Probability Sampling, Non-Probability Sampling and Nonresponse

J. Michael Brick

Westat

October 2023

The Goal of Probability Sampling

- The primary purpose of a probability sample survey is to accurately measure characteristics of a finite population at a particular point in time.
- Probability sample data may also be used for other analytic goals like studying casual relationships, but these analysis require additional assumptions or data.

Key Features of Probability Sampling

- The characteristics of the finite population are fixed.
  - *y*'s are not random variables
- The sample design (process of selecting the sample) is the random component and the basis for inference.
- The theory to support inferences requires large sample sizes.

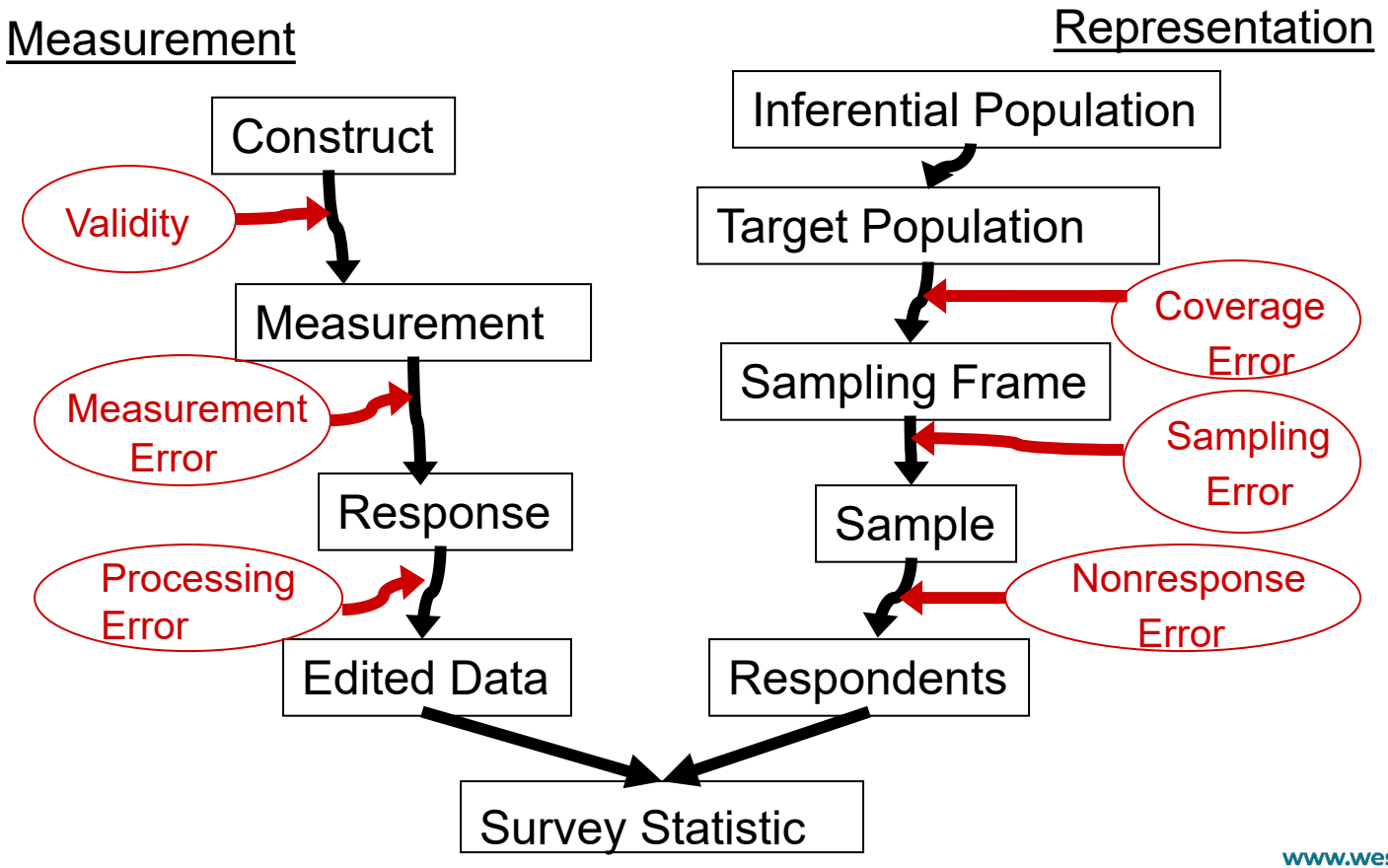
Examples of Probability Sampling

- Stratified samples of lists of providers of childcare.
- Random-digit dialing sample of telephone numbers.
- Address-based sample of households.
- Multi-stage samples of areas, segments with the areas, and households and adults within the segments.

Total Survey Error

- Collecting data is complex and errors (deviations from the ideal) can be introduced in all aspects of the process.
- Probability sampling is a unique and powerful tool to address one, albeit a very important, source of error associated with representing the population.

Total Survey Error Perspective



Non-Probability Sampling

- Any method of selecting units that does not permit computing a non-zero probability of selection for each unit in the population.
  - Examples: Mall intercept surveys, quota samples, samples of emails from aggregators, (generic) internet panel surveys
- With non-probability samples response rates and coverage rates are not well-defined.
- More generally, the causes of missingness (not having data on all units of the population) cannot be determined.
  - Example: In an internet panel sample we do not know if people are missing because they do not have internet, do not respond to requests, or were not surveyed for any other reason.

The Slippery Slope

- Methods of sampling that have some probabilistic mechanism but are not probability samples. Specifically, most of these methods may use some random component but requires assumptions to compute a "probability of inclusion."
  - Examples: Location sampling for rare groups, Respondent-driven sampling, Probability-based panels.
  - Weighting methods are often assumed to reduce the errors introduced by not knowing the real probability of selection and residual bias is difficult to evaluate.
- The effects of departures from probability sampling are harder to conceptualize and evaluate. Most evaluations are empirical.

Biggest Threat to Representation in Probability Samples

- Probability sampling solves the representation problem and provides a way to measure expected deviations from population quantities, under ideal conditions.
- Nonresponse is not ideal.
- Three types of nonresponse
  - Item, unit, and partial nonresponse (longitudinal pr panel surveys)
- Noncoverage may be a bigger threat than nonresponse, especially in a specific surveys. Coverage error is difficult to evaluate because the "missing" data is not available and special studies are needed.
- Here we focus on unit nonresponse

Nonresponse Mechanisms

- Examination of psychological and social determinants of willingness to participate
  - Tourangeau et al. (2000)
  - Goyder (1987)
- Translating these factors into data collection and weighting strategies
  - Dillman (1978)
  - Singer (2002)
  - Särndal and Lundström (2005)

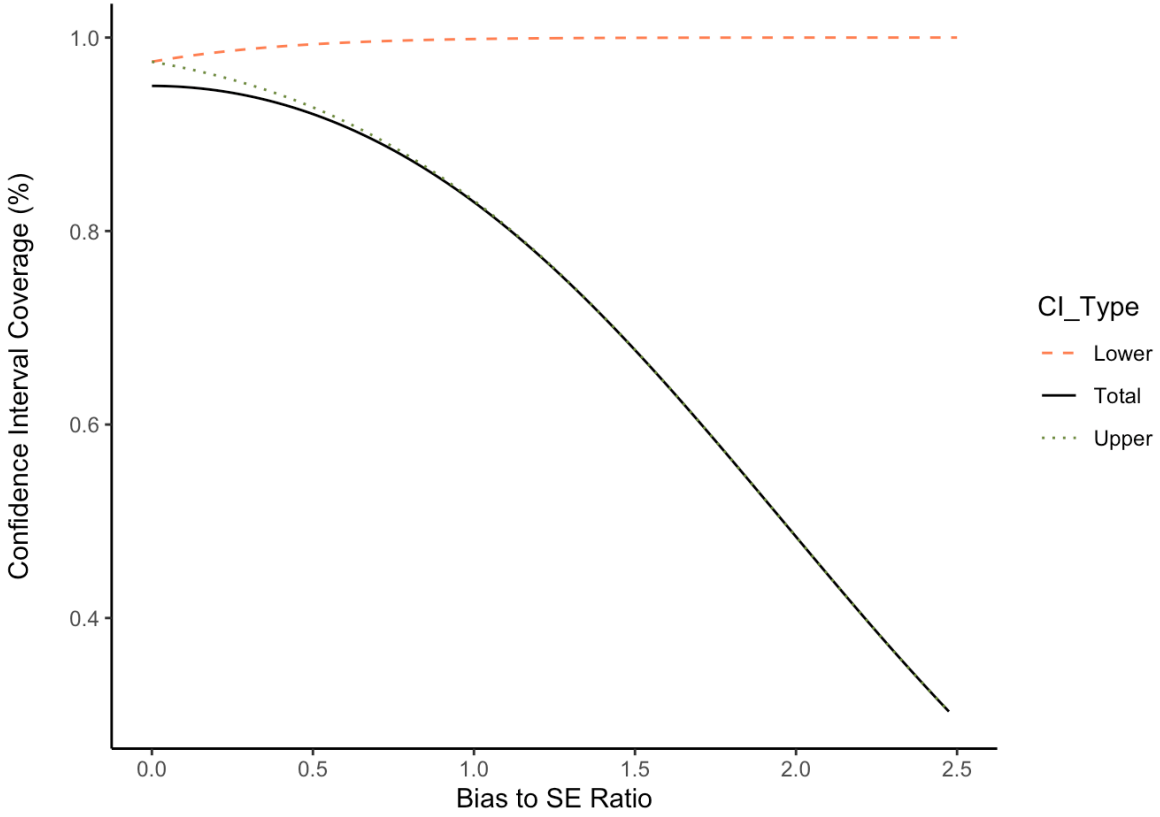
Properties of Bias

- Bias is difference between expected value of estimate with the observed response rate and the value with 100% response rate.
  - Bias is an average over all possible samples, not a specific outcome.

- Bias depends on the statistic (specific *y* variable and type such as a mean or total)
  - It depends on the specific *y* and response propensity distribution.

Perspective on Bias

- Why does bias matter?
  - It affects point estimates directly
  - It is largely unaccounted for in statements of the accuracy of the estimates
  - It causes confidence intervals to cover with less than prescribed probabilities
- Consider the bias relative to the standard error of the estimate and its effect on confidence intervals.


Confidence Interval Coverage (%)
1.0
0.8
0.6
0.4
0.0
0.5
1.0
1.5
2.0
2.5
Bias to SE Ratio
CI_Type
Lower
Total
Upper

Stochastic Representation of Bias

$$Bias(\bar{y}_r) = \frac{\sigma_{yp}}{\bar{\phi}} = \left(\frac{\rho_{y\phi}}{\bar{\phi}} \right) \sigma_y \sigma_\phi$$

where σ_{yp} = covariance between y and response propensity, ϕ

$\bar{\phi}$ = mean propensity over the sample

$\rho_{y\phi}$ = correlation between y and ϕ

σ_y = standard deviation of y

σ_ϕ = standard deviation of ϕ

Brick and Tourangeau Response Propensity (2017) Typology

- Random propensity model – largely random (respondent is sick, likes interviewer, etc.)
- Design-driven propensity model – largely due to design features (number of callbacks, modes, incentives, etc.)
- Demographic-driven propensity model – largely component related to demographics
- Correlated response propensity model – largely due to design features and non-demographic characteristics of respondents

Kreuter et al. (2010)

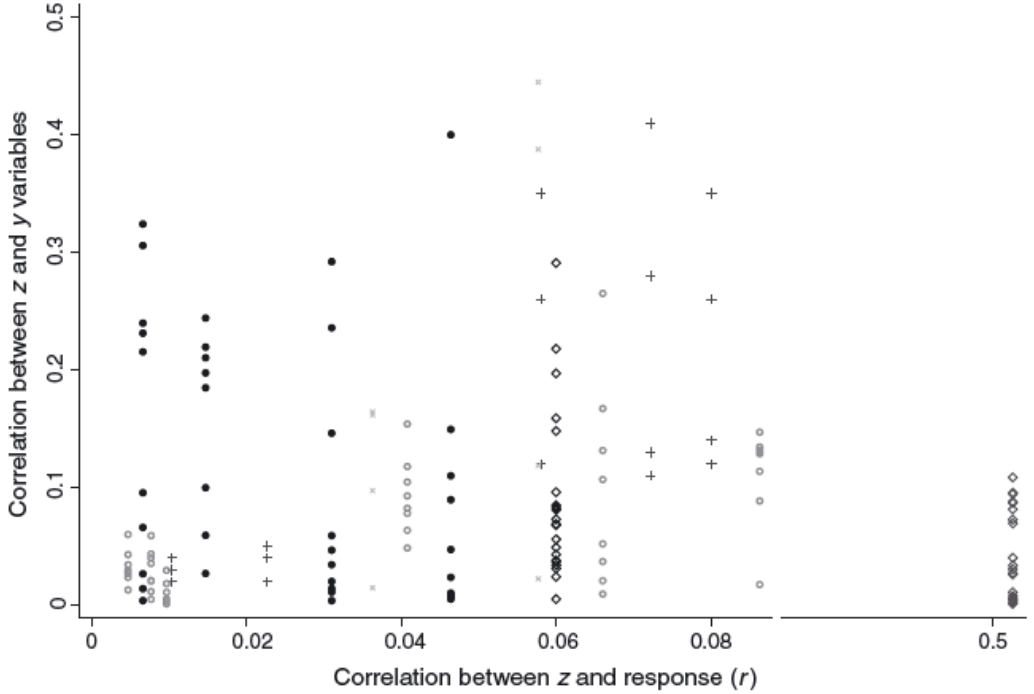


Fig. 1. Relationship between the correlation of *z*- and *y*-variables and correlation of *z* and response in five surveys (all correlations are shown as absolute values): +, UMTRI; ×, MEPS; ○, ESS; ◇, ANES; ●, NSFG

Key to Identifying Potentially Sizeable Biases

- Understand reasons for nonresponse and potential relationships between the *y*'s and the response mechanisms
- Examples:
  - Altruism is related to high response so voting, volunteering, and civic engagement may have large nonresponse biases
  - Some survey topics (sponsors) may affect participation differentially, and induce bias like cancer and health

Hedlin (2020) 'Safe Area'

- Nonresponse bias is not linear with response rates
- Large biases more likely to occur with lower response rates (less than 30 to 40%).
  - Nonresponse adjustments may keep biases relatively small with even somewhat lower response rates.
- Some specific statistics correlated to response may have large biases even in high response rate surveys.

Weighting

- Nonresponse weighting will ‘eliminate’ nonresponse bias in few situations.
  - Efforts to adjust weights are nonetheless worthwhile and may reduce ‘large’ biases substantially (by 50% or more).
  - Adjustments are generally less impactful for smaller biases.
- The key to successful nonresponse adjustment is powerful auxiliary variables and using them wisely.

Summary

- Probability sampling remains the most reliable approach to achieve accurate estimates of a finite population.
  - Requires reasonable response rates.
  - Data collection cost will be higher than with non-probability samples.
  - Weighting for nonresponse is generally useful.
- Not all missing data are equivalent.
  - Not covering units in probability samples may result in greater biases than nonresponse.
  - Missingness is not well-defined in non-probability samples

Thank you

J. Michael Brick

mikebrick@westat.com