Appendix A
Sampling and Weighting Procedures for the Survey of Sponsors

Sample

The sample universe for the *Family Child Care Homes Legislative Changes Study* consisted of family child care sponsors, family child care homes, and families participating in the CACFP. A nationally representative sample of 20 States was selected, with probability proportional to the size of each State's share of CACFP family child care home reimbursements.[1] Sponsoring agencies within States were also selected with probability proportional to size, based on the number of homes sponsored.[2]

A sample of 311 sponsors was selected within the 20 selected States, comprising a representative sample of all sponsors in the country (n=1,165). Eleven sponsors were determined ineligible to participate in the survey due to having left the CACFP, leaving an effective sample size of 300. Of this number, 268 sponsors participated in the survey, for a response rate of 89.3 percent.

Weighting

For producing population-based estimates of means and proportions of characteristics relating to sponsors, each respondent sponsor received a sampling weight. These weights combined the basic weight reflecting the probability of selection of the sponsor and an adjustment for unit nonresponse. The resulting weighted data yield estimates for all sponsors in the population.

For the selection of sponsors, a sample of sponsors was selected in each of the 20 States selected in the first stage. Therefore, the overall probability of inclusion of a sponsor is the inclusion probability of the State in which the sponsor is located multiplied by the probability of including the sponsor in the sample, given that the State was selected.

Sponsor weights were computed as follows:

1. Let W_i represent the weight for the *i*th selected State. i = 1, 2, 3, 4, ...19, 20. W_i = 1 for States selected with certainty.

2. Let W_{ij} be the weight for the *j*th selected sponsor in the *i*th State. We have

$$W_{ij} = W_i \; W_{j/i}$$

[1] Four States were included with certainty (California, Michigan, Minnesota, and Texas).

[2] Multiple "hits" were permitted, i.e., sponsors could be selected more than once.

where $W_{j/i}$ is the conditional weight of the *j*th sponsor given that the *i*th State has been selected.

We now determine $W_{j/i}$. Let the number of sponsors in the *i*th State be S_i. Let the number selected in the sample be s_i. Let the number of providers belonging to the *j*th sponsor in the *i*th State be P_{ij}.

- In 12 States, all sponsors in the State were included in the sample with certainty. In these States, we have

$$W_{j/i}=1.$$

Therefore, the overall sponsor weight in these States is $W_{ij}= W_i$.

3. The sponsors in the other eight States were selected with probability proportional to the number of providers and **with replacement**. Therefore, the same sponsor can get selected more than once. Let r_{ij} be the number of times ("hits") the *j*th sponsor gets selected in the *i*th State. The weight is therefore

$$W_{j/i}=\frac{r_{ij}\ P_i}{n_i\ P_{ij}}$$

where n_i is the total number of sponsor hits in the *i*th State and $P_i=\sum_{j=1}^{S_i} P_{ij}$ is the total number of providers.

The overall basic sampling weight for the *j*th sponsor in the *i*th State is given by:

$$W_{ij}=W_i\ W_{j/i}.$$

Adjustment for Nonresponse at the Sponsor Level

There is no nonresponse at the State level.

For sponsor nonresponse adjustment, assume that s^*_i sponsors respond to the survey out of the s_i sponsors selected in the *i*th State. Then the nonresponse adjustment to the weights of the responding sponsors is

$$A_i=\frac{\sum_{j=1}^{s_i} W_{ij}}{\sum_{j=1}^{s^*_i} W_{ij}}$$

The nonresponse adjusted conditional weight is given by

$$W^{a}_{j/i} = W_{j/i}\ A_i.$$

The overall nonresponse adjusted basic sampling weight is given by

$$W^{a}_{ij} = W_i\ W^{a}_{j/i} \quad .$$

This weight is used in all percentages and other distributional statistics presented in the report. Standard errors are estimated with adjustment for the multi-stage sample design using SUDAAN software. In addition to the standard error, tables show the unweighted number of observations on which distributional statistics are based.

Nonresponse Bias

The possibility of nonresponse bias—that is, important differences between sample members who respond to the survey and those who do not—deserves consideration in any sample survey. With a response rate of 89.3 percent, we would not expect nonresponse bias to be an important factor for the sponsor survey. Nonetheless, a series of analyses was performed to assess the extent of any bias.

The analysis is necessarily based on those few items of information that are known for the nonresponding as well as the responding sponsors. These include the number of CACFP homes sponsored, the percent of homes that are Tier 1, and geographic region (Northeast, South, Midwest, and West).[3]

The analysis compared the mean or percent for all selected sample members and the mean or percent for those responding to the survey. The difference can be viewed as the extent to which the respondents over- or under-represent the specified characteristics of the original sample. As a guide to the importance of the difference, we use a one sample *t*-test; that is, we compare the mean of the respondents with the mean of the total sample, taking into account the standard error of the mean of the respondents. The data are unweighted in this analysis because sampling weights were not computed for nonrespondents.

Differences between the responding sponsors and those selected for the original sample are very small and not statistically significant, as shown in Exhibit A.1. It is worth noting that the responding sponsors are somewhat larger, on average, than the sample as a whole. This difference contributes to the anomalous contrast between the number of sponsors in the 1995 and 1999 studies, as reported in the text.

[3] The sponsors' number of homes and number of Tier 1 homes for this analysis is based on the lists of sponsors that State agencies provided.

Exhibit A.1
Comparison of Responding Sponsors with Sample Selected

	Respondents	Original Sample	Respondent-Original Difference	Respondent Standard Error	p-value
Mean number of homes per sponsor	357.3	347.4	9.9	33.6	0.77
Mean percent of sponsor's homes that are Tier 1	67.1	68.6	-1.5	1.5	0.31
Percent of sponsors that are in region:					
Northeast	21.6	21.3	0.3	2.5	0.89
Midwest	21.6	20.0	1.6	2.5	0.52
South	28.4	29.7	-1.3	2.8	0.63
West	28.4	29.0	-0.6	2.8	0.81