APPENDIX A

SAMPLE DESIGN, SAMPLING WEIGHTS, AND ESTIMATION

This appendix describes the sampling, weighting, imputation, and estimation for the EFAS Client Survey. The sampling plan was probability-based so study results could be used to make inferences about clients in the national EFAS system. Steps involved in sample design and implementation included (1) definition of the target population, (2) construction of the sampling frame, (3) specification of sample selection procedures, (4) creation of sampling weights and adjustments for nonresponse, and (5) imputation for missing data. We also describe the methods used to estimate population totals of kitchen and pantry clients.

A. TARGET POPULATION

The target population for any survey is the entire set of population units for which the survey data will be used to make inferences. For the EFAS Client Survey, conducted in 2001, the target population was clients (households or people) who received food during the survey's data collection period from a food pantry or an emergency kitchen in the contiguous United States. "Client" was defined differently for pantries and for kitchens.

Food Pantries. The target population for the food pantry portion of the client survey included all households with at least one adult, 18 or older, receiving food packages from a food pantry, on or off site, in the contiguous United States. As described in Chapter II, the client survey used a liberal definition for "housing unit" that encompassed the types of places where homeless people reside, as well as other housing units.

Emergency Kitchens. The target population for emergency kitchens included all adults, 18 or older, receiving meals from an emergency kitchen, on or off site, in the contiguous United States. We excluded facilities that distribute food under the Nutrition Services Incentive Program (NSIP)[1] and the Child and Adult Care Feeding Program (CACFP). We also excluded

[1] Formerly known as the Nutrition Program for the Elderly (NPE).

facilities for which the meals were incidental to other activities, such as substance abuse treatment facilities, summer camps, Kids' Cafés™ (an after-school feeding program sponsored by America's Second Harvest), and senior day care facilities. Similarly, we excluded kitchens co-located with shelters serving food only to residents because distributing food is secondary to providing shelter.

B. THE SAMPLING FRAME

The sampling frame for any survey is the list or other mechanism used to enumerate target population members. The EFAS Provider Survey, conducted in 2000, served as the basis for the frame for primary sampling units (PSUs) and facility selection (Ohls et al. 2001).

PSU Frame for the Provider Survey. The provider survey used an area frame composed of PSUs containing at least 4,250 people in poverty. These PSUs were nonoverlapping land areas, that in the aggregate, span the contiguous United States. The PSUs were formed as individual counties or groups of counties that met a specified minimum size constraint. Frame building began with the "county-equivalent" records from the Area Resource File (ARF) for the 48 states and the District of Columbia. The phrase, "county-equivalent" was used because of the way ARF treats independent cities (U.S. Department of Health and Human Services 1998). Generally, the ARF combines independent cities with their original counties. Manassas City was combined with Prince William County, for instance. Some relatively large independent cities, however, were treated as county equivalents. Alexandria City is one such county equivalent.

In forming PSUs, and later in sample selection, the number of people living in poverty was used to define the size measure. The ideal measure would have been based on the number of EFAS pantries and kitchens, but this information was unknown. The number of EFAS facilities was not exactly proportional to the number of people living in poverty, leading to some unequal weighting in the EFAS provider sample. The minimum size constraint of 4,250 people was set

to yield at least five eligible responding pantries in each PSU and at least five emergency kitchens in urban PSUs.

Listing Providers Within Sample PSUs for the Provider Survey. For each PSU sampled for the provider survey, a sampling frame was constructed listing the pantries and kitchens in that PSU. In constructing these facility frames, we (1) compiled lists of kitchens and pantries, (2) combined and unduplicated these lists to develop a combined (but still partial) list of providers, and (3) supplemented information from the lists with extensive contacts of local sources in selected counties.

The first step was to assemble several different, partial sample frames on a state or regional basis. America's Second Harvest supplied lists of providers in its network, as extracted from databases maintained by its affiliate food banks. Similar lists were obtained from independent food banks not associated with America's Second Harvest (some of which belong to other networks). In addition, state TEFAP directors provided lists of EFAS providers that distribute TEFAP commodities.

The above lists were supplemented with extensive contacts with local collateral sources to obtain the names and locations of emergency kitchens and pantries. On average, we made 16 collateral contacts in each of the 360 areas we sampled. The types of organizations contacted included public and private social service agencies, libraries, churches, and similar organizations. Names obtained this way were compared with those already obtained for that area. We merged all this information to develop a list of the pantries and kitchens in each PSU, without duplicates.

PSU and Provider Frame for the Client Survey. The client survey was restricted to facilities responding to the provider survey from 60 of the original 360 PSUs. Unlike the provider survey, the client survey frame excluded kitchens co-located with shelters serving food to residents, and kitchens and pantries that served fewer than five clients per day. Because of

these restrictions, three PSUs from the provider survey contained no eligible providers for the client survey.

C. SAMPLE DESIGN

Sample selection for the client survey used a multistage sample design. The PSUs were the counties or a group of counties selected in the provider survey. The secondary sampling unit was the facility sampled for the provider survey. The third stage was a selected block of time within which clients were sampled for interview. At each facility, we selected the final sampling unit, a recipient of a meal or a food package.

Selection of PSUs for the Provider Survey. The provider survey used a multiple frame approach to sample design. From the area frame already described, a probability sample of 360 PSUs was selected for the provider survey, with probability proportional to the number of people living in poverty. Chromy's sequential sample selection procedure (Chromy 1979) was used to select the sample after sorting the area frame by U.S. Census region, metropolitan status, percent minority, and total population (including nonpoor and poor).[2] The controlled ordering of the PSUs produced implicit stratification of each stratum's PSUs, which ensured that sample PSUs were representative. Let $S(i)$ be the size measure associated with the i^{th} PSU. Then the expected relative frequency $n_1(i)$ with which the ith PSU was selected for the provider survey is given by

$$E[n_1(i)] = \frac{360S(i)}{S(+)}$$

[2] Serpentine ordering was used to preserve the similarity of adjacent PSUs in the ordered list.

where $S(+)$ is the sum of the size measures over all PSUs and 360 is the number of PSUs selected.[3]

Selection of Facilities for the Provider Survey. Within each PSU, separate samples of kitchens and pantries were selected with equal probability within each PSU for the provider survey. The sample sizes of pantries and kitchens were allocated to achieve sampling weights across PSUs that were as equal as possible. Constraints were placed on the minimum and maximum PSU sample sizes, however. Each sampled facility, hereafter referred to as a primary facility, was asked to report other facilities in the local area. Any such mentions were verified against our original frame, and, if unlisted, were entered as secondary facilities. Secondary facilities were given the same probability of selection as primary facilities of the same type from that PSU. Across primary and secondary facilities, a total of 3,157 kitchens and 2,532 pantries were selected, yielding completed interviews with 1,517 kitchens and 1,617 pantries.

Subsampling PSUs for the Client Survey. A total of 60 PSUs were selected for the client survey from the 360 PSUs in the provider survey. A PSU was eligible for the client survey if at least one kitchen or pantry responded to the provider survey. On average, about six kitchens and five pantries were sampled from each PSU. The actual number selected varied because not all PSUs contained five responding pantries and six responding kitchens. For clients of pantries and of kitchens, the goal of the sample design was to have equal-size client samples taken from each

[3] Note that some PSUs were large enough to be drawn with certainty and have expected relative frequencies of selection that are greater than one. These PSUs had multiple second-stage sample sizes assigned to them corresponding to the number of "hits." Consequently, the number of unique PSUs was somewhat less than the 360 sampling "hits."

sampled facility of a particular type and for sampled clients to have equal selection probabilities across facilities of that type.[4]

The provider survey had 360 selections or 294 unique PSUs, of which 292 had at least one responding eligible EFAS provider. Moreover, we excluded:

- Small providers that served fewer than five clients per day[5]
- Kitchens co-located with shelters
- Providers that went out of business or entered the EFAS since the 2000 provider survey

Furthermore, we designed a procedure to allow us to partially compensate for the very small number of providers in certain PSUs by increasing the number of providers, especially pantries, available for sampling. Some of the PSUs from the provider survey were geographically adjacent. In cases where two PSUs were adjacent and could logistically operate as one PSU, we collapsed the two PSUs into one. We collapsed 18 sets of PSUs in this manner, resulting in a list of 271 PSUs.

We selected 60 PSUs from this modified list of provider survey PSUs. In selecting the 60 PSUs for the client survey and then subsampling kitchens and pantries from these PSUs, our goal was twofold: (1) to select a sample of PSUs so that the probability of selection is proportional to the total number of the nation's clients served by that PSU, and (2) to select facilities within each PSU so that the probability of selection was proportional to the total number of that PSU's

[4] The original PSUs were selected with probability proportional to size sampling with the number of people in poverty used as the size measure. This size measure was the most appropriate one available, but was not perfectly correlated to the actual number of pantries and kitchens in the PSUs. This resulted in unequal weighting across facilities of a particular type.

[5] This resulted in the exclusion of three additional PSUs from sample consideration.

clients served by that facility. We used the sequential selection procedure that Chromy developed (Chromy 1979). We used the corresponding SAS procedure SURVEYSELECT, which has an option for Chromy's algorithm. We used a sequential sample selection procedure instead of simple random sampling to avoid the possibility of extreme concentrations of the selected sample in a few analytic domains. In selecting the sample, we sorted all PSUs according to whether the PSU contained any kitchens, U.S. Census region, urban status, and the total number of people in poverty. We then selected a PSU with a probability proportional to the number of times the PSU was selected for the provider survey sample. Selecting the PSUs proportional to this measure of size had the effect of bringing in two PSUs with certainty: Los Angeles with three "hits," and Chicago with two "hits." The final sample contained 57 unique PSUs.

Subsampling of Providers for the Client Survey. Many of the PSUs did not have at least six kitchens and at least five pantries. Moreover, 10 PSUs did not have any kitchens, and 3 PSUs did not have any pantries. If the remaining PSUs contained enough kitchens, we would ideally have selected six kitchens from each remaining PSU (50 x 6 = 300). However, as with pantries, that is not the case. Therefore, after allowing for the selection of all providers in PSUs with less than or equal to the required number (five pantries or six kitchens), the next step was to allocate the remaining sample to PSUs with more than five pantries or six kitchens. We allocated the remaining sample proportionally to the size of the PSU, where the size was the sum of the weighted provider sizes for all kitchens or pantries in a PSU. After the allocation was completed, we selected providers within each PSU with probability proportional to size using Chromy's sequential selection algorithm.

During the fielding period, weekly numbers of completed interviews suggested we may not complete the required number of pantry client interviews. To remedy this potential shortfall, we

supplemented the sample with the five pantries that were in the selected PSUs but were not selected as part of the provider sample. In the end, we did not sample pantries but took a census of pantries in the selected PSUs.

Sampling Meals and Time Periods. Lastly, using information collected in the provider survey, we selected a time period for each sampled facility in which sampling and interviewing would take place. This information included (1) the days and hours of operation, (2) the average number of clients per meal and day, and (3) whether the provider used a mobile food van or wagon to distribute food off site. The time period was a particular meal on a particular day for emergency kitchens or a four-hour period on a particular day for food pantries. The sample selection was designed so client visits to a particular type of facility (e.g., a pantry or a kitchen) would have approximately equal probabilities of selection across facilities. We allocated no more than three visits to any one provider to minimize the burden on providers.

Before the beginning of data collection, we verified during client sample frame development that the provider was available for interviewing during the sampled meal or time. If the provider was not available, a substitute meal or time period was selected. Substitution was made under the constraints that weekdays should be replaced by weekdays, weekends should be replaced by weekends, and, if at all possible, the same sample meal or time period would be substituted (see also Appendix B, Data Collection Methods).

Selection of Clients. Before the sampled time period, our interviewers contacted the facility managers of the kitchens or pantries, and let them know the time of our selected visit. Interviewers then visited the site to obtain a random sample of clients during that time period. The interviewers typically completed 7 to 16 client interviews per visit at a facility. The exact number depended on the difference between the number of clients expected during the time period and the number that actually arrived.

A few things should be noted about sampling clients. First, we conducted this study among adults, but the interviewers included children when counting and selecting people. When they were unsure of the selected person's age, they first asked if they were 18 or older before asking them to take part in the interview. If they selected someone under 18, they did not interview the child but did record this information on the client selection form. The client selection form included the following information: (1) the number of clients selected, (2) the number of clients completing the interview, (3) the number of clients refusing the interview, (4) the number of children selected (as ineligibles), (5) the number of clients selected but who were not interviewed due to mental, physical, or other conditions, (6) the total number of clients receiving services during the interview time, and (7) any unusual circumstances. This information allowed us to track interviews and weight the data appropriately. The interviewers returned a client selection form for each visit.

Sample Sizes. Across the 60 sampled PSUs and the associated census of 305 pantries and 300 sampled kitchens, we were able to schedule interviews at 180 pantries and 191 kitchens (see Table A.1). We were unable to schedule visits for a number of reasons: the provider saw fewer than five clients per day, the provider was no longer in business, or the kitchen was co-located with a homeless shelter. For 17 pantries and 7 kitchens, we did not schedule visits because we were unable to reach them to determine eligibility. Of the sampled kitchens, six were listed twice on the sampling frame, and therefore, the final sample size of kitchens was 294. The kitchen and pantry provider response rates are 88 percent and 83 percent (see Table A.2). At cooperating providers, we completed 2,425 interviews with kitchen clients and 2,397

TABLE A.1

FINAL STATUS OF PROVIDERS

Final Status	Reason for Eligibility Determination	Frequency	Percent
Food Pantries			
Eligible, visits scheduled	Eligible respondent	180	59.0
Eligible, visits not scheduled	Refusal	13	4.3
Eligible, visits not scheduled	Field period ended, eligible	12	3.9
Ineligible	No longer operating	48	15.7
Ineligible	Fewer than 5 clients/day	23	7.5
Ineligible	Pantry distributes off-site only	9	3.0
Ineligible	Open on emergency basis or holiday only	3	1.0
Eligibility unknown	Field period ended, eligibility not determined	17	5.6
TOTAL		**305**	**100**
Emergency Kitchens			
Eligible, visits scheduled	Eligible respondent	191	65.0
Eligible, visits not scheduled	Refusal	13	4.4
Eligible, visits not scheduled	Field period ended, eligible	7	2.4
Ineligible	No longer operating	31	10.5
Ineligible	Fewer than 5 clients/day	1	0.3
Ineligible	Kitchen co-located with shelter	40	13.6
Ineligible	Kitchen distributes to individual homes	1	0.3
Ineligible	Open on emergency basis or holiday only	1	0.3
Ineligible	Elderly or youth feeding program	2	0.7
Eligibility unknown	Field period ended, eligibility not determined	7	2.4
TOTAL		**294**	**100**[a]

SOURCE: National Emergency Food Assistance Study Client Survey (2001).

[a]Individual percentages shown do not add to 100 because of rounding.

with pantry clients. The kitchen client and pantry client cooperation rates[6] are 87 percent and 84 percent, respectively. The overall response rates (the product of the provider response rate and the client cooperation rate) for kitchen clients and pantry clients are 77 percent and 70 percent, respectively.

TABLE A.2

PROVIDER RESPONSE RATES

Provider Type	Eligibility Determination Rate[a]	Cooperation Rate[b]	Response Rate
Kitchens	98%	91%	88%
Pantries	94%	88%	83%

[a]Percent of the sample of providers screened for eligibility.
[b]Percent of eligible sampled providers that cooperated with the request to schedule appointments for client interviewing.

TABLE A.3

CLIENT RESPONSE RATES

Provider Type	Provider Response Rate	Client Cooperation Rate	Client Response Rate[a]
Kitchens	88%	87%	77%
Pantries	83%	84%	70%

[a]Product of the provider response rate and the client cooperation rate.

[6]Percent of eligible sampled clients who cooperated with the request for an interview.

D. WEIGHTING

Estimates based on the client survey must account for the survey's complex sample design and for the biasing effects that nonresponse could have. We constructed sampling weights that reflect the differential selection probabilities used to sample EFAS providers across PSUs. The sampling weights for the client survey began with the sampling of PSUs and providers.

Provider Sampling Weight. The providers for the client survey were selected from 60 PSUs. Therefore, the conditional EFAS client PSU sampling weight can be expressed as:

$$CONDPSUWT(i) = \frac{60S_2(+)}{S_2(i)}$$

where $S_2(i)$ is the EFAS client size measure of the ith PSU, and $S_2(+)$ is the sum of the EFAS client size measures across all PSUs. Therefore, the weight for the providers in the EFAS client frame is:

$$FRAMEWT\ (ijk) = EFAS1\ FINAL\ WT\ (ijk) \times COND\ PSUWT\ (i)$$

where *EFAS1 FINAL WT*(ijk) is the provider survey analysis weight for the kth provider of type j from PSU i. Here, $j = 1$ for kitchens and $j = 2$ for pantries.

Next, we calculated the conditional provider weight. Many PSUs had very few providers, and we selected all providers of a given type. In those PSUs, the conditional weight for providers is 1. To avoid a shortfall of pantry client interviews, we supplemented the sample with the remaining pantries in the sampled PSUs. Hence, the conditional provider weight for all pantries is 1. For those PSUs with sufficiently large number of kitchens from which to sample, we first removed those kitchens whose size was so large that they should be sampled with certainty. The certainty kitchens have a provider conditional weight of 1. The remaining

kitchens were selected from the PSU with probability proportional to size. The conditional provider weight for these remaining providers is

$$CONDPROVWT(ik) = \frac{PROVSIZE(i+)}{n(i)\,PROVSIZE(ijk)}$$

where *PROVSIZE*(*i*+) is the sum of the size measures for the noncertainty kitchens in PSU *i, n*(*i*) is the noncertainty sample size taken from PSU *i*, and *PROVSIZE* (*ijk*) is the size measure for the k^{th} provider of provider type $j = 1$ (kitchen) in PSU *i*.

Therefore, the unconditional provider weight for the k^{th} provider of type *j* from PSU *i* is the product of the frame weight and the conditional provider weight or

$$PROVWT(ijk) = FRAMEWT(ijk) \times CONDPROVWT(ijk)$$

At this point, we had to take into account that some of the sampled providers refused to allow for EFAS client interviews. To account for nonresponse, we implemented a weighting class adjustment followed by a poststratification adjustment.

Provider Weighting Class Adjustments. Weighting class adjustments were made by partitioning the sample into groups, called weighting classes, and then adjusting the weights of responding providers within each class so they sum to the weight total for nonrespondents and respondents from that class. Implicit in the weighting class adjustment is the assumption that if the nonrespondents had responded their responses would have been distributed in the same way as the responses of the other responding providers in their class. The client survey weighting classes were defined on the basis of type of provider, urbanicity (metro county and nonmetro county), and size of provider (three level size variable as defined in Ohls et al 2001). Two nonresponse adjustment factors were calculated.

First, we adjusted the sampling weight to account for sampled providers for which eligibility status could not be determined. The first step is to define a response and eligibility indicator. Define ELIGRESP as follows:

ELIGRESP =	1	sampled provider was identified as eligible, visits scheduled
	2	sampled provider was identified as eligible, no visits scheduled
	3	sampled provider was identified as ineligible
	4	eligibility status of the sampled provider was not identified

Note that codes 1, 2, and 3 imply that eligibility status was known, and the case was a respondent for eligibility determination. The eligibility determination adjustment $ADJ_{ed}(c)$ for respondents in weighting class c is defined as follows:

$$ADJ_{ed}(c) = \frac{\sum_{ijk \in c} PROVWT(ijk)}{\sum_{ijk \in c} \delta_{ed}(ijk)\, PROVWT(ijk)}$$

where $\delta_{ed}(ijk)$ is equal to 1 for providers where eligibility was determined and 0 otherwise. Note that eligibility determination for nonrespondents has $\delta_{ed}(ijk) = 0$ and hence an eligibility determination adjustment of 0.

Second, we adjusted for the loss of interviews from providers known to be eligible but refused to allow for EFAS client interviewing. The nonresponse adjustment $ADJ_{nr}(c)$ for respondents in weighting class c is defined as follows:

$$ADJ_{nr}(c) = \frac{\sum_{ijk \in c} PROVWT(ijk)}{\sum_{i \in c} \delta_{nr}(ijk)\, PROVWT(ijk)}$$

where $\delta_{nr}(ijk)$ is equal to 1 for providers who agreed to interviewing and 0 otherwise. Note that nonrespondents have $\delta_{nr}(ijk) = 0$ and hence a nonresponse adjustment of 0. Ineligible providers were defined to have a nonresponse adjustment of 1. The adjustments $ADJ_{ed}(c)$ and $\underline{ADJ}_{nr}(c)$

were then applied to the provider weights to obtain the nonresponse adjusted provider weight *NRADJ_PROVWT*(*cijk*) for the *ijk*th case from the *c*th weighting class as follows:

$$NRADJ\_PROVWT(cijk) = \delta_{ed}(ijk)\, ADJ_{ed}(c) \times \delta_{nr}(ijk)\, ADJ_{nr}(c) \times PROVWT(ijk)$$

Lastly, we implemented a poststratification adjustment, a common technique for adjusting survey data using external data from a sampling frame, census, or survey of higher accuracy. This technique is commonly used to reduce bias because of nonresponse and under- or overcoverage for survey data. For these data, we poststratified to the weight provider totals from the provider survey having excluded those providers ineligible for the client survey. We again created classes and, in this case, the classes were defined on the basis of type of provider and urbanicity. The poststratification adjustment $ADJ_{post}(h)$ for respondents in poststratum *h* is defined as follows:

$$ADJ_{post}(h) = \frac{\sum_{ijk \in h} PS\_PROVWT(ijk)}{\sum_{ijk \in h} NRADJ\_PROVWT(ijk)}$$

where *PS_PROVWT* is the weight from the provider survey, which is the client survey sampling frame. The adjustment ADJ_{post} was then applied to the nonresponse adjusted provider weight to obtain the poststratified provider weight *POST_PROVWT*(*hijk*) for the *ijk*th case from the *h*th poststratum as follows:

$$POST\_PROVWT(hijk) = ADJ_{post}(h) \times NRADJ\_PROVWT(hijk)$$

Client Sampling Weight. To continue toward the development of a client weight, we made a visit-level adjustment to the provider weight. To account for the number of visits to a provider, each provider weight was adjusted by the inverse of the number of visits made resulting in the visit adjusted provider weight:

$$VISIT\_PROVWT(ijk) = \frac{POST\_PROVWT(ijk)}{v}$$

where v is the number of visits made the ijk^{th} provider. The number of visits ranged from one to a maximum of three.

For each visit, the interviewer recorded the total number of clients sampled (including ineligibles and nonrespondents) and the total number of clients served during the interview period (meal or four-hour time period). The conditional client weight is the total number of clients served divided by the total number sampled:

$$COND\_CLIENTWT(ijkl) = \frac{C(ijkl)}{c(ijkl)}$$

where $c(ijkl)$ is the total number of clients sampled, and $C(ijkl)$ is the total number of clients served during the interview period on the l^{th} visit for the ijk^{th} provider. The unconditional client weight is the product of this conditional weight and the visit adjusted provider weight or:

$$CLIENTWT(ijkl) = VISIT\_PROVWT(ijk) \times COND\_CLIENTWT(ijkl)$$

Furthermore, similar to providers, we implemented a nonresponse adjustment. Within the ijk^{th} provider for the l^{th} visit, we calculate an adjustment so the weights of responding clients sum to the weight total for nonrespondents and respondents or:

$$ADJ_{nr}(ijkl) = \frac{\sum_{ijk \in l} CLIENTWT(ijk)}{\sum_{ijk \in l} \delta_{nr}(ijk)\, CLIENWT(ijk)}$$

where $\delta_{ed}(ijkl)$ is equal to 1 for responding clients and 0 otherwise. Note that nonrespondents have $\delta_{ed}(ijkl) = 0$ and a nonresponse adjustment of 0. Ineligible clients are defined to have a nonresponse adjustment of 1. The nonresponse adjusted client weight is the product of this adjustment and the unconditional client weight or:

$$NR\_CLIENTWT(ijkl) = \delta_{nr}(ijkl)\, ADJ_{nr}(ijkl) \times CLIENTWT(ijkl)$$

The next step of the client weighting process moved us toward a weekly weight, which differs for pantry and kitchen clients. At kitchens, we asked the operators how many meals per week they serve. We then multiplied the client weight by the number of meals per week to represent clients served at all meals per week.

$$WK\_CLIENTWT(ijkl) = NR\_CLIENTWT(ijkl) \times m(ijk)$$

where $m(ijk)$ is the number of meals served per week at the ijk^{th} kitchen.

At pantries, we had information on how many people were served on the day of interviewing and the number of days the pantry is open per week. Therefore, we computed a ratio (number of people served that day/number of people served during the interview period) to adjust the weight to represent all of the clients served that day or:

$$ADJ_{day}(ijkl) = \frac{D(ijkl)}{C(ijkl)}$$

where $D(ijkl)$ is the total number of clients served that day on the l^{th} visit for the ijk^{th} provider, and $C(ijkl)$ is the total number of clients served during the interview period (four-hour time period) on the l^{th} visit for the ijk^{th} provider. We then multiplied the day weight by the day adjustment and the number of days open per week, $d(ijk)$, to present clients served per week.

$$WK\_CLIENTWT(ijkl) = NR\_CLIENTWT(ijkl) \times ADJ_{day}(ijkl) \times d(ijk)$$

Lastly, we needed to adjust the weekly client weights to represent unique people. We used the client's response to question A3 (How many different times in the past seven days did you receive groceries from this or any other food pantry?) at pantries, and A7 and A7c (A7: Including the meal that you have just received, how many meals do you expect to receive today?

A7c: Counting today, on how many of the last seven days did you receive one or more meals from this or any other kitchen?) at kitchens. The inverse of the number of times in the last week, $w(ijklm)$, is used to compute the unique pantry client adjustment factor:

$$ADJ_{unq}\left(ijklm\right)=\frac{1}{w\left(ijklm\right)}$$

The inverse of the product of the number of meals today, $d(ijklm)$, and the number of times in the last week, $w(ijklm)$, is used to compute the unique weekly kitchen client adjustment factor:

$$ADJ_{unq}\left(ijklm\right)=\frac{1}{d\left(ijklm\right)}\times\frac{1}{w\left(ijklm\right)}$$

The adjustment ADJ_{unq} is then applied to the week client weight to obtain the unique client week weight $UNQ\_WK\_CLIENTWT(ijklm)$ for the m^{th} client at the ijk^{th} provider on the l^{th} visit as follows:

$$UNQ\_WK\_CLIENTWT\left(ijklm\right)=WK\_CLIENTWT\left(ijkl\right)\times ADJ_{unq}\left(ijklm\right)$$

The unique weekly client weight is the analysis weight for both pantry and kitchen clients.

E. IMPUTATION

There was considerable item nonresponse on two questions asked of pantry clients because of a skip pattern problem in the questionnaire. Because the information from these two questions was required for estimating the total number of clients served, it was necessary to impute the missing items. The two questions were: (1) A4: In the months that you received groceries from a food pantry, how many times per month did you receive them? and (2) A5: In the last year,

during how many months did you receive groceries from this or any other food pantry?[7] These two questions were asked only of pantry clients.

We imputed values for missing items by using a sequential hot deck imputation procedure (Fellegi and Holt 1976; Cox 1980). First, we identified variables to use in the classing and sorting sets. We used cross tabulations to determine which variables were correlated with the two variables to be imputed. We determined that race/ethnicity and employment status were correlated with the two variables. Second, the file was sorted by race/ethnicity, sex, and employment status. Although sex is not correlated with the two variables to be imputed, we considered it a necessary sorting variable for face validity. Third, we constructed imputation cells composed of a two-level collapsed version of race/ethnicity: "white non-Hispanic" and "all others." Within an imputation cell, we imputed values for missing items using actual survey responses from donors with complete data. Moreover, the algorithm imputed pairs of data, that is, both imputed values came from the same donor. We evaluated the distribution of both variables before and after imputation. The imputed values had no appreciable effect on the distributions of either variable.

F. ESTIMATION OF POPULATION TOTALS OF CLIENTS SERVED

We developed weekly estimates of the number of people served by kitchens or pantries. Our estimates were predominantly based on the design-based analysis weight, which was built on the probabilities of selection, the sample design, and corresponding statistical adjustments. Our operational definition of the population excluded clients at small providers, clients at kitchens co-located with a shelter, and providers no longer in operation or new entrants to EFAS.

[7] This second question was not used directly for estimation. However, because imputation for A5 was required, we imputed A4 simultaneously.

The population of interest was all clients served by all kitchens and pantries, regardless of size or situation. Because our operational definition of the population differed from the population of interest, we attempted to compensate for the shortfall with additional adjustments and extrapolations.

The estimation process relied on several data sources to derive estimates of the population total. These included:

- Information from the provider survey conducted in 2000
- Information from the sampling and data collection operations concerning observed numbers of clients served by providers and the providers' days of operation
- Information from the data collection operation on the reason a provider was ineligible for the client survey
- Information from clients related to their frequency and duration of visits to any pantry or any kitchen

Our basic approach to deriving weekly estimates of clients served was to start with the design-based estimates of clients served per week and then apply a number of adjustment factors and extrapolations to arrive at the estimate of the population of inference. However, we believe that error sources existed in extrapolating from the operational definition of the population to the population of interest and that these errors may have been large.

To adjust for the various sources of error, we calculated several adjustment factors. Each adjustment factor was calculated separately, by type, within urban and rural areas. To adjust the design-based weighted total number of clients, we applied each of these adjustment factors to the provider weight. Using this adjusted provider pseudo-weight, we calculated an adjusted client

pseudo-weight and the total number of clients under the analytic population definition.[8] These factors included:

- An adjustment to represent providers not located in the PSU during the provider survey
- An adjustment to represent providers that were no longer in business
- An adjustment factor to represent providers open very infrequently
- An adjustment factor to account for kitchens co-located with shelters

The first adjustment factor accounted for an adjustment made to the total number of providers estimated from the provider survey. The initial sample frame for the provider survey was an incomplete list of providers in the PSUs. To account for the undercoverage, we collected "secondary" sample providers identified during interviewing. Because secondary sample cases were not listed in the original sample frame, the sample weights did not fully reflect these providers. Therefore, an adjustment was added to the total number of kitchens and pantries (see Ohls et al. 2001 for more details). The adjustment factor for kitchen clients was the total number of kitchens with the secondary adjustment divided by the total number of kitchens based on sample weights. The adjustment factor for pantry clients was the total number of pantries with the secondary adjustment divided by the total number of pantries based on sample weights.

The second adjustment factor accounted for providers that were no longer in business. We assumed a steady state existed in the number of EFAS providers. That is, for each provider that went out of business, another provider opened its doors. This adjustment also assumed the distribution of providers remained constant, whether across geographic regions, type of sponsor, and any of a number of provider characteristics. The adjustment factor was the sum of the

[8] After these ratio adjustments, the provider and client weights were no longer design-based

provider weights for responding and out-of-business providers, divided by the sum of the provider weights for responding providers. This adjustment then increased the pseudo-weights of the responding providers to represent the out-of-business providers.

The third adjustment factor accounted for providers that are open only on holidays or for emergencies or provided food or groceries off site only. The adjustment factor was the sum of the provider weights for responding providers, providers open only on holiday or for emergencies, and providers that serve off site only divided by the sum of the provider weight for responding providers. When applied to the pseudo-weight, this adjustment allowed the responding providers to represent those providers that were open infrequently.

The fourth adjustment factor accounted for kitchens co-located with shelters for the homeless. We assumed the kitchens co-located with shelters were similar in size to those not co-located with shelters. The adjustment factor was the sum of the provider weights for responding providers and kitchens co-located with shelters divided by the sum of the provider weights for responding providers. A similar adjustment factor was calculated to adjust for the kitchens co-located with shelters that we excluded from the EFAS client sampling frame. In the case of the frame adjustment factor, we used the final EFAS provider weights to calculate the ratio. These factors, when applied to the pseudo-weight, permitted the responding providers to represent the kitchens co-located with shelters.

(continued)
weights. Therefore, we label them pseudo-weights.

In addition to these adjustment factors, we estimated the total number of clients excluded because of the operational definition of the population.[9] Moreover, if we learned that a sampled provider served fewer than five clients in a day, we considered them to be ineligible.

To estimate the number of clients served by these small providers, we used information from the provider survey. We assumed that each small provider served four people each day it was open. Using the reported number of days open per week from the provider survey (Ohls et al. 2001), we calculated a weighted number of clients from small providers. The number of clients estimated from small providers was added to the total derived from the factor adjustments.

The survey instrument focused on weekly usage of both kitchens and pantries. While it might have seemed reasonable to describe the total number of kitchen clients in a typical week, the patterns of pantry usage led us to estimate a monthly total. To estimate that, we calculated an adjustment factor for each responding pantry client:

Monthly adjustment factor = 4 / number of weeks per month

The number of weeks per month is a constructed variable based on question A5 (In the months that you or another member of your household received groceries from a food pantry, how many times per month did you receive them?). The construction is shown in Table A.4.

[9] To simplify data collection, we did not include in the sampling frame any providers serving fewer than five clients in a day.

TABLE A.4

MONTHLY ADJUSTMENT FACTOR

How Often Client Received Groceries	
Number of Times per Month	Number of Weeks per Month
1	1
2	2
3	3
4	4
5 or greater	4

This adjustment factor assumed, for example, if a client received groceries twice per month, then the client received these groceries in two different weeks. This monthly adjustment factor was applied to the adjusted client pseudo-weight to estimate the total number of different clients served by pantries in a typical month.

The last step in estimating the total number of clients served was to estimate the annual number. We did not derive the annual estimate by simply multiplying the monthly number by 12, which would assume that an entirely new set of clients is served each month, nor did we set the annual estimate equal to the monthly number, which would assume that no new clients are served each month. We constructed the potential range of possible numbers of different clients annually by examining the implications of alternative estimates of turnover in the system, where we defined turnover as the average percentage of the clientele that were "new" each month in the sense of not having used a pantry in the previous 12 months. If, for example, in food pantries, we assume that this turnover rate is only 4 percent per month, this would imply that the annual number of different clients is 18.0 million. That is, the estimate for number of clients served in a month (12.5 million), plus 4 percent of that total (500,000) each month for 11 months (or 5.5 million) (or 12.5 million + 5.5 million = 18.0 million). On the other hand, if, we assume a monthly turnover rate of 8 percent of the caseload, this would imply that the annual number of

different clients is 23.5 million. That is, the estimate for number of clients served in a month (12.5 million), plus 8 percent of that total (1,000,000) each month for 11 months (or 12.5 million + 11.0 million = 23.5 million).

As noted in chapters three and four, the study design limits our ability to precisely measure patterns of use over a year and estimate the total number of households and clients served by food pantries and emergency kitchens during 2001. First, data were collected for a 14-week period rather than for a year's period so the survey data do not reflect seasonal patterns of food pantry usage. Second, while we collected a limited amount of data about clients' use of pantries and kitchens for the previous 12 months, space limitations on the instrument precluded obtaining all the data necessary to fully characterize annual usage. Also, these data may contain considerable measurement error in clients' abilities to accurately report the number of months in the past 12 months that they visited a food pantry or the number of weeks in a row that they had visited one or more kitchens during the past year.