



Can Post-Stratification Adjustments Do Enough to Reduce Bias in Telephone Surveys that Do Not Sample Cell Phones? It Depends

Kathleen Thiede Call

SHADAC

AAPOR, Chicago

May 15, 2009

Funded by a grant from the Robert Wood Johnson Foundation

Acknowledgments

- Analysis inspired by Steve Cohen, AHRQ
- Coauthors:
 - Michael Davern
 - Michel Boudreaux
 - Pamela Jo Johnson
 - Justine Nelson

The Problem

- Erosion of sample coverage in traditional landline RDD (RDD) surveys due to rise in cell phone only households (CPOH)
- Conducting surveys with cell samples is...
 - Expensive
 - 2 to 4 times more expensive
 - Complex
 - Merging CPOH and RDD data to produce a single estimate is not straightforward

Why this matters

- State monitoring systems have limited budgets and expertise to include cell phone samples
- Post-stratification adjustments represent an attractive solution to account for non-coverage of CPOH in RDD surveys

Research question

1. How biased are health surveys that omit cell phone only households (CPOH)?
2. Can post-stratification adjustments reduce bias associated with not sampling CPOH in RDD health surveys?
3. How well do these adjustments work for key outcomes and segments of populations?

Methods

- Data:
 - 2008 NHIS public use data (0-64 year olds only)
- General approach:
 - Remove CPOH and re-weight non-CPOH data to NHIS control totals using an iterative process
 - Conventional: region, race/ethnicity, age
 - Less conventional: age by education, home ownership status, adult only 18-30 year old households
 - Contrast iterations of post-stratification adjustments on health related outcomes, selecting the weight that performs best based on the mean squared error (MSE) for each outcome

Overview of analysis

- Contrast CPOH and non-CPOH estimates for range of health related outcomes (using original weights):
 - Health insurance coverage, delayed care due to cost, usual source of care, drinking and current smoking status
- Reweight data: Examine extent to which the reweighted data reduce bias introduced from excluding CPOH

Definitions:

CPOH at least one member has cell service and no landline telephone in household

Non-CPOH include households with landlines, no service, and unknown service

- CPOH equal **19.5%** of the non-elderly weighted Person File
- CPOH equal **20.7%** of the non-elderly weighted Sample File

Table 1. Selected outcomes for non-elderly; original public use weights by phone status

Non-CPOH subsample (A-B) significantly underestimates all key health related outcomes

Non-CPOH and CPOH subsamples are significantly different on all health related estimates (C-B); CPOH rates are higher on all outcomes

		Total Sample (A)	CPOH Omitted (B)	CPOH (C)	Tests of significance	
					A-B	C-B
Person File	Uninsured	16.7%	14.6%	25.1%	2.0% *	10.4% *
	No care because of cost	7.1%	6.1%	11.2%	1.0% *	5.1% *
	Delayed care because of cost	9.9%	8.7%	14.8%	1.2% *	6.1% *
Sample Files	No usual source of sick care	7.9%	6.7%	12.1%	1.1% *	5.4% *
	Current heavy drinking [^]	6.0%	5.4%	8.1%	0.6% *	2.8% *
	Current smoking [^]	22.7%	21.1%	28.3%	1.6% *	7.3% *

[^]Status of adults (18+)

* T-tests corrected for overlapping observations (A-B) and are significant at the $p \leq .05$ level or more

Table 2. Contrasting adjustments for selected health outcomes for non-elderly

- Iterative post-stratification adjustments were made to the public use final person weight, sample adult and child weights

Weight	Total Sample (Original Weights)	Non-CPOH (Original Weights)	Relative bias	Iterations of Post-Stratification Adjustments					
				Census regions (wregion)	Age group (wage)	Race/ethnicity (wrace)	Age x education (wagex)	Home ownership (wtenure)	Hshld structure (whsld)
Uninsured									
Estimate	16.67%	14.64%	12.19%	14.73%	15.32%	15.48%	15.42%	16.11%	16.21%
Standard Error	0.29%	0.30%		0.30%	0.31%	0.31%	0.31%	0.31%	0.32%
Bias		2.03%		1.94%	1.35%	1.19%	1.25%	0.56%	0.46%
Percent Reduction in Bias				4.66%	33.71%	41.50%	38.37%	72.54%	77.18%
Mean Squared Error (MSE*1000)		0.42		0.38	0.19	0.15	0.17	0.04	0.03

Table 2. Contrasting adjustments for selected health outcomes for non-elderly

- Bias reduction is greatest for the home ownership adjustment; lowest for age*education
- MSEs are smallest for (whsld) which fit the data best

Weight	Total Sample (Original Weights)	Non-CPOH (Original Weights)	Relative bias	Iterations of Post-Stratification Adjustments					
				Census regions (wregion)	Age group (wage)	Race/ethnicity (wrace)	Age x education (wagex)	Home ownership (wtenure)	Hshld structure (whsld)
Uninsured									
Estimate	16.67%	14.64%	12.19%	14.73%	15.32%	15.48%	15.42%	16.11%	16.21%
Standard Error	0.29%	0.30%		0.30%	0.31%	0.31%	0.31%	0.31%	0.32%
Bias		2.03%		1.94%	1.35%	1.19%	1.25%	0.56%	0.46%
Percent Reduction in Bias				4.66%	33.71%	41.50%	38.37%	72.54%	77.18%
Mean Squared Error (MSE*1000)		0.42		0.38	0.19	0.15	0.17	0.04	0.03

Table 3. Selected outcomes by weight, non-elderly

		Total Sample (Original Weights)	CPOH Omitted (Original Weights)	Absolute Bias	MSE	Final Reweight (whsld)	Absolute Bias	MSE	Bias Reduction
Uninsured	%	16.67%	14.64%	2.03%	0.42	16.21%	0.5%	0.03	77.2%
	SE	0.29%	0.30%			0.32%			
Forgone care b/c cost	%	7.09%	6.09%	1.00%	0.10	6.51%	0.6%	0.04	42.0%
	SE	0.17%	0.17%			0.18%			
Delayed care b/c cost	%	9.89%	8.70%	1.19%	0.15	9.13%	0.8%	0.06	36.4%
	SE	0.21%	0.21%			0.23%			
No usual source OC	%	13.98%	11.67%	2.31%	0.15	13.40%	0.6%	0.06	74.8%
	SE	0.34%	0.35%			0.42%			
Current smoking	%	22.66%	21.08%	1.58%	0.27	21.74%	0.9%	0.11	41.6%
	SE	0.43%	0.48%			0.51%			
Current heavy drinking	%	5.95%	5.35%	0.60%	0.04	5.55%	0.4%	0.02	33.7%
	SE	0.25%	0.27%			0.30%			
Average absolute bias				1.45%			0.62%		
Average Mean Squared Error (MSE*1000)					0.19			0.05	
Average percent reduction in bias									50.9%

Table 3. Selected outcomes by weight, non-elderly

		Total Sample (Original Weights)	CPOH Omitted (Original Weights)	Absolute Bias	MSE	Final Reweight (whsld)	Absolute Bias	MSE	Bias Reduction
Uninsured	%	16.67%	14.64%	2.03%	0.42	16.21%	0.5%	0.03	77.2%
	SE	0.29%	0.30%			0.32%			
Forgone care b/c cost	%	7.09%	6.09%	1.00%	0.10	6.51%	0.6%	0.04	42.0%
	SE	0.17%	0.17%			0.18%			
Delayed care b/c cost	%	9.89%	8.70%	1.19%	0.15	9.13%	0.8%	0.06	36.4%
	SE	0.21%	0.21%			0.23%			
No usual source OC	%	13.98%	11.67%	2.31%	0.15	13.40%	0.6%	0.06	74.8%
	SE	0.34%	0.35%			0.42%			
Current smoking	%	22.66%	21.08%	1.58%	0.27	21.74%	0.9%	0.11	41.6%
	SE	0.43%	0.48%			0.51%			
Current heavy drinking	%	5.95%	5.35%	0.60%	0.04	5.55%	0.4%	0.02	33.7%
	SE	0.25%	0.27%			0.30%			
Average absolute bias				1.45%			0.62%		
Average Mean Squared Error (MSE*1000)					0.19			0.05	
Average percent reduction in bias									50.9%

Table 3. Summary of selected outcomes by weight, non-elderly -- Continued

- Reweighted non-CPOH compared to original unadjusted public use weight:
 - Variance increases somewhat
 - Bias is greatly reduced; by an average of 51%
 - MSE shrinks
 - Greatest bias reduction is for uninsurance (77%) and reporting no usual source of care (75%)
 - Less bias reduction for behavioral outcomes
- Reweighted non-CPOH estimates perform well
 - The magnitude of the bias is modest and the estimates are only moderately different from the total non-elderly sample (gold standard)
 - The direction of the bias is toward underestimating key health related outcomes

Table 4. Reweighted estimates for key segments of non-elderly population

	Total vs. CPOH Omitted		Total vs. Final Reweight		
	Average Absolute Bias	Average MSE	Average Absolute Bias	Average MSE	Average Percent Bias Reduction
Total	1.45%	0.19	0.62%	0.05	50.94%
Hispanic	1.43%	0.36	0.94%	0.20	45.45%
Black	1.34%	0.28	0.91%	0.18	30.38%
Young adults (18-30)	1.41%	0.28	1.14%	0.21	16.44%

Bias reduction is greatest among total non-elderly; works least well for young adults
Average absolute bias across estimates is relatively small for reweighted data

Summary

- Overall, absolute bias resulting from leaving out CPOH is small
 - Only exceeded 3% for uninsurance (one outcome for one subgroup - Hispanic)
- Can post-stratification adjustments correct for bias associated with not sampling CPOH in RDD health surveys?
 - Works well for total non-elderly population
 - Less well for Hispanic and Blacks
 - Worse for young adults (18-30)

Summary Continued

- Reweighting works better for some health related outcomes
 - Bias reduction is greatest for uninsurance and no usual source of care; worse for forgone and delayed care, drinking and smoking
- Reweighting strategy presented here is context specific; only one possible strategy
 - Driven by work with state survey

Implications and Conclusions

- Overall bias was relatively small
- Reweighting RDD data to account for CPOH
 - Cost effective for aggregate estimates of general population
 - More caution required for some segments of population: may misrepresent the level of disparities in health care access as well as health related behaviors
- Encourage addition of home ownership question to survey and weighting routine
- Must continue to monitor efficacy of this approach to dealing with coverage bias with changing telephony

Contact information

- Kathleen Thiede Call
- State Health Access Data Assistance Center (SHADAC)
 - callx001@umn.edu