EVALUATION OF THE USE OF DATA ON INTERRUPTIONS IN TELEPHONE SERVICE

J. Michael Brick, Ismael Flores-Cervantes, Westat; Kevin Wang, The Urban Institute; and Tom Hankins, Westat
Ismael Flores-Cervantes, Westat, 1650 Research Boulevard, Rockville, Maryland 20850

Key Words: **RDD Survey, Telephone Sampling, Weighting Adjustment, Keeter Adjustment**

1. Introduction

The decision to use a telephone survey introduces coverage error due to the exclusion of persons in households without telephones. For estimates correlated with socioeconomic measures such as health insurance coverage, food security, and poverty, nontelephone households represent a larger proportion of the total population. As a result, the exclusion of nontelephone households can lead to larger biases. Nevertheless, telephone surveys are still attractive even when these types of items are important because the costs of telephone surveys are much lower than for in-person surveys.

One approach to addressing this problem was prompted by the research of Keeter (1995) using panel data from the 1992-93 Current Population Surveys (CPS). He notes that 42 percent of those not having a phone in 1992 had a phone one year later. Thus, over the course of a year, "transient" telephone households comprise a substantial percentage of nontelephone households. In addition, he reports that transient telephone households bear a closer resemblance in socioeconomic characteristics to nontelephone households than to telephone households which have not experienced an interruption in phone service. Keeter proposed using data from respondents about their telephone interruption experience to create a postsurvey weighting adjustment as a method of adjusting for the exclusion of nontelephone households.

This idea is developed and evaluated in work by Brick *et al.* (1996) and Frankel *et al.* (1998). In these analyses, data from telephone households with interruptions in service were adjusted to account for nontelephone households. However, both of these analyses were limited because no data were collected in the surveys from nontelephone households and the evaluations were based on assumptions that could not be verified.

This research uses data from the 1997 National Survey of America's Families (NSAF), a survey of the well-being of children and adults under the age of 65 and their families in 13 states and the balance of the nation. The thirteen focal states are Alabama, California, Colorado, Florida, Massachusetts, Michigan, Minnesota, Mississippi, New Jersey, New York, Texas, Washington and Wisconsin. These states account for about half of the total population of the United States and represent a broad array of state economic, demographic and political characteristics. The NSAF is the household survey component of Assessing the New Federalism (ANF), an Urban Institute project designed to examine the impact of recent shifts (devolution) of much of the responsibility from the federal level to state and local governments for programs designed to assist low-income families. The NSAF has large probability samples in each of the 13 focal states and produces reliable state and national estimates of child and adult well-being. Low-income families (defined as a family income below 200 percent of the federal poverty level) are of great interest because policy changes are expected to affect these families the most.

The NSAF uses a dual-frame sample design. A random digit dialing (RDD) frame of telephone households is supplemented with an area probability sample of nontelephone households. Interviews with nontelephone households were carried out using cellular phones carried by field staff. Thus, the NSAF data allow us to make comparisons between estimates based on all households and those based only on telephone households.

In NSAF, weights were constructed to produce estimates at the study area and national levels. The weights used both area and telephone samples and were adjusted for nonresponse at both the household and in-person levels. The nonresponse-adjusted weights were then raked using two dimensions defined by site/gender/race-ethnicity/age groups and region tenure. The variance of the estimates was computed using 60 replicate weights. The replicate weights were created using the same weighting procedures as applied to the full sample NSAF weight (see Brick *et al.* 1999, for a discussion of the weighting procedures). For this study, the variances are also computed using replication.

We begin by examining 1997 NSAF estimates of the percentage of nontelephone households and describe the characteristics of telephone households that experienced a service interruption, those that did not experience such interruptions and nontelephone households. The next section presents sets of estimates from NSAF data, including estimates of only the RDD sample with no adjustment for nontelephone

households and estimates with different adjustments for the exclusion of nontelephone households. The final section summarizes the findings.

2. Characteristics

Whenever units are excluded from a survey and the undercovered population has different characteristics than the covered population, coverage bias results. The bias increases with the undercoverage rate and the size of the difference between the covered and undercovered. The NSAF collected data about both children (under 18 years) and adults (18 to 64 years), but this analysis only studies children because the percentage of children in nontelephone households is greater than percentage of adults. In the NSAF, 6.6 percent (0.4%)[1] of children are in nontelephone households while for adults it is only 4.2 percent (0.2%). Thus, undercoverage bias is likely to be a greater concern for children.

The NSAF estimate of the percentage of children in nontelephone households is considerably lower than the estimate from the CPS. For example, the March 1997 CPS estimates that 8.2 percent (0.2%) of children are in nontelephone households while the NSAF estimates 6.6 percent (0.4%). For low-income children (less than 200% of the poverty level), CPS estimates 16.0 percent (0.5%) of children are in nontelephone households while NSAF estimates 14.5 percent (1.0%).

One possible reason for the difference between the CPS and NSAF is measurement error. Ferraro *et al.* (1998) describe how measurement error might cause the CPS to overestimate the percentage of nontelephone households. This has important implications for the application of the telephone interruption procedure and is discussed in the next section. However, some basic estimates of the percentages of persons in households with interruptions and their characteristics are presented first.

The 1997 NSAF estimates that 4.4 percent of children in telephone households had an interruption of at least 1-week, and 2.6 percent had an interruption of at least 1-month. For adults, the corresponding estimates are 2.3 percent and 1.3 percent. Nontelephone households in the NSAF were also asked about interruptions in service (this indicates if the household ever had phone service in the last year) and about 34 percent of the children in nontelephone households had an interruption and about 30 percent of the adults had an interruption.

[1] Standard error in parentheses.

Table 1 shows the estimated percentage of children with various characteristics by telephone and service interruption status. The first column shows the estimated percentage of children with the characteristic (e.g., 14.4% of all children are Hispanic). We call this the *standard* estimate. The other columns show the same estimate for different subgroups. For example, in telephone households without an interruption of 1-week or more, 13.5 percent of the children are Hispanic, while in households with at least a 1-week interruption 23.9 percent are Hispanic. The last column gives the estimate for those in nontelephone households.

The characteristics of children in households with interruptions are more similar to the nontelephone households than those in households with no interruptions. This relationship was postulated by Keeter (1995) and shown to hold in practice by Brick *et al.* (1996). This relationship is the basis for reducing the coverage bias using these data. Table 1 also shows that those having longer interruptions (1 month) often have characteristics more like nontelephone households than those with shorter interruptions (1 week). For example, for the percent of children in families with incomes of less than 200 percent of the poverty rate, the standard estimate is 42.7 percent, for those with at least a 1-week interruption it is 85.2 percent, for 1-month interruption it is 90.3 percent, while for nontelephone households it is 93.7 percent.

3. Estimation Strategy

This section examines the bias and mean square error (MSE) of estimates produced under five alternatives that do not involve the use of any data collected from nontelephone households. The first alternative is to use only data from telephone households. The data is adjusted for nonresponse and raked to the same control totals used for the 1997 NSAF without any modifications. This is called the *telephone-only* estimate and it does not reflect any special adjustment. The other four adjustments can be classified into two groups depending on the method used to carry out the Keeter-type adjustments. For each of these methods, two estimates were produced – one with 1-week service interruptions and the other with 1-month interruptions. In both methods, the idea is to adjust the weights for persons in telephone households with an interruption to represent both persons in households with an interruption in telephone service and those without telephone.

The details of the first method or *Method 1* are described in Brick *et al.* (1996). In this method, an additional adjustment is included during the development of the weights. The form of the

adjustment factor is $A_c = 1 + t_4 / \hat{t}_2$ where t_4 is the percentage of children in nontelephone households in cell c, and $\hat{t}_2$ is the estimated percentage of children in telephone households with an interruption in service in cell c. In surveys that do not collect data from nontelephone households (the desired use of this adjustment), the percentage of children in nontelephone households (t_4) cannot be obtained directly. In previous studies, data from the CPS is used to estimate t_4 (Brick *et al.*, 1996). However, as noted before, the CPS estimates may be biased and using it in this situation could result in larger biases. Therefore, for the evaluation of the methods, t_4 is computed directly using the standard NSAF weights. On the other hand, the proportion of children in households with an interruption in service ($\hat{t}_2$) is always estimated using the sample. The Method 1 Keeter adjustments for the 1-week and 1-month estimates used six weighting cells defined by race/ethnicity (Hispanic, black/nonHispanic, and nonblack/nonHispanic) and tenure (own and rent home). These cells were created within each of the 13 focal states and the balance of the U.S. separately. We collapsed small size cells or when large adjustment factors were produced.

For the second method or *Method 2*, the Keeter adjustments are made by raking with an additional dimension. The standard NSAF weights were raked to two dimensions defined by site/gender/race-ethnicity/age group and tenure. For Method 2 the additional control total domain values are $(t_4 + \hat{t}_2)\, CNT$ and $(1 - t_4 - \hat{t}_2) CNT$ where CNT is the total number of children in the site. To reflect the variability in $\hat{t}_2$ (computed using the sample), replicate estimates of $\hat{t}_2$ are computed to generate variable control totals to be used for each replicate. Then the weights for each replicate were raked to the corresponding dimensions. An advantage of this method is that estimates for t_4 and t_2 are only required at the site level rather than the cell level. Computing estimates for t_4 and t_2 can be difficult for small cells. As in method 1, the adjustment can lead to large weighting adjustment factors. The size of the factor can be controlled by collapsing larger areas or by using raking with constraints (Deville *et al.*, 1993).

Table 2 shows the estimates for several characteristics from the 1997 NSAF under these alternative estimation schemes (some of these variables are described in detail in Ehrle and Moore, 1999). The estimated bias and MSE are given for each approach. These estimates are discussed in the next section.

4. Findings and Conclusions

Table 2 shows the estimated percentages and means computed from the 1997 NSAF under the alternative estimation schemes. The characteristics chosen were mainly those that we thought would have a significant bias if nontelephone households were excluded. In addition to the estimates, the estimated percent bias (the absolute bias divided by the standard estimate) and the ratio of the MSE for the estimate to the MSE for the standard estimate are given for each scheme.

The percent biases of the telephone-only estimates are relatively large for some estimated percentages, but this is almost always because the estimates are small percentages and not because the absolute biases are large. Only four telephone-only estimates in the table have absolute biases of more than 2 percentage points. The percent biases for three estimates exceed 10 percent. For example, the standard estimate of the percentage of children in families that receive food stamps is 18.0 percent and the telephone-only estimate is 15.4 percent. This gives the largest absolute bias in the table, 2.6 percentage points, and a percent bias of 14.6 percent. For the means, the absolute and percent biases are uniformly small.

When the estimated percentages and percent biases are graphed (not shown), the relationship between the size of the estimate and the percent bias is clear. Estimates possessed by less than 25 percent of children have much larger and more variable percent biases than estimates possessed by a larger proportion of children.

The adjusted estimates using all four methods greatly reduce the size of both the absolute and percent bias, especially for items with large percent biases. This is consistent with the findings of Brick *et al.* (1996). The absolute biases for all the estimates in the table are less than 1 percentage point when method 1 with a 1-week period is used. The other schemes show similar results. Continuing the food stamp example, the method 1, 1-week adjusted estimate is 17.2 percent, which has an absolute bias of 0.8 percentage points and a percent bias under 5 percent.

The MSE findings are in line with those for the bias, mainly because the MSEs are dominated by the bias contributions. The largest MSE ratios for the telephone-only scheme are for those items with large percent biases. The adjustments, again for all four schemes do well, result in substantial reductions in the MSEs of the estimates. These results make it clear that the adjusted estimates are much better than the

telephone-only estimates for national estimates from the NSAF.

When the four approaches to the adjustments are compared, it is not clear that one is superior. All four reduce the biases for the estimates considerably. The use of the 1-month classification seems to do somewhat better in terms of MSE than the 1-week period, but the conclusion is not definitive. This contrasts somewhat with Brick *et al.* (1996), who found the 1-week time period was slightly superior because it had less effect on the variances of the estimates.

An interesting and important issue is whether the survey of nontelephone households is necessary and cost-effective when it is possible to make estimates with the Keeter adjustments. Clearly, a telephone-only sample costs substantially less than one with a nontelephone sample. Unfortunately, the answer to this question is not simple and several factors must be taken into consideration.

One critical feature is the size of the sample and the standard error of the estimates from the sample. In the NSAF, the national estimates are based on large samples and the telephone samples have small standard errors. As a result, the bias from excluding nontelephone households is the dominant factor in the MSEs of the estimates. At the state level, the samples are not as large, so both the bias and the variance of the estimates contribute to the MSEs. In this case, the Keeter-adjusted estimates are much more attractive. In general, as the telephone sample size increases and the standard errors of the estimates decrease, then the introduction of a nontelephone sample becomes more attractive than using a telephone-only sample with a Keeter adjustment.

A second critical feature is the type of estimate being produced in the survey. As noted above, most of the estimates in this paper are ones we believed would be subject to significant coverage bias due to the exclusion of nontelephone households. Other estimates would have smaller biases and MSEs. Even in the estimates we chose, the biases and MSEs were not large for many estimates, especially when the estimates from the Keeter-adjustments were examined. Therefore, a nontelephone sample is only attractive if the primary emphasis of the survey is on estimates of persons with low incomes or low SES. For most statistics for other groups, either a telephone-only estimate or a Keeter-adjusted estimate would be more cost effective.

These results suggest that a different approach might be considered to increase the efficiency of the NSAF sample design. Since the NSAF is very much devoted to producing very accurate estimates of the low-income population, telephone-only estimates are not very consistent with the survey objectives. At the state level, the NSAF has relatively large sample sizes, but the Keeter-adjusted estimates are competitive at this level. At the national level, the sample sizes are larger and standard errors of the estimates are smaller, so a nontelephone sample is more attractive. These findings suggest that an alternative design might consist of state-level telephone only samples along with a national nontelephone sample. The state estimates could be produced using a Keeter-adjustment and the national sample could have a direct estimate for nontelephone households. In this design, other options such as modeling the state-level nontelephone households using the national nontelephone sample might be entertained.

Reliable estimates of the percentage of nontelephone households for this adjustment method also requires additional research.

References

Brick, J.M., Shapiro, G., Flores-Cervantes, I., Ferraro, D., and Strickler, T. (1999). 1997 NSAF snapshot survey weights. Report No. 3, *NSAF Methodology Reports: Assessing the New Federalism*, The Urban Institute, Washington, DC.

Brick, J.M., Waksberg, J., and Keeter, S. (1996). Using data on interruptions in telephone service as coverage adjustments, *Survey Methodology*, 22(2), 185-197.

Deville, J.-C., Särndal, C.-E. and Sautory, O. (1993). Generalized raking procedures in survey sampling, *Journal of the American Statistical Association*, 88, 1013-1020.

Ehrle, J. and Moore, K.A. (1999). 1997 NSAF benchmarking measures of child and family well-being. Report No. 6, *NSAF Methodology Reports: Assessing the New Federalism*, The Urban Institute, Washington, DC.

Ferraro, D., Shapiro, G., Brick, J.M., and Strickler, T. (1998). Coverage in a sample designed for interviewing only nontelephone households, *Proceedings of the Survey Research Methods Section of the American Statistical Association.*

Frankel, M.R., Ezzati-Rice, T., Wright, R.A., and Srinath, K.P. (1998). Use of data on interruptions in telephone service for noncoverage adjustment, *Proceedings of the Survey Research Methods Section of the American Statistical Association.*

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

Table 1. Differences between estimate for all children and those with interruptions in telephone service

Characteristic of child	Standard estimate	Telephone households				Nontelephone households
		1-week interrupt		1-month interrupt		
		No	Yes	No	Yes	
Hispanic	14.4	13.5	23.9	13.6	25.5	21.4
Black, not Hispanic	14.6	13.1	28.8	13.4	28.4	26.7
Not black or Hispanic	71.0	73.5	47.3	73.1	46.1	51.9
Own home	64.0	68.4	29.2	67.8	26.0	26.3
Female	48.8	48.9	51.4	48.9	51.6	46.4
Proportions						
Family gets AFDC ('96)	10.1	8.5	33.9	8.7	38.6	58.4
Family gets food stamps ('96)	18.0	15.3	57.8	15.8	63.7	67.1
Free breakfast at school* ('96)	47.7	54.6	65.9	46.2	64.0	69.1
Free lunches at school* ('96)	60.3	58.1	78.0	58.8	76.7	78.2
Skip meal due to money	14.5	12.6	42.3	13.2	41.5	33.8
Not involved in activities*	16.8	15.8	33.6	16.0	36.1	37.4
Currently covered by Medicaid	17.8	15.5	51.3	15.9	56.8	59.5
Currently uninsured	11.9	11.3	20.0	11.6	18.9	24.5
Postpone care last year	9.7	9.3	15.5	9.4	15.3	14.2
ER or no usual source of care	6.0	5.8	9.0	5.9	9.3	12.4
Postpone dental care last year	6.2	5.9	9.7	6.0	9.3	8.6
Postpone medical care last year	3.1	2.9	5.3	3.0	4.6	4.4
Less than 200% poverty	42.7	39.8	85.2	40.3	90.3	93.7
Not confident in access to care	4.6	7.9	13.6	8.0	13.5	11.6
Fair/Poor health status	4.6	4.3	8.6	4.4	8.4	11.2
Negative outings for child*	17.1	16.2	28.4	16.4	29.8	24.8
Little read to last week*	20.6	20.0	27.0	20.2	26.2	30.2
Biological dad in household	65.0	68.0	36.8	67.6	34.4	37.8
MKA** or spouse works	89.6	91.2	66.3	91.0	62.1	59.7
Means						
MKA aggravation scale score	13.93	13.96	13.51	13.95	13.60	13.51
Number of heath care visits last year	3.61	3.59	3.97	3.59	4.13	3.53
Number of well child visits last year	1.19	1.16	1.67	1.16	1.81	1.59
Age 6-11 behavioral problems index score*	16.09	16.14	15.38	16.12	15.40	15.11
Age 12-17 behavioral problems index score*	15.87	15.93	14.97	15.91	14.98	15.20
Child's engagement in school scale	13.06	13.11	12.29	13.09	12.29	12.06
MKA mental health scale score	15.98	16.08	14.52	16.04	14.68	14.92

* These estimate were computed after subpopulation such as low income children or children of a certain age.

** Most knowledgeable adult.

Table 2. Estimated percent bias and MSE ratios with different adjustments for interruptions in telephone service

		Telephone-only			1-week adj.						1-month adj.					
					Method 1			Method 2			Method 1			Method 2		
	Standard estimate	Estimate	Percent bias	MSE ratio	Estimate	Percent bias	MSE ratio	Estimate	Percent bias	MSE Ratio	Estimate	Percent bias	MSE ratio	Estimate	Percent bias	MSE ratio
Proportions																
Family gets AFDC ('96)	10.1	8.4	-17.3	15.0	9.5	-6.4	2.2	9.3	-8.3	3.6	9.8	-3.1	0.9	9.6	-5.4	1.6
Family gets food stamps ('96)	18.0	15.4	-14.6	17.5	17.2	-4.6	1.6	16.9	-5.9	2.7	17.7	-1.9	0.6	17.4	-3.3	0.8
Free breakfast at school* ('96)	47.7	45.5	-4.6	5.5	47.1	-1.4	1.2	47.1	-1.2	1.2	47.1	-1.3	1.3	47.0	-1.5	1.4
Free lunches at school* ('96)	60.3	58.7	-2.7	4.3	60.2	-0.2	1.1	60.1	-0.3	1.1	60.3	0.0	1.1	60.1	-0.3	1.1
Skip meal due to money	14.5	13.6	-6.6	4.3	15.5	6.3	4.1	15.3	5.1	2.9	15.6	7.2	5.0	15.4	5.8	3.6
Not involved in activities*	16.8	15.9	-5.3	4.3	16.1	-4.2	2.8	16.2	-3.5	2.2	16.4	-2.8	1.6	16.4	-2.7	1.5
Currently covered by Medicaid	17.8	15.6	-12.5	28.6	17.0	-4.3	3.7	16.8	-5.3	5.4	17.4	-2.3	1.3	17.2	-3.2	2.2
Currently uninsured	11.9	11.3	-4.9	3.3	11.6	-2.7	1.5	11.6	-2.8	1.5	11.5	-3.5	2.1	11.5	-3.7	2.3
Postpone care last year	9.7	9.4	-2.4	1.1	9.8	1.1	1.1	9.8	1.0	1.0	9.8	1.7	1.1	9.8	1.5	1.0
ER or no usual source of care	6.0	5.8	-4.4	2.3	5.8	-4.2	2.0	5.8	-4.6	2.2	5.8	-4.4	2.1	5.8	-4.4	2.1
Postpone dental care last year	6.2	6.0	-2.1	0.8	6.2	0.2	1.1	6.2	0.4	1.0	6.1	-0.3	1.0	6.1	-0.3	1.0
Postpone medical care last year	3.1	3.0	-2.1	1.0	3.2	4.5	1.3	3.2	3.4	1.1	3.2	3.0	1.1	3.2	2.7	1.1
Less than 200% poverty	42.7	40.4	-5.4	19.1	41.8	-2.0	3.3	41.7	-2.3	4.2	42.1	-1.3	1.9	42.0	-1.6	2.4
Not confident in access to care	8.3	8.2	-0.7	1.1	8.6	4.6	2.6	8.6	3.9	2.1	8.7	5.7	3.3	8.7	5.3	3.1
Fair/Poor health status	4.6	4.2	-7.7	3.8	4.4	-3.1	1.0	4.4	-3.2	1.1	4.5	-2.2	1.1	4.4	-2.5	1.0
Negative outings for child*	17.1	16.7	-2.4	1.1	17.5	1.9	1.2	17.3	1.0	1.2	17.7	3.0	1.3	17.5	2.3	1.3
Little read to last week*	20.6	20.1	-2.1	1.1	20.4	-1.0	1.1	20.4	-0.7	1.1	20.3	-1.3	0.9	20.4	-0.8	0.9
Biological dad in household	66.0	67.3	1.9	6.7	66.2	0.2	1.4	66.3	0.4	1.4	66.0	-0.1	1.5	66.2	0.2	1.7
MKA** or spouse works	89.6	91.4	1.9	16.9	90.4	0.8	3.1	90.5	1.0	4.2	90.1	0.5	1.4	90.3	0.7	2.3
Means																
MKA aggravation scale score	13.9	14.0	0.2	1.6	13.9	0.0	1.3	13.9	0.0	1.3	13.9	0.1	1.3	13.9	0.1	1.5
Number of heath care visits last year	3.6	3.6	0.8	1.0	3.7	1.1	1.0	3.7	1.0	0.9	3.7	2.0	1.6	3.7	1.6	0.8
Number of well child visits last year	1.2	1.2	-1.2	0.8	1.2	-1.0	0.7	1.2	-0.9	0.6	1.2	-0.2	0.7	1.2	-0.3	0.6
Age 6-11 behavioral problems index score*	16.1	16.1	0.3	2.8	16.1	0.1	1.3	16.1	0.2	1.4	16.1	0.2	1.3	16.1	0.2	1.1
Age 12-17 behavioral problems index score*	15.9	15.9	0.1	1.0	15.9	-0.1	1.0	15.9	-0.1	0.9	15.8	-0.2	1.2	15.8	-0.2	1.0
Child's engagement in school scale	13.1	13.1	0.3	1.8	13.1	0.0	1.1	13.1	0.0	1.1	13.1	0.1	1.2	13.1	0.1	1.2
MKA mental health scale score	16.0	16.0	0.3	3.0	15.9	-0.3	3.1	15.9	-0.2	1.8	15.9	-0.3	2.7	15.9	-0.2	1.2

* These estimate were computed after subpopulation such as low income children or children of a certain age.
** Most knowledgeable adult.