Adaptive Design Strategies for Addressing Nonresponse Error in NCES Longitudinal Surveys

Sarah Crissey
Elise Christopher
Ted Socha

FedCASIC
March 4, 2015

NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

Overview

- Evolution of NCES adaptive design strategies
  - From response propensity to nonresponse bias
- Progression of models across four longitudinal studies
  - Different populations, but similarity of available prior interview data and rich frame data
- Model selection, interventions, and results

Studies

- Education Longitudinal Study of 2002: 2012 Follow-up (ELS:2002)
- Baccalaureate and Beyond 2008: 2012 Follow-up (B&B:08/12)
- High School Longitudinal Study of 2009: 2013 update (HSLS:2009)
- Beginning Postsecondary Students Longitudinal Study 2012: 2014 Follow-up (BPS:12/14)

Response Propensity Focus

- Response rates declining, increasing likelihood for nonresponse bias
- Goal: identify likely nonrespondents and target in order to increase response rates
- Treatments: differential incentives, mode switching, etc.
- Use prior waves and/or field test data to predict response

Response Propensity Results

- Models developed to predict nonresponse
- Interventions increased response rates
- Did not contribute to reducing nonresponse bias – may even increase with higher response among similar cases
- Led NCES to refocus on targeting based on likelihood to contribute to nonresponse bias

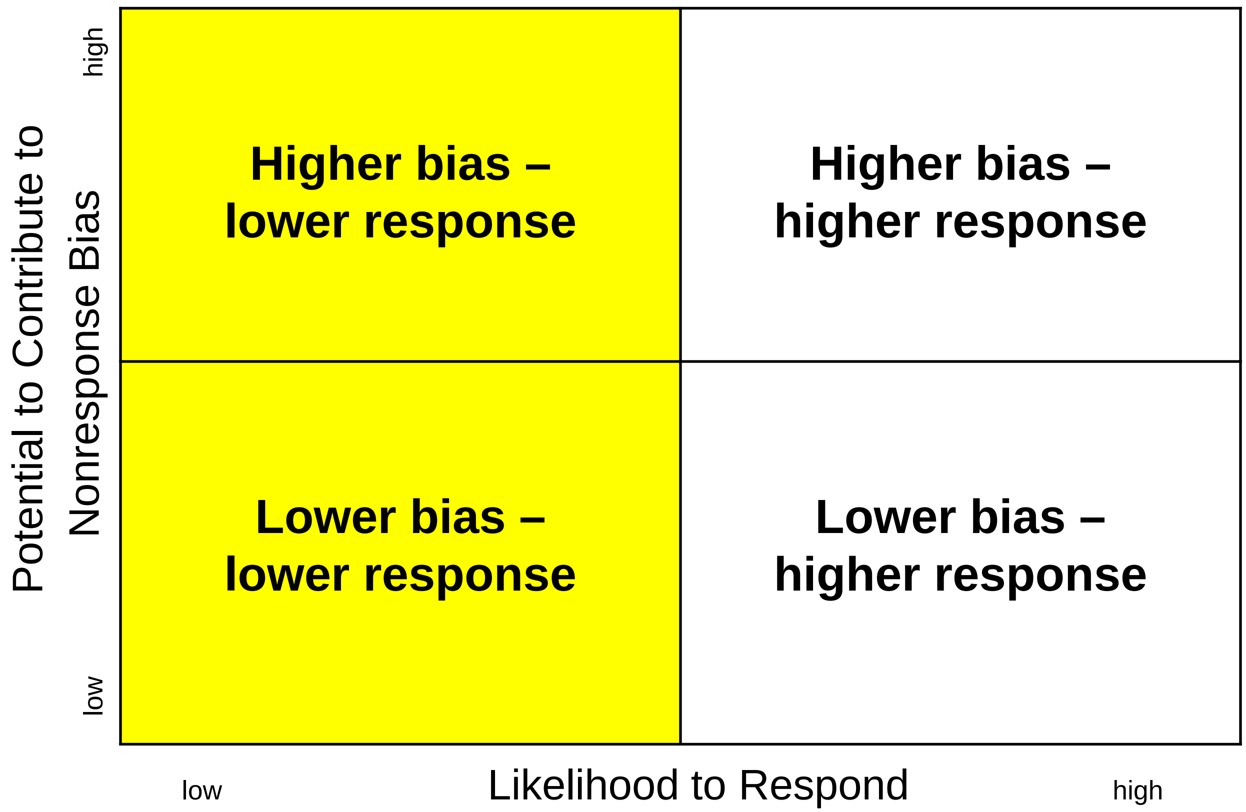
Conceptual Framework

	low	high
high	**Higher bias – lower response**	**Higher bias – higher response**
low	**Lower bias – lower response**	**Lower bias – higher response**

Potential to Contribute to Nonresponse Bias

Likelihood to Respond

Response Propensity Focus



Nonresponse Bias Focus

Potential to Contribute to Nonresponse Bias \ Likelihood to Respond	low	high
high	**Higher bias – lower response**	**Higher bias – higher response**
low	**Lower bias – lower response**	**Lower bias – higher response**

Moving to Adaptive Design

- Adaptive design: dynamic data collection strategy to reduce bias based on auxiliary information
- NCES longitudinal studies are well-suited for this
  - Paradata and substantive information from prior interviews
  - Rich frame data and administrative records
  - Mode flexibility

Mahalanobis Distance Scores

- At a given point, we identify nonrespondents who are likely to contribute to bias if they remain nonresponding
- M-score represents dissimilarity of each nonrespondent relative to mean respondent
- Each case has unique score
- Can rank nonrespondents most likely to contribute to bias

Implementing Mahalanobis

- Used in ELS:2002 and B&B:08/12
- Models used paradata and substantive variables, such as:
  - Demographic characteristics
  - Enrollment information
  - Prior round response
  - Contact attempts

Mahalanobis Results

- 3 intervention points with treatments such as:
  - Promised incentive boosts
  - Prepaid incentives
  - FedEx mailings
  - Switch to field (ELS) and abbreviated (B&B) interviews
- Some evidence of bias reduction, but varied by study and treatment type

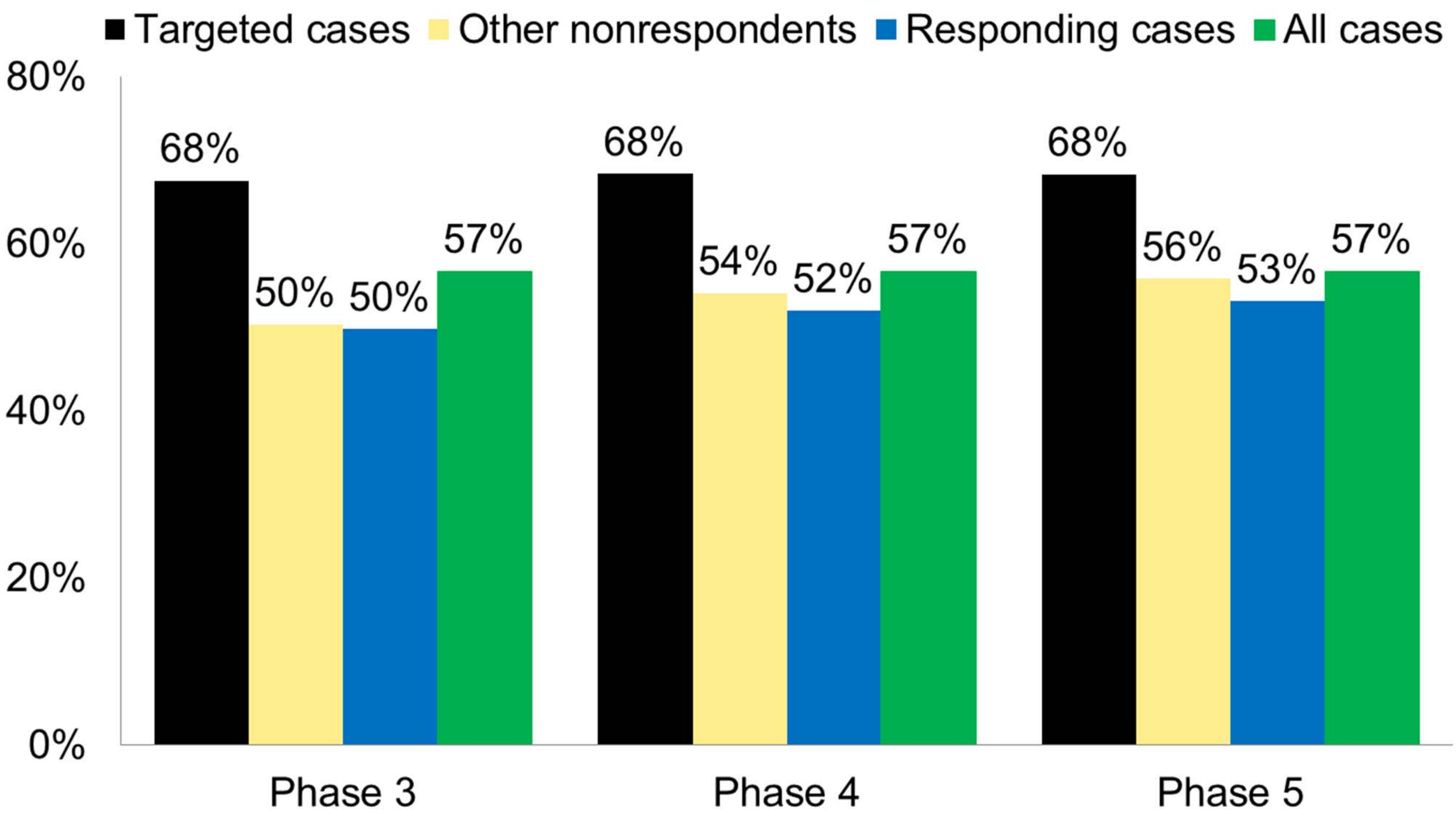
Nonresponse Bias Model

- Implemented in HSLS:2009 and BPS:12/14
- Use variables of interest to predict cases likely to introduce bias if they do not respond
  - Focus on substantive indicators rather than paradata
- Identify unique groups of cases underrepresented in respondent pool
- Provides method of prioritizing cases for treatment with finite resources

HSLS:2009 Phases

1. 3-week self-administered web period
2. 5-week computer-assisted telephone interview
3. $5 prepay for targeted cases
4. $15 offer for targeted cases
5. $25 offer for targeted cases
6. Expand cases for $5 prepay and/or $25 offer
7. Short survey for remaining cases (last 3 weeks)

HSLS:09 – Percent of cases who took Algebra 1 by respondent group and phase



HSLS:09 Results

- The green bar represents the percentage who took algebra 1 as reported in 2009 (study year 1)
- The blue bar represents the percentage who took algebra 1 as reported in 2013
- As adaptive phases progress, the respondent algebra 1 rate more closely approximates known 2009 rate

Combining Nonresponse Bias and Response Propensity

- Adaptive design can also be used for resource allocation
- Nonresponse bias model can identify cases likely to contribute to bias, but need to identify cases likely to respond to treatment
- BPS:12/14 included both nonresponse and response propensity models to generate an "importance score"

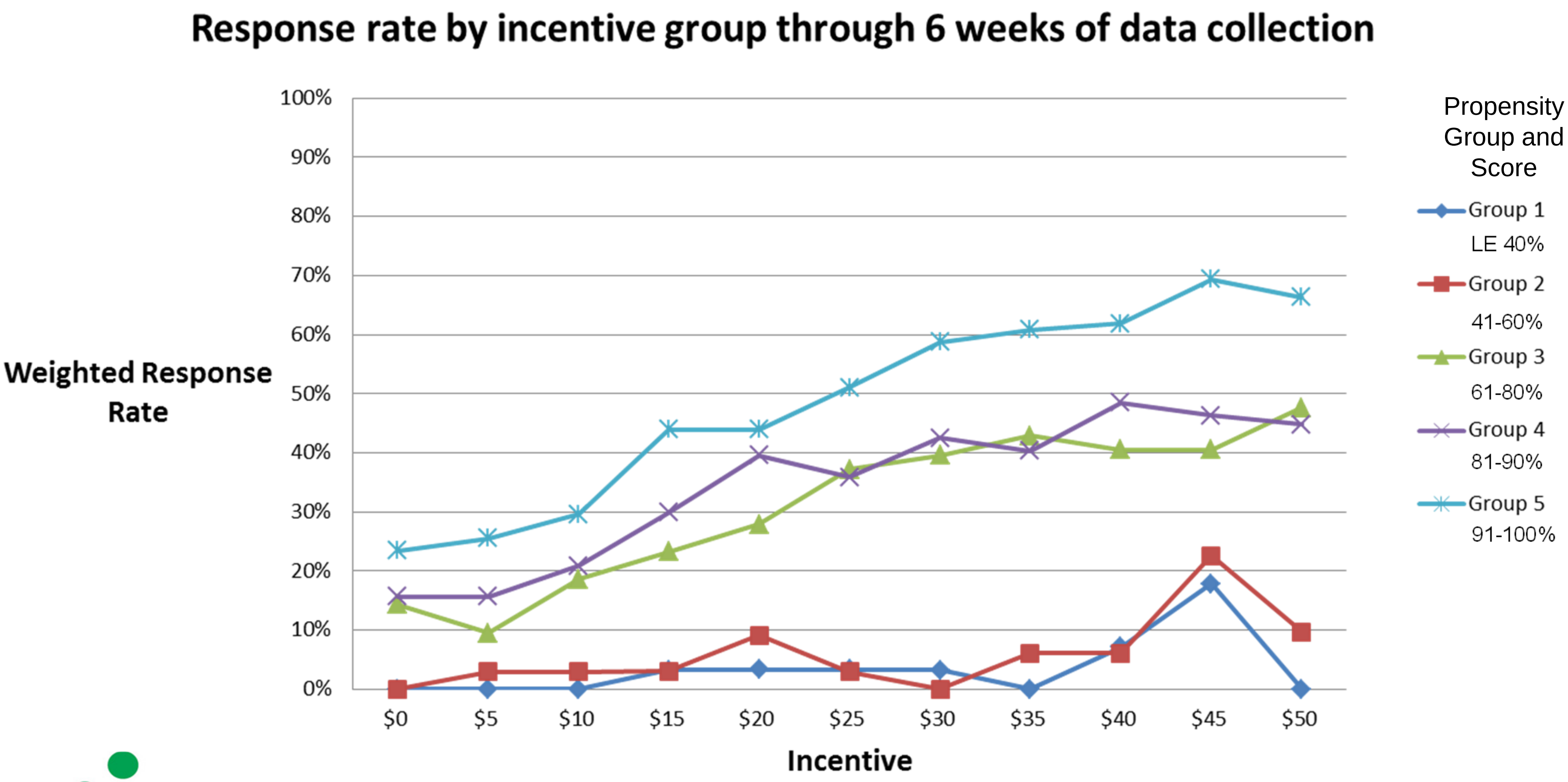
Importance Score Focus

Potential to Contribute to Nonresponse Bias / Likelihood to Respond	low	high
high	**Higher bias – lower response**	**Higher bias – higher response**
low	**Lower bias – lower response**	**Lower bias – higher response**

BPS:12/14 Experiment 1

- Ten percent subsample of the full 37,000 case national sample was drawn for testing
  - Started 7 weeks before main data collection
- Goal: Can we identify optimal initial incentive amount for a given response propensity?
- Promised incentives from $0 to $50 (in $5 increments) - compared across 5 levels of response propensity

Experiment 1 Response Rates



NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

BPS:12/14 Experiment 1 Results

- Response rate was nominally highest for groups 1– 2 at $45, but no statistical difference from lower amounts
- Response rate also nominally highest for groups 3– 5 at $45, but not significantly different from amounts of $30
- Offered $30 to remaining 90% main sample

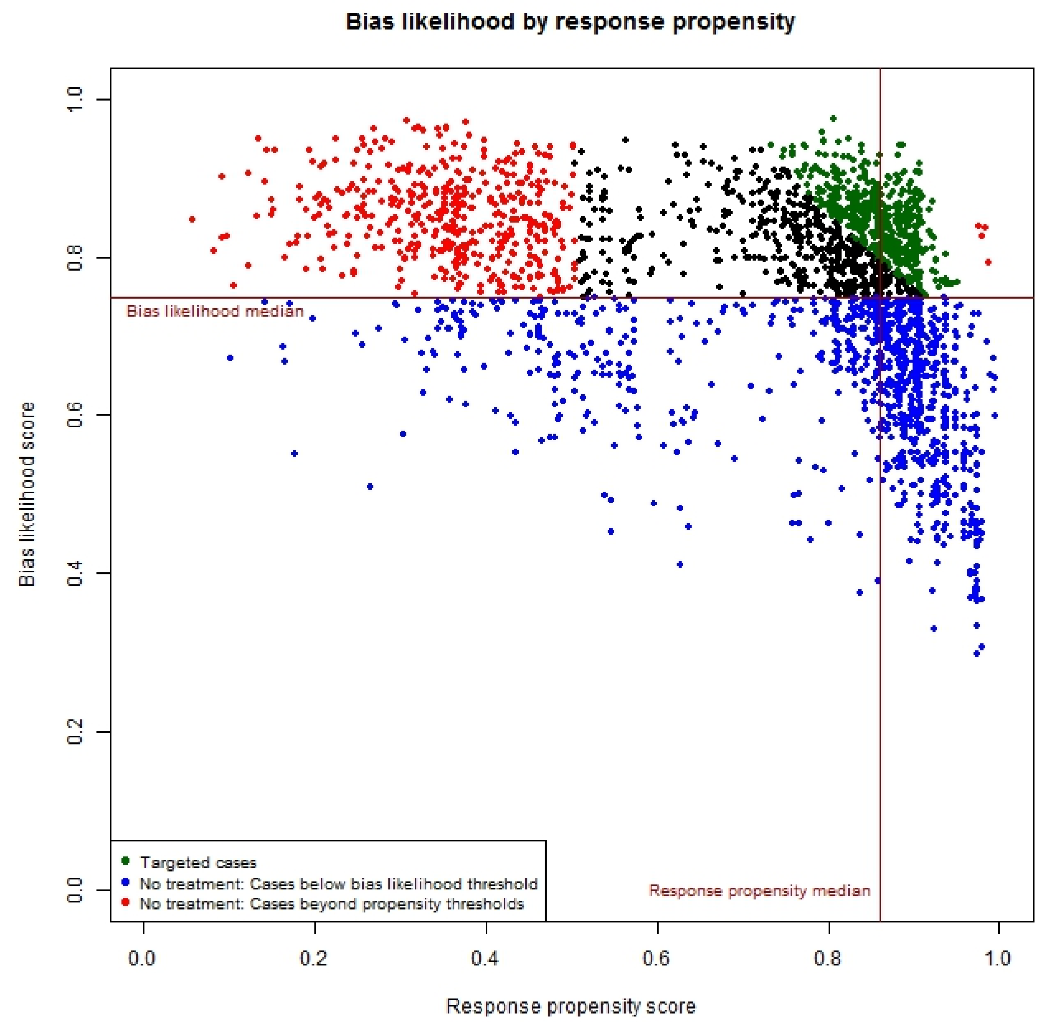
BPS:12/14 Importance Score

- Goal: identify nonrespondents that are likely to contribute to nonresponse bias and likely to respond to additional treatment
- Importance score product of two models:
  1. **Bias-likelihood model** identifies groups underrepresented at given data collection point
  2. **Propensity model** estimated prior to data collection which predicts likelihood to respond

Plot of Bias Likelihood by Response Propensity Score

Generate "importance score" as product of two models.

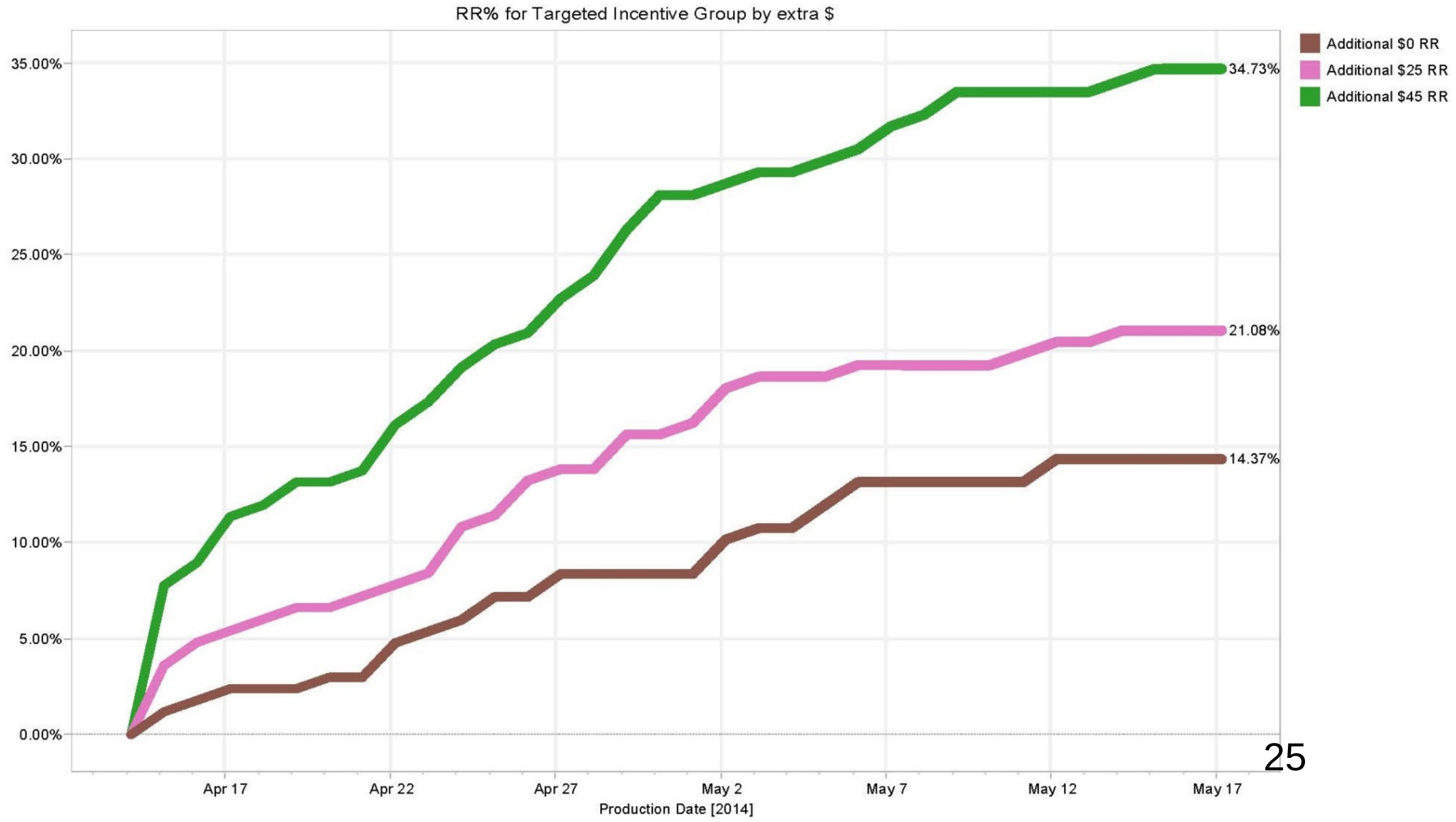
Select 500 cases with the highest importance score for experiment 2.



BPS:12/14 Experiment 2

- About 900 cases eligible for targeting, select the 500 highest importance scores
- Cases randomly divided into 3 treatments (divided across initial incentive groups):
  - $0 additional incentive offer
  - $25 additional incentive offer
  - $45 additional incentive offer

Experiment 2 Response Rate



BPS:12/14 Experiment 2 Results

- After a month of data collection during phase 3, analysis indicated:
  - $25 not significantly higher than $0
  - $45 significantly higher than $25 and $0
  - $45 led to reduction in bias in the largest number of estimates

Main Sample Implementation

- At time of intervention, response rate for main sample about 36%
  - About 21,400 nonrespondents remaining
- Targeted 30% of nonrespondent cases with $45 additional incentive
  - About 6,420 cases
- Same importance model used to select targeted cases

Summary

- NCES adaptive designs have adapted across studies
- Moving away from focus on response rate alone to focus on potential to contribute to nonresponse bias
- Leverage all sources of data to best identify cases worth targeting with finite resources