MICHIGAN

JPSM
The Joint Program in Survey Methodology
University of Maryland
University of Michigan
Westat

Nonresponse Rates and Nonresponse Bias In Household Surveys

Robert M. Groves

University of Michigan and Joint Program in Survey Methodology

Funding from the Methodology, Measurement, and Statistics Program of the US National Science Foundation, Grant 0297435

Four Mutually-Problematic Observations

1. With 100% response rates probability sampling offers an inferential paradigm with measurable uncertainties for unbiased estimates
2. Response rates are declining
3. Keeter *et al.* (2000), Curtin *et al.* (2000), Merkle and Edelman (2002) show no nonresponse bias associated with varying nonresponse rates
4. Practitioners are urged to achieve high response rates

Result: Confusion among practitioners

Nonresponse Error for Sample Mean

In simplest terms $$\bar{Y}_r = \bar{Y}_n + \left(\frac{m}{n}\right)\left[\bar{Y}_r - \bar{Y}_m\right]$$

OR

Respondent Mean = Full Sample Mean + (Nonresponse Rate)*(Respondent Mean – Nonrespondent Mean)

OR

Survey Results = Desired Results + Error

OR

Nonresponse Error = f(Rate, Difference between Respondents and Nonrespondents)

A Stochastic View of Response Propensities

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

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

$\bar{p}$ = **mean propensity over the sample**

ρ_{yp} = **correlation between** y **and** p

σ_y = **standard deviation of** y

σ_p = **standard deviation of** p

Assembly of Prior Studies of Nonresponse Bias

- Search of peer-reviewed and other publications
- 47 articles reporting 59 studies
- About 959 separate estimates (566 percentages)
  - mean nonresponse rate is 36%
  - mean bias is 8% of the full sample estimate
- We treat this as 959 observations, weighted by sample sizes, multiply-imputed for item missing data, standard errors reflecting clustering into 59 studies and imputation variance

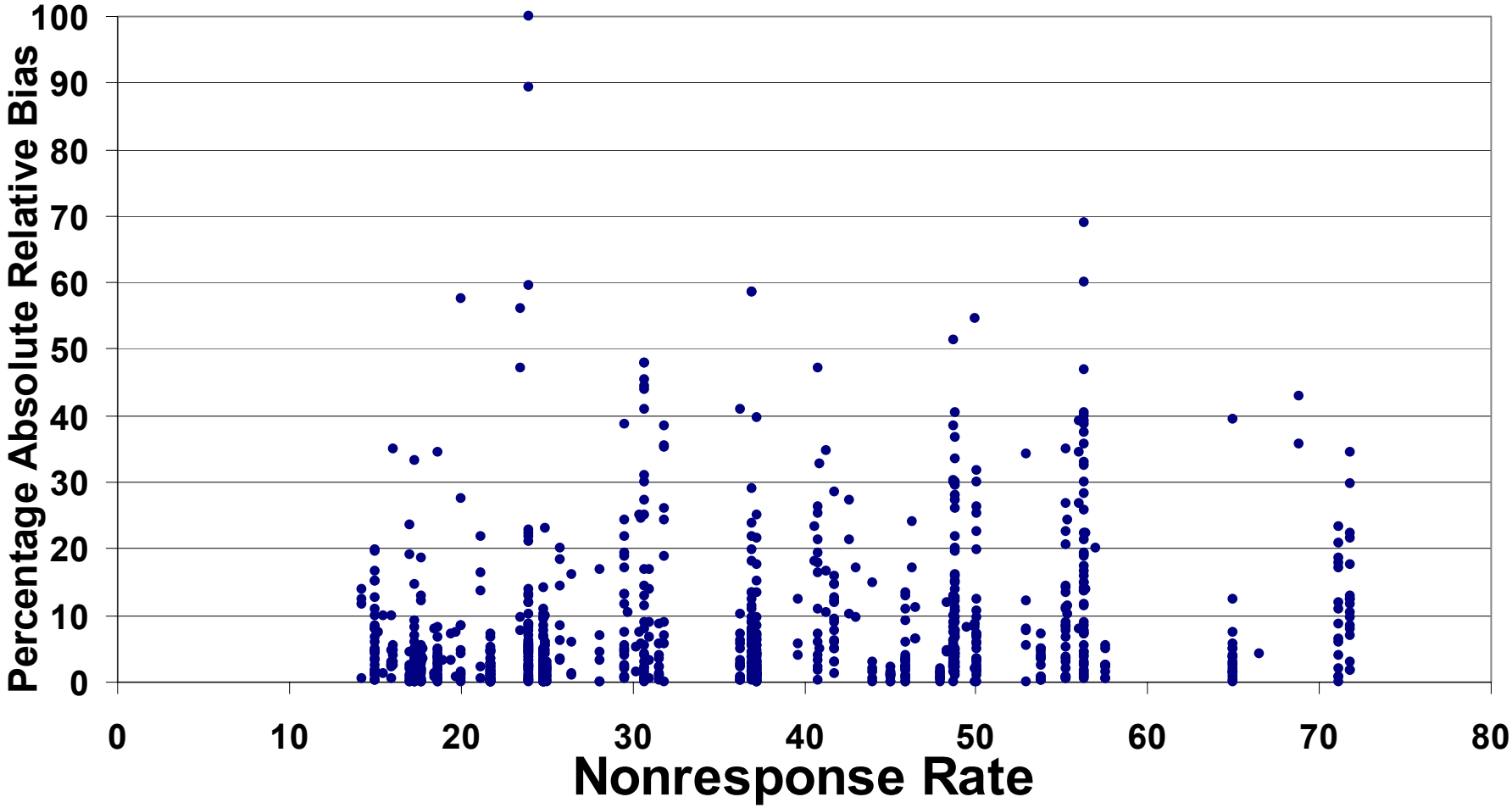
Percentage Absolute Relative Bias

$$\left| \frac{100 * (\bar{y}_r - \bar{y}_n)}{\bar{y}_n} \right|$$

where $\bar{y}_r$ is the unadjusted respondent mean

$\bar{y}_n$ is the unadjusted full sample mean

Percentage Absolute Relative Nonresponse Bias by Nonresponse Rate for 959 Estimates from 59 Studies



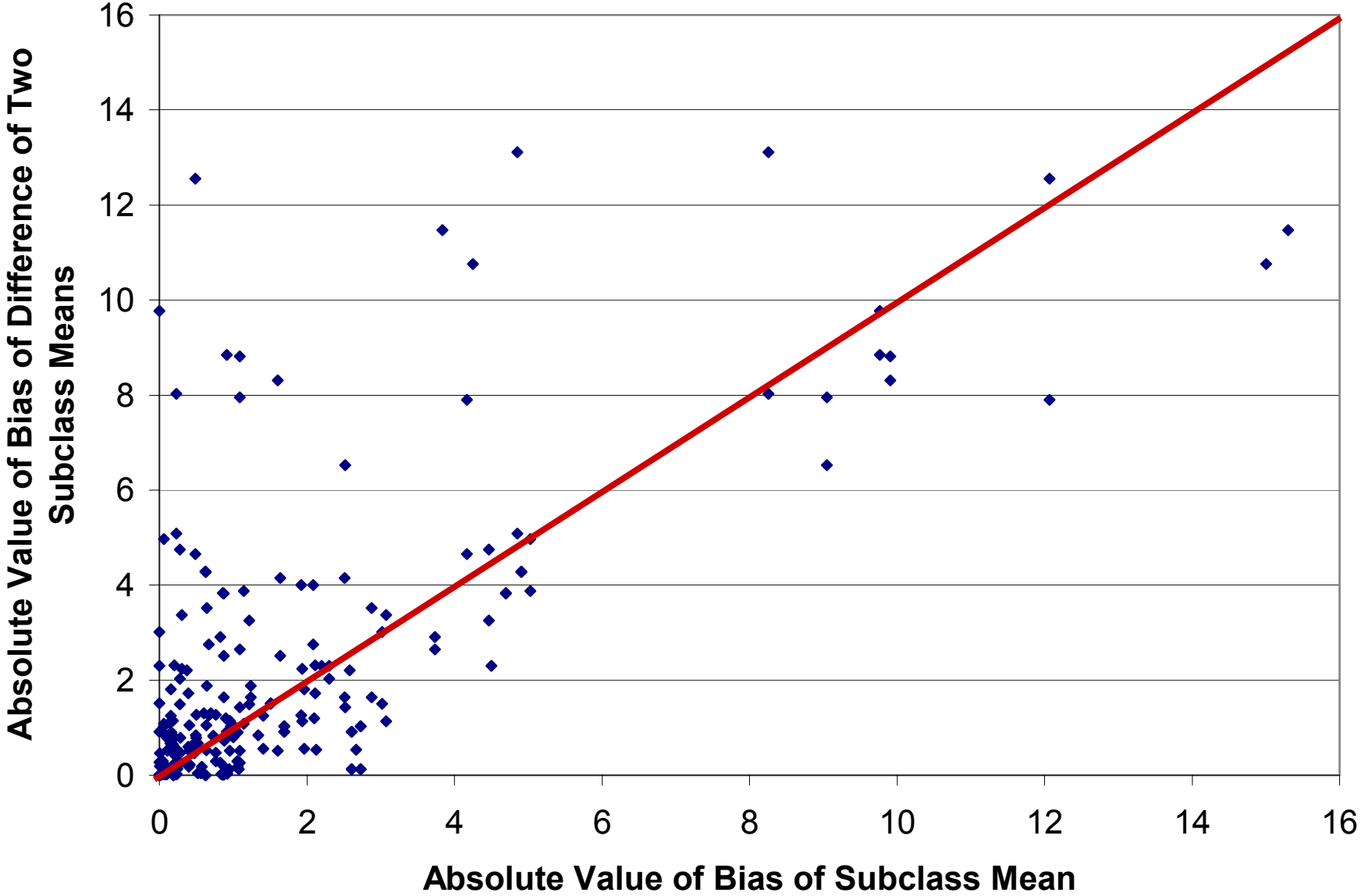
Conclusions from 959 Estimates

- Examples of large nonresponse bias exist
- Variation in nonresponse bias lies mostly among estimates *within* the same survey
- The nonresponse rate by itself is not a good predictor of nonresponse bias
- [Note: We cannot infer from the scatterplot about what would happen *within a study* if response rates were increased]

Do Differences of Subclass Means have Lower Nonresponse Bias?

- When estimating subclass differences, we hope that nonresponse biases of the two estimates cancel
- 120 reported estimates of subclass means and their differences
- Only 45 of them have bias of the differences of subclass means lower than average bias of the two subclass means
  - this comports with only 45 having two subclass means with biases of the same sign

Absolute Value of Bias of Difference of Subclass Mean by Absolute Value of Subclass Mean



Thinking Causally About Nonresponse Rates and Nonresponse Error

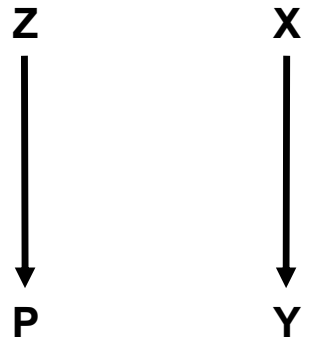
- Key scientific question concerns mechanisms of response propensity that create covariance with survey variable

$$E(\bar{y}_r - \bar{y}_n) = \frac{\sigma_{yp}}{\bar{p}}$$

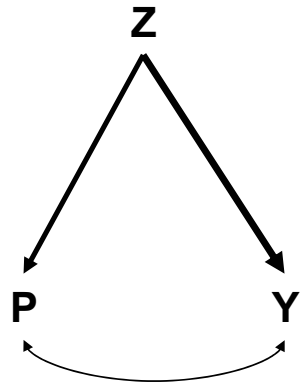
where σ_{yp} is the covariance between the survey variable, *y*, and the response propensity, *p*

- What mechanisms produce the covariance?

Alternative Causal Models for Studies of Nonresponse Rates and Nonresponse Bias

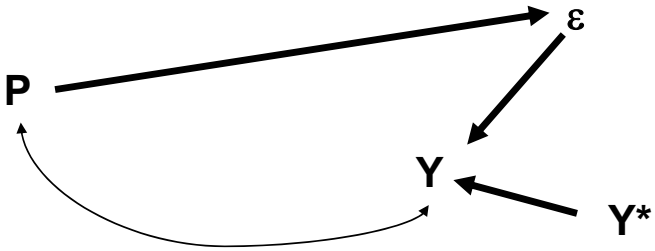


1. Separate Causes Model

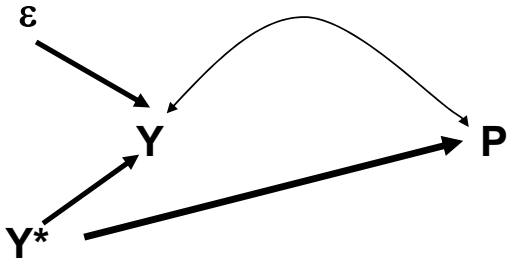


2. Common Cause Model

3. Survey Variable Cause Model



4. Nonresponse-Measurement Error Model



5. Nonresponse Error Attenuation Model

Suspects in Influencing σ_{yp}

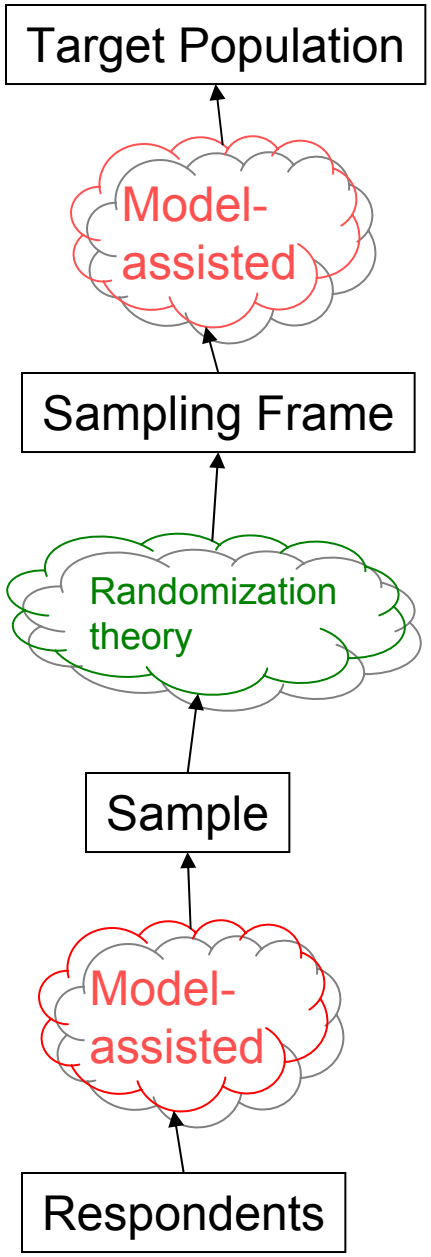
- Advanced letters (correlates of reading letters)
- Incentives (correlates of sensitivity to extrinsic benefit to participation)
- Callback rules (correlates of at-home patterns)
- Observable interviewer attributes (correlates of affect toward stranger attributes)

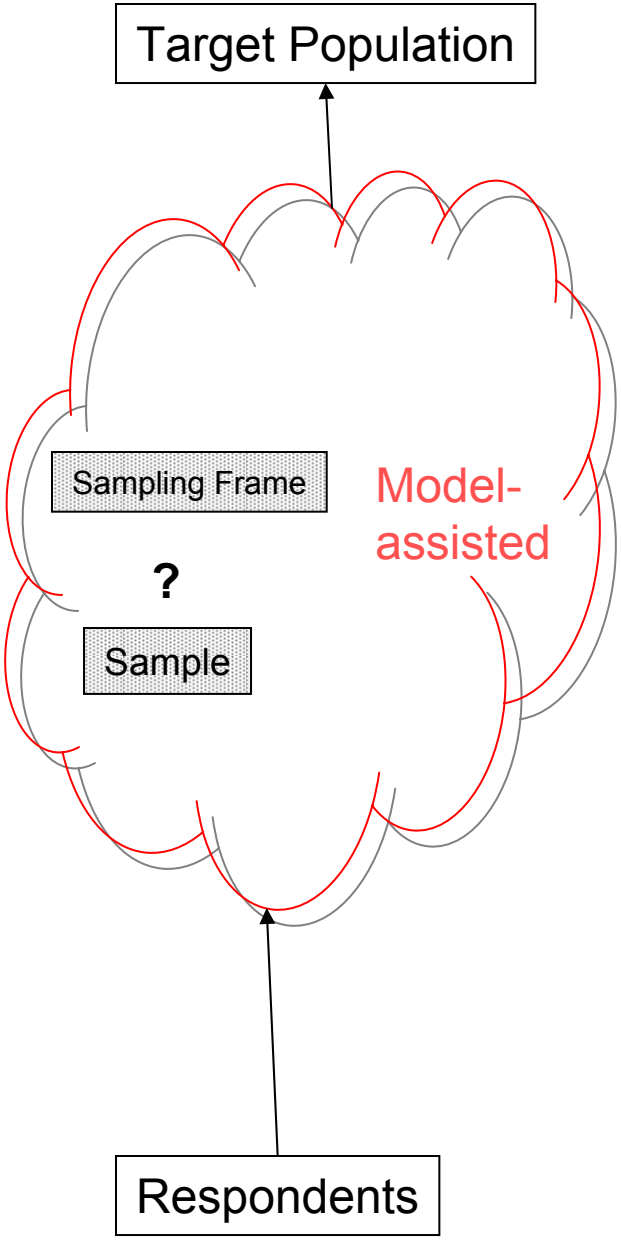
Methods of Assessing Nonresponse Bias

- Comparison of subgroup response rates
- Using rich sampling frame data
- Comparison to external estimates
- Studying variation within the respondent pool in response propensities
- Comparing effects of alternative postsurvey adjustments

Why Persist with Probability Sampling?

- Explicit sampling frames permit coverage assessments
- Randomized selection eliminates chance of biasing eligibility of measurement
- Postsurvey adjustment, when informed by useful auxiliary variables, can repair nonresponse effects


Target Population
Model-
assisted
Sampling Frame
Randomization
theory
Sample
Model-
assisted
Respondents


Target Population
Sampling Frame
Model-
assisted
?
Sample
Respondents

Recommendations to the Practitioner

1. Blind pursuit of high response rates is unwise
2. Probability samples retain the value of randomized selection from known sampling frames
3. Collecting auxiliary variables on respondents and nonrespondents is the key to useful postsurvey adjustment
4. Examining multiple postsurvey adjustments is wise
5. Studying nonresponse bias with multiple methods is wise

“Survey quality” has no practical meaning when every quality feature can vary for each estimate produced by a survey. We should be concerned only with qualities of individual estimates.

Nonresponse error varies over different estimates in the same survey. The sole response rate of a survey cannot reflect that variability.

It was the best of times… low response rate surveys don’t necessarily produce estimates with high nonresponse biases

It was the worst of times… nonresponse error is much more complicated than implied by nonresponse rates