Managing Nonresponse as a General Methodological Problem

Linda Owens
Survey Research Laboratory
University of Illinois at Chicago

Outline

- Overview of response rate issue
- Overview of non-response bias issue
- Review of available methods for identifying non-response bias
- Summary of known non-response correlates
Main Problem Confronting Us

- How accurate are survey data in an environment in which response rates continue to fall (in all developed nations)?

![Graph showing Telephone Survey Response Rate from 1997 to 2012]
Response rates are the most widely compared statistic for judging survey quality.

- Atrostic et al. (2001)
- Biemer and Lyberg (2003)

What are the reporting practices regarding survey response rates?
We conducted a brief assessment of current practice

- Survey of journal editors
- Audit of recently-published articles

Editor Survey

- sample of 18 journals
  - 9 social science journals
  - 9 health science journals
- n=10 or 18 responded (55.6%)
- n=1 soft-refusal
- IRB-approved
### Health science journals

- American Journal of Psychiatry
- American Journal of Preventive Medicine
- American Journal of Public Health
- Annals of Epidemiology
- JAMA
- Journal of Studies on Alcohol and Drugs
- Medical Care
- Preventive Medicine
- Substance Use & Misuse

### Social science journals

- American Journal of Community Psychology
- American Journal of Political Science
- American Political Science Review
- International Journal Public Opinion Research
- Journal of Community Psychology
- Journal of Health & Social Behavior
- Public Administration Review
- Public Opinion Quarterly
- Social Forces
Findings

- Many prominent journals have no clear policy regarding RR disclosure
- Some have minimum required response rate, but no clear standard as to how response rate is to be calculated

Journal article audit

- Sample of 95 papers selected from same set of 18 journals
- Published January 2000 or more recent
- Unit of analysis = individual
- Noninstitutional population
- Reporting primary survey data (no secondary analyses)
Disclosure typology

1. Report no information ............... (5.3%)
2. Report other information only ...... (23.2%)
3. Report non-AAPOR response rate .. (26.3%)
4. Report AAPOR-consistent RR ...... (24.2%)
5. Report summary of sample
dispositions ......................... (21.0%)

Findings

- Great variety of formulas being used to estimate response rates
  - Some of these are not really response rates
  - None of the papers audited underestimate response rates
  - Many appear to be over-estimating response rates
**Standard Definitions**

Final Dispositions of Case Codes and Outcome Rates for Surveys

Revised 2011

- RDD Telephone Surveys
- In-Person Household Surveys
- Mail Surveys of Specifically Named Persons
- Internet Surveys of Specifically Named Persons

**AAPOR Standards**

- have been adopted/ endorsed by many, including:
  - At least two top survey research journals:
    - *Public Opinion Quarterly*
    - *International Journal of Public Opinion Research*
  - publications such as:
    - *Encyclopedia of Survey Research Methods*
    - *Handbook of Survey Research*
  - organizations such as:
    - International Social Survey Programme
    - Organisation for Economic Co-operation and Development (OECD)
Are standard response rates good enough?

CERTAINLY NOT
What is non-response bias?

- the difference in a measureable characteristic between persons who respond to a survey and the population to which we are trying to generalize

Formula for non-response bias

(Levy & Lemeshow, 1999, p. 394):

\[ B(x) = \bar{X}_1 - \bar{X} = (\frac{N_2}{N}) (\bar{X}_1 - \bar{X}_2) \]

- \(N\)=total sample
- \(N_1\)=total respondents
- \(N_2\)=total non-respondents
- \(\bar{X}\)=mean value among respondents
- \(\bar{X}_1\)=mean value among non-respondents
- \(\bar{X}_2\)=mean value for total sample

Respondents fall into one of two discrete groups: respondents or nonrespondents.
Another formula for NR bias

- \( \text{Bias}(\bar{y}_r) \approx \frac{\sigma_{yp}}{\bar{p}} \)

- where
  - \( \sigma_{yp} \) is the population covariance between the survey variable \( y \) and the response propensity \( p \)
  - \( \bar{p} \) is the mean propensity in the target population
  - Assumes respondents have a propensity or probability (ranging from 0 to 1) of responding to a given survey

Assume 50% RR and current diabetes prevalence estimate of 20%

<table>
<thead>
<tr>
<th>Scenario</th>
<th>( \bar{X}_1 )</th>
<th>( \bar{X}_2 )</th>
<th>( \bar{X} )</th>
<th>( B(\bar{x}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>0</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>20%</td>
<td>100%</td>
<td>60%</td>
<td>40</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>20%</td>
<td>0%</td>
<td>10%</td>
<td>10</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>20%</td>
<td>35%</td>
<td>27.5%</td>
<td>7.5</td>
</tr>
</tbody>
</table>
Non-response bias as function of response rate \((t_R)\)

Empirical research in general finds no response rate effect on bias:

- Groves & Peytcheva (2008)
- Keeter et al. (2000)
- Keeter et al. (2006)
- Curtin, Presser and Singer (2000)
Example of no relationship between non-response and measurement error


Figure 16.2. 1996 response rates and signed error.

Increasing response rate will only address non-response bias when variables of interest are correlated with response propensities

- Groves (2006); Olson (2007); Groves & Peytcheva (2008)
- Example: “civic duty” may be associated with both doing volunteer work and willingness to respond to a survey request
- **Key point**: nonresponse bias has its effects at the variable level, not at the survey level

Some variables affected by non-response

- Compared to respondents, non-respondents:
  - Have poorer physical health (Cohen & Duffy, 2002)
  - Have poorer mental health (Clark et al., 1983)
  - Have worse medication adherence (Gadkari et al. 2011)
  - Have higher mortality (Barchielli & Balzi, 2002)
  - Have higher negative evaluations of health care (Perneger et al., 2005)
  - Are more likely to smoke (Cunardi et al, 2005)
  - Are more likely to drink heavily (Wild et al., 2001) or not at all (Viviënne et al., 2002)
  - Are more likely to attend church (Dunne et al., 1997)
Some variables affected by non-response

- Compared to responding employees, non-responding employees:
  - Have greater intentions to quit job
  - Have less organizational commitment
  - Have less satisfaction toward supervisors and their own jobs
  - Have more negative beliefs about how their organization handles employee data (Cohen & Duffy, 2002)

Concern that increased field efforts to minimize non-response error may lead to increased measurement error

- Motivation could be cause of both willingness to participate in a survey and willingness to perform cognitive processing necessary to produce optimal answers
- Hence, additional field efforts may do more harm than good by adding poorer quality data
Some Evidence

- **Cannel & Fowler**: Number of hospital stays and length of stay less accurately reported by those who responded to later follow-ups
- **Olson (2006)**: Found that increased field work decreased nonresponse error more than it increased measurement error
- **Kreuter, Müller & Trappmann (2010)**: Findings similar to Olson (2006)

### Table 2. Means by Stage of Sample Recruitment

<table>
<thead>
<tr>
<th>Record Value</th>
<th>Number of Months Since Divorce</th>
<th>Number of Previous Marriages</th>
<th>Length of Marriage (in Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SE</td>
<td>Mean</td>
</tr>
<tr>
<td>Target (full sample)</td>
<td>130.29</td>
<td>3.57</td>
<td>49.75</td>
</tr>
<tr>
<td>Not Contacted</td>
<td>114.46</td>
<td>7.09</td>
<td>48.74</td>
</tr>
<tr>
<td>Contacted</td>
<td>134.17</td>
<td>4.08</td>
<td>50.00</td>
</tr>
<tr>
<td>Contacted, Not Interviewed</td>
<td>134.17</td>
<td>13.16</td>
<td>46.68</td>
</tr>
<tr>
<td>Interviewed</td>
<td>134.17</td>
<td>4.29</td>
<td>50.44</td>
</tr>
<tr>
<td>Survey Report (complete cases)</td>
<td>133.92</td>
<td>4.79</td>
<td>55.74</td>
</tr>
</tbody>
</table>

Non-response Bias Assessment Methodologies

Approaches that use internal criteria
1. Comparing early vs. late Respondents
2. Follow-up surveys with initial non-Respondents
3. Comparing Respondent characteristics with estimated characteristics of non-respondents
4. Panel attrition in longitudinal studies
5. Use of vignette questions
Non-response Bias Assessment
Methodologies

Approaches that use external criteria
6. Comparisons of surveys with varying response rates
7. Comparisons using other external data sources
8. Comparisons of information from sample frame for respondents vs. non-respondents
9. Comparisons via record matching

Comparing Early vs. Late Respondents (or initial refusers)
- a.k.a. “level of effort” analysis (Olson, 2006)
- examining late respondents (or initial refusals) under the assumption that they are most similar to non-respondents
  - The “continuum of resistance” (Filion, 1976)
- Very common and inexpensive approach (Cottler et al., 1987; Crawford, 1987; Etter & Perneger, 1997; Fitzgerald & Fuller, 1982; Miller & Wedeking, 2004; Siemiatycki & Campbell, 1984; Sharp & Feldt, 1959; Stinchcombe et al., 1981; Voigt et al., 2003; Wilson, 1981)
- One variation is to include “days to respond” in regression models to determine if it is associated with key survey questions (Lindner, Murphy & Briers, 2001)
Comparing Early vs. Late Respondents (or initial refusers)

- **Limitations:**
  - Assumption that there is an underlying continuum of nonresponse has been challenged (Lin & Schaffer, 1995; Fitzgerald & Fuller, 1982).
  - Method is ad hoc
    - Requires large numbers of small decisions:
      - How to best to operationally measure “resistance”
        - Who is a later responder?
        - Who is a refuser?
      - How to distinguish “easy to locate” from “hard to locate” respondents

Follow-up Surveys with Initial Non-respondents

- Designed to collect data from persons classified as non-respondents to an original survey
- **Also a common strategy** (Casper, 1992; Crawford, 1986; 1987; Gmel, 2000; Groves et al., 2005; Hill & Roberts, 1997; Lahaut et al., 2002; Lemmens et al., 1988; Shahar et al., 1996; Turner, 1999)
- Often use different modes and more intensive methods
- Burden minimized by asking fewer questions
- Benefits maximized by increasing incentives
Follow-up Surveys with Initial Non-respondents

- Data can be used to make adjustments via the Hansen-Hurwitz (1946) technique.
  - \(\bar{y}' = \frac{n_1 \bar{y}_1 + n_2 \bar{y}_m}{n}\)
  - \(n_1=\) respondents, \(n_2=\) nonrespondents, \(m_2=\) subset of nonrespondents who participate in follow-up survey
- **Limitations:**
  - Typically obtain low response rates to follow-up survey
  - Differences in data collection mode, compared to original survey, also common
    - Mode effects may be mistaken for non-response effects

Comparing Respondent Characteristics with Estimated Characteristics of Non-respondents

- Impute the characteristics of non-respondents and compare them with respondents
- Information for imputation can come from:
  - Screening interviews
  - Direct household observations
  - Informant reports
  - Geographic and other databases
Comparing Respondent Characteristics with Estimated Characteristics of Non-respondents

- Strategies:
  1. Ask interviewers to provide estimates of the characteristics of housing units and their occupants (Fitzgerald & Fuller, 1982; Lynn, 2003; Smith, 1983)
  2. Use geography and name/surname to impute characteristics of non-respondents (DeFrank et al., 2007)

- Limitations:
  - Has generally been used to infer housing and demographic variables only
  - Strategy 2 has not been demonstrated to be accurate

Panel Attrition in Longitudinal Surveys

- Examine common baseline information for respondents and non-respondents to follow-up interview requests
- Very common approach (Barchielli & Balzi, 2002; Bucholz et al., 1996; Caetano et al., 2003; Eaton et al., 1992; Hill & Willis, 2001; Gray et al., 1996; Kalton et al., 1990; Wild et al., 2001; Wilson et al., 2001)

- Limitations:
  - Bias may still exist with regard to initial sample selection
  - Processes leading to attrition may be different from those leading to initial non-participation
Use of Vignette Questions

- Evaluates potential effects of non-response by comparing the answers of survey respondents who do and do not indicate a willingness to participate in hypothetical future interviews about specific topics
- Not as many examples (Klesges et al., 1999; McNutt & Lee, 1999; Rogelberg et al., 2000; Waltermayer et al., 2003)

Examples of questions:
- “Would you take part in a telephone survey about domestic violence by a university if you were randomly selected?” (McNutt & Lee, 1999)
- “Imagine you are at home, someone calls you, they tell you they are conducting important research, and want to ask you some questions. Would you participate?” (Klesges et al., 1999)
  - Yes
  - Yes, depending on the circumstances (e.g., they call at a convenient time, how short the call is, whether I get paid)
  - No (I’d avoid them whenever they called, or hang up on them)
Use of Vignette Questions

- **Limitations:**
  - Relies on self-reports of hypothetical rather than actual behavior
  - Based entirely on information provided by persons who have already agreed to participate in a survey.
    - No information available from persons who have refused the initial survey request

Comparisons of Surveys with Varying Response Rates

- One form of “benchmarking”
- Approaches:
  - Compare results of surveys using rigorous vs. “standard” survey procedures (Keeter et al., 2000; 2006)
  - Compare survey results with high response rate, official government surveys (Krosnick, Holbrook & Pfent, 2003)
  - Findings of these studies suggest effects of non-response on empirical results were relatively modest in size
Comparisons of Surveys with Varying Response Rates

- **Limitations**
  - Often, surveys being compared employed different modes, respondent selection methods and/or question wordings
  - Results often only provide two response rate observations
  - In some of this research, even the “rigorous” surveys have relatively low response rates (c.f., Keeter et al., 2000; 2006)

Table 1. Comparison of Rates across Studies

<table>
<thead>
<tr>
<th></th>
<th>Standard 5-Day Survey</th>
<th>Rigorous Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Rate</td>
<td>36%</td>
<td>25%</td>
</tr>
<tr>
<td>Cooperation Rate</td>
<td>43%</td>
<td>34%</td>
</tr>
<tr>
<td>Contact Rate</td>
<td>90%</td>
<td>79%</td>
</tr>
<tr>
<td>Sample Size</td>
<td>(1,000)</td>
<td>(1,000)</td>
</tr>
</tbody>
</table>

**Note:**—Figures computed according to American Association for Public Opinion Research (AAPOR 2005) standard definitions of response rate 3, cooperation rate 3, and contact rate 2.

Comparisons of Respondents and Non-respondents Using Other External Data Sources

- Survey sample compared to other data sources believed to represent population of interest
- **Limitations:**
  - Can be many unknown differences in how statistics were compiled across various sources

---

Table 1

Comparison of physician sample to the American Medical Association’s distribution statistics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>AMA(^a)</th>
<th>Survey sample(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>30.2%(^c)</td>
<td>24.4%</td>
</tr>
<tr>
<td>Metropolitan(^f)</td>
<td>74.6%(^c)</td>
<td>70.1%</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>20.9%(^d)</td>
<td>23.5%</td>
</tr>
<tr>
<td>Midwest</td>
<td>22.4%(^d)</td>
<td>24.9%</td>
</tr>
<tr>
<td>South</td>
<td>33.8%(^d)</td>
<td>31.9%</td>
</tr>
<tr>
<td>West</td>
<td>22.9%(^d)</td>
<td>19.7%</td>
</tr>
<tr>
<td>Medical school graduation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within last 10 years</td>
<td>27.4%(^d)</td>
<td>13.0%</td>
</tr>
<tr>
<td>11–20 years ago</td>
<td>29.6%(^d)</td>
<td>33.6%</td>
</tr>
<tr>
<td>21–30 years ago</td>
<td>21.1%(^d)</td>
<td>26.0%</td>
</tr>
<tr>
<td>More than 30 years ago</td>
<td>21.9%(^d)</td>
<td>27.4%</td>
</tr>
</tbody>
</table>

\(^a\)Pasko, T., Seldman, B., and Birkhead, S. Physician Characteristics and Distribution in the US. American Medical Association: 2000 (34).
\(^b\)Sample weighted for physician specialty. Sample contains nonfederal physicians only.
\(^c\)Nonfederal Physicians 1998.
\(^d\)Total Physician (Federal and Nonfederal) 1998.
\(^e\)Contains AMA’s “Metropolitan” category and survey “Urban” and “Suburban” categories.

**Source:** Johnson et al. (2005) Substance Use & Misuse 40: 1071-1084.
Comparisons of Information from Sample Frame for Respondents vs. Non-respondents

- Sample frames are often available that contain rich information regarding all members of the population
- Excellent opportunity to both validate self-reports and make direct comparisons between respondents and non-respondents
- Many examples (Clusen et al., 2002; Goldberg et al., 2001; Krueter et al., 2010; Olsen, 2006; Oropsea & Landale, 2002; Russell, et al., 2004; Solberg et al., 2002; van den Akker et al., 1998)
- Typically the best approach to investigating nonresponse bias

Limitations

- Frames often derived from sources such as administrative records or medical records which are themselves subject to errors and omissions
- Can be difficult to negotiate access to frames
Comparisons of Respondents and Non-respondents via Record Matching

- Employs auxiliary data from other sources that offer complete coverage of the population being surveyed to compare responders with non-responders

- **Classic example:**
  Groves & Couper (1998)
  - Obtained matched decennial Census data for respondents and non-respondents to several other government surveys

Comparisons of Respondents and Non-respondents via Record Matching

- Typically investigate correlates of non-response only (Cantor et al., 2005; Dixon, 2004; Eyerman et al., 2002; Gfroerer et al., 1997; Needle et al., 1985; ZuWallack et al., 2004)

- **Other examples:**
  - In the U.S., match telephone area codes with zip codes and then merge with Census level data to compare neighborhood environments of respondents and non-respondents and their associations with key survey indicators (Johnson et al., 2006a; 2006b)
Comparisons of Respondents and Non-respondents via Record Matching

- **Limitations:**
  - Geographic matching continues to be an imprecise art
  - This approach may no longer be useful for evaluation of non-response in telephone surveys given the current dominance of mobile phone technology and the decline in numbers of housing units with landline phones

Summary of known correlates of likelihood of respondent contact

- **Social environment**
  - Urbanicity/population density (-)
  - Crime rate (-)
  - Multi-unit housing structures (-)
  - Physical barriers (-)
  - Employment rate (+)

- **Survey design**
  - Number of calls to household (-)
  - Weekday evening calls (+)
  - Pre-tailoring (+)

- **Interviewer**
  - Flexible interviewer schedule (+)
  - Interviewer self-confidence/positive attitude (+)

- **Household(er)**
  - More than one adult in household (+)
  - Young children in household (+)
  - Elderly in household (+)
Summary of known correlates of likelihood of respondent cooperation

- **Social environment**
  - Employment rate (-)
  - Population density (-)
  - Crime rate (-)
  - Age of HH members (-)
  - No children in HH (-)
  - One resident in HH (-)
  - Multi-unit housing structures (-)
  - Presidential approval ratings (+)
  - Percentage of population below age 20 (+)

- **Household(er)**
  - Socio-economic status (-)

- **Survey Design**
  - Introduction tailoring (+)
  - Scheduling appointments for later HH visits (+)

- **Interviewer**
  - Frequently change jobs (-)
  - Experience with same organization (+)
  - Self-confidence (+)
  - Self-monitoring (+)
  - Friendliness (+)
  - Positive attitude (+)
  - Perception job is easy (+)
  - Highly organized (+)
Open for Discussion

lkowens@uic.edu