



# Examination of Interventions during Data Collection to Increase Response and Sample Representativeness:

A field test experiment and simulation

Daniel Pratt, Jeffrey Rosen, David Wilson, Melissa Cominole, Elizabeth Copello, Michael Duprey, and Andrey Peytchev

AAPOR 2016 Annual Conference

Austin, TX May 15, 2016

# Acknowledgments and Source

- Sponsor: U.S. Department of Education's National Center for Education Statistics (NCES)
- Study: High School Longitudinal Study of 2009 Second Follow-up
  - Fourth collection with longitudinal cohort that started as ninth-graders in 2009 – now approximately 21 years old

# Overview

- Field test experiment in 2015 to test intervention effectiveness
- Simulation of responsive design implementation
- Main study in 2016

# Field test experimental design

- Field test data collection period: April – July 2015
- Mixed-mode survey
  - Web-based, self-administration
  - E-mail, mail, telephone prompting
  - CATI administration and nonresponse follow-up
- Field test experiments to evaluate different interventions
- Goal: inform the most effective and cost-efficient treatments to be used in the main study

# Field test experimental design (continued)

- Interventions included in field test experiments:
  1. Timing of \$5 prepaid incentive (early or late)
    - Baseline prepaid incentive (sent with data collection announcement letter)
    - Late prepaid incentive (6 weeks into data collection)
  2. Baseline contingent incentive (\$15 offer at baseline or no baseline offer)
  3. Incentive boost offer
    - \$0, \$15, or \$30 added to baseline offer amount (of \$0 or \$15) 8 weeks into data collection
  4. Second boost offer (\$25 more) or abbreviated interview offer 12 days before end of data collection

# Field test experimental design (continued)

- Full factorial randomized assignment design
- Sample assigned randomly across four treatment groups
  - Group A: No baseline incentive offer; late \$5 prepaid
  - Group B: \$15 baseline incentive offer; late \$5 prepaid
  - Group C: No baseline incentive offer; baseline \$5 prepaid
  - Group D: \$15 baseline incentive offer; baseline \$5 prepaid
- Incentive boost offer (\$0, \$15, or \$30)
  - Random assignment of nonrespondents within group to one of three conditions
- Second boost offer (\$25 more) or abbreviated interview
  - Random assignment of nonrespondents within group and within incentive boost condition to one of two conditions

# Summary of Field Test Experiment Results

1. Baseline contingent incentive offer was significantly effective
2. Timing of prepaid incentive had no effect
3. Incentive boost (compared with no boost) was significantly effective, though no difference between \$15 and \$30 levels
4. Final incentive boost more effective than abbreviated interview

# Purpose of Field Test Responsive Design Simulation

1. Can we better represent population of interest by including cases in respondent pool that otherwise would be nonrespondents (sample representativeness)?
2. Are interventions effective when targeting cases using responsive design methods?



# Responsive Design Modeling and Simulation

- Field test experiments involved random assignment to treatment conditions to ensure adequate sample size
- Developed and implemented responsive design model using field test data
- Substantive variables used in model
- Ran simulations of case identification for each of final 2 interventions
- Main study will leverage responsive design methods to target interventions

# Bias Likelihood Model

- Logistic regression to identify particular nonresponding cases for targeted interventions which may improve sample representativeness and reduce potential nonresponse bias in population estimates if they participate
- Dependent variable = Response outcome at the time the model is run during current round of data collection
- Independent variables consist exclusively of substantive survey variables from prior round(s) of data collection
- Calculated before each planned intervention to target cases for special treatment

# Responsive Design Simulation Results

- Bias-likelihood score as of start of incentive-boost-offer experiment; targeted half-sample that would contribute more to nonresponse bias if remained nonrespondents
- Repeated approach as of start of abbreviated-interview-incentive-boost experiment
- Comparison of respondents to overall sample prior to these interventions and at end for model variables
  - Simulation enabled set-up and testing of procedures, including potential variable identification
  - On assorted model variable values, responding sample more closely represents overall sample at end of data collection
- Interventions: Results for simulated targeted cases consistent with overall results albeit with small Ns

# Main Study Plans

- Baseline contingent incentive for targeted cases
- Up to two incentive boosts for targeted cases
- More testing to optimize baseline contingent incentive and boost amounts
- Two models used to identify cases for targeted interventions
  - “Bias likelihood,” which estimates a case’s predicted contribution to bias in key survey variables
  - “Response likelihood,” which estimates a case’s predicted probability of participation
- Additional interventions: field follow-up; extended data collection

# Response Propensity Model

- Strategy to conserve resources
- Logistic regression that estimates the likelihood of participation in advance of data collection
- Uses independent variables that predict survey response, including paradata, frame data, survey data
- Potential uses:
  - Not offering baseline incentive to cases with very high likelihood of participation
  - Not implementing costly intervention (e.g., field follow-up) to cases with very low likelihood of participation

- **Model effectiveness:** Does the bias likelihood model successfully identify nonresponding cases that are underrepresented on key survey variables?
- **Intervention effectiveness:** Do interventions increase participation among targeted cases?
- **Sample representativeness:** Is sample representativeness – and in turn are population estimates – improved through the use of the targeted interventions for identified cases?
- **Cost effectiveness:** Does use of response propensity model conserve resources?

# More Information

**Dan Pratt**

+1-919-541-6615

djp@rti.org