APPENDIX C

The NTEWS Methodological Experiments

Research Questions and Design Rationale for

Contact Strategy, Incentive, and Seeded-Sample Experimentation

1. Introduction and overview

This appendix provides specific detail on the rationale for and design of the methodological experiments that will be incorporated into the initial National Training, Education, and Workforce Survey (NTEWS).

Because this survey is the first cycle of the NTEWS, NCSES plans to determine the optimal design features for future cycles of the survey through two methodological experiments—one on contact strategies and one on noncontingent incentives. Both experiments seek to identify data collection strategies that will minimize nonresponse bias while maintaining cost-effectiveness for the federal government. Readers should note that the range of contact strategies and incentives tested in this 2022 administration is not what will be used in future NTEWS administrations. Rather, the broad range of contact strategies and incentives tested in this administration will provide the baseline information needed to determine the optimal allocation of contact strategies and incentives for future cycles. For example, analysts will use information from these experiments to minimize the use of monetary incentives in future administrations, using them only for sample members for which they are necessary in order to cost-effectively reduce nonresponse bias. These experiments will thus provide the baseline data needed for evidence-based decision making for the methodological design of future NTEWS cycles.

Previous research, discussed in detail later in this appendix, has not clearly demonstrated the optimal design for a mixed-mode federal survey like NTEWS. The NTEWS experimentation will provide much-needed evidence on best practices for nationally representative, mixed-mode surveys. Specifically, the findings from these experiments will benefit the broader federal statistical system by providing insight into the effect of alternative contact strategies and noncontingent incentive levels on reducing the risk of nonresponse bias among the adult population. Given continued declines in response rates, the identification of data collection approaches that can produce valid estimates for the U.S. population even with relatively low response rates remains a continued priority across the federal statistical system. Because the nationally representative NTEWS sample will be drawn from respondents to the American Community Survey (ACS), a rich set of information will be available for all sample members (regardless of response status). This will permit a much more comprehensive nonresponse bias analysis than is typically possible for federal surveys, which are often drawn from frames with less detailed or accurate information about nonresponding sample members. Thus, this initial cycle of the NTEWS offers the opportunity to generate insights that will be applicable not only to future NTEWS cycles, but also to other federal household surveys.

In general, the methodological design of the NTEWS, and decisions about necessary experimentation, were informed by the experience of three surveys from which the bulk of the NTEWS items were drawn or adapted:

- The Adult Training and Education Survey (ATES), originally a topical module of the NCES National Household Education Surveys (NHES) program, from which many items on licenses, certifications, work experience programs, and education were drawn. The

ATES module of the NHES was discontinued after NCES and NCSES partnered to develop the NTEWS.

- The National Survey of College Graduates (NSCG), from which items related to work paths and some demographic items were drawn. The NTEWS sampling methodology is based on the NSCG sample design.
- The ACS, which provides the sampling frame for the NTEWS and on which many of the NTEWS demographic and employment items are based.

Note that, while the NSCG and the NTEWS are currently separate surveys targeting different populations (NSCG targets a subpopulation with bachelor's degree or above and NTEWS those with a subbaccalaureate credential), NCSES's long-term strategy is to examine the feasibility and potential benefit of merging the two surveys into one. Thus, the existing NSCG methodology, which has been tested over several survey cycles, helped guide most of the experimental decisions made for the initial NTEWS cycle.

1.1. Randomized, methodological experiments in nationally representative sample

This initial NTEWS cycle, which plans to sample 43,200 individuals from respondents to the 2018 ACS, will implement two randomized methodological experiments:

- A **contact-strategies experiment** to test six alternative sequences of the three response modes that will be offered in the NTEWS (web, paper, and telephone).
- A **noncontingent-incentive experiment** to test three alternative dollar values ($10, $20, $30) for a prepaid debit card, as well as a no-incentive control group.

Both sets of treatment groups will be randomly assigned and fully crossed in a factorial design, allowing the effects of the alternative contact strategies and noncontingent incentives to be examined independently. Also, NCSES plans to examine any potential interaction effects between the contact and noncontingent incentives treatments.

Detailed discussions of these experiments—including their rationale and methodological designs—are provided in sections 2 (for the contact-strategies experiment) and 3 (for the noncontingent-incentive experiment). Section 4 then shows expected sample sizes for each combination of contact-strategies and noncontingent-incentive treatments.

The following key research questions apply to both experiments:

1. **Nonresponse bias for key estimates**: Do key survey estimates vary across contact strategies and incentive levels? Can these differences be explained by sampling error (noise) or by differences in demographic representation, which can be mitigated by conventional weighting adjustments? If differences remain after accounting for sampling variability and demographic representation, is there an optimal contact strategy and/or incentive level that leads to the highest level of reporting/least amount of bias for key estimates?
2. **Nonresponse rates for key subpopulations**: Which population groups of interest have lower-than- and higher-than-average response rates to the NTEWS? What are the impacts

of contact strategies and incentive levels on these response rates and corresponding weight adjustments related to estimates of uncertainty (standard errors)?

3. **Predicting the likelihood of response using frame data**: To what extent can sample members for which a particular contact strategy or incentive level is effective (or ineffective) be reliably identified using the ACS frame data?

The overarching motivation for these research questions is the need to identify an optimal data collection strategy that minimizes nonresponse bias while also minimizing, as much as possible, the cost to the federal government and the burden on the public. *Nonresponse bias* occurs when the composition of the *responding* sample differs systematically from that of the target population due to non-random variations in response rates across subgroups. This bias lowers the validity of the estimates and of comparisons across groups that include the biased estimates. For this reason, the primary research questions for the NTEWS experiments encompass the effects of alternative treatments on the final composition of the responding sample (research question 1) and on the variability of response rates across subpopulations (research question 2). The experiments are also designed to determine whether nonresponse bias could be reduced in future NTEWS cycles by applying treatments differentially across key subgroups (research question 3).

In planning for future cycles of the NTEWS, NCSES will use the results of these experiments to identify a methodological design that best balances the goals of minimizing nonresponse bias, minimizing data collection costs, and producing estimates at an acceptable level of statistical reliability. In determining the proper balance, data quality will be of highest priority.

1.2.   Seeded sample evaluation

Additionally, separate from the nationally representative production sample, the initial fielding of the NTEWS will contain a convenience sample of 1,000 persons known to hold a postsecondary certificate. This sample is referred to as the **seeded sample** and is intended to evaluate measurement error in the postsecondary certificate item. Results will be compared to prior ATES seeded sample research to determine whether revisions to the postsecondary certificate item succeeded at reducing underreporting of certificates. Detailed discussion of the seeded sample evaluation is provided in section 5.

2. Contact-strategies experiment

2.1. Background

One of the most important design decisions of a survey is the contact strategy. The contact strategy makes potential respondents aware of the survey and gives them a means to respond. Research has shown that the type of contact, its timing, and the number of contacts (Dillman et al. 2014), as well as the level of personalization (Cook et al. 2000), can influence a respondent's decision to participate in a survey. The strategy that will be most effective for a particular survey will depend on many factors, including the target population, the sponsor, and the mandatory status of the survey.

As a result, the initial cycle of the NTEWS will include a contact-strategies experiment to test alternative sequences of offered response modes. The goals of this experiment are (1) to

determine which contact sequence maximizes the representativeness and size of the respondent sample, and (2) to assess the effect of response mode on responses to key survey items. Sample members will be randomly assigned to one of six contact-strategy treatment groups, described in greater detail later in this section. All treatment groups include both mail and phone contacts (e-mail contacts are not included due to a lack of e-mail addresses for sample members). All sample members will be offered the option to respond by web, paper, or computer-assisted telephone interviewing (CATI) at various points in the data collection. The contact-strategy treatment groups will vary *the sequence in which* these modes are offered.

The experimental design reflects a recognition that, although existing research supports the use of mixed-mode designs over web-only designs, there is varied evidence as to the best way to combine response modes to minimize nonresponse bias while maximizing response rates (cf. Freedman et al. 2018; Bucks, Couper, & Fulford 2019). Some studies find that offering both mail and web response modes concurrently can adversely impact the response rate (cf. Medway & Fulton 2012; Dillman et al. 2014). While limited research has been conducted to directly compare sequential and concurrent mixed-mode designs, some of the more recent studies to have done so have found acceptable results with a concurrent mixed-mode design (cf. Matthews et al. 2012; Bucks, Couper, & Fulford 2019). Other research has indicated potential success with "choice-plus" designs that incentivize sample members to respond by web (Biemer et al. 2017), but their use has been rather limited to date. More generally, the American Association for Public Opinion Research (AAPOR) (Olson et al. 2019) reviewed federal and non-federal surveys that recently implemented self-administered or mixed-mode data collection strategies. The AAPOR review concluded that there is, as yet, no definitive evidence as to the most effective approach to combining multiple modes of contact and administration within a data collection effort.

In addition, no existing research focuses specifically on the NCSES target population of the NTEWS—the STW subpopulation. Given the uncertainty of the performance of these contact strategies on this population, the initial NTEWS cycle will provide an opportunity to test several sequential and concurrent designs to determine the optimal contact strategy for future NTEWS cycles.

2.2. Treatment groups

Table C.1 identifies the six treatment groups that will make up the contact-strategies experiment and summarizes the rationale for each. Note that, throughout this appendix, italicized references to the *NSCG* and the *NHES* refer to the treatment groups in the NTEWS experiments that are based on these federal surveys, and do not refer to the separate surveys.

- The *NSCG* control group uses a similar contact strategy as the NSCG, in which contacts alternate between offering web only and offering both web and paper. This group is considered the control group because of its success in achieving a high response rate (68% in 2019) to the NSCG.
- The *NHES* treatment group uses a similar contact strategy to the NHES, where contacts begin with web only before transitioning to paper only. Both the *NSCG* and *NHES*

4

treatment groups use sequential mixed-mode designs that start by offering web only. The main difference between these treatment groups occurs after the initial switch from web. At this point, the *NHES* treatment group contacts no longer offer the option to respond by web, whereas, as noted above, the *NSCG* treatment group cycles between web-only and web-and-paper contacts. These two designs have each been applied successfully in their source surveys, but, to NCSES's knowledge, they have not been compared to each other directly. Because the NTEWS target population of interest to NCSES more closely resembles the NHES target population than the NSCG target population, it is worthwhile to test these two approaches against each other in the first cycle of the NTEWS.

- The *CHOICE* and *CHOICEPLUS* treatment groups assess whether a concurrent mixed-mode design (with web and paper offered simultaneously) may be preferable for the NTEWS (as compared to a sequential design). Though sequential mixed-mode designs are more commonly used in federal surveys, recent research suggests new promise for concurrent designs (e.g., Lesser et al. 2016; Bucks, Couper, & Fulford 2019), potentially due to the continued growth in Internet penetration (Olson et al., 2019). In particular, the *CHOICEPLUS* treatment group assesses whether, within concurrent mixed-mode designs, it is advantageous to offer an additional incentive contingent specifically on response by web. Preliminary results of an NHES:2019 experiment indicate that a similar design increased the response rate by about 7 percentage points relative to a sequential treatment (unpublished internal analysis). These preliminary results also suggest that the choice-plus treatment was more successful at encouraging response among hard-to-reach subgroups, implying that it reduced the risk of nonresponse bias (unpublished internal analysis). However, the NHES:2019 experiment did not include a "pure choice" treatment without the contingent incentive for web response, so the NTEWS experiment will help to clarify whether similar results could be obtained without the contingent incentive.

- The *PAP* and *CAT* treatment groups are included to assess whether some subpopulations may be most effectively reached with a paper-first or CATI-first design in future NTEWS cycles. NCSES does not anticipate using a paper-first or CATI-first design for the entire sample in future NTEWS cycles, because of the higher costs associated with these response modes as compared to web. However, paper-first and/or CATI-first designs could potentially be used for specific subgroups. Therefore, the inclusion of randomized paper-first and CATI-first treatments will allow for a determination of whether a tailored mode design—in which paper and/or telephone contact begins earlier for subgroups that are particularly unlikely to respond by web—should be adopted in future cycles to minimize nonresponse bias and/or reduce nonresponse follow-up costs. Separately, these treatments will also provide response data that can be used to assess the extent of mode effects in key NTEWS estimates—in other words, to assess whether responses to key items vary depending on the mode in which a person responds.

Table C.1. Overview of contact-strategy treatment groups

Treatment group	Description	Expected sample size	First offered response mode	Experiment goals
NSCG	Sequential mixed-mode, alternates between web-only and web-paper choice	9,720	Web	Evaluate success of NSCG-style contact sequence with the NTEWS target population. Serve as control group.
NHES	Sequential mixed-mode, web followed by paper	9,720	Web	Evaluate effect of the follow-up mode for the NTEWS target population: offering paper only (*NHES*) vs. choice of web and paper (*NSCG*); some literature (cf. Dillman et al. 2014) suggests that offering a choice of modes may backfire. Facilitate mode effects analysis by providing web-only data for the first 11 weeks of data collection
CHOICE	Concurrent mixed-mode, web and paper offered throughout	4,860	Web + paper	Compare a concurrent mixed-mode approach (with no additional incentivization for web response) to the NSCG and NHES-style sequential mixed-mode approaches; some recent literature suggests concurrent designs may perform as well as sequential ones (cf. Bucks, Couper, & Fulford 2019). Allow analysis of mode preference when both modes are offered but neither is incentivized.
CHOICEPLUS	Concurrent mixed-mode, web and paper offered throughout, promised $20 incentive for *web* response *by end of week 14*	4,860	Web + paper	Determine whether the proportion of web responses can be increased by incentivizing web response, and whether any resulting cost savings outweigh the cost of the added incentives; a prior general population study (cf. Biemer et al. 2017) and NHES experimentation have found this a promising approach.
PAP	Begins with paper-only, followed by web and phone	9,720	Paper	Identify any subgroups that may require a paper questionnaire to be offered earlier in data collection. Facilitate mode effects analysis by providing paper-only data for the first 11 weeks of data collection
CAT	Begins with outbound phone only, followed by web and paper	4,320	Outbound CATI	Identify any subgroups that may require CATI to begin earlier in data collection. Facilitate mode effects analysis by providing phone-only data for the first 11 weeks of data collection.

Table C.2 shows the planned contact sequence for each treatment group by week of data collection, for sample members with mailable addresses.

As noted above, all sample members will be offered the option to respond by web, paper, or CATI at various points in the data collection. In all treatment groups, inbound callers to the Census Bureau's telephone questionnaire assistance (TQA) hotline will also be able to respond by CATI.

The *CHOICE, CHOICEPLUS, PAP*, and *CAT* treatment groups will use the same contact strategy as the *NSCG* treatment group in the second half of data collection, with the transition beginning in week 12. This decision was made to reduce operational complexity in the later weeks and, in the case of *PAP*, to avoid an excessive number of paper responses, which have significantly higher keying costs and a greater risk of measurement error due to complex skips within the survey instrument. The *NHES* treatment group does not use the same contact strategy as the *NSCG* treatment group in the second half of the data collection because the *NHES* treatment group entails a purely sequential approach with only paper questionnaires offered in the second half of data collection.

The resulting design balances the goal of testing as many alternative approaches to data collection as possible (recognizing that the NTEWS differs in important ways from the existing NSCG and NHES collections[1]) with (1) the desire to minimize non-mode differences between the treatment groups at any given week and (2) the need to avoid excessive operational complexity.

---

[1] For example, although the NTEWS target population of interest to NCES is the same as in NHES (the general adult population), the STW target population of interest to NCSES is different from the NSCG and NHES target populations.

Table C.2. Contact attempts by week, by contact-strategy treatment group

Week	NSCG	NHES	CHOICE	CHOICEPLUS	PAP	CAT
1	Web invite letter	Web invite letter	Choice invite letter and questionnaire	Choice-plus invite letter[1] and questionnaire	Paper-only invite letter and questionnaire	CATI advance letter
						CATI begins[2]
2	Web perforated mailer	Web perforated mailer	Choice perforated mailer	Choice-plus perforated mailer	Paper perforated mailer	CATI perforated mailer
5	ACS-style web letter[3]	ACS-style web letter	ACS-style choice letter and questionnaire	ACS-style choice-plus letter and questionnaire	ACS-style paper-only letter and questionnaire	CATI continues
6	Generic postcard	Generic postcard	Generic postcard	Choice-plus postcard	Generic postcard	CATI postcard
8	Choice letter and questionnaire	Web letter	Choice letter and questionnaire	Choice-plus letter and questionnaire	Paper-only letter and questionnaire	CATI continues through week 11
12	Perforated web reminder	Paper-only letter and questionnaire[4]	Perforated choice reminder	Perforated choice-plus reminder	Web invite letter (transition from paper)[5]	Web invite letter (transition from CATI)[5]
14	CATI NRFU begins	CATI NRFU begins	CATI NRFU begins	CATI NRFU begins End of week: deadline for contingent incentive	CATI NRFU begins	
16	Web letter	Paper-only letter and questionnaire	Web letter	Web letter	Web letter	Web letter
20	ACS-style choice letter and questionnaire	ACS-style paper-only letter and questionnaire	ACS-style choice letter and questionnaire	ACS-style choice letter and questionnaire	ACS-style choice letter and questionnaire	ACS-style choice letter and questionnaire
23	Web letter, FedEx envelope	Paper-only letter and questionnaire, FedEx envelope	Web letter, FedEx envelope	Web letter, FedEx envelope	Web letter, FedEx envelope	Web letter, FedEx envelope
25	CATI NRFU ends	CATI NRFU ends	CATI NRFU ends	CATI NRFU ends	CATI NRFU ends	
27	End of data collection					

[1]The choice-plus invite letters, reminder letters, and perforated reminders would mention the $20 contingent gift card incentive for completing by web by the end of week 14.

[2]CATI outreach would occur continually in weeks 1 through 11.

[3]ACS-style letters are printed on a smaller, thicker piece of paper than is used for the other letters.

[4]NHES switches to paper-only in week 12 because, unlike NSCG and the other treatment groups, it uses a purely sequential mixed-mode approach.

[5]PAP and CAT require a full web invite letter (rather than a perforated reminder) in week 12 because this is the first week in which web is offered to these groups, so a full letter is required to introduce the web option.

NOTE: NRFU refers to nonresponse follow-up. CATI refers to computer-assisted telephone interviewing.

COLOR KEY FOR TABLE C.2:	Contact offers web response only	Contact offers paper response only	Contact offers both paper + web response	Contact offers phone response only	No direct way to respond

Table C.3 summarizes the number of contacts offering each mode of response in each treatment group. Counts are shown separately for weeks 1 – 11 and 12 – 27 because several treatment groups use the same contacts as the *NSCG* treatment group beginning in week 12.

Table C.3. Number of contacts offering each mode of response, by treatment group

Response mode(s) offered	NSCG	NHES	CHOICE[1]	CHOICEPLUS[1]	PAP	CAT
In weeks 1 - 11:						
Web-only	3	4	1	1	0	0
Paper-only	0	0	0	0	3	0
Web + paper	1	0	3	3	0	0
Weeks of outbound CATI	0	0	0	0	0	11
Passive reminders[2]	1	1	1	1	2	3
In weeks 12 - 27:						
Web-only	3	0	3	3	3	3
Paper-only	0	4	0	0	0	0
Web + paper	1	0	1	1	1	1
Weeks of outbound CATI	11	11	11	11	11	0
Passive reminders[2]	0	0	0	0	0	0
Total:						
Web-only	6	4	4	4	3	3
Paper-only	0	4	0	0	3	0
Web + paper	2	0	4	4	1	1
Weeks of outbound CATI	11	11	11	11	11	11
Passive reminders[2]	1	1	1	1	2	3

[1]In *CHOICE* and *CHOICEPLUS*, the web-only contacts in weeks 1 – 11 refer to the perforated reminder letters, which cannot include paper questionnaires.

[2]Passive reminders refer to contacting sample members through reminder postcards and advance letters, which do not offer a direct way to respond.

2.3. Additional operational details

As noted above, the full contact sequences shown in table C.2 will apply to sample members with mailable addresses. Sample members without mailable addresses will be routed to an "early CATI" operation if a phone number is available. For such sample members, the mailings will be paused and outbound CATI will be conducted until and unless a usable address is located. If a usable address is located, the sample member will enter back into its assigned mailing sequence at that point. For example, if an address was found in week 7, the sample member would receive the mailings associated with its assigned treatment group for week 8 and all following weeks. Thus, these individuals will no longer be eligible for the contact-strategy experiment or analysis.

All NTEWS contact and data collection materials—including the cover letters, web instrument, paper questionnaire, and CATI script—will be developed in both English and Spanish. Select sample members will receive bilingual mailings that include materials in both languages. The criteria used to target bilingual mailings in the NTEWS are based on those used to identify

households requiring "Spanish assistance" in the 2010 decennial census (Rothhaas et al. 2011). Specifically, bilingual mailings will be used for sample members who meet both of the following criteria, based on the 2018 ACS response data available on the frame:

- The person speaks Spanish at home.
- The person speaks English "well", "not well", or "not at all" (rather than "very well").

Thus, the only Spanish-speakers who will not receive bilingual mailings are those who speak English "very well". Sample members meeting the above criteria will receive the same contact sequence as others in their treatment group (table C.2) but will receive a bilingual version of each mailing package that includes both English- and Spanish-language materials.

In *CHOICEPLUS*, the $20 contingent debit card incentive will be paid if both of the following criteria are met:

- The respondent completes the survey *by web*.
- The response by web is received no later than the end of week 14 of data collection.

All else equal, a response by web is preferable to a response by paper because the web instrument can automatically route the respondent through skip patterns; thus, it reduces the amount of required data editing, a relevant consideration given the complex skip patterns on the NTEWS. Response by web also avoids the data entry costs associated with paper response and interviewer costs associated with CATI response. For these reasons, the contingent incentive is structured to encourage response specifically by web rather than by CATI or paper. The week 14 deadline is included to encourage response prior to the beginning of CATI nonresponse follow-up. Due to the expense of outbound CATI, it is best to incentivize sample members to respond before they reach this stage.

The $20 value for the contingent incentive is recommended based on early results of an NHES:2019 experiment with a similar choice-plus design. NHES:2019 tested both a $10 and a $20 contingent incentive within the choice-plus design. Based on preliminary results from the screener phase of the NHES, the $20 contingent incentive achieved a higher percentage responding by web (53.4 percent vs. 48.3 percent), relative to $10 (unpublished internal analysis). Given the relatively limited testing that has been conducted to date of contingent incentives specifically for response by a given mode, the NTEWS experiment will help to determine whether the NHES results apply specifically to the STW subpopulation that is of greatest interest to NCSES.

As discussed in section 3 of this appendix, the initial cycle of the NTEWS will include a separate experiment with *noncontingent* debit card incentives, which will primarily be sent with the week 1 mailing. The contingent incentive in *CHOICEPLUS* will be paid on top of any noncontingent incentive that the sample member receives in week 1. In other words, regardless of whether a prepaid debit card was previously received, respondents from *CHOICEPLUS* who qualify for the contingent incentive (i.e., who respond by web before the deadline) will be mailed a new debit card with the $20 contingent incentive. The card sent in week 1 will retain its original value. All

debit cards will expire approximately six months after the data collection starts; this will be made clear on the cards.

2.4. Additional research questions

The three main research questions listed in section 1 all apply to the contact-strategies experiment. This experiment will ultimately help to determine (1) which contact strategy should be used as the default in future NTEWS cycles based on nonresponse bias and response rates, and (2) Are there subgroups of the NTEWS sample for which it is beneficial to use more expensive contact strategies (e.g., *CHOICEPLUS*, *PAP*, or *CAT*) to improve data quality (sample representativeness).

This experiment also aims to answer a fourth research question: is there evidence of mode effects in key NTEWS items? Formally, mode effects occur when measurement error in an item differs depending on the mode in which the item is presented—that is, when *the same respondent* would give a different answer to *the same item* when the item is presented in a different mode (Vannieuwenhuyze & Loosveldt 2012). In mixed-mode studies such as the NTEWS, mode effects can counteract the benefits of offering multiple administration modes and therefore are an important component of total survey error. However, mode effects are difficult to evaluate because they can be confounded by selection effects, which are driven by differences in the *underlying characteristics* of respondents who choose to respond by different modes (Vannieuwenhuyze, Loosveldt, & Molenberghs 2010). For example, if persons without a high school diploma are simultaneously more likely to respond by CATI and less likely to report a certification, differences in the prevalence of certifications between web and CATI respondents could reflect not mode effects but rather the fact that these two modes capture different populations. However, the *NHES*, *PAP*, and *CAT* treatments in the NTEWS contact-strategies experiment will provide 11 weeks' worth of data from respondents who were offered only web, paper, and CATI modes of response (respectively). In comparisons of key estimates between the first 11 weeks of responses from these three treatments, selection effects are likely to be reduced[2]; therefore, these treatments will allow for a more accurate evaluation of mode effects than is typically possible in mixed-mode studies. This mode effects analysis will help determine whether any key NTEWS survey items require further revision and testing to mitigate potential mode effects. The mode effects analysis will also allow a more complete evaluation of the costs and benefits of continuing to include all three contact modes (web, paper, and CATI) in future cycles of the NTEWS.

3. Noncontingent-incentive experiment

As previously mentioned, because this is the first NTEWS cycle, it is necessary to gather baseline data to allow for evidence-based decisions on whether monetary incentives should be used to improve data quality and minimize costs. Based on the extant literature, as well as prior NSCG and NHES experimentation, NCSES hypothesizes that future cycles of the NTEWS will

---

[2] Some selection effects may still be present because, for example, sample members in the *PAP* treatment who strongly prefer to respond by web may not respond at all during the period when paper is the only offered response mode; and because, in all treatment groups, sample members will be able to respond over the phone at any point in the data collection, if they choose.

need to use noncontingent incentives for at least some portion of the sample. This is particularly true given the analytic interest in the subbaccalaureate STW. NCES's experience with the NHES (McPhee et al. 2018), along with the broader literature on self-administered and mixed-mode surveys (Olson et al. 2019), suggests that sample members with less formal education are less likely to respond to household surveys. It is possible, therefore, that incentives will be needed to ensure sufficient representation of this critical subgroup among NTEWS respondents. More generally, the NTEWS oversampling of sample members with less than a Bachelor's degree raises the possibility that a larger proportion of the NTEWS sample may need to be incentivized, relative to the NSCG, though this cannot be known with certainty until baseline data are collected.

Accordingly, the initial NTEWS cycle will include a noncontingent-incentive experiment to test multiple potential incentive levels against a no-incentive control. Because this is an experiment, the incentive structure proposed for the initial cycle of the NTEWS is not the same structure that will be used for future cycles. Rather, the purpose of the noncontingent-incentive experiment is to collect the baseline data needed to design cost-effective incentive structures for future cycles.

3.1. Background on use of incentives in federal surveys

As response rates continue to decline (cf. Meyer et al. 2015), the use of incentives has become increasingly common in non-mandatory federal household surveys. For example, To (2015) lists nine examples of federal surveys that have incorporated or experimented with incentives in recent years. Similarly, the AAPOR report on transitions to self-administered and mixed-mode surveys cites examples of approximately 15 surveys, about two-thirds of which are federally sponsored, that have used or experimented with incentive levels (Olson et al. 2019, table 6.4). Finally, both surveys that can be considered the "predecessors" to the NTEWS—the NSCG and the NHES—use monetary incentives in some form.

The current practice of the NSCG is to provide a $30 prepaid debit card with the first mailing to approximately 20 percent of new sample members. The new sample members receiving the incentive are those flagged as "highly influential" based on the combination of a large base weight and a low predicted response propensity score, implying that they are likely to cause nonresponse bias if they do not respond. Because the NSCG uses a rotating panel design, the $30 debit card incentive is also provided to returning sample members who were incentivized in a prior NSCG cycle. All other sample members receive no incentive. This incentive structure is informed by experiments incorporated into the 2010 and 2013 NSCG cycles (Zotti 2014; Thornton 2014).

The NHES is a two-phase survey in which noncontingent cash incentives are used at both the first (household screening) and second (topical module) phases. The ATES topical module was included in two NHES cycles: a 2014 feasibility study (Jackson, McQuiggan, & Megra 2016) and a 2016 full-scale administration (Jackson, McPhee, & Lavrakas 2019). Both of these NHES cycles included experiments with prepaid cash incentives at the household screening phase. Based on these experiments, the current practice in the NHES is to include a $5 cash incentive with the initial screener phase mailing for all sample members. The NHES also includes a

topical-phase cash incentive that is sent with the first topical mailing; in NHES:2019, most households were sent $5, but those that responded to the screener after the third or fourth screener mailings were sent $15. These topical-phase incentive values were based on experiments incorporated into a 2011 field test and a 2012 full-scale administration of the NHES, neither of which included the ATES (Han, Montaquila, & Brick 2013; McPhee et al. 2015).

In considering potential incentive strategies for the NTEWS—which, like the NSCG and unlike the NHES, will use debit card incentives—it is important to consider the findings of prior research (discussed in greater detail below) suggesting that the effects of debit card and cash incentives are unlikely to be equivalent for a given dollar value. In particular, to achieve similar effect as a given cash incentive, a debit card incentive of twice or greater value might be required.

3.2. Prior findings on effects of incentives on response rates

The use of payments or gifts to incentivize response is one of the most heavily studied topics in the survey research literature. Evaluations of alternative incentive types and levels have been synthesized into multiple literature reviews, most recently by Singer and Ye (2013); and meta-analyses, most recently by Mercer et al. (2015). These syntheses have demonstrated the value of incentives for both increasing response rates and (in some cases) reducing nonresponse bias.

Several key findings about the effects of incentives on response rates are well-established. First, noncontingent (i.e., prepaid, not dependent on response) monetary incentives are typically associated with higher response rates, relative to no incentive (Mercer et al. 2015). Second, for a given dollar value, the effects of contingent (i.e., promised, dependent on response) incentives tend to be smaller and less consistent than those of noncontingent incentives, particularly in mail surveys (Mercer et al. 2015).[3] Third, nonmonetary incentives (e.g., lotteries, charitable donations, pens) generally are less effective than monetary incentives (Church 1993). Fourth, the relationship between the value of the incentive and the response rate is subject to diminishing returns. While higher dollar values are generally associated with higher response rates, the effect of each additional dollar is typically less than the effect of the prior dollar (Trussell & Lavrakas 2004).

The results of prior NHES and NSCG experimentation are consistent with these general findings. The 2014 NHES experiment established that a $5 noncontingent cash incentive in the first screener mailing significantly increased the response rate relative to no incentive (69 percent vs. 62 percent, respectively), while a nonmonetary incentive (a refrigerator magnet) had no meaningful effect (Jackson, McQuiggan, & Megra 2016). The 2016 NHES experiment further established that a $5 cash incentive outperformed a $2 cash incentive (65 percent vs. 62 percent, respectively), but that a $10 cash incentive had little additional effect (relative to $5) on hard-to-reach households (Jackson, McPhee, & Lavrakas 2019). The 2010 NSCG experiment found that a $30 noncontingent debit card, provided late in data collection to hard-to-reach sample members

---

[3] A notable exception to this general finding is that Biemer et al. (2017) found a contingent incentive offered specifically for web responses to be an effective way of encouraging respondents to complete via the web when both web and paper response options are offered. This finding is a motivation for the inclusion of the *CHOICEPLUS* treatment within the contact-strategies experiment, discussed in section 2.

who had not yet responded, yielded a 30 percent response rate, compared to 24 percent for a $20 incentive and 6 percent for no incentive.[4] The 2013 NSCG experiment further found that the $30 debit card (provided to "highly influential" sample members as described previously) led to an even greater increase in the response rate (relative to no incentive) when provided with the first mailing rather than (as in the 2010 experiment) late in data collection.

Both the NHES and NSCG have also experimented with the timing of noncontingent monetary incentives, a topic that is less extensively covered in extant literature. An experiment in the 2011 NHES field test (prior to the incorporation of the ATES topical module) found that the second-phase response rate was about 10 percentage points higher when the noncontingent incentive was sent with the initial second-phase mailing than when it was sent with a nonresponse follow-up mailing (McPhee & Hastedt 2012). The 2013 NSCG experiment found that, relative both to sending no incentive and to sending a $30 incentive later in data collection, the $30 incentive for highly influential sample members led to the highest response rate when it was included with the first mailing. Furthermore, sending the incentive with the first mailing reduced data collection costs per case (relative to both the no-incentive control group and most of the treatments that sent the incentive later in data collection) because the higher response rate to the first mailing reduced nonresponse follow-up effort.

Importantly, neither the NSCG (which uses debit card incentives, the same format planned for the NTEWS) nor the NHES (which uses cash) have directly tested the relative effects of debit card vs. cash incentives on the response rate. In general, experimental testing specifically of debit card incentives is rare: in their meta-analysis, Mercer et al. (2015) located only two experimental conditions that used a debit card, compared to 118 that used cash, so it was not possible to compare effect sizes between the alternative formats.

To NCSES's knowledge, only one prior study (Bailey, Lavrakas, & Bennett 2007) has directly compared cash and debit card incentives in a randomized, controlled design. The results of that study suggest that, *for a given dollar amount, the effects of cash and debit cards are not equivalent*. Relative to a $5 cash control, a $5 debit card consistently obtained lower response rates; while a $10 debit card obtained either lower or comparable response rates to the $5 cash control, depending on the subgroup being evaluated (the experiment did not include a $10 cash control). Therefore, the relationship between the "face value" of the incentive and the response rate is unlikely to be the same for the two formats; and, in particular, the face value at which diminishing returns set in could be higher with debit cards than with cash. The results of the NHES and NSCG experimentation, though not directly comparable due to the surveys' different designs, are consistent with this hypothesis.

3.3. Prior findings on effects of incentives on nonresponse bias

As previously mentioned, nonresponse bias occurs when the composition of the *responding* sample differs from that of the *target population* as a result of systematic variation in response rates between subgroups of the sample. Thus, a given incentive structure may affect nonresponse

---

[4] These results are reported in section A of the supporting statement to the Information Collection Request for the 2019 NSCG.

bias to the extent that it changes the composition of the responding sample; or, equivalently, to the extent that its effects on the response rate vary across subgroups. This can be true even when (as will be the case with the NTEWS) an attempt is made to compensate for nonresponse through post-collection weighting adjustments. Simulation research (Sarndal & Lundquist 2014) suggests that post-collection weighting adjustments cannot fully compensate for nonresponse bias. In other words, an incentive structure that helps to achieve a pool of respondents that more closely resembles the initial sample *prior* to adjustment can help to reduce the amount of bias that remains *after* adjustment.

In discussing the potential effects of incentives on nonresponse bias, it is useful to distinguish between *uniform* incentive structures, in which all sample members receive the same incentive; and *tailored* incentive structures, in which the use and/or value of the incentive is purposefully varied between observable subgroups.[5]

The impact of a *uniform* incentive structure on nonresponse bias depends on differences in "incentive sensitivity" between subgroups. Suppose a sample can be divided into two subgroups: an "underrepresented" subgroup with lower-than-average response rates in the absence of an incentive, and an "overrepresented" subgroup with higher-than-average response rates. The introduction of a uniform incentive would mitigate nonresponse bias if the underrepresented subgroup were more sensitive to the incentive; that is, if the incentive increased the response rate from that group *more than* it increased the response rate from the overrepresented subgroup. However, if the reverse were true, the incentive would exacerbate nonresponse bias. Finally, if the response rate increase were largely uniform across subgroups, the incentive might have no impact on nonresponse bias. Consistent with this theoretical ambiguity, Singer and Ye (2013) identify examples of all three results in the empirical literature.

In contrast, a *tailored* incentive structure is, *a priori*, more likely to reduce bias because it gives the researcher some control over which subgroups receive the incentive and the resulting response rate increase. Continuing the above example, if the incentive were provided only to the underrepresented subgroup, this incentive structure would reduce bias to the extent that this subgroup was sensitive to the incentive. This would be true even if the hypothetical effect of the incentive was the same or larger in the overrepresented subgroup, since this subgroup would not be incentivized in the tailored design.

NHES experiments have generally found a reduction in bias when using both uniform and tailored incentive structures. As described above, the NHES uses a uniform incentive structure at the first (screener) phase and a tailored structure at the second (topical) phase. At the topical phase, the 2011 field test found that a uniform $5 incentive reduced nonresponse bias by improving the representation of low-socioeconomic-status subgroups (Han, Montaquila, & Brick 2013). The 2012 experiment then tested a tailored structure in which a higher ($15) topical incentive was used for late screener respondents (hypothesized to consist of harder-to-reach households) while $5 was used for early screener respondents (i.e., easier-to-reach households).

---

[5] For the purpose of this discussion, an experimental design in which the use and/or value of an incentives varies only between randomly assigned treatment groups should be understood as a uniform incentive structure, because, within each randomly assigned treatment group, all sample members receive the same incentive.

The tailored structure improved the representation of late screener respondents relative to a uniform $5 structure (McPhee et al. 2015) and therefore was adopted in subsequent NHES cycles. Finally, at the screener phase, the 2014 NHES feasibility study found that, relative to no screener incentive, the uniform $5 screener incentive improved the representation of households without an available phone number, those with a younger head of household, and households with children (Jackson, McQuiggan, & Megra 2016), all subgroups that tend to be underrepresented among NHES screener respondents (McPhee et al. 2018).

Also as described above, the NSCG uses a tailored incentive structure. The 2010 NSCG incentive experiment found that the use of an incentive for hard-to-reach sample members succeeded at obtaining additional responses from these individuals, reducing the potential for nonresponse bias.[6] Similarly, the 2013 NSCG incentive experiment found that the $30 incentive for highly influential sample members significantly reduced differences between respondents and the initial sample, an effect that was strongest when the incentive was sent with the first mailing.

Several NCES longitudinal surveys have also found potential reductions in nonresponse bias using a tailored incentive structure. The 2012/14 Beginning Postsecondary Students Longitudinal Study used a responsive design in which an incentive increase of up to $45 was offered to underrepresented subgroups at various phases of the collection. This design appeared to reduce nonresponse bias by the end of the collection (Hill et al. 2016). A similar design was used in the High School Longitudinal Study of 2009 with similar results (Duprey et al. 2018).

3.4. Need for experimentation in the initial cycle of the NTEWS

Broadly, then, because the NTEWS is the first cycle of the NTEWS, experimentation is needed to enable NCSES to develop an incentive structure for future cycles that will minimize nonresponse bias while ensuring cost-effectiveness. While the results of the NSCG and NHES experimentation are informative for generating hypotheses, the extent to which they generalize to the NTEWS is unclear. The NTEWS will use the same incentive format (debit cards) and be drawn from the same frame (prior ACS respondents) as the NSCG; however, the NTEWS sample will cover the general population with oversamples of the subbaccalaureate STW subpopulation, whereas the NSCG sample includes only college graduates. Conversely, the NTEWS will cover a similar population as the inactive ATES module of the NHES; however, the NHES uses a different sampling frame (a commercial address-based frame) and a different incentive format (cash). Finally, and possibly most importantly, some NTEWS key survey estimates (e.g., proportion of workers in skilled technical occupations) differ from both the ATES and the NSCG. NTEWS-specific experimentation is therefore required to determine the optimal incentive structure given its unique combination of target population, sampling frame, incentive format, and key estimates.

Assuming the need for incentives in future NTEWS cycles, NCSES expects to move towards an NSCG-style tailored incentive structure designed to reduce bias in key survey estimates. However, key baseline parameters necessary to design such a structure for the NTEWS are as yet

---

[6] These results are reported in section A of the supporting statement to the Information Collection Request for the 2019 NSCG.

unknown. For example, what response rates are obtained specifically from low-response-propensity sample members at various incentive levels? How does this compare to high-response-propensity sample members? Can specific sample members for which a higher incentive is effective (or ineffective) be reliably identified using the ACS frame data? In the absence of NTEWS-specific experimentation, neither the NHES nor the NSCG findings can be used to reliably address these questions. For example, a response propensity model estimated on the NSCG college-educated sample is not applicable to the NTEWS, since educational attainment itself is likely to be an important predictor of response behavior.

Finally, some key parameters that will influence the cost of a given incentive strategy remain unknown. Unlike with cash, the survey sponsor recoups a significant portion of the value of mailed debit card incentives because not all recipients use their debit cards. For example, in the 2017 NSCG, 46 percent of respondents and 97 percent of nonrespondents who received an incentive did not use their debit cards prior to expiration. These usage rates could plausibly vary with the value of the debit card—for example, sample members may be more likely to use a more valuable incentive. They could also vary across target populations, implying that the NSCG experience may not generalize to the NTEWS. Experimentation with multiple incentive levels in the initial cycle of the NTEWS will provide the data necessary to accurately project costs and to develop cost effective incentive structures for future NTEWS cycles.

For all of these reasons, even though NCSES expects to move towards a *tailored* incentive structure for the NTEWS, randomized experimentation with multiple *uniform* incentive levels is a valuable and informative first step. The noncontingent-incentive experiment, with the treatment groups described below, is designed to maximize the utility of the initial NTEWS for data-driven decision-making about future incentive structures.

3.5. Potential concerns with the use of incentives in the NTEWS

One potential risk raised by the use of incentives in the NTEWS is that of conditioning. The NTEWS will use a rotating panel design, similar to the NSCG, in which each sample will include both an "returning cohort" (consisting of respondents to one or more prior cycles) and a "new cohort" (a fresh sample drawn from the ACS frame). With such a design, there is a risk that respondents who received an incentive in one cycle might come to expect an incentive and therefore might be less likely to respond to later cycles unless also offered an incentive in those cycles. The literature on such conditioning effects is mixed. In their review, Singer and Ye (2013) identified several panel studies that evaluated conditioning effects and found no evidence that respondents who received incentives at early waves were less likely to respond to later waves. However, an incentive conditioning experiment incorporated into the 2013 NSCG did find evidence that returning sample members who had previously been incentivized were more likely to respond if again provided an incentive in later cycles (Thornton 2014). Consequently, the current practice of the NSCG is to incentivize returning sample members who had previously received an incentive. In future NTEWS cycles, NCSES will monitor response rates among returning sample members who received an incentive as part of the initial noncontingent-incentive experiment. The incentive strategy for returning sample members in future cycles will be informed in part by whether any evidence of conditioning is observed.

A second potential risk is that some sample members may react negatively to the use of incentives—for example, some sample members may consider them to be a waste of taxpayer money. However, in both sponsoring agencies' experience, complaints by sample members related to the use of incentives are rare. In the 2019 NHES, the Census Bureau received fewer than 30 complaints out of a sample of 205,000 addresses. In the 2019 NSCG, call-center staff reported no complaints about incentives; all inbound communication related to the incentives consisted of respondents seeking replacement debit cards or information about the proper use of the cards.

Finally, the fact that the sampling frame for the NTEWS consists of prior ACS respondents raises the question of whether incentives may be redundant, since all sample members will have previously demonstrated a willingness to respond to a federal survey. However, NCSES believes that experimentation with incentives is still warranted, for several reasons. First, the ACS is a mandatory survey for which individuals can be fined for not participating, whereas the NTEWS is a voluntary survey and is explicitly identified as such in respondent contact materials. Second, the ACS, unlike the NTEWS, includes an in-person follow-up phase; therefore, the NTEWS sample will include some ACS respondents who did not respond by any of the three modes that will be offered in the NTEWS (web, paper, or CATI). Third, the NTEWS (like the NSCG) will sample a single person from each household; and, in households with more than one person, the selected person may not be the same person who responded to the ACS on behalf of the household. Finally, the NSCG is also drawn from the ACS frame, and prior experimentation has found incentivization to reduce bias in the NSCG. These factors all provide reason to expect that incentives may be both needed and useful to reduce bias in the NTEWS sample. However, by including a no-incentive treatment group within the noncontingent-incentive experiment (as discussed below), NCSES will be able to determine the baseline response rate that could be achieved without an incentive, and thereby identify any subgroups for which incentivization may be unnecessary in future cycles.

3.6. Treatment groups

The noncontingent-incentive experiment will test three alternative dollar values for a prepaid debit card, as well as a no-incentive control. The NTEWS production sample will be randomly split into four equal-size treatment groups that will determine the value of the debit card sent with the first mailing: $0 (i.e., no debit card with the first mailing), $10, $20, and $30. The $0 group will be further split, with half receiving a $30 debit card with the *final* mailing (as a "late-stage incentive") and the other half receiving no incentive at any mailing. Thus, the NTEWS noncontingent-incentive experiment will have five treatment groups, as shown in table C.4.

Table C.4.   Noncontingent-incentive treatment groups

| Treatment group | Description | Expected sample size | Value of debit card sent in… | |
			First mailing	Final mailing
INC0	$0 noncontingent incentive	5,400	N/A	N/A
INC0L	$30 late-stage noncontingent incentive	5,400	N/A	$30
INC10	$10 noncontingent incentive	10,800	$10	N/A
INC20	$20 noncontingent incentive	10,800	$20	N/A
INC30	$30 noncontingent incentive	10,800	$30	N/A

NOTE: "$0 noncontingent incentive" implies that no debit card is sent.

Treatment groups receiving $0 with the first mailing (*INC0* and *INC0L*) are included to collect baseline data on the performance of the NTEWS in the absence of a prepaid incentive. If there are particular subgroups from which an acceptable response rate can be attained without an incentive, the incentive for those subgroups could potentially be omitted in future cycles. The inclusion of a randomly assigned $0 treatment group will allow any such subgroups to be identified and thus maximize the utility of the data for developing efficient tailored incentive structures in future cycles.

At the same time, while a uniform $0 treatment will provide valuable methodological data, NCSES recognizes that its use could reduce the representativeness of the responding sample and thus the validity of national estimates obtained from this administration of the NTEWS. This potential reduction in representativeness and the potential for biased estimates is the primary rationale for further dividing the $0 treatment group such that half (*INC0L*) receive a late-stage incentive at the final mailing. The late-stage incentive is intended to function, in effect, as a "safety mechanism" to mitigate any impact of the $0 treatment group on representativeness. A secondary benefit is that the use of an incentive toward the end of data collection could be considered as an adaptive design intervention in future NTEWS cycles; thus, data collected by this treatment will be useful in developing adaptive designs. The value of the late-stage incentive will be relatively high ($30) in recognition of the fact that, by definition, sample members who have not responded by the final mailing are likely to be reluctant responders, implying that a relatively valuable incentive will be needed to substantially improve response to the final mailing.

The non-$0 incentive treatments (*INC10*, *INC20*, and *INC30*) will allow the testing of multiple incentive levels in a stepwise fashion. The rationale for this stepwise approach is the recognition that the experiences of the NSCG and NHES are not entirely comparable to the NTEWS, as discussed above. Testing multiple non-$0 incentive levels will allow NCSES to quantify the effect of incentive value on the NTEWS response rate both for the entire population and for critical subgroups. This testing will also provide data that could be used to develop stepwise tailored designs for future cycles. For example, rather than a binary approach of using $0 for some subgroups and $20 for the rest of the sample, a stepwise approach of using $0 for some

subgroups, $10 for others, and $20 for the rest of the sample could turn out to be optimal. Finally, by allowing for an evaluation of the relationship between the incentive value and debit card usage rates, this approach will allow for a more complete analysis of cost-benefit tradeoffs from incentivization.

The $10, $20, and $30 incentive levels were selected based on a review of relevant literature and on previous NSCG and NHES experiences with noncontingent incentives. In general, based on the findings by Bailey, Lavrakas, and Bennett (2007), the non-$0 debit card incentives that will be tested for the NTEWS are generally larger than the cash incentives used for the NHES, and more comparable to the debit card incentives used for the NSCG. In particular, the $10 treatment (*INC10*) is included on the hypothesis that it would be approximately equivalent to the $5 cash incentives used at the first screener mailing in the NHES. The $30 treatment (*INC30*) is included because it is the default incentive level in the NSCG, and the 2010 NSCG incentive experiment found that a $30 debit card continued to offer an incremental benefit over $20. The $20 treatment (*INC20*) is included to determine whether this finding from the NSCG is replicated and whether this intermediate incentive level is sufficient for the (more general) NTEWS target population.

3.7. Additional operational details

All noncontingent debit card incentives will be included in the same envelope as the letter and/or questionnaire sent in the specified week. The value of the enclosed incentive will be stated in the letter with which it is sent and will also be printed in the "Name" field of the card itself. Subsequent mailings will *not* mention the previously provided noncontingent incentive.

All debit cards, regardless of when they are sent, will expire approximately six months after the scheduled data collection start date. The expiration date will be printed on the cards.

3.8. Research questions

The three main research questions listed in section 1 all apply to the noncontingent-incentive experiment. The analysis of the noncontingent-incentive experiment will focus on how incentives can best be employed in a tailored manner to minimize nonresponse bias at an acceptable cost to the government. NCSES will prioritize analyses that will inform the development of a cost-effective tailored incentive design. Of particular interest, therefore, will be the effectiveness of various incentive levels within subgroups that are of substantive interest (e.g., the subbaccalaureate STW), as well as among sample members with lower-than-average response propensities. NCSES anticipates that these are the types of subgroups between which the usage and value of incentives could vary in future NTEWS cycles.

Altogether, this experiment will help to determine (1) which incentive level should be used as the default in future NTEWS cycles and (2) whether there are subgroups of the NTEWS sample for which it is necessary to use a higher incentive.

4. Treatment combinations

Table C.5 shows the expected sample size for each combination of contact-strategy and noncontingent-incentive treatments. For treatment combinations that involve the use of monetary

incentives for some or all sample members, the table also shows the expected number of incentivized sample members.

Table C.5.   Expected number of sampled individuals, and expected number receiving any incentive, by experimental treatment group

Noncontingent-incentive treatment	Contact-strategy treatment					
	NSCG-style	NHES-style	Choice	Choice-plus[1]	Paper-first	CATI-first
Expected number of sampled individuals						
No noncontingent incentive	1,215	1,215	608	608	1,215	540
$30 with week 23 mailing	1,215	1,215	608	608	1,215	540
$10 with week 1 mailing	2,430	2,430	1,215	1,215	2,430	1,080
$20 with week 1 mailing	2,430	2,430	1,215	1,215	2,430	1,080
$30 with week 1 mailing	2,430	2,430	1,215	1,215	2,430	1,080
Expected number receiving any incentive (noncontingent and/or contingent)						
No noncontingent incentive	0	0	0	213	0	0
$30 with week 23 mailing[2]	122	122	61	274	122	54
$10 with week 1 mailing	2,430	2,430	1,215	1,215	2,430	1,080
$20 with week 1 mailing	2,430	2,430	1,215	1,215	2,430	1,080
$30 with week 1 mailing	2,430	2,430	1,215	1,215	2,430	1,080

[1]The choice-plus treatment within the contact-strategies experiment includes a $20 contingent (i.e., promised) incentive that will be paid to all respondents who complete the survey by web by the end of week 14 of data collection. This contingent incentive would be paid on top of any noncontingent incentive that the respondent received as part of the noncontingent-incentive experiment. The numbers in this column assume that 35 percent of sample members are eligible for the choice-plus incentive (i.e., that 35 percent complete the survey by web by the end of week 14).

[2]The numbers in this row assume that 10 percent of sample members are eligible for the week 23 noncontingent incentive (i.e., have not responded by week 23).

NOTE: Due to rounding, details may not sum to totals. Detailed information about each experimental treatment is provided earlier in this appendix. The actual number of sampled individuals in a given treatment combination may vary slightly from expectations due to rounding. The actual number receiving any incentive may vary from expectations depending on actual response patterns to the NTEWS. Shaded cells are those in which sample members would receive both a noncontingent incentive with the week 1 mailing and (if they respond by web by the end of week 14) a $20 contingent incentive.

5. Seeded sample

5.1. Background

The 1,000-person seeded sample is included to evaluate the accuracy of the postsecondary certificate attainment item on the NTEWS questionnaire (item 65 in Appendix K). Data from the seeded sample are for NCSES and NCES research purposes only. Therefore, respondents from the seeded sample will not be weighted or included in publicly released NTEWS data products. Because seeded sample respondents will be known certificate holders, the proportion who report on the NTEWS that they do *not* hold a postsecondary certificate will provide an estimate of the item's underreporting rate: the percentage of respondents who report that they do not hold a certificate when they in fact do.

During the initial development of the now-discontinued ATES module of the NHES, cognitive research found that some respondents had trouble accurately reporting whether they held a postsecondary certificate (Bielick et al. 2013, American Institutes for Research 2013). To allow for more systematic evaluation of measurement error in questionnaire items that collected postsecondary certificate data, seeded sample evaluations were conducted in 2010 (as part of a standalone ATES pilot not incorporated into the NHES), 2014 (as part of an NHES feasibility study), and 2016 (as part of a full-scale NHES cycle).[7] All three evaluations found consistently high postsecondary certificate underreporting rates (Bielick et al. 2013; Jackson, McQuiggan, & Megra 2016; Megra & Cronen 2017), leading to continued revisions and cognitive testing of the postsecondary certificate questionnaire items.

Based on this prior research, and recent rounds of cognitive testing in preparation for the initial NTEWS cycle, the postsecondary certificate items on the NTEWS instrument incorporate two major revisions designed to reduce underreporting: the main postsecondary certificate question wording has been simplified, and the postsecondary certificate items are now positioned immediately after the educational attainment section (whereas, in the ATES, these items were separated from the educational attainment section).

Therefore, the initial cycle of the NTEWS will incorporate a seeded sample of known certificate holders to evaluate whether this most recent revision of the postsecondary certificate questionnaire items succeeded at mitigating the underreporting that had existed in the ATES.

Shortly after data collection is complete, NCSES will analyze the seeded sample results to determine whether the postsecondary certificate data are of sufficient quality for inclusion in NTEWS data products, and whether further revisions to the postsecondary certificate items are required for future NTEWS cycles. NCSES will use prior ATES seeded sample results as a benchmark to determine whether the revisions to the postsecondary certificate item decreased the underreporting rate relative to prior versions. NCSES will also conduct a literature review to determine whether extant literature suggests an appropriate "cut-off" for an acceptable amount of underreporting error.

5.2. Data sources

Building on previous experience conducting seeded sample research in ATES cycles, NCES obtained lists of recent postsecondary certificate awardees, with the necessary contact information for sampling, from three public community/technical college systems and one community college.

These lists will form the frame from which the 1,000-person seeded sample will be drawn. The seeded sample will be stratified by characteristics of the known credential (e.g., the field of study and/or year awarded) to allow analyses to be broken out by these characteristics.

---

[7] The 2010 and 2014 cycles also included seeded samples of certification and license holders. Because these evaluations found acceptably low levels of measurement error in the reporting of certifications and licenses, these credentials were not included in the 2016 seeded sample and are not proposed for inclusion in the NTEWS seeded sample.

Although NCES aimed to obtain lists from as diverse a set of institutions as possible (e.g., from different geographic regions), the seeded sample will be a convenience sample and is not expected to be nationally representative. Therefore, respondents from the seeded sample will not be weighted or included in the publicly released NTEWS data products or in any national estimates derived from the NTEWS production sample. The seeded sample data are for NCSES and NCES research purposes only.

Past seeded samples included postsecondary certificate subject fields that were intended to reflect the general population of certificate holders. The NTEWS seeded sample will be designed to reflect the STW as well as the general population. The job sectors that cover the majority of the STW are installation, maintenance, and repair; health care; construction and extraction; production; and computers (Rothwell 2015). The sectors with the highest percentage of adults with a certificate are installation, maintenance, and repair; health care; protective services; administrative support; and manufacturing and farming (Cronen, McQuiggan & Isenberg 2017). However, the largest number of certificates awarded in a specific year are in the subject fields of health care, trades, and consumer services (Hudson 2018).

Based on overlap with the STW sectors, the seeded sample will contain holders of postsecondary certificates in installation, maintenance, and repair; health care; computers and information technology; manufacturing; and cosmetology (a major component of "consumer services" certificates).

NCES requested lists of certificate holders whose certificates were awarded as far back as January 1, 2016. There are three reasons for limiting the reference period to this date. First, institutions provided the addresses of credential holders at the time of enrollment or program completion. Because some seeded sample members are likely to have changed addresses since receiving their credential, NCSES anticipated that holders of relatively old credentials would be difficult to locate. For this reason, limiting the reference period will help to maximize the response rate among seeded sample members. Second, NCES has found in past seeded sample efforts that "incidental" certificate awards (which, as described below, NCES aims to exclude from the seeded sample) are more likely to be present in older institutional records. Finally, because several of the recruited institutions participated in prior seeded sample efforts, limiting the reference period will minimize duplication with the prior samples.

In past seeded sample efforts, NCES found that some institutions provided certificates of completion automatically to students who completed a certain number of *credits* during another degree program and not because a student completed a certificate *program*. NCES requested that institutions include certificates awarded only as part of a terminal certificate program and exclude these "incidental" certificates. The NTEWS is interested only in certificates that indicate the completion of a program of study.

5.3. Additional operational details

The randomized methodological experiments that apply to the NTEWS production sample will not apply to the seeded sample. All seeded sample members will receive the same contact materials as the *NSCG* treatment group (table C.2, *NSCG* column).

To ensure a sufficient number of respondents for reliable analysis, seeded sample members will receive a $10 noncontingent debit card incentive with the first mailing (i.e., the same as the *INC10* treatment group in the national sample). In the most recent seeded sample evaluation (conducted as part of the 2016 ATES), seeded sample members were sent a $5 cash incentive. As noted above, based on existing literature, NCSES hypothesizes that a $10 debit card incentive will be approximately the equivalent of the $5 cash incentive and that this is therefore the appropriate incentive level to achieve the expected response rate in the seeded sample.

5.4. Use of results

As previously mentioned, NCSES will use the seeded sample results to determine whether the NTEWS certificate data can be included in NTEWS data products, or whether further revisions to the item are needed for future NTEWS cycles to attain sufficient data quality for the production of official estimates. The analysis will be conducted shortly after the completion of data collection to allow decisions on the release of estimates and/or further revisions to the item to be made in a timely manner.

In particular, the following dimensions will be considered in evaluating the quality of the existing NTEWS certificate item:

- The **overall certificate underreporting rate**: the percentage of all seeded sample respondents who report that they do not hold a certificate. This percentage can be interpreted as an underreporting rate because the seeded sample will consist of known certificate recipients. This rate will be compared to the overall certificate underreporting rates observed in the 2010, 2014, and 2016 seeded sample efforts to assess whether the most recent revisions to the item have meaningfully reduced measurement error.
- **Certificate underreporting rates by subgroup**: the underreporting rate will be calculated within key subgroups of interest. This will allow NCSES to assess whether any observed underreporting appears to be related to specific characteristics of the postsecondary certificate and/or the respondent and thereby inform potential future revisions to the item. At a minimum, underreporting rates will be broken out by labor force status, occupational groups, whether the respondent also reported holding a certification or license, the field of the certificate, and the number of years since the certificate was awarded. Prior seeded sample efforts have found that similar characteristics were predictive of credential underreporting. Multivariate logistic regression will be used to identify characteristics that are predictive of underreporting after controlling for other characteristics.

As noted above, to provide further context for the results, NCSES will conduct a literature review to determine whether extant literature suggests an acceptable "cut-off" for the underreporting rate.

If the results of this analysis and literature review indicate that the overall underreporting rate remains unacceptably high, NCSES will consider further changes to the postsecondary certificate questionnaire items. Such revisions could include further re-wording of the items, changing the position of the postsecondary certificate items relative to other sections of the survey, and/or

adding clarification text prior to the items. The specific revisions selected will depend on the nature of the under-reporting findings; for example, if underreporting is particularly high for respondents who also hold a certification, text could be added to clarify that a certificate is different from a certification.

6. References

American Institutes for Research (2013). *Cognitive interview findings for items to measure certifications, licenses, and educational certificates among adults in the United States*. Retrieved from https://nces.ed.gov/surveys/gemena/pdf/FINAL_ATES_cog_int_report_Dec_18_2013.pdf

Bailey, J., Lavrakas, P. J., & Bennett, M. (2007, May). Cash, credit, or check: A test of monetary alternatives to cash incentives. Presented at the 62nd annual conference of the American Association for Public Opinion Research, Anaheim, CA.

Bielick, S., Cronen, S., Stone, C., Montaquila, J.M., & Roth, S.B. (2013). *The Adult Training and Education Survey (ATES) Pilot Study: Technical report* (NCES 2013-190), Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Biemer, P. P., Murphy, J., Zimmer, S., Berry, C., Deng, G., & Lewis, K. (2017). Using bonus monetary incentives to encourage web response in mixed-mode household surveys. *Journal of Survey Statistics and Methodology, 6*(2), 240–261.

Bucks, B., Couper, M. P., & Fulford, S. L. (2019). A mixed-mode and incentive experiment using administrative data. *Journal of Survey Statistics and Methodology, 0*, 1–18.

Church, A.H. (1993). Establishing the effect of incentives on mail survey response rates: A meta-analysis. *Public Opinion Quarterly, 57*(1), 62-79.

Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in web- or Internet-based surveys. *Educational and Psychological Measurement, 60*(6), 821–836.

Cronen, S., McQuiggan, M., & Isenberg, E. (2017). *Adult training and education: Results from the National Household Education Surveys Program of 2016* (NCES 2017-103rev), Washington, DC.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method,* Hoboken, NJ: Wiley & Sons.

Duprey, M.A., Pratt, D.J., Jewell, D.M., Cominole, M.B., Fritch, L.B., Ritchie, E.A., Rogers, J.E., Wescott, J.D., & Wilson, D.H. (2018). *High School Longitudinal Study of 2009 (HSLS:09) base-year to second follow-up: Data file documentation* (NCES 2018-140), Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Freedman, V. A., Mcgonagle, K. A., & Couper, M. P. (2018). Use of a targeted sequential mixed mode protocol in a nationally representative panel study. *Journal of Survey Statistics and Methodology*, *6*(1), 98–121.

Han, D., Montaquila, J.M., & Brick, J.M. (2013). An evaluation of incentive experiments in a two-phase address-based sample mail survey. *Survey Research Methods, 7*(3): 207–218.

Hill, J., Smith, N., Wilson, D., & Wine, J. (2016). *2012/14 Beginning Postsecondary Students Longitudinal Study (BPS:12/14): Data file documentation* (NCES 2016-062), Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Hudson, L. (2018). *Trends in subbaccalaureate occupational awards: 2003 to 2015* (NCES 2018-010), Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Jackson, M., McPhee, C., & Lavrakas, P. (2019). Using response propensity modeling to allocate noncontingent incentives in an address-based sample: Evidence from a national experiment. *Journal of Survey Statistics and Methodology, 0*, 1-27.

Jackson, M., McQuiggan, M., & Megra, M. (2016). *NHES:2014 Feasibility Study report*. Internal report to National Center for Education Statistics.

Lesser, V. M., Newton, L. D., Yang, D. K., & Sifneos, J. C. (2016). Mixed-mode surveys compared with single mode surveys: Trends in responses and methods to improve completion. *Journal of Rural Social Sciences, 31*(3), 7–34.

Matthews, B., Davis, M. C., Tancreto, J. G., Zelenak, M. F., & Ruiter, M. (2012). *2011 American Community Survey Internet tests: Results from second test in November 2011*. Retrieved from https://www.census.gov/content/dam/Census/library/working-papers/2012/acs/2012_Matthews_01.pdf

McPhee, C., Bielick, S., Masterton, M., Flores, L., Parmer, R., Amchin, S., Stern, S., & McGowan, H. (2015). *National Household Education Surveys Program of 2012: Data file user's manual* (NCES 2015-030), Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

McPhee, C. & Hastedt, S. (2012). More money? The impact of larger incentives on response rates in a two-phase mail survey. Presented at the 2012 research conference of the Federal Committee on Statistical Methodology, Washington, DC.

McPhee, C., Jackson, M., Bielick, S., Masterton, M., Battle, D., McQuiggan, M., Payri, M., Cox, C., & Medway, R. (2018). *National Household Education Surveys Program of 2016: Data file user's manual* (NCES 2018-100), Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Medway, R. L., & Fulton, J. (2012). When more gets you less: A meta-analysis of the effect of concurrent web options on mail survey response rates. *Public Opinion Quarterly, 76*(4), 733–746.

Megra, M. & Cronen, S. (2017). *ATES:2016 seeded sample analysis*. Internal report to National Center for Education Statistics.

Mercer, A., Caporaso, A., Cantor, D., & Townsend, R. (2015). How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly, 79*(1), 105–129.

Meyer, B. D., Mok, W. K. C., & Sullivan, J. X. (2015). Household surveys in crisis. *Journal of Economic Perspectives, 29*(4), 199–226.

Olson, K., Smyth, J. D., Horwitz, R., Keeter, S., Lesser, V., Marken, S., … Wagner, J. (2019). *Report of the AAPOR task force on transitions from telephone surveys to self-administered and mixed-mode surveys*. Oakbrook Terrace, IL: American Association for Public Opinion Research.

Rothhaus, C.A., Bentley, M. Hill, J.M., & Lestina, F. (2011). *2010 Census: Bilingual questionnaire assessment report*. Retrieved from https://www.census.gov/content/dam/Census/library/publications/2011/dec/2010_cpex_156.pdf

Rothwell, J. (2015). *Defining skilled technical work*. Retrieved from https://sites.nationalacademies.org/cs/groups/pgasite/documents/webpage/pga_167744.pdf

Sarndal, C. & Lundquist, P. (2014). Accuracy in estimation with nonresponse: A function of degree of imbalance and degree of explanation. *Journal of Survey Statistics and Methodology, 2*(4), 361-387.

Singer, E., & Ye, C. (2013). The use and effects of incentives in surveys. *The ANNALS of the American Academy of Political and Social Science, 645*(1), 112–141.

Thornton, T. (2014). *2013 National Survey of College Graduates (NSCG) incentive conditioning study*. Internal report to National Center for Science and Engineering Statistics.

To, N. (2015). *Review of federal survey program experiences with incentives*. Retrieved from https://www.bls.gov/cex/research_papers/pdf/review-of-incentive-experiences-report.pdf

Trussell, N. & Lavrakas, P. (2004). The influence of incremental increases in token cash incentives on mail survey response: Is there an optimal amount? *Public Opinion Quarterly, 68*(3), 349-367.

Vannieuwenhuyze, J.T.A. & Loosveldt, G. (2012). Evaluating relative mode effects in mixed-mode surveys: three methods to disentangle selection and measurement effects. *Sociological Methods & Research, 42*(1), 82-104.

Vannieuwenhuyze, J., Loosveldt, G., & Molenberghs, G. (2010). A method for evaluating mode effects in mixed-mode surveys. *Public Opinion Quarterly, 74*(5): 1027-1045.

Zotti, A. (2014). *Report for the 2013 National Survey of College Graduates methodological research incentive timing experiment*. Internal report to National Center for Science and Engineering Statistics.