Modeling Interviewer Effects in a Large National Health Study*

James Dahlhamer
Aaron Maitland
Benjamin Zablotsky
Carla Zelaya

Interviewers and Their Effects from a Total Survey Error Perspective Workshop

Lincoln, Nebraska
February 27, 2019

*This presentation does not represent official opinions or statistics of the National Center for Health Statistics or the Centers for Disease Control and Prevention.
Measuring Interviewer Variance (1)

- Interviewers can increase the variability of survey estimates when responses are clustered by interviewer (Kish, 1962)

- Measurement of interviewer variance in an ongoing, in-person survey is difficult

- Ideal design is fully interpenetrated where sample is randomly assigned to interviewers so one can assume that interviewers have equivalent assignments (West and Blom, 2016)
  - Partial interpenetration within smaller areas is also possible, *but modeling is often the most feasible approach*
Measuring Interviewer Variance (2)

- Multi-level models with cross-classified random effects for the interviewer and area have been used to provide estimates of interviewer variance (Schaeffer, Dykema, and Maynard, 2010)

- Models can control for differences in the characteristics of respondents across interviewer workloads and differences in the characteristics of the areas in which interviewers work
  - Interviewer level variables such as experience can be used to help explain variation due to the interviewer
  - Producing interviewer variance estimates from different types of questions helps to understand how question characteristics contribute to interviewer effects
Research Questions

- How much of the variance in National Health Interview Survey (NHIS) estimates can be attributed to interviewers?
  - Which questions/estimates are most susceptible to interviewer effects?
  - How do characteristics of questions (e.g., length, use of optional text, sensitivity) contribute to interviewer effects?
  - To what extent do interviewer effects vary by interviewer-level measures, such as experience, cooperation rates, and pace of interview?
National Health Interview Survey (NHIS)

- Multi-purpose household health survey
- In-person, CAPI interview (telephone follow-up, if necessary)
- Interviews conducted in ~35,000 households each year
Data/Methods

- Utilized 2017 NHIS data on sample adults (n=26,742)

- Estimated 125 outcomes across 102 questions using cross-classified, multi-level models
  - Respondents nested within interviewers, interviewers not uniquely nested within county
  - Compute intra-class correlation coefficients (ICCs) for interviewers for each outcome
  - For the 13 questions producing multiple outcomes, took average of ICCs

- Questions selected for inclusion based on the following criteria:
  - Difficult/complex topics (e.g., health insurance) vs. easier/less complex topics
  - Long vs. short questions
  - Definitions/clarifying text vs. no definitions/clarifying text
  - Optional text vs. no optional text
  - Sensitive vs. not sensitive
  - Demographic/factual vs. attitudinal/subjective
**Covariates Included in Multilevel Models**

<table>
<thead>
<tr>
<th>Respondent /case characteristics:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Sex</td>
</tr>
<tr>
<td>Education</td>
<td>Marital status</td>
</tr>
<tr>
<td>Cognitive difficulties</td>
<td>Mode of interview</td>
</tr>
<tr>
<td>Sample adult weights (log)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Area characteristics (county; 2012-2016 American Community Survey):</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent aged 65+</td>
<td>Percent under age 30</td>
</tr>
<tr>
<td>Percent Hispanic (quartile)</td>
<td>Percent non-Hispanic black (quartile)</td>
</tr>
<tr>
<td>Percent less than high school</td>
<td>Percent married</td>
</tr>
<tr>
<td>Percent no health insurance</td>
<td>Percent with a disability</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interviewer characteristics:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pace of interview (quartile)</td>
<td>Cooperation rate (quartile)</td>
</tr>
<tr>
<td>Order of sample adult interview (log)</td>
<td></td>
</tr>
</tbody>
</table>
Modeling Approach

- Model 1: unconditional model with random effects for interviewer
- Model 2: add random effects for county
- Model 3: add respondent/case characteristics
- Model 4: add interviewer characteristics
- **Final model: add county characteristics**
- All models estimated with SAS PROC GLIMMIX (binary outcomes) and SAS PROC MIXED (ordinal, continuous outcomes)
Summary of Interviewer ICCs for 102 Questions (Final Models)

- There were 13 questions that produced two or more outcomes
  - Averaged the interviewer ICCs for multiple outcomes from a single question

- Interviewer ICCs ranged from .0011 - .2273

- Mean interviewer ICC = .0507; median interviewer ICC = .0372

- 63.7% of questions produced an interviewer ICC < .05
RESULTS BY QUESTION CHARACTERISTICS
Median Interviewer ICC by Question Length: NHIS, 2017

- Quartile 1 (< 68 chars; n=24)ICC: 0.0291
- Quartile 2 (68-94 chars; n=26)ICC: 0.0225
- Quartile 3 (95-155 chars; n=26)ICC: 0.0358
- Quartile 4 (156+ chars; n=26)ICC: 0.0518
Example: “Shingles is an outbreak of a rash or blisters on the skin that may be associated with severe pain. The pain is generally on one side of the body or face. Shingles is caused by the chickenpox virus. A vaccine for shingles has been available since May 2006. Have you ever had the Zoster (ZOSS-ter) or Shingles vaccine, also called Zostavax®?”
Example of factual/demographic question: “Are you now married, widowed, divorced, separated, never married, or living with a partner?”

Example of an attitudinal/subjective question: “How worried are you right now about not being able to pay medical costs of a serious illness or accident? Are you very worried, moderately worried, not too worried, or not worried at all?”

Median Interviewer ICC by Factual/Demographic vs. Attitudinal/Subjective Questions: NHIS, 2017

Factual/demographic (n=85) 0.0313

Attitudinal/subjective (n=17) 0.0405
Example of a question deemed to be sensitive: “Which of the following best represents how you think of yourself? Gay, Straight, that is, not gay, Bisexual, Something else, or I don't know the answer?”
RESULTS BY INTERVIEWER CHARACTERISTICS
Median Interviewer ICC for 10 Questions by Interviewer Pace of Interview (Quartile): NHIS, 2017

- Pace (< 7.3 secs): 0.1767
- Pace (7.3 - < 8.5 secs): 0.0820
- Pace (8.5 - < 9.9 secs): 0.0644
- Pace (>= 9.9 secs): 0.0653
Median Interviewer ICC for 10 Questions by Within Interviewer Order of Interview (Quartile): NHIS, 2017
Summary/Conclusions

- Interviewer ICCs appear to be in line with other studies of in-person, interviewer-administered surveys

- Some question characteristics appear to be associated with interviewer effects
  - Larger ICCs for longer questions
  - Larger ICCs for attitudinal/subjective items (vs. factual/demographic items)

- Interviewer behavior as measured by pace of interview appears to influence responses
  - Larger ICCs for interviewers with the fastest pace of interview
  - Mixed results for within year experience and interviewer cooperation rates

- Implications for questionnaire design and interviewer training
Next Steps

- Incorporate additional interviewer characteristics and better measures of experience
- Add additional questions to the analysis
- Add paradata measures capturing interviewer effort, ease/difficulty of case
- Perform multivariate analyses to isolate question characteristics associated with interviewer ICCs
- Estimate full models, adding and removing key interviewer characteristics, to isolate their effects on ICCs
- Perform similar analysis comparing old and new questionnaire design
Discussion Questions

- How can we improve the methods/models?
  - Additional covariates?
  - Better methods?
  - Better software?

- How can we use these results to inform interviewer training?

- How can we use these results to inform questionnaire development?
Thank You!

jdahlhamer@cdc.gov