



2019 Sixth International Workshop for Advances in Adaptive and Responsive Survey Designs

Suitland, Maryland November 4, 2019

Using Preliminary Estimates to Reduce the Data Collection Period in Repeated Surveys

Louis Rizzo, Minsun Riddles, David A. Marker, Westat
DavidMarker@Westat.com

Ending Data Collection Early

- *Lao Tzu: He who knows that enough is enough will always have enough.*
- 20 years ago research showed that ending a data collection with a relatively low response rate does not necessarily lead to high nonresponse biases (Keeter et al. (2000) and Curtin et al. (2000))
- Ending data collection early *if it is warranted* can:
 - Save field costs
 - Reduce re-contact burden on sample units
 - Allow for earlier release of public-use files and reports from the survey

Warranting an Early End to Data Collection

- Tracking key estimates derived from the responding sample at time points in the field cycle is one tool for warranting (others are measuring for example R-indicators).
- A raw comparison of responding units at that time point using sample base weights is too crude a comparison.
- Proper approach is to adjust weights for nonresponse and post-stratification.

Growing Set of Respondents

- Suppose the data collection time points are $t_1 < \dots < t_k < \dots < t_K$. The corresponding sets of respondents are $r(t_1), \dots, r(t_k), \dots, r(t_K)$, with
- $r(t_1) \subset \dots \subset r(t_k) \subset \dots \subset r(t_K) \subset r(F)$
- The respondent set gradually increases over time. F is the 'final time' in which data collection is terminated.

Reweighting Each Time Period

- d_i the base weight for sample unit i
- g_i all adjustments to the base weight for nonresponse, undercoverage, etc.
- Generate sets of weights at each time point as follows:

$$\{d_i g_i^{(t_1)}\}, \dots, \{d_i g_i^{(t_k)}\}, \dots, \{d_i g_i^{(t_K)}\}$$

- Each set of $\{d_i g_i^{(t_1)}\}, \dots, \{d_i g_i^{(t_k)}\}, \dots, \{d_i g_i^{(t_K)}\}$ weights are based on an approximation to the final set of weight calibrations

Approximate Final Weighting Procedures

- Repeated survey - approximate the weight calibrations done in the last cycle
- One-time survey - anticipate the final weight calibration
- Approximate the final weighting calibration procedure as well as possible - not necessary to match the final weights perfectly, as long as major steps are approximately represented

Visual Test or Formal Stopping Rule

- Using this paradigm, data collection can be ended at point t_K if all of the key estimates “show flatness” over an extended series of time points preceding t_K
- This flatness can be checked visually in graphs by subject-matter experts, or a formal test (or ‘stopping rule’) can be devised

Avoid Null Hypothesis of No Difference

- Stopping rule *should not* be based on a simple confidence test of the null hypothesis of no difference across the time points
- Rejecting that null hypothesis requires you believe strongly there is a difference across time points, *accepting it includes when you are not sure* if there are differences or not
- To stop, we want to be sure at a high level of confidence *that any differences are negligible*

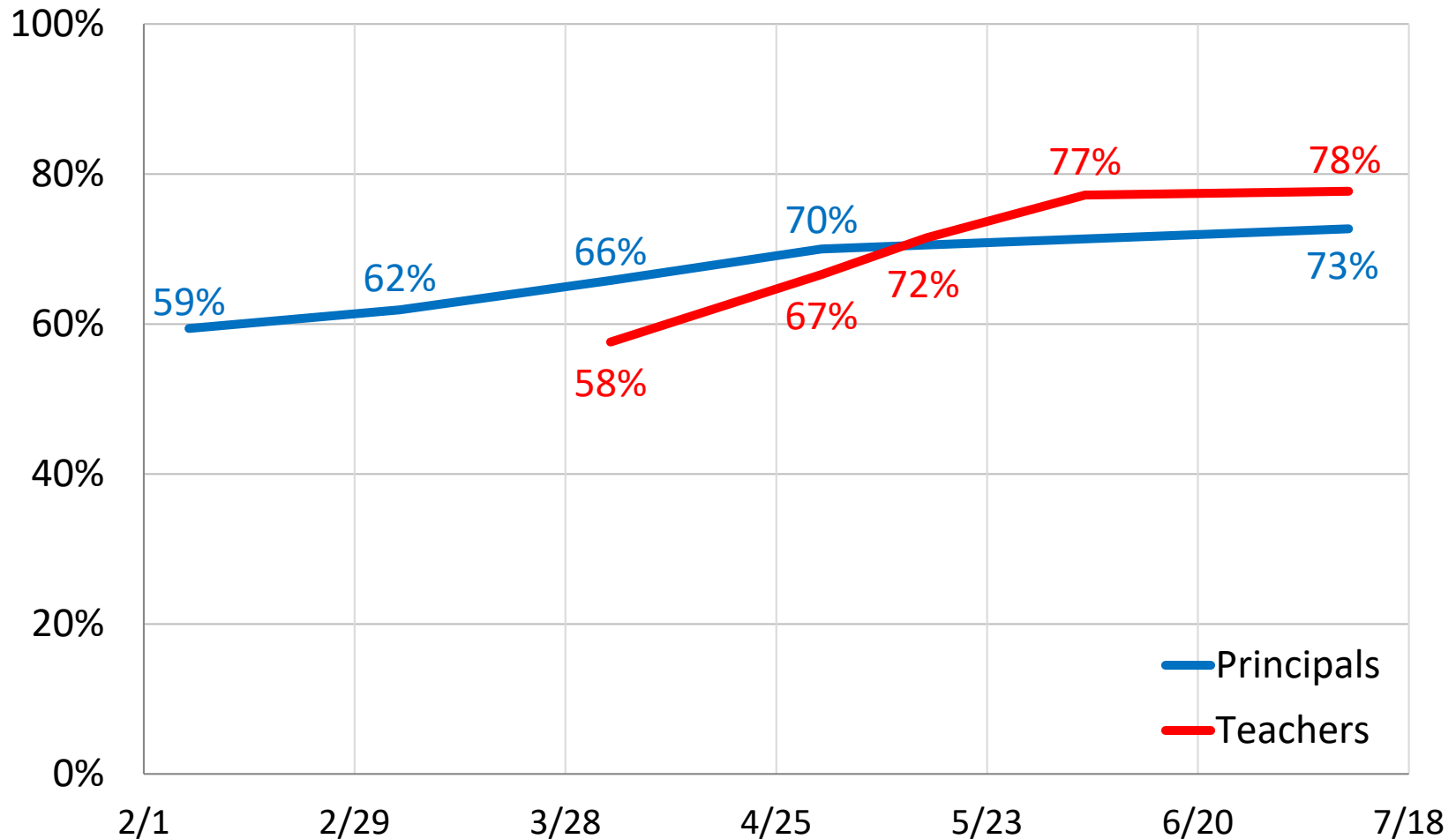
Changes Too Small to Matter

- Following Wagner and Raghunathan (2010) we define a stopping rule in terms of the probability that the difference between time periods is very small (below a cutoff value δ), and only stopping when this probability is very high.
- The cutoff value δ would be chosen based on subject-matter considerations (**practical significance 'too small' to matter**). A maximum probability of a 'false positive' needs to be set (a probability of assuming flatness when this is not valid).
- Wagner, J., and Raghunathan, T. E. (2010). "A new stopping rule for surveys." *Statistics in Medicine* 29(9)

NTPS SASS Example

- The National Teachers and Principals Survey (NTPS) is a nationally representative survey of schools, principals, and teachers in the US sponsored by the National Center of Education Statistics (NCES)
- NTPS was designed to be a replacement for the Schools and Staffing Survey (SASS), for which the last cycle was done in the school year 2011-12
- SASS design had a long field period (Sept-July) to maximize response rates
- A goal of NTPS was to shorten the period before publication of results, to make it more timely

SASS 2012 Response Rates by Date



NTPS SASS Example (cont.)

- School and principal response rates were little changed from April onwards
- Teacher response rates have little change the last two months (late May to late July)
- This extra data collection means:
 - Extra cost
 - Estimates cannot be published as quickly
- Do estimates change during these last months?

SASS Counterfactual Scenarios with Early Cutoff

- We created 8 data sets, 4 for principals, 4 for teachers, using different cutoff dates for the field period
- Reproduced initial public release tables from each data set, including a complete reweighting of the data
- Compared interim estimates to the final estimates
 - Do interim data sets produce similar estimates to the final data set?
 - Ideally identify a response level at which similar results are achieved, but earlier time frame

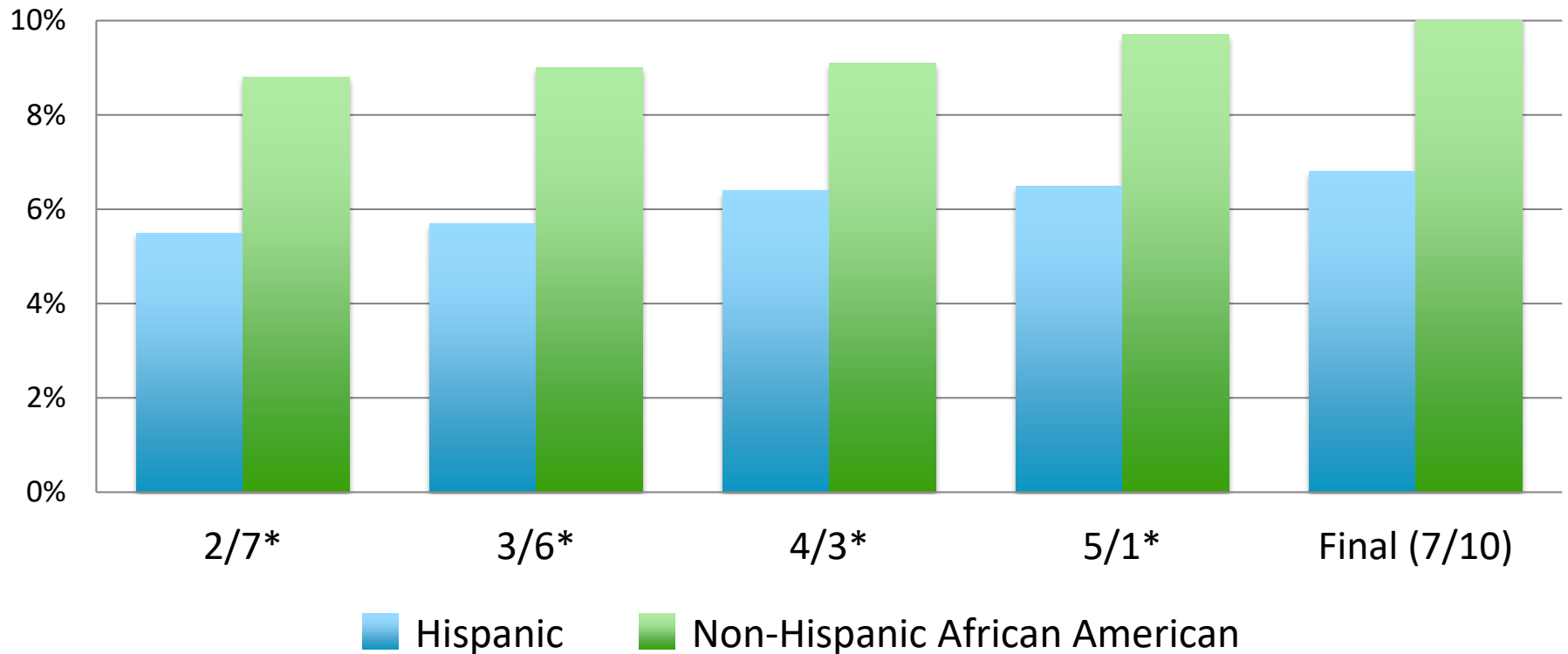
SASS Counterfactual Scenarios with Early Cutoff

- Results

- The vast majority of principal and teacher items showed very little movement in the counterfactual estimates over the four time periods
- Graphs in following slides show the few items with even limited change over the time periods

Race of Principal

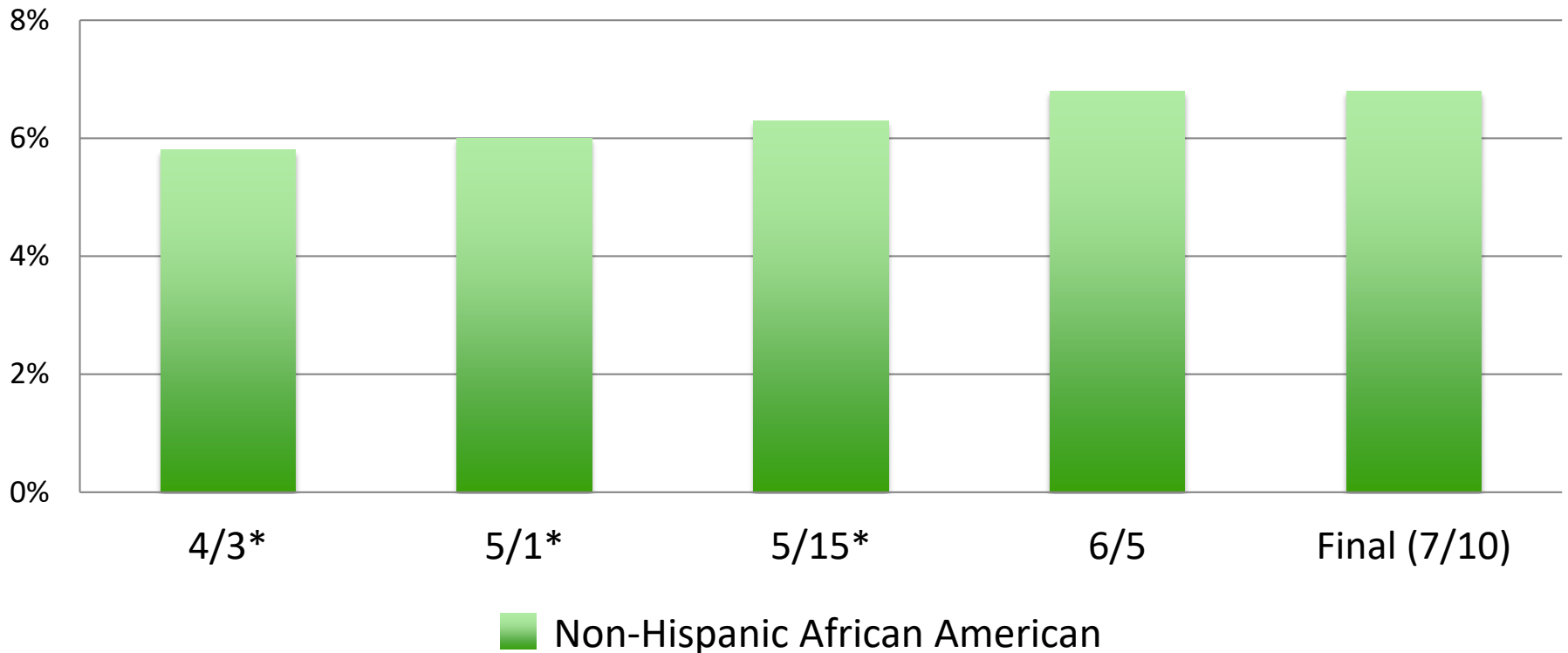
Principal Estimates



*Statistically significantly different ($p < 0.05$), t-test

Teacher Race

Teacher Estimates



*Statistically significantly different ($p < 0.05$), t-test

SASS Counterfactual Scenarios Results

- No practical change in estimates if we ended data collection:
 - Principals in early April (3+ months early)
 - Teachers in late May (almost 2 months early)
- Will other surveys show such a 'clean' result ?
- SASS 2011-12 certainly showed flatness with strong consistency, warranting in fact an early termination of data collection.

Summary

- We recommend tracking estimates over time periods, including mimicking the weighting at the end of data collection.
- These estimates should be tracked over a wide range of survey items.
- A consistent flatness of the estimates over a sufficiently long span can warrant an end to data collection.

Summary - Disrupting Phases of Data Collection

- The adaptive design community has recommended 'disrupting' data collection to bring in nonrespondents who have high nonresponse propensities
- Design phases with differing data collection protocols are geared to different sets of nonrespondents
- Our approach can allow one to discern when a particular design phase has played itself out

Summary - Organizational Flexibility

- To optimize data collection requires the organization to have flexible structures in place
- Re-assign field staff when cut-off is agreed upon
- IT system able to implement next steps (next phase of data collection, or back-end processing) on short notice

The Future

- We hope to implement this approach as part of a future cycle of NTPS or for other surveys. This requires a commitment of IT infrastructure to create the preliminary estimates and track over time.
- The particular 'stopping rules' would be survey-dependent. We hope to further develop the methodology based on a reasonable set of assumptions about cutoff points and maximum probabilities of false positives.



Improving Lives Through Research

www.westat.com