Nutrition Assistance Program Report Series
Office of Policy Support

Program Error in the National School Lunch Program and School Breakfast Program: Findings from the Second Access, Participation, Eligibility and Certification Study (APEC II)

Volume 2: Appendices

United States Department of Agriculture

May 2015

USDA is an Equal Opportunity Provider.

United States Department of Agriculture

May 2015
Special Nutrition Programs

Program Error in the National School Lunch Program and School Breakfast Program: Findings from the Second Access, Participation, Eligibility and Certification Study (APEC II)

Volume 2: Appendices

Authors:

Mathematica

Quinn Moore
Judith Cannon
Dallas Dotter
Esa Eslami
John Hall
Joanne Lee
Alicia Leonard
Nora Paxton
Michael Ponza
Emily Weaver
Eric Zeidman

Westat

Mustafa Karakus
Roline Milfort

Submitted by:
Mathematica Policy Research
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005

Submitted to:
Office of Policy Support
USDA Food and Nutrition Service
3101 Park Center Dr., Room 1014
Alexandria, VA 22302-1500

Project Director:
Eric Zeidman

Project Officer:
Joseph F. Robare

This study was conducted under GSA Contract AG-3198-C-12-0001/AG-3198-D-12-0010 with the Food and Nutrition Service. This report is available on the Food and Nutrition Service website: http://www.fns.usda.gov/ops/research-and-analysis

Suggested Citation:
U.S. Department of Agriculture, Food and Nutrition Service, Office of Policy Support, *Program Error in the National School Lunch Program and School Breakfast Program: Findings from the Second Access, Participation, Eligibility and Certification Study (APEC II) Volume 2: Appendices* by Quinn Moore, Judith Cannon, Dallas Dotter, Esa Eslami, John Hall, Joanne Lee, Alicia Leonard, Nora Paxton, Michael Ponza, Emily Weaver, Eric Zeidman, Mustafa Karakus, Roline Milfort. Project Officer Joseph F. Robare. Alexandria, VA: May 2015.

This page has been left blank for double-sided copying.

CONTENTS

APPENDIX A: SAMPLING AND DATA COLLECTION PROCEDURES A.1

APPENDIX B: CONSTRUCTING ANALYTIC WEIGHTS FOR THE SECOND ACCESS, PARTICIPATION, ELIGIBILITY, AND CERTIFICATION DATA B.1

APPENDIX C: MATCHING PROCEDURES FOR COMMUNITY ELIGIBILITY PROVISION ANALYSIS C.1

APPENDIX D: STUDENT, SCHOOL, AND SFA CHARACTERISTICS AND OTHER SUPPLEMENTAL TABLES D.1

APPENDIX E: IMPUTATIONS FOR MEAL PARTICIPATION AND INCOME SOURCES E.1

APPENDIX F: ALTERNATIVE DEFINITIONS OF CERTIFICATION ERROR F.1

This page has been left blank for double-sided copying.

TABLES

A.1 Sample summary for APEC-II A.5

B.1 Values of $W2SFAbase_{is}$ for base sample triplets of SFAs B.4

B.2 Values of $W2SFACEP_{is}$ for pairs of SFAs B.6

B.3 Reimbursements in the NSLP estimated for key subpopulations of schools B.12

B.4 Reimbursements in the SBP estimated for key subpopulations of schools B.13

C.1 Examples of combinations of variables required for a deterministic match C.6

C.2 Examples of deterministic matching variables C.6

C.3 Blocking variables for probabilistic matching C.7

C.4 Examples of rules for accepting or rejecting matches C.8

C.5 Percentage of sampled student matches by source and method C.9

C.6 National estimates of improper payments in the NSLP and SBP for schools using CEP excluding follow-up matches and other sources C.10

D.1 Characteristics of students, by certification status D.5

D.2 Characteristics of schools, by CEP status D.7

D.3 Characteristics of schools, by CEP status D.8

D.4 NSLP and SBP meal program characteristics, by CEP status D.9

D.5 NSLP and SBP meal program characteristics, by CEP status D.11

D.6 Characteristics of school food authorities, by CEP status D.13

D.7 Characteristics of school food authorities, by CEP status D.14

D.8 NSLP and SBP meal program characteristics, by CEP status D.15

D.9 NSLP and SBP meal program characteristics, by CEP status D.16

D.10 Characteristics of SFA school meal program operations, by CEP status D.17

D.11 Characteristics of SFA school meal program operations, by CEP status D.18

E.1 Availability of administrative data on NSLP participation E.4

E.2 Proportion of days student was reported to have received a reimbursable meal during survey reference period E.6

E.3 Distribution of actual versus imputed meals received among students with administrative participation records E.10

E.4 Prevalence of missing income data, by income source E.11

F.1 Household size and eligibility for free or reduced-price meals, under alternative specifications of economic unit F.4

F.2 Improper payments due to certification error in the NSLP and SBP, under alternative specifications of eligibility F.6

F.3 Improper payments due to certification error in the NSLP and SBP, using alternative meal imputation for denied applicants with erroneous certification statuses F.7

FIGURES

D.1a National estimates of improper payment rates due to meal claiming error in the NSLP by CEP status D.19

D.1b National estimates of improper payment rates due to meal claiming error in the SBP by CEP status D.20

D.2a National estimates of improper payment rates due to aggregation error in the NSLP by CEP status: point-of-sale error D.21

D.2b National estimates of improper payment rates due to aggregation error in the SBP by CEP status: point-of-sale error D.22

D.3a National estimates of improper payment rates due to aggregation error in the NSLP by CEP status: school reports of meal counts to the SFA D.23

D.3b National estimates of improper payment rates due to aggregation error in the SBP by CEP status: school reports of meal counts to the SFA D.24

D.4 National estimates of improper payment rates due to aggregation error in the SBP by CEP status: SFA reports of meal counts to the State agency D.25

This page has been left blank for double-sided copying.

APPENDIX A

SAMPLING AND DATA COLLECTION PROCEDURES

This page has been left blank for double-sided copying.

SAMPLING AND DATA COLLECTION PROCEDURES

A. Sampling procedures

The APEC-II study used a multistage-clustered sample design that first sampled School Food Authorities (SFAs), then schools served by the SFAs, and then students who attend the sampled schools. We obtained survey data for the study from the entities at each of these levels of sampling. We selected two independent multistage samples: (1) SFAs and schools participating in the Community Eligibility Provision (CEP) and (2) all other SFAs and schools. Below we refer to these groups as the CEP sample and the base (or non-CEP) sample.

For the base sample, the primary sampling unit (PSU) in the multistage design is the SFA. In the first step of selecting each sample, we selected SFAs from a sampling frame: a list of SFAs in the contiguous United States that was based on data from the Form FNS-742 (Verification Summary Report) file. After we selected the initial SFA sample, we merged it with the National Center for Education Statistics (NCES) Common Core of Data (CCD) district-level file to obtain locating and other information for public SFAs and with the NCES Private School Survey (PSS) files for private SFAs. For the CEP sample, the PSU was also the SFA. The SFA sampling frame consisted of SFAs in the six States and the District of Columbia that included at least one school participating in the CEP.

We initially selected 98 SFAs for the CEP sample and 390 for the base (non-CEP) sample. For the CEP sample, half of those sampled were randomly designated as main, and the rest were designated replacement selections; for the base sample, one-third of the selections were randomly chosen as main and two-thirds as replacement selections. Below, we describe the process for dividing the PSUs into main and replacement selections. We contacted the replacement SFA selections when main selections did not participate in the study.

Within each SFA that was sampled and agreed to participate in the study, we selected a sample of schools. The number of schools selected depended on whether the SFA had been selected with certainty and whether it had been selected for the base sample, the CEP sample, or both.[1,2] If the district included enough schools, the sampled schools were randomly divided into main or replacement selections; replacements were used if main school selections did not participate.

Students attending sampled schools were sampled from records provided by the sampled SFA. The student samples were selected from applications or lists of applicants for free or reduced-price meal benefits or students directly certified (without application) for free or reduced-price meals. The CEP student sample was also selected from a provided list of "full-price" students who never applied for free or reduced-price meal benefits (nonapplicants). The base student sample included students who were certified for free or reduced-price meals as well as those whose application was denied. Samples of certified students and denied applicants were

[1] SFAs were sampled with probability proportional to the number of their students certified for free or reduced-price meals.

[2] Some SFAs had both CEP and non-CEP schools and thus had a chance to be selected for both samples.

selected in the fall. In the spring, SFAs were contacted for an additional sample of students newly certified for free or reduced-price meals. The CEP student sample included students certified with and without an application, as well as students who were not certified, and were selected from the CEP reference year. In addition, student samples of meal transactions at schools were also selected.

The sample for the household interview was a subsample of the base student sample. All sampled students were eligible for the household interview with the exception of those sampled in a Provision 2 school in a non-base year. When only lunch or breakfast was Provision 2 in a non-base year, separate student samples were selected for each meal, and eligibility for the household interview was based on each meal's circumstance under the provision.

Below we provide additional detail on how we selected the APEC samples.

1. Selecting SFAs

The overall respondent universe includes all public and private SFAs and schools participating in the National School Lunch Program (NSLP) or School Breakfast Program (SBP) that are located in the contiguous 48 States and the District of Columbia, their students certified for free and reduced-price meals, and students that applied for and were denied certification. For the CEP sample, the universe includes SFAs with one or more schools participating in CEP, together with all CEP schools and all their students. For the base sample, the universe includes all SFAs except those where all schools participated in CEP, schools not in CEP, and their certified students and denied applicants.

The initial base sample consisted of 390 SFAs and 1,122 schools. We collected data from 130 SFAs, 392 schools, 3,761 certified student households, and 611 denied applicant households. We selected from seven participating States the sample of SFAs and schools participating in the CEP in school year (SY) 2012–2013. To produce estimates of improper payments in districts and schools participating in the CEP, we supplemented the base sample by collecting meal certification records data for 3,240 students in 135 participating schools in 45 SFAs within five selected States.

Table A.1 shows the sizes of the base and CEP samples by sampling stage.

Table A.1. Sample summary for APEC-II

	Base	CEP	Total
SFAs			
Initial sample	390	98	488[a]
Released	146	52	198[a]
Participating	130	45	175[a]
Schools			
Initial sample	1,122	204	1,326
Released	425	140	565
Participating	392	135	527
Base student sample: free and reduced-price households			
Initial sample	6,959	n.a.	6,959
Released	4,720	n.a.	4,720
Participating	3,761	n.a.	3,761
Base student sample: denied applicant households			
Initial sample	1,094	n.a.	1,094
Released	810	n.a.	810
Participating	611	n.a.	611
CEP student sample: identified students[b]	n.a.	1,350[c]	1,350
CEP student sample: certified by application	n.a.	1,080[c]	1,080
CEP student sample: not certified	n.a.	810c	810

Source: APEC-II, unweighted data.

[a] Five SFAs were selected for both base and CEP samples, and four were released and participated in both components of the study.

[b] Identified students are students certified for free meals without an application.

[c] Records data were used for both CEP sampling and analysis. There was no sample release plan or basis for participation beyond the initial sample selection.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; SFA = School Food Authority; n.a. = not applicable.

a. Base sample of SFAs

We employed a stratified design and probability-proportional-to-size (PPS) selection methods to select an initial sample of 390 SFAs (large enough to recruit 130 SFAs) for the base national sample. The first level of stratification for the base (non-CEP) sample was the States. The motivation for using States as strata and selecting a large initial sample is that the selection of the base sample is affected by the fact that the first stage of sampling (SFAs) had to be complete before we could determine which SFAs would be participating in the CEP.

The number of SFAs allocated to each State's sample had to be an integer. We first allocated a portion of the 130 SFAs (target number of participating SFAs) to each State based on its proportion of the study population of students. We then used stochastic allocation to arrive at an integer. For example, if a State had 3.65 percent of the student population, its proportional allocation would be 0.0365 x 130 = 4.745 SFAs. Before sampling, that State was randomly allocated either four or five SFAs for the final sample. [3] We tripled the initial sample size selected in the State to allow for replacement.

[3] It had 74.5 percent chance of being allocated five and a 25.5 percent chance of being allocated four.

We performed the allocation by using SAS PROC SURVEYSELECT to select a hypothetical PPS sample of 130 SFAs using sequential random sampling. [4] The allocation to a State was based on the number of actual selections made by SAS. [5] States too small to be allocated at least two selections—that is, States with less than 1.54 (0.154 x 130 = 2) percent of the study population of students—were grouped into one or more strata and were randomly allocated zero, one, or two SFAs in the sample. [6] States with fewer than two expected selections were grouped into "post-strata" for analysis.

Within each State that had an actual allocation of two or more SFAs, before sampling SFAs, we formed implicit strata based on the prevalence of schools participating in the SBP; the proportion of schools using Provision 2 or 3 (P 2/3); and the proportion of eligible students that were categorically eligible for free meals.

To account for ineligibility and noncooperation, we expanded the initial sample size. [7] Because participation in the CEP was not known at the time the base sample of SFAs was selected, the size of the base sample was three times as large as the number of SFAs to be recruited. The samples of SFAs were selected with PPS. The measure of size (MOS) was the number of students eligible for free or reduced-price meals. If the estimated number of students was not available for an SFA, we imputed it using correlated information from Form FNS-742 and CCD, such as number of schools, State, and whether public or private. We first defined certainty selections, then sampled other SFAs using sequential random sampling in SAS PROC SURVEYSELECT. From this expanded sample, we formed triplets of SFAs.

Within each triplet, one SFA was randomly designated the "main" release, and the other two as replacements to be released only if the main selection was ineligible or noncooperative. We achieved this objective in two ways:

1. We placed selections that were large enough to have been selected with certainty even if the sample had not been expanded (we called these level-one certainties) into triplets with two selected SFAs that were not level-one certainties. The level-one certainty was designated as a main selection and the other two randomly designated as first and second replacements.
2. Other triplets (that did not contain a level-one certainty) could contain other (level-two) certainties and noncertainty selections. Noncertainty selections were grouped based on adjacent zones within these triplets;[8] one SFA was randomly assigned to be the main selection and the others as first or second replacement.

[4] Also called Probability Minimum Replacement or the "Chromy Method."

[5] When calculating sampling weights, we will use the noninteger "expected hits."

[6] A State whose size is between 0 and 0.77 percent of the total would be allocated either zero or one SFA. An SFA with more than 0.77 but less than 1.54 percent would be allocated one or two.

[7] SFAs that have only CEP schools were ineligible for the main sample. It is possible that an SFA on Form FNS-742 would have no eligible students, but that is true of fewer than 1 percent of SFAs.

[8] SFAs that have only CEP schools were ineligible for the main sample. It is possible that an SFA on Form FNS-742 would have no eligible students, but that is true of fewer than 1 percent of SFAs.

b. CEP sample of SFAs

The CEP sample of SFAs was selected in a sample of five of the seven States in which the CEP was being implemented in SY 2012–2013.[9] Five States with multiple SFAs were eligible for the CEP study, and two States each had a single SFA eligible.[10] We selected one multi-SFA State with certainty. We selected three multi-SFA States from the remaining four with PPS, and one single-SFA State of two was selected randomly. States selected for the sample provided information on CEP participation. Selection of SFAs within States was with PPS and a sample large enough (98) to recruit 45 SFAs. The allocation of the 45 targeted SFAs across the five sampled States was as follows: the certainty State was allocated 14 SFAs, the three selected with PPS were each allocated 10, and the one selected at random was allocated one SFA. We selected a sample slightly larger than two times the number needed for the survey and formed pairs of selected SFAs; one was randomly designated as the main selection, and the other was the alternate. We kept a small number of pairs in reserve to allow for instances in which both SFAs in a sampled pair either decline or are ineligible to participate. SFAs that had some but not all schools participating in the CEP were eligible for selection in both the base and CEP samples.[11]

2. Sampling schools

For each SFA selected into the initial (larger) sample, we compiled a sampling frame of schools to select the school sample. The frame for public schools was the school-level CCD for SY 2009–2010. To give newer schools that were not included on the frame a chance for selection, we asked public SFAs in our sample to provide names, enrollment, and program-participation data for schools that had come into existence since the last CCD was compiled.

We selected samples designed to yield, on average, 3 schools per SFA, for a total of 398 schools (392 were actually recruited) in the base sample. In all SFAs, we stratified on level (elementary versus middle and high schools). As with SFAs, we selected a replacement sample of schools. To create the pool of replacements, we selected samples twice as large as needed in each explicit stratum and randomly assigned half of each sample to serve as substitutes in case of ineligibility or refusal to participate. Our selection method was PPS with the MOS being an estimate of the number of study-eligible students.

The sampling of schools for the CEP sample was completed using a method similar to the base sample selection. We limited selection to schools participating in CEP in SFAs in which not all schools were participating. We selected samples designed to yield, on average, 3 schools per SFA for a total of 135 schools in the CEP sample.

[9] If we had been selecting from a larger study population, we would treat the States as PSUs for the CEP sample. However, this approach would complicate estimation for the purposes of making statements about a program that was in the early stages of implementation. Therefore, we treat the States as strata and the SFAs as PSUs. Under this conception of the sample, statistical inferences for the CEP sample can be made for the five States but will be more stable relative to the alternative approach.

[10] SFAs participating in intensive components of a simultaneous study of CEP were treated as ineligible for selection to avoid excessive burden.

[11] Five SFAs were selected for both CEP and base samples.

3. Students

We sampled students in all selected schools in the study, using similar methodologies in the base and CEP samples.

Sampling students for the base sample. After we selected schools, we visited SFAs and asked them to provide student lists needed for student stratification and sampling. Selection of students for the base sample differed by whether the school participated in Provision 2. Participation in Provision 2 varied across schools in the base sample. Some schools participated in Provision 2 for one meal and not the other. Similarly, the base year for Provision 2 participation in these schools may have varied across SBP and NSLP. If the Provision 2 status and base year were not the same for both meals in a school (breakfast and lunch), we selected two independent samples of students—one for each meal. Students selected for a meal participating in Provision 2 in a non-base year were not eligible for the household survey.

Visits to the base sample SFAs for student sampling took place within the first month of SY 2012–2013. This sample consisted of students falling within two frames: (1) those certified for free or reduced-price meals and (2) those who applied whose applications were denied. Field staff met with SFA directors and obtained lists of students who were approved to receive free or reduced-price meals at each study school at the time of their visit. The lists of free-certified students included directly certified students and students certified by application. The staff then processed the lists to remove any ineligible students (for instance, denied applicants, nonapplicants, those who did not attend the school this year, and those who already had a sibling selected in the sample). Field staff then counted the total number of students certified for free or reduced-price meals and entered this information into an Excel program loaded onto their laptop computers. The program selected the sample of certified students for each study school, selecting eight main selections (and eight replacement selections).

Although SFAs keep clear records of which students were approved to receive free or reduced-price meals, subgroups of applications that were not approved were more complicated to track because SFAs did not always maintain lists of these students. When the SFA was without a list of denied applicants, field staff obtained hard copies of denied applications from the SFA or school. Staff used the same sampling program on their laptop computers that they used to select the certified sample to select denied applicants, but the program instead selected two main selections and two replacements.

After student sampling was complete, household contact information was collected for all students eligible for the household interview. Field staff requested either a roster or the application, if necessary, to obtain the names and contact information of the parents of each student selected. They then entered this information onto the sampling information forms and transmitted it securely to Mathematica's central office.

Beginning in December 2012 and continuing throughout the remainder of the school year on a rolling basis, another student sample was selected for each school for meals that were either not participating in Provision 2 or were participating in Provision 2 and in their base year. This sample consisted of students who had been certified for free or reduced-price meals in the prior two months. This sampling effort was done remotely by Mathematica staff who contacted SFAs and requested that they provide the student listing electronically. The sampling was then

conducted using the same procedures and sampling program used by the field staff in the fall. These students were also sampled and attempted for the household survey. Although households could have been selected into both samples, we chose not to complete surveys in the spring with households who completed a survey in the fall, to limit the burden on the respondent. In this approach, we checked lists of students against lists of completed cases in the fall and marked students who completed in the fall as ineligible. In total, the participating student sample consisted of 3,761 free and reduced-price students and 611 denied applicants.

Sampling students for the CEP sample. After schools were selected, we asked SFAs to provide student lists electronically from the reference year on which they based their claiming rates. We then divided these lists into three frames: (1) students who had been identified, or directly certified under CEP without an application; (2) students who were certified by application; and (3) nonapplicants or denied applicants, deemed "paid" or "full-price" students. These designations were all based on student status on April 1 of the reference year. Using the same sampling program as for the base sample, we selected 24 students from each school (10 identified students, 8 students who were certified by application, and 6 students not certified). We requested administrative data from SFAs for each sampled student.

Meal transaction observations. Prior to their arrival at each of the sampled schools, Westat staff collected information from the school food service manager (FSM) about the school's point-of-service system, to help them plan for the selection of the random sample of meal transactions for observation. When on site, Westat staff entered into their laptop computers information related to each cash register's period of operation and transaction volume prior to the meal service. The computer randomly selected the registers to be observed during specific periods and the interval samples of individuals to be observed. Field staff then used hard-copy forms to record (1) items on each tray and the amounts of each item, (2) whether the transaction involved a student or nonstudent/adult, and (3) whether the school recorded the tray as a reimbursable meal. The sampled meal transactions could include reimbursable meals obtained by students approved for free and reduced-price meals and students playing full price. Field staff were not stationed at à la carte-only lines, but if à la carte meals could be purchased in the same lines as reimbursable meals, the transactions were included as candidates for selection. Field staff observed approximately 25,000 lunch transactions from 436 schools and 23,000 breakfast transactions from 421 schools participating in SBP.

Data collection procedures

Mathematica in coordination with Westat and Decision Information Resources (DIR) completed data collection for APEC-II from August 2012 through April 2014. The procedures used to collect these data are outlined below.

- **SFA Director Survey.** The SFA Director Survey was a self-administered survey that collected information on the characteristics of each sampled SFA and sampled schools in the base and CEP samples. The requested information included institutional characteristics, such as grade span and enrollment, as well as information on participation in the meal programs, certification outcomes, and direct certification implementation. For CEP SFAs and schools, we also collected information such as the number of identified students and meal claiming percentages.

- Westat sent questionnaires via email to the person identified as the most knowledgeable about the district's administrative practices regarding the school meal programs—typically the district's food service director—beginning in March 2013. Westat requested the SFAs return the questionnaires via fax, email, or FedEx. We sent nonresponding SFAs multiple reminders by email and phone to complete the survey. To ensure that as many SFAs as possible reported the most important information for analysis, critical survey items were identified, and nonresponding SFAs were contacted with an abbreviated survey. SFAs could complete this abbreviated survey via email or phone.

- **Household survey data.** Field interviewers completed household surveys with the parent or guardian of students selected in our certified (free and reduced-price) and denied applicant samples using computer-assisted personal interviewing (CAPI) on Mathematica-provided laptops. Household surveys were conducted for all students in the base sample, except for those students selected from schools participating in Provision 2 in a non-base year (for either breakfast or lunch). The household survey included a common set of questions collecting information on household composition; income sources with supporting documentation; the sampled students' participation in SBP and NSLP; perceptions of meal program quality; participation in SNAP, TANF, and other assistance programs; and demographic information. After the student sample was selected and transmitted to the Mathematica central office, advance letters and study brochures were mailed to sampled households. The advance letters, which were printed on U.S. Department of Agriculture (USDA) letterhead, described the purpose of the study, including the time burden, incentive payment amount, and information on the confidentiality of the household responses.

  During September through December 2012, Mathematica and DIR visited sampled schools and SFAs once to complete interviews with households of students certified for free- and reduced-price meals and denied applicants. During the remainder of the school year (January to June 2013), Mathematica and DIR completed interviews with newly certified applicants from each study school during a second visit to the district. On average, each visit to an SFA lasted approximately one week.

- **Application/direct certification records data abstraction.** We collected data appearing on meal program applications and direct certification documents for all students in the base sample. Key data items abstracted included identification information such as the student name; application date; number of people in the household; foster child status; income information: participation in SNAP, TANF, or Food Distribution Program on Indian Reservations (FDPIR); and the SFA's eligibility determination. These data were collected directly from student applications and direct certification documents whenever possible. If these documents were not available on the application, we requested missing information from the SFA, which frequently provided screenshots from its data management system.

- Westat staff performed the data abstraction during its weeklong visits to each SFA, which occurred on a rolling basis throughout the school year. In addition to the application abstraction activities, Westat staff also performed during this visit the meal-counting-and-claiming activities described below. Mathematica staff requested that SFAs that had already been visited by Westat staff send applications and direct certification data electronically for any newly certified students sampled; thus, Westat did not need to make an additional in-person visit. Westat staff entered these data into an electronic standardized data abstraction

form on their Westat-provided laptop. If the SFA allowed it, staff photocopied documentation and returned it to the Westat central office.

- **Changes in certification status and enrollment.** SFAs were contacted prior to the end of the school year for information on changes in each sampled student's certification status throughout the school year, as well any changes in their enrollment status in the sampled school. This information was requested for all students in the base sample, except for those students selected for a meal participating in Provision 2 in a non-base year.
- **NSLP/SBP individual student-level participation data.** We requested data on student-level meal program participation for sampled students from each SFA. Wherever possible, we requested daily or monthly participation information for the entire year. This information was requested for all students in the base sample, except for those students selected for a meal participating in Provision 2 in a non-base year. If SFAs contained both base sample and CEP schools, they received the request for only the students in the base sample. Data were received in Excel formats as well as in hard-copy forms that required data entry into a standardized format. Also, a small number of SFAs were unable to provide these data because they were not tracked at the individual level.
- **CEP student matching data.** In schools and districts participating in the CEP, we determined the accuracy of the identified student percentage by comparing the sampled students with SNAP/TANF program participation lists collected from State or local agencies. We also requested from appropriate agencies or directly from the local education agency (LEA) lists of students in foster care, on a homeless liaison list, or income eligible for Head Start; runaways; or migrant youth. The student sample was compared with all received lists. If the LEA did not have this information and the State or local agencies could not share the lists with us for confidentiality reasons, we provided them with lists of our sampled students so they could indicate who participated in their programs. See Appendix C for additional details on the CEP student matching process.
- **Meal count and claiming data.** During their weeklong visit in each SFA, Westat staff collected information on meal transactions through on-site observations in each sampled school and data on aggregation—counting, consolidating, and claiming meal reimbursements—for all of the base sample schools as well as a subsample of the CEP sample schools. The procedures used in these data collection efforts varied from school to school based on the schools' specific procedures and systems.
    - **Meal transaction observations.** Prior to their arrival at each of the sampled schools, Westat staff collected information from the school food service manager (FSM) on the school's point-of-service system, to help them plan for the selection of the random sample for their observation. When on site, Westat staff entered information into their laptop computers for each cash register by meal and transaction volume prior to the meal service. The computer randomly selected the registers to be observed during specific periods and the interval samples of individuals to be observed. Field staff then used hard-copy forms to record (1) items on each tray and the amounts of each item; (2) whether the transaction involved a student, nonstudent, or other adult; and (3) whether the school recorded the tray as a reimbursable meal. The sampled meal transactions could include reimbursable meals obtained by students certified for free and reduced-price meals and students paying full price.

Field staff were not stationed at à la carte-only lines, but if à la carte meals could be purchased in the same lines as reimbursable meals, the transactions were included as candidates for selection.

- **Aggregation data.** Westat collected during its weeklong site visit data on sources of aggregation error, specifically the counting, consolidating, and claiming of meal reimbursements for each sampled school and SFA. These data were collected for a target week (the previous completed full week prior to the visit to the school) and target month (the prior month). Given the spread of visits throughout the year, the target weeks/months were distributed across the school year. Data were collected for each source of aggregation error using the following steps:
  - **Daily counts for target week.** Westat staff contacted each school's FSM to obtain data on the target week meal counts separately from all points of sale, as well as the total daily count recorded for the daily report the school compiles each day. Staff entered the information onto specially designed forms broken down by meal reimbursement status—free, reduced-price, and paid. Staff also validated the school's daily meal counts for the target week using the same procedure as the food service worker (for example, counting tickets in a ticket system or counting checkmarks in a roster check-off system).
  - **Monthly counts.** Westat staff requested data in the same report formats for the target month. Staff entered the information onto specially designed forms broken down by meal reimbursement status—free, reduced-price, and paid. Field staff did not validate these meal counts.
  - **District reimbursement claims for sampled school.** Westat staff also collected data from the district covering the same target week and month to determine whether the SFA accurately claimed meals for reimbursements for the sampled school when it submitted the claim to its State agency. Staff requested (1) records of the breakfast and lunch counts for the target week and month that the school submitted to the SFA, (2) documentation showing the number of breakfasts and lunches the SFA claimed for reimbursement for the sampled school when submitting the claim to the State agency, and (3) information about adjustments a sampled school or SFA intentionally makes to the counts. Staff entered the information onto specially designed forms broken down by meal reimbursement status—free, reduced-price, and paid—either weekly or monthly depending on how the claims were submitted to the State.
  - **District consolidation and claims across all schools**. In addition to the forms listed above, for the target month, staff requested data from the district on (1) the separate meal counts by type that each school submitted to the district and (2) the total meal counts reported (claimed) by the district to the State agency for meal reimbursement, to determine aggregation error from this source.

In addition to these data, Mathematica also collected data on number of students in the meal pricing categories (free, reduced-price); enrollment; daily attendance; and number of serving days to help us assess the accuracy of the meal counts.

- **Extant data.** APEC-II compiled various extant data sources for sampling and weighting purposes, as well in support of forthcoming modeling. These sources included the SFA Verification Summary Reports (Form FNS-742) and the FNS national data bank provided by FNS, public-use data from the CCD and Private Sector Survey from the National Center of Education Statistics' website, census data, and total yearly meal counts for sampled SFAs requested from State education agencies.

This page has been left blank for double-sided copying.

APPENDIX B

CONSTRUCTING ANALYTIC WEIGHTS FOR THE SECOND ACCESS, PARTICIPATION, ELIGIBILITY, AND CERTIFICATION DATA

This page has been left blank for double-sided copying.

Mathematica Policy Research constructed weights at three levels: school food authorities (SFAs), schools, and students. The weights at the three levels are not independent. In general, the final weight for the SFAs served as the initial weight at the school level, and the final school weight served as the initial weight for the student-level data. For schools, in addition to a basic weight constructed to use for national estimates and as the basis for student-level weights, we constructed eight weights at the school level for analysis of four types of non-certification error: meal claiming error, point-of-sale aggregation error, school-to-SFA report aggregation error, and SFA-to-State-agency meal claim aggregation error. For each type of non-certification error, we constructed separate weights for breakfast and for lunch. At the student level, we constructed weights for analysis of application data and for household survey data.

The rest of this appendix describes these weights. The discussion of weighting frequently refers to the sample selection process, which is described in Appendix A.

A. SFA-level weights

The SFA weights were calculated separately for the base and Community Eligibility Provision samples and then combined by post-stratifying each sample to the total size of its population (non-CEP or CEP) of SFAs so that when they were combined each sample represented its portion of the total population of SFAs.

For each sample, the initial weight for each SFA is its sampling weight, which is the inverse of the SFA's probability of selection P(is).

The SFA weight also incorporates the following:

- Adjustments to reflect the release of SFAs,[12] including backups
- Adjustments for nonresponse not accounted for by the adjustments described previously for release of backups
- Post-stratification to externally estimated totals of all SFAs

1. Base sample SFA weights

For the base sample, as discussed in Appendix A, we used stochastic allocation to determine the number of SFAs to select in each State (States were the strata) because the desired allocation was not an integer. Each stratum's allocation of SFAs for the sample was probabilistically assigned to be either the integer below or above the actual value. However, an SFA's probability of selection is based on the original, noninteger number of selections. Thus, if a State's initial allocation was a SFAs, the actual number to sampled was either

Int(a) or

Int(a) + 1

where Int(a) is the integer portion of a.

[12] As discussed in Appendix A, not all initially sampled SFAs were released for contact.

Because of the method used in the actual allocation (see Appendix A for details), the probability of selection SFA *i* in State *s* is

(1) P(*is*) = 1.0 if the SFA was large enough to be selected with certainty and

$$= \frac{a'_s(MOS_{is})}{\sum_{a'_s=1}^{A'} MOS_{is}} \qquad \text{otherwise,}$$

where a'_{s1} is the number of noncertainty SFA selections (a_s minus the number of certainty selections), and a' is the number of SFAs in *s*, minus the number of certainty selections, and MOS_{is} is the measure of size for the SFAs.

(2) $\text{W1SFA base}_{is} = 1/P(is)$

Nonresponse adjustments. The first adjustment to the base weight is for release and nonresponse within release groups. In the main sample, we assigned each sampled SFA to a release group that contained three similar SFAs (a triplet). Among the triplets of SFAs, we randomly selected one[13] as the main SFA and the others as backups to contact if the main selection did not participate. The next adjustment accounts for this subselection and for any nonresponse within the triplet.

Table B.1 shows the values of the group adjustment factor, W2SFAbase_{is}.

Table B.1. Values of $W2SFAbase_{is}$ for base sample triplets of SFAs

Within triplet		
Released	Completed	W2SFAbase
1	1	3 for the complete; 0 for the other two
2	2	1.5 for each complete; 0 for the other
2	1	3 for the complete; 0 for the other two
2	0	1.5 for each released (pending later nonresponse adjustment); 0 for the other
3	3	1 for each complete
3	2	1.5 for each complete; 0 for the other
3	1	3 for the complete; 0 for the other two
3	0	1 for each released (pending later nonresponse adjustment)

Note: The sum of $W2SFA_{is}$ for a triplet will always equal 3; when only one district in a triplet was released, $W2SFA_{is}$ reflects subsampling within the triplet; if all three were released, the weight reflects no subsampling within the triplet; but if one of the released triplet was not completed, $W2SFA_{is}$ reflects nonresponse within the triplet.

The SFA base sample weight adjusted for release of sample is

(3) $PREWTSFAbase_{is} = W1SFAbase_{is} * W2SFAbase_{is}$

[13] Some SFAs, called level-one certainty selections, were large enough to be selected with certainty even if we had selected no back-up sample. These level-one certainties were always the main selection in their release group.

The release adjustment accounts for nonresponse within the triplet (Table B.1). If there are no completes in a triplet, this nonresponse is adjusted for in the next step.

The next step planned to form cells to adjust for nonresponse (not already accounted for in the triplet adjustments). However, the group adjustments accounted for all nonresponse for the base sample because each triplet had at least one SFA participating in the studies.[14]

Because there was no additional nonresponse, the base sample SFA weight adjusted for nonresponse is

(4) $SFANRWTbase_{is} = PREWTSFAbase_{is}$ if SFA is complete, and 0 otherwise.

2. CEP SFA weights

We constructed the weights for the CEP sample in a similar manner to that used for the base sample SFAs, with three differences:

- There was no allocation to States as for the base sample.
- There was an adjustment because only five of seven CEP States were selected to be part of the study.
- The release adjustment factor was based on pairs rather than triplets.

Thus, the weights for CEP SFA were constructed as

(5) $W1PSFACEP_{is} = 1/P(S \text{ in CEP sample}) * (1/PCEP_{is})$

where P(*S* in CEP) is the probability of a CEP State being selected for the study and $PCEP_{is}$ is the probability of SFA_{ks} being selected in State *s*.

P(*S* in CEP sample) = 5/6

$PSFACEP_{is}$ = 1 if the SFA was large enough to be selected with certainty and

$= a'(CEP_s)\ MOS_{is}) / \sum_{is \in CEP, noncert} MOS_{is}$ otherwise.

$a'(CEP_s)$ = number (noncertainty) of CEP SFAs to be selected in State *s*.

MOS_{is} is the SFA measure of size, and

$\sum_{is \in CEP, noncert} MOS_{is}$ is the sum of the MOS_{is} for all CEP SFAs not selected with certainty.

[14] We retain discussion of this step because there were such nonresponse adjustments for the CEP sample (discussed later in this appendix).

The next adjustment was for release within groups. The release groups for the CEP comprised pairs of SFAs; Table B.2 shows those adjustments.

Table B.2. Values of $W2SFACEP_{is}$ for pairs of SFAs

Within a pair		
Released	Completed	W2SFACEP
1	1	2 for the complete SFA (based on *1/p; p = 1/2*); 0 for the other
1	0	2 for the SFA released (pending nonresponse adjustment); 0 for the other
2	2	1 for each of the complete SFAs
2	1	2 for the completed SFA (*1/p x 1/rr* where *p = 1 and rr = 1/2*); 0 for the others
2	0	1 for each of the released SFAs (pending nonresponse adjustment)

Note: The sum of $W2SFA_{is}$ for a pair will always equal 2; when only one district in a pair was released, $W2SFA_{is}$ reflects subsampling within the pair; if both were released, the weight reflects no subsampling within the pair; but if one of the pair was not completed, $W2SFA_{is}$ reflects nonresponse within the pair.

The pair adjustments did not account for all nonresponse, so we constructed another adjustment.

(6) *RRCEP* = (number of CEP SFAs completed)/(number of CEP SFAs released and not ineligible)

(7) $W3SFACEP = 1/RRCEP$

(8) $SFACEPNRWT_{is} = W1SFACEP_{is} * W2SFACEP_{is} * W3SFACEP$

If an SFA responded for both the base and the CEP sample, its final weight was the average of the two weights it would receive as part of the base sample and as part of the CEP sample.

3. Weights to combine the base and CEP samples

After we computed the weights separately for the base and CEP samples of SFAs, we post-stratified each sample so that the weights reflected the estimated totals of CEP and non-CEP SFAs. After this post-stratification, the combined samples can be used to make estimates for all SFAs (in other words, the combined samples are now representative of all SFAs).

We made the post-stratification adjustment within the Food and Nutrition Service (FNS) region for the base sample and overall for the CEP.

(9) $PSTRAT = N'NON_r / \sum_{is\,complete} SFANRWTbase_{is}$ if in base sample only,

$= N'_{CEP} / \sum_{is=complete} SFANRWTWJ_{is}$ if in CEP sample only,

$= N'both / \sum_{is=complete} SFANRWT_{is}AVG$ if in both samples,

where

$N'NON_r$ is the estimated number of SFAs in region R that are not CEP,

$N'CEP$ is the estimated number of SFAs participating in CEP (where all the SFA's schools are in CEP), and

$N'both$ is the estimated number of SFAs that have both CEP and non-CEP schools. $SFANRWT_{is}AVG$ is the average weight discussed above.

The final SFA weight is

(10) $FINSFAWT_{is} = SFANRWTbase_{is} * PSTRAT$ for the base sample

$= SFANRWTCEP_{is} * PSTRAT$ for the CEP sample

$= SFANRWT_{is}AVG * PSTRAT$ for SFAs in both samples

B. School-level weights

We created nine different weights at the school level. In addition to a basic school-level weight, we created eight weights for estimating non-certification error.

- Two meal claiming error weights: one for lunch and one for breakfast
- Six weights for estimating aggregation error: breakfast and lunch weights for point-of-sale, school-to-SFA, and SFA-to-State agency errors

1. Basic school-level weights

The initial weight ($WOSCH_{jhis}$) for any school j in stratum h[15] and SFA_{is} is the variable $FINSFAWT_{is}$ for the SFA of which the school is part. The first adjustment factor, $W1SCH_{ijk}$, is the inverse of the probability of the first phase of selection of the school within its SFA (and within its stratum). We selected schools with probability proportional to size (and within its stratum), and some were large enough to select with certainty. Thus,

(11) $W1SCH_{jhis} = 1/PSCH_{jhis}$

$PSCH_{jhis} = 1$ if school is selected with certainty, and

$$PSCH_{jhis} = \frac{n'_{jhis} MOS_{jhis}}{\sum_{jhis} MOS_{jhis}}, \text{ otherwise,}$$

where

[15]The notation is general, as not all samples were explicitly stratified within SFAs. As discussed in Appendix A, in most SFAs, samples of schools were stratified by level (elementary, middle, or high schools). When there was no stratification, j is a constant and the weights are calculated as if there was one stratum.

MOS_{jhis} is the measure of size for the school *j* in stratum *h* in SFA *i* in State *s*.

n'_{jhis} is the number of noncertainty selections made in stratum *j*, $SFA_{k.}$

N'_{jhis} is the number of schools available for noncertainty selection with PPS in *j* and *k.*

The next factor, $W2SCH_{ijk}$, accounts for the subselection of schools into the main and replacement samples and for release of schools.

In all SFAs we computed the following:

(12) $W2SCH_{ijk} = 1/Prel_{jhis}$

$Prel_{jhis} = nrel_{his}/ninit_{his}$

$nrel_{his}$ is the number of schools released in stratum *h*, SFA_{is}

$ninit_{his}$ is the number of schools initially selected in stratum *h*, SFA_{is}

(13) $WPRELIM_{jhis} = WOSCH_{jhis} * W1SCH_{jhis} * W2SCH_{jhis}$

We then post-stratified the school weights so that the sum of weights for completed schools is consistent with our best estimate of the number of study-eligible schools in SFAs using the NSLP or SBP.

We computed separate post-stratification factors for four groups defined as private schools and three groups of public schools:

- Those with enrollments fewer than 500
- Those with enrollments from 500 to 999
- Those with enrollments larger than 999

For each group,

(14) $PSTRAT_G = N'_G \sum_{jhis \in g} WPRELIM_{jhis}$

(15) $WSCHPS_{jhis} = WPRELIM_{jhis} * PSTRAT_G$,

where

N'_G is the estimated number of schools in the group.

2. Weights for estimating non-certification errors

In addition to the basic school weight, additional school-level weights were constructed for the four types of non-certification error: meal claiming error, point-of-sale error, school-to-SFA

reporting error, and SFA-to-State agency reporting error. Each type of non-certification error had two sets of weights, which were used to generate separate estimates of school-level characteristics that either were not related or were related to student meal program participation.

For analyses of school characteristics that were unrelated to student meal program participation, such as calculations of the national percentage of schools using Offer-versus-Serve, the weight for error type *e* and meal *m* (separate weights were constructed for breakfast and lunch) is the post-stratified school weight adjusted for nonresponse. Thus,

(16) $ERRWT_{jhis \in m} = WSCHPS_{jhis} * RRADJ_{cem}$,

where

(17) $RRADJ_{cem} = 1/RR_{cem}$

(18) $$RR_{cem} = \sum_{\substack{jhis \text{ complete} \\ \text{for } em \in c}} WSCHPS_{jhis} / \sum_{\substack{jhis \text{ basic} \\ \text{complete} \in c}} WSCHPS_{jhis}$$

Nine cells were defined based on geography and CEP status:

- Seven cells were defined, each comprising the schools in an FNS region that were in a State that did not have CEP.
- One cell included base sample schools in CEP States.
- Another cell included CEP schools (in CEP States).

Within each of these cells, we constructed an adjustment for breakfast and one for lunch.

For analyses of school characteristics that were related to student meal program participation, such as estimates of reimbursements and improper payments, we implemented further adjustments to the school-level weights described in equation 16. These adjustments helped align the distribution of students in our school sample to the distribution of students across schools nationally. We defined adjustment factors for nine categories of schools based on their student enrollment (fewer than 500 students; 500–999 students; 1,000 or more students) and the percentage of students that were certified for free meals (low; medium; high). The "low," "medium," and "high" bins of percentages of students certified for free meals were defined within each school enrollment grouping so that the low, medium, and high bins each contained a roughly equal number of students. The adjustment factor for each category was defined as the number of students in schools within that category according to the Common Core of Data, divided by the estimated number of students in schools within that category in our study sample using the weights from equation 16.

C. Student-level weights

We constructed six sets of student-level weights, creating separate weights for breakfast and lunch for each of the three samples selected:

1. Certified applicants selected during the early part of the school year
2. Denied applicants selected during the early part of the school year only
3. Certified applicants selected during the latter part of the school year

The samples of certified and denied applicants included samples in schools using Provision 2 and Provision 3 that were not in their base year; in these schools, the student samples were selected for application data abstraction only.

The initial weighting factor for all the student-level weights was the post-stratified weights for the student's school. So for student *l* in school *j,h,i,s,*

(19) $W0STU_{ljhis} = WSCHPS_{ljhis}$.

The first adjustment for the three groups (*Z*) is the inverse of within-school probability of selection.

(20) $WISTU_{ljhis} = 1/P_{ljhis}$

(21) $P(Z)_{ljhis} = n(Z)_{jhis} / M(Z)_{jhis}$

where

$n(Z)_{jhis}$ is the number of applications sampled for *Z* in school *j,h,i,s*

$M(Z)_{jhis}$ is the estimated total number of applications for *Z* in school *j,h,i,s*

We computed the probability of selection in one step, $n(Z)_{jhis}$, representing the total number for which data collection was attempted. We then defined a preliminary weight adjusted for nonresponse and post-stratified to population totals.

The nonresponse adjustment is as follows:

(22) $PRWTSTU_{ljhis} = WOSTU_{ljhis} * W1STU_{ljhis}$

(23) $RRADJSTU_c = 1/RRSTU_c$

(24) $$RRSTU_c = \sum_{\substack{\text{students} \\ \text{observed} \in c}} PRWSTU_{ljhis} / \sum_{n_{ljhis}=1 \in c}^{n(Z)_{ijk}} PRWSTU_{ljhis}$$

(25) $RRADJWT_{ljhis} = PREWTSTU_{ljhis} * RRADJSTU_c$

The response rate cells were the following:

- All private schools were in one cell.
- For public schools, cells were defined based on whether the school used CEP.

Student weights for APEC II were trimmed using procedures developed for the National Assessment of Educational Progress (see Potter 1990). Trimming cells were defined by meal and sample (certified early, certified late, and denied).

D. Post-stratifying weights

We post-stratified the student- and school-level weights across all analyses of certification and non-certification error in order that our sample-based sums of dollar amounts of SBP and NSLP meal reimbursements would equal national totals based on administrative data maintained by FNS. However, the post-stratification process was complicated somewhat because administrative data do not cover the same time period or present total reimbursements specifically for schools covered by the APEC-II study sample.

For both the NSLP and SBP, we used FNS administrative data for FY 2013 (October 2012 through September 2013) to estimate total reimbursements in the contiguous United States during SY 2012–2013 (August 2012 through July 2013). These calendar periods overlap except in August and September; FNS data on dollar reimbursements in FY 2013 include August 2013 and September 2013, whereas SY 2012–2013 includes August 2012 and September 2012. The dollar amounts in these periods could differ because of any change in meal participation, including random variation, or because meals were reimbursed at slightly higher average reimbursement rates in 2013.

The rest of this section describes the process that was applied to estimate total reimbursements nationally and to subpopulations of schools for analysis (for example, CEP schools, Provision 2 and 3 non-base year schools, and so on). The general approach is the same for NSLP and SBP reimbursements.

1. NSLP reimbursements

Based on FNS administrative data for the NSLP, total reimbursements to schools in FY 2013 equaled $12,214,627,854, and total Section 11 reimbursements for free and reduced-price lunches were $9,352,973,395. This estimate of total reimbursements included the following types of reimbursements that are included in our estimates of improper payments:

- Additional 6-cent reimbursement per meal in districts that claimed at least 60 percent of lunches at the free and reduced-price rates during SY 2010–2011
- Additional 6-cent reimbursement per meal in schools that were certified to be in compliance with the school meal pattern
- The value of donated USDA Foods
- Reimbursements in Provision 2 and 3 schools
- Reimbursements in CEP schools

However, this total also included reimbursements in Alaska, Hawaii, U.S. territories, and Department of Defense, which our estimates of improper payments do not cover. It was straightforward to remove these reimbursements because the FNS administrative data included separate totals by State and U.S. territory. After removing reimbursements in Alaska, Hawaii,

U.S. territories, and Department of Defense, total reimbursements in the 48 contiguous States and the District of Columbia were $11,977,397,682, and Section 11 reimbursements were $9,158,789,330. We further adjusted these numbers to exclude estimated reimbursements made to residential childcare institutions (RCCIs). This adjustment yields a final NSLP reimbursement value of $11,801,419,647 and Section 11 reimbursements of $9,019,695,284.

After estimating reimbursements for non-RCCI NSLP participants in the contiguous United States, we adjusted the school-level analysis weights so that a weighted estimate of total reimbursements for an analysis sample would equal our estimate of total reimbursements from FNS administrative data. We used these calibrated school-level weights to calculate reimbursements for the analyses and corresponding subpopulations of schools shown in Table B.3.

Table B.3. Reimbursements in the NSLP estimated for key subpopulations of schools, SY 2012–2013

Subpopulation	Reimbursements	Analysis of error
Community Eligibility Provision (CEP) schools	$286,089,120	Certification error and non-certification error
Non-CEP schools	$11,515,331,584	Certification error and non-certification error
Provision 2/3 schools in a non-base year (P 2/3 non-base year)	$471,991,200	Combinations of error
Non-CEP and non-P 2/3 non-base year schools	$11,043,339,327	Certification error

NSLP = National School Lunch Program; P 2/3 = Provision 2 or 3 schools; SY = school year.

2. SBP reimbursements

Based on FNS administrative data, total reimbursements under the SBP to schools in FY 2013 were $3,514,024,251. This estimate of total reimbursements included the following types of reimbursements that are included in our estimates of improper payments:

- Additional 30-cent reimbursement per free and reduced-price meal in schools that claimed at least 40 percent of lunches at the free and reduced-price rates during SY 2010–2011
- Reimbursements in Provision 2 and 3 schools
- Reimbursements in CEP schools

However, as with the NSLP reimbursement calculations, this total also included reimbursements in Alaska, Hawaii, U.S. territories, and Department of Defense. After removing these reimbursements, total SBP reimbursements in the 48 contiguous States and the District of Columbia were $3,456,584,472. We further adjusted these numbers to exclude estimated reimbursements made to residential childcare institutions (RCCIs). This adjustment yields a final SBP reimbursement value of $3,339,545,125. We post-stratified the school-level weights to match this reimbursement amount and then these used calibrated school-level weights to calculate reimbursements for the subpopulations of schools shown in Table B.4.

Table B.4. Reimbursements in the SBP estimated for key subpopulations of schools, SY 2012–2013

Subpopulation	Reimbursements	Analysis of error
Community Eligibility Provision (CEP) schools	$117,916,784	Certification error and non-certification error
Non-CEP schools	$3,221,628,416	Certification error and non-certification error
Provision 2/3 schools in a non-base year (P 2/3 non-base year)	$410,546,720	Combinations of error
Non-CEP and non-P 2/3 non-base year schools	$2,811,081,621	Certification error

P 2/3 = Provision 2 or 3 schools; SBP = School Breakfast Program; SY = school year.

REFERENCE

Potter, Francis J. "Methods for Extreme Weights in Sample Surveys." Institute of Statistics Mimeo Series No. 1875T. Chapel Hill, NC: University of North Carolina at Chapel Hill, Department of Biostatistics, 1990.

This page has been left blank for double-sided copying.

APPENDIX C

MATCHING PROCEDURES FOR COMMUNITY ELIGIBILITY PROVISION ANALYSIS

This page has been left blank for double-sided copying.

INTRODUCTION

Chapter V presented findings on improper payments due to certification error in schools using the Community Eligibility Provision. A key component of the CEP analysis is assessing the accuracy of the identified student percentage (ISP)—the proportion of enrolled students directly certified or certified to receive free meals without an application that determines the free and paid rate percentages schools use for claiming meal reimbursements. To calculate an estimated actual ISP (which is then compared with the school's calculated ISP to generate estimates of improper payments), the analysis team attempted to match sampled students to SNAP/TANF lists and other program data sources. This process included several data preparation steps, deterministic or "exact" matching, and probabilistic or "fuzzy" matching. This appendix provides more detail on these steps.

A. Data overview

To support the CEP matching analysis, the analysis team collected student-level data from CEP schools in the five States included in the analysis. These files include all sampled students and, when possible, other children within the sampled student's household. Eligibility for free meals is extended to all children within a household if at least one child participates in SNAP, TANF, or FDPIR. The team also collected program participation data from State agencies and districts. We describe both types of data in Chapter V, Section C. This section provides additional details on data items included in these files and the data preparation steps we implemented.

1. Variables for matching records across data files

Records were matched when specific data elements matched exactly or when a group of variables in the records likely represented the same person but did not match exactly. In either case, multiple variables were required to accurately match records. Matching validity is known to be higher when individuals shared unique identifiers, such as Social Security Number (SSN) or SNAP or TANF case IDs. However, these types of identifiers were rarely available in the student data from the SFA and were subject to error (for example, missing data, typos, and fraud). Therefore, additional personally identifiable information (PII) was used as confirmatory variables in the matching process. When more matching variables were available across the files and included in the matching process, it was easier to make accurate matches. In addition, the quality of the data greatly influenced the number and accuracy of matches. Matching variables that were frequently missing or contained typos, misspellings, or partial information yielded fewer, lower quality matches. Different combinations of matching variables were used depending on the variables available on the two files. Rules for combinations of matching variables were established to determine what was considered to be sufficient for accepting matched pairs. The following is a list of variables we used for matching records across files:

- Administrative identifiers: SSN, SNAP or TANF case ID; parent SSN
- Names: First, middle, last; parent first and last names
- Dates: Date of birth

- Mailing or residence addresses (current, previous): Street address, city, State, zip code
- Phone numbers: Home, (parent) work, mobile
- Demographics: Gender, race, ethnicity

2. Data preparation and quality control

After we collected data and received program data extracts from State agencies, we performed extensive quality control (QC) checks on all files to verify the quality of the data. Ongoing QC checks of both the data and the programs used to process the data ensured the integrity of the matched records. We produced the following documentation, QC checks, and reports when processing the original data files used to match students to program data:

- **Sample sizes**. We checked the record count and number of data items to ensure that the files matched the data request. We also verified unique program IDs, record counts, and data documentation.
- **Data item distributions**. For each file, we tabulated frequencies of categorical variables and distributions (means, minimum, maximum, and other summary statistics) of numerical variables. For categorical variables, we confirmed the documentation of all necessary code definitions. For numeric variables, we examined extreme values and converted missing values (for example, 0-fills and 9-fills) to a standard missing value.
- **File completeness**. We confirmed that the agency provided complete data and did not inadvertently exclude a portion of the sample or requested time period.

B. Key steps for data matching

The process of matching records included several steps and used three software packages:

1. **Preparing data files for linking.** We prepared data for linking tasks by renaming variables to match across files and to standardize variables (described below in Section B.1). We performed this task using SAS code.
2. **Deterministic and probabilistic matching with State list of children receiving SNAP or TANF.** We used SAS to deterministically link records within or across files (described below in Section B.2.a). All deterministically matched records were given an ID number to reflect this link. We exported files to a text (.txt) format for later use in LinkageWiz, a probabilistic matching software. After deterministic matching, we attempted to link records through probabilistic matching (described below in Section B.2.b). The probabilistic matching method used weighted scores to identify matches based on the likelihood that the variables used in the matching identified the same unique individual across files. We loaded the text files that were created at the end of the deterministic matching step into LinkageWiz.
3. **Reviewing matches.** We reviewed matches systematically and manually. Rules were created in SAS to automatically accept or reject matches. Matches that SAS could not automatically accept or reject were exported to an Excel file for manual review (described below in Section B.3.a). Matches that did not meet any automatic acceptance or rejection rules were manually reviewed by a member of the analysis team (described in Section

B.3.b). Matches considered acceptable were flagged and kept together. Conversely, matches considered unacceptable were flagged and broken apart.

4. **Additional matching with other sources of data.** We calculated an estimated actual ISP for each CEP group based on the outcome of the deterministic and probabilistic matching of student records to State SNAP and TANF files and processing sibling matches (Steps 1–3). If the estimated actual ISP was 62.5 or above after Step 3, no additional follow-up on unmatched students was needed because the free claiming percentage would be 100 percent and would not be affected by further matching. If the estimated actual ISP was below 62.5, unmatched sampled students were compared with additional data from State-, county-, and SFA- level programs to determine whether there were any additional matches. (See Section B.4 for more details.)

All of these steps aid in maximizing the likelihood of correctly identifying all true matches and minimizing the likelihood of identifying false matches. Steps 1–3 were performed only using SNAP/TANF program files. The additional sources of data were used only in Step 4. Next, we describe each of the matching steps.

1. Data file preparation

Before we linked the records, the original data files were carefully cleaned and prepared for linking. Once high-level diagnostics were complete, variables used for linking were cleaned and standardized across files. That is, we named and stored the same variable in the same way in all data files. In addition, programmers used the SAS functions LENGTH, UPCASE, TRIM, LEFT, and COMPRESS to standardize character variables. Additional cleaning was also done for street addresses to standardize street designations and abbreviations. For example, “St.” “st” and “STR” would be changed to “Street.” Lastly, any recoding schemes for categorical variables were developed and applied. For example, a consistent coding scheme was used for race variables across sampled student data and program data.

Several additional variables were also created at this stage. To allow for misspellings or typos in names, we created name variables that used the SOUNDEX standardization scheme, which attempts to distill the sound of a name. Additionally, because some of the administrative data files contained truncated names (for example, “Martinez-Joh” instead of “Martinez-Johnson”), we created truncated versions of names for matching.

2. Deterministic and probabilistic data matching

Once data preparation was complete, data linkage began. Two common approaches to data linking are deterministic matching (which requires observations to match exactly on several variables) and probabilistic matching (which is based on the likelihood of a match). These two methods are not mutually exclusive, and we opted to use a hybrid approach—that is, deterministic and probabilistic matching used in conjunction—to enhance matching results. Before discussing each approach, we consider a simple example of using deterministic and probabilistic matching sequentially to match three records (Table C.1). Under deterministic matching, the first two records are grouped together under a new ID (D111); the third record is considered a different person (D222) because the day of birth is transposed. Under probabilistic matching, which allows for minor differences and considers additional confirmatory variables,

the third record is recognized as the same person and grouped with the other two records under a new, single ID (P111).

Table C.1. Examples of combinations of variables required for a deterministic match

Probabilistic matching variables							
Deterministic matching variables							
First name	Last name	DOB	County	Deterministic match ID	Gender	Hispanic	Probabilistic match ID
John	Doe	01/01/1991	Adams	D111	--	--	P111
John	Doe	01/01/1991	Adams	D111	M	Y	P111
John	Doe	01/10/1991	Adams	D222	M	Y	P111

a. Deterministic matching

Deterministic matching required observations to match exactly on several variables (see Table C.2 for examples). It is a relatively simple process in SAS. The matching was performed by merging the sampled student file to the program data file using all of the deterministic matching variables. In combination, these variables uniquely identified individuals. Most of the files did not include a unique identifier such as SSN. Therefore, confirmatory variables were used extensively to ensure the accuracy of the match. In addition, criteria for deterministic matching are very strict, often including address information, because matches created at this stage are automatically accepted as true matches.

Table C.2. Examples of deterministic matching variables

	Matching variable
Combination 1	Date of birth, first name, last name, middle initial, gender, zip code
Combination 2	Date of birth, first name, last name, middle initial, parent last name, zip code
Combination 3	Date of birth, first name, last name, gender, street address, zip code

b. Probabilistic matching

The probabilistic matching method (also known as stochastic or fuzzy matching) used weighted scores to identify matches based on the likelihood that the variables used in matching identified a unique individual. Probabilistic matching was used to increase the match rate and because of the potential for data errors in PII (misspellings, truncations, typos). We selected LinkageWiz software to perform probabilistic matching.

Across two files, every record in the first file was paired with every record in the second file. Each pair of records was assigned a probabilistic match score based on whether the matching variables contained the same value and on the weights assigned to the matching variables. Each matching variable had a positive agreement weight (score) and a negative disagreement weight (penalty). LinkageWiz provides default weights, which we used for our matching process. Larger weights were assigned to variables that are more specific and therefore more likely to uniquely

identify individuals (date of birth, family name), whereas smaller weights were assigned to variables that are less likely to uniquely identify individuals (gender, race, county of residence). Unique identifiers, such as SSN, were assigned large scores and large penalties. A sum of the scores and penalties assigned to the matching variables determined the total probabilistic matching score. Matches with higher total scores were considered more likely to be true matches than matches with lower total scores.

LinkageWiz automatically rejected matches that did not meet a predefined scoring threshold. Again, LinkageWiz provides a default value for the scoring threshold, but we opted to set the score relatively low, to evaluate as many potential matches as possible, including relatively poor matches. For example, matched pairs with the lowest scores might have only a matching first name, city, State, and race. In most cases, these poorly matched pairs were rejected by the analysis team after a closer review.

Depending on the size of the files, probabilistic matching can be a lengthy process. To reduce processing time, we "blocked" data on several variables. With this approach, only records with the same value for the blocked variable were compared with every other record that had the same value for the blocked variable. For example, if we blocked records on last name, records with last name SMITH were compared only with other SMITHs; records with different last names were not compared within that block. Similarly, blocking on county of residence would cause only records in the same county to be compared. To avoid false matches or missed matches, we blocked data multiple times using different variables. Table C.3 shows the blocking variable we used in all States. All matches were thus required to match exactly with at least one of these variables.

Table C.3. Blocking variables for probabilistic matching

Student's last name
Student's first name
Student's last name as NYSIIS code[b]
Date of birth
Street address
Zip code

[b] New York State Identification and Intelligence System (NYSIIS) algorithms are produced in LinkageWiz.

Probabilistic matching identified good matches that were not captured at the deterministic matching step. However, unlike deterministic matching, probabilistic matching yielded some false matches as well. In particular, probabilistic matching was less accurate when individuals in the data file shared many of the same values for matching variables. For example, twins and siblings were often mismatched because last names, address information, and demographic information are identical. These cases were later rejected in the systematic or manual review process.

3. Review of matches

After each probabilistic matching step, a combination of review methods was used to determine whether the matched records were accurately grouped together. In essence, matched records were placed into one of three categories: (1) strong matches to be automatically kept, (2)

weak matches to be automatically dropped, and (3) mediocre matches to be manually reviewed. Again, to minimize the potential for false matches, the analysis team used relatively conservative criteria for automatically keeping or rejecting matched records. This approach meant a large percentage of matches were manually reviewed, but given the small samples, the task remained manageable. For these matched pairs, both a systematic and a manual review were performed before matches were considered final.

a. Systematic review

We used a systematic approach to reduce the burden of reviewing thousands of matches. This type of review worked well for strong and weak matches because they are easily defined. Specifically, we created an SAS program to evaluate probabilistic matches using a set of rules that automatically accepted or rejected records within a group. Table C.4 shows examples of rules that we used for different States based on the available variables and patterns in the data. We manually reviewed any matches that these rules failed to capture.

Table C.4. Examples of rules for accepting or rejecting matches

Decision	Rule
Accept	Probabilistic matching score above a specified threshold.
Accept	First name, last name, date of birth match. Address or house number match and race or zip code match.
Reject	Probabilistic matching score below a specified threshold and first name, last name, and date of birth all do not match.

b. Manual review

We output to an Excel file any matches that could not be systematically accepted or rejected, based on the rules established in the SAS program, for manual review. The Excel file included all of the matching variables from both records, side by side, in the matched pair. An analysis team member then reviewed this information and evaluated whether the match was acceptable. An indicator flag (1/0) was populated within the Excel file to show the matches to accept or reject.

The manual review files included several dozen to several hundred pairs. Therefore, the process used conditional formatting and workbook sharing to facilitate the review and make it more efficient. For example, color coding was used to show when matching variables, such as last name and date of birth, matched exactly across the two files. Applying this formatting to the whole file reduced the time and strain of comparing every cell for a given pair.

Following the manual review, the Excel files were read back into SAS to compile final matching rates. We then used these match rates to calculate the estimated actual ISP. CEP groups above the 62.5 threshold were considered complete because at that point, all meals can be reimbursed at the free rate. CEP groups below the 62.5 threshold required follow-up with additional data sources and proceeded to Step 4.

4. Additional matching

If the estimated actual ISP was below 62.5, we used other data from State-, county-, and SFA-level programs to determine whether there were any additional matches for that CEP group. The sources, described in Chapter V, Section C, used in subsequent rounds of matching varied somewhat by State and SFA according to the availability of data sources. The most common sources were lists of Head Start or homeless youth. We also requested additional information on "partial matches" to strengthen weak but likely matches. For example, some SFAs were able to provide us with additional contact information and SNAP or TANF case IDs that the initial data file did not include. After identifying any new matches with each additional data source, we recalculated a new estimated actual ISP to determine whether further matching was necessary.

5. Summary of matching results

The vast majority of matches came from the SNAP and TANF files. Probabilistic matching was a key step to matching many students. Not many sampled students were matched based on extending benefits from another child within the household. Table C.5 shows the percentage of matches that came from the various data sources. It also shows the percentage of matches based on the different matching steps.

Table C.5. Percentage of sampled student matches by source and method

	State A		State B		State C		State D		State E	
	Identified students	Other students	Identified students	Other students	Identified students	Other students	Identified students	Other students	Identified students	Other students
Not matched	3.5	91.0	10.2	72.9	3.5	72.7	2.0	83.7	5.9	59.9
SNAP/TANF										
Deterministic matches	41.9	2.7	23.7	1.4	14.8	1.6	4.0	0.0	35.2	12.8
Probabilistic matches	51.4	5.7	64.4	24.3	80.4	25.5	89.6	14.3	57.5	23.7
Benefit extension[a]	0.0	0.0	0.0	0.0	0.4	0.3	0.8	1.1	0.2	1.9
Follow-up matches and other sources	3.2	0.6	1.7	1.4	0.9	0.0	3.6	0.9	0.9	1.6

Source: APEC-II study, weighted data.

[a] Indicates the number of sampled students who were identified for free meals because another student within the household was matched to SNAP/TANF.

APEC = Access Participation, Eligibility, and Certification; SNAP = Supplemental Nutrition Assistance Program; TANF = Temporary Assistance to Needy Families.

The relatively small number of matches found during follow-up with SNAP and TANF on "partial matches" and through matching against additional data sources had limited overall effect on the national CEP estimates. We ran the main analysis excluding these additional matches (Table C.6) and the results were little changed.

Table C.6. National estimates of improper payments in the NSLP and SBP for schools using CEP excluding follow-up matches and other sources, SY 2012–2013

	NSLP	SBP
Total reimbursements (millions of dollars)		
Total reimbursements	286.1	117.9
Improper payment amounts (millions of dollars)		
Overpayments	1.0	0.5
	(0.3)	(0.2)
Underpayments	4.1	2.0
	(2.7)	(1.2)
Gross improper payments	5.1	2.5
	(2.7)	(1.2)
Net improper payments	-3.6	-1.5
	(2.7)	(1.2)
Improper payment rates (percentages)		
Overpayments	0.3	0.4
	(0.1)	(0.2)
Underpayments	1.6	1.7
	(1.0)	(1.0)
Gross improper payments	1.9	2.1
	(1.0)	(1.0)
Net improper payments	-1.2	-1.3
	(1.0)	(1.0)
Sample size (schools)	**135**	**135**

Source: APEC-II study, weighted data.

Note: Analysis weights are calibrated based on total national reimbursements reported in the FNS national data file. Standard errors in parentheses.

APEC = Access Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; FNS = Food and Nutrition Service; NLSP = National School Lunch Program; SBP = School Breakfast Program; SY = school year.

APPENDIX D

STUDENT, SCHOOL, AND SFA CHARACTERISTICS AND OTHER SUPPLEMENTAL TABLES

This page has been left blank for double-sided copying.

The APEC-II study collected information on the administrative and operational structure of SFAs and schools sampled for the study that when weighted can be tabulated to provide descriptive summaries that are representative of SFAs and schools participating in the school meal programs nationally during SY 2012–2013. Tables D.1–D.11 provide summary statistics on the characteristics of students (certified students and denied applicants), schools, and SFAs. These data are weighted to be nationally representative. Characteristics of SFAs and schools are presented two ways: (1) weighted by the SFA or school and (2) the SFA or school weight adjusted for the number of enrolled students with access to the school meal programs. The latter show findings in terms of the percentages of students in the SFA (or attending schools) with characteristics indicated in the tables.

Because the primary objective of the APEC-II study was to generate precise national estimates of the dollar amounts and rates of improper payments in the NSLP and SBP due to certification error, and not to estimate characteristics of SFAs and schools precisely, some caution should be exercised when using the data to examine SFA and school characteristics. In particular, the samples of SFAs and schools are smaller than what would be considered ideal for that purpose, meaning the estimates of characteristics are subject to greater sampling variability.

This appendix also includes several supplemental tables on non-certification error. Specifically, Table D.12 presents claiming error findings by CEP status, and Tables D.13–D.15 present aggregation error findings by CEP status.

This page has been left blank for double-sided copying.

Table D.1. Characteristics of students, by certification status (percentages of students)

	Certified students and denied applicants		
	Free and reduced-price certified	Denied applicants	All students
Child's grade			
Pre-K to K	10.3	6.3	9.8
1 to 3	24.1	26.3	24.4
4 to 5	15.1	21.6	16.0
6 to 8	25.4	27.4	25.6
9 to 12	24.0	17.9	23.1
Ungraded	1.1	0.5	1.0
Gender			
Male	50.2	53.8	50.7
Female	49.8	46.2	49.3
Race/ethnicity			
White, non-Hispanic	30.2	54.8	33.6
Black, non-Hispanic	22.4	12.6	21.1
Hispanic	38.3	21.7	36.0
Other	9.0	10.9	9.3
Urbanicity			
Urban	36.3	42.5	37.0
Suburban	39.8	34.5	39.2
Town	9.6	12.2	9.9
Rural	14.3	10.8	13.9
Household headed by			
Two parents	50.5	71.5	53.4
Single parent	42.1	26.4	40.0
Other relative	7.2	2.1	6.5
Nonrelative	0.2	0.0	0.2
Parent's education			
Less than high school	23.8	7.0	21.5
High school or GED	46.0	25.9	43.3
Some college	20.2	29.8	21.6
College graduate	10.0	37.3	13.7
Program participation[a]			
TANF	5.4	0.2	4.7
Food stamps	44.8	2.4	39.0
FDPIR	0.5	0.3	0.5
Medicaid	64.5	14.4	57.7
Number of children < 18 years			
1	18.5	42.4	21.8
2	32.7	28.5	32.1
3	26.4	17.1	25.1
4 or more	22.4	12.0	21.0
Age of youngest child			
Younger than 5	30.9	24.0	30.0
5 to 8	31.0	29.3	30.8
9 to 13	28.3	38.1	29.7
14 to 18	9.8	8.6	9.6
Household size			
1 to 3	24.3	32.0	25.4
4 to 6	64.3	59.1	63.6
7 to 9	10.6	8.9	10.4
10 or more	0.8	0.0	0.7

Table D.1 (*continued*)

	Certified students and denied applicants		
	Free and reduced-price certified	Denied applicants	All students
Income relative to poverty (percentage)			
Less than 50	22.5	7.8	20.5
50 to 99	33.1	9.8	29.9
100 to 129	15.3	8.2	14.3
130 to 184	17.3	12.1	16.6
185 to 249	6.3	30.2	9.6
250 to 399	4.7	24.0	7.3
400 or more	0.9	7.8	1.8
Sample size	**3,608**	**582**	**4,190**

Source: APEC-II study, household survey data.

Notes: Data are weighted by student weight. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding.

[a] Does not sum to 100 percent because students' household may participate in more than one program.

FDPIR = Food Distribution Program on Indian Reservations; TANF = Temporary Assistance to Needy Families.

Table D.2. Characteristics of schools, by CEP status (percentages of schools)

Characteristic	CEP status		All schools
	CEP	Non-CEP	
Urbanicity			
District covers urban area	70.3	36.3	37.5
District covers suburban area	8.6	34.2	33.3
District serves a town	3.4	6.7	6.6
District covers rural area	17.7	22.8	22.6
Region			
Northeast	6.1	11.6	11.4
Mid-Atlantic	23.3	13.8	14.1
Southeast	0.0	18.0	17.4
Midwest	70.6	17.1	18.9
Southwest	0.0	13.0	12.6
Mountain Plains	0.0	11.3	10.9
Western	0.0	15.1	14.6
School level			
Elementary	76.6	69.2	69.4
Middle	6.6	14.9	14.6
High	16.8	15.9	16.0
School enrollment			
Fewer than 400 students	69.1	32.4	33.5
400 to 799 students	26.1	52.0	51.2
800 to 1,200 students	4.4	8.9	8.7
More than 1,200 students	0.5	6.7	6.6
Median	326	477	472
Mean	371.7	611.1	604.2
Sample size	**135**	**392**	**527**

Source: APEC-II study, data on study schools from SFA Director Survey.

Notes: Data are weighted by school weight. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding.

Table reads: "37.5 percent of all schools offering the NSLP and/or SBP are located in urban areas."

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; NSLP = National School Lunch Program; SBP = School Breakfast Program; SFA = School Food Authority.

Table D.3. Characteristics of schools, by CEP status (percentages of students in schools with characteristics indicated in row headings)

	CEP status		
Characteristic	**CEP**	**Non-CEP**	**All schools**
Urbanicity			
District covers urban area	75.5	36.7	37.6
District covers suburban area	7.7	38.2	37.5
District serves a town	2.8	6.3	6.2
District covers rural area	13.9	18.9	18.7
Region			
Northeast	6.7	9.8	9.7
Mid-Atlantic	17.6	13.1	13.2
Southeast	0.0	17.8	17.4
Midwest	75.8	14.4	15.8
Southwest	0.0	12.6	12.3
Mountain Plains	0.0	13.3	13.0
Western	0.0	19.0	18.5
School level			
Elementary	73.0	56.6	57.0
Middle	6.6	17.0	16.7
High	20.4	26.4	26.3
School enrollment			
Fewer than 400 students	44.4	17.7	18.2
400 to 799 students	40.7	48.1	47.9
800 to 1,200 students	11.0	13.5	13.5
More than 1,200 students	3.8	20.7	20.4
Median	473	686	682
Mean	536.5	890.2	883.5
Sample size	**135**	**392**	**527**

Source: APEC-II study, data on study schools from SFA Director Survey.

Notes: Data are weighted by school weight adjusted for number of enrolled students with access to the school meal programs. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding. Table reads: "34.7 percent of all students with access to the school meal programs are in schools located in urban areas."

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; SFA = School Food Authority.

Table D.4. NSLP and SBP meal program characteristics, by CEP status (percentages of schools)

Characteristic	CEP status		All schools
	CEP	Non-CEP	
Types of meal programs offered			
NSLP only	5.9	8.4	8.3
SBP only	0.0	1.4	1.4
Both NSLP and SBP	94.1	90.2	90.3
Provision 2 or 3 status			
NSLP only, NP 2/3	0.0	12.7	12.3
SBP only, NP 2/3	0.0	1.5	1.4
SBP NP 2/3, NSLP NP 2/3	97.0	79.0	79.6
NSLP P 2/3, SBP NP 2/3	0.0	0.0	0.0
SBP P 2/3, NSLP NP 2/3	0.0	4.2	4.1
SBP P 2/3, NSLP P 2/3	3.0	2.6	2.6
Uses offer versus serve (OVS)[a]			
Uses OVS	92.8	97.8	97.7
Does not use OVS	7.2	2.2	2.3
Student certification (mean percentages)			
Certified for free meals	89.5	47.8	48.3
Certified for reduced-price meals	n.a.	8.0	7.9
Certified for free or reduced-price meals	89.5	55.8	56.2
Percentage of NSLP lunches by type			
Free	85.7	59.4	60.0
Reduced-price	n.a.	9.0	8.8
Paid	14.1	31.6	31.3
Receives NSLP 60 percent subsidy[b]			
Yes	81.4	74.8	75.0
No	18.6	25.2	25.0
NSLP participation (percentages)			
Average daily participation rate			
Among all students	70.1	62.8	63.0
Among students certified for free meals	72.6	75.7	75.7
Among students certified for reduced-price meals	n.a.	69.4	69.3
Among students not certified (paid)	32.7	40.8	40.8
Percentage of breakfasts by type			
Free	83.8	72.5	72.7
Reduced-price	n.a.	9.2	9.0
Paid	15.9	18.3	18.2
Receives severe needs SBP subsidy[c]			
Yes	95.0	90.8	91.0
No	5.0	9.2	9.0
SBP participation (percentages)			
Average daily participation rate			
Among all students	55.6	34.6	35.2
Among students certified for free meals	66.2	43.2	43.4
Among students certified for reduced-price meals	n.a.	36.4	36.3
Among students not certified (paid)	20.7	19.7	19.7
Certified for additional 6 cents[d]			
Yes	89.4	89.4	89.4
No	10.6	10.6	10.6
Sample size	**135**	**392**	**527**

Table D.4 (*continued*)

Source: APEC-II study, data on study schools from SFA Director Survey.

Notes: Data are weighted by school weight. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding.

Table reads: "8.3 percent of all schools participating in school meal programs offer the NSLP only."

[a] School uses offer-versus-serve option in one or both school meal programs.

[b] School receives extra 2 cents for each lunch served because 60 percent of more of total lunches are to students certified for free or reduced-price meals.

[c] School receives extra 24 cents for each free or reduced-price breakfast served.

[d] School receives additional 6-cent subsidy for each NSLP lunch served because SFA certified it compliant with new meal pattern and nutrition requirements by State agency.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; NSLP = National School Lunch Program; NP 2/3 = non-Provision 2 or 3; P 2/3 = Provision 2 or 3 schools; SBP = School Breakfast Program; SFA = School Food Authority

Table D.5. NSLP and SBP meal program characteristics, by CEP status (percentages of students in schools with characteristics indicated in row heading)

Characteristic	CEP status		All schools
	CEP	Non-CEP	
Types of meal programs offered			
NSLP only	7.0	10.3	10.2
SBP only	0.0	1.7	1.6
Both NSLP and SBP	93.0	88.1	88.2
Provision 2 or 3 Status			
NSLP only, NP 2/3	0.0	13.9	13.6
SBP only, NP 2/3	0.0	1.7	1.7
SBP NP 2/3, NSLP NP 2/3	96.9	78.0	78.3
NSLP P 2/3, SBP NP 2/3	0.0	0.1	0.1
SBP P 2/3, NSLP NP 2/3	0.0	3.6	3.5
SBP P 2/3, NSLP P 2/3	3.1	2.8	2.8
Uses offer versus serve (OVS)[a]			
Uses OVS	93.5	98.4	98.4
Does not use OVS	6.5	1.6	1.6
Student certification (mean percentages)			
Certified for free meals	91.9	44.5	45.0
Certified for reduced-price meals	n.a.	7.7	7.6
Certified for free or reduced-price meals	91.9	52.2	52.6
Percentage of NSLP lunches by type			
Free	85.6	57.7	58.0
Reduced-price	n.a.	8.9	8.8
Paid	14.1	33.4	33.2
Receives NSLP 60 percent subsidy[b]			
Yes	84.2	72.8	73.0
No	15.8	27.2	27.0
NSLP participation (percentages)			
Average daily participation rate			
Among all students	72.8	58.0	58.3
Among students certified for free meals	73.6	72.2	72.3
Among students certified for reduced-price meals	n.a.	65.2	65.1
Among students not certified (paid)	28.7	37.1	37.1
Percentage of breakfasts by type			
Free	84.4	73.5	73.6
Reduced-price	n.a.	9.2	9.1
Paid	15.2	17.3	17.3
Receives severe needs SBP subsidy[c]			
Yes	95.3	86.5	86.7
No	4.7	13.5	13.3
SBP participation (percentages)			
Average daily participation rate			
Among all students	53.6	28.7	29.1
Among students certified for free meals	62.1	37.7	37.8
Among students certified for reduced-price meals	n.a.	30.3	30.3
Among students not certified (paid)	17.4	15.5	15.5
Certified for additional 6 cents[d]			
Yes	89.3	89.5	89.5
No	10.7	10.5	10.5
Sample size	**135**	**392**	**527**

Source: APEC-II study, data on study schools from SFA Director Survey.

Table D.5 (*continued*)

Notes: Data are weighted by school weight adjusted for number of enrolled students with access to the school meal programs. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding. Table reads: "10.2 percent of students with access to the school meal programs attend schools that offer the NSLP only."

[a]School uses offer versus serve option in one or both school meal programs.

[b]School receives extra 2 cents for each lunch served because 60 percent of more of total lunches are to students certified for free or reduced-price meals.

[c]School receives extra 24 cents for each free or reduced-price breakfast served.

[d]School receives additional 6-cent subsidy for each NSLP lunch served because SFA certified it compliant with new meal pattern and nutrition requirements by State agency.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; NSLP = National School Lunch Program; NP 2/3 = non-Provision 2 or 3; P 2/3 = Provision 2 or 3 schools; SBP = School Breakfast Program; SFA = School Food Authority.

Table D.6. Characteristics of school food authorities, by CEP status (percentages of SFAs)

Characteristic	CEP status		All SFAs
	CEP	Non-CEP	
Public vs. private SFA			
Administers public schools only	92.1	87.2	87.2
Administers private schools only	4.2	12.3	12.2
Administers both public and private schools	3.7	0.6	0.6
Single- vs. multiple-district SFA			
Administers single district	74.7	98.4	98.1
Administers multiple districts or entities	25.3	1.6	1.9
Urbanicity			
District covers urban area	55.5	26.6	27.0
District covers suburban area	6.4	25.6	25.3
District serves a town	7.4	18.8	18.6
District covers rural area	30.7	29.1	29.1
District size			
Total number of schools (mean)	34.6	53.2	52.9
Total number of schools (median)	8.0	7.0	7.0
Percentage of schools by type of school			
Elementary	63.4	57.0	57.1
Middle	13.7	21.2	21.2
High	19.9	18.9	18.9
Other	3.0	2.9	2.9
Student enrollment			
Fewer than 1,000	30.2	21.9	22.0
1,000 to 4,999	54.9	33.6	33.8
5,000 to 9,999	7.6	17.0	16.9
10,000 to 19,999	4.1	19.4	19.2
20,000 to 49,999	1.8	5.6	5.5
50,000 or more	1.3	2.6	2.6
Median	2,795	3,344	3,313
Mean	4,255.4	8,789.3	8,730.0
Sample size	**36**	**127**	**163**

Source: APEC-II study, SFA Director Survey data.

Notes Data are weighted by SFA weight. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding. Table reads: "87.2 percent of SFAs administer the NSLP and/or SBP in public schools only."

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; SFA = School Food Authority.

Table D.7. Characteristics of school food authorities, by CEP status (percentages of students in SFAs with characteristics indicated in row headings)

Characteristic	CEP status		
	CEP	Non-CEP	All SFAs
Public vs. private SFA			
Administers public schools only	87.3	94.3	94.3
Administers private schools only	9.1	4.9	4.9
Administers both public and private schools	3.6	0.8	0.8
Single- vs. multiple-district SFA			
Administers single district	88.8	97.1	97.1
Administers multiple districts or entities	11.2	2.9	2.9
Urbanicity			
District covers urban area	55.5	29.5	29.6
District covers suburban area	9.9	33.6	33.5
District serves a town	3.8	16.9	16.8
District covers rural area	30.5	20.0	20.0
District size			
Total number of schools (mean)	64.1	36.2	36.4
Total number of schools (median)	8.0	10.0	10.0
Percentage of schools by type of school			
Elementary	65.2	60.5	60.6
Middle	15.0	18.6	18.6
High	16.9	18.5	18.5
Other	2.8	2.4	2.4
Student enrollment			
Fewer than 1,000	13.1	7.7	7.8
1,000 to 4,999	49.5	27.6	27.7
5,000 to 9,999	16.5	19.5	19.5
10,000 to 19,999	7.2	31.2	31.1
20,000 to 49,999	8.2	8.5	8.5
50,000 or more	5.4	5.5	5.5
Median	3938	7,574	7,574
Mean	9,403.1	14,454.7	14,429.5
Sample size	**36**	**127**	**163**

Source: APEC-II study, SFA Director Survey data.

Notes: Data are weighted by SFA weight adjusted for number of enrolled students with access to the school meal programs. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding. Table reads: "94.3 percent of students are in SFAs that administer the NSLP and/or SBP in public schools only."

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; SFA = School Food Authority.

Table D.8. NSLP and SBP meal program characteristics, by CEP status (percentages of SFAs)

Characteristic	CEP status		All SFAs
	CEP	Non-CEP	
Percentage of schools by type of meal program offered			
NSLP only	0.6	18.5	18.3
SBP only	0.6	0.5	0.5
Both NSLP and SBP	98.7	81.0	81.2
Percentage of enrolled students by type of meal program offered			
In schools offering NSLP only	1.6	20.6	20.4
In schools offered SBP only	0.6	2.3	2.2
In schools offering both NSLP and SBP	97.7	77.1	77.4
Student certification status (percentages)			
Certified for free meals	51.7	34.3	34.4
Certified for reduced-price meals	3.7	7.2	7.1
Certified for free or reduced-price meals	55.4	41.5	41.6
Percentage of NSLP lunches by type			
Free	86.3	49.1	49.5
Reduced-price	2.2	9.2	9.2
Paid	11.5	41.6	41.3
NSLP participation (percentages)			
Average daily participation rate			
Among all students	68.9	58.3	58.4
Among students certified for free meals	75.6	75.8	75.8
Among students certified for reduced-price meals	60.0	68.8	68.7
Among students not certified (paid)	24.1	37.6	37.5
Percentage of SBP breakfasts by type			
Free	87.3	72.4	72.6
Reduced-price	2.2	9.8	9.7
Paid	10.5	17.8	17.7
SBP participation (percentages)			
Average daily participation rate			
Among all students	50.4	26.5	26.8
Among students certified for free meals	53.4	39.2	39.3
Among students certified for reduced-price meals	25.4	28.2	28.2
Among students not certified (paid)	15.5	10.0	10.1
Sample size	**32**	**122**	**149**

Source: APEC-II study, SFA Director Survey data.

Notes: Data are weighted by SFA weight. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding. Table reads: "18.3 percent of SFAs offer NSLP only."

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; NSLP = National School Lunch Program; SBP = School Breakfast Program.

Table D.9. NSLP and SBP meal program characteristics, by CEP status (percentages of students in SFAs with characteristics indicated in row headings)

Characteristic	CEP Status		All SFAs
	CEP	Non-CEP	
Percentage of schools by type of meal program offered			
NSLP only	1.3	10.7	10.7
SBP only	1.3	0.5	0.5
Both NSLP and SBP	97.3	88.8	88.8
Percentage of enrolled students by type of meal program offered			
In schools offering NSLP only	2.2	13.2	13.2
In schools offered SBP only	1.3	2.0	2.0
In schools offering both NSLP and SBP	96.5	84.8	84.9
Student certification status (percentages)			
Certified for free meals	56.8	36.0	36.1
Certified for reduced-price meals	2.4	7.3	7.3
Certified for free or reduced-price meals	59.3	43.3	43.4
Percentage of NSLP lunches by type			
Free	85.5	52.3	52.5
Reduced-price	1.5	9.7	9.6
Paid	13.0	38.0	37.9
NSLP participation (percentages)			
Average daily participation rate			
Among all students	64.6	58.3	58.4
Among students certified for free meals	60.7	75.7	75.7
Among students certified for reduced-price meals	61.7	68.8	68.8
Among students not certified (paid)	31.8	37.8	37.8
Percentage of SBP breakfasts by type			
Free	85.3	72.9	73.0
Reduced-price	2.0	9.8	9.7
Paid	12.7	17.3	17.3
SBP participation (percentages)			
Average daily participation rate			
Among all students	47.9	24.0	24.2
Among students certified for free meals	50.6	37.5	37.5
Among students certified for reduced-price meals	26.1	28.9	28.9
Among students not certified (paid)	24.6	9.2	9.2
Sample size	**32**	**122**	**149**

Source: APEC-II study, SFA Director Survey data.

Notes: Data are weighted by SFA weight adjusted for number of enrolled students with access to the school meal programs. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding. Table reads: “6.8 percent of students are in SFAs that offer NSLP only.”

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; NSLP = National School Lunch Program; SBP = School Breakfast Program.

Table D.10. Characteristics of SFA school meal program operations, by CEP status (percentages of SFAs)

Characteristic	CEP status		All SFAs
	CEP	Non-CEP	
Uses of food service management company			
Yes	23.6	10.1	10.3
No	76.4	89.9	89.7
Uses direct certification			
Yes	97.0	100.0	100.0
No	3.0	0.0	0.0
Direct certification method[a]			
Central matching	92.0	86.9	86.9
Local matching	29.8	46.8	46.6
Letter method	31.4	18.4	18.6
Number of years using direct certification			
1 to 2 years	26.2	2.7	3.0
3 to 5 years	24.1	20.8	20.9
6 to 8 years	11.0	27.9	27.7
More than 8 years	38.7	48.6	48.5
Percentage of students certified for free meals by certification method			
Direct certification	70.7	61.2	61.2
Other certification not by application	16.3	1.4	1.5
Application—categorically certified	2.2	9.5	9.4
Application—income certified	10.8	27.9	27.9
Percentage of approved application by type of approval			
Free, categorically certified	11.1	15.8	15.8
Free, income certified	47.5	42.6	42.6
Reduced-price, income certified	41.4	41.6	41.6
District tracks and maintains data on student participation in NSLP and/or SBP at individual level			
Yes	100.0	99.8	98.5
No	0.0	0.2	1.5
Sample size	**41**	**130**	**171**

Source: APEC-II study, SFA Director Survey data.

Notes: Data are weighted by SFA weight. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding. Table reads: "10.3 percent of SFAs use a food management company."

[a] Items can sum to more than 100 because districts can use more than one method.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; NSLP = National School Lunch Program; SBP = School Breakfast Program.

Table D.11. Characteristics of SFA school meal program operations, by CEP status (percentages of students in SFAs with characteristics indicated in row headings)

	CEP status		
Characteristic	**CEP**	**Non-CEP**	**All SFAs**
Uses of food service management company			
Yes	22.6	12.0	12.0
No	77.4	88.0	88.0
Uses direct certification			
Yes	96.7	100.0	100.0
No	3.3	0.0	0.0
Direct certification method[a]			
Central matching	88.6	85.1	85.1
Local matching	45.5	50.9	50.9
Letter method	23.5	18.4	18.4
Number of years using direct certification			
1 to 2 years	13.9	2.3	2.4
3 to 5 years	9.5	31.7	31.5
6 to 8 years	6.0	18.6	18.6
More than 8 years	70.6	47.4	47.5
Percentage of students certified for free meals by certification method			
Direct certification	70.7	57.5	57.6
Other certification no by application	17.8	1.8	1.8
Application—categorically certified	2.1	8.6	8.6
Application—income certified	9.3	32.1	32.0
Percentage of approved application by type of approval			
Free, categorically certified	11.2	15.3	15.3
Free, income certified	50.2	49.3	49.3
Reduced-price, income certified	38.6	35.4	35.4
District tracks and maintains data on student participation in NSLP and/or SBP at individual level			
Yes	100.0	99.9	99.2
No	0.0	0.1	0.8
Sample size	**41**	**130**	**171**

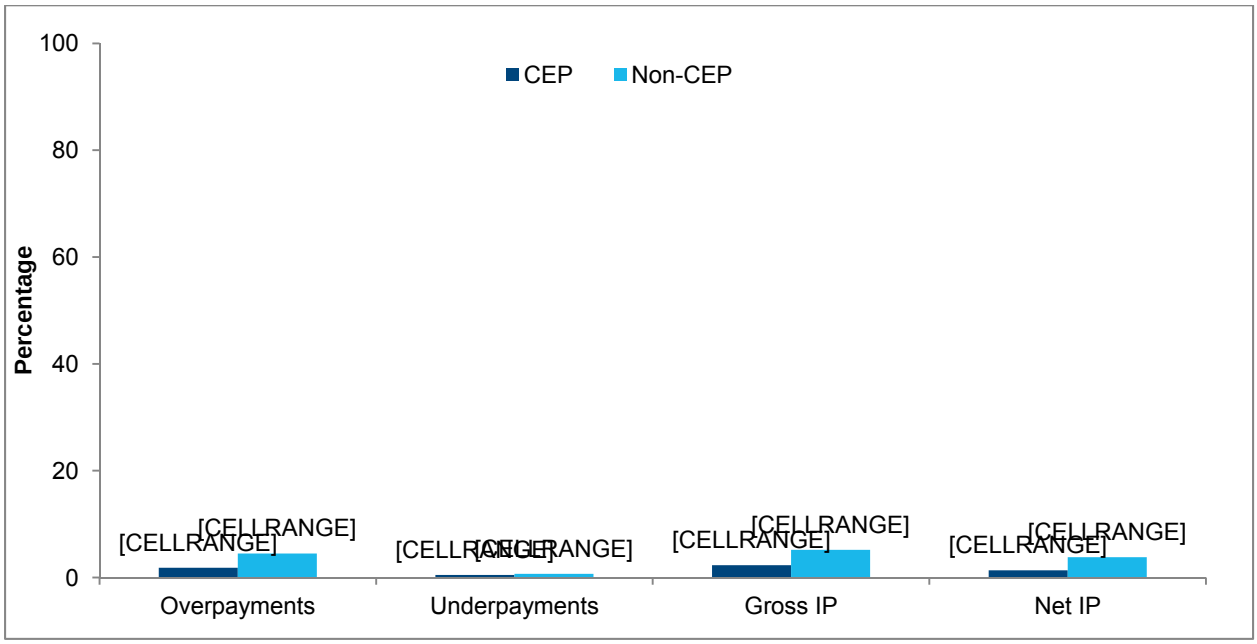
Source: APEC-II study, SFA Director Survey data.

Notes: Data are weighted by SFA weight adjusted for number of enrolled students with access to the school meal programs. Sample size varies because of item nonresponse. Items may not sum to 100 percent because of rounding. Table reads: "12.0 percent of students are in SFAs that use a food management company."

[a] Items can sum to more than 100 because districts can use more than one method.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; NSLP = National School Lunch Program; SBP = School Breakfast Program.

Figure D.1a. National estimates of improper payment rates due to meal claiming error in the NSLP by CEP status, SY 2012–201

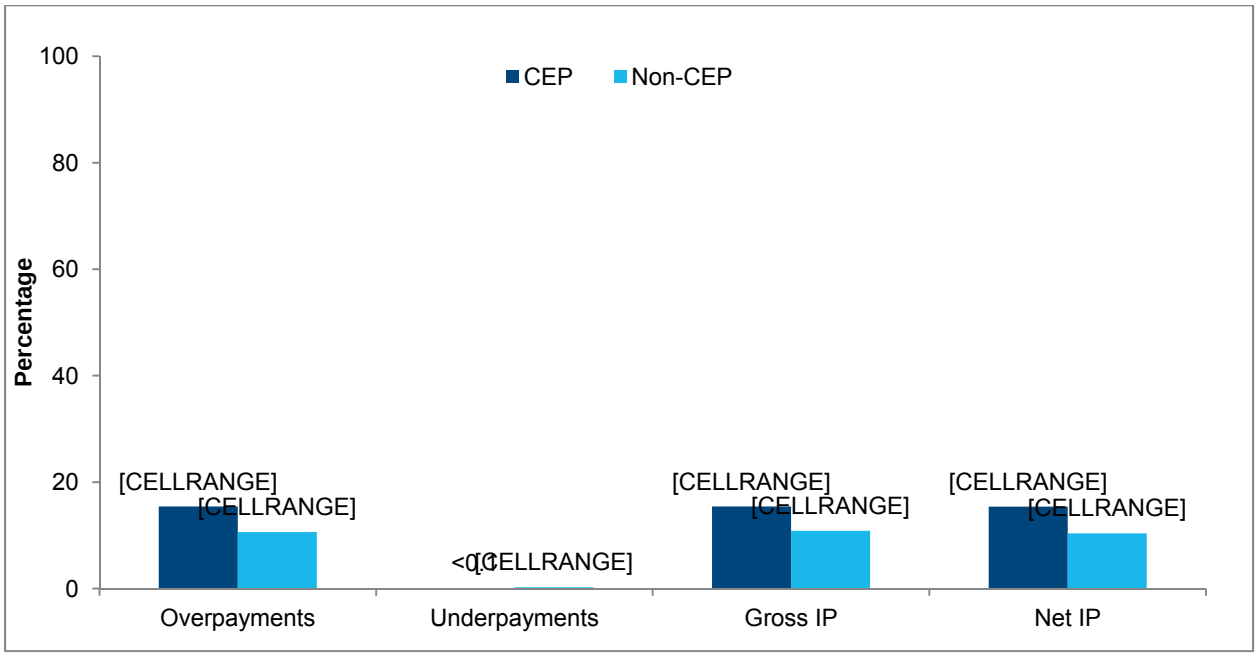


Source: APEC-II study, weighted data.

Note: Analysis weights are calibrated based on total national reimbursements reported in the Food and Nutrition Service (FNS) national data file. Standard errors in parentheses. The sum of overpayments and underpayments does not equal gross improper payments because of rounding. Sample sizes were 46 CEP schools and 392 non-CEP schools.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; IP = improper payments; NSLP = National School Lunch Program.

Figure D.1b. National estimates of improper payment rates due to meal claiming error in the SBP by CEP status, SY 2012–2013

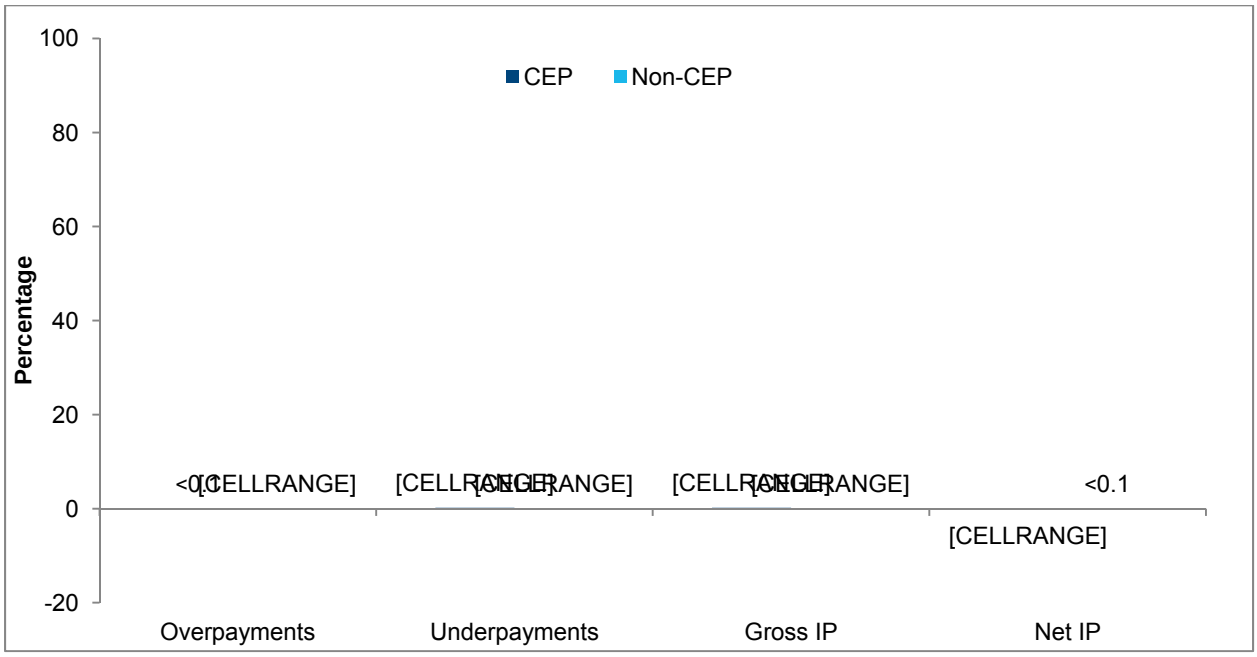


Source: APEC-II study, weighted data.

Note: Analysis weights are calibrated based on total national reimbursements reported in the Food and Nutrition Service (FNS) national data file. Standard errors in parentheses. The sum of overpayments and underpayments does not equal gross improper payments because of rounding. Sample sizes were 46 CEP schools and 392 non-CEP schools.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; IP = improper payments; SBP = School Breakfast Program.

Figure D.2a. National estimates of improper payment rates due to aggregation error in the NSLP by CEP status: point-of-sale error, SY 2012–2013

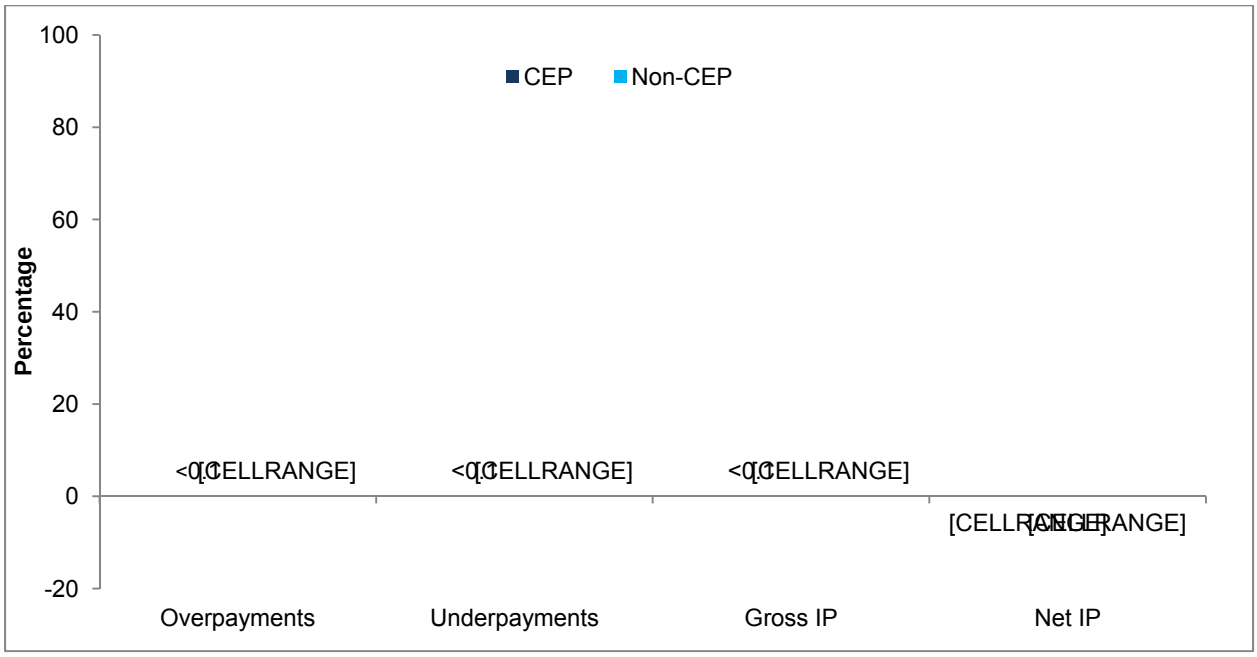


Source: APEC-II study, weighted data.

Note: Analysis weights are calibrated based on total national reimbursements reported in the Food and Nutrition Service (FNS) national data file. Standard errors in parentheses. The sum of overpayments and underpayments does not equal gross improper payments because of rounding. Point-of-sale aggregation error occurs when the sum of daily meal count totals from school cafeteria points of sale differs from the total meal counts reported by a school to the school district office that prepared the claim for reimbursement. Sample sizes were 41 CEP schools and 344 non-CEP schools.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; IP = improper payments; NSLP = National School Lunch Program.

Figure D.2b. National estimates of improper payment rates due to aggregation error in the SBP by CEP status: point-of-sale error, SY 2012–2013

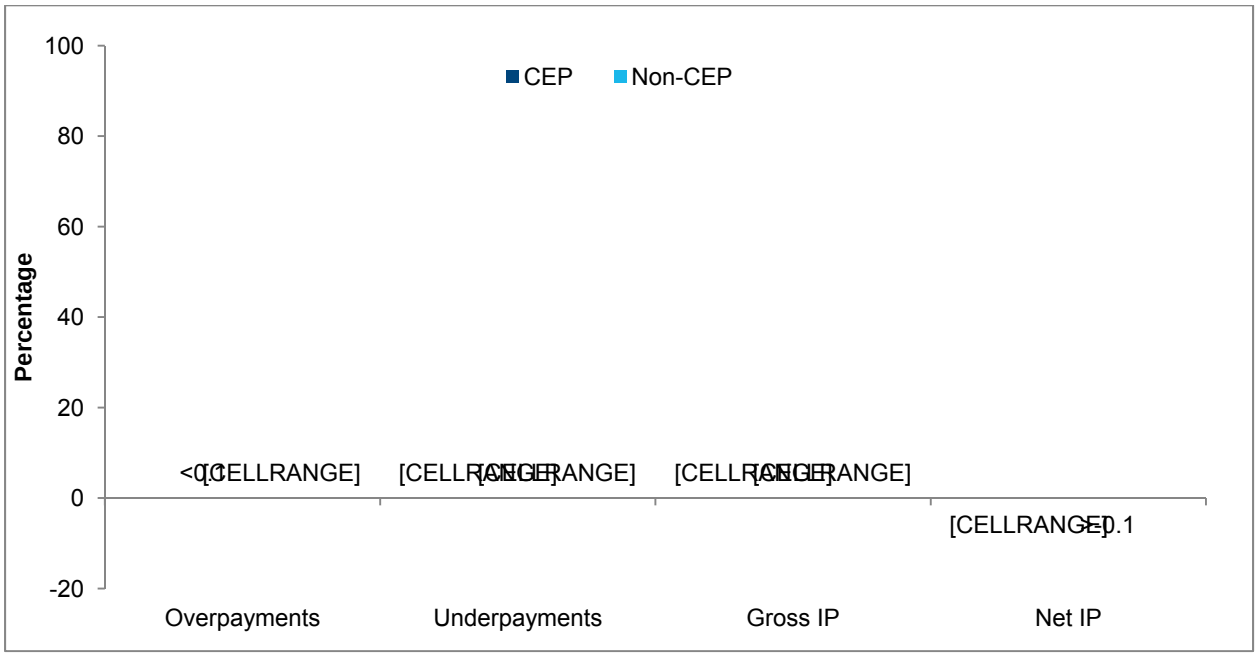


Source: APEC-II study, weighted data.

Note: Analysis weights are calibrated based on total national reimbursements reported in the Food and Nutrition Service (FNS) national data file. Standard errors in parentheses. The sum of overpayments and underpayments does not equal gross improper payments because of rounding. Point-of-sale aggregation error occurs when the sum of daily meal count totals from school cafeteria points of sale differs from the total meal counts reported by a school to the school district office that prepared the claim for reimbursement. Sample sizes were 41 CEP schools and 344 non-CEP schools.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; IP = improper payments; SBP = School Breakfast Program.

Figure D.3a. National estimates of improper payment rates due to aggregation error in the NSLP by CEP status: school reports of meal counts to the SFA, SY 2012–2013

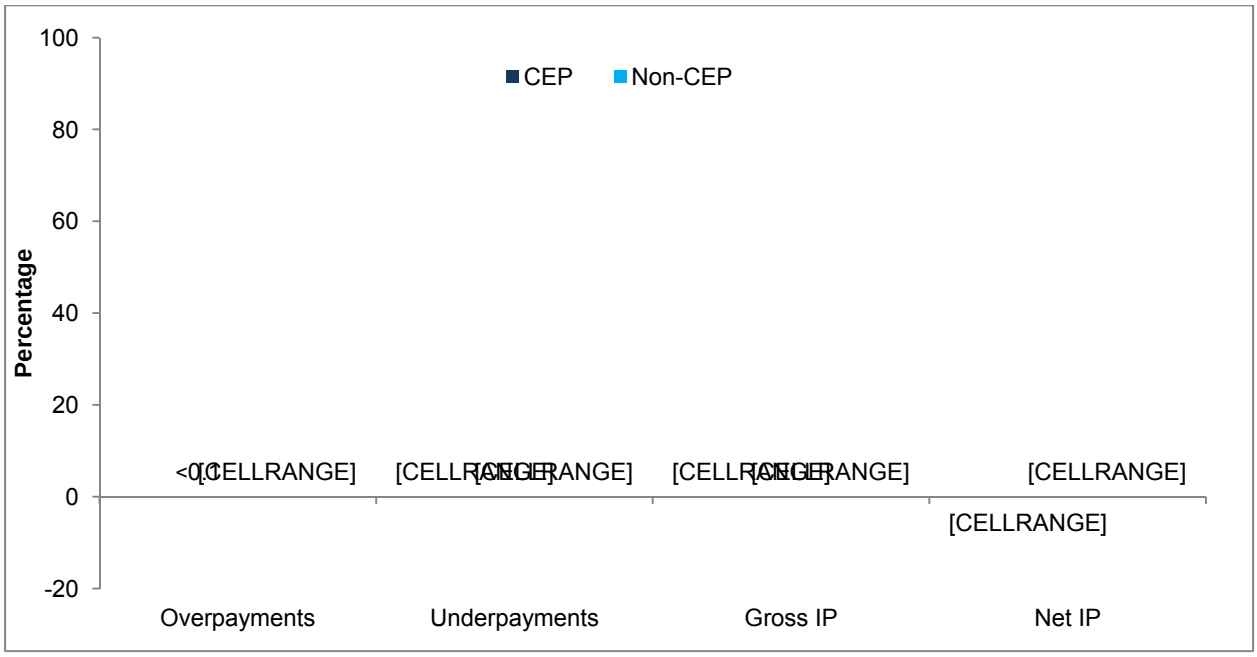


Source: APEC-II study, weighted data.

Note: Analysis weights are calibrated based on total national reimbursements reported in the Food and Nutrition Service (FNS) national data file. Standard errors in parentheses. The sum of overpayments and underpayments does not equal gross improper payments because of rounding. School-to-SFA aggregation error occurs when the sum of monthly meal count totals from schools differs from the total meal counts recorded at the SFA office that prepared the claim for reimbursement. Sample sizes were 42 CEP schools and 369 non-CEP schools.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; IP = improper payments; NSLP = National School Lunch Program; SFA = School Food Authority.

Figure D.3b. National estimates of improper payment rates due to aggregation error in the SBP by CEP status: school reports of meal counts to the SFA, SY 2012–2013

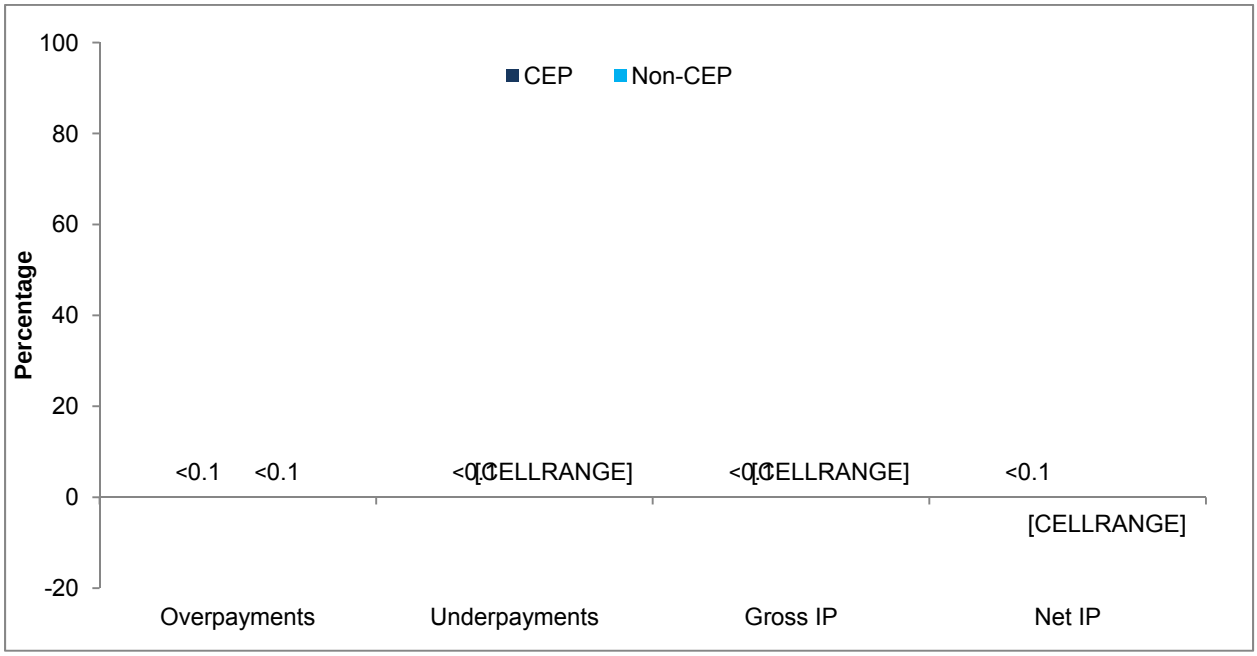


Source: APEC-II study, weighted data.

Note: Analysis weights are calibrated based on total national reimbursements reported in the Food and Nutrition Service (FNS) national data file. Standard errors in parentheses. The sum of overpayments and underpayments does not equal gross improper payments because of rounding. School-to-SFA aggregation error occurs when the sum of monthly meal count totals from schools differs from the total meal counts recorded at the SFA office that prepared the claim for reimbursement. Sample sizes were 42 CEP schools and 358 non-CEP schools.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; IP = improper payments; SBP = School Breakfast Program; SFA = School Food Authority.

Figure D.4. National estimates of improper payment rates due to aggregation error in the SBP* by CEP status: SFA reports of meal counts to the State agency, SY 2012–2013



Source: APEC-II study, weighted data.

Note: Analysis weights are calibrated based on total national reimbursements reported in the Food and Nutrition Service (FNS) national data file. Standard errors in parentheses. The sum of overpayments and underpayments does not equal gross improper payments because of rounding. SFA-to-State agency aggregation error occurs when the district's records of the number of reimbursable meals differ from the State agency's records of the number of reimbursable meals. Sample sizes were 46 CEP schools and 332 non-CEP schools.

*Figure of NSLP improper payment rates for this type of error are omitted because all improper payment rate estimates were less than 0.01 percent for both CEP and non-CEP schools.

APEC = Access, Participation, Eligibility, and Certification; CEP = Community Eligibility Provision; IP = improper payments; SBP = School Breakfast Program; SFA = School Food Authority.

This page has been left blank for double-sided copying.

APPENDIX E

IMPUTATIONS FOR MEAL PARTICIPATION AND INCOME SOURCES

This page has been left blank for double-sided copying.

INTRODUCTION

To estimate improper payments due to certification errors, we needed measures of the numbers of meals students received during SY 2012–2013 and of the income sources of student households at the time of their application for school meal benefits. Although the APEC-II study has data on these measures for most sampled students, they are not available for all students. This appendix describes the procedures we used to impute student meal participation and household income sources for sampled students missing that information.

A. Imputation of meal participation

A key input in calculating improper payments is the number of meals received by each sample member in each month of SY 2012–2013. For a portion of the sample in a given month, we have high quality administrative data on the number of meals received by the individual in the month. Schools tracked these meal counts electronically as students received their school meals. For other sample members, however, we were unable to obtain administrative data because the district does not track participation at the individual student level (either electronically or manually). For these students, we only have survey data on their participation status. These survey data obtain the students' reported participation status during the school day and school week just before the administration of the household survey (which was typically a month or two into the school year). Parents or guardians reported their students' participation. When possible, the respondent asked the student to confirm the number of meals reported. For these students, we have no other information on the number of school meals they obtained during other portions of the school year. Thus, we had to impute the actual number of meals received by these students in each month for which we had no administrative data.

This section describes the basic approach we used to impute students' monthly school lunch and school breakfast participation (that is, the number of meals received during the month) in cases in which we have no administrative data for a given student in a given month. First, we describe the overall approach to imputing these missing values. Next, we describe the details of the imputation model we estimated to determine the relationship between actual participation in a given month (as determined by the administrative data) and the response to the participation questions on the household survey (and other individual characteristics) among those with both administrative and survey data. We also describe the results of the estimation of this model. The final section describes the process we used for imputing monthly participation among sample members for whom we have survey data but no administrative data on participation.

1. Overall approach

We attempted to obtain participation data from schools' administrative records for all students in the household survey sample (in other words, all those whose parent or guardian completed a household survey). As shown in Table E.1, we successfully obtained the data for all months of the school year for more than half of this sample (55 percent in the case of NSLP participation). Students with incomplete participation data fall into two groups. In the first group, for about 20 percent of all students in the sample, we obtained participation from administrative data for some but not all months of the school year. The second group of students includes those for whom we have no administrative data on NSLP or SBP) meals received during the year. This

group makes up 25 percent of the sample in the case of the NSLP. The figures for SBP participation are similar.

Table E.1. Availability of administrative data on NSLP participation

	Percentage of students
Available for all months	55.3
Available for more than half of all months	16.2
Available for fewer than half of all months	3.4
Not available	25.1

Source: APEC-II study data.

APEC = Access, Participation, Eligibility, and Certification; NSLP = National School Lunch Program.

The process for imputing participation status for those with missing data for a given month consisted of three steps. The first involved estimating the imputation model using sample members with both administrative and survey data on participation. In particular, we estimated the relationship between the number of meals the student received in the month and a set of variables we believe can predict the number of meals received. The key predictor variable is the student's participation level as reported on the household survey.

In the second step, we calculated a predicted number of meals received in the month among students for whom we did not have administrative data on participation in that month. We calculated predicted participation by multiplying the estimated coefficients from the model estimated in the first step by the values of the model's independent variables for a given individual.

The third step in the imputation process involved accounting for the fact that the imputation model did not perfectly explain variation in individuals' monthly school meal participation. In other words, there was some variation in the number of meals received among students who had the same predicted participation levels. Therefore, after calculating each student's predicted number of meals received (among those missing administrative data on participation in a given month), we calculated the student's imputed number of meals received. This imputed number of meals was set equal to the predicted number of meals plus an imputed value of the stochastic error term (which included one part that varied only across individuals and a second part that varied across both individuals and months). The inclusion of this error term ensures that our imputation process does not artificially reduce the overall variation of monthly participation.

2. Estimating the imputation models

We estimated separate imputation models for the number of school breakfasts and school lunches received in a given month. Each of these imputation models was estimated using data from sample members for whom we have both administrative data on meals received in a given month and survey data that could be used to define the key explanatory variables included in the model.

a. Imputation model

This imputation models can be represented as

$$(1)\quad M_{it} = \alpha_0 + X_i\beta + Z_{it}\delta + MP_t\theta + \alpha_1 P_i + \alpha_2 P_i^2 + \alpha_3 P_i * MS_{it} + \alpha_4 P_i * H_i + u_i + e_{it},$$

where M_{it} = number of meals (breakfast or lunch) received by student i in month t

X_i = time-invariant student characteristics related to participation in month t

Z_{it} = time-varying student characteristics related to participation in month t

P_i = survey-based indicator of student i's meal program participation

MP_t = set of binary variables indicating current month (t)

MS_{it} = number of months between survey and month t

u_i = student-level error term

e_{it} = student/month error term

In the model, the outcome (or dependent variable) is the number of school meals received during a given month by a given student. We regressed this outcome on a set of factors that vary by student but not by month (X_i), a set of factors that vary by student and month (Z_{it}), the month in which participation is being measured (MP_t), and a set of terms representing a function of the student's participation as reported on the household survey (P). We describe this set of terms in greater detail next.

The error structure in the model consists of two components—individual-specific and random error terms. The individual-specific term (u_i) represents unobserved factors that explain why a given individual received more or fewer meals in a month than would be expected given his or her survey-reported characteristics and other observed factors. This term is constant for a given individual across months of the school year. The random term (e_{it}) represents factors that are specific to both the individual student and month that cause the student to receive more or fewer meals in a month than would be expected given his or her survey-reported characteristics and other observed factors. One of the model's assumptions is that e_{it} is independent and identically distributed across sample members and months, whereas u_i is independent and identically distributed across sample members.

Household survey-reported school meal participation is a key predictor of a student's actual monthly participation (based on administrative data). The measure is based on a survey question about whether the student received a school lunch (breakfast) on the previous school day as well as questions about the student's participation on each day of the most recently completed full school week before the survey. Thus, we have information on whether students received school breakfasts and lunches on as many as six school days during the month in which the survey was conducted. However, we do not have six data points for each student in the sample for two

reasons. First, for some students, the previous school day was part of the most recent full school week, so there are participation data for only five days for these students. Second, students might not have attended school during every day of the most recent school week (or during the previous school day). These students could not possibly have eaten a school meal on those days that they did not attend school.

To account for these various situations, the key survey-reported participation variable measured the proportion of days that the student received a school lunch (breakfast) among days on which he or she could possibly have done so. If, for example, a student with complete data received a school lunch on the previous day and on three of the five days of the most recent full school week, the value of this variable would be 4/6, or 0.67. If, on the other hand, the previous school day was the Friday of the most recent school week (for this same student), the value of the variable would be 3/5, or 0.60.

Table E.2 shows the distribution of the survey-reported NSLP and SBP participation variables. Because parents or guardians responded to the household survey, we had some concerns about the accuracy of the information they reported on their children's school meal participation. Parents might not be truly aware of how frequently their child eats a school meal. In some cases, sample members (students) sat in on that part of the household survey interview and helped their parents answer those questions; in those cases, we expect the information to be more accurate than if students did not help their parents answer the questions on school meal participation. Thus, the imputation model includes an interaction term that allows the estimated relationship between the survey-based participation variable and the number of meals received during the month (the dependent variable) to vary according to whether the sample member helped his or her parent answer those questions.

Another issue with using survey-based participation as a predictor of the actual number of meals received during a given month is that the survey-based variable represents participation during a single week just before the survey was conducted—typically very early in the school year. By contrast, the dependent variable represents meals received during any given month during the school year. It seems reasonable to expect that the survey-based participation variable would more accurately predict meals received early, rather than later, in the school year. Thus, the model also includes an interaction term that allows the estimated relationship between survey-based participation and the number of meals received to vary according to the number of elapsed months between the month in which the survey was conducted and the current month (the month in which the dependent variable was being measured).

Table E.2. Proportion of days student was reported to have received a reimbursable meal during survey reference period (percentage of students)

	NSLP	SBP
Zero days during reference period	11.6	49.6
Some, but not all, days during reference period	6.9	11.1
All days during reference period	81.5	39.3

Source: APEC-II study data.

APEC = Access, Participation, Eligibility, and Certification; NSLP = National School Lunch Program; SBP = School Breakfast Program.

Survey-based participation enters the imputation model in two additional ways. First, to allow for the possibility that the relationship is nonlinear, the model includes a squared value of survey-based participation. Second, the model also includes a survey-based variable representing the student's participation in the other meal program. In other words, the NSLP lunch participation model includes the proportion of days during the reference period the student was reported to receive a school breakfast (in addition to all of the survey-based lunch participation variables). Similarly, the SBP breakfast participation model includes a survey-based lunch participation variable.

The model includes various other explanatory variables in the vectors X_i and Z_{it}. Each of these variables is potentially related to the number of reimbursable school meals a student received during a given month and is available for students for whom we have to impute participation. In addition to the survey-based participation variables described earlier and a set of binary variables indicating the current month, the model includes the following variables:

- Age and age squared
- Gender
- Race/ethnicity
- Household size and structure
- Free/reduced-price income eligibility
- Free/reduced-price certification status
- Whether directly certified or categorically eligible
- Whether the respondent reported that the student did not eat a school lunch (breakfast) on every day of the previous week because he or she ate at home or does not always like the food served[16]
- Whether the respondent reported that the student did not eat a school lunch (breakfast) on *any* day of the previous week because he or she regularly eats at home or never likes the food served
- Index of the student's overall level of satisfaction with school meals
- Index of the parent's overall level of satisfaction with school meals
- Proportion of days absent from school

b. Estimation method

To determine the appropriate technique to estimate the model, we first had to assess the extent to which the value of the dependent variable in this model is limited or censored, in the sense that it could not exceed the total number of days the student attended school in the month or be less than zero. If a large enough proportion of observations are censored at either of these

[16]This variable was designed to control for differences between students who skip a day or two of school lunch (breakfast) because of persistent issues (such as not liking the food or making it a habit to eat at home) versus transitory issues (such as having a one-time conflict with another school event).

limits, then estimates from Equation (1) could be biased. For example, although an increase in survey-reported participation presumably is associated with an increase in the number of meals received during the month, changes in survey-reported participation cannot have this effect among students who already received a school meal on each school day during the month (that is, who already have the maximum value of the dependent variable).

We have incomplete data on the extent to which the administrative data on meals received during a month are censored. We know the proportion of students who do not receive any school meals during the month—it is reasonably small (21 percent) for lunch participation but much larger (53 percent) for breakfast participation—but we cannot determine the exact extent of censoring at the upper limit on meals received. This upper limit varies across students, both because students attend different schools with different numbers of school days during a given month and because students are absent from school for different (and unobserved) numbers of days during a month. Thus, although we have an estimate of the number of school days in each month of the school year, this information is not specific to particular students or particular schools.

Based on the extent of censoring at zero, we imputed NSLP participation based on Equation (1) alone. In the case of SBP participation, however, we estimated a two-stage model to account for the substantial censoring at zero. The first stage is a logistic regression model predicting whether students received any SBP breakfasts at all during a given month. The second stage is an ordinary least squares (OLS) model of the number of breakfasts received during the month, conditional on consuming some positive number. In other words, we used the full sample of students to estimate the logistic regression model of "any participation" but used only the portion (slightly less than half) of the sample with some participation to estimate the OLS model of the number of meals actually received.

3. Imputing the number of meals received

To impute the number of NSLP lunches received in a given month among students for whom that information is missing, we first used the coefficient estimates from the imputation model to calculate predicted meals received. This calculation involved multiplying the values of the independent variables included in the imputation model by the appropriate coefficient estimates from the model for each student for whom we wished to make an imputation. To calculate predicted school lunches received in September for a given student, for example, we used

$$(2) \quad \hat{M}_{i,Sept} = \hat{\alpha}_0 + X_i\hat{\beta} + Z_{i,Sept}\hat{\delta} + MP_{Sept}\hat{\theta} + \hat{\alpha}_1 P_i + \hat{\alpha}_2 P_i^2 + \hat{\alpha}_3 P_i * MS_{i,Sept} + \hat{\alpha}_4 P_i * H_i$$

The next step involved simulating the error terms of the imputation model (u_i and e_{it}) for individual students. Relying on the predicted values calculated in Equation (2) alone would provide a reasonable set of estimates of the number of meals received by students, on average, but would not accurately reflect the variability in this outcome across the population. By adding the two error terms to the predicted values, we generated a set of imputed values that accurately estimates actual meals received on average and accurately represents the true variability of meals received among this population of students.

To simulate the error terms, we made some assumptions about their behavior in the population. In addition to assuming that each is independent and identically distributed and that the error terms are independent of each other and of the explanatory variables included in the imputation model, we also assumed that both u_i and e_{it} are normally distributed.

Thus, we drew a single value of u_i for each individual student for whom we had to impute meals. This value was drawn at random from a normal distribution with mean 0 and standard deviation $\widehat{\sigma_u}$. This single draw of u_i for a given individual was used for each of that individual's monthly observations. We then drew a value of e_{it} at random from a normal distribution with mean 0 and standard deviation $\widehat{\sigma_e}$. We drew separate values of this error term for each individual and for each month we wished to make an imputation for that individual. Finally, for a given month t and a given individual i, the imputed number of meals received was set equal to the predicted number of meals plus the sum of the two error terms:

$$(3) \quad M_{it}^{*} = \hat{M}_{it} + u_i^{*} + e_{it}^{*}$$

We made one final adjustment to this imputation. We bottom-coded the imputed value at 0 and top-coded the value at our estimate of the number of school days in the month. In other words, if the value resulting from Equation (3) was less than 0, we set the imputed number of lunches to 0 because a student could not receive a negative number of meals during the month. Similarly, we assumed that the student could not receive more lunches during the month than the number of days school was in session, so if the value resulting from Equation (3) was greater than the number of school days, we set the imputed number of lunches equal to the number of school days. These adjustments affected a small proportion of cases.

Because we used a two-stage model to impute school meals received for the SBP, the process for imputing the number of breakfasts received was a bit different. The first step involved imputing whether a particular student had received any breakfasts. We did this by calculating the predicted probability that a given student had received any breakfasts using the coefficient estimates from the logistic regression model along with values of the independent variables. If this predicted probability was less than 50 percent, we assumed that the student did not participate in the SBP and imputed 0 breakfasts received during the month. If the predicted probability was 50 percent or more, we moved to the second-stage OLS equation. In particular, we used the same process for imputing the specific number of breakfasts received for this group as we used for imputing the number of lunches described earlier.

4. Imputation results

To assess the accuracy of our imputation process, we imputed values of the monthly lunch and breakfast totals not only for students lacking administrative data on participation but also for students for whom we had the administrative data. For students with administrative data, as a result, we have both an actual and an imputed value of the number of reimbursable meals received. Table E.3 shows the distributions of the actual and imputed monthly number of NSLP lunches (and SBP breakfasts) received among students for whom we have both administrative and survey data.

The imputation process generates a distribution of meals received that differs somewhat from the actual distribution at the extremes of the distribution, but it does capture the typical number of meals received fairly accurately. For example, the proportion of students who did not receive any NSLP lunches during the month (21 percent) is more than twice as large as the proportion with an imputed value of 0 (8 percent). Similarly, the proportion who actually received 16 or more lunches per month (33 percent) is greater than the proportion with imputed values this high (29 percent). On the other hand, the mean imputed number of lunches received (10.6) is very close to the mean actual number received (11.0) for this group of students.

Because the purpose of the imputation process is to allow us to more accurately estimate the dollar amount and rate that free or reduced-price meal reimbursements that are improper, its most important attribute is that it be relatively accurate for both students certified in error and those whose certification status is accurate. In particular, we would like the mean number of imputed meals to equal the mean number of actual meals for both groups of students. If, by contrast, the imputation process overestimated meals consumed among those certified in error and underestimated meals consumed among those certified accurately, the resulting estimate of the rate of improper payments would be biased.

By this measure, the imputation process is successful. Among students who have been certified accurately, the mean number of lunches actually consumed was 10.9 and the imputed mean was 11.3. Among those certified in error, the mean number actually consumed was 9.6 and the imputed mean was 10.2.

Table E.3. Distribution of actual versus imputed meals received among students with administrative participation records

	NSLP		SBP	
	Actual	Imputed	Actual	Imputed
Number of meals				
0	21.4	7.7	52.5	57.0
1 to 5	8.4	13.7	13.9	6.5
6 to 10	12.6	23.3	11.2	13.2
11 to 15	25.0	26.4	12.3	13.4
16 or more	32.7	28.9	10.2	9.7
Median number	13.0	10.9	0.0	0.0
Mean number	10.6	11.0	4.6	4.7

Source: APEC-II study data.

APEC = Access, Participation, Eligibility, and Certification; NSLP = National School Lunch Program; SBP = School Breakfast Program.

The imputation process is somewhat more accurate in the case of SBP breakfasts than in the case of NSLP lunches, primarily because a large proportion of students do not participate in the SBP and the model is fairly successful in identifying those students. For example, the administrative data suggest that 52.5 percent of students in a typical month do not receive school breakfasts, whereas the imputation process resulted in 0 meals received for 57.0 percent of students (Table E.3). Among students with positive numbers of meals received, the model again somewhat underestimates the proportion of students at either end of the distribution, but the mean number of imputed breakfasts received (4.7) is very close to the actual mean (4.6). Again,

we found that the mean number of meals received was close to the actual number for students certified erroneously and for those certified accurately.

B. Imputation of household income sources and amounts

The survey contained a comprehensive set of questions about who was in the sampled student's household and how much income each person in the household had during the reference month (month covered by the household's meal benefit application). For household members older than age 16, we asked the respondents whether they received income—and, if so, the dollar amount received—from each of 23 possible sources. For children ages 16 and younger, we asked the respondents whether they received income from four different sources and the amount from each source. Information on income sources and amounts were missing for some sample members because of survey item nonresponse. Table E.4 shows that rates of item nonresponse for income items were extremely low; item nonresponse is less than 1 percent for all income sources.

Table E.4. Prevalence of missing income data, by income source (percentage)

	Missing income responses
Adult income sources	
Job	0.13
Unemployment compensation	0.09
Worker's compensation	0.07
Social Security	0.07
Private pensions	0.15
Veteran's benefits	0.06
Supplemental Security Income	0.09
Alimony	0.06
Child support	0.07
Interest and dividends	0.06
Rental income	0.07
Nonfarm business, partnership, or professional practice	0.06
Farm	0.06
Financial aid for college	0.16
Savings withdrawals	0.19
Regular contributions from people outside the household	0.15
Other cash income	0.07
General assistance	0.07
Nonmilitary housing subsidies	0.03
Black lung benefits	0.07
Other public assistance	0.02
Child income sources	
Child support	0.69
Social Security	0.79
People outside the household	0.55
Other	0.51

Source: APEC-II study data.

Note: Other child income includes adoption assistance, State assistance, and other sources specified by the survey respondent.

APEC = Access, Participation, Eligibility, and Certification.

Our imputation strategy for replacing these missing values mirrors the strategy described earlier for imputing SBP meal participation; we first model whether the household received the

income source at all and then model the amount of income received for households that did receive the income source. The rest of this section describes our imputation methods and the results of the imputation.

1. Imputation models

To impute missing income amounts, we estimated a two-stage model for each type of income included on the survey. This approach enabled us to account for the fact that most income categories are relatively uncommon, leading to substantial censoring at zero. The first stage of the model is a logistic regression model predicting whether the household member received income from the source during the target month. The second stage is an OLS model of the amount of income received from the income source, conditional on receiving some positive amount. In other words, although the full sample of household members was used to estimate the logistic regression model of "any income received from the source," we used only the portion of the sample with some income to estimate the OLS model of the amount of income received.

The imputation models include variables that are potentially related to income receipt. These variables include the following:

- Age of household member and age squared
- Education level of survey respondent
- Household size and structure
- Race/ethnicity
- Free/reduced-price certification status of target child
- Proportion of sample household members in the school district receiving the income source

In addition to these variables, the models of the amount of income received include the average amount of income received among sample household members in the same school district who received any income from that source.

2. Imputing income received

The first step of imputing income from each source involved imputing whether a household member had received that income source at all. We did this by calculating the predicted probability that a given household member had received the income source using the coefficient estimates from the logistic regression model along with values of the independent variables. If this predicted probability was less than 50 percent, we assumed that the household member did not receive the income source. If the predicted probability was greater than or equal to 50 percent, we moved to the second stage of the imputation process. In particular, we used coefficient estimates from the model of income amounts (conditional on receiving any income) along with values of the independent variables to estimate a predicted value of the amount received from the income source. Next, we simulated the income amount error term of the imputation model for individual household members and added it to the predicted income amount value. We estimated the simulated error term values assuming that the error term is independent of the explanatory variables included in the model and that it is normally distributed. By adding the simulated error terms to the predicted values, we generated a set of

imputed values that both accurately estimates actual income received on average and accurately represents the true variability of income received among this population.

3. Imputation results

To assess the accuracy of our imputation process, we imputed values of income sources for household members for whom the survey data were missing and for those for whom we had the income source and amount data. For household members with nonmissing income data, as a result, we have both an actual and an imputed value for each income source. Table E.5 shows the actual and imputed mean income from each source for household members with nonmissing survey data for that source. Actual and imputed mean income estimates are similar for most income sources. Mean imputed income from jobs for adult household members in the sample is about $200 more than their actual income from jobs as reported by them. The positively skewed distribution in actual income among those with positive earnings drives this difference, which leads to higher imputed values on average. Imputed mean income differs by no more than $2 from actual mean income for all other data sources. These findings suggest that our methods for imputing values for the small number of household members with missing income information are acceptable.

This page has been left blank for double-sided copying.

APPENDIX F

ALTERNATIVE DEFINITIONS OF CERTIFICATION ERROR

This page has been left blank for double-sided copying.

ALTERNATIVE DEFINITIONS OF CERTIFICATION ERROR

The base measurement of improper payments used in the analysis of certification error for APEC-II is methodologically consistent with the approach used in APEC-I. As was the case in APEC-I, we incorporated, to the extent possible, policies and procedures specified by the USDA Food and Nutrition Service when determining student eligibility for free or reduced-price meals for students who applied for meal benefits.

To determine the extent to which alternative definitions and assumptions affect our base estimates of the rates of improper payments, we developed a series of sensitivity checks. These checks test the effect of changing criteria for inclusion in the economic unit as well as assumptions about the eligibility of students who had missing or incomplete applications or direct-certification documentation or who reapplied later in the school year. To reduce complexity, our sensitivity analyses exclude Provision 2 or 3 schools in a non-base year.

To determine whether a student meets income eligibility requirements for school meal benefits, school officials compare household size and total household income with income eligibility guidelines. According to the August 2013 Eligibility Manual for School Meals issued by USDA, the household or economic unit is defined as:

> A group of related or unrelated individuals who are not residents of an institution or boarding house but who are living as one economic unit, and who share housing and/or significant income and expenses of its members. Generally, individuals residing in the same house are an economic unit. However, more than one economic unit may reside together in the same house. Separate economic units in the same house are characterized by prorating expenses and economic independence from one another.

To determine the eligibility status of a student's household, we use information provided in the household survey. This approach requires determining which of the persons reported by the respondent as residing in the household at the time of application was actually part of the economic unit. The base measurement of improper payments considers all relatives—including relatives by marriage or adoption—of the sampled student living in the same household to be in the same economic unit, but considers any nonrelatives living in the household to be outside the economic unit. Excluding unrelated persons from the economic unit affects the measure of a student's eligibility in two ways: (1) it reduces household size, and (2) it reduces household income by the amount of any income these nonrelatives have. Because household size and household income affect eligibility status in opposite directions, omitting nonrelatives could theoretically result in either increasing or decreasing the student's eligibility. To test the sensitivity of eligibility—and therefore our estimate of amounts and rates of improper payments—to our assumption about nonrelatives, we computed an alternative version of eligibility that redefines an economic unit to include all household members reported in the survey.

Table F.1 shows how the different assumptions about who is in the economic unit affect measures of household size, income, and students' eligibility for free or reduced-price meals. Under the main definition of an economic unit used for this report, the mean household included about 4.6 members and had a monthly income of $2,538. Including unrelated persons in the economic unit had only a marginal effect on mean household size, which still rounded to 4.6

people, and mean monthly household income, which increased by only $4. Not surprisingly, these minor differences between household size and income resulted in very small shifts in the level of benefits for which a student was eligible. Including unrelated household members in the economic unit increased the total number of students eligible for free meals and reduced-price meals by only 0.02 and 0.06 percentage points, respectively, and decreased denied applicants by just 0.08 percentage points.

Table F.1. Household size and eligibility for free or reduced-price meals, under alternative specifications of economic unit

	Definition of economic unit	
	Main definition: excludes unrelated household members	Alternative definition: includes unrelated household members
Household size (mean)	4.58	4.62
Household income (mean)	$2,538	$2,542
Percentage of students:		
Eligible for free meals	70.59	70.61
Eligible for reduced-price meals	13.12	13.18
Ineligible for free or reduced-price meals	16.29	16.21

Notes: The eligibility estimates in this table are not adjusted for carryovers, incomplete applications, nonrespondents to verification, re-applicants, or newly certified cases.

The measurements of certification error and improper payments used in APEC-I and APEC-II require making assumptions about the eligibility of students in certain circumstances. To test these assumptions, we conducted tests of the sensitivity of the results on the way we treated (1) the eligibility of students in which the applicant failed to properly complete meal benefit applications, (2) applications and direct certification documents missing from SFA files, and (3) the circumstances of students who reapplied for NSLP or SBP benefits later in the school year.

We created four versions of the eligibility variables. The first version is the base measure of eligibility and is methodologically consistent with that used for the base measurement of eligibility in APEC-I. The base measurement was used to compute the certification error and improper payment measurements presented in this report and assumes the following:

- Students are considered to be eligible for benefits based on the information on household circumstances provided in the household survey even if the applicant did not properly complete necessary NSLP/SBP paperwork—specifically, if the applicant either (1) submitted an incomplete application or (2) failed to respond if selected for verification. If, for a particular student, an application is found that does not include all required information, it is considered an administrative error. However, if the certification status is correct based on the household's circumstances reported on the application and household survey, then it is not considered a certification error and thus is not included in the calculations of improper payments. (Therefore, administrative errors do not always equate to certification errors and thus improper payments.)
- Students are considered to be ineligible for free or reduced-price meals if the application for free or reduced-price meals or direct certification information could not be found in the

SFA's files. This definition corresponds to FNS rules, which require SFAs to keep all applications on file, and conforms to how the State's Coordinated Review Effort (CRE) reviewers handle these situations. When CRE reviewers encounter a student receiving meal benefits without an application on file or a student on the directly certified list, the State agency is required to recover the free or reduced-price portion of the reimbursement paid for the meals served to that student. In this study, we consider missing applications an administrative error and a certification error, and they are included in the estimate of improper payments.

- Students who reapplied for benefits later in the year after their initial application are still assumed to be eligible for the same level of benefits reflected in the data reported in their initial household survey. It is possible that the reapplication was prompted by a change in household circumstances, but we did not collect survey data on the households at the later date, so information from the initial survey is the best estimate.

Each of the three alternative versions we implemented reverses one of these assumptions:

- **Alternative 1: Ineligible if application is incomplete.** Under this alternative, we consider students to be ineligible for meal benefits if they submitted an incomplete application or did not respond to a verification request, regardless of whether the certification decision is consistent with the information on household circumstances reported on the application or in the study's household survey. This alternative would result in a lower rate of overcertification, and a slightly higher rate of undercertification, than the base definition.
- **Alternative 2: Missing application or direct certification documentation does not automatically mean ineligible.** Under this alternative, households are considered to be eligible for benefits based on the information on their household circumstances provided in the household survey, even if their application could not be found in the SFA's files. This alternative would result in a sizable decrease in the overcertification rate and a slight increase in the undercertification rate.
- **Alternative 3: Re-applicants certified without error.** The final alternative version assumes the information reported by re-applicants on their new application was accurate and the SFA assessed it correctly. Because this alternative assumes away the occurrence of reporting and administrative error after a reapplication, it would result in a lower rate of certification error than the base definition.

Table F.2 provides estimates of improper payments under these four specifications for students who applied for meal benefits in schools that do not use Provision 2 or 3 or that are Provision 2 or 3 schools in their base year. Alternative 1 and Alternative 3 resulted in gross improper payment estimates that differed just slightly from the NSLP and SBP base estimates presented in the main text, which were 9.6 percent and 10.6 percent, respectively. Considering students for whom incomplete applications were submitted to be ineligible (Alternative 1) resulted in gross improper payment rates of 9.7 percent for the NSLP and 10.6 percent for the SBP, and assuming that students reapplying later in the school year were certified without error (Alternative 3) yielded gross improper payment rates of 9.2 percent for NSLP and 10.2 percent for SBP.

Alternative 2 resulted in gross improper payment estimates that differed more from the NSLP and SBP base estimates. Removing the requirement that students whose applications were not found in the SFA's files were ineligible for benefits (Alternative 2) resulted in gross improper payment rates of 8.7 percent for the NSLP and 9.3 percent for the SBP.

Table F.2. Improper payments due to certification error in the NSLP and SBP, under alternative specifications of eligibility

	NSLP	SBP
Base measure		
Improper payments as percentage of free or reduced-price reimbursements		
Overpayments	6.91	7.19
Underpayments	2.68	3.39
Gross improper payments	9.60	10.58
Net improper payments	4.23	3.80
Alternative 1: Ineligible if application incomplete		
Improper payments as percentage of free or reduced-price reimbursements		
Overpayments	6.98	7.23
Underpayments	2.68	3.37
Gross improper payments	9.66	10.60
Net improper payments	4.29	3.85
Alternative 2: Missing application or direct certification documentation does not automatically mean ineligible		
Improper payments as percentage of free or reduced-price reimbursements		
Overpayments	5.14	4.71
Underpayments	3.58	4.57
Gross improper payments	8.72	9.28
Net improper payments	1.56	0.14
Alternative 3: Re-applicants certified without error		
Improper payments as percentage of free or reduced-price reimbursements		
Overpayments	6.53	6.85
Underpayments	2.65	3.32
Gross improper payments	9.18	10.16
Net improper payments	3.88	3.53

Notes: These estimates exclude schools that are Provision 2 or 3 schools in their non-base year.

NSLP = National School Lunch Program; SBP = School Breakfast Program.

In addition to the alternative estimates of improper payments presented in Table F.2, we produced a set of alternative estimates of based on alternative meal counts for denied applicants with improper certification statuses. Table F.3 provides estimates of improper payments computed by replacing the actual meal counts for improperly denied applicants with alternative meal counts that represent the imputed number of meals they would have received if their certification status had been correct.[17] For example, free-eligible students who were erroneously denied benefits would likely have received more meals if their certification status had been free. These alternative estimates result in identical estimates of overpayment rates to those using the base measure of meal imputation. Underpayment rates increased just slightly by less than one percentage point for both the NSLP and SBP, resulting in slightly higher gross improper payments.

Table F.3. Improper payments due to certification error in the NSLP and SBP, using alternative meal imputation for denied applicants with erroneous certification statuses

	NSLP	SBP
Base measure		
Improper payments as percentage of free or reduced-price reimbursements		
Overpayments	6.91	7.19
Underpayments	2.68	3.39
Gross improper payments	9.60	10.58
Net improper payments	4.23	3.80
Alternative meal imputation for erroneously denied applicants only		
Improper payments as percentage of free or reduced-price reimbursements		
Overpayments	6.91	7.19
Underpayments	3.35	3.69
Gross improper payments	10.26	10.88
Net improper payments	3.56	3.50

Notes: These estimates exclude schools that are Provision 2 or 3 schools in their non-base year.
NSLP = National School Lunch Program; SBP = School Breakfast Program.

[17] See Appendix E for a detailed description of the meal imputation process.

This page has been left blank for double-sided copying.

This page has been left blank for double-sided copying.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ WASHINGTON, DC

MATHEMATICA
Policy Research