A Total Survey Error Analysis of an Address-Based Sampling (ABS) Survey

Ting Yan

Survey Research Center, University of Michigan

Rupa Datta

NORC at the University of Chicago

ITSEW 2012

Address-based Sampling (ABS)

- Extensive use of the US Postal Service (USPS) Delivery Sequence File (DSF)
  - Single-stage sampling design
  - Multi-stage sampling design
- Better coverage of US households
- Enabling the use of multiple modes to contact, recruit and interview
  - Phone (Some DSF addresses can be matched to telephone numbers)
  - Field
  - Mail

ABS (2)

- Error properties of estimates from ABS surveys are unknown
  - Especially when mixed-mode design employed
  - Relative contribution of component bias to the total survey bias
    - Nonresponse bias (by mode)
    - Measurement bias (by mode)
    - Total bias (by mode)

Research Questions

- In an ABS survey with mixed-mode design
  - What is the total bias by mode?
  - What is the relative contribution to total bias by
    - Nonresponse bias?
    - Measurement bias?
  - How does total bias (and component biases) move when sample progresses through the mixed-mode sequence?

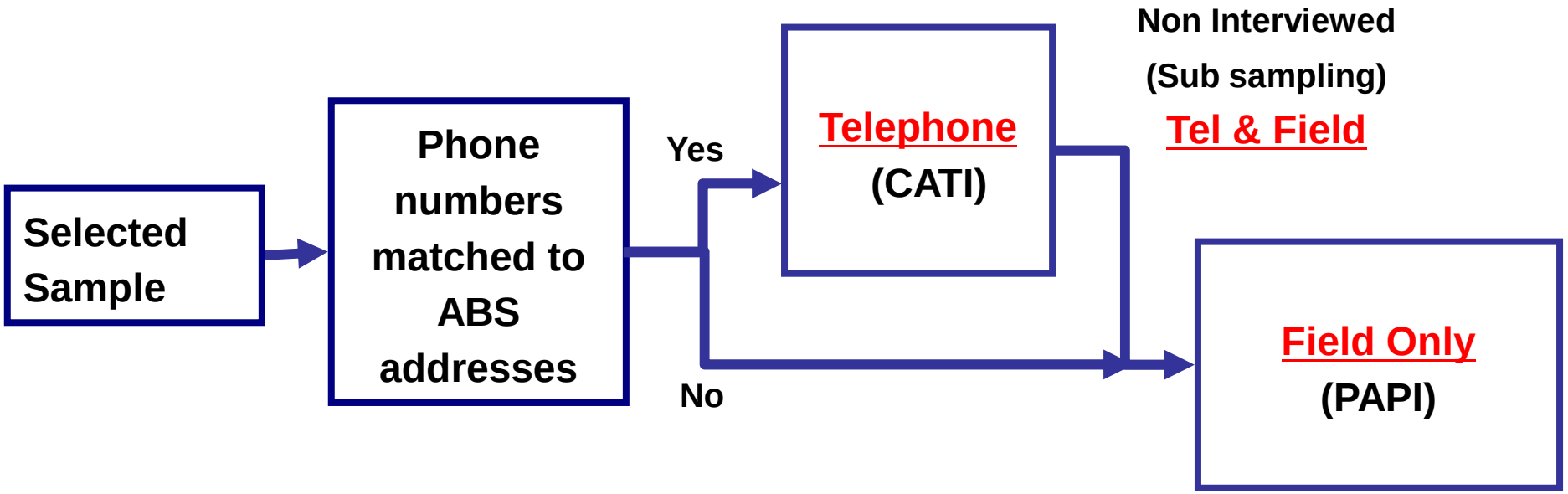
Data

- 2010 Census Integrated Communications Program Evaluation (2010 CICPE)
  - Conducted by NORC at the University of Chicago
  - Sponsored by the Census Bureau
  - to assess the extent to which the 2010 Census Integrated Communications Campaign achieved a variety of specific goals related to:
    - increased mail returns
    - improved accuracy through reduced differential undercount
    - improved cooperation with enumerators

2010 CICPE

- Three waves of interviewing
  - Wave 1: early partner activity/before paid media
  - Wave 2: peak activities
  - Wave 3: mid-April to mid-July, 2010 (Non-Response Follow-Up)
- Waves 2 and 3 consisting of
  - Cross-sectional sample
  - Panel sample
- Multi-mode data collection
  - CATI, PAPI, Paper SAQ (panel only, w2 and w3), Web (panel only, w3)
- Wave 3 fresh cases used for this analysis

2010 CICPE (2): Cross-sectional Samples (ABS+Mixed Modes)



Telephone Only: tel #'s matched and completed on the phone

Tel & Field: tel #'s matched but not completed on the phone (Harder cases)

Field Only: tel #'s not matched, sent straight to the field

2010 CICPE (3)

- Census form returned before April 18?
  - Self-report: Wave 3 asked R to report whether or not his/her household returned the Census Form
  - True value: Administrative data provided by the Census Bureau
  - So we can compare self-report against true value
  - Reasons for mismatch:
    - HH returned late
    - R not aware that another adult in HH returned
    - Deliberate misreporting
      - False positive

Decomposing Total Bias

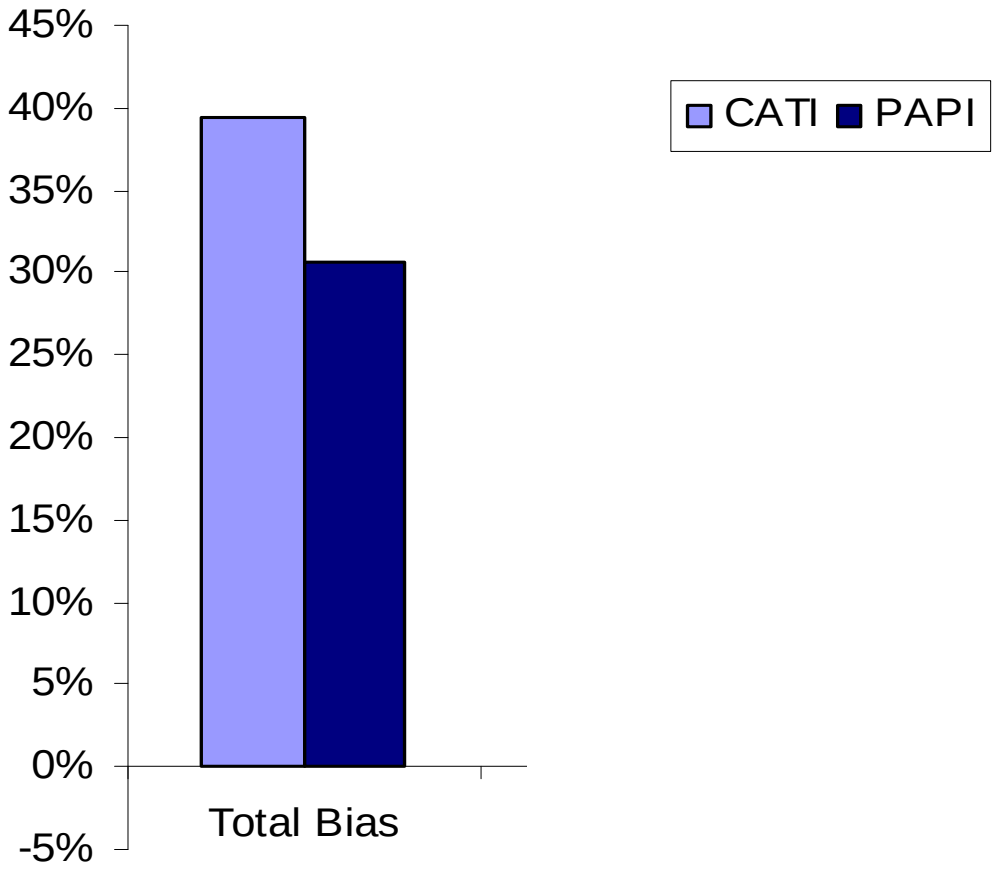
$$(\overline{y_R} - \overline{Y}) = (\overline{Y_{unitR}} - \overline{Y}) + (\overline{Y_R} - \overline{Y_{unitR}}) + (\overline{y_R} - \overline{Y_R})$$

Total Bias=Unit Nonresponse Bias +

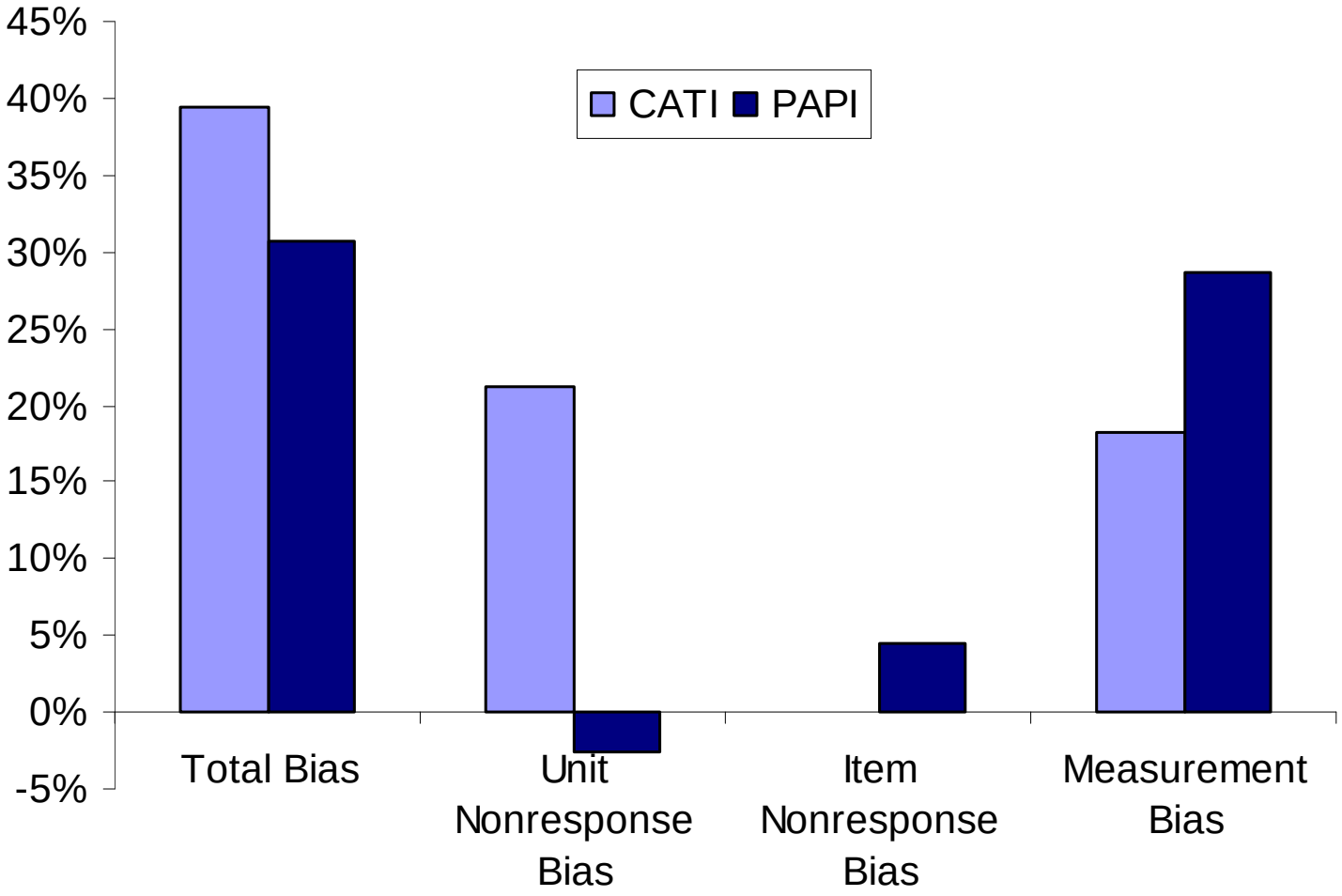
Item Nonresponse Bias +

Measurement Bias

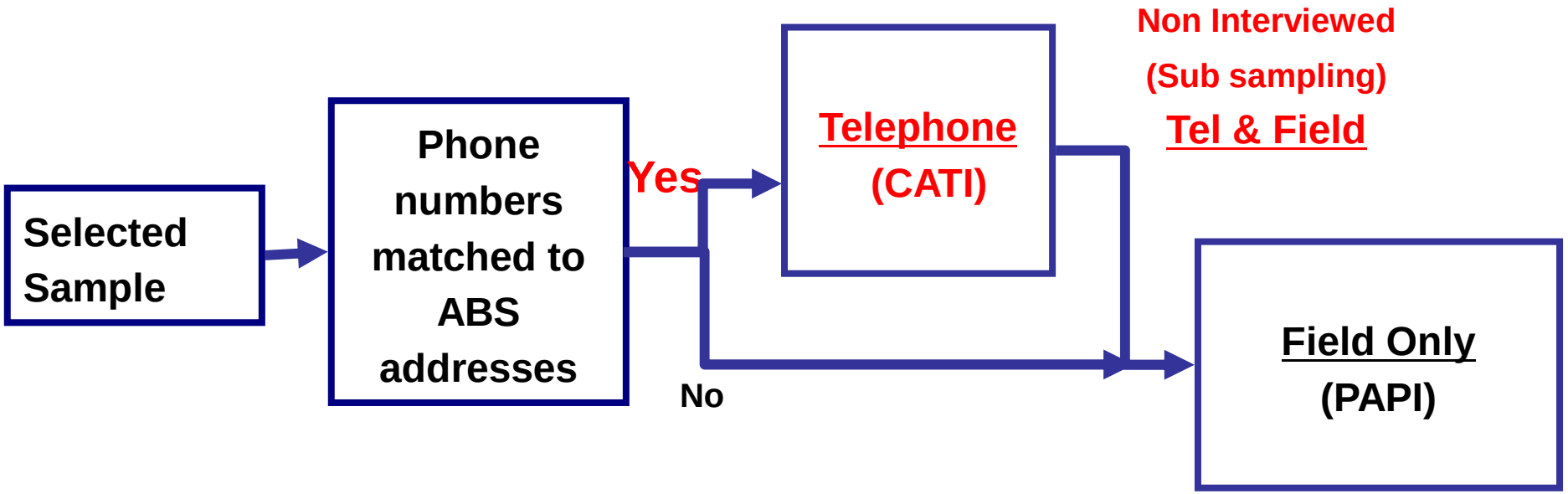
Total Bias by Mode



Component Biases by Mode



2010 CICPE
(ABS+Mixed Modes)

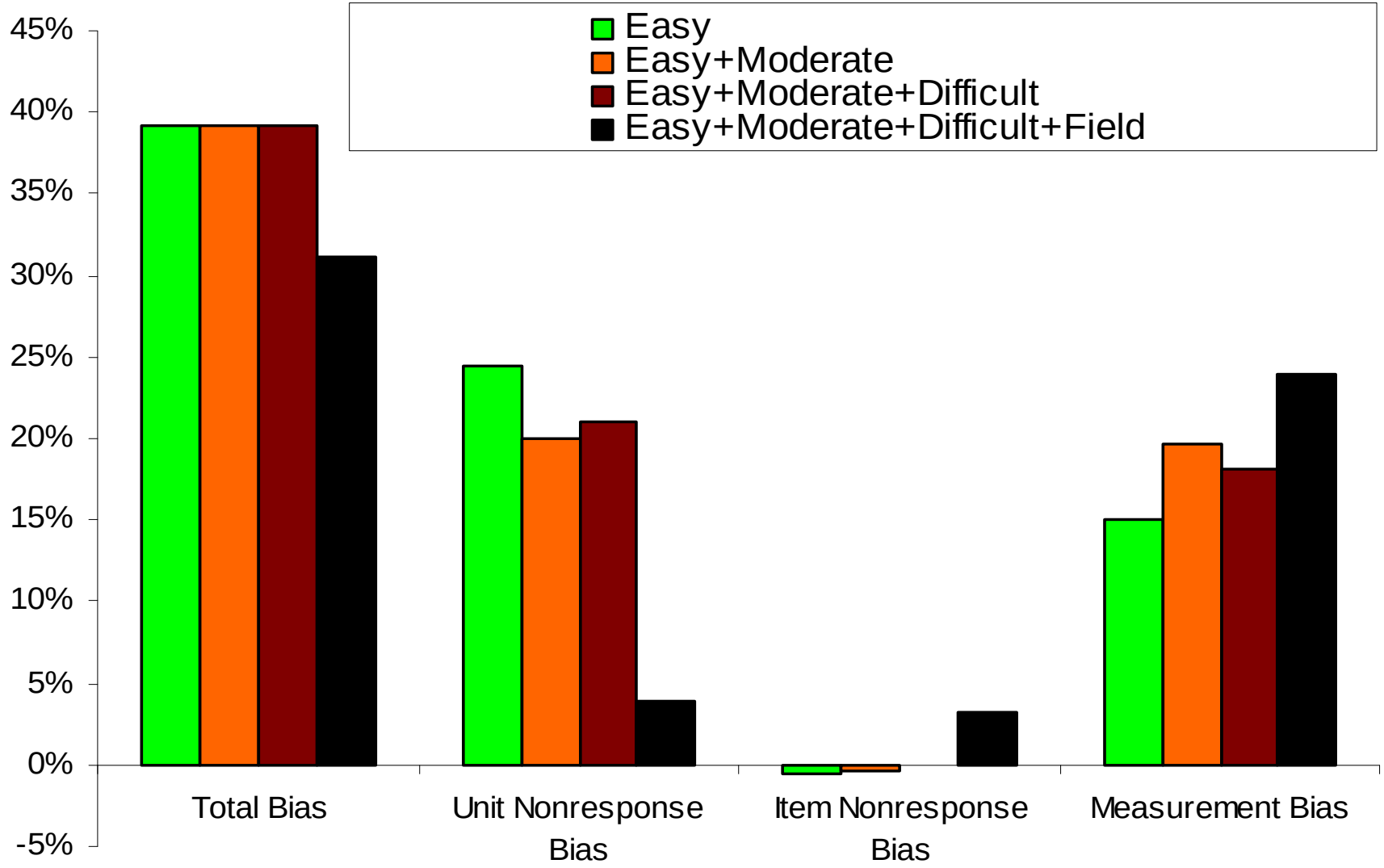


Telephone Only: tel #'s matched and completed on the phone

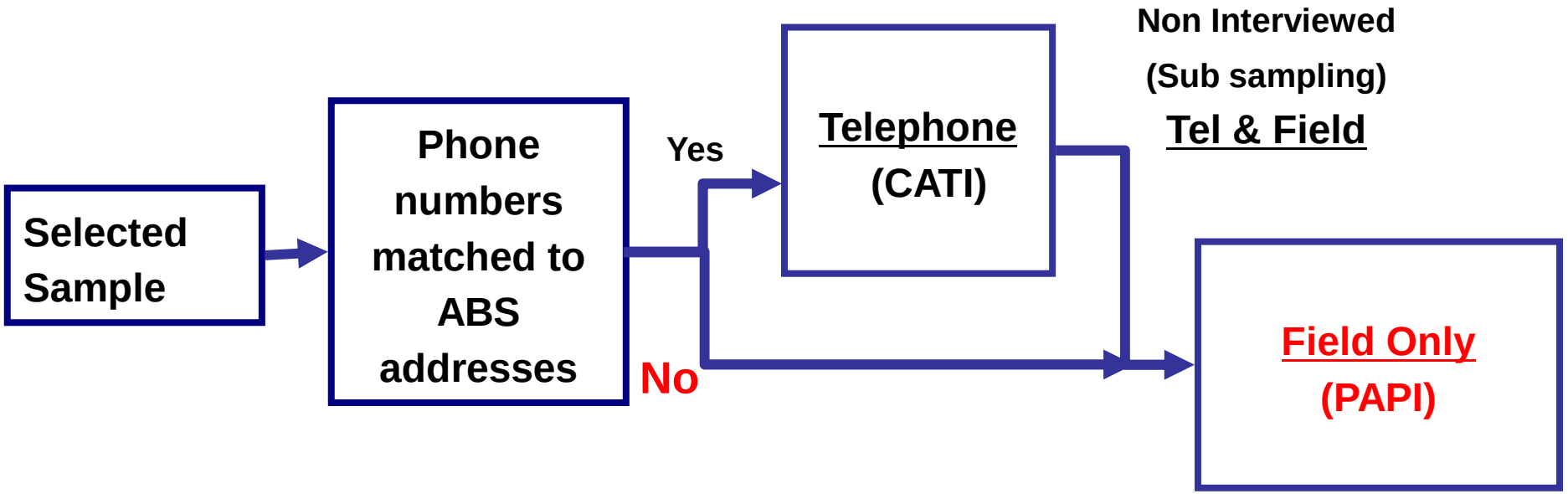
Tel & Field: tel #'s matched but not completed on the phone (Harder cases)

Field Only: tel #'s not matched, sent straight to the field

Biases by Sample Progress – Phone Matches



2010 CICPE
(ABS+Mixed Modes)

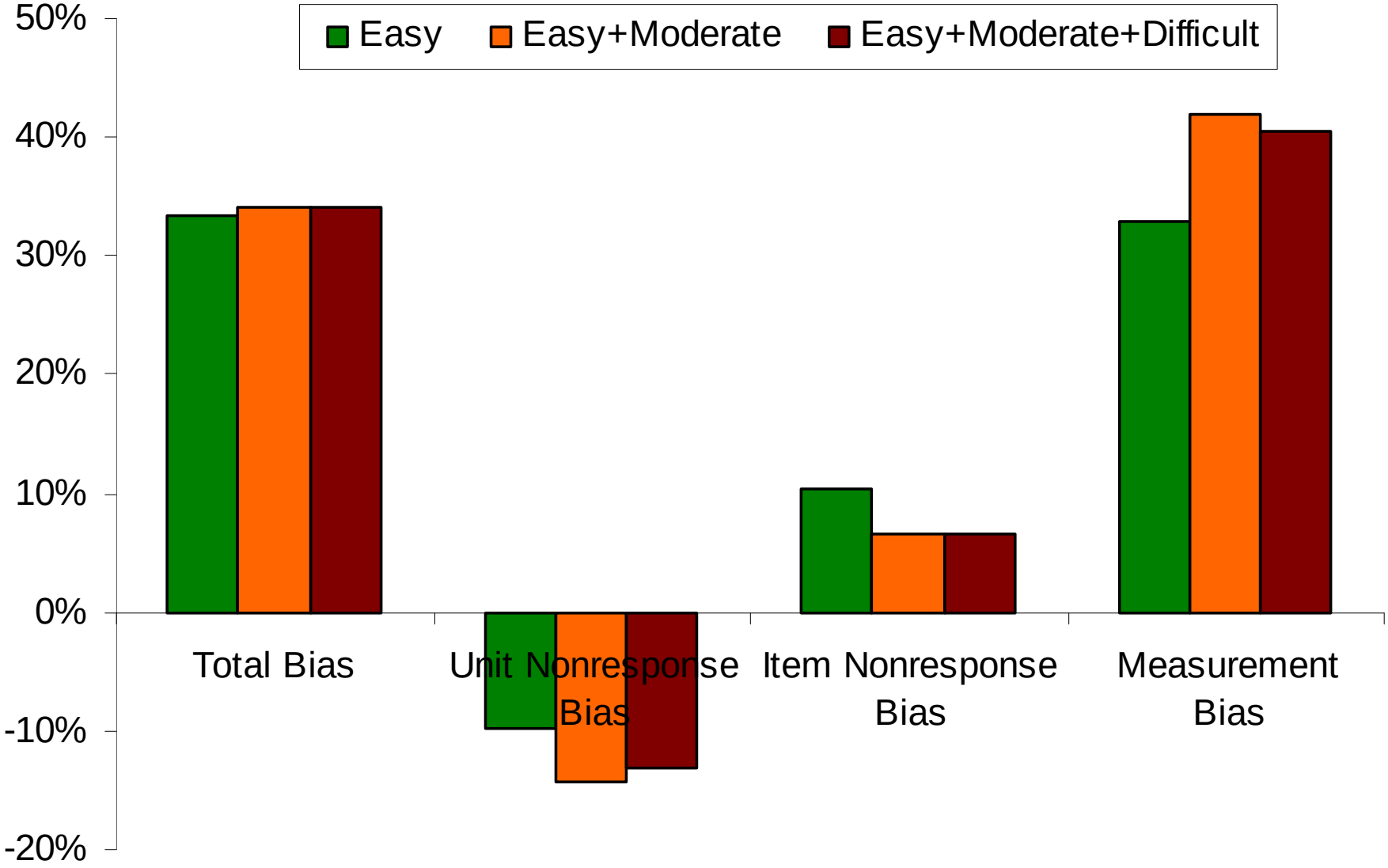


Telephone Only: tel #'s matched and completed on the phone

Tel & Field: tel #'s matched but not completed on the phone (Harder cases)

Field Only: tel #'s not matched, sent straight to the field

Biases by Sample Progress – Non-Phone Matches



Conclusions

- For sampled DSF addresses matched to a phone number, increasing recruiting effort (before subsampling)
  - Unit nonresponse bias ↓
  - Item nonresponse bias ↓
  - Measurement error ↑
  - Total bias —

Conclusions (2)

- For sampled DSF addresses matched to a phone number, subsampling and moving to field
  - Unit nonresponse bias ↓
  - Item nonresponse bias ↑
  - Measurement error ↑
  - Total bias ↓

Conclusions (3)

- For sample not matched with a telephone number, increasing recruiting effort
  - Unit nonresponse bias
  - Item nonresponse bias
  - Measurement bias
  - Total bias

Conclusions (4)

- CATI mode:
  - Item nonresponse bias close to 0
  - Measurement and Unit nonresponse bias about the same magnitude
- PAPI mode:
  - Item nonresponse bias much higher than in the CATI mode
  - Measurement bias about 3 to 4 times as big as nonresponse bias

Conclusions (5)

- For variable “Census form returned”
  - Positive measurement bias suggesting social desirability bias
  - Unit nonresponse bias
    - Positive for sample matched with a telephone number
    - Negative for sample NOT matched with a telephone number

Discussion

- ABS+Mixed-Mode
  - Different R responded to different modes
    - Phone matches vs. non-phone-matches
  - Different sizes of measurement bias associated with different modes
    - Bigger SDB in PAPI than in CATI
  - Total bias comparable though
- What is the deal about differential item nonresponse bias by mode?

Thank you!

tingyan@umich.edu