

Class 4: Survey Error & Bias Correction

MAST5953: Creating Your Own Data

Dr. Miriam Sorace

www.miriamsorace.net

01 December 2021

Outline of Today's Class

Survey Design: Revision

Total Survey Error

Survey Weighting

Survey Design: Revision

Exercises

- ▶ In an effort to gauge public support for energy conservation, the (fictional) Andrews Foundation conducted a recent poll that found that 72% of Americans agreed with the following statement (phrased as an "Agree/ Disagree" question):
 - ▶ "I would support President Obama's decision to use the U.S. military to help local cities achieve energy independence by installing more energy efficient public lighting."
- ▶ Does it meet the analytic goal of assessing public support for installing more energy efficient public lighting? Why or why not?

Exercises

- ▶ Describe two ways you could reduce the effects of social desirability when asking respondents to report their attendance of religious services.
- ▶ Describe two different kinds of analyses one could do to assess the validity of answers to questions designed to measure a subjective state, such as happiness or anxiety.

Total Survey Error

'Survey quality is more than a margin of error' ¹

- ▶ The following two requirements must be met:

¹Wolf et al., 2016: Chapter 3

'Survey quality is more than a margin of error' ¹

- ▶ The following two requirements must be met:
 1. Answers accurately describe respondent's characteristics, behaviours, attitudes (*measurement*).

¹Wolf et al., 2016: Chapter 3

The Total Survey Error Framework

'Survey quality is more than a margin of error' ¹

- ▶ The following two requirements must be met:
 1. Answers accurately describe respondent's characteristics, behaviours, attitudes (*measurement*).
 2. Respondents must be representative of the population (*representativeness*).
- ▶ But ... in reality survey analysts often have to grapple with:

¹Wolf et al., 2016: Chapter 3

Measurement Error

The Total Survey Error Framework



Support us

Contribute

Subscribe

NEWS INDEPENDENT TV CLIMATE SPORT VOICES CULTURE PREMIUM INDY/LIFE INDYBEST INDY100 MY INDEPENDENT VOUCHERS COMPARE CYBER M

'Shy Tory' effect, says leading pollster

'My sense is that certain approaches to polling [prior to election day] will have underestimated the amount of people likely to vote Conservative,' says ComRes pollster

Maya Oppenheim | Thursday 08 June 2017 17:27 | comments



Sampling Error

The Total Survey Error Framework

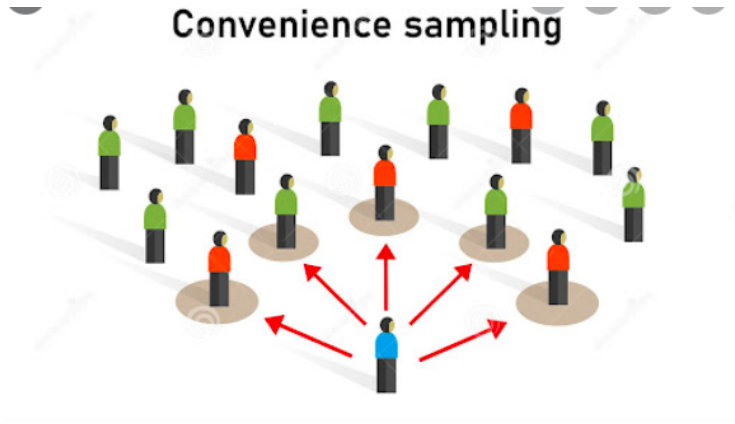


Image from: [ResearchArticles.com](https://www.researcharticles.com)

Coverage Error

The Total Survey Error Framework



Nonresponse Error

The Total Survey Error Framework

RESPONSE RATES

187

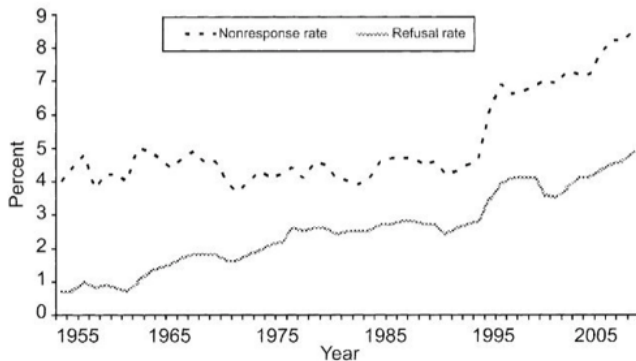


Figure 6.2 Nonresponse and refusal rates for the Current Population Survey by year. (Source: U.S. Census Bureau.)

Interviewer Effect

The Total Survey Error Framework

Journal of Survey Statistics and Methodology (2017) 5, 175–211

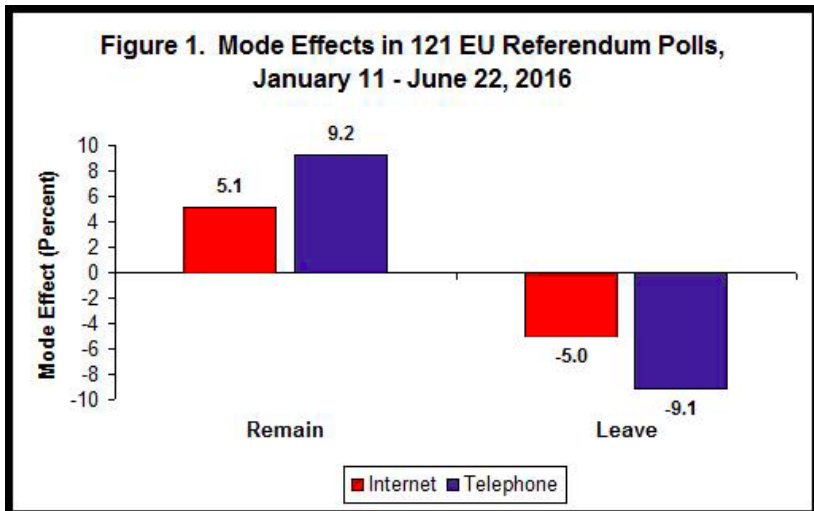
EXPLAINING INTERVIEWER EFFECTS: A RESEARCH SYNTHESIS

BRADY T. WEST*
ANNELIES G. BLOM

A rich and diverse literature exists on the effects that human interviewers can have on different aspects of the survey data collection process. This research synthesis uses the Total Survey Error (TSE) framework to highlight important historical developments and advances in the study of interviewer effects on a variety of important survey process outcomes, including sample frame coverage, contact and recruitment of potential respondents, survey measurement, and data processing. Included in the scope of the synthesis is research literature that has focused on *explaining* variability among interviewers in these effects and the different types of variable errors that they can introduce, which can ultimately affect the efficiency of survey estimates. We first consider common tasks with which human interviewers are often charged and then use the TSE framework to organize and synthesize the literature discussing the variable errors that interviewers can introduce when attempting to execute each task. Based on our synthesis, we identify key gaps in knowledge and then use these gaps to motivate an organizing model for future research investigating explanations for interviewer effects on different aspects of the survey data collection process.

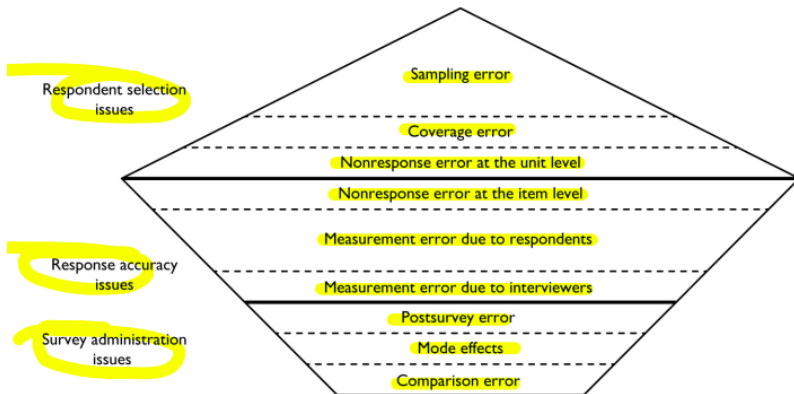
Mode Effect

The Total Survey Error Framework



The Total Survey Error Framework

Figure 3.1 The different types of survey error source. Source: Adapted from Weisberg (2005, p. 19).



Are We Doomed?

- ▶ “The job of a survey designer is to minimize error in survey statistics by making design and estimation choices that minimize the gap between two successive stages [measurement and representativeness, *ndr*] of the survey process.” (Groves et al 2009: 49)
- ▶ “Despite all these potential sources of error, carefully designed, conducted, and analyzed surveys have been found to be uniquely informative tools” (Groves et al 2009: 41)

Error-Minimisation Strategies

Measurement Error & Item Non-Response

- ▶ Survey pilots/cognitive interviewing
- ▶ Telephone/face-to-face modes allow for clarifications
 - ▶ But be wary of interviewer characteristics for sensitive topics!

Error-Minimisation Strategies

Interviewer effects

- ▶ Self-administered surveys
- ▶ Interviewer training

Error-Minimisation Strategies

Sampling/Coverage Error

- ▶ Probability sampling
- ▶ Adequate sampling frames (ideally official/government lists)

Error-Minimisation Strategies

Nonresponse

- ▶ highlight academic sponsorship - purpose is to help research
... Universities generally perceived more favourably
- ▶ Cash incentives
- ▶ Choose a long data collection period - to give time to everyone
- ▶ Put sensitive questions at the end of the questionnaire
- ▶ Keep questionnaire short
- ▶ Contact non-responders repeatedly - using different mediums
(pre-notification letter, persuasion letters, change to
face-to-face mode ...)

Exercises

Nonresponse

- ▶ Identify a method used to increase response rates
- ▶ Thinking of the adult household population, identify a subgroup for whom the method does not appear to be as effective as for other subgroup and explain why.

Exercises

Minimising nonresponse

- a) Assume that your goal is to maximize cooperation rates. Choose whether to use the sponsoring organization's letterhead (the organization paying for the survey) or the data collection organization's letterhead (the organization collecting the data, which you can assume to be relatively unknown among sample persons); then explain the reason for your decision.

Sponsor	Data Collection Organization	Target Population	Which organization's stationery for advance letter?	Reason for answer
Commercial credit company	Academic survey center	U.S. households		
Federal government	Commercial market research firm	Low-income households		
Highway Construction lobbying organization	Nonprofit research organization	Presidents of highway construction companies		

Exercises

Minimising nonresponse

b) Now assume that you are interested in minimizing the nonresponse bias of the following estimates for each survey.

Sponsor	Data Collection Organization	Target Population	Which organization's stationery for advance letter?	Reason for answer
Commercial credit company	Academic survey center	U.S. households	Mean amount of credit card debt	
Federal government	Commercial market research firm	Low-income households	Percent with favorable attitudes about food stamp program	
Highway Construction lobbying organization	Nonprofit research organization	Presidents of highway construction companies	Percentage judging lobbying effort effective	

Post-Survey Statistical Adjustments

- ▶ weighting
- ▶ multiple imputation

Weighting

- ▶ “ [...] the adjustment of computations of survey statistics to counteract harmful effects of noncoverage, nonresponse, or unequal probabilities of selection into the sample” (Groves et al. 2009: 331)
- ▶ “Weighting of survey data is required to ‘map’ the sample back to an unbiased representation of the survey population” (Heeringa et al. 2017: 38)

Differential Selection Probability Weighting

Example

- ▶ Need to collect a sample of 125k of the US population (total 199,500,000)
 - ▶ The proportion of Latinos in the population is $1/8$ (or 12.5%)
- ▶ Need to over-sample Latinos (raising their probability of selection to $1/2$, or 50%) to carry out some group-specific analysis
 - ▶ Instead of 15,625 Latinos in the sample, we will have 62,500
- ▶ **Problem:** the aggregate descriptive statistics are not representative of the US population unless respondents are weighted to bring back the Latino probability to $1/8$. (any analysis separate by sub-group would be fine though).

Differential Selection Probability Weighting

Example

- ▶ What to do?
 - ▶ Simply put: Weight each observation by its original proportion in the population.
 - ▶ Latinos weight = 0.125
 - ▶ Non-Latinos weight = $(1-0.125) = 0.875$
 - ▶ Which more simply means that each Non-Latino observation needs to be multiplied by 7 (i.e. assign Latinos a weight of 1 and Non-Latinos a weight of 7):
 - ▶ Non-Latinos weight = $0.875/0.125 = 7$

Differential Selection Probability Weighting

Example

- ▶ What to do?
 - ▶ More technically: weight each observation by the *inverse of its probability of selection*
 - ▶ Latinos: $1/(\text{their } N \text{ in sample} / N \text{ tot Latino population}) = 1/[62500/(199,500,000 \cdot 0.125)] = 1/(62,500/24,927,500) = 399$
 - ▶ Non-Latinos: $1/(\text{their } N \text{ in sample} / N \text{ tot Non-Latino population}) = 1/[62500/(199,500,000 \cdot 0.875)] = 1/(62,500/174,562,500) = 2,793$
 - ▶ Which also translates to assigning a weight of 7 to Non-Latinos and a weight of 1 to Latinos $(2793/399) = 7$

Non-Response Adjustment Weighting

Example

- ▶ Imagine that Latinos respond at an 80% rate and Non-Latinos respond at a 99% rate, if one assumes that - within Latinos and Non-Latinos - respondents are a random sample of all sampled persons in the group (missing at random assumption), this 0.80 or 0.99 probabilities can be thought of as *sampling rates*
- ▶ What to do?
 - ▶ Weight each observation by the *inverse of its sampling rate*:
 - ▶ Latinos: $1/0.80 = 1.25$
 - ▶ All the rest: $1/0.99 = 1.01$
 - ▶ Which then can be cumulated to the original sampling weights via multiplication, to form the below weights:
 - ▶ Latinos: $1 * 1.25 = 1.25$
 - ▶ Non-Latinos: $7 * 1.01 = 7.07$

Weighting

- ▶ “Generally, the *final survey weights* in survey data sets (sometimes referred to as *estimation weights*) are the product of sample selection weights, non-response adjustment factor, and the poststratification factor” (Heeringa et al. 2017: 38)
 1. Apply *design weights* - i.e. adjustments for differential sampling probabilities
 2. Apply *non-response corrections* - i.e. response propensities
 3. Apply *calibration methods*: i.e. raking or poststratification to make sample conform to known auxiliary (extra-survey) population variable distributions
- ▶ “The computation of [estimation weights] is therefore an accounting function, requiring only multiplication of the probabilities of selection at each stage of sampling and then taking the reciprocal of the product of the probabilities” (Heeringa et al. 2017: 39)

Applied Example

ESS

► ESS Weights

Exercise

- ▶ What type(s) of survey errors could be adjusted for in the process of post-stratification?

Multiple Imputation

- ▶ This method is used for *item non-response*
- ▶ Typically analysts/statistical softwares ignore the missing cases, i.e. engage in *casewise deletion* of missing data
- ▶ Effectively, casewise deletion is a form of imputation: it simply assumes that each of the missing cases have average values on each covariate
- ▶ An alternative is building an explicit multiple imputation model, although many consider this alternative akin to “data fabrication”.

Multiple Imputation

- ▶ It entails regression prediction for the missing value.
- ▶ It requires that all variables used as predictors in the analytical model are also included on the right-handside as predictors of the missing value.
- ▶ It generates series of multiply imputed datasets each with slightly different realizations of the multiple imputation model
- ▶ Variation in estimates across the multiple datasets then allows for the estimation of overall variation

What did we learn today?

- ▶ The various errors that can threaten survey quality
- ▶ How to minimise the total survey error
- ▶ What are survey weights
- ▶ What is multiple imputation