Survey Methodology, December 1996
Vol. 22, No. 2, pp. 185-197
Statistics Canada

Using Data on Interruptions in Telephone Service as Coverage Adjustments

J. MICHAEL BRICK, JOSEPH WAKSBERG and SCOTT KEETER[1]

ABSTRACT

Telephone surveys in the U.S. are subject to coverage bias because about 6 percent of all households do not have a telephone at any particular point in time. The bias resulting from this undercoverage can be important since those who do not have a telephone are generally poorer and have other characteristics that differ from the telephone population. Poststratification and the other usual methods of adjustment often do not fully compensate for this bias. This research examines a procedure for adjusting the survey estimates based on the observation that some households have a telephone for only part of the year, often due to economic circumstances. By collecting data on interruptions in telephone service in the past year, statistical adjustments of the estimates can be made which may reduce the bias in the estimates but which at the same time increase variances because of greater variability in weights. This paper considers a method of adjustment using data collected from a national telephone survey. Estimates of the reductions in bias and the effect on the mean square error of the estimates are computed for a variety of statistics. The results show that when the estimates from the survey are highly related to economic conditions the telephone interruption adjustment procedure can improve the mean square error of the estimates.

KEY WORDS: Coverage; Bias; Weighting adjustment; Telephone sampling; RDD surveys.

1. INTRODUCTION

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

Weighting that includes poststratification based on demographic variables known to be associated with telephone coverage is effective in mitigating some of the consequences of coverage bias in telephone surveys. Postsurvey weighting is also generally used to compensate for nonresponse and other biases. But even when effective, weighting to known demographic totals only partially solves the problem of coverage bias, undercompensating for some variables (Massey and Botman 1988) and overcompensating for others (Brick, Burke, and West 1992).

This article describes a study of an alternative method for adjusting telephone survey data to compensate for coverage bias. The method, suggested by Keeter (1995), is based on the observation that telephone subscription is a dynamic 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. Despite considerable information on the size and characteristics of the nontelephone population, little 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. Households may have a telephone when they can afford it, but the telephone may be turned off when times are harder, or when the bills get too large to manage, (Federal Communications Commission 1988). It is not known how many households change their telephone status and how long they stay in a particular status.

Keeter (1995) examined two household panel surveys to obtain estimates of the dynamics of telephone service subscription. Those households that changed telephone status (presence of a telephone in the household) are called 'transient' households. For data from one panel survey that collected data 12 months apart, half of the 6 percent of all households without a telephone at either time were transient. For the other panel survey in which data were collected only two months apart, one-fourth of the 6 percent of households without telephones at either point in time were transient. Since these estimates were based on observations at two points in time rather than continuous measurement, they underestimate the percent of households that are transient. Nevertheless, these results show that a substantial proportion of households without a telephone at a specific point in time is transient.

Another important condition that must be satisfied if the transient telephone households are to be useful in reducing

[1] J. Michael Brick and Joseph Waksberg, Westat, Inc., 1650 Research Blvd., Rockville, MD 20850, U.S.A.; Scott Keeter, Virginia Commonwealth University, Survey Research Laboratory, Richmond, VA 23284, U.S.A.

coverage bias involves the characteristics of transient households and nontelephone households. If the two groups are not similar, then the adjustments will not be effective. Using the panel data and data from several Virginia surveys, Keeter (1995) showed that the characteristics of the transient households are much more consistent with nontelephone households than telephone households.

These findings suggest 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 approach to adjusting the weights, questions were added to two national surveys conducted in 1993 by Westat. Both of these surveys were random digit dial (RDD) and computer assisted telephone surveys, 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 NHES:93 was conducted for the National Center for Education Statistics of the Department of Education in the spring of 1993 to study issues related to school readiness of young children and school safety and discipline of children in school. The other survey was the National Survey of Veterans (NSV) which was conducted in the second half of 1993 for the U.S. Department of Veterans Affairs. In this survey, adults were screened to determine if they were veterans, and the veterans were then asked about a variety of topics including their health, education, and financial status.

Below, we present estimates of the percentage of persons that experienced some interruption of telephone service, describe procedures for adjusting the survey weights using these data, and discuss the statistical implications of using the adjusted weights. The final section summarizes the findings and gives some considerations for using this technique in RDD telephone surveys.

2. 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. Questions 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 a sampled adult was a veteran, then a more detailed interview was attempted. The results reported here are those asked about the adults enumerated in the screening interview which included only a few characteristics of the adults and 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 and about 12,700 parents of children in grades 3 through 12 were interviewed for the SS&D component. Data on interruptions in telephone service were collected from households in which at least one SR or SS&D interview was completed.

Since the responses to the questions in the NHES:93 were only obtained for those households that completed either an SR or SS&D interview, many characteristics of the children can be analyzed, but the data do not apply to as broad a population as the NSV. The NSV applies to all adults, but only limited data were collected on most of the adults. For all households that had completed an interview (a screening interview in the NSV and a more detailed interview in the NHES:93), a member of the household was asked if the household had experienced an interruption in telephone service in the last 12 months and how long it lasted.

Estimated Service Interruptions in the NSV and NHES:93

The estimated percentage of persons in households that had a telephone interruption of one day or more during the last 12 months varies substantially from survey to survey. Only 2.3 percent of adults had an interruption of one 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 (grade 3 through 12) it was 9.2 percent.

Figure 1 shows estimates and 95 percent confidence intervals of the percentage of persons that had interruptions of one day or more along with estimates for those with interruptions of telephone service that lasted for at least one week and at least 4 weeks. While the percentages vary from sample to sample, the patterns of increase by length of interruption are relatively stable. The percentage with interruptions of one week or longer is less than half the percentage with any interruption, and the percentage with interruptions of 4 weeks or more is about one-fourth the percentage with any interruption.

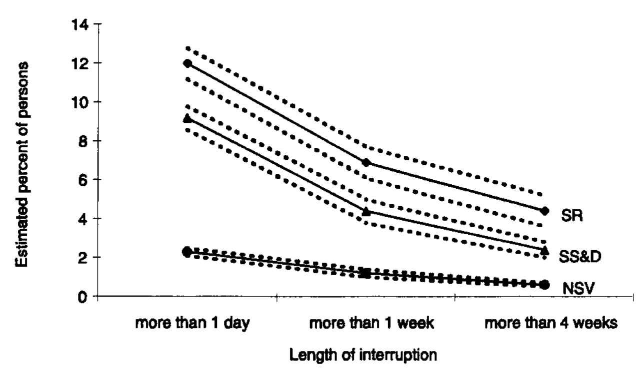


Figure 1. Estimated percentage of persons with interrupted telephone service from the three populations

The large difference in the estimates from the NSV and the NHES:93 comes from at least two important sources. The first source is that the populations were different. We would expect young children to live in households that experience more interruptions than older children and adults. Thornberry and Massey (1988) estimated that the telephone coverage rate for young children was lower than for any other age group. Thus, the difference of about 3 percent in the estimates of the percentage with an interruption between the younger (SR) and older (SS&D) children from the NHES:93 is reasonable.

The difference in the populations does not completely account for the large difference between the NSV and the NHES:93 estimates. An important reason for this difference is related to the way the questions were asked in the two surveys. The NHES:93 interview began by asking, "During the past 12 months, has your household ever been without telephone service for more than 24 hours?". In the NSV interview, respondents were asked if, "At any time during the past 12 months, has your household *not* had telephone service?". This was followed by a question that asked if the interruption was for at least 24 hours. Thus, the NSV version was a screening item followed by a more detailed question. This type of construction often depresses reports of subsequent activities, which is consistent with the lower NSV estimates.

A more important reason for the difference is probably due to the wording of the questions. With the NSV question, a 'no' response may have confused respondents because the question asks if they did *not* have telephone service. Converse and Presser (1986) discuss the problems that arise with this type of question construction. The wording for the NHES:93 is less confusing. The combination of the wording and the use of a screening item in the NSV is likely to be the main reason for the smaller estimate using the NSV questionnaire.

The difference in the estimates associated with the different ways of asking the interruption questions is evident from the estimates from two surveys conducted in Virginia by Virginia Community University. In a November 1993 survey, the items about telephone interruptions were asked using the NSV wording; in April 1994 the items were changed to 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. Thus, it is clear that the different ways of asking the questions heavily influenced the size of the estimates, and it suggests 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 in the NSV.

Characteristics of Persons With Service Interruptions

Estimates of the percentage of persons who had a telephone interruption are examined below by the characteristics of the person to evaluate the potential of using these data to adjust for coverage bias. We estimated the percentage of persons in households with any interruption in service by characteristics collected in both the NSV and the NHES:93. These estimates are shown in the first part of table 1. Some 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.

All subsequent analysis is restricted to NHES:93 data for two reasons. First, more data on the characteristics are available from the NHES:93 detailed SR and SS&D interviews than the NSV screening interview. Second, the telephone interruption estimate from the NSV is biased due to the wording of the item, as discussed earlier. Of course, the NHES:93 estimates apply to households with children which have higher nontelephone rates than the general population, and in that sense they do not reflect the situation for the total population.

Using the NHES:93 data, we find that the percents 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 larger for those from households with lower household income than for those from households with higher income. Similarly, children participating in public assistance programs (WIC or free meals) have much higher rates of service interruptions than nonparticipants. However, 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. Additional characteristics for both populations were examined and presented in Brick, Keeter, Waksberg and Bell (1996), but are not shown here. 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.

3. WEIGHT ADJUSTMENTS

In almost all sample surveys, the data collected from respondents are adjusted to account for nonresponse and noncoverage and to reduce the variability in the estimates by using auxiliary data from other data sources. One of the most important benefits of this type of adjustment in telephone samples is that it often reduces the bias associated with the undercoverage of persons living in households without telephones.

Kalton and Kasprzyk (1986) discuss adjustments to the base weights, classifying the adjustments into four categories: population weighting adjustments, sample weighting adjustments, raking ratio adjustments, and response probability

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[1]						
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[2]						
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[3]						
Yes			18.2	1.3		
No			8.0	0.6		
Free meal at school or center[4]						
Yes			21.1	1.2		
No			7.6	0.5		
Birth weight						
5.5 pounds or less			12.0	1.6		
Greater than 5.5 pounds			12.0	0.4		
School control						
Public					9.4	0.4
Private					7.5	1.1
Ease of obtaining marijuana at school[5]						
Very or fairly easy					9.7	0.6
Hard					8.0	0.8
Nearly impossible					9.0	0.7

[1] Race/ethnicity is reported for the oldest member in the NSV and for the child in the NHES:93.
[2] Education is for the oldest household member in the NSV and the most educated parent of the child in the NHES:93.
[3] Estimate restricted to preschoolers.
[4] Estimate applies to children except preschoolers.
[5] Estimate applies only to children in grades 6 through 12.
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.

adjustments. In the NHES:93, sample weighting adjustments and raking ratio adjustments were used. Sample weighting adjustments were used to account for differential nonresponse from sampled persons. Raking ratio adjustments were then used to make the specified marginal distributions of the sample correspond to totals from the October 1992 Current Population Survey (CPS). One of the most important benefits of the type of raking ratio adjustment used in the NHES:93 is that it reduces the bias associated with the undercoverage of persons living in households without telephones because the CPS covers persons in both telephone and nontelephone households.

The data on telephone service interruptions can be used to make a response probability adjustment. Response probability adjustments are constructed by assuming that each sampled unit has a probability of responding to the survey, estimating that probability, and then using the inverse of the estimated response probability as a weighting adjustment. The Politz and Simmons (1949) method is probably the best known application of the response probability adjustment procedure, and Kalton and Kasprzyk (1986) discuss others.

To apply this type of adjustment using the telephone service interruption data, assume that living in a telephone household is a dynamic phenomenon and that a probability distribution can be associated with this status. Conceptually, a survey is conducted by sampling from this distribution and observing only those members that live in telephone households at the time of the survey. The probability of living in a telephone household (the equivalent of the response probability) must then be estimated for each respondent. The inverse of the 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 between zero and one (but not equal to zero).

The data on whether or not a household had an interruption in telephone service and the length of that interruption are the basis for this type of adjustment. Persons are divided into two categories: those in households with interruptions in service and those in households without interruptions in service. The probability is assumed to be equal to one for persons in households without interruptions and their weights are not adjusted. The weights of persons in households with at least some interruptions in the last 12 months are adjusted to account for other households that have a probability of being covered of less than one. The adjustments may vary depending on the length of time they lived in nontelephone households and on other characteristics of the household. The purpose of having different adjustments is to account for the fact that some persons are more likely to live in nontelephone households than others.

Although the weighting adjustments may reduce the undercoverage bias, introducing adjustments also typically increases the variances of the estimates. Kish (1992) discusses the reasons for unequal weights as well as the consequences from using them in a variety of situations. He advocates a common statistical approach of balancing the bias reductions against the variance increases. If the weights reduce the bias of the estimates significantly, then it may be worthwhile accepting the variance increases. On the other hand, small reductions in bias associated with large variance increases are not recommended.

In the remainder of this section, the specific weighting adjustment procedures are described. The statistical properties of the weights developed under four alternative adjustment schemes are presented. The alternative weights are applied to the NHES:93 data and the decrease in the bias of the estimates is compared with the increase in the variance of the estimates due to the unequal weighting.

Adjustment Schemes

The first step was to decide how to classify the length of interruption in telephone service. Various lengths of interruptions were examined to determine cut-offs that discriminated between temporary interruptions, not due to economic causes and others. It was decided to use two categories for forming adjustment cells: one week or more, and one month or more.

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. These cells were chosen because the percentage of persons with interruptions varied by these characteristics and the corresponding data were also available from the CPS. Four adjustment schemes were defined using these items:

Scheme A1 – children in households that had a telephone service interruption of one week or more within categories defined by parental education (less than high school, high school diploma, college diploma or above) and race/ethnicity (Hispanic, black/non-Hispanic, white and other/non-Hispanic);

Scheme A2 – children in households that had a telephone service interruption of one month or more within categories defined by parental education and race/ethnicity;

Scheme B1 – children in households that had a telephone service interruption of one week or more within categories defined by tenure (own/other, rent) and race/ethnicity; and

Scheme B2 – children in households that had a telephone service interruption of one month or more within categories defined by tenure and race/ethnicity.

The adjustment factors for these schemes could not be obtained directly from the NHES:93 data because no data were collected from households without telephones. Instead, the adjustments were developed using both CPS and NHES:93 data and then applied to the NHES:93 weights.

To motivate the adjustment of the weights, consider partitioning the universe of persons into four components: t_1 is the number of persons in *telephone* households with *no telephone interruptions* in the past year; t_2 is the number of persons in *telephone* households with *some telephone interruptions* in the past year; t_3 is the number of persons in *nontelephone* households with *no telephone interruptions* in the past year (*i.e.*, persons who lived in nontelephone households throughout the entire year); and t_4 is the number of persons in *nontelephone* households with *some telephone interruptions* in the past year. As noted above, the response probability model assumes $t_3 = 0$.

Using the CPS it is possible to estimate $t_1 + t_2$ and t_4 (assuming $t_3 = 0$); designate these estimates as $\hat{t}_1 + \hat{t}_2$ and $\hat{t}_4$, respectively. From the NHES:93, t_1 and t_2 can be estimated separately; call these estimates t_1^* and t_2^*, respectively. The bias in the NHES:93 estimates arises because they are from a telephone survey and do not include persons in nontelephone households (t_4).

A weight adjustment of $A = 1 + t_4/t_2$ would result in unbiased estimates of totals; however, this adjustment involves unknown, population quantities that must be estimated. Since t_2 can only be estimated from the NHES:93 and t_4 can only be estimated from the CPS (assuming $t_3 = 0$), the adjustment is expressed in ratios to reduce the bias due to estimating the totals from different surveys. The revised weight is

$$w_i' = w_i \left(1 + \delta_i \frac{\dfrac{\hat{t}_4}{\hat{t}_1 + \hat{t}_2}}{\dfrac{t_2^*}{t_1^* + t_2^*}} \right), \qquad (1)$$

where w_i is the NHES:93 weight adjusted for nonresponse of sampled persons but not yet raked to October 1992 CPS totals, $\delta_i = 1$ if the person lives in a household that had an interruption of telephone service in the last year and is zero otherwise. The quantity in parenthesis in (1) is an estimate of A, the weight adjustment.

Revised weights were computed separately for the SR and SS&D components. Rather than the overall adjustment as given in (1), the weight adjustments were computed within the cells defined for each of the four weighting schemes (A1, A2, B1, and B2). Table 2 shows the resulting adjustment factors for the SR and SS&D components. The adjustments in the first column are those for schemes A1 and B1. The second column contains the adjustment factors for schemes A2 and B2. The adjustment factors for the schemes based on the one month or more interruptions are greater than those based on the one week or more because the denominator of the ratio is, by definition, smaller for this classification (see Figure 1 for estimates of the percentage of persons with interruptions for each scheme).

The last weighting step rakes the four alternative weights to the same October 1992 CPS totals used in raking the standard NHES:93 person-level weights. The result of this process is the standard NHES:93 weight and four alternative weights based on different adjustment schemes. All five of the weights conform to the same marginal totals. The only difference in the weights is the adjustment for the telephone

Table 2
Weighting Cell Adjustments Factors, Based on Length of Interruption of Telephone Service

Factor	SR		SS&D	
	Length of service interruption			
	One week or more	One month or more	One week or more	One month or more
Cells defined by parental education and race/ethnicity (Schemes A1 and A2)				
Less than high school; Hispanic	5.75	16.35	4.89	8.52
Less than high school; black, non-Hispanic	5.10	6.72	4.26	5.95
Less than high school; white and other, non-Hispanic	4.98	5.37	3.81	4.86
High school diploma; Hispanic	2.31	2.76	2.67	4.51
High school diploma; black, non-Hispanic	2.65	3.73	3.06	4.71
High school diploma; white and other, non-Hispanic	2.16	2.79	2.18	3.09
College degree or more; Hispanic	1.34	2.33	1.96	8.22
College degree or more; black, non-Hispanic	1.77	2.64	1.35	8.83
College degree or more; white and other, non-Hispanic	1.58	2.09	1.91	3.48
Cells defined by tenure and race/ethnicity (Schemes B1 and B2)				
Renter; Hispanic	3.74	5.15	3.58	6.08
Renter; black, non-Hispanic	3.23	4.54	3.38	4.95
Renter; white and other, non-Hispanic	2.43	2.96	2.99	4.00
Owner/other; Hispanic	2.00	3.06	2.81	5.66
Owner/other; black, non-Hispanic	2.53	3.46	2.90	6.11
Owner/other; white and other, non-Hispanic	2.26	3.45	2.03	3.10

service interruption prior to raking. The standard weights are not further adjusted while the alternative weights have different adjustments depending on the scheme.

4. FINDINGS

As noted above, adjustment of the weights to reduce the bias increases the variability of the weights, thus increasing the variance of the estimates. Kish (1992) gives an approximate expression for this increase in variance arising from unequal weights. We call this expression for the increase in variance due to differential weights the variance inflation factor (*VIF*). The *VIF* can be written as

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

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 variability in the standard weights. The *VIF* for the combined SS&D file is somewhat larger (1.5) because it includes youth who were sampled at different rates.

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. These ratios show how much greater the variances of estimates produced using the alternative weights are expected to be as compared to the variances of the standard NHES:93 weights.

Overall, the increase in variance due to the telephone interruption coverage adjustment are from 9 to 13 percent for schemes A1 and B1 in the SS&D component but up to 20 percent for the SR component. The ratios are larger for the schemes A2 and B2, ranging from 24 to 35 percent, with the largest ratio for Scheme A2 for the SR component. The larger ratios (hence *VIF*s) for the schemes based on interruptions of one month or more are a consequence of the larger and more variable factors shown in the second column of table 2. The ratios for the SR population are higher than the SS&D ratios.

4.1 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, benchmarks comparable to the estimates from the two components of the NHES:93 do not exist and other methods are needed to assess the bias-reducing potential of the coverage adjustments.

Due to of the lack of a benchmark, some model assumptions are required to assess the effectiveness of the adjustments. For this evaluation we assume that the adjustment procedures reduce the coverage bias. As a result of this assumption, the difference between the standard estimate and the adjusted estimate is considered an unbiased estimate of the decrease in the coverage bias resulting from using the procedures. Clearly, the coverage bias is not completely eliminated by any of the adjustment procedures. Even if the model were correct, the bias reductions from the data would still be subject to sampling error. Despite the problems with this assumption, this type of assumption is necessary to obtain some idea of the effectiveness of the adjustment. If the adjustment eliminates the bias, the mean square errors of the adjusted estimates are equal to the variances of the estimates, with no contribution from coverage bias. Therefore, the model assumption is favorable to the adjusted estimates, positing the adjusted estimates to be unbiased. The impact of this assumption is discussed critically after evidence of the effectiveness of the method is presented.

The estimate from each scheme can be compared to the standard NHES:93 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 bias reduction are possible. The estimated reduction in bias is

$$b_a = \hat{p}_s - \hat{p}_a, \qquad (3)$$

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.

Table 3
Ratios of Variance Inflation Factor Due to Coverage Adjustment

Component	Sample size	*VIF** standard 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	1.20	1.35	1.16	1.26
School Safety and Discipline						
3rd through 5th graders	2,563	1.37	1.12	1.25	1.13	1.26
6th through 12th graders	10,117	1.39	1.13	1.27	1.09	1.24
3rd through 12th graders	12,680	1.49	1.12	1.26	1.11	1.25

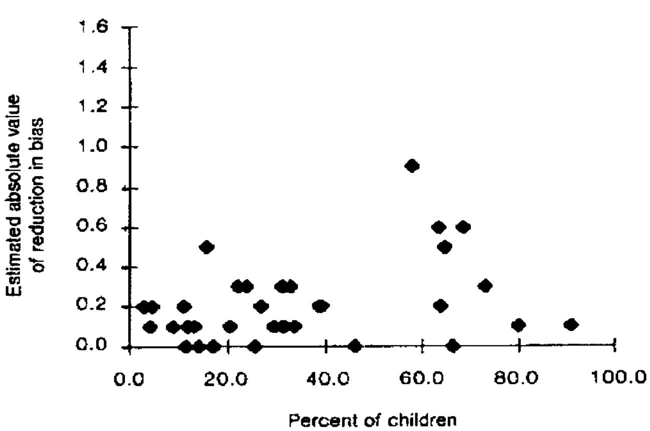
* *VIF* is the standard inflation factor. It is the coefficient of variation of the weights squared plus one.
Source: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey, spring 1993.

The estimated reductions in bias under each adjustment weighting scheme are given in table 4. Estimates for additional characteristics are given in Brick *et al.* (1996). The bias reductions in the standard estimate assume each adjustment scheme eliminates the coverage bias.

The bias reduction estimates for most of the items in Table 4 are less than one percent and consistent in direction across the schemes. Before summarizing the estimates, we must account for the fact that the total number of children is constant for all the estimates due to the raking of the estimates to the CPS totals. The fixed total number of children across response categories has two consequences: it creates a negative correlation in the estimated reduction in bias across response categories; and it gives a false impression of the number of independent pieces of information in the tabled values.

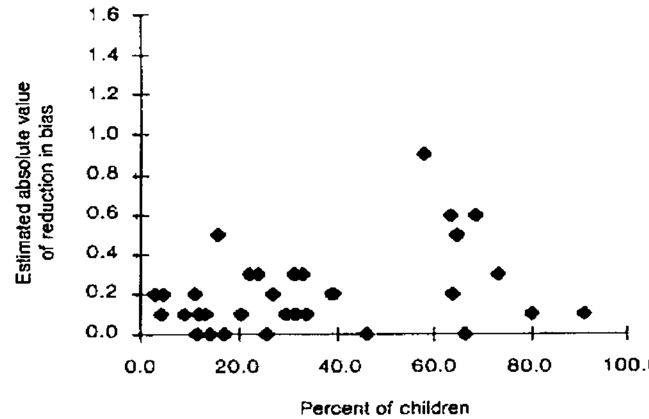
The approach taken to address to this problem in summarizing the bias estimates is to delete the estimate for one of the response categories for each item. The "no" response category for all items with "yes" and "no" response categories was deleted. For other types of variables, the response category with the smallest estimate was deleted.

Figure 2 presents the absolute value of the reduction in bias estimated using scheme A1 for the SR characteristics, and figure 3 is the same representation for the SS&D. These figures use all the estimates presented in Brick *et al.* (1996), rather than just those shown in table 4. For both components, the bias reductions are small. The largest absolute bias is 1.3 percent for SR and 0.9 percent for SS&D. The mean and median of the bias reductions and the absolute values of the bias reductions were also computed for each scheme and each component. For the SR component, the mean and median of the absolute value of the estimated bias reductions for the four schemes are between 0.2 and 0.4 percent. For the SS&D, the mean and median of the absolute values are between 0.1 and 0.3.



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

Figure 2. Estimated reduction in absolute 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 of bias for School Safety and Discipline characteristics (scheme A1)

Bias Ratio

The size of the absolute reduction in bias is not a very useful statistical measure of the impact of the bias because it does not take the magnitude of the sampling error of the estimate into account. Cochran (1977) discusses the impact on confidence intervals as the ratio of the bias to the sampling error varies. For each scheme the bias ratio is given by

$$r_a = \frac{b_a}{\mathrm{se}(\hat{p}_s)}, \tag{4}$$

with the standard error of the standard estimate as the denominator. As the bias ratio increases, the chance of covering the population value departs significantly from the nominal confidence interval.

The bias ratios for selected characteristics are shown in Table 4. Many of the bias ratios for the SR items are large, even though the average and median ratios are near zero. Nearly half of the ratios for all the items examined are larger than 0.4 in absolute value. A ratio of 0.4 is large enough to reduce a nominal confidence interval from 95 percent to about 93 percent. For the SS&D items, the bias ratios are smaller, with only 15 percent of all the items having bias ratios greater than 0.4.

4.2 Mean Square Error

Since the variance is not an adequate measure of 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

$$\mathrm{MSE}_a = \mathrm{var}(\hat{p}_s) + b_a^2 \tag{5}$$

Table 4
Estimated Reduction in Bias and Bias Ratio for Selected Characteristics of the NHES:93

Characteristic	Standard estimate		Estimated reduction in bias				Bias ratio			
	Estimate	Standard error	Scheme A1	Scheme A2	Scheme B1	Scheme B2	Scheme A1	Scheme A2	Scheme B1	Scheme B2
School Readiness (SR) population										
Parental educational level										
Less than high school graduate	8.6	0.3	-1.7	-1.9	0.1	0.1	-5.7	-6.3	0.3	0.3
High school graduate or equivalent	33.9	0.8	0.4	0.3	-0.7	-1.0	0.5	0.4	-0.9	-1.3
Some college	57.5	0.7	1.3	1.6	0.6	0.9	1.9	2.3	0.9	1.3
Mother's employment status										
No mother in household	2.4	0.2	-0.1	-0.1	-0.1	-0.1	-0.5	-0.5	-0.5	-0.5
Employed 35 hours/week or more	34.3	0.5	0.5	0.8	0.2	0.5	1.0	1.6	0.4	1.0
Employed less than 35 hours/week	20.9	0.5	-0.1	-0.2	0.0	-0.2	-0.2	-0.4	0.0	-0.4
Seeking employment	6.6	0.4	0.0	-0.1	-0.1	-0.1	0.0	-0.3	-0.3	-0.3
Not in labor force	35.8	0.6	-0.4	-0.3	0.0	0.0	-0.7	-0.5	0.0	0.0
Father's employment status										
No father in household	26.3	0.5	-0.4	-0.6	0.0	-0.1	-0.8	-1.2	0.0	-0.2
Employed 35 hours/week or more	63.4	0.6	0.3	0.5	0.1	0.2	0.5	0.8	0.2	0.3
Employed less than 35 hours/week	3.8	0.3	0.0	-0.1	0.0	0.1	0.0	-0.3	0.0	0.3
Seeking employment	3.2	0.3	0.0	0.0	-0.1	-0.2	0.0	0.0	-0.3	-0.7
Not in labor force	3.3	0.2	0.1	0.2	0.0	0.1	0.5	1.0	0.0	0.5
Time since doctor visit for routine care										
Less than 1 year	84.1	0.4	0.4	0.4	0.2	0.1	1.0	1.0	0.5	0.2
Over 1 year	15.9	0.4	-0.4	-0.5	-0.2	-0.1	-1.0	-1.3	-0.5	-0.2
Birth weight										
5.5 pounds or less	93.3	0.3	-0.1	0.0	0.0	0.1	-0.3	0.0	0.0	0.3
Greater than 5.5 pounds	6.7	0.3	0.1	0.0	0.0	-0.1	0.3	0.0	0.0	-0.3
Child attending center-based program[1]										
Yes	52.6	0.8	0.9	0.3	0.8	0.6	1.1	0.4	1.0	0.8
No	47.4	0.8	-0.9	-0.3	-0.8	-0.6	-1.1	-0.4	-1.0	-0.8
Child ever attended center-based program[1]										
Yes	62.9	0.8	0.5	0.3	0.4	0.3	0.6	0.4	0.5	0.4
No	37.1	0.8	-0.5	-0.3	-0.4	-0.3	-0.6	-0.4	-0.5	-0.4
Attended center-based program prior to school[2]										
Yes	73.5	0.5	0.6	0.7	0.5	0.6	1.2	1.4	1.0	1.2
No	26.5	0.5	-0.6	-0.7	-0.5	-0.6	-1.2	-1.4	-1.0	-1.2
Women, Infant, and Children program participant[1]										
Yes	33.8	1.0	-0.6	-0.1	-0.8	-0.7	-0.6	-0.1	-0.8	-0.7
No	66.2	1.0	0.6	0.1	0.8	0.7	0.6	0.1	0.8	0.7
Free meal at school or center[2]										
Yes	35.8	0.6	-0.9	-1.1	-0.5	-0.5	-1.5	-1.8	-0.8	-0.8
No	64.2	0.6	0.9	1.1	0.5	0.5	1.5	1.8	0.8	0.8
Repeated kindergarten[3]										
Yes	5.7	0.4	-0.3	-0.5	-0.2	-0.2	-0.8	-1.3	-0.5	-0.5
No	94.3	0.4	0.3	0.5	0.2	0.2	0.7	1.3	0.5	0.5
School Safety and Discipline (SS&D) population										
Parental educational level										
Less than high school graduate	9.4	0.5	-1.2	-1.3	-0.3	-0.6	-2.4	-2.6	-0.6	-1.2
High school graduate or equivalent	32.7	0.6	0.3	0.0	-0.2	-0.6	0.5	0.0	-0.3	-1.0
Some college	57.9	0.5	0.9	1.3	0.5	1.1	1.8	2.6	1.0	2.2
Mother's employment status										
No mother in household	3.5	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Employed 35 hours/week or more	46.2	0.5	0.0	0.1	-0.1	0.1	0.0	0.2	-0.2	0.2
Employed less than 35 hours/week	20.3	0.5	0.1	0.0	0.0	-0.1	0.2	0.0	0.0	-0.2
Seeking employment	4.5	0.3	-0.2	-0.2	-0.2	-0.2	-0.7	-0.7	-0.7	-0.7
Not in labor force	25.5	0.5	0.0	0.1	0.2	0.2	0.0	0.2	0.4	0.4

Table 4
Estimated Reduction in Bias and Bias Ratio for Selected Characteristics of the NHES:93 – Concluded

Characteristic	Standard estimate		Estimated reduction in bias				Bias ratio			
	Estimate	Standard error	Scheme A1	Scheme A2	Scheme B1	Scheme B2	Scheme A1	Scheme A2	Scheme B1	Scheme B2
Father's employment status										
No father in household	26.8	0.6	-0.2	-0.2	-0.1	-0.2	-0.3	-0.3	-0.2	-0.3
Employed 35 hours/week or more	63.2	0.5	0.6	0.9	0.6	0.8	1.2	1.8	1.2	1.6
Employed less than 35 hours/week	3.1	0.2	-0.2	-0.2	-0.2	-0.2	-1.0	-1.0	-1.0	-1.0
Seeking employment	2.6	0.2	-0.2	-0.3	-0.2	-0.3	-1.0	-1.5	-1.0	-1.5
Not in labor force	4.3	0.3	-0.1	-0.1	-0.1	-0.1	-0.3	-0.3	-0.3	-0.3
School control										
Public	91.2	0.3	-0.1	-0.1	-0.1	-0.1	-0.3	-0.3	-0.3	-0.3
Private	8.8	0.3	0.1	0.1	0.1	0.1	0.3	0.3	0.3	0.3
Visitors required to sign in at school										
Yes	79.9	0.5	0.1	0.4	0.0	0.2	0.2	0.8	0.0	0.4
No	20.1	0.5	-0.1	-0.4	0.0	-0.2	-0.2	-0.8	0.0	-0.4
Had drug or alcohol ed program this year										
Yes	68.5	0.7	0.6	0.8	0.7	0.9	0.9	1.1	1.0	1.3
No	31.5	0.7	-0.6	-0.8	-0.7	-0.9	-0.9	-1.1	-1.0	-1.3
Students in fighting gangs at school[4]										
Yes	22.3	0.5	-0.3	-0.4	-0.3	-0.5	-0.6	-0.8	-0.6	-1.0
No	77.7	0.5	0.3	0.4	0.3	0.5	0.6	0.8	0.6	1.0
Ease of obtaining marijuana at school[4]										
Very or fairly easy	39.2	0.6	-0.2	-0.3	-0.2	-0.3	-0.3	-0.5	-0.3	-0.5
Hard	29.7	0.5	0.1	0.1	0.2	0.2	0.2	0.2	0.4	0.4
Nearly impossible	31.1	0.6	0.1	0.1	0.0	0.1	0.2	0.2	0.0	0.2
Fear of incident of crime at school										
None	66.1	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Fear of theft or robbery[5]	11.9	0.5	-0.1	-0.2	0.0	-0.2	-0.2	-0.4	0.0	-0.4
Fear of bullying or assault[5]	8.6	0.3	-0.1	-0.1	-0.1	-0.1	-0.3	-0.3	-0.3	-0.3
Fear of two or more types of incidents[5]	13.3	0.5	0.1	0.3	0.1	0.2	0.2	0.6	0.2	0.4
Knowledge of crime at school										
None	38.7	0.6	0.2	0.1	0.2	0.1	0.3	0.2	0.3	0.2
Fear of theft or robbery[5]	14.1	0.5	0.2	0.3	0.2	0.3	0.4	0.6	0.4	0.6
Fear of bullying or assault[5]	15.6	0.4	-0.5	-0.4	-0.4	-0.4	-1.3	-1.0	-1.0	-1.0
Fear of two or more types of incidents[5]	31.6	0.6	0.1	0.0	0.0	0.0	0.2	0.0	0.0	0.0
Victimization by crime										
Not victimized	73.0	0.5	0.3	0.2	0.3	0.2	0.6	0.4	0.6	0.4
Victim of theft or robbery[5]	10.9	0.3	-0.2	-0.1	-0.1	0.0	-0.7	-0.3	-0.3	0.0
Victim of bulling or assault[5]	8.9	0.3	-0.1	0.0	-0.2	-0.1	-0.3	0.0	-0.7	-0.3
Victim of two or more types of incidents[5]	7.2	0.3	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	-0.3
Witnessed crime at school										
None	63.8	0.8	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Witnessed robbery[6]	0.6	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Witnessed bulling or assault[6]	24.1	0.8	-0.3	-0.3	-0.3	-0.3	-0.4	-0.4	-0.4	-0.4
Witnessed two or more types of incidents	11.4	0.4	0.0	0.1	0.0	0.0	0.0	0.2	0.0	0.0

[1] Applies to preschoolers only.
[2] Applies to all children except preschoolers.
[3] Applies to children in primary school only.
[4] Applies to students in grades 6 through 12 only.
[5] For the fear of incident, knowledge of crime, and victimized by crime variables, the second response category is used if either theft or robbery was reported but not both, the third response category is used if either bullying or assault was reported but not both.
[6] This response category is used if either bullying or assault was reported, but not both, was reported.
Note: Percents may not add to 100 because of rounding.
Source: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey, spring 1993.

where $\hat{p}_s$ is the estimated proportion under the standard approach and b_a is the reduction in bias under scheme a. Because of the high correlation in the estimates of the bias from the four adjustment schemes, only the mean square errors for scheme A1 were computed. In Brick *et al.* (1996), the estimates using other schemes are shown to have negligible effects.

The mean square errors of the adjusted estimates are now contrasted with the variability in the standard NHES:93 estimates. The variance increase from adjusting the weights using the telephone service interruption data was expressed as a *VIF* in table 3. Multiplying the variance estimates of the standard estimates by the appropriate adjustment factor yields an approximate variance for the adjusted (presumably unbiased) estimates which are then compared to the mean square error of the standard estimates.

To aid in comparing the weighting procedures, ratios of the variance of the adjusted estimate to the mean square error for the standard estimate were tabulated (see Brick *et al.* 1996). The ratio is called the mean square ratio and can be written as

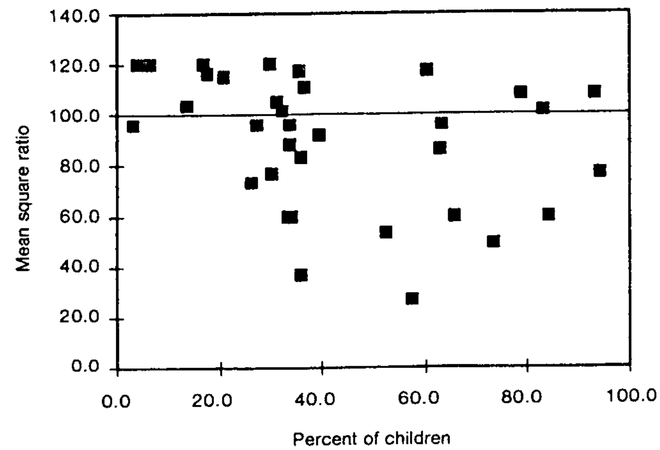
$$\mathrm{msr}_a(\hat{p}) = \frac{100 \times \mathrm{relative} VIF_a \times \mathrm{var}(\hat{p}_s)}{\mathrm{mse}_{A1}(\hat{p})}. \qquad (6)$$

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. As noted above, this simplification does not have much effect on the mean square ratios because the bias estimates are approximately the same across schemes.

The mean square ratios include contributions from the 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, standard 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.

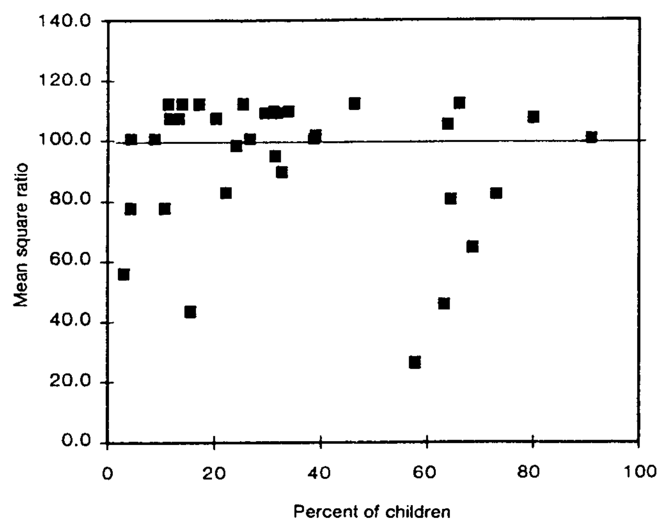
Figures 4 and 5 graphically present the msr for the two component surveys using scheme A1. In addition, Table 5 shows summary statistics for the msr for all four adjustment schemes. The distributions of mean square ratios for both components are very similar with the mean square ratios slightly lower for the SR component. The medians for schemes A1 and B1 (those based on interruptions of one week or more) are near the break-even point of 100. The means for these schemes are close to 90 and the figures confirm that the difference between the mean and medians is due to the skewed distributions of the mean square ratios.

A striking feature of the distributions of the mean square ratios for schemes A1 and B1 is the size of the ratios at the extremes of the distribution. The maximum mean square ratios for both components is 120, while some ratios are as small as 26. This means the maximum increase in the mean square error of the estimates is 20 percent, while the reductions in



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

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)

mean square error for a number of other estimates are quite large. Thus, the penalty associated with adjusting even when the estimate is not biased is modest, but the benefits of adjusting when it is needed are impressive.

The distributions for the mean square ratios for schemes A1 and B1 are very similar, and the choice of which of these schemes should be used may be determined by nonstatistical issues, such as availability of data and the other types of adjustments required in the survey. The mean square ratios show that the adjusted weights reduce the mean square error for about half the estimates considered below those derived from the standard weights. The distributions of the mean square ratios for schemes A2 and B2 (those based on interruptions of 1 month or more) have medians and means

Table 5
Summaries of Distribution of Mean Square Ratios for Selected Characteristics of School Readiness and School Safety and Discipline Components

	Adjustment scheme			
	A1	A2	B1	B2
School Readiness				
Mean	89.8	101.0	86.8	94.2
Median	96.0	108.0	92.8	100.8
Minimum	27.0	30.3	26.1	28.3
Maximum	120.0	135.0	116.0	126.0
School Safety and Discipline				
Mean	93.3	104.9	92.2	103.9
Median	100.8	113.4	99.9	112.5
Minimum	26.4	29.7	26.2	29.5
Maximum	112.0	126.0	111.0	125.0

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

that are greater than 100. Essentially, these mean square ratios are shifted upward when compared with those of schemes A1 and B1, and are not recommended.

5. CONCLUSIONS

If the percentage of the target population living in non-telephone 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 adjust the weights to reduce the coverage bias. In this study, the weights for persons in households reporting an interruption in telephone service were increased to account for those without telephones.

The bias reduction estimates computed under the assumed model showed that the coverage adjustments for the SR component improved some of the estimates substantially, and did not do much harm to any statistics. The bias reduction estimates from the SS&D component, on the other hand, were not as substantively important. The adjustments reduced bias for both components, but they also increased the variability of the estimates. The distributions of the mean square ratios show that about half the estimates could be improved using the telephone service interruption adjustments. Furthermore, even for those estimates that were less accurate due to the variance increases associated with the differential weights, the magnitude of the increases were not large. In other words, the penalty for adjusting when it did not reduce the coverage bias was not very great. These findings suggest that the adjustments should be seriously considered.

The alternative weighting schemes performed differently with respect to the mean square ratios. The schemes based on interruptions of telephone service of one week or more were better than the schemes based on interruptions of one 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 size of the sample is a relevant factor that should be taken into account when evaluating the use of the telephone service interruption adjustment. Bias ratios increase with the sample size because the bias is not affected while the sampling error of the estimate (the denominator of the bias ratio) decreases. Thus, the adjustments should be more beneficial in surveys with large sample sizes where the bias ratios might be expected to be large.

While the results of this study suggest that the adjustments could be useful for many estimates from telephone surveys, confirmation is needed before the adjustments are recommended. As discussed earlier, the estimates of the mean square errors in this study were based on the assumption that the adjusted estimates eliminated the bias of the estimates. This model assumption could not be verified because of the lack of benchmark data for comparison. The assumed model is very beneficial to the adjusted estimates in the sense that it results in lower bounds on the mean square errors for the adjusted estimates. Thus, the findings of this study should be taken as an indication that adjustment using data on interruptions in telephone service is a feasible method, but requires further study and evaluation.

The questions about interruptions in telephone service were recently added to the National Health Interview Survey, a survey conducted by the Census Bureau for the National Center for Health Statistics. The findings from this survey should be very useful in evaluating this method because the survey covers households without telephones by in-person interviews, eliminating the need for the critical model assumption used in this study.

ACKNOWLEDGEMENTS

The authors would like to thank the referees and editor for comments that substantially improved the methods and presentation of the material.

REFERENCES

BRICK, J., BURKE, J., and WEST, J. (1992). Telephone Undercoverage Bias of 14- to 21-year-olds and 3- to 5-year-olds. National Household Education Survey Technical Report No. 2, Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, NCES 92-101.

BRICK, J., KEETER, S., WAKSBERG, J., and BELL, B. (1996). Adjusting for Coverage Bias Using Telephone Service Interruption Data. National Household Education Survey Technical Report, Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, NCES 96-336.

COCHRAN, W. (1977). *Sampling Techniques*. New York: John Wiley and Sons, 12-15.

CONVERSE, J., and PRESSER, S. (1986). *Survey Questions, Handcrafting the Standardized Questionnaire*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-063. Beverly Hills: Sage Publishers.

FEDERAL COMMUNICATIONS COMMISSION (1988). Monitoring Report: CC Docket No. 87-339. Prepared by the Staff of the Federal-State Joint Board in CC Docket No. 80-286, Washington DC.

KALTON, G., and KASPRZYK, D. (1986). The treatment of missing survey data. *Survey Methodology*, 12, 1-16.

KEETER, S. (1995). Estimating noncoverage bias from a phone survey. *Public Opinion Quarterly*, 59, 196-217.

KISH, L. (1992). Weighting for unequal P_i. *Journal of Official Statistics*, 8, 183-200.

MASSEY, J., and BOTMAN, S. (1988). Weighting adjustments for random digit dialed surveys, Chapter 9 in *Telephone Survey Methodology*. Eds. Groves, Biemer, Lyberg, Massey, Nicholls, and Waksberg. New York: John Wiley and Sons, 143-160.

POLITZ, A., and SIMMONS, W. (1949). An attempt to get the 'not at homes' into the sample without callbacks. *Journal of the American Statistical Association*, 44, 9-31.

THORNBERRY, O., and MASSEY, J. (1988). Trends in United States telephone coverage across time and subgroups, Chapter 3 in *Telephone Survey Methodology*. Eds. Groves, Biemer, Lyberg, Massey, Nicholls, and Waksberg. New York: John Wiley and Sons, 25-50.