APPENDIX H

2015 NSCG Adaptive Design Experiment Goals, Interventions, and Monitoring Metrics

**2015 NSCG Adaptive Design Experiment Goals, Interventions,
and Monitoring Metrics**

The 2015 NSCG will include an adaptive design experiment to identify appropriate data collection interventions and monitoring methods for the NSCG. This appendix discusses adaptive design goals that NSCG will pursue as part of the 2015 NSCG, what interventions would allow the NSCG to achieve those goals, and what monitoring metrics would inform those interventions. Since adaptive design is a data-driven approach for contact tailoring to encourage response, the specific interventions used in the 2015 NSCG adaptive design experiment will be determined based on data monitoring results.

The 2013 NSCG Adaptive Design experiment employed only a few interventions, and had only one intervention available to apply to under-represented cases: moving a case to CATI [1]. The full set of possible interventions for the 2015 NSCG adaptive design experiment are discussed below. It should be noted that all of these interventions may not be used depending on the data monitoring results – this is a comprehensive list detailing the functionalities we plan to have available for 2015.

Goal 1: Balance Sample / Reduce Nonresponse Bias
Sampling balancing and/or reducing nonresponse bias relate to maintaining data quality in the face of shrinking budgets and falling response rates. Nonresponse bias arises when the outcomes of interest (the survey y-values) for respondents are different from those of nonrespondents. This difference results in a bias because the resulting estimates only represent a portion of the total target population. Surveys often try to correct for this *after* data collection using weighting, post-stratification, or other adjustments. Adaptive design interventions during data collection attempt to correct for nonresponse bias *during* data collection by actually changing the respondent population to be more balanced on frame characteristics related to response and outcome measures.

While discussing R-indicators, Schouten et al., provides reasons why balancing on variables related to response status and outcome variables is desirable. "In fact, we view the R-indicator as a lack-of-association measure. The weaker the association the better, as this implies that there is no evidence that non-response has affected the composition of the observed data." [3] This suggests that "selective forces…are absent in the selection of respondents" out of the sample population [2], and so nonresponse approaches missing at random, reducing the risk of non-response bias.

Interventions: Interventions are used to change the type or quantity of contacts targeted at specific subgroups or individuals. Interventions used in the 2013 experiment that would continue in the 2015 Test include:

- Sending an unscheduled mailing to sample persons[21];

---

[21] This was not actually used as an *ad hoc* intervention during data collection; rather, we sent apology letters to cases that experienced poor internet server performance. However, this intervention could be used in an adaptive design setting as well.

- Sending cases to CATI prior to the start of production CATI non-response follow up (NRFU), to target cases with an interviewer-assisted method rather than limiting contacts to self-response methods;

- Putting cases in CATI on hold, to reduce contacts in interviewer-assisted modes, while still requesting response in self-response modes;

- Withholding paper questionnaires while continuing to encourage response in the web mode to reduce the operational and processing costs associated with certain groups of cases; and

- Withholding web invites to discourage response in certain groups of cases, while still allowing these cases to respond using previous invitations.

Additional, new interventions being considered for use in the 2015 experiment include:

- Sending paper questionnaires to web nonresponders earlier than the scheduled mail date to provide two modes of self-response rather than one; and

- Changing the CATI call time prioritization to increase or decrease the probability a case is called during a specific time.

Monitoring Methods:
- R-indicators [2], [3], [4];

- Mahalanobis Distance or other distance measure [5];

- Response influence [6]; and

- Uncertainty/influence of imputed y-values [7].

We used R-indicators in the 2013 experiment, and plan to continue using them in the 2015 effort. As a metric, R-indicators were useful for measuring response balance, and served their purpose as a proof of concept for data monitoring. However, employing more metrics during data collection allows us to assess the usefulness of each monitoring metric and provides more confidence that data collection interventions were targeted in the most efficient way possible. That is, if R-indicators identify *subgroups* that should be targeted to increase response balance, and another metric (e.g., weighted response influence, Mahalanobis distance, etc.) identifies *specific cases* in those subgroups that also are likely to have an effect on nonresponse bias, then we have more confidence that those identified cases are the optimal cases for intervention, both from a response balance and non-response bias perspective. One particular metric of interest that we are exploring deals with uncertainty/influence of imputed y-values was discussed by Wagner at the 2014 Joint Statistical Meetings.

Goal 2: Increase Timeliness of Data Collection

Analysts and other data users that need relevant, up-to-date information to build models, investigate trends, and write policy statements rely on timely survey data. NCSES specifically focused on timeliness as a goal for the 2013 NSCG [4], and reduced the length of time from the beginning of data collection to the time of data release from 28 months to 12 months. This required a reduction in the data collection from ten months to six months. In the future, NCSES is interested in further reducing data collection, specifically, from six months to five months.

Interventions: Interventions will attempt to either encourage response to the NSCG earlier than the standard data collection pathway or will be used to stop data collection if new respondents are not changing key estimates. This could be achieved by introducing modes earlier than the standard data collection pathway, sending reminders that elicit response more quickly, or stopping data collection for all or a portion of cases and reallocating resources. Possible interventions include:

- Sending cases to CATI prior to the start of production CATI non-response follow up (NRFU), to target cases with an interviewer-assisted method rather than limiting contacts to self-response methods;

- Sending paper questionnaires to web nonresponders earlier than the scheduled mail date to provide two modes of self-response rather than one;

- Sending email reminders earlier than the scheduled dates in data collection; and

- Stopping data collection for the sample or for subgroups given a sufficient level of data quality. For example, we could stop data collection if:
    - key estimates have stabilized and standard errors fall within acceptable ranges, or
    - the coverage ratio for a subgroup of interest reaches a pre-determined threshold.

Monitoring Methods:

- Propensity to Respond by Modes [8];
- Change Point Analysis [9];
- Stability of Estimates [10]; and
- Coverage Ratios.

Ongoing NSCG research conducted by Chandra Erdman and Stephanie Coffey [8] could inform appropriate times to introduce new modes to cases ahead of the standard data collection schedule. Another possibility involves exploring change point analysis. If respondents per day as a metric changes over time, showing fewer responses in a given mode, there may be cause to introduce a new mode ahead of schedule. In addition, we will be able to calculate key estimates on a weekly or semi-weekly basis. As a result, we will be able to track stability of estimates during data collection to identify times when the data collection strategy has peaked, resulting in fewer responses or similar information that was already collected.

<u>**Goal 3: Reduce Cost**</u>

Controlling costs are always a survey management goal. More recently however, "the growing reluctance of the household population to survey requests has increased the effort that is required to obtain interviews and, thereby, the costs of data collection…[which] has threatened survey field budgets with increased risk of cost overruns" [10]. As a result, controlling cost is an important part of adaptive design. By allowing survey practitioners to reallocate resources during the data collection period, surveys can make tradeoffs to prioritize cost savings over other goals.

Interventions: Interventions will be used to encourage survey response via the web while discouraging response in more expensive modes (mail, CATI), or to eliminate contacts that may be ineffective. Possible interventions include:

- Putting CATI cases on hold, to reduce contacts in interviewer-assisted modes, while still requesting response in self-response modes;

- Withholding paper questionnaires while continuing to encourage response by web to reduce the operational and processing costs associated with certain groups of cases;

- Withholding web invites to discourage response from certain groups of cases, while still allowing these cases to respond using previous invitations;

- Prioritizing or deprioritizing cases in CATI during certain call times to increase or decrease the probability a case is called during a specific time frame without having to stop calling the case entirely; and

- Stopping data collection for the sample or for subgroups if key estimates and their standard errors have stabilized.

Monitoring Methods:
- R-indicators;

- Mahalanobis Distance or other distance measure;

- Response influence;

- Uncertainty/influence of imputed y-values;

- Stability of estimates; and

- Numbers of trips to locating.

The same indicators that are valuable for monitoring data quality also could measure survey cost reduction. If cases are in over-represented subgroups, or have low response influence, we may want to reduce or eliminate contacts on those cases.

In addition, the key estimates valuable to increasing timeliness, are also valuable for controlling cost. When estimates stabilize and their standard errors fall within acceptable limits for subgroups or the entire survey, new respondents are providing similar information to that which we have already collected. If continuing data collection would have little effect on estimates and

their standard errors, stopping data collection to all or subgroups of cases would be an efficient way to control costs.

Another potential cost-saving intervention would be to limit the number of times a case could be sent to locating. If we have no contact information for a case, or previously attempted contact information has not been useful for obtaining contact, a case is sent to locating where researchers attempt to identify new, more up-to-date contact information. This operation can be time intensive, especially for cases repeatedly sent to locating. We could track the number of times a case is sent to interactive locating, or the length of time it spends in locating. Cases repeatedly send to locating or spend a large amount of time being researched may not be ultimately productive cases. Reallocating effort spent on these cases to those in locating for a fewer number of times may be a sensible cost-saving measure that allows us to attempt contact on more cases, rather than spending large amounts of time (money) on the same cases.

References:

[1] Coffey, S. "Report for the 2013 National Survey of College Graduates Methodological Research Adaptive Design Experiment". Census Bureau Memorandum for NCSES. April, 2014.

[2] Schouten, B. Cobben, F. Bethlehem, J. "Indicators for representativeness of survey response." *Survey Methodology.* 35.1 (June 2009): pp 101 – 113.

[3] Schouten, B. Shlomo, N. Skinner, C. "Indicators for monitoring and improving representativeness of response." *Journal of Official Statistics.* 27.2 (2011): pp 231 – 253.

[4] Coffey, S. Reist, B. White, M. "Monitoring Methods for Adaptive Design in the National Survey of College Graduates (NSCG)." *2013 Joint Statistical Meeting Proceedings,* Survey Research Methods Section. Alexandria, VA: American Statistical Association.

[5] de Leon A.R., Carriere K.C. "A generalized Mahalanobis distance for mixed data." *Journal of Multivariate Analysis.* 92 (2005). 174-185.

[6] Särndal, C., Lundström, S. (2008). Assessing auxiliary vectors for control of nonresponse bias in the calibration estimator. *Journal of Official Statistics.* 24, 167-191.

[7] Wagner, J. (2014). "Limiting the Risk of Nonresponse Bias by Using Regression Diagnostics as a Guide to Data Collection**."** Presentation at the 2014 Joint Statistical Meetings. August, 2014

[8] Erdman C., Coffey S. (2014). "Predicting Response Mode During Data Collection in the NSCG." Presentation at the 2014 Joint Statistical Meetings. August, 2014

[9] Killick, R. Eckley, I. "Changepoint: An R Package for Changepoint Analysis". Downloaded from http://www.lancs.ac.uk/~killick/Pub/KillickEckley2011.pdf on August 8., 2014.

[10] 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.