
36-303: Sampling, Surveys and Society

Non-response
Brian Junker
Department of Statistics
132E Baker Hall
brian@stat.cmu.edu

Handouts

- These Notes

Announcements

- Due Today – On Blackboard:
 - II.5: Revised questionnaire/protocol based on try-outs!
 - IRB forms (if you have not already submitted them!)
 - Also Due Today – In Email to me.
 - Peer evaluations
 - One email per person to me, containing forms for all of your team-mates.
 - Forms online at <http://www.stat.cmu.edu/~brian/303>
 - Spring Break!
-

Outline

- Sampling Details Handout
- Non-response and missingness
 - Unit Nonresponse
 - Item Nonresponse
 - Dealing with Nonresponse
 - Types of Missingness
 - Imputation

Sampling Details Handout

Non-response and Missingness

- Non-response is the what
 - Unit non-response: no information at all from that unit (that respondent)
 - Item non-response: (no response on one or more items from that unit; you do get info on other items from that unit)
- Missingness is the why
 - Ignorable missingness: basically got a smaller sample than you expected, but no bias
 - Nonignorable missingness: smaller sample **and** **bias**

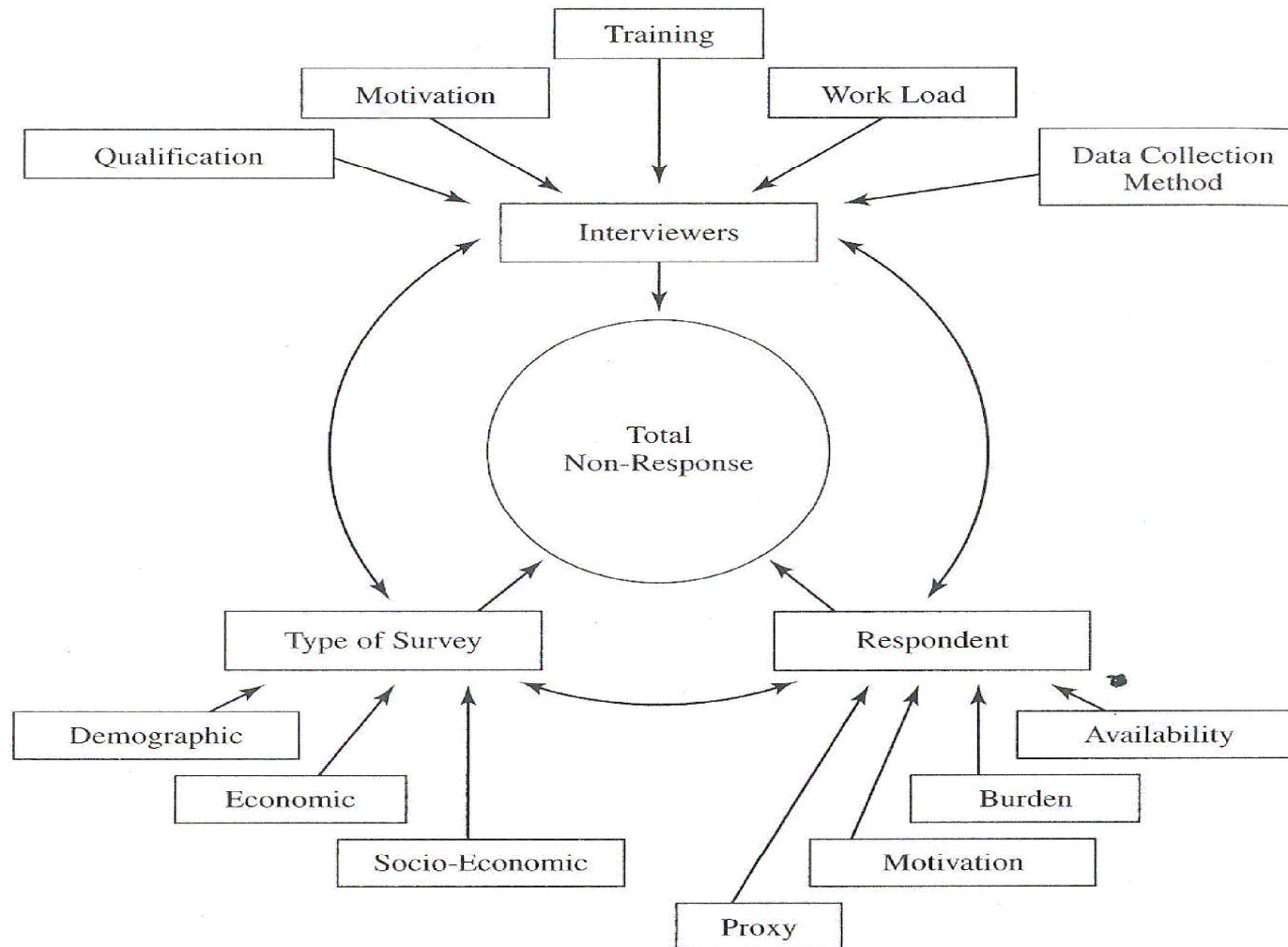
Unit Non-response

- Why might we get a high percentage of unit non-response?
- What's the problem with unit non-response?

Some Types of Unit Non-response

- Non-contact
- Refusal
- Unable to participate
 - How can you remedy these?
 - How can you learn about why they happen?
 - How can you learn how they affect your results?

Factors Affecting Unit Non-response



SOURCE: "Some Factors Affecting Non-Response," by R. Platek, 1977, *Survey Methodology*, 3, 191-214. Copyright © 1977 Survey Methodology. Reprinted with permission.

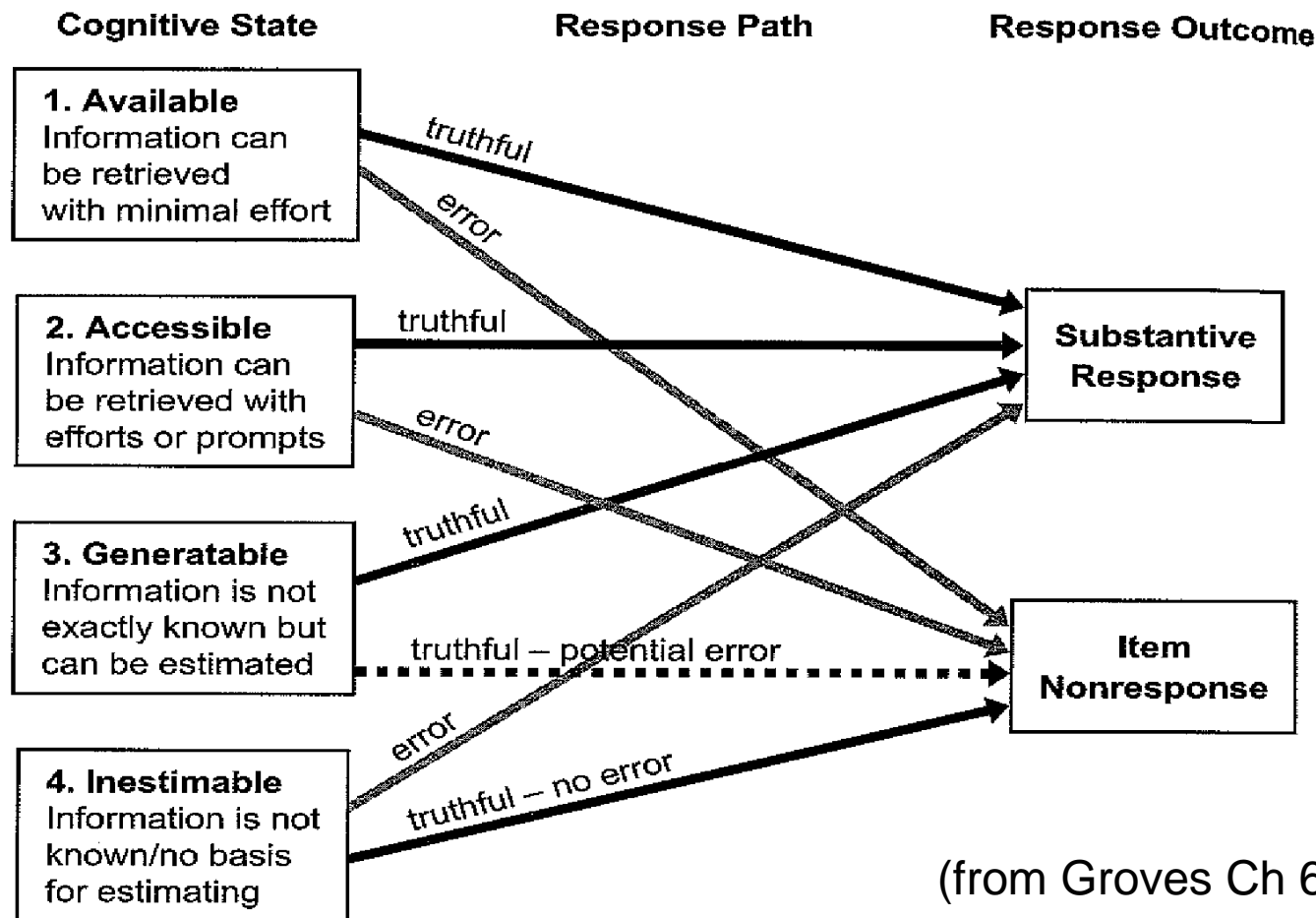
Item Non-response

- Why might we get a high percentage of item non-response?
- What's the problem with item nonresponse?

Some Types of Item Non-response

- Refusal
 - Blank
 - Unintelligible
 - Off-topic
-
- How can you remedy these?
 - How can you learn about why they happen?
 - How can you learn how they affect your results?

Factors Affecting Item Non-response



Best Approach: Prevent Non-response Through Good Design

- Survey Content
- Time of Survey
- Interviewers
- Data Collection Method
- Questionnaire Design
- Burden on Respondent
- Survey Introduction
- Incentives/Disincentives
- Followup

Dealing with Non-response: Ignore It?

- What assumption does this make about our sample and our population?
- What effect does it have on our estimates?
 - bias?
 - variability?
- What if we just asked more people?

Bias Due to Non-response

$$N = N_R + N_M = N_{responding} + N_{missing}$$

$$\bar{Y}_{pop} = \frac{N_R}{N} \bar{Y}_R + \frac{N_M}{N} \bar{Y}_M$$

$$\begin{aligned}(bias) &= \bar{Y}_R - \bar{Y}_{pop} = \bar{Y}_R - \frac{N_R}{N} \bar{Y}_R - \frac{N_M}{N} \bar{Y}_M \\ &= \frac{N_M}{N} (\bar{Y}_R - \bar{Y}_M)\end{aligned}$$

- What makes this bias big?
- What makes it small?

Types of Missingness

- Can we survey the non-respondents to find out why?
 - How? They didn't respond!?
 - Multi-phase sampling & call-backs
- Types of Missingness
 - Missing Completely at Random (MCAR)
 - Just got a smaller sample – only SE's affected
 - Missing at Random (MAR)
 - Different strata may have different missingness but within strata it is MCAR
 - Non-ignorable missingness (MNAR)
 - “Who is missing” is related to what they would have said!

Imputation (predicting / filling-in missing values)

- Some possible methods:
 - Fill in the mean or median of everyone else
 - Find another unit like this one in every way except for the missing value, and copy that value over
 - From the “complete units” build a regression model to predict the missing value
- These methods all depend on MAR
- How does imputation affect survey estimates?

Summary

- Due Today - Blackboard
 - ❑ II.5: Revised questionnaire/protocol based on try-outs!
 - ❑ IRB forms
 - Due Today in Email to me:
 - ❑ Peer evaluation forms
 - ❑ One email from each student, containing forms for all your teammates.
 - Non-response and Missingness
 - Spring Break!
-