Interrupted Telephone Service Adjustment (*ITSA*) in RDD Telephone Surveys: an assessment using the Behavioral Risk Factor Surveillance System (BRFSS)

Mansour Fahimi, Lily Trofimovich, Paul Levy, Henry Wells,
Ali Mokdad, Lina Balluz, William Garvin, and Machell Town

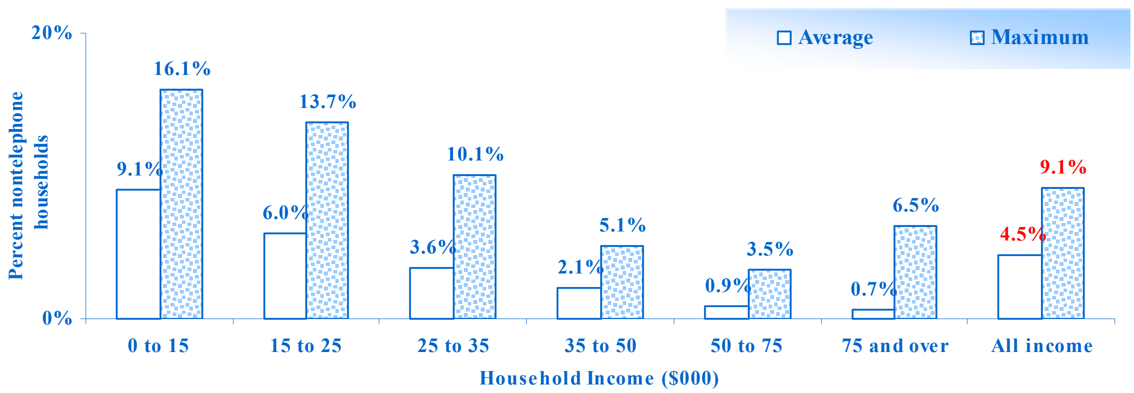
Mansour Fahimi, RTI International, 6110 Executive Blvd., Rockville, MD 20852 (fahimi@rti.org)

Key words: RDD, weighting, undercoverage, and variance inflation

1. Introduction

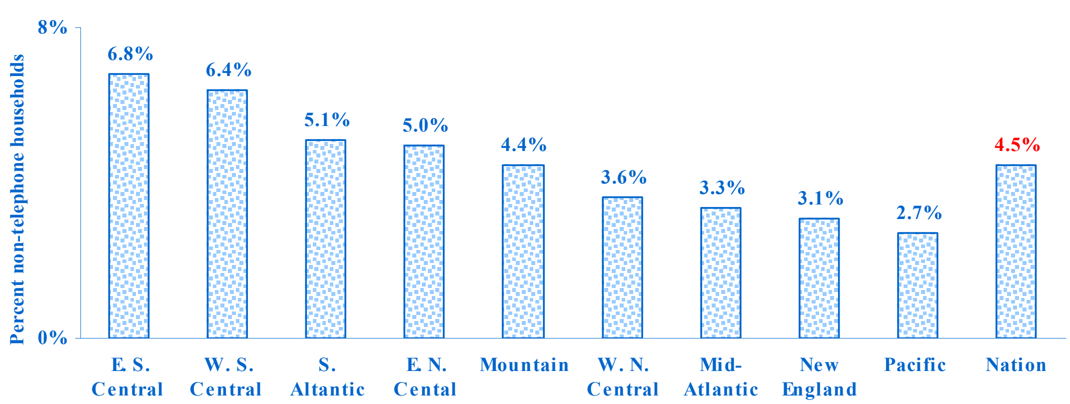
Exclusion of households without telephone is a source of systematic bias in telephone surveys. While lack of telephone service among households is about five percent nationally, this rate is higher among households of lower levels of socioeconomic status. The following chart shows the extent of this variation as a function of household income, using data from the March Supplement of Current Population Survey (CPS) 2002. Accordingly, in certain states (Alabama, Georgia, and Louisiana) the percent of nontelephone households for low to moderate income households can exceed 16 percent. On the national level, this percent across all income categories can exceed nine.

Chart 1. Average and maximum percent nontelephone households by household income (CPS 2002)



Moreover, as depicted in the following chart, this differential coverage pattern is also of note across different Census divisions, which is expected to be a correlate of household income.

Chart 2. Percent households without telephone by the Census division (CPS 2002)



Using the interruption in telephone service as a surrogate for lack of service, in this work a weighting adjustment methodology is used to compensate for non-telephone coverage in the 2003 Behavioral Risk Factors Surveillance System (BRFSS). Inherent to this methodology is the assumption that persons living in households with interrupted telephone service are similar to those living in households without a telephone. In this presentation, we discuss the history and rational for using such weighting adjustments, describe the procedure for developing Interrupted Telephone Service Adjustment (*ITSA*) weights for BRFSS, and evaluate its merits. Specifically, we will examine the extent of variance inflation that results from applying this additional layer of adjustment and contrast that against the potential gains in bias reduction on a number of key BRFSS outcome measures.

2. Background

Coordinated by the Centers for Disease Control and Prevention and all state health departments, BRFSS is the world's largest telephone survey. Each month, state health departments conduct telephone surveys of randomly selected, noninstitutionalized adults who are18 years and older, to obtain data on personal behaviors associated with increased risk for chronic disease and other health-related factors. Monthly data are aggregated to provide prevalence of risk behaviors and preventive health practices on an annual basis. The results are used by public health officials to determine the problem areas in their states, to develop prevention policies and intervention strategies, and to evaluate success in reducing prevalence of behaviors that endanger public health.

The BRFSS uses state-based random digit-dialed (RDD) samples in which all households with telephones have a known, non-zero probability of inclusion in the study. Currently, all states employ disproportionate stratified sampling in which listed residential telephone numbers are sampled at a higher rate than not-listed residential telephone numbers. One adult is then chosen at random to be interviewed from each selected household. Annual sample sizes vary from state to state, although in all states sample sizes are adequate to provide reasonable estimates for key outcome measures. In 2002, the sample sizes ranged from 2,408 in the District of Columbia to 13,491 in Pennsylvania, with a median state sample size of 4,401.

Survey data from BRFSS are weighted to reflect the employed sample design, as well as to compensate for differential nonresponse and undercoverage. Briefly, base weights are constructed and adjusted to account for oversampling of listed telephone numbers, households with multiple telephone lines, and subsampling of adults in each household. Moreover, the resulting base weights are ratio-adjusted (post-stratified) to CPS counts within age, gender, and race/ethnicity cells for each state. As detailed in the next section, these weights will be used as the starting weights for the *ITSA* weighting and to produce benchmark estimates for comparisons of the two weighting methodologies.

Literature promoting an *ITSA* type procedure first appeared in an article by Keeter (1995), leading to a number of recommended weighting adjustment methods. While research conducted by Frankel et. al. (1998) indicates that adjusting the weights of individuals from households with interruptions in telephone service can reduce the bias due to telephone noncoverage, other research conducted by Brick, et. al. (1996) seems to offer a less enthusiastic support for the general utility of this type of adjustment for all RDD surveys. That is, the resulting inflation in variance of survey estimates due to this adjustment can outweigh the gain in bias reduction for many estimates. However, there are consistent endorsements for using an adjustment of this form for survey estimates that are related to socioeconomic status of individuals where the coverage problem is most pronounced. This uniquely qualifies BRFSS as a candidate for which this type of weighting adjustment might improve the validity and reliability of many of its generated estimates, since this survey focuses on risk factors and prevalence estimates for disease, both of which are highly related to the socioeconomic status of individuals.

3. Methodology

The methodology detailed here is a variant of that suggested by Frankel et. al. (2003), which involves increasing the weights attached to individuals from households with interruptions in telephone service to account for non-telephone households. For illustration purposes, use will be made of the BRFSS:03 survey data to assess the bias reduction capability of this methodology against the potential loss in precision that will result from variance inflations. In order to describe the computational details, the following notations will be used throughout this section. It should be noted that the current methodology is explained for application within income categories, however, due to a high rate of missing income values for this investigation it was decided to collapse over all income categories and apply this method simply at the sate level.

w_{ijk}^{B}: base weight for the k^{th} respondent in the i^{th} state and j^{th} household income category

w_{ijk}^{P}: post-stratified weight for the k^{th} respondent in the ij^{th} cell based on the 2003 methodology

n_{ij}: number of respondents in the ij^{th} cell

I_{ij}: subset of respondents in the ij^{th} cell with telephone service interruptions

N_{ij}^{T}: number of adults in telephone households in the ij^{th} cell

$N_{ij}^{\bar{T}}$: number of adults in non-telephone households in the ij^{th} cell

N_{ij}^{I} number of adults in households with telephone service interruptions in the ij^{th} cell

$N_{ij}^{\bar{I}}$: number of adults in households with no telephone service interruptions in the ij^{th} cell

Unfortunately, none of the above four population numbers are readily known. However, the first two can be estimated from the CPS, while the latter two can be estimated from the survey data. In that sense, let $\hat{N}_{ij}^{T}$ and $\hat{N}_{ij}^{\bar{T}}$ be the estimates of N_{ij}^{T} and $N_{ij}^{\bar{T}}$ based on the CPS, respectively. The other two population numbers, N_{ij}^{I} and $N_{ij}^{\bar{I}}$, can be estimated by $\hat{N}_{ij}^{I}$ and $\hat{N}_{ij}^{\bar{I}}$ as follows.

$$N_{ij}^{\bar{I}} \cong \hat{N}_{ij}^{\bar{I}} = \hat{N}_{ij}^{T}\left(1-\frac{\sum_{k\in I_{ij}} w_{ijk}^{P}}{\sum_{k=1}^{n_{ij}} w_{ijk}^{P}}\right) = \hat{N}_{ij}^{T} \times \frac{\sum_{k\notin I_{ij}} w_{ijk}^{P}}{\sum_{k=1}^{n_{ij}} w_{ijk}^{P}} \qquad (1)$$

$$N_{ij}^{I} \cong \hat{N}_{ij}^{I} = \hat{N}_{ij}^{T}\left(\frac{\sum_{k\in I_{ij}} w_{ijk}^{P}}{\sum_{k=1}^{n_{ij}} w_{ijk}^{P}}\right) = \hat{N}_{ij}^{T} - \hat{N}_{ij}^{\bar{I}} \qquad (2)$$

The *ITSA* weighting procedure consists of two main steps. In the first step, the design weight of all respondents in the ij^{th} cell, w_{ijk}^{B}, are adjusted for telephone non-coverage to obtain the adjusted weights, according to the following:

$$w_{ijk}^{A} = \begin{cases} w_{ijk}^{B} \times \dfrac{\hat{N}_{ij}^{I} + \hat{N}_{ij}^{\bar{T}}}{\sum_{k \in I_{ij}} w_{ijk}^{B}}, \forall k \in I_{ij} \\ w_{ijk}^{B} \times \dfrac{\hat{N}_{ij}^{\bar{I}}}{\sum_{k \notin I_{ij}} w_{ijk}^{B}}, \forall k \notin I_{ij} \end{cases} \tag{3}$$

In the second step, the resulting telephone non-coverage adjusted weights go through the existing post-stratification process to be adjusted to CPS counts within age, gender, and race/ethnicity cells for each state to produce the final *ITSA* weights.

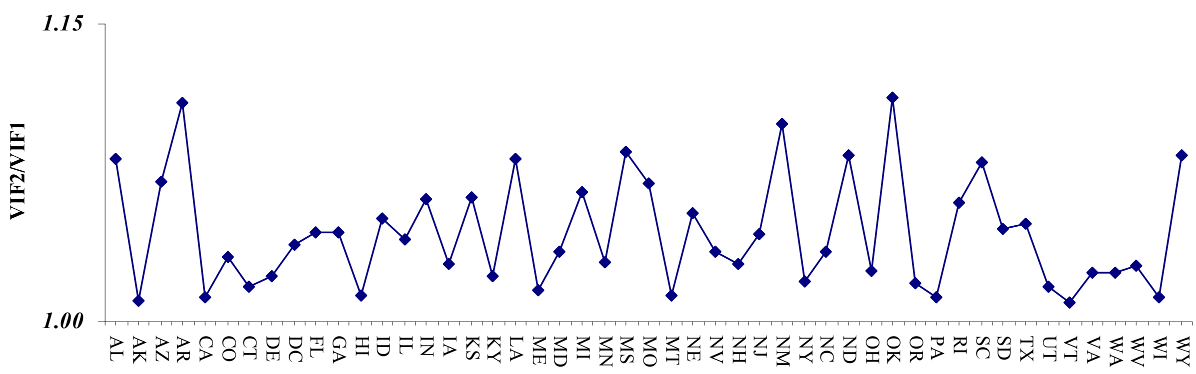
4. Evaluation

As stated earlier, a potential downside of this methodology is that the resulting weights will be subject to a larger variability as compared to those obtained under the standard methodology. However, the hope is that the reductions in bias will more than offset inflation in variances. With w_i representing the final *ITSA* weight of the i^{th} respondent, the inflation in variances due to unequal weighting effect (UWE) can be approximated by:

$$UWE = 1 + \left[CV(w_i)\right]^2 = 1 + \frac{\sum_i \frac{(w_i - \overline{w})^2}{n-1}}{\overline{w}^2} \tag{4}$$

The following chart shows the ratio of UWE under the *ITSA* weighting to that under the standard weighting methodology for each state, using the 2003 survey data. As expected, the variability in weights has increased, as all calculated ratios are greater than one. It is worth noting that the average UWE across all states for the standard methodology is 1.63 while that for the *ITSA* methodology is 1.70, with an average ratio of 1.04.

Chart 3. UWE ratios under *ITSA* and standard weighting methodologies for key outcome measures by state



A more telling indicator, however, that takes into account the negative effect of variance inflation as well as the positive gain due to bias reduction for this situation is the so-called Mean Square Ratio (MSR).

This metric is defined as the ratio of the mean square error of a point estimate under the two weighting methodologies. That is, for a given point estimate of percentage, $\hat{p}$, its MSR will be calculated by:

$$MSR(\hat{p}) = \frac{MSE(\hat{p}_{ITSA})}{MSE(\hat{p}_{Old})} \tag{5}$$

In the above, $MSE(\hat{p}_{ITSA})$ and $MSE(\hat{p}_{Old})$ represent the mean square error of $\hat{p}$ obtained under the *ITSA* and standard (old) weighting methodologies, respectively. Furthermore, assuming that the bias reduction under the *ITSA* methodology will render the resulting point estimate unbiased, the above metric will reduce to:

$$MSR(\hat{p}) = \frac{V(\hat{p}_{ITSA})}{V(\hat{p}_S) + (\hat{p}_S - \hat{p}_{ITSA})^2} \tag{6}$$

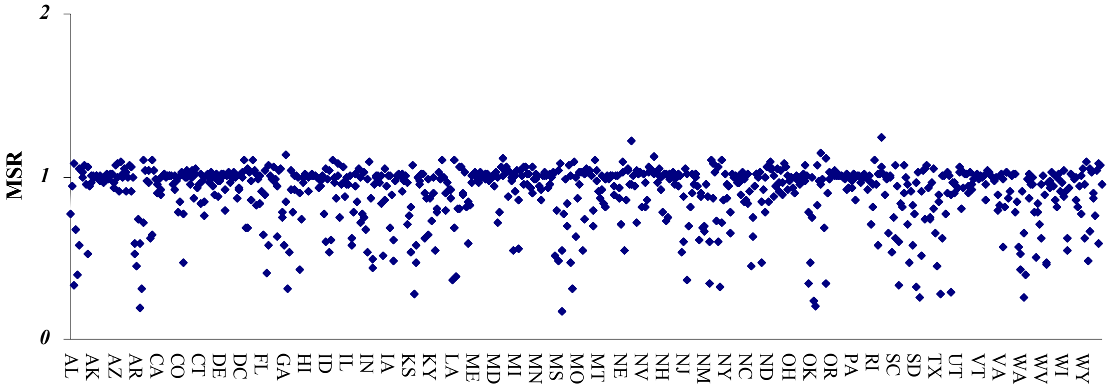
In order to evaluate the impact of *ITSA* weighting on variances, the following 17 key outcome measures were included in this investigation. For each measure, state level point estimates (prevalence) and associated standard errors were generated based on the weights calculated under the standard and *ITSA* methodologies. The needed standard errors were calculated using SUDAAN via the linearization method.

Table 1. Outcome measures included in the evaluation of *ITSA* weighting methodology

Outcome Measure	Variable
General health status	*GOODHLTH*
Any kind of health care coverage	*HLTHCOV*
Cost prevented dr. Visit, past 12 months	*COSTPREV*
Any exercise, past month	*EXERCISE*
Diagnosed diabetes, excluding pregnancy	*EVERDIAB*
Diagnosed high blood pressure, excluding pregnancy	*EVERBP*
Diagnosed high blood pressure, excluding pregnancy, and currently taking medicine	*CURBPMED*
Ever had blood cholesterol checked	*BP*
Currently trying to lose weight	*CURLOSEW*
Currently have asthma, Dr. Diagnosed	*CURASTH*
Had flu shot, past 12 months	*FLU*
Ever had pneumonia shot, people 65+	*PNEUM*
Current smoking status	*CURSMK*
Obesity	*OBESE*
Binge drinking	*BINGE*
Ever tested for HIV, excluding tested when donating blood, people 18-64	*HIVTEST*
Any activities limited due to physical, mental, or emotional problems	*LIMACT*

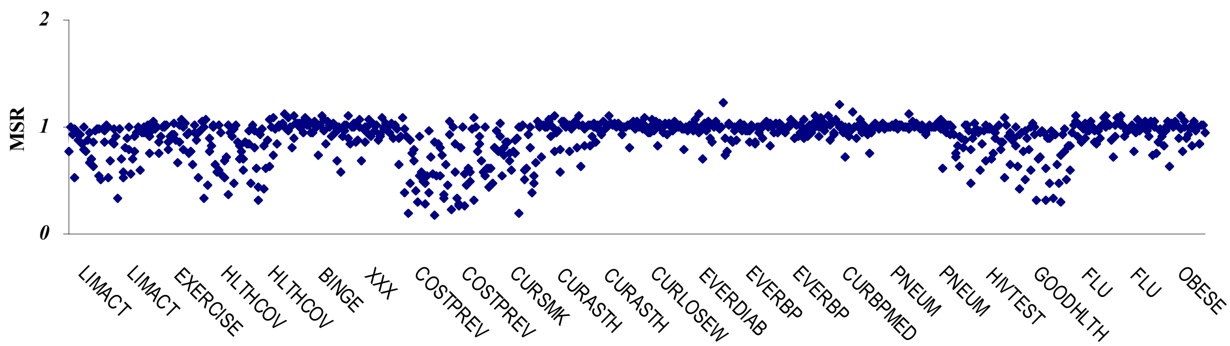
The following chart shows the MSR for the prevalence estimate for each of the above outcome measures. As seen, for the majority of these estimates their associated MSR tend to be smaller than one. Specifically, the average MSR across all estimates in the nation is 0.9, indicating that overall there appears to be a positive net gain in bias reduction in comparison to the inflation that results in variances due to application of the *ITSA* weighting methodology.

Chart 4. Mean Square Ratio (MSR) of key outcome measures by state



Moreover, the above examination conducted for each outcome measure reveals a more segmented impact. As seen in the following chart the MSR for the prevalence estimate of certain outcome measures tend to be smaller than the others. For instance, outcome measures that are known to be highly correlated with income, such as current smoking status, lack of doctor visit due to cost, and health care coverage seem to benefit significantly from this type of weighting adjustment.

Chart 5. Mean Square Ratio (MSR) by key outcome measure across states



5. Conclusion

Virtually, all surveys aim to provide reliable estimates for a host of population parameters germane to the particular subject matter under the study. Much like the design of a survey, where it is impractical to develop a sampling plan that can accommodate all precision and cost constraints without any compromise, post-survey data enhancement procedures are subject to trade-offs as well. Oftentimes, survey data are ratio-adjusted to known population totals with the hope that a sizable amount of bias – be it due to undercoverage or differential nonresponse – could be eliminated. This gain, however, is always achieved at the expense of added variability in the survey estimates. While for certain key estimates this tradeoff can easily be justified, for other estimates the potential gain in bias reduction can be trivialized in comparison to the variance inflation resulting from the imposed adjustments.

In line with previous investigations, for survey results that are influenced by the socioeconomic status of respondents, this work suggests that a refined weighting methodology similar to what is outlined here can reduce the bias due to lack of telephone coverage without increasing the variance of estimates significantly. As seen in Chart 3, the state level mean square ratio for most key estimates tend to be less than 1, suggesting a positive net gain in precision of estimates due to application of the *ITSA* weighting adjustment. Chart 4 indicates that this trend holds for a large number of estimates, with certain outcome measures receiving a sizeable benefit from this adjustment.

As mentioned earlier, due to a high prevalence of missing values for household income it was deemed unadvisable to conduct the *ITSA* weighting in conjunction with household income in each state. Instead, the needed adjustments were applied at the state level irrespective of household income. Knowing that this methodology is most effective when the socioeconomic status of respondents is taken into consideration, future applications of this methodology to the BRFSS data should be carried out within weighting cells indexed by categories of household income. Alternatively, one can use educational attainment categories to develop adjustment cells, since in most surveys this item is subject to a lesser rate of missing values as compared to household income.

Finally, it should be noted that there are other sources of bias for RDD surveys, such as the emerging prevalence of cell-only households. Since such households are typically excluded from RDD sampling frames – much like nontelephone households – it is important to develop comprehensive adjustment procedures that can address all sources of bias, to the extent possible.

6. References

Keeter, S. (1995). Estimating Noncoverage Bias from a Phone Survey. Public Opinion Quarterly, 59, 196-217.

Frankel, M., Ezzati-Rice, T., Wright, R., and Srinath, K.P. (1998) “Use of Data on Interruptions in telephone Service for Noncoverage Adjustment,” *1998 Proceedings of the Section on Survey Research Methods,* Alexandria, VA: American Statistical Association, pp. 290-295.

Brick, J., Waksberg, J., and Keeter, S. (1996). Using Data on Interruptions in Telephone Service as Coverage Adjustments, Survey Methodology, Vol. 22. Pp. 185-197.

Martin R. Frankel, K.P. Srinath, David C. Hoaglin, Michael P. Battaglia, Philip J. Smith, Robert A. Wright, and Meena Khare (2003). Adjustments for non-telephone bias in random-digit-dialing surveys. Statistics in Medicine Statist. Med. 2003; 22:1611–1626.