New Methodologies for the Study and Decomposition of Interviewer Effects in Surveys

Brady T. West
Michael R. Elliott
Survey Research Center
University of Michigan-Ann Arbor

Acknowledgements

- Thanks to Peter Miller, Alan Karr, and the ITSEW 2014 committee for accepting our work
- Thanks to the various junior faculty at SRC for fruitful discussions about these methods
- Thanks to the TSE15 committee for their continued hard work on developing and promoting what promises to be an excellent international conference on TSE!

Overview: Research Questions

- **RQ1:** What models can we use to study interviewer effects in the absence of interpenetrated designs?
  - Inflating variances for intra-interviewer correlations arising entirely from sample assignment leads to erroneous inferences!
- **RQ2:** If "total" interviewer variance arises from a combination of nonresponse error variance and measurement error variance among interviewers, what models can we use to study both error sources *simultaneously?*

Review of the Literature

- Statistical Methods for estimating correlated components of variance in the presence of non-interpenetrated designs:
  - Biemer and Stokes (1985)
  - Kleffe et al. (1991)
  - Gao and Smith (1998)
  - Von Sanden and Steel (2008)
- These studies assume *semi*-interpenetrated designs
- Common practice involves fitting multilevel models including area-specific covariates (Schaeffer et al. 2010)

Review of the Literature

- Decomposing total interviewer variance into measurement error variance and nonresponse error variance:
  - West and Olson (2010): telephone surveys
  - West, Kreuter, and Jaenichen (2013): FTF surveys
- These studies have followed a very simple descriptive approach using multilevel models, assuming interpenetration and independence of the two error sources within interviewers

Gaps in the Literature

- Methods for the estimation of interviewer variance components are needed in designs where there is complete (or nearly complete) lack of interpenetration (most FTF surveys)
- More elegant modeling methods are also needed for studying the decomposition of total interviewer variance into the separate error variance components

Method 1: Anchoring

- Consider a simple random sample…
- If cases with correlated values on a variable of interest are assigned to interviewers in a non-random fashion, we are just re-ordering the random sample given agents of the data collection process
- We have not altered anything about the actual data: no interviewer effects, no variance inflation
- Adjusting variance estimates for “interviewer” effects would lead to anti-conservative inferences

Method 1: Anchoring (cont’d)

- **Basic idea:** Identify an ancillary variable (“anchor”) that
  1. Is unlikely to be subject to interviewer effects in measurement (e.g., age)
  2. Is correlated with a key survey variable of interest that may be subject to interviewer effects
- Next, fit a model allowing the two variables to have correlated residuals, and including random interviewer effects ONLY for the survey variable
- This removes the portion of the within-interviewer correlation due to non-random assignment, leaving a “clean” estimate of the between-interviewer variance

Method 1: Anchoring (cont'd)

- In the simplest case, we have two variables, one (Y_1) treated as measurement error-free, and one (Y_2) treated as possibly having interviewer-induced measurement error
- We fit a model of the form

$$y_{ijk} = \mu_k + I(k=2)b_i + \varepsilon_{ijk}$$

where *i* indexes interviewers, *j* indexes respondents within interviewers, *k* indexes the variable, $b_i \sim N(0,\tau^2)$ and $\begin{pmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right)$.

Method 1: Anchoring (cont'd)

- Standard linear mixed model software can be used to obtain a REML point estimate of the mean for the second variable, along with an estimated variance component (we used PROC MIXED)
- High correlation between the residuals will lead to a more accurate estimate of the variance component, and thus of the true impact of the interviewer-induced measurement error on the variance of the estimated mean

Anchoring: An Illustration

- In a preliminary example using data from the 2012 BRFSS, we assume that the measurement error free variable is **age** and the variable of interest is **perceived health status** (1 = poor, …, 5 = excellent)
- We choose age as an anchoring variable because:

  a) We believe it is likely to be reported without differential measurement error,

  b) It is associated with interviewer assignment, as interviewers tend to work shifts at different times of the day, and interview time of day is associated with age, and

  c) It is also associated with health status.

Anchoring: An Illustration (cont'd)

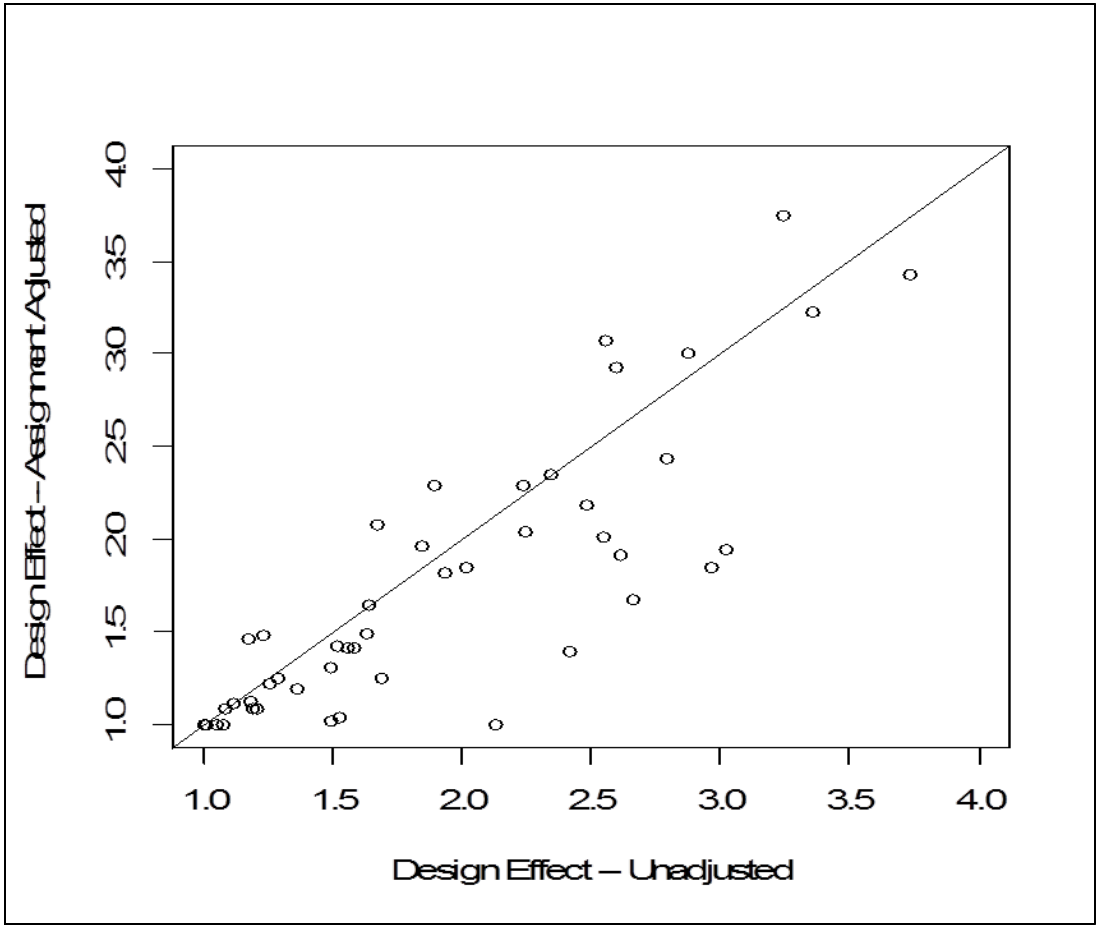
- We compute the interviewer effect for mean health status in each of the 50 states:

  a) assuming an interpenetrated design, and

  b) using the assignment-adjusted random effect, where age is assumed to be reported without measurement error

- The mean interviewer effect was 2.09 in the unadjusted analysis vs. 1.86 in the adjusted analysis
- Here is a picture of the results…

Anchoring: An Illustration (cont'd)



- While most states were unchanged in their interviewer effect estimates, 9 states had decreases of more than 25%, including one of 53%!

Multivariate Extensions

- For analysts interested in regression models, we can use standard software to fit the model

$$y_{ijk} = \beta_{0k} + \sum_{l=1}^{p} \beta_{lk} x_{ij} + b_{iK0} + \sum_{l=1}^{p} b_{iKl} x_{ij} + \varepsilon_{ijk},$$

with $i = 1,..,I,\ j = 1,...,J_i, k = 1,...,K$, where $y_{ij1},....,y_{ijK-1}$ are assumed free from measurement error, $(b_{iK0} \cdots b_{iKp})^T \sim N_{p+1}(0,D)$, $(\varepsilon_{ij1} \cdots \varepsilon_{ijK})^T \sim N_K(0,\Sigma)$, and the variance-covariance matrices are unstructured.

Method 2: Assignment Propensity

- An alternative possible approach to dealing with this problem adapts propensity score adjustment methods to develop “assignment weights,” equal to the inverse of the probability of assignment to a given interviewer
- The assignment probability is estimated as a function of covariates known to be (or treated as) free of measurement error

Method 2: Assignment Propensity

- Use of these weights in estimation “re-creates” an approximate interpenetrated design
- The approximation will improve to the degree that the covariates capture all of the assignment mechanism that is correlated with the measurement error
- The weights would be estimated by fitting a multinomial model to respondent data, with interviewer ID as the dependent variable

Method 2: Assignment Propensity

- If the covariates are unrelated to assignment, then the weights will be approximately equal (i.e., the design is essentially interpenetrated)
- If there are strong associations of the covariates with assignment, the weights should remove the impact of the assignment from analysis
- If sampling weights are present, these would be used for estimation of the multinomial model, and the sampling weights would be multiplied by the assignment weights to obtain final analysis weights

Method 3: MLMI

- MLMI = A Multi-Level Multiple Imputation approach to decomposing total interviewer variance
- We need a *simultaneous* modeling method for decomposing total interviewer variance into measurement error and nonresponse error variance, while also enabling estimation of the covariance of these two error sources (among interviewers)

Method 3: MLMI

- Outline of the proposed approach:

1. Form a data set using the full sample, including true values of Y (for the full sample) and reported values of Y for respondents (note: generally **rare** to have true Y values for the full sample)
2. Create a (1, -1) variable (X), where 1 = respondent, and -1 = non-respondent. Ignoring interviewer effects, fit a regression model to the true values of Y using the (1, -1) variable, which will produce the full sample mean (B0) and the nonresponse error (B1)

Method 3: MLMI

3. Allowing B0 to randomly vary across interviewers will capture variance in assignment (should be zero for interpenetrated assignment); allowing B1 to vary across interviewers will capture variance in the nonresponse errors; we can also allow the random effects to covary!
4. Non-respondents will have missing values on *reported* Y; use true values, interviewer effects, and other auxiliary variables to *impute* reported values of Y for non-respondents

Method 3: MLMI

5. In each imputed data set, fit a model to the *reported values* of Y (including the imputed values), allowing the intercept and the “response” effect to randomly vary across interviewers
6. **Assuming interpenetration**, we can now estimate the variance of the intercepts (measurement error), the variance of the “response” effects (nonresponse error), and the covariance of these effects across interviewers

- NOTE: we could consider the assignment propensity approach if interpenetration was not evident…

Method 3: MLMI

- Clearly the success of this method depends on some key features of the available data:
  - A rich sampling frame including true values on selected variables
  - Auxiliary variables that are strongly predictive of survey reports (for good imputation models)
- We are eager to apply this approach to existing data sets (which again will be rare)…this is the next (methodological) step

Discussion Points

- What does everyone think about these ideas?
- How would you suggest that we refine these approaches?
- Are we missing any other developments in the literature in these areas?
- Empirical applications are certainly needed! We are just at the idea phase right now.