Measurement Error in Proxy Measures of Key Survey Variables to Inform Responsive and Adaptive Survey Designs

Andy Peytchev, RTI International

Emilia Peytcheva, RTI International

Doug Currivan, RTI International

Matt Jans, University of California and Abt Associates

Advances in Adaptive Survey Design Workshop

Ann Arbor, November 2017

Outline

- Need for auxiliary information to address survey nonresponse
- Desirable properties of auxiliary data
  - Associations
  - Measurement properties
- Objectives of Adaptive and Responsive Designs
- Proposed approach for collection of such data in random digit dialed (RDD) surveys
- Empirical test and evaluation

Need for Auxiliary Information to Address Nonresponse

- Response rates are declining, increasing the risk of nonresponse bias in survey estimates
- Increasing need to:
  - Measure nonresponse bias
  - Reduce nonresponse bias and the risk of nonresponse bias
  - Adjust for nonresponse
- Information on the full sample is essential
  - Deficient in common general population sampling frames, such as RDD and ABS

Desirable Properties of Auxiliary Information

- Available for respondents and nonrespondents
- Associated with:
  - Response outcome
  - Main survey variables (ideally, proxy “Y”s)
- Measured at the sample household and sample person level

Objective of Responsive Design

- Objective—measure and track nonresponse bias to intervene as needed

- Additional objectives—adjust for nonresponse

Responsive Design (Groves and Heeringa, 2006)

(a) Preidentify a set of design features potentially affecting costs and errors of survey estimates

(b) Identify a set of indicators of the cost and error properties of those features and monitor those indicators in initial phases of data collection

(c) Alter the features of the survey in subsequent phases based on cost–error trade-off decision rules

(d) Combine data from the separate design phases into a single estimator.

Adaptive Survey Design

- Within a phase of data collection, target a select subset of sample cases
- Need indicators at the sample member level that are predictive of nonresponse, or other sources of error, or cost

California Health Interview Survey (CHIS)

- State-wide telephone survey producing county and state level estimates
- Landline and cell phone numbers
- Three target populations, selecting up to one of each per household:
  - Adults
  - Adolescents (12-17 years of age)
  - Children (0-11 years of age by proxy)
- Two-stage sampling design
  - Screening questions and respondent selection
  - Main interview with selected person
- Two-phase data collection with subsampling for nonresponse
  - Phase 1: Main phase of data collection (approximately 12 weeks)
  - Phase 2: Nonresponse follow-up of a stratified random sample of Phase 1 nonrespondents, with substantially increased incentives

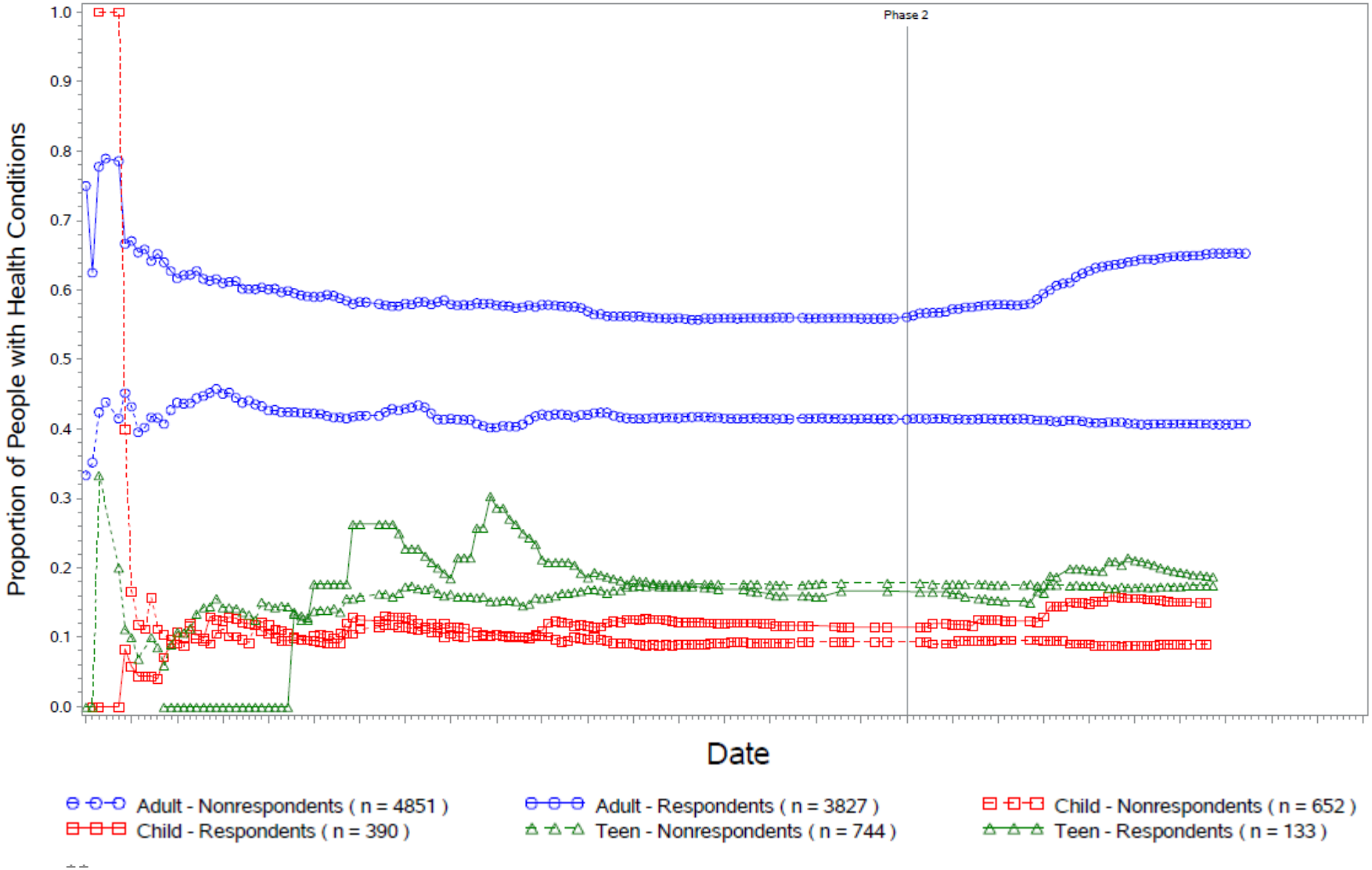
Responsive Design Plan

- Identify key survey variables that are associated with other key variables using prior data
- Add these survey questions to the screening instrument
- Track these variables for respondents and nonrespondents, by sample type
- Target nonrespondents based on these variables to reduce nonresponse bias
- Evaluate whether the design is effective at reducing estimated nonresponse bias
- **Major assumption: measurement error in these proxy indicators is not correlated with nonresponse**

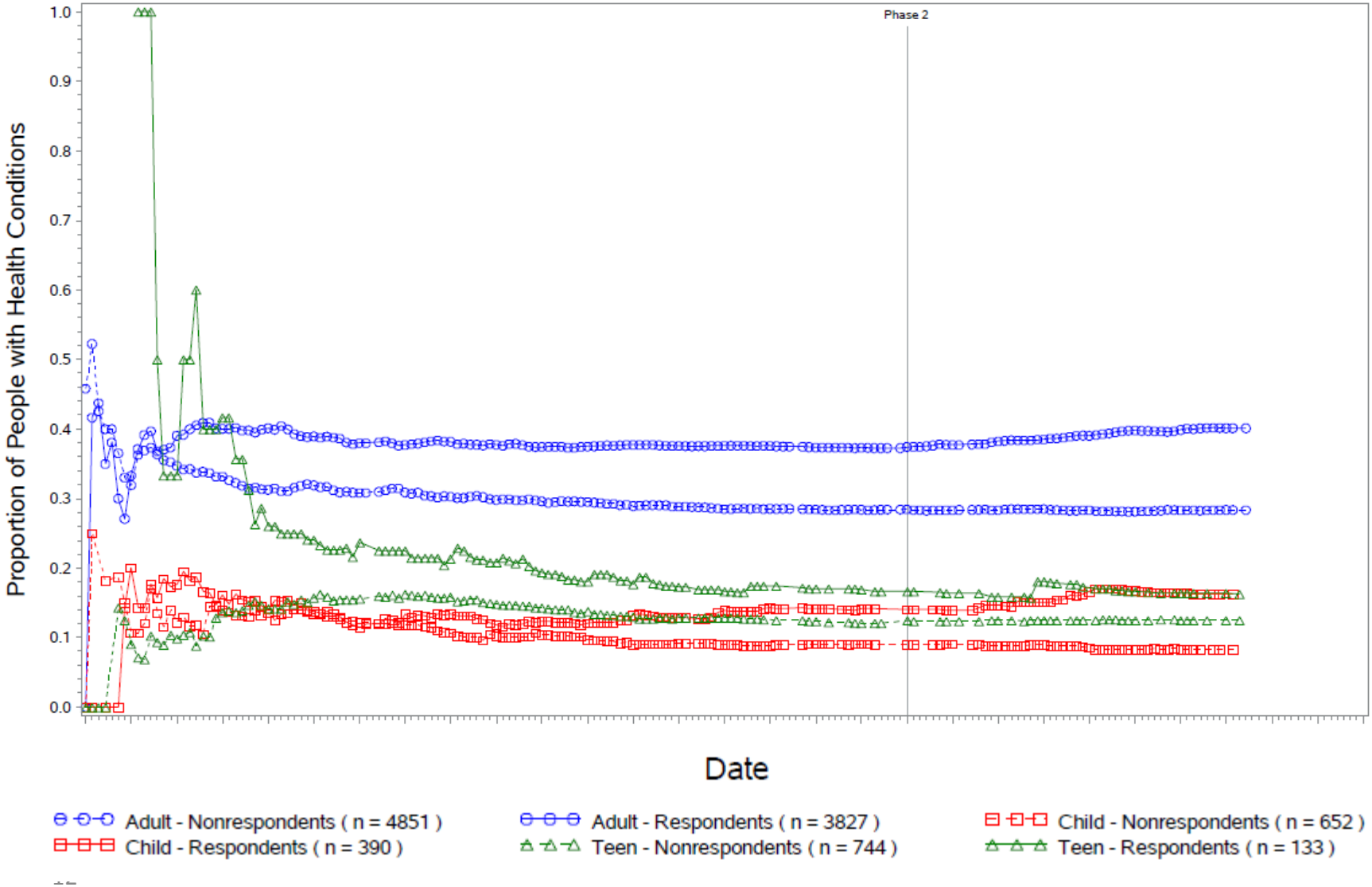
Proxy Measures in CHIS

- MediCal coverage (household-level)
- Any health conditions (person-level)
  - Main interview questions:
    - Has a doctor ever told you that you have asthma?
    - Has a doctor ever told you that you have diabetes or sugar diabetes?
    - Has a doctor ever told you that you have high blood pressure?
    - Has a doctor ever told you that you have any kind of heart disease?
  - Question asked in the screener (“Proxy Y”):
    - To the best of your knowledge, does [SELECTED ADULT] have any health conditions?

Graphs for Health Conditions—Landline



Graphs for Health Conditions—Cell Phone

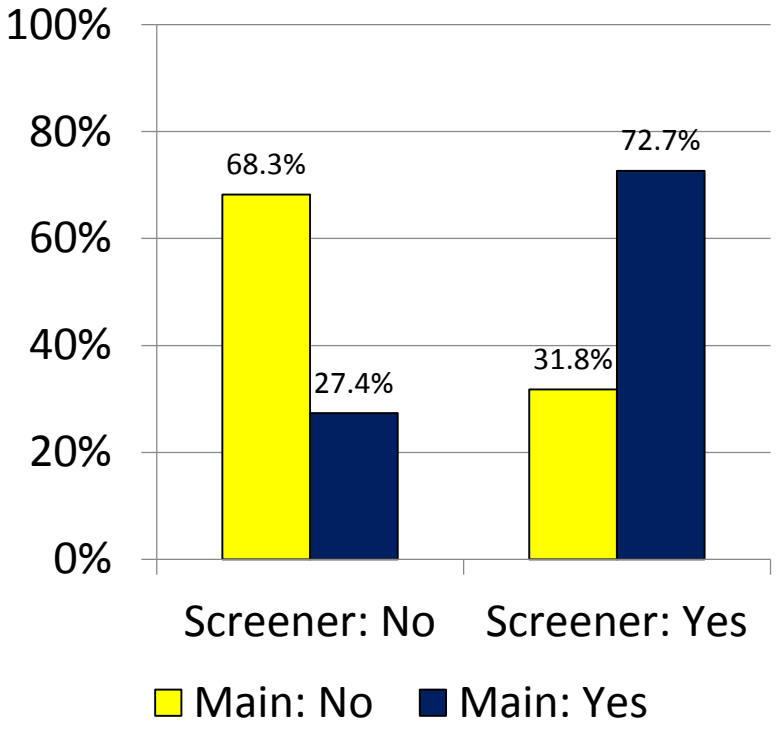


Possible Association between Measurement Error in the Proxy Measures, and Nonresponse

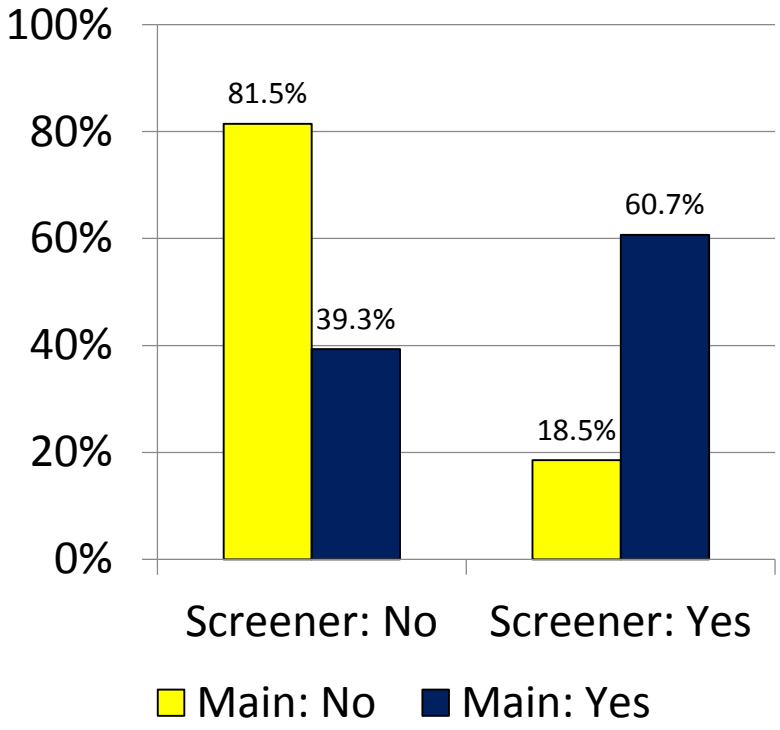
- More reluctant sample members may provide inaccurate responses in the screening instrument
- Proxy responses are more likely to disagree with the selected persons' responses (i.e., the screener respondent is not the selected adult respondent), and the proportion of proxy respondents can be increasing over the course of data collection

1. Is There Measurement Error in the Auxiliary Data (Health Conditions)?

Landline



Cell Phone



2. Is Measurement Error in the Auxiliary Data Associated with Nonresponse?

	Landline		Cell Phone	
	Odds Ratio	**Lower CL**	**Odds Ratio**	**Lower CL**
Call Attempts	*			
3-5 calls (vs 1-2)	1.0	(0.84, 1.25)	1.0	(0.84, 1.30)
6-10 calls (vs 1-2)	1.0	(0.82, 1.25)	1.1	(0.84, 1.33)
11-15 calls (vs 1-2)	1.4	(1.10, 1.77)	1.0	(0.81, 1.35)
16-20 calls (vs 1-2)	1.1	(0.73, 1.57)	1.6	(1.07, 2.47)
21+ calls (vs 1-2)	1.2	(0.76, 1.82)	1.1	(0.67, 1.94)
Phase 2, NRFU	0.9	(0.67, 1.19)	0.9	(0.65, 1.30)
Ever refused	1.1*	(0.97, 1.36)	0.9	(0.75, 1.07)
Proxy	1.2*	(0.98, 1.39)	N/A	N/A

Note: Wald χ^2, * denotes .10, ** denotes .05 level of significance.

3. Is *Underreporting* in the Auxiliary Data Associated with Nonresponse?

	Landline		Cell Phone	
	Odds Ratio	**Lower CL**	**Odds Ratio**	**Lower CL**
Call attempts	*			
3-5 calls (vs 1-2)	1.0	(0.77, 1.29)	1.0	(0.76, 1.30)
6-10 calls (vs 1-2)	1.2	(0.95, 1.60)	1.2	(0.88, 1.52)
11-15 calls (vs 1-2)	1.5	(1.09, 1.96)	1.3	(0.95, 1.72)
16-20 calls (vs 1-2)	1.2	(0.75, 1.88)	1.8	(1.10, 2.87)
21+ calls (vs 1-2)	1.1	(0.64, 1.88)	1.5	(0.79, 2.66)
Phase 2, NRFU	0.9	(0.67, 1.29)	0.9	(0.62, 1.36)
Ever refused	1.2**	(1.01, 1.50)	1.0	(0.8, 1.21)
Proxy	1.2*	(0.97, 1.47)	N/A	N/A

Note: Wald χ^2, * denotes .10, ** denotes .05 level of significance.

4. Could We Correct for Measurement Error in the Auxiliary Data?

- We cannot explain much of the measurement error (previous models) for informing Adaptive Survey Designs or evaluating the effectiveness of a Responsive Design phase, but it is a step in the right direction:

4. Measurement Error by Phase with Correction

- Zero model

	Landline		Cell Phone	
	Odds Ratio	**Lower CL**	**Odds Ratio**	**Lower CL**
Phase 2, NRFU	1.1	(0.88, 1.30)	1.1	(0.87, 1.34)

- Model with number of call attempts and prior refusal

	Landline		Cell Phone	
	Odds Ratio	**Lower CL**	**Odds Ratio**	**Lower CL**
Phase 2, NRFU	0.9	(0.67, 1.29)	0.9	(0.62, 1.36)

Not a Unique Problem

- Proxy Y-measures collected from other individuals and using different procedures are prone to measurement error, which can be correlated with nonresponse
- Interviewer observations are another example of this problem:

West and Kreuter (2014)

Table 3. Accuracy Rates for the Interviewer Judgments of Current Sexual Activity in the Last Four Quarters of the NSFG (2006–2010).

	Quarter 13	Quarter 14	Quarter 15 (I)	Quarter 16 (I)
Accuracy rate	0.7709	0.7700	0.7809	0.7915
False positive rate	0.5732	0.5989	0.4514	0.5380
False negative rate	0.1328	0.1249	0.1488	0.1124

Note: I = Intervention Quarters; NSFG = National Survey of Family Growth.

Summary, Conclusions, and Implications

- The Bad News
  - There is substantial measurement error in the “designed” auxiliary data (“proxy” measures of “Y”)
  - Measurement error is correlated with nonresponse
  - Can not use these variables for nonresponse adjustments
  - Using paradata such as number of call attempts, ever refused, and proxy report, we can construct a measurement error-adjusted version of the “proxy Y” auxiliary variables, but is unlikely to be sufficient
- The Good News
  - At least for this survey, the nonresponse bias indicated by the “proxy Y” variables is biased upward

Discussion Questions

- Should we start collecting measures that are predictive of measurement error in paradata?
- How can we evaluate to what extent observed differences are due to error in the auxiliary data?
- In some instances measurement error in the paradata may make the adaptive survey design less effective, but in others, it could lead to the wrong design (e.g., not use a multiphase design). How do we gauge acceptable levels for measurement error in auxiliary data:
  - For adaptive and responsive survey designs
  - For weighting adjustments