



Adaptive Design for ABS in a National CATI Survey of Households with Children

Charles DiSogra, Heather Hammer,
Stas Kolenikov, Rachel Martonik
(Abt SRBI)

David Finkelhor, Heather Turner
(University of New Hampshire)

Federal Committee on Statistical Methodology
Washington, DC
November 4-6, 2013

Motivation

- The Third National Survey of Children's Exposure to Violence (NatSCEV III)
- Contact HHs with children (33.4% nationally)
- Initial contact by mail (ABS)
- Topic interview by phone
- Optimize the mailout performance

Mail ABS design

- Would like to target HHs with children
- First mailout (n=40,000)
 - Consider as a pilot/experiment
 - Use Census data for stratification
 - Target high incidence areas to collect more returns for the incidence model
- Three subsequent mailouts
 - Census + field data for stratification

Census data for Mailout 1

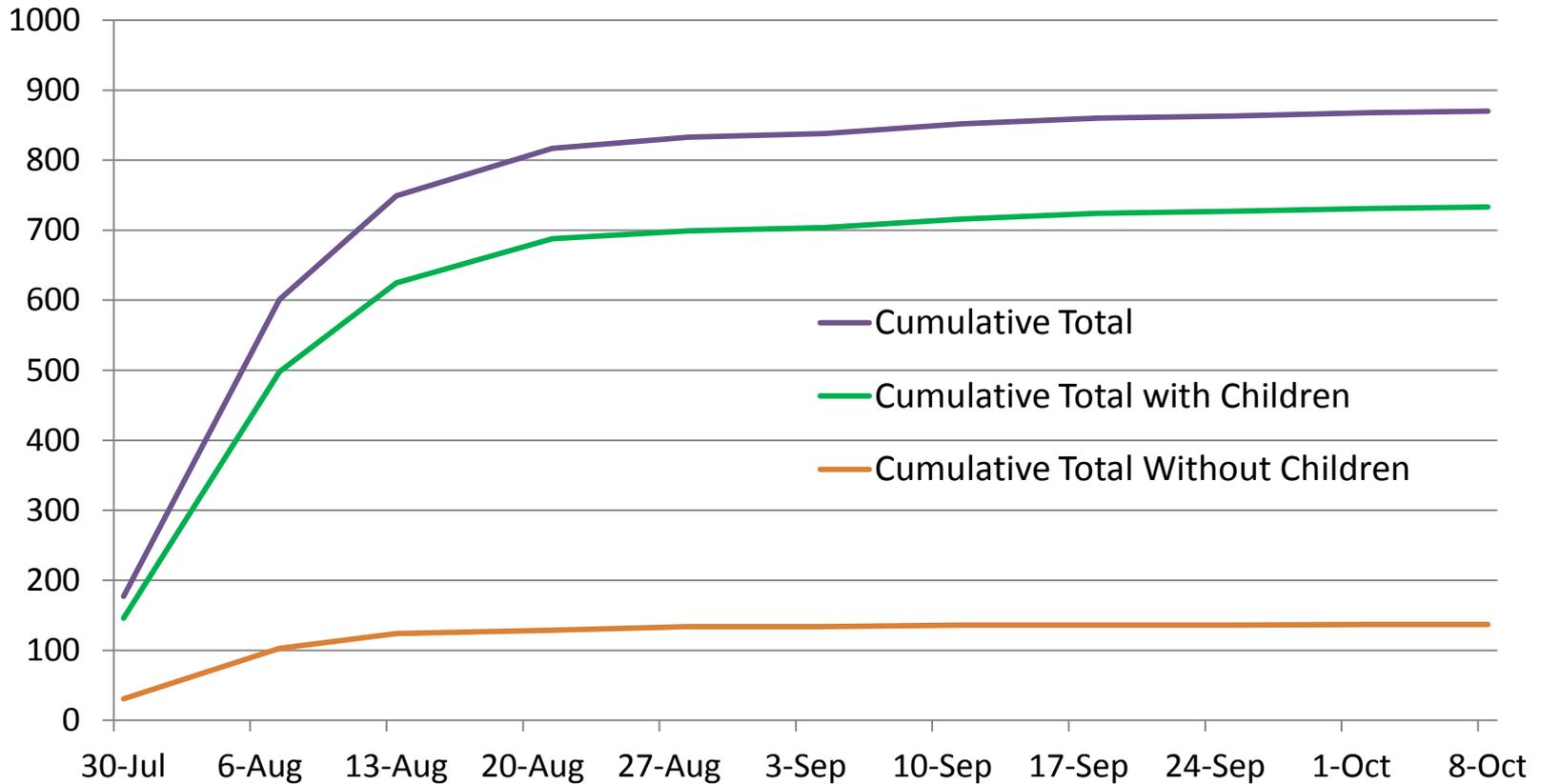
- Summary File 2 block group level data
 - P20i1 = # of HHs
 - P20i2 = # of HHs with people under 18
 - H1i1 = total # of HUs
 - H3i3 = # of vacant HUs
- Incidence = $P20i2 / H1i1$

Mailout 1 Stratification

| | Stratum 1 | Stratum 2 | Stratum 3 | Stratum 4 | Stratum 5 |
|--------------------------|-----------|--------------|--------------|--------------|-------------|
| Incidence | [0-25.4%] | [25.5-30.7%] | [30.8-36.0%] | [36.1-43.5%] | [43.6-100%] |
| # HHs w/kids | 7.77M | 7.77M | 7.77M | 7.77M | 7.77M |
| # of HUs | 45.2M | 27.5M | 23.3M | 19.7M | 15.2M |
| # of census block groups | 68.7K | 48.6K | 40.0K | 32.9K | 25.0K |
| % frame | 34.5% | 21.0% | 17.8% | 15.0% | 11.6% |
| % population | 20% | 20% | 20% | 20% | 20% |
| # sampled | 6037 | 7037 | 8140 | 8940 | 9846 |

Mailout 1 Field

Mailing done July 18th-24th and returns were scanned on a weekly basis.



Mailout 1 results: contact and in-scope data

- 870 total returns
- 90.1% (n=784) of returns provided a phone number
 - 73.5% (n=577) provided a cell phone number
- 84.2% (n=733) of returns indicated presence of children
 - Of these, 93.7% (n=687) provided a phone number

Mailout 1 results: field vs. vendor data

- 69.6% (n=39) of the 56 returns on Spanish side of form had an Hispanic MSG flag
- 53.8% (n=468) of returns did not have a matched phone on file
 - Of those, 91.7% (n=429) provided a phone number
 - Of those, 85.5% (n=367) provided a cell phone number
- 46.2% (n=402) of returns had a matched phone on file
 - Of those, 56.0% (n=225) provided a different phone number than matched phone number on file

Sample Provider Flags

| Mailed out | No evidence of children in HH* | Children present in HH | No info about children in HH | Total |
|------------|--------------------------------|------------------------|------------------------------|--------|
| Stratum 1 | 4378 | 637 | 1022 | 6037 |
| Stratum 2 | 4945 | 1099 | 993 | 7037 |
| Stratum 3 | 5783 | 1352 | 1005 | 8140 |
| Stratum 4 | 6392 | 1616 | 932 | 8940 |
| Stratum 5 | 6813 | 2037 | 996 | 9846 |
| Stratum 1 | 72.5% | 10.6% | 16.9% | 100.0% |
| Stratum 2 | 70.3% | 15.6% | 14.1% | 100.0% |
| Stratum 3 | 71.0% | 16.6% | 12.3% | 100.0% |
| Stratum 4 | 71.5% | 18.1% | 10.4% | 100.0% |
| Stratum 5 | 69.2% | 20.7% | 10.1% | 100.0% |
| Overall | 70.8% | 16.9% | 12.4% | 100% |

*Other flags may exist, but not the children flag

Mailout 1 Returns with Children

| # screeners returned | No evidence of children in HH* | Children present in HH | No info about children in HH | Total |
|----------------------|--------------------------------|------------------------|------------------------------|-------|
| Stratum 1 | 39 | 25 | 9 | 73 |
| Stratum 2 | 51 | 38 | 14 | 103 |
| Stratum 3 | 85 | 46 | 8 | 139 |
| Stratum 4 | 80 | 55 | 15 | 150 |
| Stratum 5 | 105 | 68 | 12 | 185 |
| Total | | | | 650 |

*Other flags may exist, but not the children flag

Productivity by Stratum and Match

| % mailings returned with children | No evidence of children in HH* | Children present in HH | No information about the HH | Total |
|-----------------------------------|--------------------------------|------------------------|-----------------------------|-------|
| Stratum 1 | 0.89% | 3.92% | 0.88% | 1.21% |
| Stratum 2 | 1.03% | 3.46% | 1.41% | 1.46% |
| Stratum 3 | 1.47% | 3.40% | 0.80% | 1.71% |
| Stratum 4 | 1.25% | 3.40% | 1.61% | 1.68% |
| Stratum 5 | 1.54% | 3.34% | 1.20% | 1.88% |
| Overall | | | | 1.63% |

This reported productivity is a conflation of genuine incidence and response rate – these two components must be teased apart!

Other Factors Impacting Response (?)

- Logistic regression to predict mail return with children
 - % Hispanic HHs (SF1, quadratic)
 - % black HHs (SF1, quadratic)
 - Education (MSG flag)
 - Own/rent (MSG flag)
 - Marital status (MSG flag)

How Accurate is the Vendor Flag?

- HHs flagged as having children were notably more productive, but...
- HHs without this flag also produced a non-negligible number of returns with children
- Some HHs with no children mailed the screener back even though they were instructed not to

How Accurate is the Vendor Flag?

- National incidence = [% kids] =
[SF1 # HH w/kids] / [SF1 #occupied HUs] = 33.4%
- OR for mail return with kids = [OR kids] =
$$\frac{[\# \text{ returns with kids}] / [\# \text{ returns w/o kids}]}{[\% \text{ kids}] / (1 - [\% \text{ kids}])}$$

= 10.35
- Incidence among the flagged HHs =
[% kids | MSG flag kids > 0] =
$$\frac{[\# \text{ returns with kids | flag}]}{\{ [\# \text{ returns with kids | flag}] + \text{OR } [\# \text{ returns w/o kids | flag}] \}} = 69.2\%$$

Cross-classified Incidence

| Incidence of HHs with children | No evidence of children in HH | Children present in HH | No info about children in HH |
|--------------------------------|-------------------------------|------------------------|------------------------------|
| Stratum 1 | 12.1% | 69.2% | 6.5% |
| Stratum 2 | 21.7% | 69.2% | 15.4% |
| Stratum 3 | 27.3% | 69.2% | 19.5% |
| Stratum 4 | 34.2% | 69.2% | 23.9% |
| Stratum 5 | 48.0% | 69.2% | 35.6% |

Vacant households: split between the “no children” and “no info” sub-strata

Mailout 2 Design

| Sampling rate $\times 10^4$ | Mailout 1 sampling rate | No evidence of children in HH | Children present in HH | No info about children in HH |
|-----------------------------|-------------------------|-------------------------------|------------------------|------------------------------|
| Stratum 1 | 1.3 | 2.7 | 4.2 | 2.1 |
| Stratum 2 | 2.6 | 3.1 | 4.1 | 2.5 |
| Stratum 3 | 3.5 | 3.0 | 3.9 | 2.8 |
| Stratum 4 | 4.5 | 3.3 | 3.6 | 2.5 |
| Stratum 5 | 6.5 | 3.3 | 3.2 | 2.9 |

Optimal design:

$$\# \text{ sampled} \propto \text{stratum size} \sqrt{\frac{\text{incidence}}{\text{screener cost} + \text{incidence} \times \text{main cost}}}$$

Current Status

- Mailout 2 is in the field
- Mailout 1 callbacks are in the field
 - Phone numbers from mail returns
 - Matched numbers from the sample provider



s.kolenikov@srbi.com
h.hammer@srbi.com

References

Kalton, Graham and Dallas W. Anderson. 1986. Sampling Rare Populations. *Journal of the Royal Statistical Society--A* 149(1):65-82.

Barron, M., Davern, M., Montgomery, R., Tao, X., Wolter, K., Zeng, W., Dorrell, C., Black, C. (2012) Can Information from Market Research Companies be Used to Develop an Efficient Sampling Strategy for a Rare Population? *Proceedings of the Hard-to-Reach (H2R) Population Conference, ASA.*