

Early Experience of Adaptive Design Work in the NSCG

Ben Reist

Federal Economic Statistics Advisory Committee Meeting

June 2014



U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

National Survey of College Graduates

- Sponsored by National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation (NSF)
- Part of the Science & Engineering Statistical Data System (SESTAT)
- Person-level survey sampled from American Community Survey (ACS)
- Target population is college graduates
- Occurs every 2-3 Years

2013 Data Collection

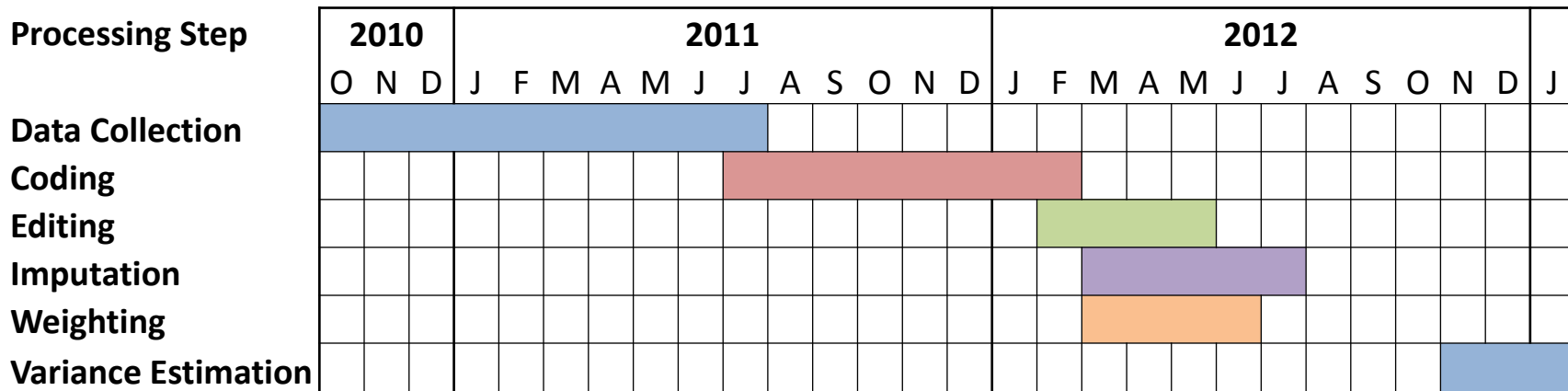
- February 21 – August 25
- Sample Size ~143,000 cases
 - 83,000 in New Cohort (2011 ACS)
 - 60,000 in Old Cohort (2009 ACS + 2010 NSRCG)
- Data collection modes include: internet, mail, phone
 - Different costs and effort

Methodology Studies

- What Strategies Work?
 - Incentive Timing
 - Priority Mail vs. First Class Mail
 - Mode Switching
 - Incentive Conditioning

Motivation for Adaptive Design

- NSCG Priority:
 - Reduce the time from start of data collection to delivery of finished product.



Needs to be done **without** sacrificing data quality!

Motivation for Adaptive Design

- Additional Goals
 - Allocate data collection resources efficiently
 - Avoid exhausting money and time
 - Move beyond response rate as the major metric of survey quality

Challenges to Implementation

- System:
 - Independent data collection systems
- Processing:
 - Move processing
 - Make assumptions
- Data Quality:
 - What measures do you use?
 - How do you use them in the decision-making process?

Adaptive Design Components Targeted for 2013 NSCG

Challenges Served as a Roadmap for 2013

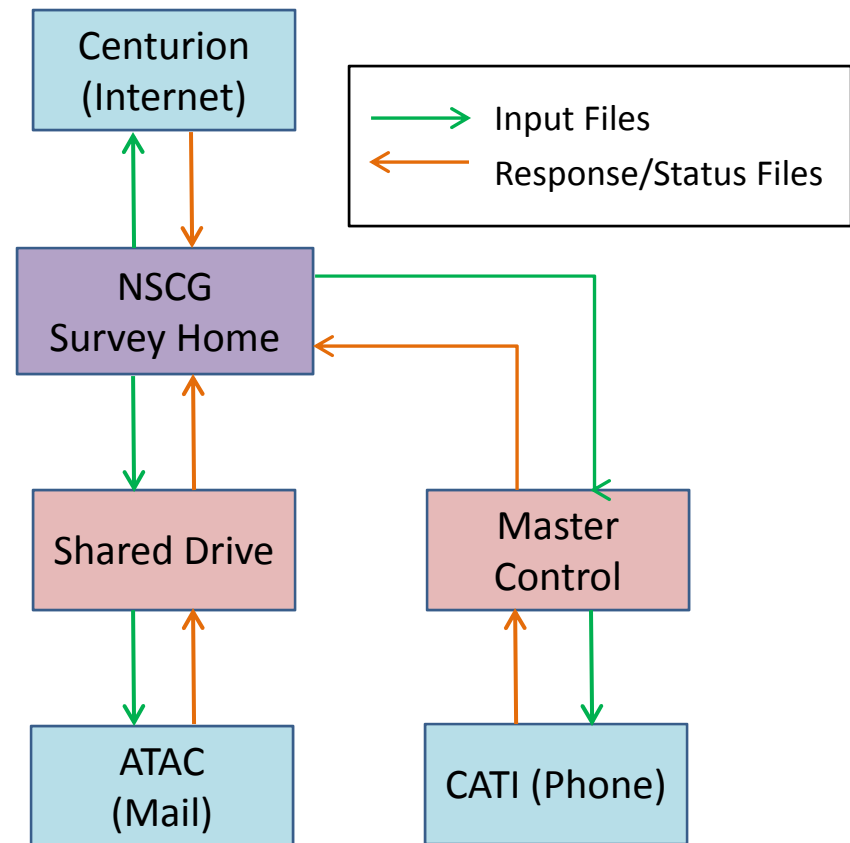
- Integrate Disparate Data Collection Systems
 - Integrated Systems
 - Integrated Reporting
- Institute Flow Processing
- Data Monitoring Methods
 - Increase Access to Paradata
 - Implement Methods
- Determine Possible Interventions

Integration of Systems

Integration of Systems

Baseline (2010)

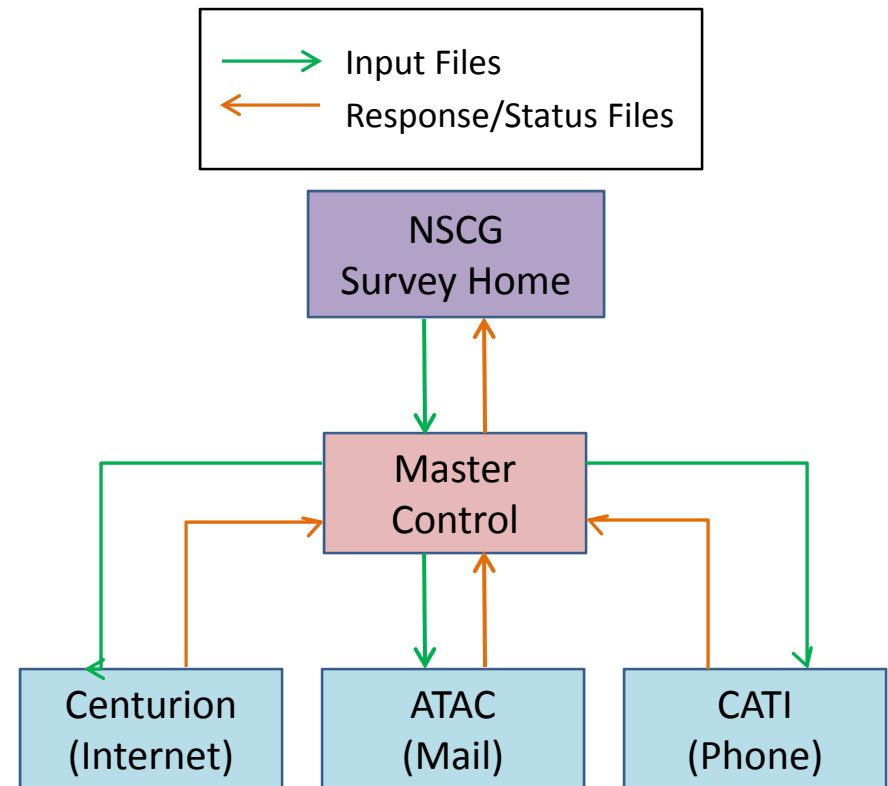
- Input files must be delivered to several different locations
- Many unrelated handoffs
- Separate intermediaries for mail and telephone
- Response files located in several different locations
- No mode-level interventions or communication without data flow to/from NSCG
- Different contact paths by mode



Integration of Systems

New Version (2013)

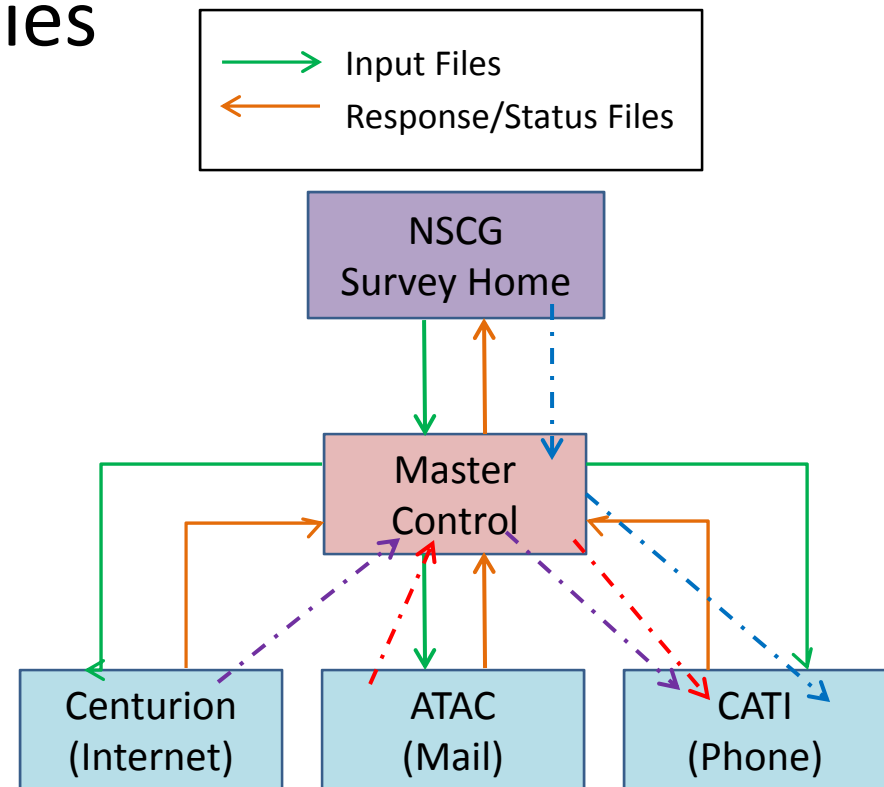
- Input files now delivered to one location
- Response files are now all in one location
- Single intermediary
 - Aware of all modes
 - Can pass info between modes
 - No need to wait for NSCG to affect action/interventions
- Single contact path for all modes



Integration of Systems

New System Functionalities

- CATI Holds from Internet
 - Every 2 Hours
- Mail Processing Holds
 - Daily
- Data Monitoring Holds
 - Weekly
- Integrated Reporting
 - Daily



Integrated Reporting

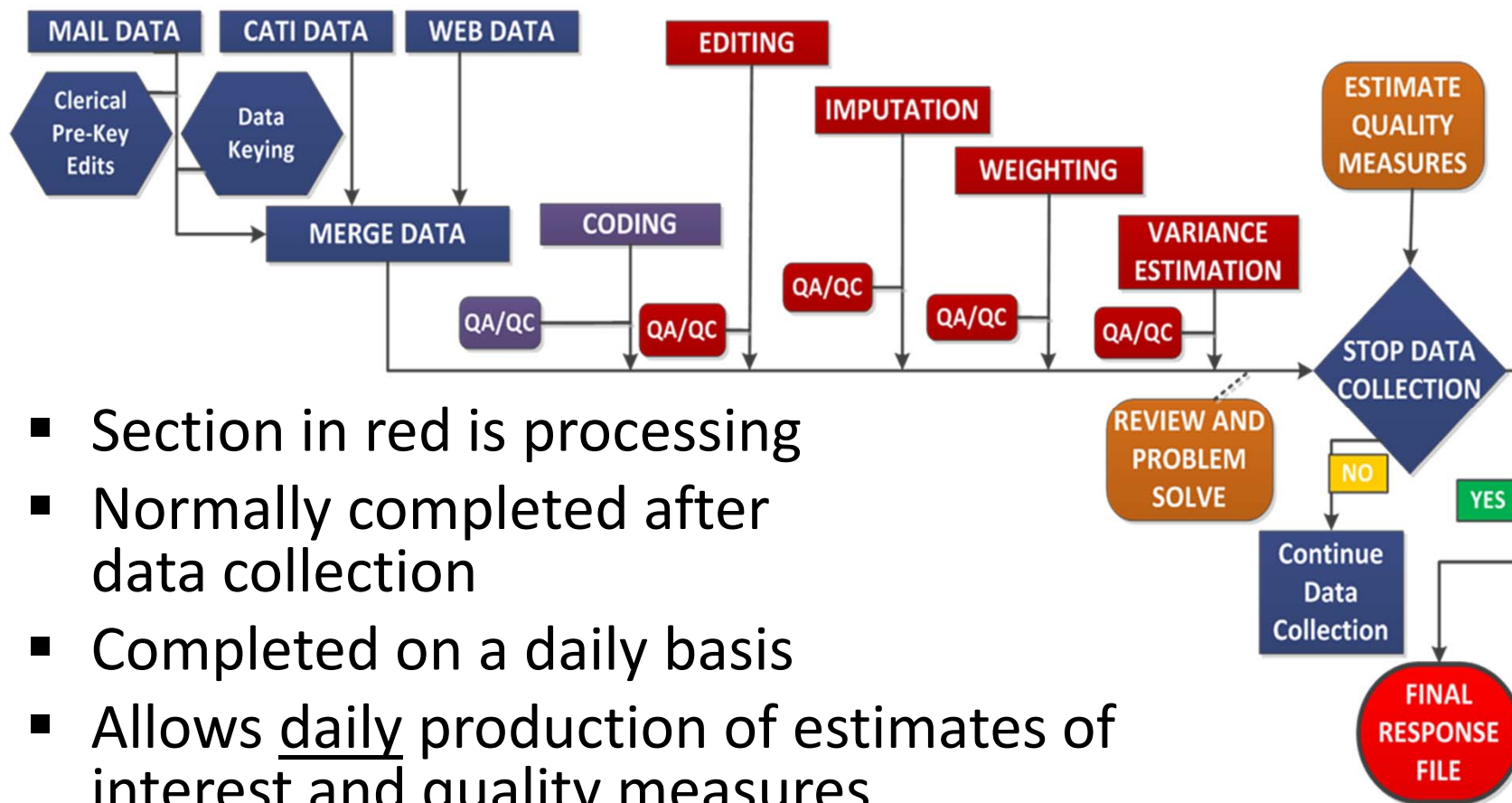
- Universal Tracking System (UTS)
 - Census Bureau enterprise-wide reporting system
 - Combines data streams from various systems
 - Met two major NSCG needs for adaptive design
 - Full Contact Path Report
 - Chronological report of all contacts for a sample person
 - Allowed us to respond to a specific sample person request
 - Contact Aggregation Report
 - Total contacts by category for a sample person
 - Include in data monitoring

Flow Processing

Flow Processing

- Complete most/all parts of processing
- NSCG has a goal of daily processing
- Make some assumptions
 - Less editing or less manual review
- Need coding, editing, imputation, weighting, and variance estimation

Flow Processing



- Section in red is processing
- Normally completed after data collection
- Completed on a daily basis
- Allows daily production of estimates of interest and quality measures

Flow Processing Benefits

- Operational Benefits
 - Processing programs completed earlier
 - Real-world testing opportunities
- Data Benefits
 - See effects of changes in editing or imputation rules immediately in the data
 - Daily views of “final” data and data quality
 - This information is important for data monitoring

Data Monitoring

Data Monitoring & Intervention

- Data-driven view of “what’s going on?”
- Make data-driven data collection interventions
- Propensity models
 - Uses frame, 2010, and 2013 NSCG data
 - Determine propensity to be in the respondent population
- R-indicators^{[1],[2]} (initial monitoring metric):
 - Great sampling frame (ACS)
 - What “type” of cases are responding?
 - Identify under-/over- represented groups

Data Monitoring & Intervention

- Benchmarking to frame and sample totals
 - Evaluate non-response propensity model
- Stability of estimates^[3]
 - Help develop stopping rules^{[4][5]}: Are new respondents moving the estimates/variance? Is it “worth it” to continue?
- Fraction of missing information^[6]
 - Help develop stopping rules: Measures uncertainty surrounding imputed values (Requires multiple imputation)

Interventions

- Data Monitoring provides information
 - Watch it or act on it?
 - 2013 NSCG includes mode-switching test
 - Monitoring methods help identify target cases
 - Move case to mode with the highest response propensity
 - Hold a case in web if it is a “low impact” case
 - Put a CATI case on hold (no contacts) if R-indicator indicates the group is over-represented
 - Need to identify more possibilities
 - Interventions are part of cost/quality tradeoff in adaptive design

Interventions

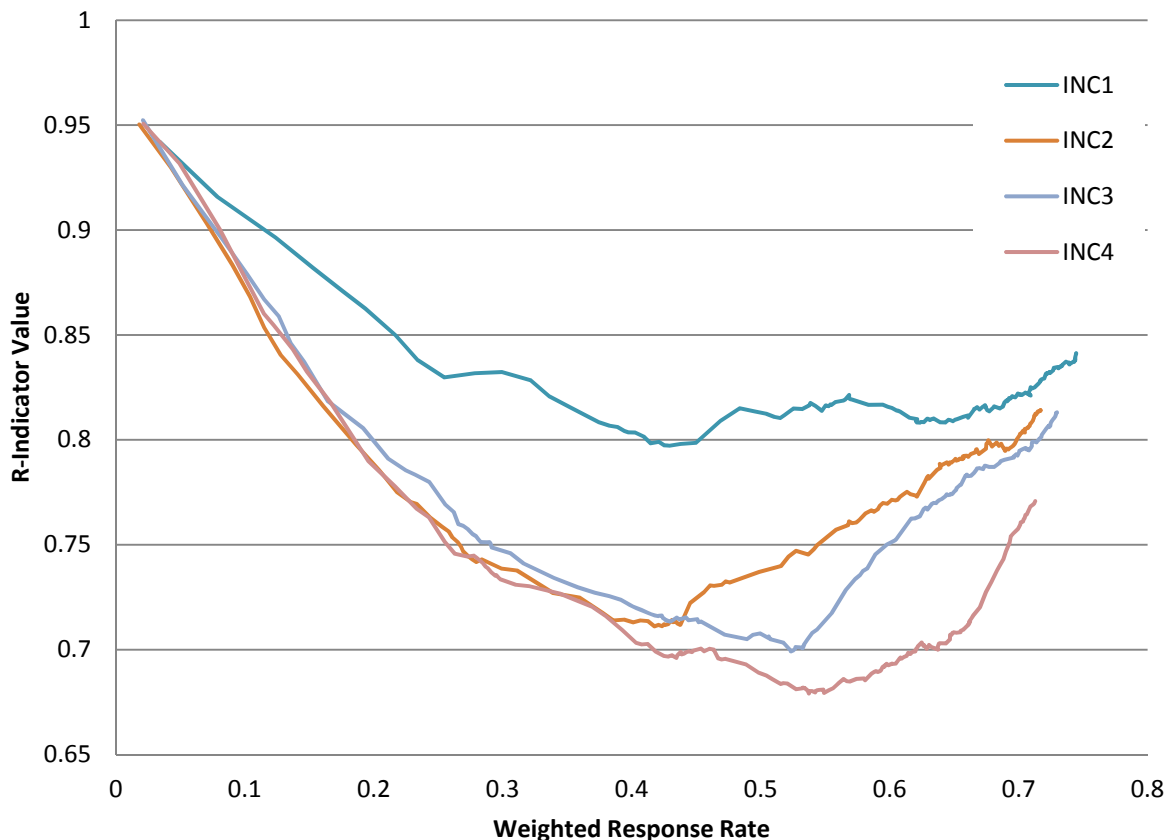
- Other types of interventions
 - Investigate and react to issues in data collection
 - Web server was extremely slow during first week of data collection
 - Used web paradata to identify time frame of slow service
 - Identified respondents affected by slow service
 - Mailed apology letter

R-Indicators Overview

- Sample R-Indicators
 - Evaluate representativeness of respondent population as compared to the sample population, given a set of balancing variables
- Unconditional Partial R-Indicators
 - Variable-Level
 - Evaluate which variables are driving the variation in propensities
 - Category-Level
 - Evaluate which subgroups of a variable or a cross of variables are over- or under-represented

R-Indicators Overview

Sample R-Indicators (Balancing Model) for Incentives Study Groups vs. Weighted Response Rate



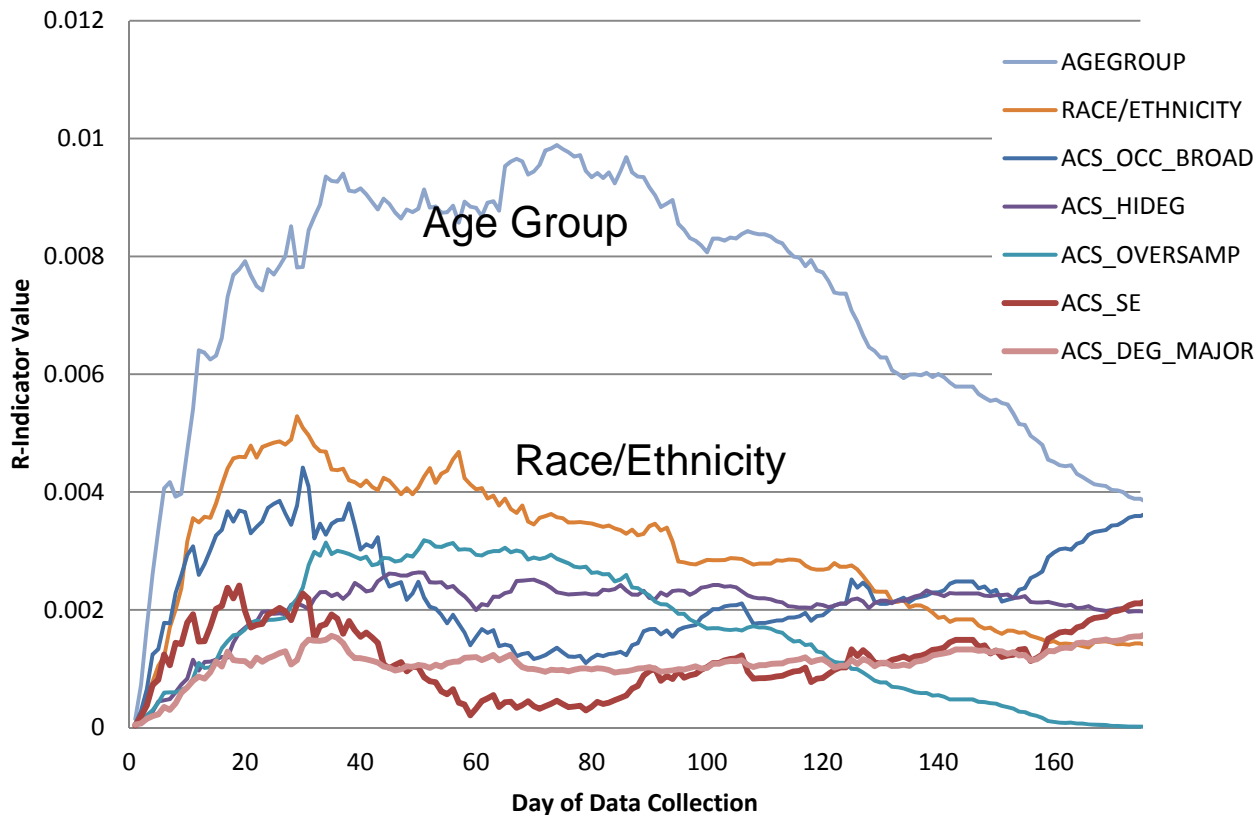
$$R(\hat{\rho}) = 1 - 2 \left(\sqrt{\frac{1}{N-1} \sum_{i=1}^N \frac{s_i}{\pi_i} (\hat{\rho}_i - \hat{\rho})^2} \right)$$

$$0 \leq R(\hat{\rho}) \leq 1$$

- $R(\rho) = 1$ means that the respondent population is fully representative of the sample population (all cases have the same propensity to respond)
- A decreasing R-Indicator means an increase in the variation in propensities.
- Can compare different samples (as here) provided the same variables are used in the balancing propensity model.

Data Monitoring Example

Unconditional R-Indicators for Variables in the Balancing Propensity Model (with Data Through 8/17) - MOSW



$$R_u(\text{var}, \hat{\rho}) = \sum_{k=1}^K \frac{N_k}{N} (\hat{\rho}_{x,k} - \hat{\rho}_x)^2$$

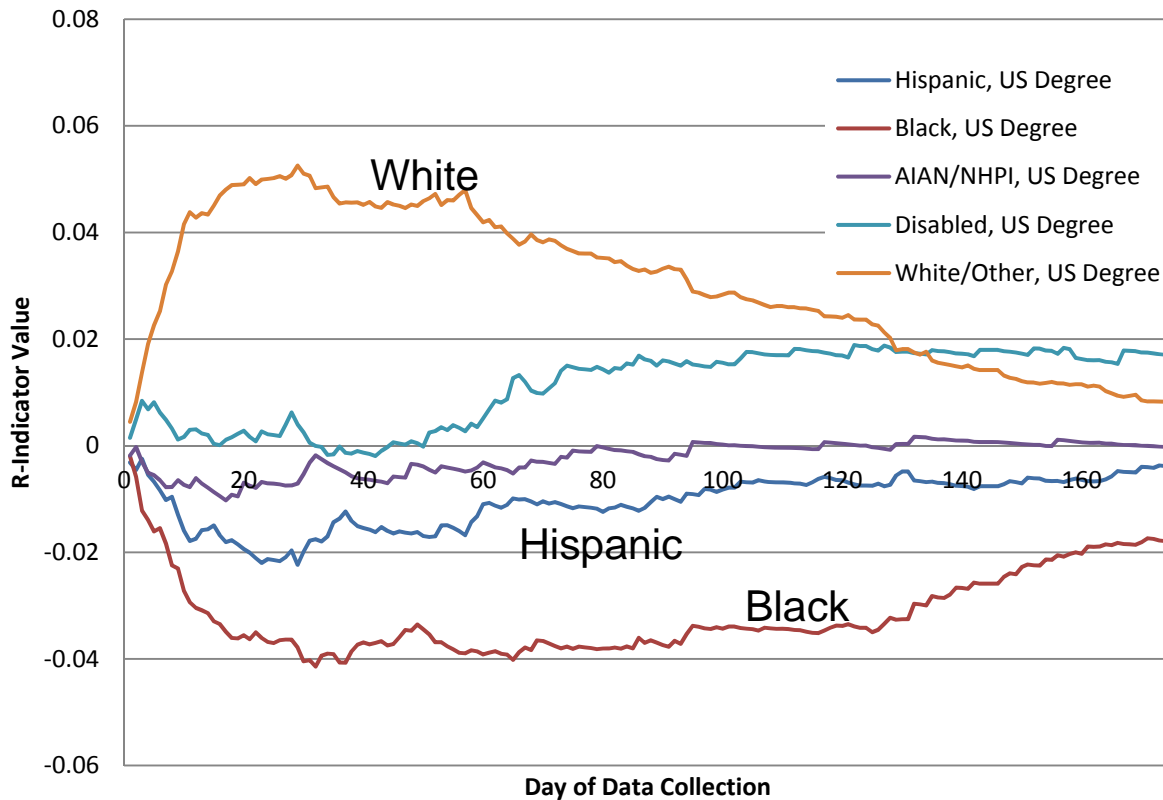
$$0.00 \leq R_u \leq 0.50$$

Variable –Level Unconditional Partial R-Indicators:

- **Identify variables** that drive variation in propensity.
- $R_u = 0$ means the variable does not drive variation in propensities

Data Monitoring Example

Partial Unconditional R-Indicators for Race/Ethnicity
(Data Through 8/17) - MOSW



$$R_u(\text{var}, k, \rho) = \sqrt{\frac{N_k}{N}} (\bar{\rho}_{x,k} - \bar{\rho}_x)$$

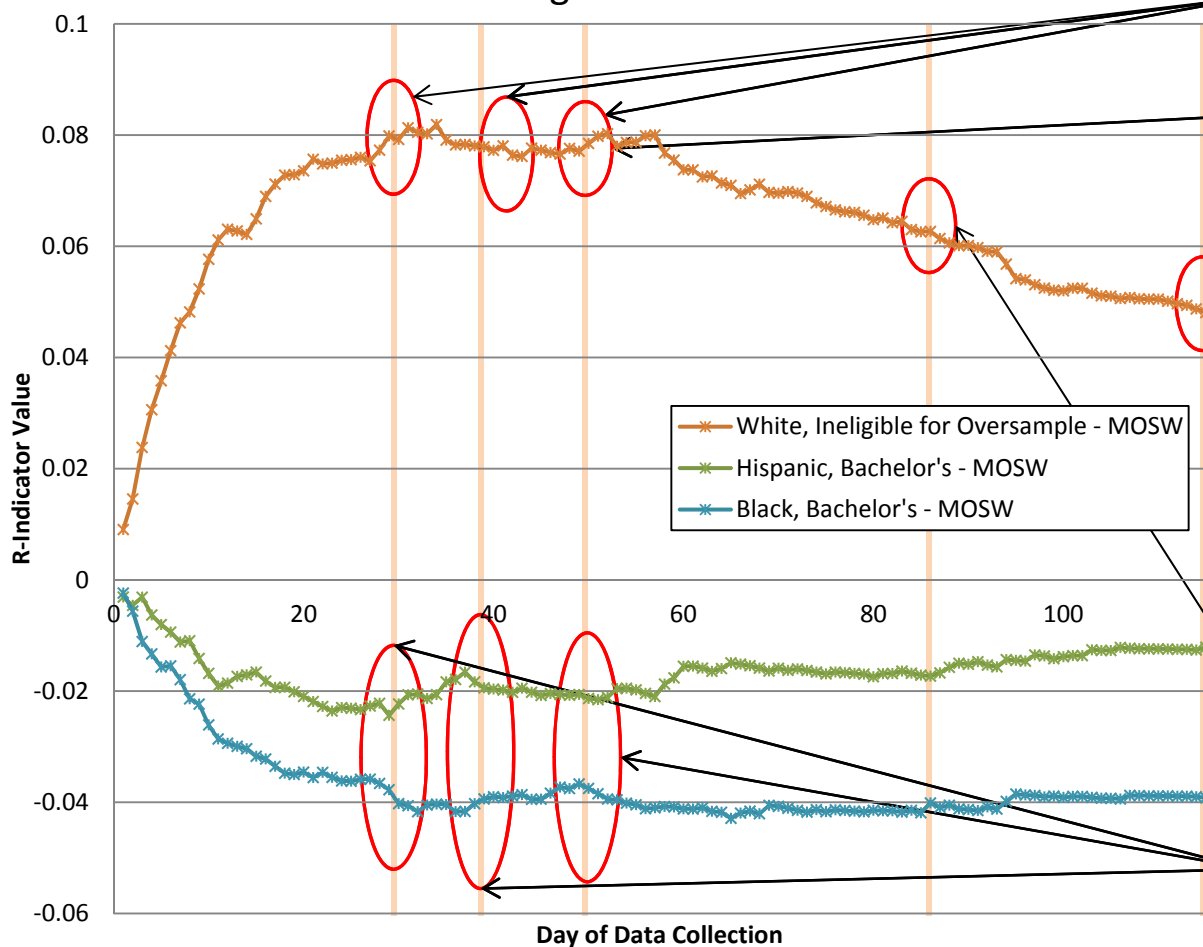
$$-0.50 \leq R_u(\text{var}, k, \rho) \leq 0.50$$

Category –Level Unconditional Partial R-Indicators:

- **Identify subgroups** that are over- or under- represented.
- This information can be used for targeting cases

Intervention Example

Unconditional Partial R Indicators for Targeted Subgroups
Mode Switching Interventions Schedule



Cases in the over-represented group & in CATI were put on hold to reduce contact attempts/shift resources to other cases. (Total of 40 cases)

For this intervention, cases in the over-represented group were identified. 50% of cases will only receive a web invite instead of a full questionnaire packet. Results in cheaper mailings, and reduction in future resources needed for keying. (Total of 498 cases)

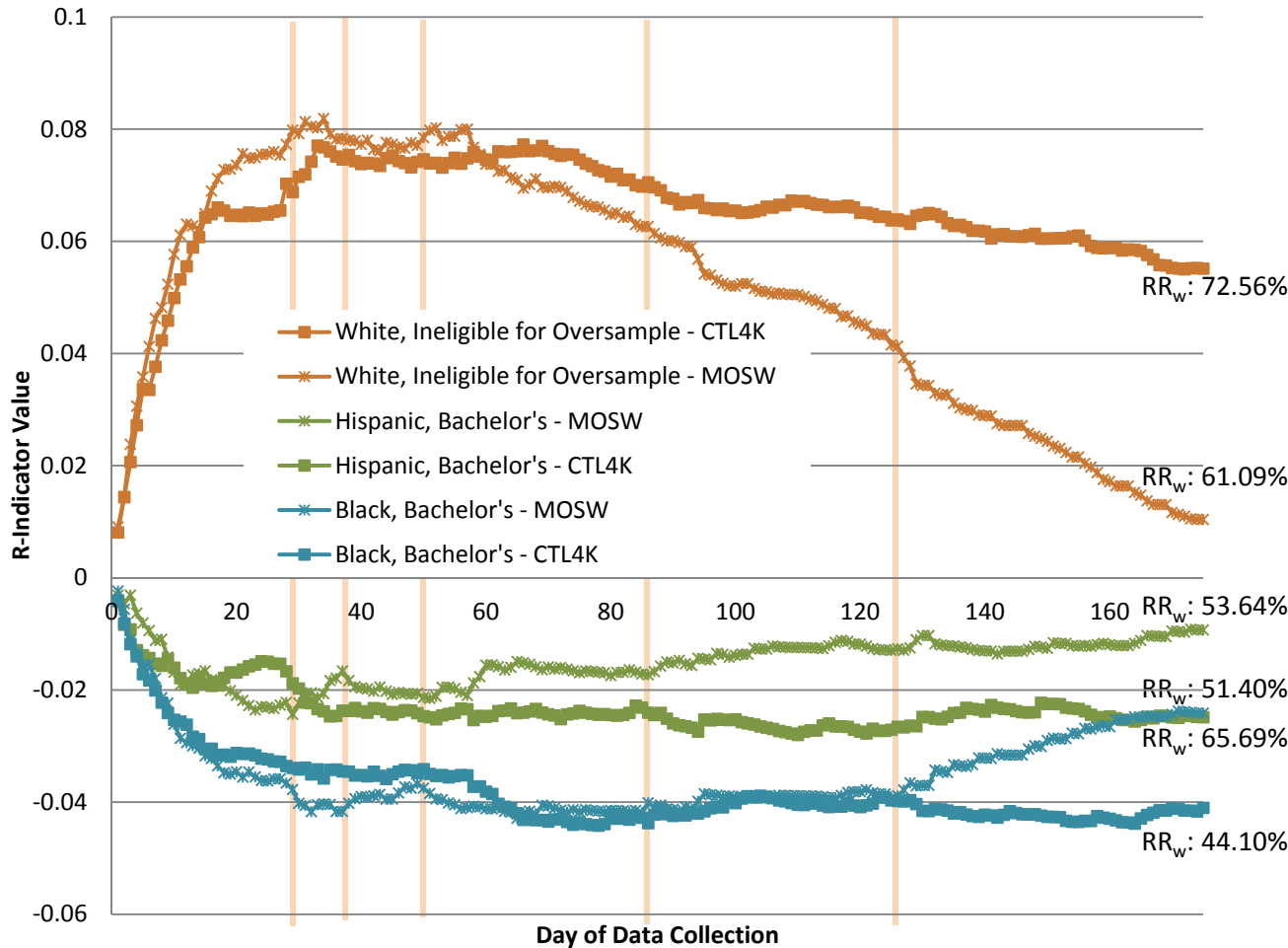
Cases in over-represented group were not sent week 18 questionnaire or week 23 final mailing. (Total of 508 cases)

Cases in over-represented group & not in CATI were held out of CATI to reduce contact attempts. (Total 495 cases)

Cases in under-represented groups moved to CATI to pursue those cases more aggressively. (Total of 85 cases)

Intervention Example – Is It Working?

Unconditional Partial R Indicators for Targeted Subgroups
(Data Through 8/17) Mode Switching vs. Control



- All interventions improved representativeness vs. a control where no mode switching occurred.

- Sending a web-invite only to over-represented cases resulted in fewer responses and reduced over-representation. (Tradeoff between Response/Representativeness)

- Moving cases to CATI in the under-represented groups resulted in increased response rates and representativeness as compared to the control.

- Until the end of data collection, the black bachelor population behaves nearly identically in both the mode switching and control group.

Questions

- What information needs to be provided to data users about interventions taken?
- How should we balance the quality of key estimates and quality of the microdata?
- How much adaptation is too much adaptation and how will we know?

References

- [1] Shouten, B., Cobben, F. & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology* **35**, 101-113.
- [2] Shouten, B., Shlomo, N., Skinner, C. (2011). Indicators for monitoring and improving representativeness of response. *J. Offic. Statist.* **27**, 231-253.
- [3] Groves, Robert M., and Steven Heeringa. (2006). Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society Series A: Statistics in Society*, **169**, 439-457.
- [4] Wagner, J., Raghunathan, T.E. (2009). A new stopping rule for surveys. *Statistics in Medicine*, **29**, 1014-1024.
- [5] Rao, R.S., Glickman, M.E., Glynn, R.J., (2008). Stopping rules for surveys with multiple waves of nonrespondent follow up. *Statistics in Medicine*, **27**, 2196-2213.
- [6] Wagner, J., (2010). The fraction of missing information as a tool for monitoring the quality of survey data. *Public Opinion Quarterly* **74**, 23-243.