

### **7.8.2 Analysis of Nonresponse Bias in the NSCAW Wave 1**

One investigation assessed nonreporting bias in the NSCAW. The NSCAW Wave 1 CPS data were analyzed to address the following questions:

- Is the language in the consent forms discouraging respondents from giving complete and accurate information?
- Are item missing rates indicative that current caregivers are concerned about the repercussions of honest and complete answers?
- The results of that investigation concluded the following:
  - Although respondents may have been concerned about the privacy of their answers, there is no evidence to suggest a tendency for respondents to either falsify or withhold information, either as a result of the consent form or information from the interviewer. In addition, interviewers appear to be neutral collaborators in the interview, whose presence does not seem to have had a detrimental effect on honest reporting.
  - Sensitive items are subject to significantly greater item nonresponse than non-sensitive items (98.2 vs. 99.8). However, for sensitive items, the item nonresponse rate is still less than 2 percent, which is negligible for most analyses. Therefore, the tendency for respondents to either actively or passively refuse to answer sensitive questions is quite small in the study.

Another investigation has been conducted in order to provide additional information on the extent of the bias arising from unit nonresponse—the failure to obtain an interview from a NSCAW sample member. An estimate of the nonresponse bias is the difference between the sample estimate (based only on respondents) and a version of the sample estimate based upon respondents and nonrespondents. In the NSCAW, a number of distinct data sources are used to obtain information on the sample child. When the sample child or caregiver did not respond to the survey, other data sources (such as the frame and caseworker data) can be used to provide information about them. Thus, it is possible to compare nonrespondents and respondents for some characteristics in order to investigate the potential nonresponse bias in the NSCAW results. In the remainder of this section, we briefly summarize the results of an investigation of the bias in the NSCAW results due to nonresponse using the data on nonrespondents available from other data sources.

An overall indicator of the severity of the bias due to nonresponse in the NSCAW is simply to count data items in our analysis for which respondents and nonrespondents differ significantly. Although this measure does not take into account either the type of comparisons that are significant or their importance for future analysis, it can be used as an indicator of the extent of the bias for general analysis objectives.

Variables used in this analysis were those that were also collected in the Wave 1 caseworker interview for the nonrespondents. However, only about 60 percent of the nonrespondents had a caseworker interview available. In this regard, the estimates of nonresponse bias are themselves subject to a bias due to incomplete information from caseworkers. However, we did not attempt to account for this potential bias in the analysis. These results assume that nonrespondents for whom caseworker information is unavailable are similar to nonrespondents for whom caseworker data is available.

Using the data collected for CPS and LTFC sample members from caseworkers at Wave 1, we estimated the bias due to using only the data for those with a key respondent interview. Let  $\pi$  denote the true average of the characteristic based upon the entire target population; i.e.,  $\pi$  is the average value of C that we would estimate if we conducted a complete census of the target population. Thus,  $\pi$  is the target parameter that we intend to estimate with  $\bar{y}_R$ . Then bias in  $\bar{y}_R$  as an estimate of  $\pi$  is simply the difference between the two, viz.,

$$B(\bar{y}_R) = \bar{y}_R - \pi \quad (1)$$

The bias can be estimated as follows. Let  $\bar{y}_{NR}$  denote the estimate of the average value of C for the unit nonrespondents in the sample; i.e.,  $\bar{y}_{NR}$  is computed as  $\bar{y}_R$  but over the nonrespondents in the sample rather than the respondents. For example, we may have information on the characteristic C that is measured in the child interview from some other source such as the caseworker or caregiver interview or the sampling frame. If that is true, then  $\bar{y}_{NR}$  can be computed. From this, we can form an estimate of  $\pi$  using the following formula:

$$\hat{\pi} = (1 - \eta) \bar{y}_R + \eta \bar{y}_{NR} \quad (2)$$

where  $\eta$  is the unit nonresponse rate for the interview corresponding to the characteristic C. Thus, an estimator of the bias in  $\bar{y}_R$  is obtained by substituting  $\hat{\pi}$  in (2) for  $\pi$  in (1). This results in the following estimator

$$\hat{B}(\bar{y}_R) = \bar{y}_R - \hat{\pi} \quad (3)$$

or, equivalently,

$$\hat{B}(\bar{y}_R) = \eta(\bar{y}_R - \bar{y}_{NR}) \quad (4)$$

That is, the estimator of the nonresponse bias for C is equal to the nonresponse rate for the interview that collects C times the difference in the average of C for respondents and nonrespondents.

We estimated these means and their standard errors using the weights and accounting for the survey design, as described in **Section 7.3**. We estimated  $\hat{\pi}$  using the unadjusted base weight. We estimated the mean for respondents,  $\bar{y}_R$ , in two ways: (1) using the unadjusted base weight, and (2) using the final adjusted analysis weight. This allowed us to see if the bias was reduced by applying the nonresponse and post-stratification adjustments to the weights.

We first tested the null hypothesis that the bias is 0 with  $\alpha=0.05$ , i.e.,  $H_0: \text{Bias}=0$  vs.  $H_a: \text{Bias} \neq 0$ . We used a t-statistic for the test, and Taylor series linearization to estimate the standard errors. Variables with fewer than 20 cases in the denominators of the proportions or means were excluded from the analyses. Because of the dependencies in the tests, we used the largest k-1 categories when a variable had k levels. We counted the number of times that the null hypothesis was rejected.

**Exhibit 7-2** summarizes the results of this analysis. The analysis for children is for those who were key respondents (i.e. age 10 or older); this group of children was eligible to be interviewed and assent from them was necessary in order for the interview to proceed. In the CPS data, for the child interview, the number of tests that were deemed significant is slightly more than the number expected purely by chance (6.9 percent using the final analysis weight). This analysis indicates for the caregiver that

there are more variables with significant bias than would be expected by chance (13.8 percent).

We examined the variables with significant bias. The biases, while statistically significant due to the large NSCAW sample size, were generally small and not practically significant. For this reason, we also tested a hypothesis of practical significance. A “practically important” bias was defined as the smallest difference that researchers would be concerned about in their analyses. A consensus opinion among Technical Working Group members was that an absolute relative bias exceeding 5 percent would be considered practically important; otherwise, it would not be of any practical importance. Practical significance then is defined as an absolute relative bias that is significantly different from 5 percent; i.e., if the hypothesis  $H_0: |\text{Relative Bias}| < 5 \text{ percent}$  is rejected in favor of  $H_a: |\text{Relative Bias}| > 5 \text{ percent}$ , the bias is considered practically significant.

Therefore, we tested the null hypothesis  $H_0: |\text{Relative Bias}| < 5 \text{ percent}$  vs  $H_a: |\text{Relative Bias}| > 5 \text{ percent}$ , where the relative bias is calculated as  $100 * \text{Bias} / \hat{\tau}$ . **Exhibit 7-2** shows the number of times that the null hypothesis was rejected at  $\alpha = 0.05$ , using both sets of weights. Note that, at  $\alpha = 0.05$ , we would expect  $H_0$  to be rejected for about 5 percent of all tests of significance when  $H_0$  is true. The exhibit shows that for the CPS sample, with the final analysis weight, the number of variables with practically significant relative bias is four percent, or within the range of what would be expected by chance. Thus, we conclude that nonresponse bias in the CPS sample is unlikely to be consequential for most types of analyses.

Variables showing practically significant bias in the CPS sample were variables related to the type and severity of abuse/neglect, relationship of the primary caregiver to the child, likelihood of abuse/neglect in the next 12 months without services, child placement in a group home, and the outcome of the investigation being substantiated. The actual bias in these variables was small (less than 10%).

**Exhibit 7-2 also shows the results for the LTFC sample. When using the final response adjusted analysis weight, approximately four percent of the tests that the bias is zero**

**were significant at a five percent alpha, and less than one percent of the tests that relative bias is small were significant at a five percent alpha. This analysis also suggests that the bias was reduced by applying the nonresponse adjustment to the weights. Thus, there is no evidence of nonresponse biases in the LTFC data.**

Exhibit 7-2. Number of Significant Biases Observed by Type of Respondent for the CPS and LTFC Samples

Caregiver	CPS Sample		LTFC Sample	
	Base Weight	Final Analysis Weight	Base Weight	Final Analysis Weight
Items with more than 20 cases in the denominator	500	500	1,107	1,107
Items where $H_0$ : Bias=0 was rejected	83 (16.6%)	69 (13.8%)	187 (16.9%)	50 (4.5%)
Items where $H_0$ :  Relative Bias <5% was rejected	33 (6.6%)	19 (3.8%)	32 (2.9%)	4 (0.4%)
Child	Base Weight	Final Analysis Weight	Base Weight	Final Analysis Weight
Items with more than 20 cases in the denominator	478	478	802	802
Items where $H_0$ : Bias=0 was rejected	48 (10.0%)	33 (6.9%)	108 (13.5%)	33 (4.1%)
Items where $H_0$ :  Relative Bias <5% was rejected	45 (9.4%)	19 (4.0%)	26 (3.2%)	8 (1.0%)

**Exhibit 7-3** indicates that the response rate tends to be slightly lower for children in the LTFC sample component aged 11 to 14 than for children 10 or younger. This suggests that the potential for nonresponse bias is greater for older children and their caregivers. This effect of age on nonresponse was not apparent in the previous analysis because those data were analyzed separately by key respondent type: child and caregiver. (For NSCAW, the caregiver was the key respondent when the child was less than 11 years old.) Therefore, the nonresponse bias results for children included only children who were at least 11 years old. Still, the lack of evidence for nonresponse bias in the previous analysis suggests that the greater relative bias for older children was quite small.

Exhibit 7-3. Response Rates by Age of Child for the LTFC Sample at Wave 1

Age	# of respondents	% unweighted response rate	% weighted response rate
0 - 2 years old	246	76.64	78.94
3 - 5 years old	122	71.35	64.37
6 B 10 years old	196	73.41	76.07
11- 14 years old	163	69.07	69.41
TOTAL	727	73.07	73.41

## 7.9 Nonresponse Bias Analysis for Wave 2

The total bias due to nonresponse at Wave 2 has two components—the bias due to nonresponse at Wave 1, and the additional bias due to nonresponse at Wave 2. **Section 7.4.2** describes the results of the analysis of the bias due to Wave 1 nonresponse.

While there may be some additional bias due to nonresponse at Wave 2, the amount of that bias was expected to be small since the Wave 2 response rates were near 90%. Further, many of the Wave 2 nonrespondents were interviewed at Wave 3, and some of the variables missing at Wave 2 (e.g., child/family service utilization history) were collected. This will further reduce the bias in the Wave 2 variables due to nonresponse.

An analysis of the additional bias due to Wave 2 nonresponse was conducted which examined the characteristics for some key variables for the Wave 1 respondents. Variables in the analysis include the variables considered for use in the Wave 2 weighting response adjustment, such as: the sampling strata, the sampling domains, gender, race, age, whether the child was receiving services, child setting, type of abuse, type of insurance coverage, overall health status of the child, urbanicity of the PSU, and size of the PSU. The formulas in **Section 7.8.2** were adapted for this analysis in order to evaluate the additional bias due to Wave 2 nonresponse. For this analysis, estimates based upon the Wave 1 respondent sample who also responded in Wave 2 were compared to the corresponding estimates based upon all Wave 1 respondents regardless of their Wave 2 response status. The difference between these two estimates for some characteristics is an estimate of the Wave 2 nonresponse bias for the characteristic. Further, in order to see if the nonresponse weight adjustments reduced or eliminated the nonresponse bias, the analysis was then repeated using the respondent sample and the

response-adjusted Wave 2 analysis weight (NANALWT2). A t-test was used to determine which variables had significant nonresponse bias at the 5 percent level. To adjust for multiple comparisons, the p-value was compared to  $0.05/(k-1)$  where k is the number of levels of the primary variable. **Tables 1 and 2** in **Appendix III-D** present the results of this analysis for the CPS and LTFC samples.

For the CPS sample, the variables with one or more categories showing statistically significant bias at the 5 percent level of significance were: sampling domain, age of child, whether receiving services based on frame data, whether the case was substantiated based on frame data, child setting, type of abuse, Wave 1 caregiver relationship to the child, standardized child Social Skills and Child Behavior Checklist scores, active alcohol use by caregiver, active drug use by caregiver, whether the caregiver had serious mental health problems, whether the caregiver had intellectual or cognitive impairments, whether the caregiver was considered to have poor parenting skills, and whether the family had trouble paying for basic necessities. In all cases the bias was small (less than 3%). In order to eliminate the bias, these variables were considered as dependent variables in the models for adjusting for Wave 2 nonresponse, and included if necessary to reduce the bias. Compared to the full sample:

- Among the sampling domains, a higher percentage of the Wave 2 respondents were caregivers of infants receiving services, or were children in out-of-home placement.
- A higher percentage of respondents had had substantiated cases or were receiving services (based on frame data).
- Among the age groups, a higher percentage of those with a Wave 2 response were ages 3-5 years old; there were no statistically significant differences for the other age groups.
- By type of setting, a higher percentage of those with a Wave 2 response were in some type of out-of-home setting (foster care, kin care, or other out-of-home). There were no differences in the percentage for those who were in-home and receiving services.
- By type of abuse, a higher percentage of respondents had “physical abuse - failure to provide” as the most serious type of abuse. There were no statistically significant differences between the other types of abuse.

- A higher percentage of respondents had a Wave 1 caregiver who was a foster mother, or a relative other than the natural father or mother.
- By the standardized Social Skills Rating scores, children ages 3-5 and 11-16 at Wave 1, a higher percentage of those with Wave 2 response had low standardized social skills.

For the LTFC sample, the following variables had one or more categories showing statistically significant bias: caregiver race, age of the Wave 1 caseworker, whether the caregiver had serious mental health problems, whether the family had trouble paying for basic necessities, and whether there was more than one supportive caregiver in the home at Wave 1. In all cases, the bias was small (less than 2%). These variables were included in the models that adjusted the weights for nonresponse. Compared to the total sample:

- By caregiver race, a lower percentage of Wave 2 responding children had a Black caregiver at Wave 1.
- By age of the Wave 1 caseworker, a higher percentage of respondents had a caseworker age 40 years or older and a lower percentage had a caseworker age 30-39 years old.
- A lower percentage of Wave 2 respondents had a primary caregiver with serious mental health problems.
- A lower percentage of respondents were from families that have trouble paying for basic necessities.
- A higher percentage of respondents had more than one supportive caregiver in the home.

When using the Wave 2 response adjusted weight (NANALWT2), none of the variables considered had statistically significant bias.

### **7.10 Nonresponse Bias Analysis for Wave 3**

This section presents characteristics of the CPS Wave 3 total sample, both respondents and nonrespondents in order to determine the extent to which nonrespondents differ from respondents. The analysis follows that for Wave 2 (described in *Section 7.9*). *Tables 3 and 4 in Appendix III-D* present the weighted distributions of all Wave 1 cases, Wave 3 respondents, and Wave 3 nonrespondents, for the variables that were considered for the Wave 3 nonresponse adjustment. These distributions were



computed using the Wave 1 analysis weight (NANALWT). These tables also present the bias in each variable due to considering only the Wave 3 respondents. Then, the tables present the distribution of the Wave 3 respondents using the Wave 3 analysis weight (NANALWT3), and the remaining bias in the variables after the weight adjustment. A t-test was also used to compare the individual percentages before and after the adjustment; this test was adjusted for the multiple comparisons within each variable.

For the CPS sample, prior to the Wave 3 nonresponse adjustment, the following variables showed differences between the total distribution and the Wave 3 respondent distribution:

- By sampling domain, a higher percentage of respondents were ages 1-14 and in out of home care.
- By child's health status, a larger percentage of respondents were in fair or poor health.
- By child's insurance coverage, a larger percentage of respondents were covered by Medicaid or another state-funded program.
- By type of setting, a higher percentage of those with a Wave 3 response were in some type of out-of-home setting (foster care, kin care).
- By caregiver work status at Wave 1, a higher percentage of the respondents had a caregiver who did not work.
- A larger proportion of respondents had a Wave 1 caregiver who was a foster mother.
- By the standardized Social Skills Rating scores, children ages 3-5 and 11-16 at Wave 1, a higher percentage of those with Wave 3 response had low standardized social skills.
- By the Child Behavior Checklist score for children ages 4 and older, a smaller proportion of respondents had a CBCL score below the normal range.

For LTFC, prior to the Wave 3 nonresponse adjustment, the following variables showed differences between the total distribution and the Wave 3 respondent distribution:

- By race/ethnicity, a smaller percentage of respondents were Black, and a larger percentage was White.

- By child setting, a higher percentage of respondents were in out-of-home care in some type of kin care arrangement.
- By Wave 1 caregiver race/ethnicity, a smaller percentage of respondents were Black, and a larger percentage of respondents either White or Hispanic.
- By relationship to the Wave 1 caregiver, a higher percentage of respondents had a caregiver who was a nonrelative.
- By Wave 1 caseworker age, a higher percentage of respondents had a caseworker who was ages 30-39, and a lower percentage had a caseworker who was 40-49 years old.
- A lower percentage of respondents had a primary caregiver with serious mental health problems.

The bias was small (4 percent or less) in all cases. When using the Wave 3 response-adjusted weight (NANALWT3), none of the variables reviewed had statistically significant bias.

### **7.11 Nonresponse Bias Analysis for Wave 4**

This section presents characteristics of the CPS Wave 4 total sample, both respondents and nonrespondents, in order to determine the extent to which nonrespondents differ from respondents. The analysis is similar to that for Wave 2 (described in *Section 7.4*). *Tables 5 and 6 in Appendix III-D* present the weighted distributions of all Wave 1 cases, Wave 4 respondents, and Wave 4 nonrespondents, for the variables that were considered for the Wave 4 nonresponse adjustment. These distributions were computed using the Wave 1 analysis weight (NANALWT). These tables also present the bias in each variable due to considering only the Wave 4 respondents. Then, the tables present the distribution of the Wave 4 respondents using the Wave 4 analysis weight (NANALWT4), and the remaining bias in the variables after the weight adjustment. A t-test was also used to compare the individual percentages before and after the adjustment; this test was adjusted for the multiple comparisons within each variable.

For the CPS sample, prior to the Wave 4 nonresponse adjustment, the following variables showed differences between the total distribution and the Wave 4 respondent distribution:

- By sampling domain, a higher percentage of respondents were ages 1-14 and in out of home care.
- By child's insurance coverage, a larger percentage of respondents were covered by Medicaid or another state-funded program, and a smaller percentage of respondents were covered by Private insurance.
- By type of setting, a higher percentage of those with a Wave 4 response were in some type of out-of-home setting.
- By caregiver age, a higher percentage of respondents had a caregiver age greater than 54 years old.
- A larger proportion of respondents have a Wave 1 caregiver who was a foster mother or other nonrelative.
- By the standardized daily living skills (Vineland) scores for children ages 6-10 Wave 1, a higher percentage of those with Wave 4 response had low standardized daily living skills
- By the Child Behavior Checklist score for children ages 4 and older, a smaller proportion of respondents had a CBCL score below the normal range.
- A higher percentage of respondents had had their Wave 1 caregiver reported as having poor parenting skills.
- Wave 4 respondents were less likely to have another supportive caregiver in the home at Wave 1.

For LTFC, prior to the Wave 4 nonresponse adjustment, the following variables showed differences between the total distribution and the Wave 4 respondent distribution:

- By child setting, a higher percentage of Wave 4 respondents were in kin care or out-of-home care other than foster care.
- By type of abuse, a higher percentage of respondents had physical neglect.
- A higher percentage of respondents had a caregiver with White race/ethnicity.
- By the standardized daily living skills (Vineland) scores for children ages 6-10 Wave 1, a higher percentage of those with Wave 4 response had moderate standardized daily living skills

- By the Child Behavior Checklist scores of children ages 4-18, a lower percentage of Wave 4 respondents had below normal CBCL scores.