APPENDIX A

DESCRIPTION OF DATA COLLECTION

This appendix describes the primary data collection activities for the Direct Certification Study. Two primary sources of data were used in the study: (1) the School Food Authority (SFA) survey and (2) State administrative records data. Section A outlines the procedures used to conduct the SFA survey and summarizes the results of efforts to interview a nationally representative sample of school districts. Section B describes the collection of State administrative records data on the FS/TANF/FDPIR receipt of families living in the areas covered by the districts in the sample.

A. SFA SURVEY

1. Questionnaire Development and Pretesting

The development of the SFA questionnaire began in November 2000. The questionnaire drew heavily on the 1996 Study of Direct Certification so that it would be as similar to this previous study as possible. The questionnaire also drew on MPR's previous experience with surveys of school districts, such as the first School Nutrition Dietary Assessment (SNDA-1) Study. To address the objectives of the Direct Certification Study, the SFA questionnaire was organized into eight sections:

- ***Section 1:*** This section introduced the respondent to the interviewer and provided the respondent with an overview of the study's purpose.
- ***Section 2:*** This section was designed to ensure that the interviewer talked to the appropriate person: the one most knowledgeable about the NSLP and certification for free meals in the program.
- ***Section 3:*** This section provided both contextual information about how districts implement aspects of the NSLP and information on a district's actual levels of enrollment, certification, and participation.
- ***Section 4:*** This section collected information on whether the district uses direct certification and, if so, how it is implemented.
- ***Section 5:*** This section collected information on the application and verification process for free and reduced-price certification.

- ***Section 6:*** The questions in this section collected information on the timing and other aspects of direct certification activities.
- ***Section 7:*** This section collected additional descriptive information on direct certification, such as problems that districts had in implementing this process, districts' views on the costs and benefits of using direct certification, and changes in the process since the previous study was conducted in 1996. For districts that do not use direct certification, this section collected information on their reasons for not using it.
- ***Section 8:*** This section collected descriptive information about the districts in the sample and the person who was the primary respondent in each.

After the questionnaire was developed, it was revised on the basis of comments from the USDA's Economic Research Service (ERS). It was then pretested to detect any remaining problems, such as poorly worded questions, terms that were not defined, missing or inadequate response categories, difficult transitions between questionnaire topics, or erroneous or unclear interviewer instructions. The pretest was also used to assess the flow of the entire instrument and the burden on respondents. It was conducted with nine respondents in February and March 2001. The questionnaire was revised during spring 2001, with revisions based on the results of the pretest and additional comments from ERS.

This revised version of the SFA questionnaire was submitted to the Educational Information Advisory Committee (EIAC) for consideration at its May 2001 meeting. After MPR revised the questionnaire based on EIAC comments, EIAC granted its approval for data collection activities in June 2001. The questionnaire was next submitted (along with supporting materials) to the Office of Management and Budget (OMB). OMB approval was granted in November 2001.

Computer-assisted telephone interviewing (CATI) programming of the survey instrument began in September 2001, during the OMB clearance process. This kept the time between receipt of OMB clearance and the start of interviewing as short as possible. All changes required

by OMB were incorporated into the final version of the instrument. Programmers, survey staff members, and interviewers extensively tested the program for accuracy.

2. Sample Selection and Weights

The target population for this study consisted of all SFAs in public school districts participating in the NSLP in the 50 States and the District of Columbia. The sampling frame consisted of 14,571 school districts included on the National Center of Education Statistics (NCES) Common Core of Data (CCD).[64] The frame included information on the number of students in each school district, address information, and some demographic characteristics of the districts' students.

a. Stratification

Stratification is used to form relatively homogeneous groups from which separate samples are selected. These separate samples are then combined to form the overall sample. Stratification was used in the SFA survey to facilitate oversampling of subgroups of SFAs and increase the precision of some estimates.[65] The sample employed a stratified design that was modeled on the sample design of the 1996 study and consisted of nine strata. One stratum (9) was reserved for the largest districts, which were selected with certainty (that is, all of these districts were selected). The other strata were formed based on the number of students in the

[64]The CCD is an electronic database of school districts available from the NCES. No frame of SFAs was available. In a large majority of cases, however, there is one SFA per district. The CCD-based sample frame of 14,571 includes a small proportion of districts that do not participate in the NSLP; thus, the true sample frame is slightly smaller than this.

[65]Where strata are formed that are homogenous with respect to the variance of a statistic, the sampling error of that statistic can be reduced (see Kish 1965). In this study, for example, one key estimate is of the prevalence of direct certification. Since the variance of a proportion is higher near 0.50 and lower in the tails, stratifying on the expected prevalence of direct certification and oversampling from the high variance strata may increase the precision of this estimate.

school district and on the estimated percentage (in 1998) of SFAs using direct certification in the State in which the district is located. The nine strata are summarized in Table A.1.

TABLE A.1

DESCRIPTION OF SAMPLE STRATA

Stratum[a]	Estimated Prevalence of Direct Certification in State (Percent)	Number of students
1	0 to1	< 28,000
2	99 to 100	< 28,000
3	More than 1, up to 8.5	< 28,000
4	More than 91.5, less than 99	< 28,000
5	More than 8.5, up to 22	< 28,000
6	More than 78, up to 91.5	< 28,000
7	More than 22, up to 78	< 5,000
8	More than 22, up to 78	5,000-27,999
9	Any	>= 28,000

[a]The stratum numbers are those used in sampling. Stratum 9 consists of districts to be selected with certainty. The other (noncertainty) strata are in ascending order of expected sampling error for an estimate of the prevalence of direct certification.

b. Sample Allocation Plan

Districts were selected at different rates from the different strata. In particular, districts were sampled at the highest rates from strata containing the largest districts and also from strata contained districts in States in which the estimated prevalence of direct certification was closest to 50 percent. As noted above, all districts from stratum 9 were selected into the sample. The sampling rates were lowest in strata 1 and 2 (because these contained districts in States with an estimated prevalence of direct certification very close to either 0 or 100 percent) and in stratum 7 (because it contained the smallest districts). Aside from stratum 9, the sampling rate was highest in stratum 8, because it contained no small districts and only districts in States with estimated prevalence of direct certification relatively close to 50 percent. In addition (as described in

greater detail below), from within each stratum, districts were selected into the sample with probability proportional to size.

To evaluate the implications of alternative sample designs and choose the best of these alternatives, we conducted simulations involving sample allocation options using five different measures of size (MOS) for purposes of the probability proportional to size sampling. Using the square root of the number of students as the MOS gave us the best results, and was the process used in selecting this sample.[66] The final sample design was also similar to the design of the 1996 study in that it defined sample strata on the basis of district size and estimated prevalence of direct certification in the State.

While we estimated that only 1,547 districts would be needed to achieve the desired number of SFA interviews, we selected an overall sample size of 1,847 to allow for additional sample if survey response was lower than expected. Stratum 9 contained 204 districts, and all of these were selected into the sample. From the remaining eight (noncertainty) strata, districts were selected with probability proportional to the measure of size described above (the square root of the number of enrolled students in the district). We used the formula given below to determine the specific number of districts to be selected in each of the eight noncertainty strata.

$n_h = n_{noncert} * (W_h)$ where h = 1, 2,…, 8

n_h = sample size in stratum h

W_h = stratum h's MOS/Sum (of the MOS of all eight noncertainty strata)

$n_{noncert}$ = the total sample size across all noncertainty strata
= 1,847 - 204 = 1,643

[66]Details of the alternative sampling allocations evaluated and the simulation results are available from the authors upon request.

We used the Chromy (1979) procedure to generate the sample of 1,847 school districts using a serpentine sort. The Chromy procedure is a systematic selection procedure that sorts the cases in each sampling stratum in a serpentine fashion based on specified characteristics. This procedure imposes implicit stratification beyond the explicit strata to ensure that the sample is balanced by specified characteristics. For the SFA survey sample, we used State and poverty rate as the implicit strata so that within each of the explicit strata districts were sorted first by State and then by poverty rate.

After selecting the overall sample of 1,847 school districts, we then used the Chromy procedure three more times to create four subsamples for sequential release. The first release consisted of 1,547 cases (all the certainty cases included), and the remaining three releases included 100 cases each. The additional releases were not needed, as the first release resulted in the desired number of 1,200 completes.

Table A.2 provides the total number of districts in each stratum and the number sampled from each stratum.

TABLE A.2

SAMPLING RATES, BY STRATUM

Stratum	Number of Districts in Sample Frame	Number of Districts Selected for Overall Sample	Number of Districts Released into Final Sample
1	656	56	46
2	4,084	393	321
3	2,329	261	213
4	1,314	171	140
5	1,700	225	184
6	676	110	90
7	3,216	314	257
8	392	113	92
9	204	204	204
Total Sample	**14,571**	**1,847**	**1,547**

To restore the sample to the proportional distribution of the population that it represents, sample weights were constructed. Each SFA's weight is the product of two factors:

- The inverse of the probability of selection of the district associated with the SFA
- The inverse of the response rate calculated for the SFA's sampling stratum

Combined, these factors correct for both under- and oversampling and for differences in response rates among strata.[67] Use of the sample weights will reduce, if not eliminate, the bias that would arise from using unweighted survey data.[68] The use of sample weights will also affect the estimated sampling error for estimates based on the sample. We used the SUDAAN statistical package to correctly estimate standard errors taking into account sample weights and the complex sample design.

In addition to these sample weights that make the SFA sample of districts representative of the population of public school districts (offering NSLP lunches) nationally, we constructed weights that allowed us to estimate the prevalence of particular district characteristics among all public school students nationally (or among particular groups of public school students, such as certified students). These weights were constructed by starting from the initial district-level weights described above and within each district, multiplying the weight by the relevant count of the number of students in the district.

[67]Other weighting steps often used in surveys—poststratification and trimming—were not employed for this sample. Poststratification was not used because there was no reliable external data source to use as a standard for adjusting the weights. Trimming was not called for because, in our judgment, there were no weights with extreme values.

[68]If there are unobserved factors that contribute to differences in SFAs' probability of survey response, use of the stratum-specific response rates in constructing weights might not fully eliminate nonresponse bias.

3. Conducting the Survey

To secure the cooperation of SFAs selected for the final sample and to facilitate the interviewing process, MPR took several steps before calling the school districts to conduct the survey:

- ***Chief State School Officer (CSSO) Letter.*** In December 2001, letters were sent to each State's CSSO—the State superintendent of schools—to explain the objectives of the study and data collection procedures.
- ***State EIAC Representative Letter.*** At the same time that the CSSO letters were mailed, a copy of the letter was sent to the EIAC representative from the same State.
- ***Child Nutrition Director Letter and Follow-Up Call.*** In December 2001, letters were sent to the State child nutrition directors explaining the study's objectives and data collection process. These letters asked each director to name a State-level study liaison and said that someone from MPR would call the director to obtain the name of the State liaison. As an attachment to this letter from MPR, a letter from USDA signed by a representative of ERS was included to assure the child nutrition directors of the legitimacy of the study and encourage their cooperation. The child nutrition directors were each contacted in early January 2002, and they all agreed to support the study and provide a State liaison for the study.
- ***Local Superintendent Letter.*** In mid-January 2002, letters were mailed to the local school district superintendents associated with each SFA in the study sample. These letters explained the study. They also provided the name of the State liaison for the study so the superintendents could contact the liaison to verify the study's legitimacy.
- ***SFA Advance Letter and Fact Sheet.*** Shortly after the superintendent letters were sent, letters were mailed to the directors of each SFA included in the sample. These letters described the study and alerted the SFA directors that they would be contacted during the next few weeks to participate in the study. Accompanying the SFA letters were fact sheets that the SFA directors (or their staff) were asked to complete before the interview. This fact sheet was designed to guide SFA directors in gathering information from their files before the interview.

Immediately after these supporting materials had been sent, and just before the interviews were to begin, interviewer training was held at MPR's central survey operations center near Princeton, New Jersey (January 22 to 24, 2002). The training included an overview of the project, a question-by-question review of the survey instrument, a discussion of frequently asked

questions and their answers, and practice interviewing and role playing. In addition, the interviewers provided input on the questionnaire, which was then revised as appropriate.

Telephone interviewing for the SFA survey began on January 25, 2002, and continued until March 22, 2002. Calls were placed to SFA directors from 8:00 A.M. to 4:00 P.M., local time, with appointments and callbacks scheduled at the convenience of the food service directors (usually early in the morning or after the lunch sessions). Specially trained monitors, who attended the interviewer training session, monitored all interviewers. In addition, the survey director and two assistants continually monitored interviews. When a respondent from one of the 15 largest school districts was interviewed, a member of the project supervisory staff monitored the interview for quality assurance.

One of the purposes of the extensive monitoring of interviews was to detect and correct problems with the survey instrument or with the way in which questions were asked. One problem discovered during the early stages of interviewing was that the percentage of districts reporting that they have schools that operate under USDA Provision 2 or 3 was much higher than expected. This finding led us to carefully examine the status of the districts that reported having Provision 2 or 3 schools, by checking their responses to survey questions and calling back the districts to verify their answers. This re-examination led us to conclude that respondents were confused by the wording of the question related to Provision 2 or 3 funding, which led some districts to incorrectly report that they had Provision 2 or 3 schools. For districts that corrected their responses to the Provision 2 or 3 questions on the survey when we called them back, we replaced their initial responses with these corrected responses on the data file. We also asked these districts questions that had been skipped previously because the district had reported receiving Provision 2 or 3 funding. Finally, we clarified the wording of the Provision 2 or 3 questions on the SFA survey to be administered to subsequently interviewed districts.

Several new questions were added to the SFA questionnaire midway through the January-March interviewing period. At the request of ERS, questions about electronic point of sale, or point of service (POS), systems were added to the end of the survey. Sample members were first asked if any schools used a POS system and, if so, how many schools did so and in what school year they started using it. Respondents who had completed the interview before these questions were added (about 500) were not called back to answer the new questions.

We used two tools to monitor each interviewer's productivity: (1) CATI reports and (2) interviewer productivity reports. CATI reports provided the results of each telephone attempt, on both a daily and cumulative basis, and showed, by interviewer, the numbers of calls made, completes, refusals, and minutes-per-complete. Interviewer productivity reports provided each interviewer's daily and cumulative hours worked and the number of interviews completed. Interviewers with low productivity or high refusal rates were closely monitored and retrained. Interviewers whose performance was unsatisfactory, even after remediation efforts, were removed from the project.

As interviewing neared completion, MPR analyzed the response rate by sample strata. Special reports were created that showed the disposition of all 1,547 sample districts, by stratum. We used this information to better target the survey resources to achieve similar response rates in each stratum. We focused particular attention on ensuring an adequate response rate in stratum 9, so that the largest districts nationally would not be underrepresented in the analysis sample.

4. Survey Response Rates

Overall, interviews were attempted with 1,547 school districts. Of this full sample, 35 districts (or about 2.3 percent) reported that they did not participate in the NSLP, which made them ineligible for the study. Excluding these districts, the full sample size was 1,512. The

response rate on the SFA survey was 81 percent, as full interviews were completed with 1,223 of the 1,512 districts in the sample.

There were several reasons for nonresponse. Just under eight percent of districts (or 114 of 1,512) refused to participate in the study. The reasons that these districts gave for their refusal included:

- They were too busy/had no time to get the information/were already overwhelmed with audits and paperwork.
- They were too new on the job to be able to answer the questions.
- They believed the USDA could obtain the requested information from other sources.
- They were not interested/did not want to participate because it was not mandatory.

Other districts did not participate because they could not be reached and did not return calls. Some districts were reached initially, could not complete the interview at that time, then could not be reached later. These two categories of cases accounted for 11 percent of the overall sample.

Response rates were fairly consistent across the nine sample strata. Response rates in the strata ranged from 66 percent to 86 percent, including a response rate of 86 percent in stratum 9 (which contained the largest districts).[69] Differences in response rates across the strata did not appear to be correlated with the size of districts in the strata. Response rates were also fairly consistent across States. Among the 32 States from which at least l5 districts were sampled, the minimum response rate was 62 percent, in New Jersey. The maximum response rate was 100 percent, in Oregon. Twenty-seven of the 32 States had response rates between 70 and 95 percent.

[69]In our analysis, we used the sample weights described above to adjust for these response rate differences by strata.

For each completed interview, an average of seven calls were made to districts to attempt interviews. The interview itself took an average of 29 minutes. In addition, respondents who completed interviews reported the amount of time they spent completing the fact sheet before the interview; the mean amount of time spent completing the fact sheet was two hours, although this mean value includes few districts that spent an extremely long time on the fact sheet. The median amount of time spent on the fact sheet was just one hour.

B. STATE ADMINISTRATIVE RECORDS DATA

The second major data source used in the Direct Certification Study is administrative records data from State food stamp or welfare agencies. For each of the districts in our sample that reported using direct certification, we attempted to collect State administrative data on receipt of FS/TANF/FDPIR benefits among children living in the areas covered by these districts. In particular, we asked for lists of school-aged children receiving benefits at two times around the 2001-2002 school year: (1) summer 2001, and (2) December 2001. For the summer time period, we requested that State welfare agencies submit data from the same month they had originally generated data for direct certification. For example, if a State welfare agency had generated the list of children receiving benefits in late July (using July data), we requested data from July. Although we initially requested data on FDPIR receipt, the relevant data on FDPIR benefits was typically kept separately from FS and TANF benefits, so we focused our data request on receipt of FS/TANF alone.[70]

[70]Most states provided us with information on recipients of FS and TANF. A few states gave us information on FS receipt only. In California, we used data from the MEDS system. This system was designed to track participation in California's public health insurance program, but it also has information on food stamp receipt (and nearly all—although not all—of California's food stamp recipients are on this system). We used data from the MEDS system instead of attempting to collect food stamp administrative records because the latter are kept only at the county level, and MEDS data are available in a single state-level file.

The information on the children/families on FS/TANF as of summer 2001 and as of December 2001 was collected from States to assess caseload turnover. In particular, the children eligible for direct certification based on FS/TANF receipt in the summer but not in December might be considered administratively ineligible for free meal certification as of this later date. We also conducted an analysis using data from the Survey of Income and Program Participation to determine the proportion of this group that may have remained income-eligible for free meal certification via written application, even if their families no longer received food stamps or TANF.

We took three steps to collect State administrative data: (1) we sent introductory letters to State food stamp directors, (2) we telephoned these directors to secure cooperation, and (3) we followed up with either the directors or data processing staff in their office to work out the details of the data provision. We sent the letters to food stamp directors in December 2001, targeting them because we thought they were most likely to be able to authorize the provision of the data we were requesting. The letters introduced the study, explained our data request, and alerted the directors that we would be calling them to seek their approval of the request. To help secure the cooperation of the food stamp directors, the central office of the USDA's Food and Nutrition Service (FNS) sent letters to its regional offices explaining the study and asking them to send letters or call the State food stamp directors in their region to encourage them to participate.

Starting in January 2002, study team members began contacting the State food stamp directors by telephone. They introduced the study, described the study's goals and the data being requested from States, and asked the directors if they had questions or concerns about the research. They then asked the directors to give permission for their States to participate in the

study and provide administrative data. These initial calls to State food stamp directors continued until May 2002.

If, during the telephone call, the food stamp director granted permission for the State to participate in the study, a detailed data request letter was sent to the director of the food stamp program or to a contact person in the State agency designated by the food stamp director. Often, the State directors named a data processing staff member as a contact person for the data request, and a copy of the detailed data request letter was sent to this contact person. The letter explained in detail the variables, formats, and time frame of the data we were requesting. In particular, the letter asked for lists of school-age children in families receiving FS/TANF/FDPIR at two times—summer 2001 and December 2001—along with each child's age, zip code, and some type of identification number. States were encouraged to provide the data in whatever form was most convenient for them, as long as it met the needs of the study.

In March 2002, we began working with States that agreed to participate to answer their questions about the data being requested and to determine the best way for the State to transfer the data files. In May, follow-up email and telephone contacts were attempted with States that had agreed to provide data for the study but whose data had not yet been received.

While calling State food stamp directors to request participation in this study, study team members encountered some concerns about the confidentiality of the data and permission to release the data. Because several directors requested written verification of the study from the USDA, the project director requested that the USDA send a verification letter to States, and this was done in February 2002. In addition, in response to specific requests from States, FNS headquarters recontacted regional FNS and asked them to speak with food stamp directors in certain States to confirm the legitimacy and importance of the study. MPR entered into formal data-sharing agreements with three States to obtain data for this study.

During the telephone calls to food stamp program directors, 43 States agreed to provide us with data (Table A.3). Of these States, 37 actually provided the data. Efforts to obtain data from one State were stopped when the costs of obtaining the data were determined; given the project budget, these costs were judged to be too high.[71]

Seven States refused to provide administrative records data. One refused because of concerns about the privacy of individuals receiving food stamp benefits in the State. Another State had recently suffered a disaster and had to put an emergency program in place, which left that State with inadequate resources to comply with the data request. The remaining States also cited a lack of time and resources to produce the data requested for the study. The results of our data request in the 50 States and the District of Columbia are summarized below:

TABLE A.3
STATE RESPONSES TO ADMINISTRATIVE DATA REQUEST

Number of States That Have:	
Agreed to Participate and Provided Data	37
Agreed to Participate but Did Not Provide Data	6
Agreed to Participate but Only at a High Cost	1
Refused to Participate	7

The overall State response rate was 72 percent. Among districts using direct certification in our sample, 77 percent were in States that have provided data and as a result were used in the analyses based on this data source. Finally, among students enrolled in direct certification districts nationally, an estimated 83 percent were in States that provided administrative data.

[71]The food stamp director in this state initially agreed to participate if MPR would pay a contractor to extract the data and transfer the data file. The estimate of these costs turned out to be higher than had been anticipated (and budgeted) for the collection of administrative data from this state. Since few districts in this state were in the study sample, we decided to abandon efforts to collect the administrative records data from the state.