Methods and Considerations in Transitioning RDD to ABS: An Assessment of ABS Nonresponse Bias

David Dutwin, PhD, SVP NORC

March 4, 2020
Research Problem

• Many studies are transitioning from telephone to address-based designs
  • Response rate concerns for RDD in the age of telephone spamming and blocking
  • Cost
• Will ABS have differential nonresponse properties?
• Modal changes will have some impact
• Will the trend persist?
Uh….We Can Definitely Say This is a Downward Trend!!

After brief plateau, telephone survey response rates have fallen again

Response rate by year (%)

Note: Response rate is AAPOR RR3. Only landlines sampled 1997-2006. Rates are typical for surveys conducted in each year.

PEW RESEARCH CENTER
Spam Flagging and Blocking Impact

- 4.7 billion robocalls placed in the U.S. in January 2020
- 1,800/second
- 14.4 calls per person
They Don’t Want Us to Reach Them

Google Pixel call screen now blocks any call from known robocallers in Google’s database.

Apple iOS Silence Unknown Callers feature blocks phone numbers that you’ve never been in contact with.

Future possibilities:
1. Blocking technology continues to grow, and is “white list” agnostic.
2. Blocking technology continues to grow, but is not white list agnostic.
3. If #2, how does the research community develop white listings?
The Most Telling Statistic: The Rise of the No Answer/Voice Mail

Answering Machines/No Contacts, Cell Phones
Bias Is Generally Stable

Unweighted Overall Mean Absolute Biases, ABC, CBS, and Pew Polls

Year | Bias (%)
--- | ---
1996 | 3.9
1997 | 3.9
1998 | 4.0
1999 | 4.3
2000 | 4.4
2001 | 5.0
2002 | 5.3
2003 | 5.7
2004 | 5.6
2005 | 5.6
2006 | 5.3
2007 | 5.6
2008 | 5.2
2009 | 4.8
2010 | 4.8
2011 | 5.2
2012 | 5.3
2013 | 5.6
2014 | 5.6
2015 | 5.2
Should We Care About Trend?

• We care about trend if:
  • RDD and ABS are both equally representative and have similar nonresponse properties
  • If mode changes have negligible impact

• We do not care about trend if:
  • ABS has better nonresponse properties than RDD
  • If we do not care about trend, what, nevertheless, is the optimal approach to ABS weighting?
Thinking About Differential

- **RDD**
  - Dual-frame sample requires frame correction (Buskirk and Best, 2012)
  - Nonresponse particular to phone ownership
- **ABS**
  - Push-to-web alone will hold Internet coverage bias
  - Push-to-web combined with mail and/or phone will likely oversample Internet users
- **Both RDD and ABS**
  - Thought to under-represent the young, non-Whites, low education…but are they the same?
Frame-Independent Nonresponse

- Traditionally, a difficult assessment in RDD:
  - Hard to reach telephone analysis
  - Comparing point estimates against gold standard benchmarks.
  - Although less common, compare to appended data universe counts.
- ABS possibilities:
  - Comparing point estimates against gold standard benchmarks.
  - Compare to appended data universe counts.
  - Compare to Census Planning Database universe counts.
  - Compare to ABS frame universe counts.
Frame-Comparative Nonresponse

• Model-based comparison of mode
  • Random forest or regression to uncover most important variables in terms of modal differences
    • Would be variables one can calibrate
    • RF Can uncover key interactions
  • Similar random forest model on key survey estimates can allow for a “key driver-like” analysis
    • Variables important to key estimates and on mode will have the greatest impact on differences by mode, if calibrated to.
Process

• Weight each frame independently and similarly.
• Assess whether key point estimates are significantly different by frame.
• If so, assess nonresponse of each frame.
• Make decision as to whether to reweight telephone, abs, or both.
• Identify new metrics that will minimize nonresponse as much as possible and also possibly doing the most to preserve trend.
Data

- Minnesota Health Access Survey
  - Since 2001
  - ~11,000 interviews every other year
- Massachusetts Health Interview Survey
  - Since 2000
  - ~5,000 interviews, often annually
- Colorado health Access Survey
  - Since 2008-9
  - ~10,000 interviews, every other year

- All measure health insurance, accessibility, and full household roster demographics
Random forest interaction detection procedures can go deeper in detecting differential nonresponse.

<table>
<thead>
<tr>
<th></th>
<th>MA</th>
<th>CO</th>
<th>MN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Sig.</td>
<td>B</td>
</tr>
<tr>
<td>Number of People</td>
<td>0.05</td>
<td>.001</td>
<td>0.02</td>
</tr>
<tr>
<td>Born in US</td>
<td>0.05</td>
<td>.031</td>
<td>-</td>
</tr>
<tr>
<td>Male</td>
<td>0.00</td>
<td>.833</td>
<td>0.00</td>
</tr>
<tr>
<td>African American</td>
<td>-0.09</td>
<td>.002</td>
<td>0.00</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.03</td>
<td>.260</td>
<td>0.01</td>
</tr>
<tr>
<td>Other Race</td>
<td>-0.04</td>
<td>.167</td>
<td>-0.01</td>
</tr>
<tr>
<td>Age Under 30</td>
<td>-0.03</td>
<td>.217</td>
<td>-0.03</td>
</tr>
<tr>
<td>Age 30-49</td>
<td>0.05</td>
<td>.010</td>
<td>-0.03</td>
</tr>
<tr>
<td>Age 50-64</td>
<td>0.07</td>
<td>.001</td>
<td>-0.02</td>
</tr>
<tr>
<td>HS or Less</td>
<td>-0.13</td>
<td>.001</td>
<td>-0.08</td>
</tr>
<tr>
<td>Some College</td>
<td>-0.02</td>
<td>.163</td>
<td>-0.02</td>
</tr>
<tr>
<td>Pop Density</td>
<td>0.00</td>
<td>.702</td>
<td></td>
</tr>
<tr>
<td>Use Internet</td>
<td>0.37</td>
<td>.001</td>
<td>0.12</td>
</tr>
</tbody>
</table>
How do *baseweighted* interviews compare to the ABS on insurance status? (note: ACS person weighted)

* Limited to correcting differential sampling strata
ABS Appended Counts Example

How do baseweighted interviews compare to the ABS sample universe appended data punches?

Note: MA Caucasian uses different indicator than CO or MN
ABS – CPD Comparison Example

How do baseline weighted interviews compare to the Census Planning Database?
How do baseweighted interviews compare to the ABS sample universe appended data punches?

**Route Type = H**

<table>
<thead>
<tr>
<th></th>
<th>CO</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Delivery Point = A**

<table>
<thead>
<tr>
<th></th>
<th>CO</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Dwelling Type = Multi Family**

<table>
<thead>
<tr>
<th></th>
<th>CO</th>
<th>MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusion

• Researchers must recognize there can be significant changes in trend when transitioning from RDD to ABS.
  • Nonresponse patterns can be different
  • Modal changes
• If they exist and are attributable to nonresponse, ABS can be potentially recalibrated to nonresponse found and there are many more opportunities to discover nonresponse in ABS.
Conclusion

• Studies conducted thus far suggest that there are differences in estimates, but that they are likely modal in nature.
• Studies conducted thus far find very low nonresponse error in ABS research by a range of comparative sources:
  • Census planning database
  • ABS universe
  • ABS appended variables
  • Typical Census metrics (notably with the typical modest exceptions of age, education, and race).