# Appendix A
Sample Design and Analysis Weights

This study analyzes data from a nationally representative sample of local food stamp offices and households served by those offices. To obtain these data, the design used a two-stage cluster sampling approach. The first stage led to the selection of a national probability sample of local food stamp offices. Within the sampled offices, second stage samples were drawn and data were collected in two domains: policies and practices, and participation (figure A.1). Data on *policies and practices* came from surveys of local office supervisors and caseworkers, and from unobtrusive observations in and around the offices.[1] Data on *participation* came from a survey of eligible nonparticipants living in the catchment areas of the sampled offices who did not apply for food stamp benefits in the month prior to the study, and from surveys and case record abstractions of five groups of applicants and participants in June 2000: approved applicants, denied applicants, approved recertifications, case closures, and TANF diversions.[2] (This last category could overlap with any of the other participant groups.)

Figure A.1—Data collection
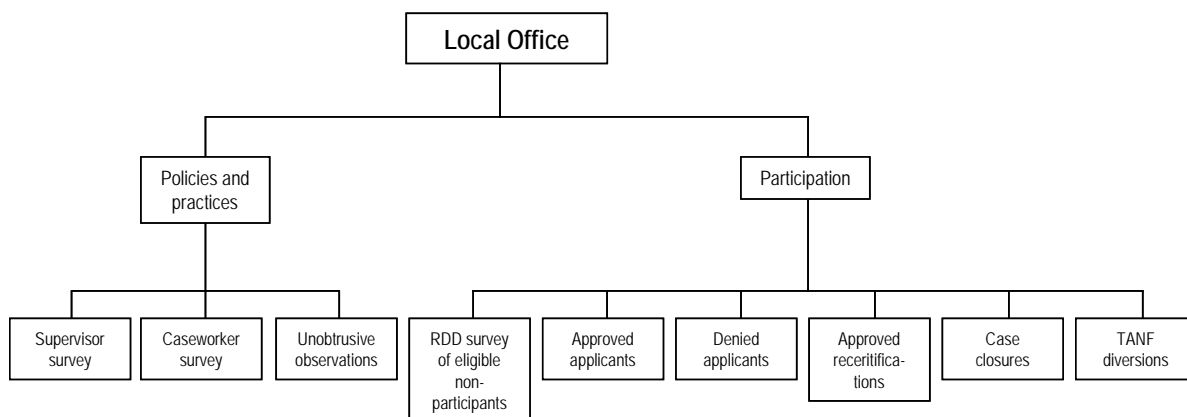
The sampling objectives were to achieve national representation, to encompass substantial variation in administrative practices both between States and within States, and to support office-level analyses of the effects of administrative practices on caseload entry and exit. The sampling plan was designed to ensure that most States (among the 48 in the continental US and the District of Columbia) were represented. Maximizing the representation of States was deemed an important consideration, because welfare reforms under PRWORA created distinctive administrative regimes in each State, which may have had important implications for food stamp participation. Appropriate weights were used to make the household samples represent the universe of food stamp-eligible households in its entirety:

---

[1]   These policies and practices are described in Chapter 7 and Gabor et al. (2003).

[2]   The eligible nonparticipant survey was conducted between February and June 2001. Households were neither current FSP recipients nor applicants for benefits in the month prior to the survey. We assume that they are similar to eligible nonparticipant households in June 2000.

households that were not participating and did not apply in a given month, approved and denied applicants, ongoing cases in recertification months, and ongoing cases in interim months.

The following sections in this appendix describe the two-stage sample and the construction of the weights used in the analyses:

- The first stage sample (local food stamp offices)
- The second stage sample
    - Caseworker supervisors
    - Caseworkers
    - New applicants, recertifying cases, closed cases
    - TANF diverted cases
    - Expansion sample of ongoing cases in interim months
    - Eligible nonparticipants
- Office, supervisor, and caseworker weights
- Participant household weights
- Eligible nonparticipant household weights

First-Stage Sample: Local Food Stamp Offices

In the first stage sampling, the universe of 3,359 local food stamp offices[3] was stratified by FNS region, and State within region to maximize the diversity of the distribution of sampled offices throughout the various States. Within States, sub-stratification was based on degree of urbanicity, using the OMB definition of Metropolitan Statistical Areas (MSA) to define metropolitan *versus* nonmetropolitan areas. In those instances where the expected number of office selections for an individual State fell below one, neighboring States were placed in frame proximity in order to assure the proportionate representation of the first stage sample with respect to groups of neighboring States. This approach helped ensure that most States had at least one sample office.

Probability-proportional-to-size (PPS) sampling was used to draw a sample of 120 local food stamp offices from the strata, with caseload as the measure of size.[4] Using this methodology, larger offices within a stratum had a greater probability of being selected than smaller offices. The sampled offices were located in 40 States and the District of Columbia. All selected States, with the exception of New York State, agreed to participate in the research study. New York declined to participate due to a pending lawsuit in New York City concerning access to the Food Stamp Program that was scheduled

---

[3] A local office is defined as the smallest geographic jurisdiction within which food stamps are administered to all segments of the program population. Thus, within any geographic area, separate sites that served discrete segments of the caseload (for example, cash assistance cases or the elderly) were combined to form one office. We excluded the 430 offices with monthly caseloads below 150 from the sampling frame because of the difficulties in obtaining an adequate sample of applicants. These small offices accounted for only 0.44 percent of the total food stamp caseload.

[4] In some States caseload information was not available by office, and summary caseload data by county was used instead. For those counties that were selected, office and caseload lists were then obtained, and specific offices chosen.

for trial during the data collection period. The final research sample included 109 local food stamp offices, located in 39 States and the District of Columbia.[5]

Second Stage Sample: Caseworker Supervisors

The supervisor survey was designed to collect information on office policies that may affect access to the Food Stamp Program. These policies included those regulating food stamp intake, application processing, and ongoing case maintenance for all different types of households. The objective of the data collection effort was to complete one survey in each sampled office. The level of worker specialization determined the number of supervisors interviewed in each office.

In offices with generic workers, one supervisor was able to answer all the survey questions. In offices in which the caseworkers did not have generic roles and supervisors were therefore specialized based on the types of clients their workers served (for example, households receiving TANF, or elderly cases) or by the portion of the application and eligibility process for which their workers were responsible (for example, intake *versus* ongoing cases), multiple supervisors were interviewed, each about his or her areas of specialization.

Supervisors were selected for interviewing based on their responsibilities. When multiple supervisors had the same responsibilities, the supervisor designated by the office manager as most knowledgeable was interviewed. If the office manager did not designate a specific individual, then the supervisor who had been working at the office the longest was interviewed. A total of 201 supervisors was selected for interviews.

Second Stage Sample: Food Stamp Office Caseworkers

The caseworker survey was designed to collect information regarding caseworker practices. The goal was to obtain two caseworker responses for each set of questions covering particular types of cases. (This was not possible in small offices that did not have two caseworkers handling particular types of cases.) As with the supervisors, the level of worker specialization determined the number of workers interviewed in each office

Caseworkers hired after April 1, 2000 were excluded from the sample frame because of their limited experience in the office. In addition, they may have undergone a training period and not worked a full caseload for very long. All other caseworkers were included in the sampling frame. The list of caseworkers for each office was divided into mutually exclusive, collectively exhaustive groups defined by responsibility. Caseworkers were then randomly selected from each group. A total of 509 caseworkers, an average of 5 workers per office, were sampled.

---

[5] Three of the offices initially sampled were determined to be specialized offices—processing initial applications only, or serving only elderly/persons with disabilities or institutional clients—and therefore ineligible for the study. They were replaced with three additional offices. See table 2.1 for summary characteristics of the sampled offices.

Second-Stage Sample:
New Applicants, Recertified Cases, Closed Cases

The second stage participant household samples were drawn directly from pools of new applicants, reapproved cases, and case closures within the sampled local food stamp offices in June 2000. The universe for these food stamp cases included:

- *New applicants:* households who filed a food stamp application in June 2000. These households included two subcategories of interest: those whose application was subsequently approved for assistance; and those whose application was not approved, because it was denied by the agency (for failure to meet either a circumstantial or procedural requirement of initial eligibility) or was voluntarily withdrawn by the client prior to an agency decision.

- *Recertified cases:* households whose food stamp certification expired in June 2000 and were reapproved for continuing benefits.

- *Closed cases:* households whose food stamp benefits ended in June 2000. These included cases that reached the end-month of their certification period and were not approved for continued assistance, and cases that were closed by the agency in the midst of a certification period (interim case closures)—either for failure to comply with some participation requirement such as ABAWD work requirements or monthly reporting, because the household voluntarily withdrew from the program, or because income and/or assets exceeded eligibility limits.

In order to have analytic samples of sufficient size to answer the research questions, a target initial sample size of 2,400 applicants and recipients was needed. The sample allocation relied on stratified sampling within each office to obtain samples of the domains of interest. The allocation oversampled the two new applicant domains in order to have a sufficient number for analysis purposes. Table A.1 shows the sampling plan and the target initial sample sizes by type.

Table A.1—Initial sampling plan

Domain	Target initial sample size	Actual initial sample size
New applicants—approved	800	882
New applicants—not approved	400	400
Recertifying applicants—approved	600	652
Closures	600	821
Total	2,400	2,755

The sampling of four different domains per office required careful control over the sample selection equations. For each domain we first computed an approximate overall sampling fraction. Given the probability of selection of the office, we divided the overall sampling fraction for each domain by the office selection probability to obtain a within-office sampling fraction for each domain. The

reciprocal of this within-office sampling faction was used as a sampling interval to draw a systematic random sample of applicants from a given domain. The final sample comprised an average of 20 to 25 households in each of the 109 offices.

The samples of new applicants, recertified cases, and closures in June 2000 were drawn from lists provided by the sampled States (or counties). In some instances, due to the structure of a State's management information system, the State was unable to provide files conforming exactly to the required specifications. We anticipated, in some States, that the files would contain some ineligible households. In these cases, we increased sampling rates in order to obtain the desired number of cases in the final sample. The actual initial sample sizes are also shown in table A.1.

Second Stage Sample: TANF Diversions

TANF-diverted applicants, defined as households who entered the TANF application process and were diverted before completing the application process, are a group of particular concern to policy makers to the extent that they may not receive food stamp benefits for which they are eligible. Three types of TANF diversion policies are commonly used to encourage families to avoid becoming TANF recipients:

- *Lump-sum cash benefits* provide a cash payment to help families with short-term emergency needs.

- *Applicant job search* requires some TANF applicants to begin to look for work as a condition of eligibility.

- *Referral to alternative assistance sources*, where caseworkers encourage households to look for other sources of assistance before filing a TANF application.

The TANF-diverted applicants sampled for the current study were those that received lump-sum cash benefits in June 2000. The sample was restricted to this group of households as management information systems in the sampled States did not keep track of applicants assigned to job search or those referred to other assistance sources.

In June 2000, eighteen of the sampled States reported that they had lump-sum diversion policies in effect. Seven States reported that no one in any of the sampled offices received a payment in June 2000 and one other State could not provide any information on payments. The remaining ten States provided lists of households in the sampled offices that received cash payments. The 179 households listed comprised the TANF-diverted sample and represent the universe of households who were diverted from applying for TANF by the use of lump-sum cash payments in the sampled offices in June 2000. These cases were removed from the lists of new applicants, recertified cases, and closed cases before the respective samples were drawn.

Second Stage Sample: Expansion Sample

The samples described above were not sufficient to represent the entire active caseload in June 2000, for use with multivariate models involving interim closures (Appendix D) and in the descriptive analysis of households leaving the FSP (Chapter 6). They cover cases that were in their first month of receipt (approved applicants), in their last month of receipt (closures), or in the final month of a

recertification period that is not the last month of receipt (approved recertifications). Omitted is the great majority of active cases in *ongoing interim months*, for whom no case action was taken in June 2000. These cases must be accounted for correctly in order to be able to represent the full caseload by the sampled cases.

This gap was filled by noting that in a steady state, the cohort of cases with six-month recertification periods ending in June 2000 looked very similar to five other cohorts of cases with six-month recertification periods that ended in July, August, September, October, and November 2000. The first cohort was sampled for this study, although the other five cohorts were not. We could represent the unsampled cases, therefore, by "expanding" the sample that we already have, creating five additional observations for every one we drew. (In general, we expanded by a factor of (L–1), where *L* is the certification length.)

This solution needs to be elaborated upon in three regards. First, while this expansion provides representation of cases which will reach their next *recertification* in July, August, etc., it does not represent cases that will experience *interim closures* in that period. To cover these cases as well, we similarly expanded the sample of interim closures, by a factor of (E–1), where *E* is the time elapsed since the last certification or recertification. For example, for a cohort of cases that was recertified in March and closed in June, we assumed that there were two similar cohorts of cases, recertified in April and May, respectively, that would experience interim closures in July and August, respectively. We expanded the sampled cohort by a factor of two, to represent the two cohorts that were not sampled.

The second issue was the need to avoid double counting of initial months of a spell. The expansion sample generated from *cases in their first certification period* included the first month of that period; but those months were, in fact, already represented by approved applications. It is not generally possible to tell if a particular sampled case is in its first or subsequent certification period. Therefore Month 1 of the expansion samples was given a relative weight of (R/(R+A)), where R was the estimated number of cases coming up for recertification in June 2000, and A was the estimated number of approved initial applications in June.

The final issue was the need to determine the length of the certification period for each sampled case, in order to know how many replicates it should contribute to the expansion sample. This turned out to be surprisingly difficult. The *case records* of a substantial proportion of closed and recertified cases did not include the date of the last certification, and many of the dates that were included were implausible (e.g. in the same month as the current certification, or many years previously).[6] A second data source that was considered was the *usual certification period length reported by the office supervisor* for cases of each type (e.g. all members elderly and disabled, TANF cases with earnings, ABAWDs subject to time limits, and so on). This approach was also unsatisfactory because the supervisor reports contained numerous anomalies, such as cases with all elderly and members with disabilities being assigned 3-month certification periods while TANF cases in the same office had 12-month certification periods, or cases with earnings being assigned 24-month certification periods. While the central tendencies of the reported certification lengths were reasonable, there was concern

---

[6]   The case record abstraction form included an item that was to record the start date of the certification period ending June 2000 (or the most recent certification, for cases closed in June 2000). This item was missing for 21 percent of cases, and represented a date less than one month previous or more than 24 months previous to the June recertification date for an additional 17 percent of cases.

that random inaccuracies could lead to biases in the composition of the expansion sample. The chosen solution was to determine certification lengths from an external source, namely the FY 2000 FSP Quality Control Database. For each State in the study sample, the frequency distribution of certification lengths was calculated for each of ten case profiles.[7] For example, in Alabama, 61 percent of active cases with both TANF and earnings had certification lengths of 12 months, another 13 percent had certification lengths of 6 months, etc. Certification lengths were assigned to recertifying cases by drawing probabilistically from these frequencies, given the State and the case characteristics.

To determine the number of replicates for the much smaller expansion sample generated from interim closures, no alternative existed to compare the recorded date of last recertification with June 2000. When this information was missing or implausible, elapsed months since last recertification was imputed probabilistically, based on the certification length appropriate to the type of case and the observed distribution of time elapsed for other interim closure cases with the same certification length. For example, if a case was assigned a six-month certification length based on its State and characteristics, and among cases with six-month certification periods that experienced interim closures 40 percent experienced the closure in the second interim month, then that case had a 40 percent chance of being assigned a second interim month closure.

The characteristics of the expansion sample households are similar to the characteristics of similar households in the food stamp Quality Control data (table A.2)—evidence of the validity of the procedures used to create the expansion sample.

Second Stage Sample: Eligible Non-Participants

Eligible nonparticipants are households who are circumstantially eligible for the Food Stamp Program but are not participating in it. While Food Stamp Program applicants and participants in an office can be enumerated, no list or sampling frame exists for eligible nonparticipants. As described in Chapter 2, the sample frame was created from a random-digit-dialing telephone survey of the entire population in the catchment areas around the 109 sampled offices. Catchment areas were defined by the telephone prefixes (or zip codes, when telephone numbers were unavailable) of the food stamp new and recertifying applicant households that were included in the applicant and participant sampling frames discussed above. The use of a random-digit-dial sampling approach limits the sample to those with working telephones. Adjustments were made to correct for this limitation during creation of sample weights (see page A-13).

---

[7]   The ten household types were child only, all adults elderly and/or disabled but with earnings, all adults elderly and/or disabled with no earnings, ABAWD-like with children, ABAWD-like without children, other non-TANF cases with earnings, other non-TANF cases with no earnings but government benefits, other non-TANF cases with neither earnings nor government benefits, other TANF cases with earnings, and other TANF cases without earnings. It is not possible to determine with certainty whether a case with given characteristics would have been subjected to ABAWD time limits in a given office in June 2000. Offices differed in how they exempted cases based on presence of dependent children, employment status, and other factors. We identify "ABAWD-like" cases as childless households containing at least one able-bodied adult aged 18 to 50, and households with dependent children containing at least two able-bodied adults aged 18 to 50.

Table A.2—Sample Characteristics: Food stamp access study vs. June 2000 quality control data

Characteristics	Food stamp cases in interim months, June 2000[a]		Active food stamp cases, June 2000[b]	
	Food stamp access study data	FSP quality control data	Food stamp access study data	FSP quality control data
Characteristics of household head				
Male	22.5	22.8	22.5	21.3
Female	77.5	77.2	77.5	78.7
Race of household head				
White	57.9	45.0	56.5	44.7
Black	24.5	34.3	25.2	34.9
Hispanic	13.3	17.0	14.2	16.6
Other	4.2	3.8	4.1	3.8
Citizenship of household head				
US citizen	95.2	91.4	94.9	91.9
Non-citizen	4.8	8.6	5.1	8.1
Household composition				
Single adult without children	44.5	43.0	41.5	40.6
Multiple adults without children	5.0	5.6	5.0	5.4
Single adult with children	37.2	34.5	38.3	37.1
Multiple adults with children	12.2	12.0	14.0	12.4
Children only	1.2	4.8	1.2	4.6
Elderly in household	15.8	22.9	14.1	20.4
Income sources				
Earned income	23.3	26.2	26.6	28.7
Cash assistance (TANF or GA)	23.5	33.9	27.9	30.1
SSI	36.5	34.7	38.4	31.6
Social security	23.6	27.9	26.1	25.1
Income as a percent of poverty				
0-50	41.0	29.1	44.8	31.8
>50 –100	45.1	58.9	43.0	56.0
>100-130	10.4	10.5	9.3	11.0
>130	2.8	1.4	2.9	1.2

a   Interim months include all active cases except those in the first month of a certification or recertification in June 2000. For the food stamp access study data, this includes the expansion sample and the interim closure sample (who received benefits in June 2000).

b   Active cases are those that received food stamp benefits in June 2000, including approved applicants, ongoing cases, and terminating cases.

A list-assisted RDD sample of 72,711 telephone numbers was drawn using the GENESYS sampling system, which helped ensure that a high percentage of the sampled numbers belonged to residential households. Once a household was reached, a short screening questionnaire was administered to determine whether the household was apparently eligible for food stamps, but not currently receiving benefits. Households were screened out if they were current FSP participants or had applied for benefits in the prior month. Households were deemed to be presumptively eligible for food stamps if:

- Their gross household income was no more than 130 percent of the federal poverty level;

- Their financial assets were less than $3,000 if the household included an elderly member and less than $2,000 if there were no elderly members of the household; and

- All vehicles owned were at least five years old, unless they were used for business or to transport disabled persons.

This set of screening questions has been shown to be quite accurate in predicting FSP eligibility (McConnell, 1997).[8] All respondents included in the complete survey responded that their total income was less than the threshold amount. The main body of the survey collected more detailed information on income, including earnings and receipt of various types of unearned income. Using these responses, a second measure of household income was calculated and, as expected, some households appeared to have incomes above 130 percent of the poverty level. We excluded households whose reported detailed incomes exceeded 200 percent of the federal poverty level, though we retained those with incomes between 130 percent and 200 percent. Even if we excluded the latter group, we would still not have a true measure of FSP eligibility. The survey did not collect all the information necessary for a full determination of eligibility, and undoubtedly not all the reported information is accurate, as we did not require proof, such as wage receipts for verification of earnings. In addition, some households that were excluded on the income screening question would have provided detailed income information that resulted in a measure of total income equal to or less than 130 percent of poverty. In order to make certain that survey results were not affected by this aspect of the sampling methodology, all analyses reported in the eligible nonparticipant report (Bartlett and Burstein, 2004) were also conducted on a sample that excluded households with incomes greater than 130 percent of the poverty level. No substantial differences were noted in the results between the two approaches and thus the tables present analyses using the full sample of 1,220 households.

Office, Supervisor, and Caseworker Weights

First-stage office weights were calculated for the final sample of 109 offices. The base sampling office weight equals the reciprocal of the probability of selection of the office. (When the office had been chosen within a county, the base sample weight associated with the selection of the county was multiplied by the inverse of the probability of selection for the office within the county.) Office nonresponse was taken into account at this stage.[9] For each office in the sampling frame, a set of

---

[8]   The screening procedures did not take into account changes in eligibility rules instituted with PRWORA (e.g., immigrant status).

[9]   The nonresponse adjustments made to the sampling weights accounted for the nonparticipation of New York State, mentioned above. This involved increasing the weights of offices in the Northeast and Mid-Atlantic regions that were similar in size and degree of urbanicity to the New York State offices. All the selected offices in New York City were large, urban offices, and thus the weights of other large, urban offices were increased. To the extent that offices in New York City are similar to other large urban offices in the Northeast and Mid-Atlantic along the dimensions measured in the study, bias will be minimized. If food stamp policies and practices in New York City are very different from other large, urban offices, the study will not accurately represent the practices in New York. The other four offices in New York were a mix of small, medium, and large offices, and thus the weights of similar offices in the Northeast and Mid-Atlantic were appropriately increased.

ratio-adjustment cells was formed by three variables: FNS region (five categories, with New England and the Mid-Atlantic regions combined and the Western and Mountain Plains regions combined); MSA vs. non-MSA status; and office size (four categories).[10] The weights of the offices that participated in the study were then ratio-adjusted so that the weighted total number of offices was in agreement with the known totals from the sampling frame.

A *caseload-adjusted* office weight was also created by multiplying the office weights by the office caseloads. These could be used to determine the fraction of the national caseload subject to particular policies (in contrast to the fraction of offices implementing particular policies). These weights were ratio-adjusted using office caseload to create the ratio-adjustment cells. A final adjustment was made to these weights so that the weighted caseload by region was in agreement with total caseload for November 1999, the month used for sampling local offices.[11]

The caseload-adjusted office-level weights were used directly in analyzing office-level data in the description of policies and practices in Gabor et al. (2003) and in Chapter 7 of this report. The supervisor survey was designed so that a single supervisor answered each item even if multiple supervisors were interviewed. Hence the weights for this survey were simply the caseload-adjusted office weights. Multiple caseworkers in a office were asked the same questions about procedures (for example, "If someone comes in at least 30 minutes late for their food stamp appointment with you, do you usually reschedule their appointment for that same day, or reschedule the appointment for another day, or automatically deny the application?") and could respond differently. Their responses were given equal proportional weight, adding up to the office weight—that is, the office weight was divided up among the caseworkers in that office answering each question.

Office weights (not caseload-adjusted) were also a component of the weights used to analyze data on individual households. The household-level weights were conceptually equal to the office-level weights divided by the probability that the household would be included in the sample once the office had been selected. These weights were intended to add up to the corresponding national populations for each group. This calculation was done separately by analysis sample, as described below.

Participant Household Weights

Participant households were sampled from lists provided by the 109 sites in five categories, corresponding to events occurring in June 2000:

- approved applications
- denied applications
- approved recertifications
- case closures
- TANF diversions

---

[10] Size categories were caseloads of 1,651 and under, caseloads of 1,652 to 4,261, caseloads of 4,262 to 7,834, and caseloads of 7,835 and over.

[11] The total caseload in the sample frame was 7.29 million, compared with an actual total of 7.4 million based on more accurate FNS data. The FNS total was available by State. We aggregated this total by region (using the five categories) and scaled the caseload weights to get the correct totals.

In general, the sample comprised 6 to 8 approved applications, 3 to 5 denied applications, 5 to 6 approved recertifications, and 5 to 6 case closures from each site.[12] In Indiana, Tennessee, and Ohio, the lists provided by States did not conform to our requested specifications. Cases in offices from those States were therefore oversampled, so that the number of completes would be adequate. Also, three offices in Ohio were unable to provide any lists of case closures. It was not possible to obtain closure samples for the Ohio offices, so these sites were dropped from the corresponding analyses. All identified TANF diversions in the sample were included—a total of 179 households.

In principle the weights assigned to each household were equal to the size of the list divided by the number of sampled cases in each category and site. Several adjustments to this procedure were needed, however, following the steps outlined below.

Step 1: Due to the inability of some sites to provide separate listings of approved and denied applications, occasionally cases in these two categories could not be distinguished. In these instances the categories for new applications were collapsed within an office. In addition, other offices could not provide separate lists of denied applications and case closures. These categories were similarly collapsed. This type of adjustment occurred in 13 offices.

Step 2: Once the proper adjustments to the sample categories were made, base sampling weights for the first four categories were constructed using the actual numbers sampled and numbers in the lists provided by each office for each category. If s_{hj} is the number of cases sampled and p_{hj} is the number of cases on the case list in site h and category j, then the base sampling weight $BSW1_{hj} = p_{hj} / s_{hj}$.

Alternate procedures were required in the Kansas office, which was unable to provide a list of case closures. In this instance, samples were physically drawn from the file cabinets, sampling at every 62 inches. The Kansas office estimated an average case file is one inch thick, so $BSW1$ for the case closures in Kansas is 62.

Step 3: As noted above, all identified TANF-diverted cases were selected into the sample. (Only a few offices practiced TANF diversion.) These households could appear on any of the other four lists, as they could be approved applicants, denied applicants, and so on. Their identity as a TANF diversion trumped their other identity with regard to sampling. Since these households were selected with certainty, $BSW1$ for these households equals 1.

Step 4: The goal was to have the sample drawn represent the total number of applicants, recertifications, closed cases and TANF-diverted cases in the country for June 2000. $BSW2_{hj}$ was defined as $BSW1_{hj}$ multiplied by the caseload weight for each office described in section 1. If completed data was obtained on all of the sampled cases, this was the analysis weight used. However, abstraction data for a few of the sampled cases did not exist, so a non-response adjustment was made to $BSW2$. If a is the number of cases sampled in an office, and b is the number of cases with complete data, then the final analysis weight $W_{hj} = BSW2_{hj} \times (a/b)$. The non-response adjustments generated

---

[12]   The exact numbers drawn varied from office to office for two reasons. First, the desired sample size per office in each category was not an integer, so in some offices the number drawn was the next lowest integer and in others the next highest integer. Second, after New York State declined to participate in the study, a supplemental sample was drawn to replace it, leading to additional cases being taken from offices whose samples had already been drawn.

weights outside of the median plus 6 times the interquartile range for three denied applicants, who consequently had their weights trimmed.

Once the samples were drawn, the record abstraction sometimes found that a case was of a different type than its sample type (e.g. a case that was on the list of approved applicants might actually have been an approved recertification). Such cases were reclassified for the purpose of analysis, but their weights still reflected the basis on which they were selected into the sample.

Some groups of participant households were also surveyed, in addition to having their case records abstracted: approved applicants, denied applicants, incomplete recertifications, and TANF diversions. The survey weights were the same as the abstraction weights, with an additional factor representing survey nonresponse.

Eligible Nonparticipant Household Weights

Eligible nonparticipants were identified in the random digit dialing (RDD) survey. For each site h, a list was constructed of telephone exchanges of households that would be served by that office (based on the actual telephone exchanges of cases served by that office).[13] From these exchanges, a list of phone numbers was randomly generated.

The final disposition codes for the released phone numbers were as follows:

1: Nonworking, business, or nonresidential
2: Unresolved; unknown whether in category 1, 3, 4, 5, or 6
3: Known to be a household, but not known if eligible for RDD interview (i.e. might be a current FSP participant or ineligible for food stamps)
4: Known household, ineligible for RDD interview
5: Known household, known eligible for RDD interview, but interview incomplete
6. Known household, eligible for RDD interview, interview complete (the analysis sample).

The goal of weighting was to make households in Category 6 represent not only themselves, but also: (a) Category 5; (b) that fraction of Category 3 that was eligible for the interview; and (c) that fraction of Category 2 that was eligible for the interview. This was done as described below.

Step 1: Let N_h = the total number of phone numbers in the site (= number of exchanges × 10,000); m_h = the number of sample phone numbers released; t_h = the number of phone numbers drawn in *all* sample replicates (including those not released); and s_h = the number of phone numbers remaining in all sample replicates after GENESYS ID removed identifiable nonworking and business numbers.[14] The ratio between s_h and t_h was assumed to be the same for all replicates and released replicates. The base sampling weight for each site, BSW_h, was then equal to $(N_h / m_h) \times (s_h / t_h)$. This weight was assigned to each household in Category 6.

---

[13]   Some offices were unable to provide this information. In these cases the offices' catchment area was defined in terms of covered zipcodes—either based on case records or as stated by office staff—and then the corresponding telephone exchanges were determined.

[14]   GENESYS ID only captures a portion of these numbers. For example, it could identify the Abt main switchboard number (492-7100), but not the individual office phone numbers at Abt (e.g. 349-2796).

Step 2: For each of the completes, the number of voice-use lines in the household was known from the survey (or randomly imputed when missing, based on the distribution of this variable in the site). The average number of voice-use lines per eligible complete in each site was calculated as v_h, and the baseweight was adjusted for this factor: $BSW\_P_h = BSW_h / v_h$.[15]

Step 3: Three nonresponse adjustments (NRAs) were then applied to take account of Categories 2, 3, and 5 above. Let Nj be the number of phone numbers in Category j in a site. In each site,

- The completes were scaled up so they also represent the known eligibles who did not complete: NRA1 = (N5+N6) / N6 = known eligibles / eligible completes.

- The known eligibles were scaled up further so they also represent households of unknown eligibility—under the assumption that the same proportion of Category 3 are eligible as among Categories 4, 5, and 6 (for whom eligibility is known):

    NRA2 = {(N5+N6) + [N3 × (N5+N6) / (N4+N5+N6)]} / (N5+N6)
    = {known eligibles + [households, eligibility unknown
    × (known eligibles / households, eligibility known)]} / known eligibles

- The numerator of NRA2 corresponds to total estimated eligibles among phone numbers known to be residential. For the final nonresponse adjustment, the count was scaled up so it also included an appropriate fraction of those phone numbers for which it was not known whether they were nonworking, business, nonresidential, or households. This fraction, the estimated proportion of unresolved numbers that are residential, was taken to be 0.27.[16]

    NRA3 = [{(N5+N6) + [N3 ×(N5+N6)/(N4+N5+N6)]} + ((N5+N6)/ (N4+N5+N6))
    × N2 × 0.27}/ {(N5+N6) + [N3 ×(N5+N6)/(N4+N5+N6)]}
    = {total estimated eligibles, residential status known
    + (known eligibles/ households, eligibility known) × unresolveds × 0.27}
    / total estimated eligibles, residential status known

The nonresponse-adjusted base sampling weight applied to the completes in each site is then given by W = BSW_P × NRA1 × NRA2 × NRA3.

Step 4: An adjustment was made to account for households without telephones, which according to the March 2001 CPS comprises 12.4 percent of households under 130 percent of poverty. These households were represented by sampled households that experienced interruptions in phone service.[17] The adjustment was performed by proportionately increasing the weight on that subgroup of households.

---

[15] Recall that the weights are intended to sum to the number of households. Thus if eligible households in a site have on average 1.1 voice-use lines, then the number of their voice-use lines (represented by BSW_h) should be divided by 1.1, to get the number of eligible households per se.

[16] The source for this estimate is the National Immunization Survey (NIS) run by Abt Associates. For the NIS we drew a large sample of unresolved numbers and called local telephone company business offices to determine if the numbers were residential, business, or not assigned.

[17] Sixteen percent of the sample reported that they had experienced an interruption in phone service of a week or more in the past 12 months. For further discussion of this issue, see Frankel et al. (2003).

Step 5: An adjustment was made to account for the exclusion from the sampling frame of telephone exchanges which accounted for only a few active cases in an office.[18] The rule for constructing the sampling frame was as follows. Starting with the set of exchanges that accounted for at least 75 percent of cases in an office, as many exchanges as necessary were added to bring in an additional 10 percent of cases, as long as that addition to the set was no greater in number than the original set. For example, 4 exchanges might comprise 81 percent of the cases, while another 4 exchanges might bring the total to 89 percent of the cases. The final adjustment increased the sum of the weights in each site to include households not in the sample frame of exchanges. Weights were then truncated at the median plus 6 times the interquartile range.

Step 6: The weights constructed up to this point were constant within a catchment area. As a final step, these weights were raked to correspond to national control totals for 2000 from the Current Population Survey (CPS) and the FSP Quality Control (QC) Database, as reported in Cunnyngham (2002).[19] Characteristics that were taken into account in the raking were household type (presence of children, single *versus* multiple adults), presence of elderly, presence of earnings, presence of noncitizens, and household income relative to FPL (0 to 50 percent, 50 to 100 percent, over 100 percent).

A subsample of eligible nonparticipants of special interest in the analysis was *near-applicants*: households that contacted the local FSP office but did not formally apply. As this was a rare event, households were asked about their contacts with the local office in the past year. It could thus be supposed that the sample included 12 times as many near applicants as would have been identified for any one month. When combined with applicants in June 2000 to represent all households contacting the local office in a month, the weights on the nonparticipant near-applicants were therefore divided by 12.[20]

---

[18]   This adjustment was not made for offices whose catchment area was defined by zipcodes.

[19]   The number of eligible nonparticipants is calculated as the difference between the number of eligible households (from the CPS) and the number of participating households (from QC data).

[20]   While the assumption that near-applicants contacted the office only once in 12 months is questionable, it seemed the most reasonable procedure. The weight on near-applicants who responded early in the course of the survey was adjusted further, to account for the fact that data on near-applicant status was initially collected only for the past 6 rather than the past 12 months. To make the adjustment it was assumed that the ratio of non-participants who were near applicants 1 to 6 months ago to those who were near applicants 7 to 12 months ago was the same in the earlier and the later segments of the sample; this ratio could be calculated for the later respondents.