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

A field test experiment and simulation

Presenter: Dan Pratt

4th Workshop of Advances in Adaptive and Survey Design
University of Manchester, November 10, 2015

Acknowledgments and Source

- Collaborators: Melissa Cominole, Liz Copello, Andy Peytchev, Emilia Peytcheva, Jeff Rosen, Dave Wilson
- Sponsor: U.S. Department of Education's National Center for Education Statistics
- 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
  1. Baseline incentive offer
  2. Timing of incentive prepay
  3. Incentive boost offer
  4. Comparison of second incentive boost to abbreviated interview offer
- Simulation of responsive design implementation
- Refine procedures in preparation for main study in 2016

Field test experimental design

- Field test experiments to evaluate different interventions
- Interventions included in field test experiments:
  1. Timing of $5 prepaid incentive (early or late)
     - Early prepaid incentive (sent with data collection announcement letter)
     - Late prepaid incentive (6 weeks into data collection)
  2. Baseline incentive offer ($15 offer at baseline or no baseline offer)
  3. Incentive boost offer amount
     - $0, $15, or $30 added to 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
- Sample assigned randomly across treatment groups

Field test phases and treatments

Phase	Group A	Group B	Group C	Group D
Phase 1 (Apr 13): Web only, $5 prepaid for selected cases	No baseline incentive offer; late $5 prepaid	$15 incentive offer; late $5 prepaid	No baseline incentive offer; early $5 prepaid	$15 incentive offer; early $5 prepaid
Phase 2 (May 4): Telephone interviewing added	Telephone added	Telephone added	Telephone added	Telephone added
Phase 3 (May 26): $5 prepaid for selected cases	$5 prepaid	$5 prepaid	(Prepaid at baseline)	(Prepaid at baseline)
Phase 4 (June 8): Increased incentive for selected cases	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost
Phase 5 (July 6): Increased incentive or abbreviated	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview

Field test experiment results overall and within phase

Phase	Group A	Group B	Group C	Group D
	No baseline incentive offer; late $5 prepaid	$15 incentive offer; late $5 prepaid	No baseline incentive offer; early $5 prepaid	$15 incentive offer; early $5 prepaid
Overall Response Rate	45.5	52.4	47.6	56.4
Within Phase 1 (Apr 13): Web only, $5 prepaid for selected cases	6.9	15.6	8.7	17.8
Within Phase 2 (May 4): Telephone interviewing added	9.4	14.2	12.0	15.0
Within Phase 3 (May 26): $5 prepaid for selected cases	9.5	10.1	7.7	10.9

Field test experiment results (continued)

Phase	Group A	Group B	Group C	Group D
	No baseline incentive offer; late $5 prepaid	$15 incentive offer; late $5 prepaid	No baseline incentive offer; early $5 prepaid	$15 incentive offer; early $5 prepaid
Overall Response Rate	45.5	52.4	47.6	56.4
Phase 4 (June 8): Increased incentive for selected cases	Within phase: 17.6 No boost: 11.4 $15 boost: 23.2 $30 boost: 18.3	Within phase: 16.8 No boost: 11.9 $15 boost: 23.7 $30 boost: 14.8	Within phase: 12.7 No boost: 4.5 $15 boost: 14.7 $30 boost: 18.8	Within phase: 21.5 No boost: 21.4 $15 boost: 27.6 $30 boost: 15.5
Phase 5 (July 6): Increased incentive or abbreviated	Within phase: 13.3 Abbrev: 10.3 $25 boost: 16.3	Within phase: 12.1 Abbrev: 8.0 $25 boost: 16.2	Within phase: 19.1 Abbrev: 17.0 $25 boost: 21.1	Within phase: 11.1 Abbrev: 4.5 $25 boost: 17.6

1. Baseline incentive offer: B and D received $15 offer

Phase	Group A	Group B	Group C	Group D
Phase 1 (Apr 13): Web only, $5 prepaid for selected cases	No baseline incentive offer; late $5 prepaid	$15 incentive offer; late $5 prepaid	No baseline incentive offer; early $5 prepaid	$15 incentive offer; early $5 prepaid
Phase 2 (May 4): Telephone interviewing added	Telephone added	Telephone added	Telephone added	Telephone added
Phase 3 (May 26): $5 prepaid for selected cases	$5 prepaid	$5 prepaid	(Prepaid at baseline)	(Prepaid at baseline)
Phase 4 (June 8): Increased incentive for selected cases	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost
Phase 5 (July 6): Increased incentive or abbreviated	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview

1. Baseline incentive offer: AC vs. BD

- Baseline incentive was significantly effective overall
  - No baseline offer (AC) vs. $15 baseline offer (BD): Chi-square = 6.72, p = 0.009

Experiment group	Final response rate
AC: No baseline offer	46.5
BD: $15 baseline offer	54.4

2. Timing of $5 prepaid incentive: at baseline for C and D

Phase	Group A	Group B	Group C	Group D
Phase 1 (Apr 13): Web only, $5 prepaid for selected cases	No baseline incentive offer; late $5 prepaid	$15 incentive offer; late $5 prepaid	No baseline incentive offer; early $5 prepaid	$15 incentive offer; early $5 prepaid
Phase 2 (May 4): Telephone interviewing added	Telephone added	Telephone added	Telephone added	Telephone added
Phase 3 (May 26): $5 prepaid for selected cases	$5 prepaid	$5 prepaid	(Prepaid at baseline)	(Prepaid at baseline)
Phase 4 (June 8): Increased incentive for selected cases	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost
Phase 5 (July 6): Increased incentive or abbreviated	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview

2. Timing of $5 prepaid incentive: AB vs. CD

Experiment group	Final response rate
AB: Late $5 prepaid	48.9
CD: Early $5 prepaid	52.0

- Prepaid timing had no effect
  - Late prepaid (AB) vs. baseline prepaid (CD): Chi-square = 1.05, p = 0.31

3. Incentive boost offer

Phase	Group A	Group B	Group C	Group D
Phase 1 (Apr 13): Web only, $5 prepaid for selected cases	No baseline incentive offer; late $5 prepaid	$15 incentive offer; late $5 prepaid	No baseline incentive offer; early $5 prepaid	$15 incentive offer; early $5 prepaid
Phase 2 (May 4): Telephone interviewing added	Telephone added	Telephone added	Telephone added	Telephone added
Phase 3 (May 26): $5 prepaid for selected cases	$5 prepaid	$5 prepaid	(Prepaid at baseline)	(Prepaid at baseline)
Phase 4 (June 8): Increased incentive for selected cases	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost
Phase 5 (July 6): Increased incentive or abbreviated	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview

3. Incentive boost: comparison overall and by amount

Group	Within phase response rate
No boost	11.9
Any boost	19.5
$15 boost	22.0
$30 boost	17.0
No boost	11.9
$15 boost	22.0
No boost	11.9
$30 boost	17.0

- Incentive boost

Significant effect of boost to no boost:
Chi-square = 6.90, p = 0.009

No significant difference between $15 and $30 conditions:
Chi-square = 2.09, p = 0.15

Significant effect of $15 boost to no boost:
Chi-square = 9.22, p = 0.002

No significant difference between no boost and $30 boost:
Chi-square = 2.67, p = 0.10

4. Abbreviated interview vs. second incentive boost offer

Phase	Group A	Group B	Group C	Group D
Phase 1 (Apr 13): Web only, $5 prepaid for selected cases	No baseline incentive offer; late $5 prepaid	$15 incentive offer; late $5 prepaid	No baseline incentive offer; early $5 prepaid	$15 incentive offer; early $5 prepaid
Phase 2 (May 4): Telephone interviewing added	Telephone added	Telephone added	Telephone added	Telephone added
Phase 3 (May 26): $5 prepaid for selected cases	$5 prepaid	$5 prepaid	(Prepaid at baseline)	(Prepaid at baseline)
Phase 4 (June 8): Increased incentive for selected cases	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost	$0 or $15 or $30 boost
Phase 5 (July 6): Increased incentive or abbreviated	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview	$25 boost or abbreviated interview

4. Abbreviated interview vs. second incentive boost offer

Group	Within phase response rate
Abbreviated	10.4
$25 boost	17.9

- Abbreviated or $25 boost
  - Significant effect of $25 boost over abbreviated: Chi-square = 7.37, p = 0.007

Summary of Field Test Experiment Results

1. Baseline $15 incentive offer was significantly effective
2. Timing of $5 prepaid incentive had no effect
3. Incentive boost was significantly effective, though no difference between $15 and $30 levels
4. Final incentive boost more effective than abbreviated interview

Purpose of the Field Test Responsive Design Simulation

1. Can we better represent population of interest (fall 2009 ninth-graders as of 2016) by including cases in respondent pool who otherwise would be nonrespondents (sample representativeness)?

2. Are interventions tested with field test random-assignment implementation effective when targeting cases using responsive design methods (intervention effectiveness)?

Responsive Design Modeling and Simulation

- Field test experiments involved random assignment to experiment groups for all experiments to ensure adequate sample size
- Main study will leverage responsive design methods to target interventions
- Developed and implemented responsive design modeling using field test data
- Substantive variables used in the model
- Ran simulations for last 2 interventions (experiment 3: $0/$15/$30 incentive boost offer; experiment 4: abbreviated interview offer or $25 incentive boost offer) to identify cases based on responsive design model (incentive boost level; abbreviated- versus second incentive boost)

Responsive Design Model Variables

- Approximately 3 dozen variables included in modeling
- Variables from 2013 Update responsive design model – prior round of data collection included responsive design implementation (e.g., demographics; timing of algebra 1; grade in algebra 1; highest education expected)
- Additional variables from 2013 Update and high school transcripts (e.g., high school credential status; postsecondary enrollment as of November 2013; employment as of November 2013)

Responsive Design Simulation Results: Case Selection

- Assigned bias-likelihood score for all nonrespondents as of start of phase 4; targeted cases were half-sample who would contribute greatest to nonresponse bias if remained nonrespondents
- Repeated approach as of start of phase 5
- Comparison of overall respondents to overall sample prior to phase 4 and at end for model variables
  - Simulation enabled set-up and testing of procedures, including potential variable identification
  - Interventions applied based on random assignment, not responsive design
  - On assorted model variable values, responding sample more closely represents overall sample at end of data collection

Simulated impact of responsive design interventions on sample representativeness

Student/school indicator	Percent among respondents before phase 4	Percent among respondents before phase 5	Percent among all field test respondents	Percent among overall field test sample
Male	51.8	48.7	49.6	48.1
Asian	5.1	4.6	4.0	4.1
Hispanic	2.8	3.2	3.5	3.7
Suburban base-year school	43.9	44.1	42.9	38.5
Town base-year school	4.0	5.2	5.2	6.4
Rural base-year school	15.4	16.1	17.1	19.7
Northeast base-year school	15.4	17.8	18.3	20.7
Midwest base-year school	22.9	21.8	20.8	19.0

Responsive Design Simulation Results: Interventions

- Phase 4: incentive boost offer ($0, $15, $30)
- Phase 5: abbreviated interview or $25 additional boost offer
- Numbers of cases and respondents very small
- Simulated targeted cases were the lower-half among nonrespondents using model variables at each stage
  - Overall 9.4% of cases selected in phase 4 simulation participated during phase 4
  - Overall 7.1% of cases selected in phase 5 simulation participated during phase 5
- Results among simulated targeted cases in same general direction as overall results albeit with very small Ns

Simulated effectiveness of responsive design interventions: incentive boost

Intervention	Within phase response rate for all cases	Within phase response rate for targeted half-sample of nonrespondents
Total	17.0	9.4
No incentive boost	11.9	6.1
$15 incentive boost	22.0	10.0
$30 incentive boost	17.0	11.8

Simulated effectiveness of responsive design interventions: abbreviated interview versus additional incentive boost

Intervention	Within phase response rate for all cases	Within phase response rate for targeted half-sample of nonrespondents
Total	14.2	7.1
Abbreviated interview	10.4	4.1
$25 incentive boost	17.9	10.2

Main Study Plans

- Calibration samples
  - Small sub-sample to be used for tests of interventions (incentive amounts)
  - Will be worked approximately 5-7 weeks in advance of the main sample
- 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

Bias Likelihood Model

- Dependent variable = Second follow-up response outcome at the time the model is run (at the beginning of each phase)
- Independent variables consist exclusively of substantive survey variables from prior wave(s)
- Calculated before each of 4 phases (after initial web-only and initial CATI phases) to target cases for intervention

Response Likelihood

- Estimates the likelihood of participation a priori
- Uses independent variables that predict survey response, including paradata, frame data, survey data
- Calculated once in advance of data collection
- Used as potential phase-specific filter for interventions
  - To ensure that we don't target cases that have either a very high likelihood of participation (e.g., for incentive boost) or very low likelihood (e.g., for field follow-up)
- Could consider using model results to refine classification of cases
  - For example, cases with very high response propensity could be treated as "ultra-cooperative" cases

Closing

Thank you!