Dual-Frame Weights (Landline and Cell) for the 2009 Minnesota Health Access Survey

Kanru Xia[1], Steven Pedlow[1], Michael Davern[1]
[1]NORC/University of Chicago, 55 E. Monroe Suite 2000, Chicago, IL 60603

Abstract
In recent years, the increasing undercoverage of random-digit-dial (RDD) landline frames has driven surveys to employ landline plus cell phone dual-frame designs. The 2009 Minnesota Health Access Survey, a large-scale health insurance survey conducted jointly by the Minnesota Department of Health and the University of Minnesota, collected 9,811 landline interviews and 2,220 interviews via a cell phone sample (regardless of landline/cell phone usage). This paper compares weights produced under four different screening strategies (RDD-only, RDD plus cell-only, RDD plus cell-only/cell-mostly, and RDD plus cell-any) and five different weighting adjustments for combining landline and cell-phone interviews. Results show the lowest mean-squared errors result from using all interviews with an effective sample size adjustment factor.

Key Words: Composite Weighting

1. Introduction

The Minnesota Health Access Survey (MNHA), a general population health insurance survey conducted jointly by the Minnesota Department of Health and the University of Minnesota's State Health Access Data Assistance Centre (SHADAC), tracks health access in the state of Minnesota. In 2009, the MNHA began collecting interviews via a cell phone frame regardless of landline/cell phone usage in addition to landline random-digit-dial interviews.

This paper compares twelve different weight calculations for the 2009 MNHA. In particular, it focuses on two weighting choices: whether to allow a full or partial overlap in the landline and cell phone sample frames, and what weighting adjustment factor to use in any overlap.

2. Weighting Choices

2.1 Sample Frame Choices

Although the 2009 MNHA considered all cell phone respondents to be eligible, regardless of their landline and cell phone usage, all respondents from either frame were asked whether they had access to both a landline and a cell phone. If they had access to both, they were asked whether they used mostly their landlines or mostly their cell phones. Therefore, the landline respondents could be categorized into: landline-only (no cell phone), cell-mostly (they rarely make or answer calls on their landline), or landline-mostly. Meanwhile, the cell phone respondents could be categorized into: cell-only (no landline phone), cell-mostly, or landline-mostly. Table 1 shows the sample sizes in each category.

Table 1: Sample Sizes by Sample and Phone Status.

Phone Status	*Sample Type*		*Sample Sizes*		
	Landline	*Cell*	*Landline*	*Cell*	*TOTAL*
Landline-only	√		1,967		1,967
Cell-only		√		890	890
Cell-mostly	√	√	1,283	342	1,625
Landline-Mostly	√	√	6,561	988	7,549
TOTAL			9,811	2,220	12,031

In considering the weights for the 2009 MNHA, we first examined how much overlap to allow in the landline and cell phone sample frames. We considered four possible choices:

1) All landline respondents but no cell phone respondents (no cell phone coverage, no overlap)
2) All landline respondents plus cell-only cell phone respondents (full coverage, no overlap)
3) All landline respondents plus cell-only and cell-mostly cell phone respondents (full coverage, overlap in the cell-mostly category)
4) All landline respondents plus all cell phone respondents (full coverage, overlap in the cell-mostly and landline-mostly categories)

2.2 Overlap Adjustment Options

When a sample frame overlap happens, the households in the overlap are double-covered since they could have been selected from the landline frame or the cell phone frame. The sum of the weights for these households in each frame will estimate the number of such households in Minnesota. Without a weight adjustment, our estimate of households in the overlap will be double the appropriate estimate.

The most common method of adjusting weights in sample overlaps is to multiply the weights in one sample (landline in this case) by a weighting adjustment factor λ, and multiply the weights in the other sample (cell phone in this case) by $(1-\lambda)$. This is often referred to as composite weighting and corrects the double-counting problem. Even though respondents in the overlap still have a chance to be in either the landline or cell phone samples, their weights are corrected so that their phone usage category is not over-represented through the weights.

Unfortunately, it is often the case that we have limited information about the actual size of the overlap population. For the MNHA, we have population totals by phone usage category for the Midwest (from the National Health Interview Survey), but not for the state of Minnesota alone. Therefore, there are many possible ways to calculate a weighting adjustment factor λ, and the literature suggests that different methods will be optimal for different projects.

For the 2009 MNHA, we used 5 methods to calculate λ, and we calculated separate λ's for the cell-mostly and landline-mostly categories when they are both part of the overlap in sample frame choice 4.

2.2.1 Option 1: Set λ=0.5

This is the simplest assumption possible. This method assumes that, like a household with two landline telephones, a household with a landline phone and a cell phone should have its weight divided by two because of the doubled chances of inclusion. However, this assumption will lead to inefficiency if the probability of selection is much different for an overlap case in the landline and cell phone samples.

2.2.2 Option 2: Sample Size

This method calculates λ proportional to the sample size in the landline sample. The logic of this choice is that the landline and cell phone surveys are both trying to represent the same population of households. Ignoring undercoverage (of cell-only households in the landline frame and landline-only households in the cell phone frame), if the landline and cell phone samples have different sizes, the sample with the smaller sample size will tend to have larger weights. This method will reduce the larger weights from the smaller sample more than the smaller weights from the larger sample. In the 2009 MNHA, the landline sample size is larger so the λ's are all greater than 0.5.

Let

n_L be the landline sample size in the overlap

n_C be the cell phone sample size in the overlap

Then

$$\lambda = \frac{n_L}{n_C + n_L}$$

2.2.3 Option 3: Effective Sample Size

This option adds another consideration to option 2, the variance in weights among the two samples being combined. Option 2 assumes that if the sample sizes are even, the optimal λ is 0.5, but this option considers the estimated information gained from each sample rather than just the sample size. When the weights are more variable, the information is less per completed interview. This option adjusts λ to measure the estimated information gained from each sample rather than just the sample size. This method calculates λ proportional to the estimated effective sample size in the landline sample (O'Muircheartaigh and Pedlow 2002).

Let

W_i be the person weight prior to overlap adjustment for interview *i*

d_L be the design effect due to weighting in the landline sample

d_C be the design effect due to weighting in the cell phone sample

Then

$$\lambda = \frac{n_L/d_L}{n_C/d_C + n_L/d_L}$$

And d_L and d_C can be estimated by

$$\hat{d}_L = 1 + [CV(W_i \in landline\ RDD)]^2$$
$$\hat{d}_C = 1 + [CV(W_i \in cell\ phone)]^2,$$

where CV is the coefficient of variation, which is the standard deviation of the weights divided by the mean of the weights.

2.2.4 Option 4: Weighted Sample Size

In this option, we use the population represented by the sample (the sum of the weights) rather than the actual sample size to determine λ. The motivation for this method is that the coverage might be different between the two samples, and the coverage can be approximated by the sum of the weights. The sample with the larger representation is then given more weight in the calculation of λ. This method calculates λ proportional to the weighted sample size in the landline sample.

Let

$\sum_L W_i$ be the sum of person weight over landline sample

$\sum_C W_i$ be the sum of person weight over cell phone sample

Then

$$\lambda = \frac{\sum_L W_i}{\sum_C W_i + \sum_L W_i}$$

2.2.5 Option 5: Bias Correcting

The idea for this overlap adjustment method is that the landline sample is more likely to find landline-mostly households than cell-mostly households, whereas the cell phone sample is more likely to find cell-mostly households than landline-mostly households. This method first adjusts the sample sizes within sample to match assumed population percentages (of landline-mostly and cell-mostly households). Then, based on these revised sample sizes, we apply the “sample size” sample combination method. So, λ is proportional to the sample size in the landline sample but we correct the sample sizes within category by the ratio of true proportions of phone usage categories to the observed proportions.

Let

P_{uL} be the assumed proportion of the landline frame that is in phone usage category *u*

P_{uC} be the assumed proportion of cell phone frame that is in phone usage category *u*

p_{uL} be the observed proportion of the landline frame that is in phone usage category *u*

p_{uC} be the observed proportion of the cell phone frame that is in phone usage category *u*

Then

$$\lambda_u = \frac{\left(\frac{P_{uL}}{p_{uL}}\right) * n_L}{\left(\frac{P_{uL}}{p_{uL}}\right) * n_L + \left(\frac{P_{uC}}{p_{uC}}\right) * n_C}$$

3. Weight Evaluation

Combining the four sample frame choices and the five overlap adjustment choices produced twelve weighting options (see Table 2). The construction of these weights involved the following steps: 1) compute the person weights for landline sample; 2) compute the person weights for cell phone sample; 3) combine the person weights from both samples; 4) calculate and apply the appropriate weighting adjustment λ to the person weights; and 5) post-stratify.

Table 2: Weights Produced by Combining Sample Frames and Overlap Adjustments.

Weight	*Sample Frame*	*Sample Combination Method*
1	landline + no cell	N/A
2	landline + cell-only	N/A
3	landline + cell-only an mostly	0.5
4	landline + cell-only and mostly	sample size
5	landline + cell-only and mostly	effective sample size
6	landline + cell-only and mostly	weighted sample size
7	landline + cell-only and mostly	bias-correcting
8	landline + all cell	0.5
9	landline + all cell	sample size
10	landline + all cell	effective sample size
11	landline + all cell	weighted sample size
12	landline + all cell	bias-correcting

3.1 Weight Properties

Table 3 shows the calculated λ values for the third and fourth sample frame choices. For both sample frame choices, three of the λ values (sample size, effective sample size, and bias-correcting) are very similar, ranging from 0.79 to 0.81 for the cell-mostly and 0.82 to 0.87 for the landline-mostly. These λ values reduce cell phone weights much more than landline weights. The weighted sample size is the only λ that reduces landline weights more than cell phone weights.

Table 4 shows a summary of the twelve weights after overlap adjustment, and Table 5 shows the same twelve weights after post-stratification. The number of observations depends on the choice of sample frame. Both tables show that the number of observations increases as we accept more of the cell phone cases into the sample. After post-stratification, the means are all the same for the same sample frame, but the standard deviations of the weights differ. In particular, the standard deviations of the weighted sample size method are largest, while the effective sample size method has the smallest standard deviations (though the two other sample combination methods with similar λ's have standard deviations that are almost as small). In general, the smaller the standard deviation among the weights, the smaller the resulting variance (or standard error or confidence interval) in estimates will be.

Table 3: λ Values for the Third and Fourth Sample Frames.

Weight	*Sample Frame*	*Weighting Adjustment*	*Phone Category*	*λ value*
3	landline + cell-only and mostly	0.5	Cell Mostly	0.50
4	landline + cell-only and mostly	sample size	Cell Mostly	0.79
5	landline + cell-only and mostly	effective sample size	Cell Mostly	0.81
6	landline + cell-only and mostly	weighted sample size	Cell Mostly	0.35
7	landline + cell-only and mostly	bias-correcting	Cell Mostly	0.79
8	landline + all cell	0.5	Cell Mostly	0.50
8	landline + all cell	0.5	LL Mostly	0.50
9	landline + all cell	sample size	Cell Mostly	0.79
9	landline + all cell	sample size	LL Mostly	0.87
10	landline + all cell	effective sample size	Cell Mostly	0.81
10	landline + all cell	effective sample size	LL Mostly	0.86
11	landline + all cell	weighted sample size	Cell Mostly	0.35
11	landline + all cell	weighted sample size	LL Mostly	0.45
12	landline + all cell	bias-correcting	Cell Mostly	0.79
12	landline + all cell	bias-correcting	LL Mostly	0.82

Table 4: Weights after Overlap Adjustment**.**

Weight	*Observations*	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*
1	9,811	517.0	437.9	6.3	1924.1
2	10,701	830.3	1672.3	6.3	83250.0
3	11,043	836.4	1698.2	3.5	83250.0
4	11,043	818.0	1651.8	5.6	83250.0
5	11,043	816.9	1650.7	5.7	83250.0
6	11,043	846.1	1743.7	2.4	83250.0
7	11,043	817.7	1651.5	5.6	83250.0
8	12,031	801.2	1757.7	3.2	83250.0
9	12,031	759.6	1587.1	5.5	83250.0
10	12,031	758.9	1586.3	5.5	83250.0
11	12,031	813.5	1830.6	2.4	83250.0
12	12,031	762.9	1591.1	5.2	83250.0

Table 5: Post-Stratified Weights.

Weight	*Observations*	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*
1	9,811	517.0	575.8	3.9	9460.4
2	10,701	474.0	562.5	3.9	26344.8
3	11,043	459.3	585.9	1.7	25518.7
4	11,043	459.3	546.7	3.2	25878.6
5	11,043	459.3	545.7	3.3	25908.6
6	11,043	459.3	612.7	1.1	25398.0
7	11,043	459.3	546.4	3.2	25885.9
8	12,031	421.6	677.5	1.6	27386.4
9	12,031	421.6	523.1	3.2	26493.6
10	12,031	421.6	522.4	3.1	26547.7
11	12,031	421.6	725.7	1.1	27362.7
12	12,031	421.6	527.9	2.9	26729.5

3.2 Evaluation

Using the twelve post-stratified weights, we computed Mean Squared Error (MSE), which equals the squared bias plus the variance, for four key variables in the 2009 MNHA. The four key variables are employment status, health status, confidence in receiving care, and dental insurance coverage.

The variance in the estimates is directly related to the variance of the weights, and is easily calculated. However the bias is not calculable because we don't know the true population mean. Instead, we need to make an assumption on which estimate is the most unbiased. The choice of which weight is unbiased is not straightforward. To compare the weights under different assumptions, we calculated MSE's under two different assumptions as to which estimate was unbiased.

For the first set of MSE calculations, we assumed that Weight 2 (landline plus cell-only, no adjustment) was unbiased and used its estimate to calculate biases in the other weights. This choice represents the whole population without any overlap and the theoretically most pure sample design for calculating known probabilities of selection for each household and person. For the second set of MSE calculations, we assumed that Weight 12 (landline plus all cell combined with bias-correcting adjustment) was unbiased. This choice represents the full set of data with the overlap adjustment method that tries to combine the samples solely based on reducing bias. These choices are unfair to the other weights, but we cannot calculate estimated MSEs without a bias assumption. We chose our two assumptions to be at opposite ends of the sample frame choices (2 and 4). It is unreasonable to assume sample frame choice 1 could be unbiased since this excludes all Minnesota households without a cell phone.

The calculated MSE's are shown in Tables 6-9. These numbers all look small, but they are on the scale of variances. For example, the MSE for Weight 2 (when it is considered unbiased) for the employment status variable is 4.30×10^{-5} (Table 6). The square root of this is .007= 0.7 percent, which is the standard deviation of this variable.

As noted above, there are two decision points examined in this report. First, how many of the cell phone cases should be included in the sample frame. As might be expected, sample frame 4, which includes all cell phone interviews, generally results in the lowest MSE's. Sample frame 1 generally does the poorest, indicating that the bias caused by collecting only landline interviews is significant. Sample frame 2 is outperformed by sample frames 3 and 4 except when Weight 2 is assumed to be unbiased for dental insurance. Similarly, sample frame 4 has lower MSE's than sample frame 3 except when sample frame 2 is assumed unbiased for dental insurance. We believe that sample frame 4 outperforms the others because when sample size increases, the weights get smaller. Though this is not necessarily going to be true, the coefficient of variation shrank when the weights did, resulting in smaller variances for sample frame 4. Even when we assumed that sample frame 2 had the unbiased estimate, the bias was not large enough to outweigh the difference in variances (except for the dental insurance variable).

Having chosen sample frame 4, the second decision point is the overlap adjustment method. As shown in Table 6-9, for all four variables, regardless of which of the two weights were considered unbiased, the best-performing weight in sample frame 4 with the lowest MSE was Weight 10 (landline plus all cell combined with the effective sample size weighting adjustment). This weight has the lowest variability in the weights and does not have enough bias (under either unbiased assumption) to offset its advantage in variance. We believe that Weight 10 performs best because it has the largest values of λ. Since it is the cell phone weights that are the largest, these are the weights that determine the variability of the weights. The largest values of λ result in the smallest values of $(1-\lambda)$, which is the multiplier for the cell phone weights. Thus, the effective sample size reduces the cell phone weights the most, which results in the smallest standard errors. It can be seen that the λ's (and MSE's) are very similar for the effective sample size and sample size methods. This implies that effective sample sizes used for Weight 10 are very similar to the sample sizes used in Weight 9. Nevertheless, Weight 10 does have lower MSE's than Weight 9 under all scenarios, so Weight 10 is preferred. The simplest $\lambda = 0.5$ has larger variance and larger bias, but the worst performing λ was the weighted sample size λ. It was thought that the bias-correcting λ was the theoretically best choice, but it was outperformed by the effective sample size λ. Even when the bias-correcting λ is assumed to result in unbiased weight, the assumed bias in the effective sample size estimate is not large enough to offset the larger difference in the variability of the estimates. The effective sample size overlap adjustment method is superior even under the best case scenario for the bias-correcting adjustment, so we conclude that Weight 10 is optimal.

Table 6: MSE's for Employment Status.

Weight	*MSE (Weight 2 unbiased)*	*MSE (Weight 12 unbiased)*
1	5.450E-05	5.000E-05
2	4.300E-05	4.410E-05
3	4.970E-05	4.720E-05
4	4.300E-05	4.240E-05
5	4.280E-05	4.230E-05
6	5.420E-05	5.090E-05
7	4.300E-05	4.240E-05
8	6.400E-05	6.000E-05
9	4.220E-05	4.120E-05
10	4.190E-05	4.100E-05
11	7.430E-05	6.880E-05
12	4.260E-05	4.160E-05

Table 7: MSE's for Health Status.

Weight	*MSE (Weight 2 unbiased)*	*MSE (Weight 12 unbiased)*
1	1.110E-04	1.385E-04
2	5.880E-05	6.140E-05
3	6.400E-05	6.180E-05
4	5.800E-05	5.770E-05
5	5.770E-05	5.770E-05
6	6.770E-05	6.500E-05
7	5.790E-05	5.770E-05
8	9.060E-05	8.030E-05
9	5.760E-05	5.560E-05
10	5.730E-05	5.550E-05
11	1.038E-04	9.110E-05
12	5.880E-05	5.620E-05

Table 8: MSE's for Confidence in Receiving Care.

Weight	*MSE (Weight 2 unbiased)*	*MSE (Weight 12 unbiased)*
1	2.135E-04	2.166E-04
2	5.600E-05	5.610E-05
3	6.050E-05	6.040E-05
4	5.500E-05	5.490E-05
5	5.480E-05	5.480E-05
6	6.410E-05	6.390E-05
7	5.500E-05	5.490E-05
8	7.100E-05	7.100E-05
9	5.370E-05	5.370E-05
10	5.360E-05	5.350E-05
11	7.750E-05	7.750E-05
12	5.420E-05	5.420E-05

Table 9: MSE's for Dental Insurance Coverage.

Weight	*MSE (Weight 2 unbiased)*	*MSE (Weight 12 unbiased)*
1	5.530E-05	6.770E-05
2	5.000E-05	5.850E-05
3	7.100E-05	5.580E-05
4	5.410E-05	4.930E-05
5	5.330E-05	4.930E-05
6	8.030E-05	6.150E-05
7	5.390E-05	4.930E-05
8	1.023E-04	7.660E-05
9	5.480E-05	4.720E-05
10	5.390E-05	4.710E-05
11	1.214E-04	9.050E-05
12	5.610E-05	4.770E-05

These results are for the 2009 MNHA. While not unanimous, we recommend the use of Weight 10 (sample frame 4, effective sample size overlap adjustment) for the 2009 MNHA. It is important to note that this paper and these comparisons do not reflect all of the decisions made in creating the 2009 MNHA weights. In particular, the base weights prior to sample combination (with λ) do include post-stratification steps in a particular order (although post-stratification is also done as the last weighting step after sample combination). Alternative base weights were not compared for this report, but the same methodology could be used to compare alternative base weights to the base weight used.

References

O'Muircheartaigh, C., and S. Pedlow (2002). Combining Samples vs. Cumulating Cases: A Comparison of Two Weighting Strategies in NLSY97. 2002 Proceedings of the American Statistical Association, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association, pp. 2557-2562.

Pedlow, S., K. Xia, and M. Davern (2010). Dual-Frame Sample Sizes (RDD and Cell) for Future Minnesota Health Access Surveys. 2010 Proceedings of the American Statistical Association, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.