT-R data, and 86 percent had all three data components. Table IV.9 provides the same information as is

shown in Table IV.8, but broken down by age group. Although the level of nonresponse in the survey was relatively low, the examination sampling weights account for nonresponse to the examination component.

Table IV.8—Sample sizes of children aged 6 to 16 years, by SBP participation and dietary and cognitive test score data

SBP participation	Initial sample	Sample size with dietary data	Sample size with WISC-R/ WRAT-R data	Sample size with dietary and WISC-R/ WRAT-R data
SBP participant	1,816	1,678 (92.4%)	1,646 (90.6%)	1,577 (86.8%)
SBP nonparticipant	1,818	1,667 (91.7%)	1,626 (89.4%)	1,583 (87.1%)
No SBP offered	1,916	1,699 (93.6%)	1,667 (87.0%)	1,603 (83.7%)
Total[a]	**5,550**	**5,044 (90.9%)**	**4,939 (89.0%)**	**4,763 (85.8%)**

Source: NHANES III, 1988-1994

Note: Percentages are weighted to reflect the population based on the examination sample weight. If the analysis is restricted to only intakes on school days, the overall sample size will be reduced by 30 percent or more.

SBP = School Breakfast Program; WISC-R = Wechsler Intelligence Test for Children, Revised; WRAT-R = Wide Range Achievement Test, Revised.

[a]Excludes 186 children (3.2 percent) with missing information on school breakfast program participation.

The study sample with dietary and cognitive test score data includes a higher proportion of 6- to 11-year-olds than 12- to 19-year-olds. However sample sizes are adequate to examine both groups separately.

Table IV.9—Sample sizes of children aged 6 to 16 years—SBP participation, dietary and cognitive test score data

	Sample size with dietary data				Sample Size with WISC-R/ WART-R data[a]				Sample size with dietary and WISC-R/WRAT-R data[a]			
	Age 6–11		Age 12–16		Age 6–11		Age 12–16		Age 6–11		Age 12–16	
SBP participation	N	(%)[b]	N	(%)[b]	N	(%)[b]	N	(%)[b]	N	(%)[b]	N	(%)[b]
SBP participant	1,161	22.1	517	14.2	1,148	21.9	498	14.3	1,094	21.6	483	13.9
SBP nonparticipant	929	30.2	738	35.7	914	30.1	712	36.0	886	30.7	801	35.9
No SBP offered	1,039	47.7	660	50.1	1,022	38.0	645	49.7	977	47.7	626	50.2
Total	**3,129**	**100**	**1,915**	**100**	**3,084**	**100**	**1,855**	**100**	**2,957**	**100**	**1,806**	**100**

Source: NHANES III, 1988-1994.
SBP = School Breakfast Program; WISC-R = Wechsler Intelligence Scale for Children, Revised; WRAT-R = Wide Range Achievement Test, Revised.
[a]A child is classified as having WISC-R/WRAT-R data if he or she completed any one of the four subtests.
[b]Weighted to reflect the population based on the examination sample weight.

b. Measurement Issues

SBP Participation. One analytic objective is to examine the mean dietary intakes in the SBP participation groups described earlier. Because parents report children's SBP participation, the measurement may be subject to error. One crosscheck is the 24-hour dietary recall, which captures what was consumed for breakfast and the location where breakfast was consumed. This information could determine whether breakfast was eaten at school on the day of the 24-hour recall.

A significant weakness of this approach is that the information would not be available on all children who also have WISC-R and WRAT-R data—some children are interviewed about their Saturday and Sunday intakes, and some are interviewed during the summer, when school is out of session. School day dietary intakes would be unavailable for at least 30 percent of the total sample. However, for children whose 24-hour dietary recall interviews referred to a school day, we could compare SBP participation on this sample day with the "usual participation" reported by the parents.

Alternative methods could be used to distinguish "usual participation" based on the information collected and based on the desired analysis. One option would define "usual participants" as children who ate a school breakfast on at least 3 of 5 school days. Another option would define SBP participation based on the consumption of two or more food items from the school cafeteria at breakfast.

Participation could be reported based on the sample day that the 24-hour dietary recall measures, or by the weekly or monthly frequency of participation. The day-to-day variability in participation measures must be considered for analysis conducted at the individual-level, in particular (Gleason and Suitor, 2000).

Dietary Intake. Dietary intake is assessed using 24-hour dietary recalls. The proposed analysis proposes to estimate average breakfast intakes and total dietary intakes for groups of children (such as SBP participants, nonparticipants, and nonbreakfast eaters), and to estimate the relative contribution of dietary intakes to nutritional status. The NHANES design includes one 24-hour recall per child and a second 24-hour recall on a subsample.[61]

The second day's intake provides information to adjust nutrient intake distributions using statistical software that takes into consideration the day of the week and within- and between-person variability (Nusser et al., 1996). Adjusted distributions of nutrient intake could be used to estimate the proportion of SBP participants and nonparticipants who met dietary recommendations and dietary adequacy. This approach provides information for comparing group dietary data but does not provide a better measure of individual students' usual dietary intakes for use in regression analysis.

Nutrition and Health Status. The 24-hour dietary recall provides information on current dietary

[61] It should be noted that the meal pattern requirements for a school breakfast changed during the data collection for NHANES III. The requirement to offer two grains/breads, two meat/meat alternates or a combination of both went into effect for school year 1998-1990. The final rule was published on March 30, 1989.

intake at the group level. A single day's intake is insufficient for estimating an individual's "usual intake"; longer-term nutritional status is reflected in biochemical assessments of blood and urine, hematologic determinations, and anthropometric measurements. Nutrition and health outcomes, such as growth, overweight[62] and iron deficiency anemia may relate to children's readiness to learn in school (Briefel at al., 1999; and Chapter II).

Behavioral Measures. Certain behaviors may contribute to a child's overall health status and school attendance. Protective behaviors, such as more physical activity and less time watching television, may improve health and readiness to learn. Certain risk behaviors, such as alcohol and drug use by older children, provide information that may be related to poor school performance, days absent from school, suspensions, and expulsions. These types of lifestyle and behavioral information collected in NHANES III can be used as control or explanatory variables when comparing learning outcomes between SBP participants and nonparticipants.

WISC-R and WRAT-R. Two subtests of the WISC-R, the Block Design and the Digit Span, are indicators of cognitive functioning and academic performance (Kramer et al., 1995). The Block Design is a performance examination; the Digit Span is a verbal test. The WRAT-R was used in the NHANES III to assess academic performance in reading and mathematics. After the WISC-R was given, trained interviewers administered the WRAT-R in the mobile examination center. An automated data collection system ensured that responses were within acceptable ranges. As described in Chapter III, the scores for all four subtests used a common scale and were derived for each child relative to his or her age group based on test-specific samples created by the test developers (Wechsler, 1974; and Jastak and Wilkinson, 1984).

3. Previous Research

Although the NHANES III contains a wealth of information pertinent to the research objectives, little information has been published that describes dietary intakes from breakfast or school breakfast, the cognitive functioning/academic performance tests, or methods to relate the two variables. Kramer et al. (1995) reported on the disparities in cognitive functioning across sociodemographic and health characteristics of children in the United States.

The authors found that lower income, lower education level, and minority status of an adult reference person were independently related to poorer performance on all cognitive subtests. Although less consistent as predictors of test performance, overall health status, a history of birth complications, and sex also were related to cognitive functioning and academic achievement. These results were based on NHANES III data for only the first phase of data collection (1988–1991). Measures of dietary intakes, food sufficiency, or nutritional status were *not* included as variables in the models.

Alaimo et al. (2001a) studied the relationship among poverty, food sufficiency, and health in U.S. schoolchildren, using NHANES III data. Although the study did not assess dietary intakes or school breakfast participation, it found that food-insufficient children were significantly more likely than food-sufficient children to have poorer health status, more frequent stomachaches, and more frequent

[62] Growth and overweight can be assessed using children's height and weight measurements relative to the BMI for age, height-for-age, and weight-for-age percentiles found in the CDC growth charts (Kuczmarski et al., 2000).

headaches. Another analysis of NHANES III relates food insufficiency to cognitive, academic (WISC-R scores and WRAT-R scores), and psychosocial outcomes in children 6–11 years old and 12–16 years old (Alaimo et al., 2001b). After adjusting for confounding variables, 6–11 year old children who were food insufficient were found to have significantly lower math scores and were more likely to have repeated a grade.

4. Analysis Plans

The primary NHANES III analysis for this design would compare mean differences in dietary, nutrition, health, and learning outcomes among the three groups of interest: (1) SBP participants, (2) SBP nonparticipants, and (3) children who do not eat breakfast. Additional subgroups could be analyzed based on whether the child consumes or does not consume a breakfast at home.[63] Descriptive analysis would include, but would not be limited to, comparisons of the following outcomes:

- Mean nutrient and food group intakes for breakfast and for the total day
- Mean dietary quality and variety score, assessed by the Healthy Eating Index
- The proportion meeting current dietary recommendations or dietary requirements for nutrient adequacy, as defined by the Recommended Dietary Allowances, the Dietary Reference Intakes, and the Dietary Guidelines for Americans[64] (U.S. Department of Agriculture and U.S. Department of Health and Human Services, 2000)
- The proportion defined as underweight, at a healthy weight, or overweight based on height and weight measurements and the revised CDC growth charts (Kuczmarski et al., 2000)
- The proportion with iron deficiency anemia or some other vitamin or mineral deficiency, based on biochemical test results
- Mean and distributions of cognitive test scores
- Mean number of missed days of school, tardiness, suspensions, expulsions, and skipped grades

To estimate the effects of SBP participation on these outcomes, the study would use regression analysis to compare the outcomes in participants and nonparticipants after controlling relevant, measurable factors. When comparing academic and learning outcomes, these factors would include food insufficiency; poor nutritional status; and factors that relate to prenatal nutrition, such as low birth weight and exposure to cigarette smoke.

[63] SBP participants and nonparticipants may consume a breakfast at home in addition to or in the place of a school breakfast.

[64] Software from Iowa State University could be used to adjust nutrient intake distributions. The software takes into account the within- and between-person variability in intake and the skewness of nutrient intake distributions (Nusser et al. 1996). However, as described above, it cannot make regression-adjusted comparisons of the proportion of the groups that meet dietary recommendations because the procedure does not generate estimates of individuals' usual dietary intakes.

Environmental exposures, such as to lead, demonstrated by elevated blood lead levels, should be considered when interpreting the results of cognitive tests and academic performance. Thus, the set of control variables that the NHANES III data could provide would be richer in some ways than those available in the ECLS-K. Unlike the ECLS-K, however, the NHANES III dataset provides few school-level control variables (or indicators of which schools sample members attend). Thus, this analysis would not be able to measure or control for some of the institutional factors associated with the schools students attend.

In particular, the model to be estimated to generate estimates of the impact of SBP participation on learning and other outcomes is the following:

$$(7) \qquad Y_{is} = \alpha A_s + \gamma X_{is} + \delta P_{is} + u_s + e_{is},$$

where the variables are defined as in the ECLS-K section of this chapter. (Y_{is} represents the value of a particular outcome for student i in school s; A_s is an indicator of whether the student's schools offers the SBP; X_{is} is a vector of characteristics of student i that influence this outcome; P_{is} is an indicator of whether student i participates in the SBP; and u_s and e_{is} are, respectively, unmeasured school-specific and student-specific factors that influence the outcome.) This equation is analogous to equation (2) from this chapter except that it does not include a vector of school-specific characteristics. (In this case, however, the vector X_{is} includes a richer set of characteristics.)

As long as we assume that there is no selection bias (that is, that the terms u_s and e_{is} are uncorrelated with the independent variables in the model), the estimation of this model is relatively straightforward. If Y_{is} is a continuous variable, OLS regression techniques will produce unbiased estimates of the impact of participation (δ). Because we do not know the extent to which there are multiple observations of students from the same schools, estimating fixed- or random-effects models with NHANES III data is not possible.[65]

Under the equation (7) specification, the model would be estimated on a sample of all NHANES III students, including those not attending SBP schools. To increase the flexibility of the specification somewhat, the participation indicator (P_{is}) could be interacted with other individual characteristics in the model. This specification would allow researchers to investigate whether eating a school breakfast had different impacts on students with different characteristics, such as age or gender.

As was the case with the ECLS-K analysis, the possibility of selection bias would have to be considered. There is possible selection bias related to the availability of the SBP in schools, parents' knowledge about the availability of the program, and parents' decisions about their children's participation in it.[66]

[65] Furthermore, the NHANES III sample is unlikely to contain large numbers of students from the same schools. Thus, estimating fixed- or random-effects models is unlikely to be necessary in this context.

[66] Although schools participating in the SBP must notify parents, parents might not have received or might not have recalled receiving notification.

Accounting for SBP participation decisions in the analysis is important because there is a risk that nonrandom selection into the sample of SBP participants would bias estimates of the impact of SBP participation on learning. To account for selection bias, the approaches described in earlier in this chapter, such as IV models, could be used.

If we believed that selection bias was not present in the analysis, the impacts of SBP participation on nutritional and learning-related outcomes could be estimated with a reasonable level of statistical precision. Using the approach described in the previous chapter, we calculated minimum detectable difference (MDD) sizes (measured as a percentage of a standard deviation in the outcome variable) under the following assumptions:

- A design effect of 2.5
- Control variables in the regression explain about 10 percent of the variation in the outcome
- About 30 percent of the variation in students' participation status can be explained by the control variables
- Sample sizes of 1,259 SBP participants and 1,017 nonparticipants aged 6–11 years (and sample sizes of 1,816 and 1,818, respectively, for those aged 6–16 years), as reported in Table IV.6

Under these assumptions, the NHANES III analysis would have the power to detect effects of SBP participation of as low as 15 percent of a standard deviation of the outcome measure. With respect to the dietary and nutritional outcomes in particular, effects of this size would be reasonable based on other analysis of these variables (National Center for Health Statistics, 1996). Given selection bias, however, MDDs would be much larger (in the range of 0.4 to 0.9); in other words, true SBP participation effects would be much more difficult to detect statistically.

5. Summary

The analytic approach in using data from NHANES III is to describe and compare mean differences in dietary, nutritional, health, and learning outcomes among SBP participants, SBP nonparticipants, and students who do not eat breakfast. To estimate the effects of SBP participation, we propose that regression analysis be used to compare the outcomes of participants and nonparticipants after controlling for all measurable relevant factors (for example, prenatal exposure to smoke, low birthweight, iron deficiency anemia, and elevated blood lead levels).

The NHANES III analysis offers the advantage of an existing national survey of school-aged children with comprehensive information on family background, SBP participation, dietary intakes, nutritional status, and health, as well as cognitive and academic performance test scores. The NHANES III provides a framework for linking cognitive functioning, academic performance, and school behavior with SBP participation, and for assessing such intermediate variables as dietary intakes and nutrition and health status. The NHANES III analysis would capture many of the important domains necessary to link SBP participation and learning. It would therefore be important for informing and supplementing the ECLS-K design and analysis plans.