Nonresponse Bias Mitigation Strategies

Andy Peytchev

FCSM Conference
November, 2021
Focus

- Nonresponse bias, not response rates
- Emphasis on selected findings and more recent developments

1. Study design features
   - Level of effort
   - Mode
   - Incentives

2. Strategies during data collection
   - Responsive Design
   - Adaptive Survey Design

3. Postsurvey adjustments
   - Modeling approach
   - Auxiliary information
Study Design Features: Level of effort

- Curtin, Presser, and Singer (2000)
  - 1-2 vs. 3+; 1-5 vs. 6+

- Numerous studies following Curtin, Presser, and Singer (2000)
  - In general, little if any relationship between effort and survey estimates
  - Keeter et al. (2000), with an experimental design, an exception

- Lin and Schaeffer (1995)
  - Nonresponse bias could even be in the opposite direction (we return to this under adjustment)

- Yet, level of effort measured as contact attempts has risen over time (e.g., Curtin, Presser, and Singer, 2000; 2005)
Survey Design Features: Mode of Data Collection

- A very expansive topic, but two key points with regard to nonresponse bias:
  - Modes vary in resulting nonresponse bias (e.g., Rothbaum and Bee, 2021, CPS ASEC)
  - Modes that reduce nonresponse bias may induce greater measurement bias for some estimates (e.g., Sakshaug, Yan, and Tourangeau, 2010)
  - Designs with multiple modes rarely include random assignment to allow estimation of mode-specific errors
Survey Design Features: Incentives

- Leverage-salience theory and related experiments to produce nonresponse (Groves, Singer, and Corning, 2000; Groves, Presser, and Dipko, 2004; Groves et al., 2006)
  
  - Those who are less interested or involved in the topic are less likely to participate
  
  - Monetary incentives are particularly effective for these groups of individuals
Survey Design Features: Level of Effort and Incentives

National Intimate Partner and Sexual Violence Survey Pilot

Source: Peytchev, Baxter, and Carley-Baxter, 2009
Strategies During Data Collection

- Heterogeneity
  - Leverage-salience theory (Groves, Singer, and Corning, 2000)

- Responsive Design (Groves and Heeringa, 2006)
  - Multiple phases and multiple protocols
  - Error-sensitive indicators play a central role

  - Tailoring or protocols at the individual level
  - Often involve models to assign protocols to sample members during data collection
Strategies During Data Collection: Dutch Labor Force Survey

**Strata**
- Registered unemployed
- 65+ households without employment
- Young household members without employment
- Non-western without employment
- Western without employment
- Young household member and employed
- Non-western and employed
- Western and employed
- Large households

**Modes**
- Web
- Telephone
- Face to Face
- Only mode vs. follow-up mode
- Standard vs. extended effort (contact attempts)

Source: Calinescu and Schouten, 2015
## Strategies During Data Collection: Dutch Labor Force Survey

Estimated response propensities by mode and stratum

<table>
<thead>
<tr>
<th>Mode</th>
<th>Stratum 1</th>
<th>Stratum 2</th>
<th>Stratum 3</th>
<th>Stratum 4</th>
<th>Stratum 5</th>
<th>Stratum 6</th>
<th>Stratum 7</th>
<th>Stratum 8</th>
<th>Stratum 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>23.2%</td>
<td>23.6%</td>
<td>15.5%</td>
<td>10.8%</td>
<td>27.9%</td>
<td>27.7%</td>
<td>17.5%</td>
<td><strong>36.7%</strong></td>
<td>22.4%</td>
</tr>
<tr>
<td>TS</td>
<td>12.2%</td>
<td>31.4%</td>
<td>8.5%</td>
<td>4.7%</td>
<td>19.7%</td>
<td>13.3%</td>
<td>7.2%</td>
<td>18.1%</td>
<td>21.2%</td>
</tr>
<tr>
<td>TE</td>
<td>20.8%</td>
<td>41.3%</td>
<td>15.2%</td>
<td>8.6%</td>
<td>31.1%</td>
<td>23.8%</td>
<td>14.3%</td>
<td>33.3%</td>
<td>37.5%</td>
</tr>
<tr>
<td>F</td>
<td>43.5%</td>
<td>53.5%</td>
<td><strong>42.2%</strong></td>
<td>34.1%</td>
<td>45.1%</td>
<td>45.3%</td>
<td>35.9%</td>
<td><strong>46.7%</strong></td>
<td>54.6%</td>
</tr>
<tr>
<td>FE</td>
<td>52.4%</td>
<td>58.3%</td>
<td><strong>51.0%</strong></td>
<td>41.2%</td>
<td>51.2%</td>
<td>54.9%</td>
<td>46.0%</td>
<td><strong>56.8%</strong></td>
<td>61.4%</td>
</tr>
<tr>
<td>W→TS</td>
<td>28.3%</td>
<td>41.0%</td>
<td>20.2%</td>
<td>13.9%</td>
<td>36.3%</td>
<td>34.0%</td>
<td>20.8%</td>
<td>44.5%</td>
<td>23.1%</td>
</tr>
<tr>
<td>W→TE</td>
<td>32.8%</td>
<td>48.4%</td>
<td>23.8%</td>
<td>17.5%</td>
<td>42.1%</td>
<td>41.1%</td>
<td>25.8%</td>
<td>52.1%</td>
<td>24.4%</td>
</tr>
<tr>
<td>W→FS</td>
<td>46.3%</td>
<td>57.7%</td>
<td>38.6%</td>
<td>32.7%</td>
<td>50.0%</td>
<td>51.0%</td>
<td>39.3%</td>
<td><strong>58.9%</strong></td>
<td>50.0%</td>
</tr>
<tr>
<td>W→FE</td>
<td>49.8%</td>
<td>58.3%</td>
<td>43.4%</td>
<td>36.6%</td>
<td>52.6%</td>
<td>54.7%</td>
<td>44.3%</td>
<td><strong>62.0%</strong></td>
<td>54.2%</td>
</tr>
</tbody>
</table>

“W”, “T” and “F” refer to Web, telephone and face-to-face; “S” and “E” refer to standard and extended effort. The strata are based on age, size of the household, number of registered unemployment and ethnicity.

Source: Calinescu and Schouten, 2015
Concept of Bias Propensity (Peytchev, Pratt, and Duprey, 2020)
  • Exclude paradata that are strong predictors only of nonresponse
  • Include demographic and substantive variables of interest
  • Allows identification of nonrespondents who contribute to nonresponse bias

Intervene on underrepresented sample members based on the bias propensity
  • Prepaid $5 incentive

Evaluate nonresponse bias reduction
  • Frame and prior round data
  • Additional phases of data collection
3.1
3.0
2.3

Variables Available Only in the Survey

End of Phase 1
End of Phase 2, Control
End of Phase 2, Treatment

Average Absolute Bias (in Percentage Points)

Compared to Estimate from Additional Follow-up Phases

Source: Peytchev, Pratt, and Duprey, 2020
Postsurvey Adjustments

- Relatively small gains from alternative statistical approaches and estimation methods
  - Machine learning, e.g., tree-based methods (for reducing nonresponse bias)
    - Some can lead to increase in variance without a commensurate reduction in bias
  - Replication-based variance estimation methods (for reducing variance estimates)
  - It would have to be a very rudimentary statistical method to find a substantial improvement

- Substantial gains are possible from additional auxiliary information
  - A shift in approach from demographic characteristics to substantive variables
    - Designed paradata (Groves and Heeringa, 2006)
    - Data from other surveys, at lesser risk of nonresponse bias
    - Administrative data
Postsurvey Adjustments: Observations in Several Studies

Kreuter et al. (2010). Using proxy measures and other correlates of survey outcomes to adjust for non-response: examples from multiple surveys. JRSS-A.

Fig. 1. Relationship between the correlation of z- and y-variables and correlation of z and response in five surveys (all correlations are shown as absolute values): +, UMTRI; ×, MEPS; ○, ESS; ◊, ANES; ●, NSFG
Postsurvey Adjustments

- Auxiliary variables informed by social science
  - Civic duty
    - Political participation is strongly related to survey participation
  - Altruism
    - Charitable activities are strongly related to survey participation
Postsurvey Adjustments: General Social Survey

Source: Peytchev, Presser, and Zhang, 2018
Postsurvey Adjustments: General Social Survey

Source: Peytchev, Presser, and Zhang, 2018
Postsurvey Adjustments: GSS—Voting and Preference

Relative Difference in Weighted Estimates

-5% -4% -3% -2% -1% 0% 1% 2% 3% 4% 5%

- Fair or poor health
- Very or pretty happy
- Life exciting
- Most people try to be helpful
- Most try to take advantage
- Most people can be trusted
- Family income below average
- Read newspaper every day *
- No religion
- Donated blood in past year
- Donated to charity in past year *
- Support birth control to teens 14 to 16 years old
- Same-sex female couple can bring up child just as well *
- Oppose capital punishment
- Courts are too harsh with criminals

Source: Peytchev, Presser, and Zhang, 2018
Postsurvey Adjustments: GSS—Voting and Preference

Source: Peytchev, Presser, and Zhang, 2018
Summary Remarks

- NRB is a counterfactual that is difficult to study
  - Contributes to a focus on response rates
- We usually know what design features can reduce NRB
  - But it often involves tradeoff with other sources of survey error
- We can have much smarter data collection
  - Added effort that needs to be planned
  - Requires statistical expertise
- Postsurvey adjustments
  - Over-emphasis of statistical methods
  - Relying on demographic characteristics alone may be insufficient
  - Needed variables related to both nonresponse and survey variables of interest
Andy Peytchev
apeytchev@rti.org