

Appendix A

Sampling and Weighting Procedures for the Survey of Former Providers

The *Family Child Care Homes Legislative Changes Study* involved several surveys, including surveys of sponsors, current CACFP providers, parents of children currently served by CACFP providers, and former CACFP providers. The surveys of former providers are the data source for the section on “Former CACFP Homes,” including Exhibits 9-11. The sample design for these surveys and the weighting procedures used in the analysis are described below. The sampling and weighting for other surveys are discussed in other reports in this series.

The sample universe for the 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 All selected State agencies agreed to participate in the study and provided lists of the CACFP sponsors in their State. Sponsors were also selected within States with probability proportional to size, based on the number of homes sponsored.²

Each selected sponsor was asked for a list of the family child care homes sponsored, including three groups of homes: Tier 1 homes active (i.e., receiving CACFP reimbursement) in January 1998; Tier 2 homes active in January 1998; and all homes active in January 1997.³ A sample frame for “dropout” providers was defined to include all homes active in January 1997 that were not active in January 1998. Within each sponsor’s list of dropout homes, a random sample of five was drawn (for sponsors with five or fewer dropouts, all were drawn).⁴

A sample of 300 sponsors was selected within the 20 States.⁵ Of the selected sponsors, 289 supplied lists of current and former providers, and 280 of these had at least one former provider meeting the definition required for inclusion in the survey, for a response rate of 93.3 percent.⁶ From those lists, a sample of 1,971 former providers was selected.

¹ Four states were included with certainty (California, Michigan, Minnesota, and Texas).

² Sponsors were sampled with replacement, meaning that a sponsor could be selected more than once.

³ Homes received tier designations only when tiering was implemented, in July 1997.

⁴ The number of dropouts selected depended on the number of times the sponsor was selected – i.e., if the sponsor was selected twice, 10 dropouts rather than 5 would be selected from the sponsor’s list.

⁵ A total of 311 were selected, but 11 were not eligible because they had left the CACFP.

⁶ The data submitted by sponsors do not allow us to distinguish between a sponsor who had no homes leave the CACFP between January 1997 and January 1998 and a sponsor who could not identify the dropouts. For this calculation, we take the conservative approach of assuming that these 11 sponsors are all nonrespondents with regard to the list of former providers. If we assume that none of them actually had any dropouts, the response rate would be 96.3 percent.

Telephone “screening interviews” were attempted with these providers. The purpose of the screening interview was to determine the current status of the provider and, for those still providing care but not in the CACFP, to recruit them for a further survey of operations and meal service, as discussed below. The screening interview itself provides the data on the former provider’s status used in Exhibits 9 and 10.

The former provider’s current status was determined for 1,275 providers, or 64.6 percent of the sample, through the telephone screening survey. This includes five individuals who were not actually interviewed, but who were determined to have moved or died. In-person screening was then attempted for subsample of 195 of the 701 providers who could not be reached by telephone. Of these, a current status was determined for 123, or 63.1 percent (including 16 who had either moved or died).

Former providers who were identified during the telephone survey as still providing child care and not in the CACFP were asked to participate in a second survey, a mail survey with two self-administered components. One component, the Former Provider Operations Survey, gathered information about the providers’ current child care operations and their reasons for leaving the CACFP. The second component was the Former Provider Menu Survey, in which the providers kept a record of all foods served to children in their care during a selected week. The Operations Survey is the source of the information reported in Exhibit 11. Both the Operations Survey and the Menu Survey are analyzed more extensively in other reports (Zotov *et al.*, E-FAN-02-004; Crepinsek *et al.*, E-FAN-02-006).

Among the respondents reached in the telephone screener survey, 153 were determined to be eligible for the Operations and Menu surveys. Of those, 85 provided usable responses to the Operations survey. This represents a response rate of 55.6 percent among those screened. It represents a response rate of 48.2 percent among all members of the original sample estimated to be still providing care but not in the CACFP.⁷

It is sometimes useful in multi-stage samples to consider the compound response rate, which is the product of the response rates at each stage. The compound response rate for the screening survey is 66.2 percent, based on the sponsor response rate of 93.3 percent and a 70.9 response rate within the provider sample.⁸ The Operations Survey compound response rate is 45.0 percent, based on the sponsor response rate of 93.3 percent and the response rate of 48.2 percent within the former provider sample.

Nonresponse bias is always a potential concern in sample surveys, and the relatively low response rate for the Operations Survey makes it particularly salient for analyses based on that sample. This issue is discussed further below.

⁷ Among all subsample members whose status was determined, 3.4 percent were still providing child care and not in the CACFP. Applying this percentage to the 696 sample members whose status was not determined by the telephone survey yields an estimate of 23 providers. This is added to the 153 determined by the telephone survey to be still providing child care but not in the CACFP.

⁸ Responses for the telephone and in-person surveys are summed in this response rate.

Weighting

For producing population-based estimates of means and proportions of characteristics relating to former providers, each former provider received a sampling weight. These weights combined three elements: a weight reflecting the probability of selection of the sponsor; an adjustment to account for the fact that not all sponsors provided lists of dropouts; and a within-sponsor weight reflecting the probability that a particular dropout provider would be selected. The resulting weighted data yield estimates for all former providers in the population.

Basic Sponsor Selection *Weight*

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:

- Let W_i represent the weight for the i th selected State. $i = 1, 2, 3, 4, \dots, 19, 20$. $W_i = 1$ for States selected with certainty.
- Let W_{ij} be the weight for the j th selected sponsor in the i th State. We have

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

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 $W_{ij} = W_i$.

- The sponsors in the other eight States were selected with probability proportional to the number of providers. The conditional sponsor selection weight is:

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

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.

“Nonresponse” for sponsors includes sponsors who failed to supply any list of providers and those who supplied a list but the list indicated no dropouts between January 1997 and January 1998.

Let the number of sponsors responding to the provider lists be s^{**}_i out of the s_i selected. Then the nonresponse adjustment to the sponsor weight is

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

and the adjusted sponsor weight is

$$W^b_{j/i} = W^p_{j/i} A^*_i.$$

The overall sponsor weight is given by

$$W^b_{ij} = W_i W^b_{j/i}.$$

Dropout Provider Selection Weight

For the selection of providers from a selected sponsor, we stratify the providers by Tier 1, Tier 2, and dropout. Let P_{ijk} denote the number of providers sponsored by the j th sponsor in the k th stratum ($k=1,2,3$). Let p_{ijk} be the number of providers selected. Then the basic conditional weight for the l th selected provider in the k th stratum belonging to the j th sponsor in the i th State is

$$W_{l/ijk} = \frac{P_{ijk}}{p_{ijk}}.$$

This weight must be adjusted for nonresponse. If one or more of the providers for a particular sponsor fail to respond, the weights for the responding providers are inflated such that the sum of the adjusted weights for the responding providers equals the sum of the unadjusted weights of all originally selected providers for that sponsor. Thus, if out of p_{ijk} providers in the sample, only p^*_{ijk} respond, the nonresponse-adjusted conditional provider sampling weight is

$$W^a_{l/ijk} = \frac{P_{ijk}}{p^*_{ijk}} W_{l/ijk}.$$

The overall provider weight is therefore

$$W^{a_{l/ijk}} = W_i W^{b_{j/i}} W^{a_{l/ijk}}$$

This weight is used in all percentage distributions shown in Exhibits 9-11.

Use of the In-person Screening Subsample of Former Providers

Response patterns for the subsample were qualitatively similar to those for the telephone respondents, but did show some potentially important differences. For example, 63.1 percent of the former providers reached by telephone were not currently providing child care, compared with 71.3 percent of the subsample respondents. And while 13.1 percent of the telephone respondents were providing child care but not in the CACFP, only 3.3 percent of subsample respondents fell in that category. Estimates based solely on the telephone respondents would therefore be expected to understate the number of providers no longer providing child care and overestimate the number still providing child care but not in the CACFP.

Subsample respondents were therefore combined with telephone respondents in the analyses presented here. The weights for subsample respondents were adjusted such that the sum of the subsample respondents' adjusted weights equals the sum of the unadjusted weights of the 701 original sample members whose status was not determined by the telephone interview.

Nonresponse Bias

In order to assess the possibility of nonresponse bias, we examined those few bits of information that are available for both responding and nonresponding former providers. The only information available for the nonrespondents is their location and their sponsor. The analysis therefore focused on the percent of providers in each geographic region (Northeast, South, Midwest, and West) and two sponsor characteristics: the average number of homes sponsored, and the percent of sponsored homes that were Tier 1.

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 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.

Two analyses were performed. The first compared the respondents to the screening survey with the overall sample of former providers. The second compared respondents to the Operations survey with all providers who were identified as still providing child care but not participating in the CACFP.

Neither analysis revealed any bias by geographic region. In all instances the proportion of responding providers in the region was within two percentage points of the proportion for the full sample.

Somewhat larger differences were observed for sponsor characteristics. In particular, the former providers who responded to the survey at each stage tended to come from larger sponsors. The sponsors of former providers who responded to the screening survey had an average of 602 providers, about 9 percent greater than the average of 554 for the sample as a whole (statistically significant).⁹ Former providers responding to the Operations survey had sponsors that were about 5 percent larger, on average, than the total pool of providers eligible for the survey, but the difference is not statistically significant. The differences compound, however, so that the average sponsor size of the former providers responding to the survey is estimated at about 13 percent greater than would be the case if there were no nonresponse bias.

Some difference is also observed in the proportion of Tier 1 homes sponsored by the former provider's sponsors. The difference for screening survey respondents is fairly small (65.5 percent Tier 1 for sponsors of respondents, vs. 63.8 percent Tier 1 for the overall sample) but statistically significant. The difference for respondents to the Operations survey is in the same direction but not statistically significant. Because larger sponsors tend to have smaller proportions of Tier 1 homes, it is likely that the difference on this variable simply reflects the difference in sponsor size.

⁹ Note that this mean is calculated at the provider level, and large sponsors would be expected to have more former providers than small sponsors. As a result, the mean sponsor size in this calculation is much larger than the mean reported in analyses of sponsors.