EVALUATING THE USE OF DATA ON INTERRUPTIONS IN TELEPHONE SERVICE FOR NONTELEPHONE HOUSEHOLDS

J. Michael Brick, Joseph Waksberg, Westat, Inc., Scott Keeter, Virginia Commonwealth University
J. Michael Brick, Westat, Inc., 1650 Research Blvd., Rockville, MD 20850

Key Words: coverage, bias, weighting adjustment, telephone sampling

1. Introduction

Telephone surveys provide a relatively economical method of data collection compared with personal interviewing. However, telephone surveys are subject to an important source of bias that does not affect household surveys conducted with face-to-face interviewing: only 94 percent of households nationally have telephone service at any given time. Moreover, for some populations such as children, coverage rates are lower than 94 percent.

Weighting that includes poststratification based on demographic variables known to be associated with telephone coverage is effective in mitigating many of the consequences of coverage bias in telephone surveys generally.[1] But even when effective, weighting to known demographic totals does not completely solve the problem of coverage bias, undercompensating for some variables and overcompensating for others (Massey and Botman, 1988).

This report describes a study of an alternative method for adjusting telephone survey data to compensate for coverage bias. This method, developed by Keeter (1992), is based on the observation that telephone subscription is a variable condition not just across households in the population, but also within many households over time. A sizable number of U.S. households lose and gain telephone status during a given year. Because of this phenomenon, the telephone population at a given time includes households that have recently been in the nontelephone population. The existence of such households suggests the possibility that weighting adjustments that use the data from households that have telephones only sometimes during the year might be an improvement over the current practice.

To evaluate this new approach to adjusting the sampling weights, specific questionnaire items were added to two national RDD surveys conducted in 1993 by Westat. Both of these surveys used computer assisted telephone methods and the data were collected in the telephone research centers of Westat. One of the surveys is the National Household Education Survey of 1993 (NHES:93). The National Survey of Veterans (NSV) is the other survey.

We describe the estimates from the two surveys of the percentage of persons that experienced some interruption of telephone service, procedures for adjusting the survey weights using these data, and the statistical implications of using the adjusted weights. The final section summarizes the findings and gives some recommendations for use of this technique in telephone surveys.

2. Background

Despite considerable research on the size and characteristics of the nontelephone population, almost nothing is known about its dynamics over shorter time periods. Evidence from social workers, telephone companies, and others who deal with indigent households suggests that for many families, telephone subscription is episodic. They have telephones when they can afford them; the telephones are turned off when times are harder, when the bills get too large to manage, or both (Federal Communications Commission, 1992).

The Current Population Survey (CPS), conducted monthly by the Bureau of the Census, provides information on some aspects of the problem. The CPS employs a design in which households at a sample of addresses are interviewed monthly for 4 months, dropped for 8 months, then picked up for 4 more months of interviewing. The presence of a telephone in the household is verified in the first of each 4-month series of interviews. Using the 1991 and 1992 March CPS data files, 22,255 households were identified as part of the panel spanning those years.[2]

Another panel survey provides additional information on the dynamics of the telephone population. The 1988 National Election Study (NES), conducted by the Survey Research Center at the University of Michigan, was a two-wave panel study that interviewed a national sample of adults before and after the presidential election in 1988. The mean elapsed time between interviews was 51 days. Definitive data on telephone status were obtained at both time points for 1,729 out of the original 2,040 first-wave respondents.

While the estimates from the two surveys are not comparable because of the differences in the time between interviews, these data suggest that a nontrivial proportion of nontelephone households at any given

1 Postsurvey weighting is also used to compensate for nonresponse and other biases.

2 The CPS sample is of dwelling units rather than families. Some or all of the inhabitants of the dwelling unit may have changed during the time the unit was in the sample. We only included households for which the designated householder was reported to have remained the same from the previous year's interview. Consequently, estimates of the size of the "telephone transient" population are somewhat low.

time are transient. In the NES survey, 9 percent of nontelephone households had obtained telephone service 2 months later, while in the CPS, 40 percent had a telephone 1 year later.

Comparison of Transient and Chronic Nontelephone Households

Several conditions should be satisfied if the transient telephone households are to be useful in reducing coverage bias. One of the most important of these is the comparability of the characteristics of transient households and other nontelephone households.

Panel households from the CPS Study were divided into three groups: households reporting no telephone at either interview, transient telephone households whose telephone status changed from one interview to the next, and households with telephone service at both interviews. For comparison, CPS households that were not part of the 1992 panel and did not have telephone service at the time of the 1992 interview were also included.

The data show that transient households are more affluent than the 1991-92 panel households with no service and the nontelephone households in the new CPS sample, but their incomes are much closer to nontelephone households than to those with continuous coverage. In terms of labor force participation and unemployment of the householder, the transient group is quite similar to nontelephone households in the new CPS sample. The differences in reported utilization of various means-tested public programs are small between the transients and the nontelephone households in the new CPS sample. On the other hand, the differences in utilization of these public programs between either the transient or the 1991-92 panel households with no service with the continuous service households are quite striking.

Telephone Identification

The fact that a sizable number of transient telephone households exist, and that they are reasonably similar to all nontelephone households, does not guarantee that they can reliably be identified in a telephone survey. However, several small-scale telephone surveys conducted in Virginia between 1990 and 1993 (Keeter, 1992) provide evidence for the efficacy of the technique.

These statewide telephone surveys conducted by the Virginia Commonwealth University Survey Research Laboratory between 1990 and 1993 included questions asking respondents if they had lost telephone service at any time during the past 12 months. Four other general population surveys in Virginia yielded an average of 4 percent reporting a service interruption. Another study that interviewed a competent household participant about health insurance found a 3 percent rate of interruption.

On a variety of characteristics, the transient telephone households in the Virginia surveys were much more like nontelephone households than like those in the telephone population. The pattern of findings across these telephone surveys suggested that a larger test of the method was warranted. The results from the studies were the catalyst for the inclusion of items on the interruption of telephone service in the NHES:93 and the NSV.

3. Estimates of Interruptions of Telephone Service

Estimates of the percentage of persons with interruptions of telephone service from national surveys were needed to further examine the potential of reducing coverage biases using these data. As noted before, items were added to the NSV and the NHES:93 for this purpose.

In the NSV, about 23,000 households were screened and interviews were completed with over 5,500 eligible veterans. In the screening interview, all household members 14 years and over were enumerated and questions were asked about their characteristics and their veteran status. If the adult was a veteran (using a relatively broad definition) then an extended interview was attempted.

The results reported here are those asked about the adults enumerated in the screening interview. The items available for analysis from the NSV are limited to age, sex, race, education, and region, plus items about military service. Race and education were only obtained for the oldest person in the household.

In the NHES:93, 64,000 households were screened and nearly 30,000 interviews were conducted within those screened households. Two survey components were included: School Readiness (SR) and School Safety and Discipline (SS&D). Approximately 11,000 parents of 3- to 7-year-olds completed interviews on SR topics. About 12,700 parents of children in grades 3 through 12 and about 6,500 youth in grades 6 through 12 were interviewed for the SS&D component.

Data on interruptions in telephone service were collected from households in which at least one interview was completed with the parent most knowledgeable about the sampled child. The parent was asked if the household had experienced an interruption in telephone service for more than 24 hours in the last 12 months. If the respondent said yes, he or she was asked how many days, weeks, or months the household was without service.

Since the responses to these questions in the NHES:93 were only obtained for households with children that completed either an SR or SS&D interview, the data do not apply to as broad a population as the NSV. The data from this study pertain only to persons in certain types of households: those in which there was at least one child from preschool age (at least 3 years old) to the end of high school. Since the two eligible populations in the NHES:93 are not overlapping, the estimates are presented separately for the SR and SS&D.

Estimated Service Interruptions in the NSV and NHES:93

The estimated percentage of persons in households that had a telephone interruption of 1 day or more during the last twelve months varies substantially from survey to survey. Only 2.3 percent of adults had an interruption of 1 day or more based on the data from the NSV, while the percentage from the NHES:93 for younger children (the SR population of 3- to 7-year-olds) was 12.0 percent and for the SS&D population of older children it was 9.2 percent. Note that these are estimates of persons not households because both surveys were designed to produce these types of statistics.

Figure 1 shows estimates of any interruption along with estimates for those with interruptions of telephone service. The dotted lines are 95 percent confidence intervals on the estimated percentages. While the percentages vary from sample to sample, the patterns of increase by length of interruption are relatively stable.

The difference of about 3 percent in the estimates of the percentage with an interruption between the younger and older children from the NHES:93 seems reasonable, given the different telephone penetration for these groups. However, the difference between the NHES and NSV are not explained by this. Thornberry and Massey (1988) give estimates of telephone penetration by age of person.

The NSV interview asked if "At any time during the past 12 months, has your household not had telephone service?" This was followed with an item that asked if the interruption was for at least 24 hours. The NHES:93 interview began by asking if the household ever was without telephone service for more than 24 hours. A screening item such as that used in the NSV often depresses reports of subsequent activities, and this would be consistent with the observed results.

A second reason for the difference is due to the negative construction in the NSV screening question. With the NSV question, a 'no' response is a double negative because the question asks if they did not have telephone service. The combination of the negative construction and the screening item is likely to be a major contributor to the smaller estimate using the NSV questionnaire.

The difference in the estimates associated with the different ways of asking the questions is clearly shown by examining results from two surveys conducted in Virginia by VCU. In a November 1993 survey, the items about telephone interruptions were asked using the NSV wording, while in April 1994 the items were changed to be more like the NHES:93 wording.

The results from the surveys parallel the differences in the estimates between the NSV and the NHES:93. The November 1993 Virginia study estimated that 3 percent had an interruption in service in the last 12 months, while in April the estimated percentage was 9 percent. This evidence suggests that the different ways of asking the questions heavily influenced the size of the estimates.

These circumstances indicate that the estimates from the NSV are biased downward. Some adults who did experience an interruption in telephone service during the previous 12 months probably responded incorrectly to the NSV screening item.

Characteristics of Those With Service Interruptions

The telephone interruption estimates are examined below by the characteristics of the person to evaluate the potential of using these data to adjust for nontelephone coverage bias.

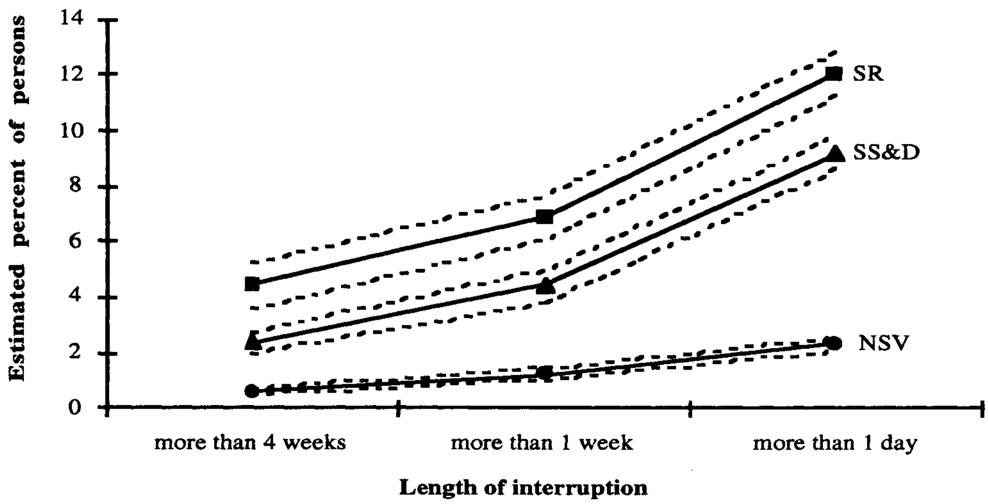
We estimated the percentage of persons in households with any interruption in service by three characteristics in both the NSV and the NHES:93. These estimates are shown in table 1. Some of the differences in the distributions may be due to the different ways of asking the questions. For example, the education classification is different in the two surveys. In the NSV education is recorded for the oldest person in the household, while in the NHES:93 education is the highest for either of the parents of the child.

The subsequent analysis is restricted to NHES:93 data for two reasons. First, more data on the characteristics are available from the NHES:93 extended interviews than the NSV screening interview. This enhances the analysis. Second, the telephone interruption estimate from the NSV is biased due to the wording of the item.

Using the NHES:93 data, we find that the percentage distributions of persons with some interruption are relatively consistent for the SR and the SS&D populations (see table 1). The characteristics generally associated with lower economic status have the highest percentage with interruptions. For example, the percentage of children with interruptions in both the SR and SS&D populations is higher for those from households with lower household income than for those from households with higher income.

For most of the items, the percentages of children in households with telephone interruptions are less variable for characteristics related to school readiness and school safety and discipline than for the socioeconomic items. Striking differences are for the estimates of the SR children participating in the Women, Infant, and Children (WIC) program and those participating in the free meal or lunch program at school. The higher estimates of the percentage with telephone interruptions for children participating in these programs designed for the economically disadvantaged is consistent with the estimates for the other economically related items.

Figure 1. Estimated percentage of persons with interrupted telephone service from the three populations



For most of the other substantive items, the differences in the percentage of persons with some interruption in telephone service were either not statistically significant or not large enough to be of great practical importance.

4. Weighting Adjustments

In almost all sample surveys, the data collected from respondents are adjusted to make the estimates more representative of the population surveyed. One of the most important benefits of this type of adjustment is that it often reduces the bias associated with the undercoverage of persons living in households without telephones.

To apply an adjustment using the data on telephone interruption, the probability of living in a telephone household is estimated for each respondent. The inverse of this estimated probability is the coverage adjustment. This model assumes that each person can be assigned a probability of being in a household with a telephone and that the probability is greater than zero. The model does not contain any provision for persons who always live in nontelephone households.

The data collected in the telephone surveys on telephone service interruptions and the length of those interruptions are the basis for making a similar type of adjustment. Instead of attempting to assign coverage probabilities to all respondents, we instead divide them into two categories: in households with interruptions of service and in households without interruptions. The weights for persons in households with no interruptions are not adjusted. The weights for those in households with interruptions are adjusted to account for those currently living in nontelephone households both those who have telephone service in the year and those who never had service in the year.

In the remainder of this section, we describe the specific weighting adjustment procedures examined using the telephone service interruption data.

Adjustment Schemes

For the SR and SS&D populations, we examined the various lengths of telephone interruption before deciding on two categories for forming adjustment cells: 1 week or more and 1 month or more. Those with interruptions of less than 1 week were not included in the adjustments because short-term interruptions are likely to have been caused by factors that are different from the longer term interruptions.

Within each of the length-of-service interruption categories, the children were classified into adjustment cells based on either parental education or tenure (home ownership). Race/ethnicity was used to form cells within the parental education and tenure categories. Four adjustment schemes were defined using these items.

- **Scheme A1**---children in households that had a telephone service interruption of 1 week or more within categories defined by crossing parental education and race/ethnicity;
- **Scheme A2**---children in households that had a telephone service interruption of 1 month or more within categories defined by crossing parental education and race/ethnicity;
- **Scheme B1**---children in households that had a telephone service interruption of 1 week or more within categories defined by crossing tenure (renter, own/other) and race/ethnicity; and
- **Scheme B2**---children in households that had a telephone service interruption of 1 month or more within categories defined by crossing tenure and race/ethnicity.

The adjustments were computed by first tabulating the proportion of children living in households without telephones from the March 1992 CPS within each of the cells defined for a scheme. The estimates from CPS were restricted to persons fulfilling

Table 1. Estimated percentage of persons with any interruptions in telephone service in last 12 months for three populations

	NSV		NHES:93 SR		NHES:93 SS&D	
	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Total	2.3%	0.1%	12.0%	0.4%	9.2%	0.3%
Region						
Midwest	2.3%	0.2%	11.0%	1.0%	7.3%	0.7%
Northeast	2.0%	0.2%	9.5%	1.2%	9.0%	0.8%
South	2.6%	0.2%	13.6%	0.7%	10.8%	0.6%
West	2.4%	0.2%	12.5%	0.9%	9.2%	0.8%
Race/ethnicity*						
White	2.0%	0.1%	9.3%	0.5%	7.2%	0.3%
Black	3.5%	0.4%	19.8%	1.5%	14.7%	1.1%
Hispanic	3.9%	0.5%	17.2%	1.5%	14.1%	1.1%
Other	2.6%	0.6%	11.7%	2.6%	9.3%	1.5%
Education**						
Less than high school diploma	3.2%	0.2%	18.4%	1.8%	17.4%	1.6%
High school graduate	2.0%	0.2%	15.4%	0.8%	11.0%	0.8%
Some college	2.3%	0.2%	11.8%	0.7%	8.6%	0.5%
Bachelor's degree	1.6%	0.2%	5.5%	0.8%	5.3%	0.8%
Graduate school	2.2%	0.3%	5.2%	0.7%	4.5%	0.6%
Household income						
$10,000 or less			22.8%	1.3%	19.0%	1.3%
$10,001 to $20,000			19.9%	1.4%	15.7%	1.1%
$20,001 to $30,000			9.3%	0.8%	7.9%	0.6%
More than $30,000			5.5%	0.5%	5.0%	0.3%
Women, Infant, and Children program participant†						
Yes			18.2%	1.3%		
No			8.0%	0.6%		
Free meal at school or center††						
Yes			20.5%	1.2%		
No			7.4%	0.4%		

* Race/ethnicity is reported for the oldest member in the NSV and for the child in the NHES:93.

**Education is for the oldest household member in the NSV and for most educated parent of the child in the NHES:93.

†Estimate restricted to preschoolers.

††Estimate applies to children except preschoolers.

SOURCE: U.S. Department of Veterans Affairs, National Survey of Veterans, summer/fall 1993, and U.S. Department of Education, National Household Education Survey, spring 1993.

the eligibility requirements for the SR and SS&D components separately.

The next step was to use the NHES:93 data for the two components to estimate the proportion of children living in households that had telephone interruption of the specified length, depending on the adjustment scheme. The adjustment factor for a cell was then defined as the ratio of the proportion of children in households without telephones in the cell (estimated from the CPS) to the proportion of children in households with a service interruption in that same cell (estimated from the NHES:93). Table 2 shows the resulting adjustment factors for the SR and SS&D components.

For each component, we applied the adjustment factors to NHES:93 person-level weights and computed four alternative weights. The last step was to rake the four alternative weights to the same known CPS totals as used in raking the standard NHES:93 person-level weights. The only difference in the weights is the adjustment for the telephone service interruption. The standard weights had no adjustment and the alternative weights had different adjustments depending on the scheme.

5. Findings

The adjustment of the weights to reduce the bias will also increase the variance of the estimates. Kish (1992) gives an approximate expression for this. We will call this expression the variance inflation factor (*VIF*). The *VIF* can be written as

$$VIF = 1 + CV^2(\text{weights}) \qquad (5.1)$$

Table 2. Weighting cell adjustments factors, based on length of interruption of telephone service

Factor	SR		SS&D	
	Length of service interruption			
	1 week or more	1 month or more	1 week or more	1 month or more
Cells defined by parental education and race/ethnicity (Schemes A1 & A2)				
Less than high school; Hispanic	4.42	12.04	4.12	7.03
Less than high school; black, non-Hispanic	3.26	4.15	3.21	4.36
Less than high school; white and other, non-Hispanic	3.86	4.14	3.16	3.97
High school diploma; Hispanic	2.15	2.54	2.51	4.17
High school diploma; black, non-Hispanic	2.33	3.19	2.77	4.20
High school diploma; white and other, non-Hispanic	2.08	2.67	2.13	3.00
College degree or more; Hispanic	1.34	2.32	1.95	8.14
College degree or more; black, non-Hispanic	1.76	2.61	1.34	8.70
College degree or more; white and other, non-Hispanic	1.57	2.09	1.91	3.46
Cells defined by tenure and race/ethnicity (Schemes B1 & B2)				
Renter; Hispanic	3.07	4.14	3.10	5.13
Renter; black, non-Hispanic	2.62	3.57	2.86	4.09
Renter; white and other, non-Hispanic	2.23	2.69	2.73	3.62
Owner/other; Hispanic	1.95	2.95	2.73	5.46
Owner/other; black, non-Hispanic	2.34	3.16	2.75	5.71
Owner/other; white and other, non-Hispanic	2.23	3.39	2.01	3.05

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey, spring 1993.

Table 3. Ratios of variance inflation factor due to coverage adjustment

Component	Sample size	VIF* std. weight	Ratio of scheme's VIF to standard weight's VIF			
			Scheme A1	Scheme A2	Scheme B1	Scheme B2
School Readiness	10,888	1.36	1.12	1.19	1.11	1.17
School Safety and Discipline						
6th through 12th graders	10,117	1.39	1.09	1.17	1.08	1.16
3rd through 5th graders	2,563	1.37	1.09	1.16	1.10	1.19
3rd through 12th graders	12,680	1.49	1.09	1.16	1.09	1.17

*VIF is the variance inflation factor. It is the relative variance of the weights plus one.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey, spring 1993.

where *CV* is the coefficient of variation of the weights.

Table 3 shows the *VIF* for the standard NHES:93 weights for each component. The SS&D component is broken down by the grade of the student, because youth were selected at different rates for these grade levels. The *VIF* for each of the components is about 1.4, indicating the variance is inflated by about 40 percent due to the differences in the standard weights. The *VIF* for the combined SS&D file is somewhat larger.

The other factors given in table 3 are the ratios of the *VIF* for the four alternative weights to the *VIF* for the standard weight. Overall, the increase in variance due to the telephone interruption coverage adjustment is about 8 to 12 percent for schemes A1 and B1. The ratios are somewhat greater for the schemes A2 and B2, ranging from about 16 to 19 percent. The larger ratios for the schemes based on interruptions of 1 month or more are a consequence of the larger and more variable factors shown in table 2. The ratios for the SR population are slightly higher than the SS&D ratios.

Since four adjustments were constructed, five different estimates can be computed: one from the standard NHES:93 weights and one for each of the four adjusted weights.

5.1 Coverage Bias Reduction

If estimates of the same characteristics as those produced from the NHES:93 were available from an independent source and these benchmark estimates were free of telephone coverage bias, then it would be possible to compare the five estimates to the benchmark. However, benchmark estimates of the NHES:93 do not exist. Consequently, other methods are needed to assess the bias-reducing potential of the coverage adjustments.

Although benchmark data are not available, in order to evaluate the approximate size of the bias, we tentatively assume that the adjustment procedures reduce the coverage bias of the standard estimates. The bias estimates under the given assumption provide an estimate of the decrease in the coverage bias resulting from using the procedures. The size of the bias reduction from using the adjustment is a critical feature of the assessment of the potential gains of the method.

The estimate from each scheme can be compared to the standard estimate, and the difference between the standard estimate and the adjusted estimate is an estimate of the reduction in the coverage bias. With four adjusted estimates, four different estimates of the bias reduction are produced. The estimated reduction in bias can be written as

$$b_a = \hat{p}_s - \hat{p}_a, \qquad (5.2)$$

where b_a is the estimated bias reduction using adjustment scheme a (a=A1, A2, B1, or B2), $\hat{p}_s$ is the estimate of the proportion using the standard estimate, and $\hat{p}_a$ is the estimated proportion using adjustment scheme a.

It is important to understand that the estimates are of the amount of reduction in bias in the **standard estimates**, assuming each adjustment scheme reduces the coverage bias. For example, the estimated reduction in the standard estimate of the percentage of preschool children attending a center-based program is 0.4 percent if scheme A1 is the assumed benchmark, and 0.2 if scheme A2 is used instead.

The bias reduction estimates for nearly all the items are relatively small and consistent across the schemes. In examining the estimates, it is important to realize that the total number of children is constant for all the estimates due to the raking of the estimates to the CPS totals. Therefore, the estimated reductions in bias across different response categories of an item must sum to zero. The fixed total number of children across response categories has two consequences. It creates a negative correlation in the estimated reductions in bias across response categories and gives a false impression of the number of independent pieces of information.

The approach taken to address this problem for summaries of the estimates is to delete the estimate for one of the response categories for each item. For example, a dichotomous item would have only one estimate reported.

Figure 2 presents the reduction in bias estimated using scheme A1 for the SR characteristics and figure 3 of the SS&D estimates. As a more formal summary of the reduction in bias, the mean and median were computed for each scheme and each component. For the SR component, the mean and median of the absolute value of the estimated bias reduction were 0.2 or smaller. For the SS&D, the estimates were comparable, with the mean and median of the absolute values almost all equal to 0.1.

The size of the absolute reduction in bias is not a directly useful statistical measure of the impact of the bias, mainly because it does not take the magnitude of the standard error of the estimate into account. The bias ratio, defined as the bias divided by the standard error of the estimate, does this. As the bias ratio increases, the chance of covering the true value departs significantly from the nominal confidence interval.

For the SS&D items, only a few of the bias ratio estimates are greater than 0.5. The bias ratios are larger for the SR items. The bias ratios are large across most of the items, with ratios exceeding 0.6 being common. These bias ratios are large enough to have a substantial impact on the confidence intervals.

These findings show that the biases could have an important effect on the inferences made from the survey estimates. The effect on the inferences is a greater problem for the SR component than for the SS&D component. The confidence intervals based on the standard estimates for some characteristics are not likely to attain the nominal confidence intervals due to the undercoverage bias.

The correlations between the bias reduction estimates under the four schemes are a measure of the consistency across the schemes. In general, the correlations are very high. For the SR component, the correlation between the estimates goes from a low of 0.6 between estimates for schemes A1 and B2, to a high of 0.8 between schemes B1 and B2. The correlations for the bias reduction estimates from the characteristics of the SS&D component are uniformly high, with the lowest correlation 0.8 between schemes A2 and B1 and the highest over 0.9 between schemes A2 and B2.

The consistency in the estimates indicates that bias adjustments from defining the cells by educational attainment within race/ethnicity are roughly equivalent to those defined by tenure within race/ethnicity.

So far, the discussion has only been about the consistency of the estimates of bias reduction. The variance implications are presented below.

5.2 Variance Implications

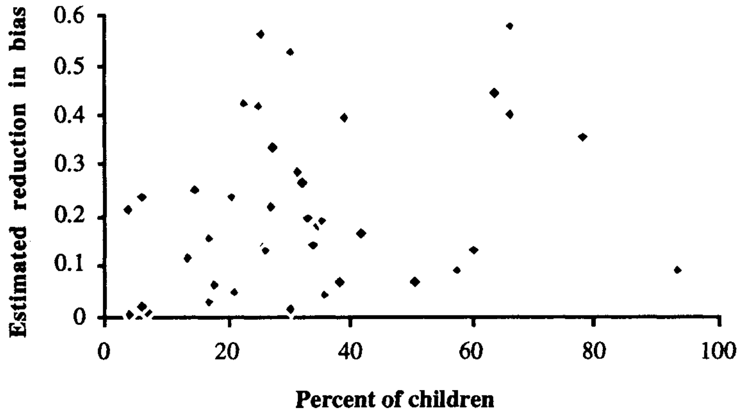
The results described above show that the standard estimates from the NHES:93 are subject to some coverage bias. The variance does not adequately measure error for biased estimates; the mean square error of the estimates is used instead. The mean square error (MSE) is the sum of the variance and the square of the bias of the estimate.

The MSE can be estimated for the NHES:93 estimates by using the standard variance estimates and the bias reduction estimates presented above. The estimated MSE can be approximated as

$$\text{MSE}_a = \text{var}(\hat{p}_s) + b_a^2 \qquad (5.3)$$

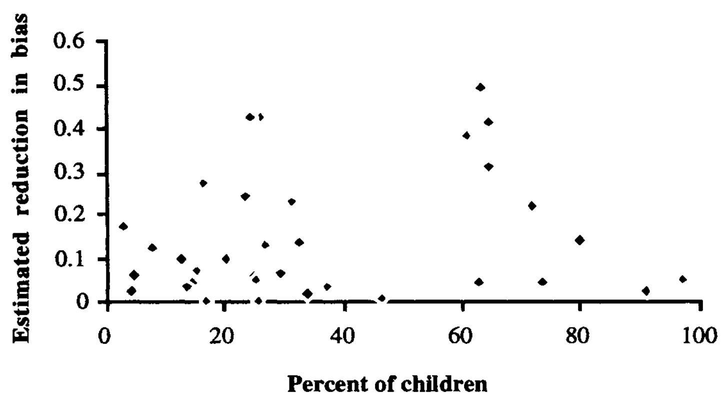
where $\hat{p}_s$ is the estimate proportion under the standard approach and b_a is the estimated reduction in bias under scheme a. Because of the high correlation in the estimates of the bias from the four adjustment schemes,

Figure 2. Estimated reduction in absolute value bias for School Readiness characteristics (scheme A1)



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey, spring 1993.

Figure 3. Estimated reduction in absolute value bias for School Safety and Discipline characteristics (scheme A1)



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey, spring 1993.

only the estimated mean square errors for scheme A1 were computed.

The estimated mean square errors of the estimates will now be used to contrast the variability in the standard NHES:93 estimate with the variability in the adjusted estimates.

The size of the variance increase from adjusting the weights using the telephone service interruption data was expressed as the VIF in table 3. Multiplying the standard variance estimates by the appropriate adjustment factor yields an approximate variance for the adjusted estimates.

To aid in comparing the procedures, the ratio of the estimated variance of the adjusted estimate to the estimated mean square error for the standard procedure was tabulated. This estimate is called the mean square ratio. It can be expressed as

$$MSR_a(\hat{p}) = 100 \times \frac{\text{variance of estimate under scheme a}}{\text{mean square error of standard estimate}}$$

$$= \frac{100 \times VIF_a \times \text{var}(\hat{p}_s)}{mse(\hat{p})} \qquad (5.4)$$

where the var($\hat{p}_s$) is the variance of the standard estimate, VIF_a is the adjustment factor for scheme *a*, and mse($\hat{p}$) is the estimated mean square error as discussed above. Note that the mean square error is derived using the bias estimated from scheme A1 only, but it is used to compute the mean square ratios for all four schemes. This is necessary to make the ratios comparable across the schemes.

The mean square ratios include contributions from both the reduction in bias (in the mean square error estimates) and the variance (in the *VIF*). When the mean square ratio is 100, the variance of the adjusted estimate is exactly equal to the mean square error of the biased, unadjusted estimate. A ratio less than 100 indicates that the bias reduction of the adjustment is greater than the variance increase that comes with it. A mean square ratio over 100 means that the variance increase associated with the adjustment is greater than the bias reduction.

Figures 4 and 5 show the MSR for each of the components. In addition, table 4 shows the summary statistics for these estimates.

For the SS&D items, the mean square ratios are typically greater than 100, and the increases in the variances due to the adjustment are larger than the bias reductions. These conclusions are true for all four adjustment schemes, but the schemes that require the longer interruptions (A2 and B2) have higher ratios.

This finding is consistent with earlier results for the SS&D items that showed the bias reductions associated with using the telephone interruption data for most items were relatively small. The penalty paid in terms of added variability is generally small, and none of the estimates of the mean square ratios are as large as 120. Furthermore, the use of adjustment schemes A2 and B2 result in larger VIFs and larger variances for the adjusted estimates.

The mean square ratios for the SR items present a more ambiguous picture. The medians of the mean square ratios are generally less than 100 for schemes A1 and B1, showing improvements from using the telephone service interruption adjustments.

Comparing the four schemes for the estimates from the SR component leads to conclusions consistent with those from the SS&D component. The schemes that use adjustments based on interruptions of telephone service of 1 week or more (A1 and B1) are better than the schemes based on interruptions of 1 month or more (A2 and B2). Schemes A1 and B1 have equivalent outcomes and the choice between them may be determined by nonstatistical issues.

6. Conclusions

If the percentage of the target population living in nontelephone households is relatively large and the characteristics of those persons are different from those who live in telephone households, then the estimates may be susceptible to significant coverage bias.

One method of addressing this problem without resorting to other modes of data collection is to statistically adjust the estimates to reduce the coverage bias. The weights for persons in households reporting an interruption in telephone service are increased to adjust for those without telephones.

The bias reduction estimates computed under the assumed model showed that the coverage adjustments for the SR component could have an important effect on the inferences from the survey estimates. The bias reduction estimates from the SS&D component, on the other hand, were generally small and not as substantively important.

When the variance implications of the adjustments were considered, the effectiveness of the procedure was less, even for the SR estimates. Although the adjustments reduced bias, they also increased the variability of the estimates.

The alternative weighting schemes performed differently with respect to the mean square ratios. The schemes based on interruptions of telephone service of 1 week or more generally outperformed the schemes based on interruptions of 1 month or more. The bias adjustments resulting from using educational attainment by race/ethnicity categories were roughly equivalent to those using tenure by race/ethnicity.

The SS&D items generally had mean square ratios greater than 100, indicating the increases in the variances due to the adjustment were larger than the bias reductions. However, the losses in precision were typically small.

The median of the mean square ratios was less than 100 for SR, suggesting that improvements could be obtained using the telephone service interruption adjustments. The characteristics improved the most by the alternative weight adjustments were those related to the economic status of the household, such as participating in the WIC or free meal programs.

The estimates have generally been based on assumptions that are favorable to the adjusted estimates. For example, the assumption that the adjustments always reduce the bias is not likely to be true, especially for variables that are not related to telephone coverage.

The adjustments also generally increase the variability in the estimation weights and this increase in variability is not beneficial unless it also decreases the bias of the estimates. We found that estimates related to economic status generally benefited from the adjustment, sometimes substantially. Major benefits were not found for most other types of estimates. If many of the important estimates are related to economic status, the adjustments are likely to improve the accuracy of the estimates.

References

Federal Communications Commission (1992). "Telephone penetration and household and family characteristics," prepared by Alexander Belinfante, Washington, DC.

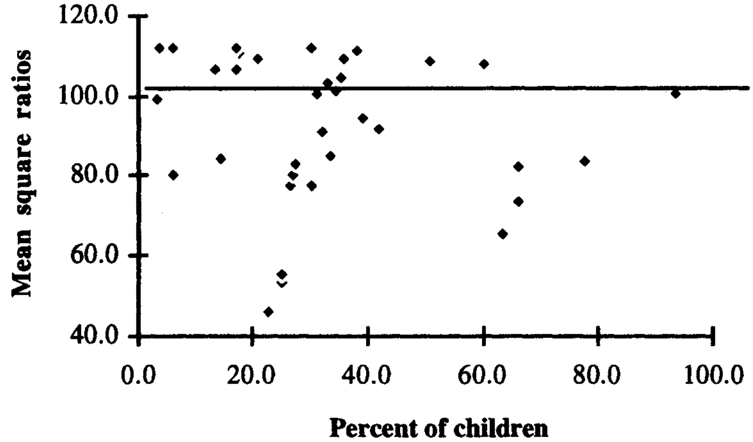
Keeter, S. (1992). "Estimating telephone noncoverage bias with a telephone survey" (Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Petersburg, FL).

Kish, L. (1992). "Weighting for unequal Pi," *Journal of Official Statistics, 8.*

Massey, J., and Botman, S. (1988). "Weighting adjustments for random digit dialed surveys," in *Telephone Survey Methodology*, eds. Groves, Biemer, Lyberg, Massey, Nicholls, and Waksberg. New York: John Wiley and Sons.

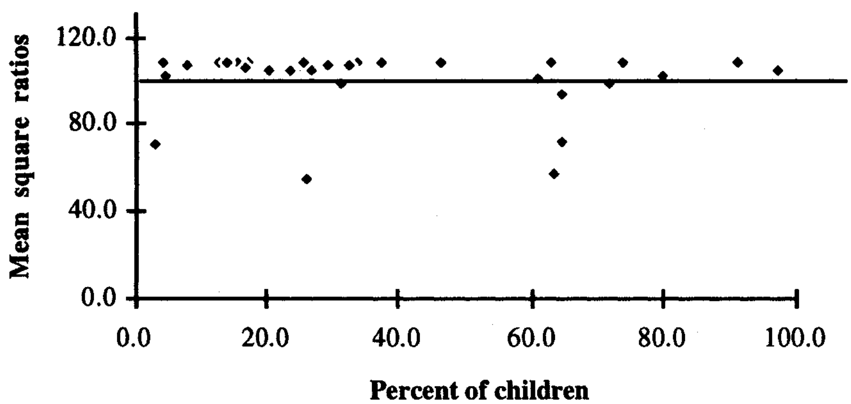
Thornberry, O., and Massey, J. (1988). "Trends in United States telephone coverage across time and subgroups," in *Telephone Survey Methodology*, eds. Groves, Biemer, Lyberg, Massey, Nicholls, and Waksberg. New York: John Wiley and Sons.

Figure 4. Estimated mean square ratios for selected School Readiness items (scheme A1)



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey, spring 1993.

Figure 5. Estimated mean square ratios for selected School Safety and Discipline items (scheme A1)



SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey, spring 1993.

Table 4. Summary statistics of mean square ratios

School Readiness	Scheme A1	Scheme A2	Scheme B1	Scheme B2
mean	92.8	98.6	92.0	97.0
median	99.6	105.8	98.7	104.0
minimum	46.2	49.1	46.8	48.3
maximum	112.0	119.0	111.0	117.0
School Safety and Discipline	**Scheme A1**	**Scheme A2**	**Scheme B1**	**Scheme B2**
mean	101.1	107.6	101.1	108.5
median	107.4	114.3	107.4	115.3
minimum	55.6	59.2	55.6	59.7
maximum	109.0	116.0	109.0	117.0