Class 4: Survey Error & Bias Correction

MA5953: Creating Your Own Data

Dr. Miriam Sorace

www.miriamsorace.github.io

20 November 2023

Outline of Today's Class

Total Survey Error

Survey Weighting

Total Survey Error

The Total Survey Error Framework

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

- The following two requirements must be met:

[1] Wolf et a;. 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*).

[1] Wolf et a;. 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*).

[1] Wolf et a;. 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:

[1] Wolf et a;. 2016: Chapter 3

Measurement Error

The Total Survey Error Framework

Sampling Error

The Total Survey Error Framework

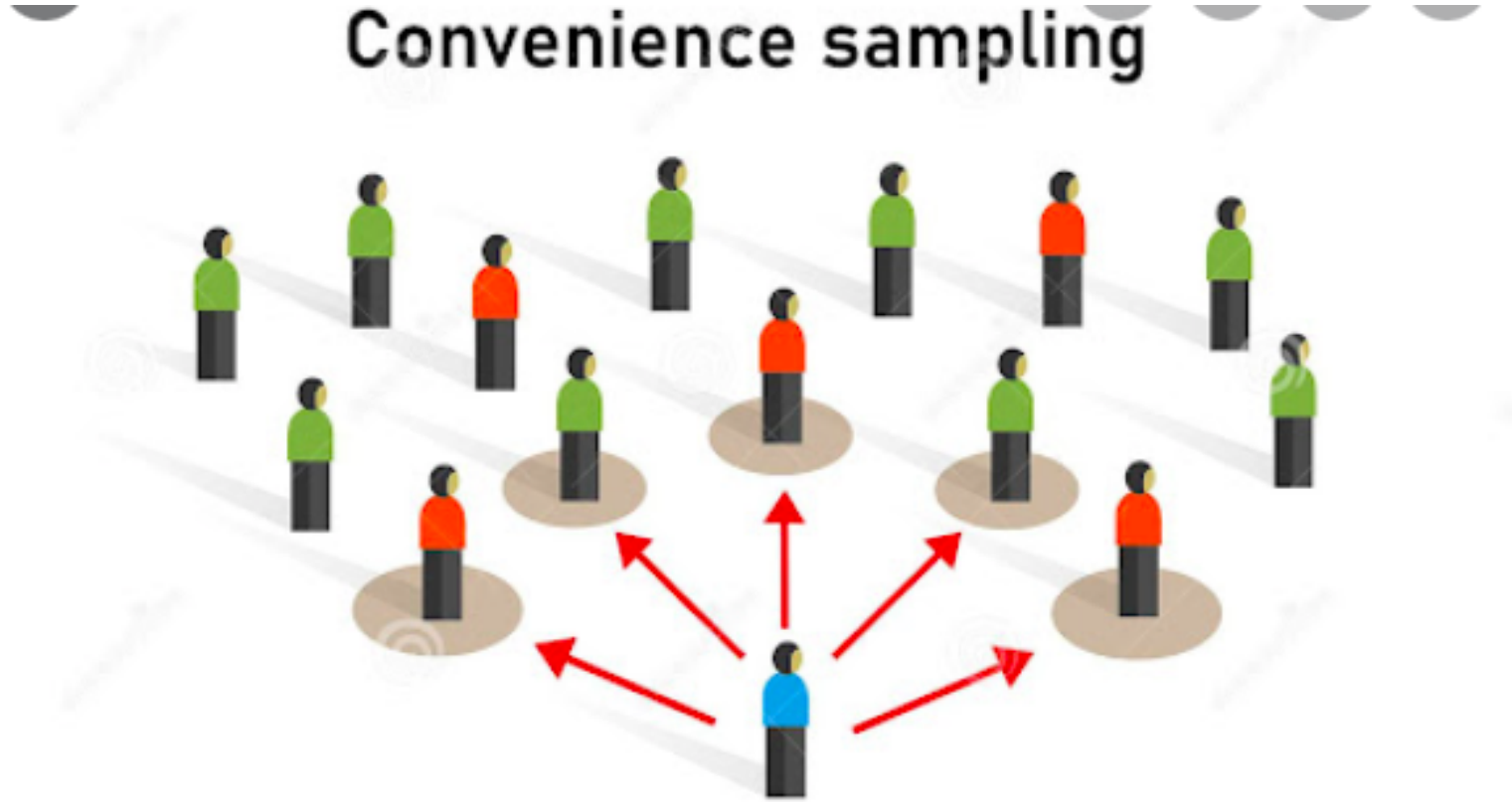


Image from: ResearchArticles.com

Coverage Error

The Total Survey Error Framework

Nonresponse Error

The Total Survey Error Framework

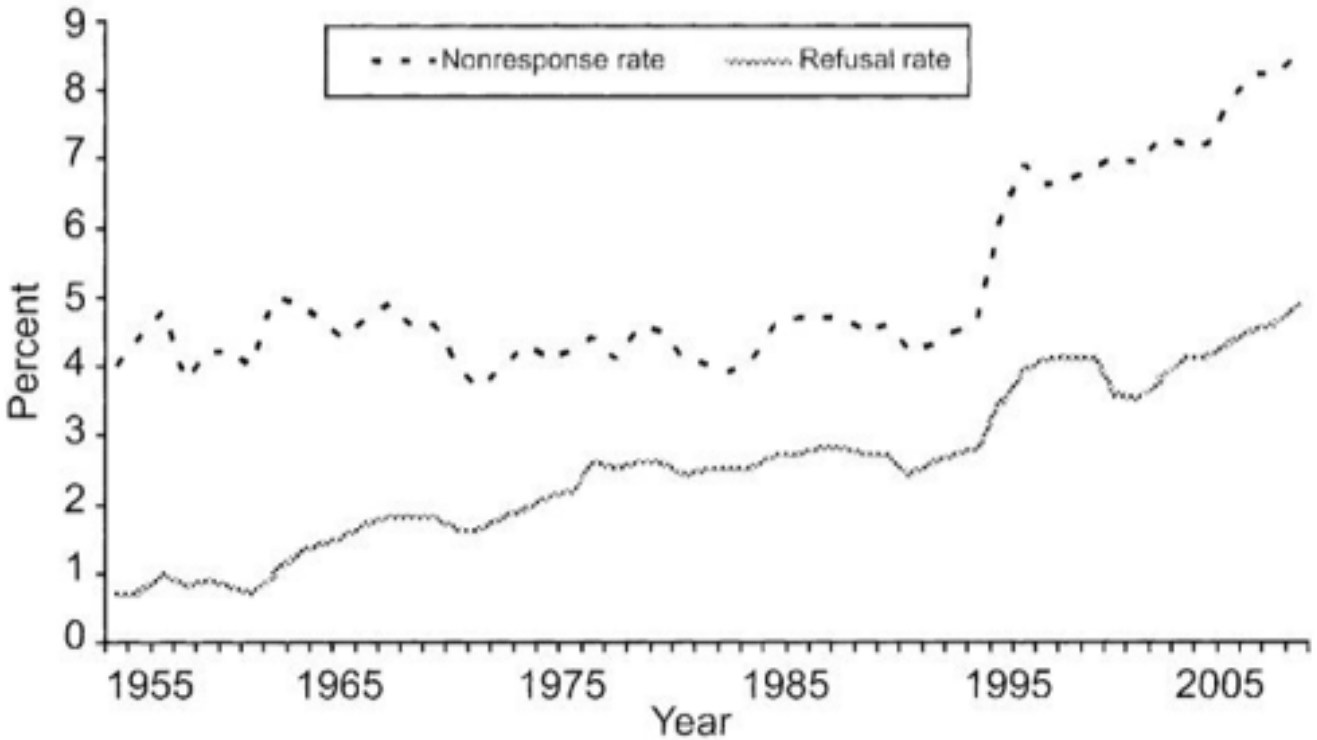


Figure 6.2 Nonresponse and refusal rates for the Current Population Survey by year. (Source: U.S. Census Bureau.)

Interviewer Effect

The Total Survey Error Framework

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

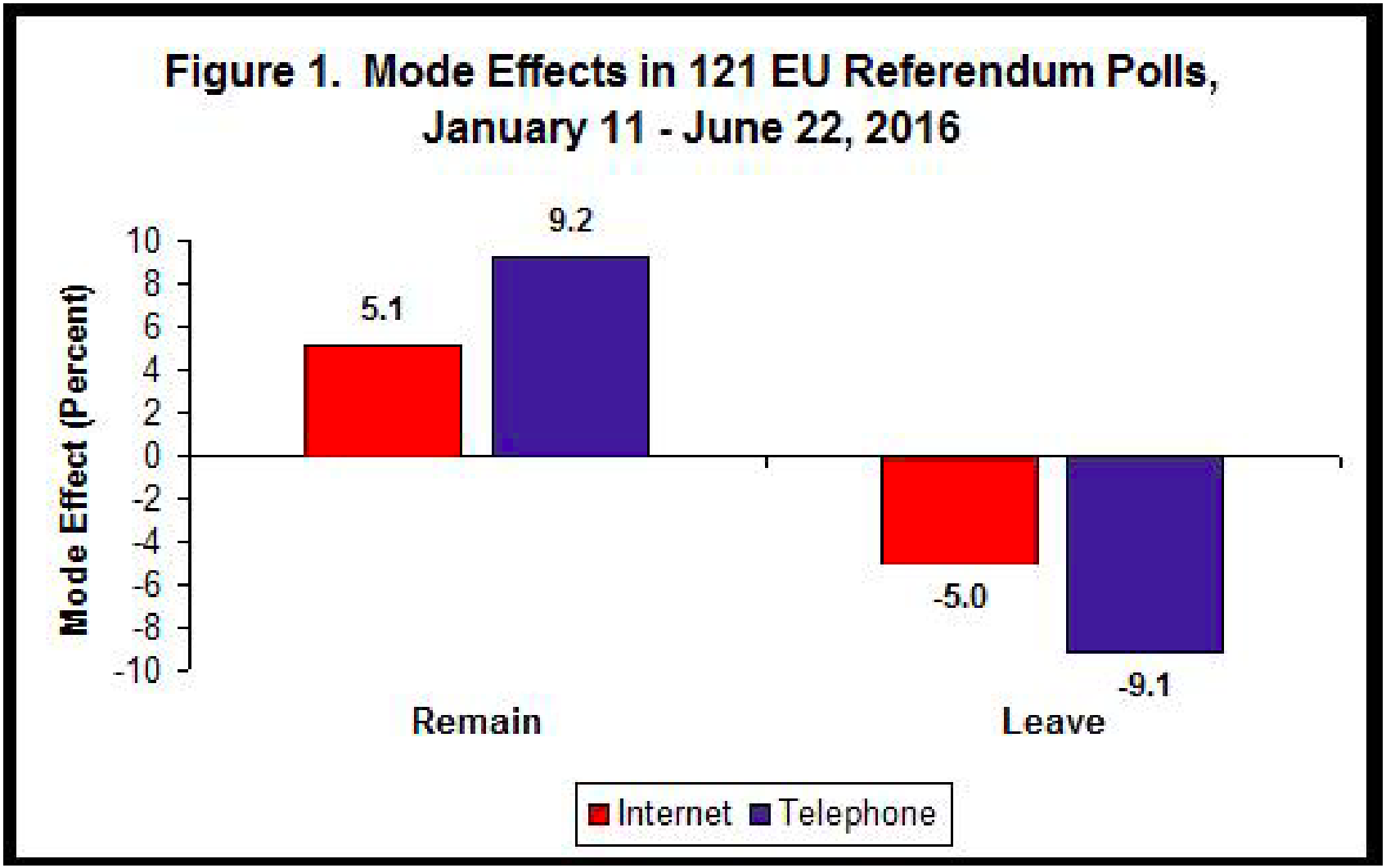
EXPLAINING INTERVIEWER EFFECTS: A RESEARCH SYNTHESIS

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

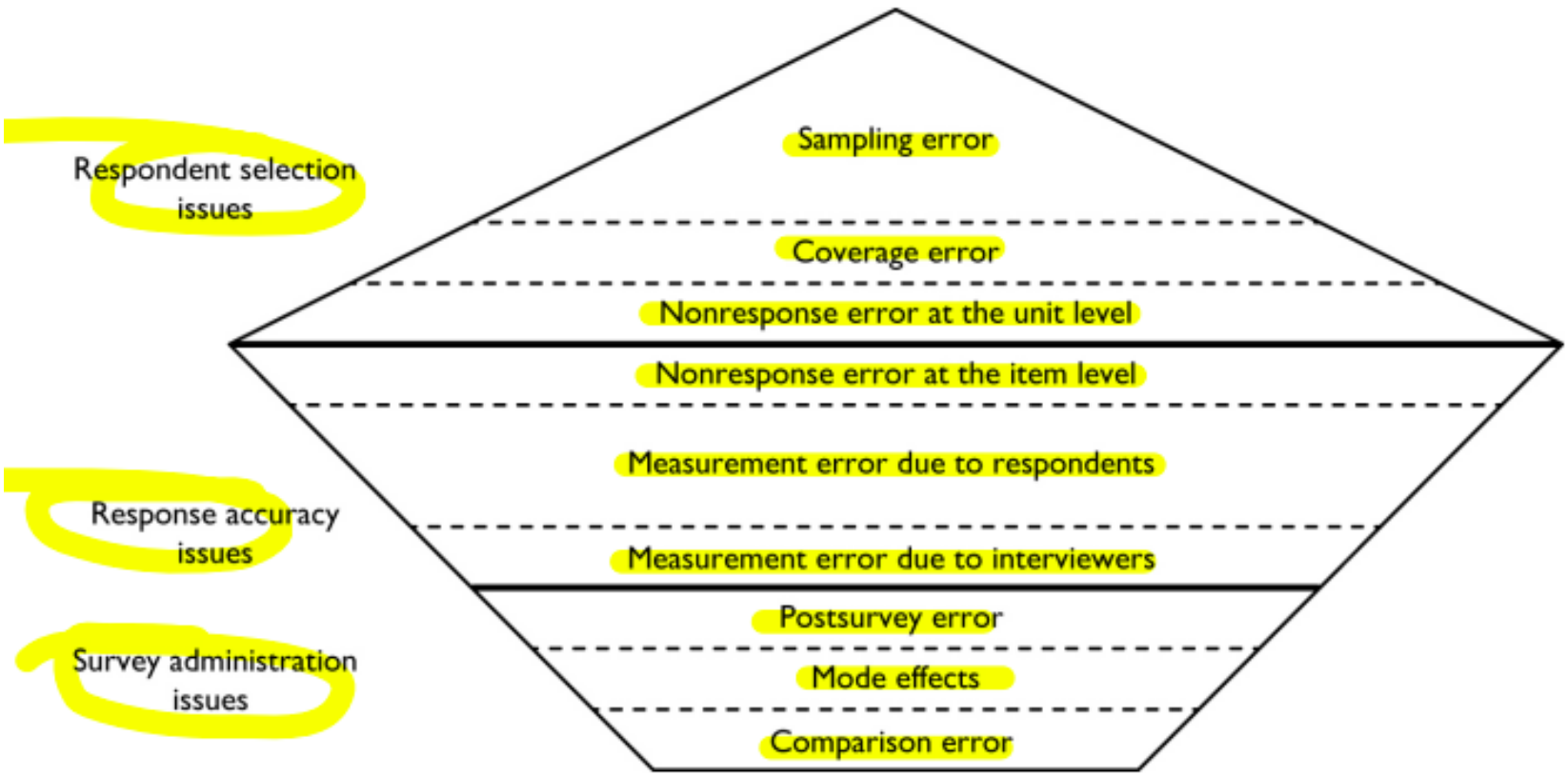
The Total Survey Error Framework



From: UK in a Changing Europe

The Total Survey Error Framework

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



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 ...)

Survey Weighting

Post-Survey Statistical Adjustments

- weighting
- multiple imputation

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/(their N in sample / N tot Latino population) = 1/[62500/(199,500,000*0.125)] = 1/(62,500/24,927,500) = 399
    - Non-Latinos: 1/(their N in sample / N tot Non-Latino population) = 1/[62500/(199,500,000*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)

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