Dual-Frame Weighting: Issues and Approaches for Incorporating an Undersampled Cell Phone Frame in a Dual-Frame Telephone Survey

Elizabeth Ormson¹, Kennon R. Copeland¹, Stephen J. Blumberg², N. Ganesh¹, and Wei Zeng¹

¹NORC at the University of Chicago
²National Center for Health Statistics, Centers for Disease Control and Prevention

The findings and conclusions in this paper are those of the author(s) and do not necessarily represent the views of the Centers for Disease Control and Prevention.
Objectives

- Identify alternative weighting approach given cost constraints related to including cell samples in telephone surveys
  - Attenuate cell sample weights to reduce variance
  - Consider impact on bias
- Assess performance on example survey
Current data show 87% of households (HHs) have cell telephone (Blumberg & Luke 2011).

Cell-phone-only (CPO) HHs constitute 32% of population.

- CPO population has higher proportions of young adults, Hispanics, low income, renters.
- Traditional RDD landline surveys subject to bias due to noncoverage of CPO population.

---

**Percentages of adults and children living in households with only wireless telephone service or no telephone service: United States, 2003–2011**

- **Children with wireless service only**
  - 1.7% (2003)
  - 36.4% (2011)

- **Adults with wireless service only**
  - 30.2% (2003)
  - 36.4% (2011)

- **Children with no telephone service**
  - 1.8% (2003)
  - 1.7% (2011)

- **Adults with no telephone service**
  - 5.0% (2003)
  - 11.2% (2011)

**NOTE:** Adults are aged 18 and over; children are under age 18.

**DATA SOURCE:** CDC/NCHS, National Health Interview Survey.
Inclusion of Cell Sample in Telephone Surveys

• Allocation tends to be smaller than representation in population
  • Due to higher costs of cell phone interviews

• Leads to differential weights for cell, landline sample
  • Large impact on variances
  • Attenuating cell sample weights can reduce variance
    ➔ need to consider impact on bias

• Overlap of landline and cell samples
  • Full cell sample overlaps with landline sample
  • Consideration of cell sample other than CPO, landline sample with cell phone
National Survey of Children with Special Health Care Needs

• National Survey of Children with Special Health Care Needs
  • Sponsored by National Center for Health Statistics
  • Estimate prevalence of children with special health care needs (CSHCN)
  • Assess impact on family and child
  • Eligibility requirements
    – HHs with children <18 yrs
    – Longer questionnaire for CSHCN

• Cell sample inclusion
  • Added during last two quarters of data collection
  • Screened for CPO and Cell-phone-mainly (CPMa)
  • Smaller representation than population distribution dictates
  • 14.7% of sample released from cell, yielding 9.4% of completes from cell sample
• Original Weighting Method:
  • Cell sample was weighted to represent CPO/Ma population (9% of sample weighted to represent 33% of population)
  • Landline sample weighted to represent remainder of population (including phoneless population)
  • Created very high variance in the estimates due to large cell sample weights compared to landline sample weights
• Baseweights are inverse of probability of selection for all sampled telephone numbers and directly related to the relative sample size within each frame

• Cell sample baseweights significantly larger than those for landline sample, due to smaller sample size

• Simple integration of landline, cell samples leads to large variability in weights

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL</td>
<td>6,643,010</td>
<td>2.6</td>
<td>28.4</td>
<td>1264.7</td>
</tr>
<tr>
<td>Cell</td>
<td>1,140,661</td>
<td>37.6</td>
<td>294.0</td>
<td>9563.4</td>
</tr>
<tr>
<td>LL+Cell</td>
<td>7,783,671</td>
<td>2.6</td>
<td>35.9</td>
<td>9563.4</td>
</tr>
</tbody>
</table>
Weighting Methodology: Original

• Original weighting resulted in large variances of the state-level estimates
  • Example: Prevalence rate of CSHCN
    – 372,698 children in sample, weighted to represent the full population of non-institutionalized children ages 0-17
  • Average Confidence Interval (CI) half-widths were 10% of estimate,
    – 2005-06 average CI half-widths were 7% of estimate
    – Wider CIs than expected given overall sample sizes

<table>
<thead>
<tr>
<th>State</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>CI Half-Width</th>
<th>CI Half-Width as Percent of Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC</td>
<td>15.43%</td>
<td>1.29</td>
<td>2.53%</td>
<td>16.4%</td>
</tr>
<tr>
<td>ND</td>
<td>14.33%</td>
<td>1.08</td>
<td>2.12%</td>
<td>14.8%</td>
</tr>
<tr>
<td>AR</td>
<td>21.10%</td>
<td>1.18</td>
<td>2.31%</td>
<td>10.9%</td>
</tr>
</tbody>
</table>
Weighting Methodology: Original

- Original weighting resulted in large Design Effects (DEFF) for CSHCN prevalence rates
- Landline sample shows much lower DEFF than combined sample

![Distribution of State DEFFs: Original Method](chart1)

![Distribution of State DEFFs: LL Sample](chart2)
Weighting Methodology: Alternative

• Alternative method sought to reduce CIs, DEFFs
• Seek to attenuate cell sample weights
  – Reduce variability
• Leverage the landline sample
  • Model CPO status using Logistic Regression
    – NS-CSHCN had rich demographic information from interviews such as education, income, number of people, etc.
    – Identify landline cases that are similar to CPO cases based on model (“Proxy Landline”)
• Need to control for potential bias
Weighting Methodology: Alternative

• Combine proxy landline and CPO samples
  • Cell sample weights attenuated with the goal of minimizing the Mean Square Error (MSE) which combines bias and variance
  • Use compositing approach for combining cell, landline sample to derive estimate for cell population
    – Proxy landline sample also used to represent portion of landline population
  • Balance between introduction of potential bias from proxy landline sample versus reduction in variance
• Alternative method resulted in cell sample representing smaller portion of population based on compositing factor (now weighted to represent just 16% of population)
• Alternative method resulted in a smaller DEFFs for CSHCN prevalence rate
Comparison of Methods: Prevalence Rate

- Change in state level prevalence estimates
  - Large reduction in DEFFs
    - Median DEFF reduced by 38%
  - Introduction of slight bias using new method
    - Median bias of -0.52 percentage points

<table>
<thead>
<tr>
<th>Change in DEFF</th>
<th>Bias (Change in Estimate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>-0.92</td>
</tr>
<tr>
<td>Maximum</td>
<td>2.13</td>
</tr>
<tr>
<td>Minimum</td>
<td>-5.57</td>
</tr>
</tbody>
</table>
Comparison of Methods: Key Estimates for CSHCN

• Results more variable for state level key estimates related to CSHCN
  • Due to smaller sample sizes

• Results consistent across measures
  • States with large reductions in variance for one of the key estimates from survey typically had large reduction in all estimates

• Focused on two key estimates in addition to CSHCN prevalence rate
  • “Currently Insured CSHSN whose insurance is not adequate”
  • “CSHCN whose families experienced financial problems due to child’s health needs”
Comparison of Methods: North Dakota

- ND: One of two states with largest DEFFs and standard errors under original methodology
- Saw substantial reduction in CIs with alternative method
  - 45 states saw reduction in CI for prevalence

![Comparison of Methods: North Dakota](image.png)
• DC: Second state with largest DEFFs and Standard Errors under original methodology, again saw large reductions with alternative method
• Handful of states performed worse under alternative method
  • These states typically had the smallest DEFFs and standard errors under the original method
Discussion

• Cell phone samples tend to be small
  • Higher cost to field cell sample than landline sample
  • Leads to undersampled cell phone frame

• Differential between landline and cell weights
  • Undersampled cell phone frame leads to higher baseweights than landline sample
  • Increases overall variance of estimates
Discussion

• Attenuating cell sample weights is one way to account for large differences in weights between samples
  • Increases bias, reduces variance
  • Can reduce overall RMSE even with addition of bias

• Technique may not be applicable to all surveys
  • Threshold has not been tested for minimal amount of cell sample needed to use this approach
  • Survey may be more concerned with potential introduction of bias
• Balance between bias and variance of weighted estimates is a significant consideration
  • Determine degree of dampening so as to minimize MSE
    – Run simulations to determine optimization point
    – Degree of dampening may be somewhat subjective, and balance between bias and variance
• User views of bias vs. variance
  – End user may be more concerned with bias of the estimates than the confidence intervals around the estimates
  – User may not understand conceptual difference between bias and error
  – May require additional knowledge for user to understand and accept attenuated weighting approach