Dual-Frame Sample Sizes (RDD and Cell) for Future Minnesota Health Access Surveys

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

Abstract
The Minnesota Health Access Survey (MNHA), a large-scale health insurance survey conducted jointly by the Minnesota Department of Health and the University of Minnesota, tracks health insurance coverage and access in Minnesota. The 2009 round of the MNHA is the first to use data collected from a cell phone sample as well as a random-digit-dial (RDD) sample. This paper explores the trade-offs between screening cell samples based on landline and cell phone usage patterns and accepting all cell phone users as respondents. After concluding that screening is the optimal solution, the paper then explores the optimal percentages of MNHA interviews that should be attempted from the landline and cell phone frames.

Key Words: Sample Efficiency, Design Effects, Cell phone, Landline, Screening

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 Center (SHADAC), tracks health insurance coverage and 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 (RDD) interviews. The 2009 MNHA consists of 12,031 total interviews, 9,811 of which were collected from landline telephones through the RDD frame, while 2,220 interviews were collected via a cell phone frame (regardless of landline/cell phone usage). Xia, Pedlow, and Davern (2010) compared weights for the 2009 MNHA produced under four different screening scenarios (include the RDD cases only, include the RDD plus cell-only cases, include the RDD plus cell-only/cell-mostly cases, and include the RDD plus all cell phones cases) and five different weighting adjustments for adjusting for any overlap between the RDD and cell phone interviews for the 2009 MNHA. Section 2 reviews this work.

The 2009 MNHA work described in Xia, Pedlow, and Davern (2010) made weighting decisions with data collection already completed. This paper considers the sample design for the 2011 MNHA and other future rounds. In particular, we explore two separate questions.

The first question is whether to allow a full or partial overlap in conducting surveys through the landline and cell phone sample frames. We consider a full overlap to occur when all cell phone respondents are interviewed regardless of their landline/cell phone usage. An overlap can be prevented by not having a cell phone frame sample or by screening out respondents from the cell phone frame that also have a landline telephone. A partial overlap can be allowed if cell phone respondents are interviewed if they live in a cell-only household (no landline) or a cell-mostly household, and if they are screened out of the cell phone interview if they live in a landline-mostly household. This decision is discussed in Section 3.

The second question concerns the proportion of interviews that should be conducted from each frame. We consider this question through the percentage of interviews that should be collected from the cell phone frame, and this decision is discussed in Section 4. Finally, we summarize this paper and consider its generalizability in Section 5.

2. Review of 2009 MNHA Work

The 2009 MNHA considered all adult cell phone respondents residing in the state of MN to be eligible for inclusion in the survey, regardless of their landline and cell phone usage. However, for research and weighting purposes, 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 questions to categorize the respondent as a cell-mostly user or a landline-mostly user. Therefore, the landline respondents could be landline-only (no cell phone), cell-mostly (they receive very few or no calls on their landline), or landline-mostly (they receive very few or none of their calls on their cell phone). Meanwhile, the cell phone respondents could be cell phone only (no landline phone), cell-mostly, or landline-mostly. The 2009 MNHA collected 9,811 landline interviews through a RDD sample and 2,220 from a cell phone frame, so 18 percent of the 12,031 interviews were collected from the cell phone frame.

For the 2009 MNHA, we first examined the first question above: how much overlap should we allow in the frames. We considered four possible overlap frame choices: 1) use all landline interviews and no cell phone interviews (no overlap, no cell phone coverage), 2) use all landline interviews and only those cell phone respondents who do not have a landline (no overlap, but full coverage), 3) use all landline interviews and exclude only those cell phone respondents who are classified as landline-mostly (partial overlap), and 4) use all landline and cell phone interviews (full overlap). It is important to note that all four choices use all landline interviews and a subset, all, or none of the cell phone interviews.

We created separate weights for each overlap frame choice. The MNHA weights were developed separately for the landline and cell phone samples, starting with the inverse of their selection probability as the base weight. 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 an appropriate estimate. We used a simple composite weight adjustment in which the overlap interviews have their weight multiplied by λ in the landline frame, and the overlap interviews in the cell phone frame have their weights multiplied by $1\text{-}\lambda$. We actually compared five different methods of determining λ, and further details are given by Xia, Pedlow, and Davern (2010).

We compared these four overlap frames by comparing the mean-squared errors (which is the variance of the estimate plus the bias of the estimate squared) for a set of key variables. The variance in the estimates was easily calculated, and is directly related to the variance of the weights. However, the bias is not calculable because we don't know the true population mean for any of the key variables. Instead, we needed to make an assumption on which estimate is most unbiased. The choice of which weight is unbiased is not straightforward. To compare the weights under different assumptions, we calculated mean-squared errors under two different assumptions. In one, the second overlap frame option (no overlap, but full coverage) was assumed to be unbiased and we used this estimate to calculate the bias in other estimates. In the other assumption, we assumed that one of the λ choices under the full overlap scenario provided unbiased estimates, and we used these estimates to calculate the bias of the other options.

When we compared the four overlap choices, the lowest mean-squared errors resulted from the full overlap option, which involved keeping all of the interview data. The other overlap choices all required some interview data to be discarded. If anything other than the full overlap option had been chosen before data collection, the costs of carrying out data collection would have been less. Therefore, choosing the full overlap option for the 2009 MNHA was to choose the most expensive data collection option. This paper is concerned with choosing among the four overlap options when costs are equal; in other words, when one static budget is available. We had a secondary question that we also explored which is what would be the optimal split in the proportion of cases allocated to the landline telephone sample versus the cell phone sample. This analysis takes into account the cost differentials between cell phone completes and landline completes since cell phone completes are more expensive.

Regarding the other question in this paper, for the 2009 MNHA, we were constrained by the data that had already been collected. Therefore, we could not determine whether the 18 percent of interviews via the cell phone frame was the right percentage. However, our work did suggest that the 2009 MNHA cell phone frame base weights were too high relative to the landline frame. Increasing the percentage of interviews from the cell phone frame would have reduced this base weight ratio, which suggests that the optimal percentage of interviews from the cell phone frame is more than 18 percent.

3. Frame Overlap Exploration

We consider four possible overlap options between the landline and cell phone frames:

Option 1. No cell phone interviews (landline RDD only)
Option 2. Cell phones, but no overlap (all landline + cell-only)
Option 3. Cell phones and partial overlap (all landline + cell-only + cell-mostly from cell frame)
Option 4. Full overlap (all landline + all cell interviews)

Option 1 is to collect all interviews from a landline RDD sample. Option 2 screens out all cell phone respondents who also have a landline telephone. Option 3 screens out only those cell phone respondents who have a landline and use primarily their landline. The overlap in Option 3 is that cell-mostly respondents are interviewed in both the landline and cell phone frames. Option 4 collects interviews from all cell phone respondents. The overlap in Option 4 is that cell phone respondents with a landline are interviewed in both the landline and cell phone frames.

Since our goal is to compare these four options when costs are equal, we developed cost models based on the 2009 MNHA cost data supplied by the survey vendor, Social Science Research Solutions (SSRS). These cost models show that cell phone interviews are more costly to collect than landline interviews. We also assume that there is a cost to screening out portions of the cell phone sample based on telephone usage. In effect, we assume that in Options 2 and 3, there is an additional screening cost constant for every cell phone respondent reached. Table 1 presents scenarios that our cost models show as having an equal cost to collecting 10,000 landline RDD interviews.

Table 1: Scenarios with Equal Cost to 10,000 Landline RDD Interviews

Overlap Choice	*Landline Completes*	*Cell Completes*	*TOTAL Completes*
No cell	10,000	0	10,000
No overlap	8,036	729	8,765
Partial Overlap	7,688	965	8,653
Full Overlap	7,325	1,657	8,982

Since cell phone interviews are more expensive in our cost model, the maximum number of interviews for a fixed total cost is achieved by not completing any cell phone frame interviews. The full overlap option shown in Table 1 assumes that 18 percent of the interviews are done with cell phone respondents (like the 2009 MNHA). Collecting 1,657 cell interviews reduces the number of landline completes by 2,675. Screening these 1,657 cell interviews reduces the number of cell completes, but increases the number of landline interviews that can be completed for the same total cost. Nevertheless, the total number of completes is less with screening because of the fixed screening cost. The 729 cell completes in the no overlap option are all cell-only respondents while the 965 cell completes in the partial overlap are all cell-only or cell-mostly respondents. Our cost model shows that more total interviews are collected under the no overlap option than under the partial overlap option because cell phone interviews are more expensive than landline interviews.

To compare these four overlap frame options at equal cost, we use the 2009 MNHA data. The 2,220 cell phone interviews were distributed the following way: 890 cell-only, 342 cell-mostly, and 988 landline-mostly respondents. We know the cost for the 2009 MNHA. In order to create options with equal costs for all four overlap choices, we either need to increase certain sample sizes through bootstrapping or another augmentation method; or we need to decrease certain sample sizes through subsampling. We chose to decrease the sample sizes through subsampling.

For our first experiment (Experiment 1), we kept all 9,811 landline interviews and subsampled the cell interviews randomly. In one run, we kept 75 percent of the cell interviews. In another, we kept the other 25 percent of the cell interviews. We also split the cell interviews into two halves and calculated the cost (which is the same under our cost model) and mean-squared error for both separately. Table 2 below summarizes the different runs for Experiment 1.

Consistent with Xia, Pedlow, and Davern (2010), we calculated separate weights under each overlap frame option and used these weights to calculate the mean-squared errors. When subsampling was complete, the base weights were adjusted so that the sum of the base weights did not change. For example, when we kept only half of the cell phone interviews, the base weights for those kept were doubled.

Table 2: Subsampling Runs for Experiment 1

Run Description	*Landline Completes*	*Cell-Only Completes*	*Cell-Mostly Completes*	*Landline-mostly (Cell) Completes*
No subsampling	9,811	890	342	988
75 percent kept	9,811	667	257	741
50 percent kept	9,811	445	171	494
25 percent kept	9,811	223	85	247

For Experiment 1, we calculated mean-squared errors for three key binary variables from the 2009 MNHA: 1) EMPLOYED = employment status (employed or not employed), 2) HEALTHSTAT = health status (excellent/very good or good/fair/poor), and 3) CONFIDCARE = the respondent's confidence that they could receive any health care they might need (very confident or some/little/no confidence).

As with the work reviewed in Section 2, we make two different bias assumptions to calculate bias (leading to two separate mean-squared error calculations). One assumption is that the no overlap option provides an unbiased estimate while the second assumption is that the full overlap option provides an unbiased estimate. For the work reviewed in Section 2, we compared five different composite weighting methods, but for this paper, we only used the composite weighting method that provided unbiased estimates under the second bias assumption: the bias-correction method.

The bias-corrected composite weighting method starts with the basic idea of the sample size composite weighting method:

$$\lambda = \frac{n_L}{n_L + n_C}, \quad 1-\lambda = \frac{n_C}{n_L + n_C},$$

where n_L is the sample size from the landline sample and n_C is the sample size from the cell phone sample. Instead of using the sample sizes themselves, however, the bias-correction method adjusts for the fact that the landline sample is more likely to find landline-mostly households (since cell-mostly households don't usually answer their landline calls), while the cell phone sample is more likely to find cell-mostly respondents (since landline-mostly respondents are more likely to have their cell phone turned off). This method first adjusts the sample sizes within sample to match assumed population percentages for each usage type. Then, based on these revised sample sizes, we apply the sample size composite weighting method. Here is the formula for λ for the bias-correction composite weighting method:

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

where: P_{uL} is the assumed proportion of the landline RDD frame that is in phone usage category *u*,

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

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

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

Summarizing the above, we examined three key variables under two different bias assumptions, so there were six different scenarios. For each scenario, we compared the four different frame overlap options. For Experiment 1, we calculated the cost and mean-squared errors under each frame overlap option for each of the subsamples in Table 2. Figures 1 and 2 below show two of the six scenarios.

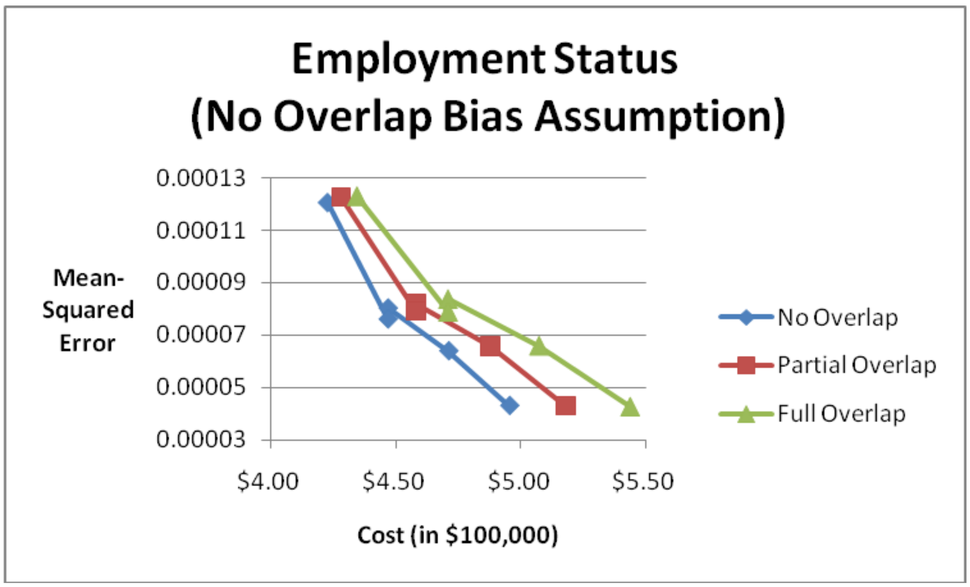

Figure 1: Mean-Square Errors for the Employment Status Variable under the No Overlap is Unbiased Assumption.

Neither figure shows the no cell interviews option because all the data points for the no cell option are above and to the left of the areas shown in the figures. This option has higher mean-squared errors, but lower costs so we did not succeed at making a fair comparison with the other frame choices. Conventional wisdom suggests that a cell phone component is now important because of the potential bias in a landline-only RDD survey, so we assume that future MNHA rounds would have a cell phone component.

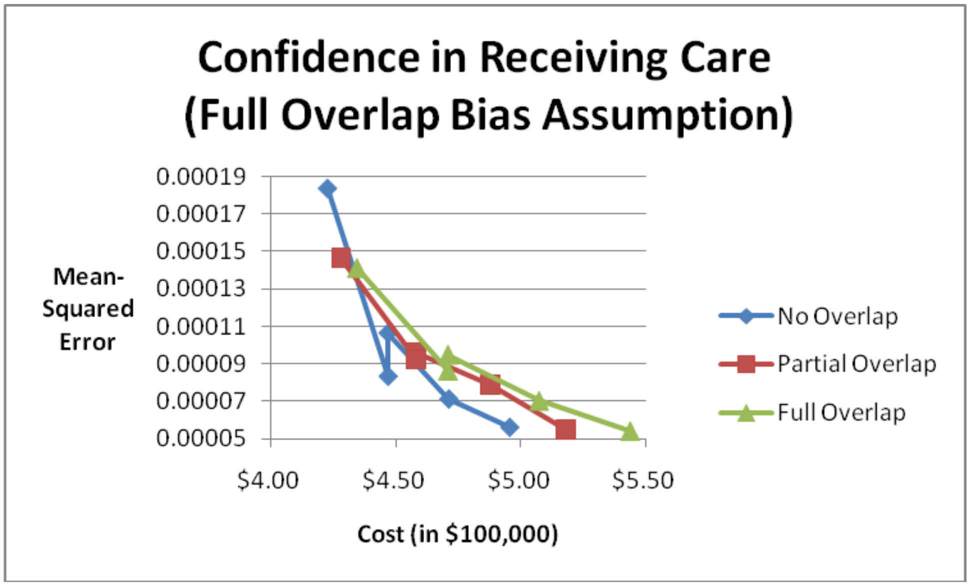

Figure 2: Mean-Square Errors for the Confidence in Receiving Care Variable under the Full Overlap is Unbiased Assumption.

In Figures 1 and 2, the blue lines refer to the no overlap option, the red lines refer to the partial overlap option, and the green lines refer to the full overlap option. Figure 1 shows that for employment status under the no overlap unbiased assumption, the no overlap frame option has the lowest mean-squared errors for the same cost because the blue line is below the red and green lines. Five out of the six scenarios look like Figure 1. Figure 2 shows the same result for the confidence in receiving care variable under the full overlap unbiased assumption, though the blue line is not as clearly separated from the red and green lines. In both figures, each line has a hitch in the middle. To explain this hitch, recall that when we subsampled half of the cell phone interviews, we analyzed both halves. The two halves had the same cost, but different mean-squared errors, so the two points are connected by a vertical line.

Experiment 1 clearly shows a preference for the no overlap frame option because the blue lines travel under the red and green lines, showing that for the same cost, the no overlap option results in the lowest mean-squared errors. This option does result in the largest number of landline interviews (except for the no cell interviews option), but it also results in the lowest number of cell phone completes and total completes for a known cost, so the explanation of why it is best is not straightforward. Our hypothesis is that screening makes each cell interview more expensive, and that the overlap cases have less independent information.

4. Balance of Landline/Cell Phone Interviews

The second question we explore in this paper is the proper balance in interviews collected from the two frames. We consider this question through the percentage of interviews that should be collected from the cell phone frame. The 2009 MNHA collected 18.5 percent of its interviews from the cell phone frame. However, Experiment 1 shows us that the no overlap frame will be our recommendation for future MNHA rounds. To study this question, we need to drop the cell-mostly and landline-mostly interviews from the cell phone frame. This leaves us with 9,811 landline interviews and 890 cell-only interviews, so only 8.3 percent of the remaining interviews are cell-only. From this base of 8.3 percent, Experiment 1 subsampled the cell phone interviews, resulting in percentages of cell-only interviews lower than 8.3 percent. We now describe Experiment 2, which will subsample the landline interviews, resulting in percentages of cell-only interviews that are above 8.3 percent.

For Experiment 2, we kept all 890 cell-only interviews and subsampled the 9,811 landline interviews randomly. In one run, we kept 75 percent of the landline interviews. In another, we kept the other 25 percent of the landline interviews. We also split the landline interviews into two halves and calculated the cost (which is the same under our cost model) and mean-squared error for both separately. Table 3 summarizes the different runs for Experiment 2.

Table 3: Subsampling Runs for Experiment 2

Run Description	*Landline Completes*	*Cell-Only Completes*	*TOTAL Completes*	*Percentage Cell-Only*
No subsampling	9,811	890	10,701	8.3%
75 percent kept	7,359	890	8,249	10.8%
50 percent kept	4,906	890	5,796	15.4%
25 percent kept	2,452	890	3,342	26.6%

Table 3 shows that as we subsampled the landline interviews, the percentage of cell-only interviews increased. We again calculated separate weights for each run and used these weights to calculate the mean-squared errors. Once again, when subsampling was done, the base weights were adjusted so that the sum of the base weights did not change. For example, when we kept only half of the landline interviews, the base weights for those kept were doubled.

For Experiment 2, we calculated mean-squared errors for the same three key binary variables used in Experiment 1, and used the same two bias assumptions to calculate bias (leading to two separate mean-squared error calculations).

The data from both experiments are combined in Figures 3 and 4, which give us nine different subsamples. The scenarios from Experiment 1 range in the percentage of cell-only interviews from 2.2 percent to 8.3 percent. The scenarios from Experiment 2 range in the percentage of cell-only interviews from 8.3 percent to 26 percent. We have calculated the cost of interviewing for each of the nine subsamples. The maximum cost is the subsample where we keep all 9,811 landline interviews and all 890 cell-only interviews. As we subsample more cases of either frame, the costs decrease. Each subsample with less than 8.3 percent cell-only interviews will achieve equal costs with one subsample with more than 8.3 percent cell-only interviews. These two subsamples can be compared on mean-squared error.

Once again, there are six possible scenarios because there are two bias assumptions used for three key variables. All six of the graphs look like Figure 3 below.

In Figure 3, the blue line shows the calculated mean-squared errors and costs when the landline interviews are subsampled and the percentage of cell-only interviews is above 8.3 percent, while the red line shows the calculated mean-squared errors and costs when the cell-only interviews are subsampled and the percentage of cell-only interviews is below 8.3 percent. Both lines start at the same minimum mean-squared error value and the maximum cost. As cases are subsampled, the costs decrease while the mean-squared errors increase. Since the mean-squared errors rise much more rapidly for the red line, it seems clear from Figure 3, if costs are equal, the mean-squared errors will be smaller for cell-only percentages greater than 8.3 percent.

Since all six scenarios show the same results as Figure 3, there is clear evidence that future MNHA rounds under the no overlap option should have more than 8.3 percent of the interviews from the cell phone sample. However, Figure 3 doesn't show how much greater the cell-only percentage of interviews should be.

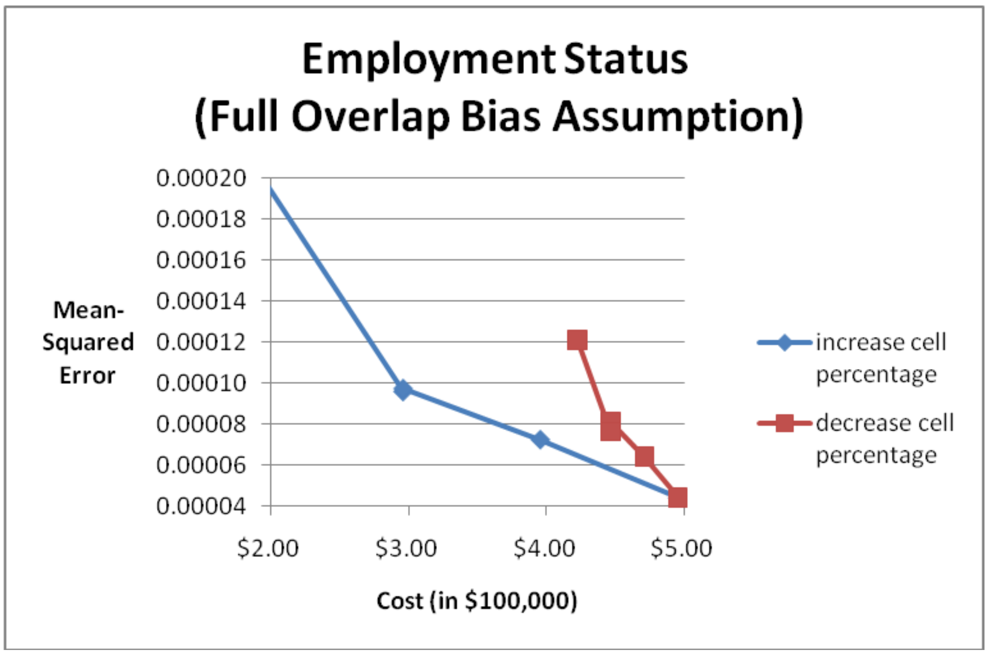

Figure 3: Mean-Square Errors for the Employment Status Variable under the Full Overlap is Unbiased Assumption.

To explore this issue further, we built a linear regression of mean-squared error on cost and the percentage of cell-only interviews for each of the six scenarios. Since we did not expect the relationships to be linear, we also included quadratic terms in the model. Each model was built using the nine different observations from the two experiments. Using these regression models, we then kept cost constant and studied the relationship between mean-squared errors and the percentage of cell-only interviews.

Figure 4 shows the resulting curves for all three key variables under the no overlap unbiased assumption.

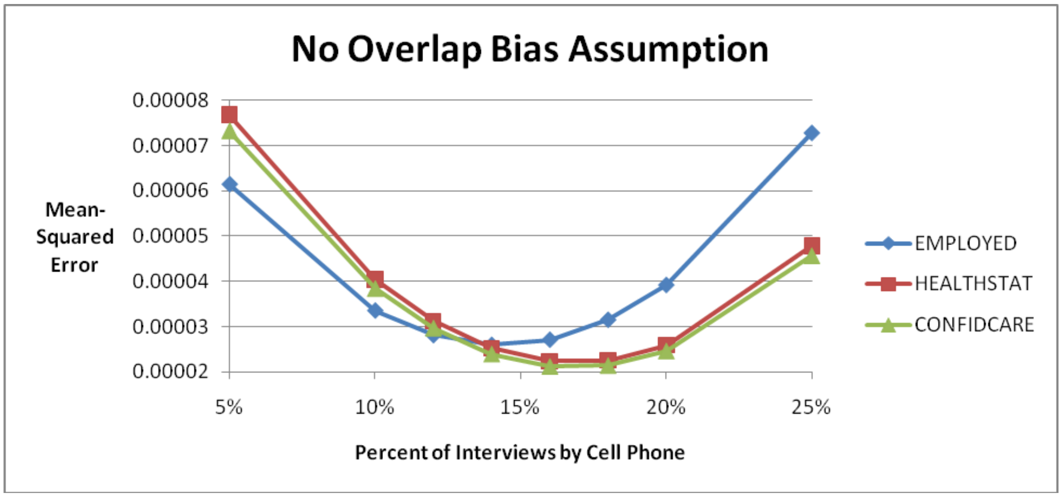

Figure 4: Mean-Square Errors for Equal Costs by the Cell-Only Percentage under the No Overlap is Unbiased Assumption.

Every point in these three lines has the same cost, so we can compare the mean-squared errors by the percentage of cell-only interviews. The curves for health status and confidence in care are similar, and both show a minimum mean-squared error at 16 percent. The curve for employment status is shifted up and to the left; its minimum is at 14 percent. Figure 5 shows the similar curves for all three key variables under the full overlap unbiased assumption.

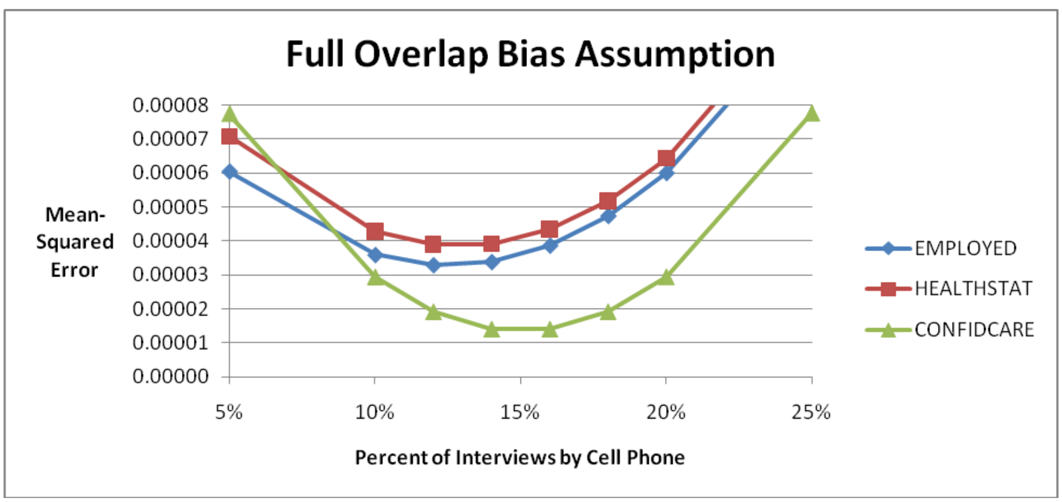

Figure 5: Mean-Square Errors for Equal Costs by the Cell-Only Percentage under the Full Overlap is Unbiased Assumption.

In Figure 5, the employment status and health status curves are similar, and both show a minimum mean-squared error at 12 percent. The curve for confidence in care is shifted down and to the right; its minimum is at 15 percent.

All six scenarios agree on having a minimum mean-squared error for the same cost at around 14 percent, so that is what we recommend for the 2011 Minnesota Health Access

Survey. Table 4 shows many options with the same total cost as the 2009 Minnesota Health Access Survey.

Table 4: Options for the 2011 MNHA Equal in Cost to 2009 MNHA

Frame	*Percentage of Cell-Only Completes*					
	0%	10%	12%	14%	16%	18%
Landline	13,395	10,313	9,800	9,316	8,857	8,009
Cell-Only	0	1,146	1,336	1,516	1,687	2,002
TOTAL	**13,395**	**11,459**	**11,136**	**10,832**	**10,544**	**10,011**

Table 4 shows that if no cell interviews were conducted, the 2009 MNHA could have collected 13,395 landline interviews for the same cost as its 9,811 landline and 2,220 cell (12,031 total) interviews. Our recommendation is that 14 percent of the interviews be collected from cell-only respondents. For the same cost as the 2009 MNHA, this would imply 9,316 landline and 1,516 cell-only (10,832 total) interviews.

5. Summary and Generalizability of Our Results

We recommend that the 2011 Minnesota Health Access Survey screen out all cell phone households that also have a landline telephone and keep only the cell-only respondents. We also recommend that 14 percent of the interviews be collected from cell-only households. Table 5 below compares the 2009 MNHA with our recommendation for 2011.

Table 5: Comparison of 2009 MNHA with Our Recommendation for 2011

Telephone Status	*2009 MNHA*	*Recommendation For 2010 MNHA*
Landline	9,811	9,316
Cell: Cell-Only	890	1,516
Cell: Cell-Mostly	342	0
Cell: Landline-Mostly	988	0
TOTAL	**12,031**	**10,832**

Though this work was conducted for the Minnesota Health Access Survey, it is natural to consider how general these findings may be. With respect to the recommendation against allowing the two telephone frames to overlap, the subsampling analyses that we have described seem valid, but do depend on the cost models we used. Nevertheless, the relationships seem to be robust in showing the no overlap choice to be best. We expect this recommendation may be widely applicable.

With respect to the percentage of interviews to collect from the cell phone frame, our work with the 2009 Minnesota Health Access Survey showed that if the cell phone sample size were increased enough, the cell weights would no longer be significantly larger than the landline weights. Therefore, we expected that the optimal choice of the percentage of cell-only interviews would be larger than in the 2009 survey. The six simple regression models all agreed, but it is hard to know the quality of these models.

The minimums of these models are probably quite sensitive to our cost models, so 14 percent may not be optimal for the Minnesota Health Access Survey, therefore, we would expect the optimal percentage to vary for other surveys as well.

We plan to try these recommendations for the 2011 Minnesota Health Access Survey and do the necessary follow-up work to see if they really are the optimal decisions.

Reference

Xia, K., S. Pedlow, and M. Davern. 2010. Dual-Frame Weights (RDD and Cell) for the 2009 Minnesota Health Access Survey. In JSM Proceedings, Survey Research Methods Section. Alexandria, VA: American Statistical Association.