



**National Household Education  
Survey 2019 (NHES:2019)  
Full-scale Data Collection**

**OMB# 1850-0768 v.14**

**Part B  
Description of Statistical  
Methodology**

April 2018  
revised July 2018

## TABLE OF CONTENTS

<b>DESCRIPTION OF STATISTICAL METHODOLOGY</b> .....		<b>1</b>
B.1	<i>Respondent Universe and Statistical Design and Estimation</i> .....	1
B.1.1	Sampling Households.....	1
B.1.2	Within-Household Sampling.....	2
B.1.3	Expected Yield.....	4
B.1.4	Reliability.....	5
B.1.5	Estimation Procedures.....	6
B.1.6	Nonresponse Bias Analysis.....	6
B.2	<i>Statistical Procedures for Collection of Information</i> .....	8
B.3	<i>Methods for Maximizing Response Rates</i> .....	23
B.4	<i>Tests of Procedures and Methods</i> .....	24
B.4.1	Tests Influencing the Design of NHES:2019.....	24
B.4.2	Tests Included in the Design of NHES:2019.....	27
B.5	<i>Individuals Responsible for Study Design and Performance</i> .....	32

### List of Tables

<u>Table</u>		<u>Page</u>
1	Expected percentage of households with eligible children or adults, by sampling domain	4
2	Expected numbers sampled and expected numbers of completed screeners and topical surveys	5
3	Numbers of completed topical interviews in previous NHES surveys administrations	5
4	Margins of error for screener	5
5	Margins of error for topical surveys	6

### List of Figures

<u>Figure</u>		<u>Page</u>
1	Screener Data Collection Plan Flow 1	12
2	Screener Data Collection Plan Flow 2	15
3	Screener Data Collection Plan Flow 3	16
4	Topical Mail Out Data Collection Plan	17
5	Topical Web Data Collection Plan	20

## DESCRIPTION OF STATISTICAL METHODOLOGY

### B.1 Respondent Universe and Statistical Design and Estimation

Historically, an important purpose of the National Household Education Surveys Program (NHES) has been to collect repeated measurements of the same phenomena at different points in time. Decreasing response rates during the past decade required NCES to redesign NHES. This redesign involved changing the sampling frame from a list-assisted Random Digit Dial (RDD) to an Address-Based Sample (ABS) frame. The mode of data collection also changed from telephone, interviewer administered to web and mail, self administered.

The last NHES data collection prior to the redesign was conducted in 2007 and included the PFI and the School Readiness (SR) survey<sup>1</sup>. NHES:2012, comprised of the PFI and ECPP topical surveys, was the first full-scale data collection using an addressed-based sample and a self-administered questionnaire. The overall screener plus topical response rate was approximately 58 percent in 2012 and 49 percent in 2016 for both the PFI and the ECPP, compared to the 2007 overall response rate of 39-41% (depending on the survey). The results suggest that the new methodology can increase response rates and address coverage issues identified in the 2007 data collection.

In addition, NHES:2016 had a web component to help increase response rates and reduce costs. Sample members who were offered web in 2016 were offered paper surveys after the first two web survey invitations were unanswered; therefore, we refer to those cases in 2016 as receiving a “sequential mixed-mode” survey protocol. For this mixed-mode experimental collection, the overall screener plus topical response rate was approximately 52 percent for both the ECPP and PFI. The web component allowed for higher topical response rates because the screener and topical survey on the web appear to the respondent to be one seamless instrument instead of two separate survey requests.

NHES:2019 will be an address-based sample covering the 50 states and the District of Columbia and will be conducted by web and mail from January through September 2019. The household target population is all residential addresses (excluding P.O. Boxes that are not flagged by the United States Postal Service [USPS] as the only way to get mail) and is estimated at 128,882,098 addresses.

Households will be randomly sampled as described in section B.1.1, and an invitation letter to complete the survey on the web or a paper screening questionnaire will be sent to each sampled household.<sup>2</sup> Demographic information about household members provided on the screener will be used to determine whether anyone is eligible for the Early Childhood Program Participation (ECPP) survey or the Parent and Family Involvement in Education (PFI).

The target population for the ECPP survey consists of children age 6 or younger who are not yet in kindergarten. The target population for the PFI survey includes children/youth ages 20 or younger who are enrolled in kindergarten through twelfth grade or who are homeschooled for kindergarten through twelfth grade. Ages will be calculated based on information provided in the screener.

#### B.1.1 Sampling Households

A nationally representative sample of 205,000 addresses will be used. The sample will be drawn in a single stage from a file of residential addresses maintained by a vendor, Marketing Systems Group (MSG), based on the USPS Computerized Delivery Sequence (CDS) File. MSG provided the sample for NHES:2012, the NHES Feasibility Study (NHES-FS) in 2014, NHES:2016, and the NHES:2017 Web Test. The initial sample will consist of 225,500 addresses selected from an ABS frame by MSG, based on the United States Postal Service (USPS) Computerized Delivery Sequence (CDS) File. After invalid addresses are removed from the 225,500-household sample, it will be subsampled by NCES’s sample design contractor, the American Institutes for Research (AIR), yielding a 205,000-household final screener sample. The remaining households will remain unused to protect respondent confidentiality.

As in past NHES surveys, NHES:2019 will oversample black and Hispanic households using Census data. This oversampling is necessary to produce more reliable estimates for subdomains defined by race and ethnicity. NHES:2019 will use the stratification methodology that was used for the NHES:2012 and NHES:2016 administrations. The sampling design concentrates on achieving a significant sample of black and Hispanic households. The universe consists of Census tracts likely to contain relatively high proportions of these subgroups. The three strata for NHES:2019 are defined as households in:

---

<sup>1</sup> NHES:2007 also included an adult education module, but data collection was stopped due to low response rates.

<sup>2</sup> In the web mode, each household will receive a letter with a user identification code and URL address asking someone in the household to complete the survey by web rather than a letter with a screening questionnaire. In the choice-plus condition, households will receive a letter inviting them to complete the survey by web but will also include a paper screener questionnaire. In the paper condition, households will receive a paper screener with a cover letter.

- 1) Tracts with 25 percent or more Black persons
- 2) Tracts with 40 percent or more Hispanic persons (and not 25 percent or more Black persons)
- 3) All other tracts

The Hispanic stratum contains a high concentration of households in which at least one adult age 15 or older speaks Spanish and does not speak English very well and will be used for the assignment of bilingual screener mailing materials. The sample allocation to the three strata will be 20 percent to the black stratum, 15 percent to the Hispanic stratum and 65 percent to the third stratum. This allocation will provide improvement in the precision of estimates by race/ethnicity compared to use of a uniform sampling rate across each stratum, and will protect against unknown factors that may affect the estimates for key subgroups, especially differential response rates.

In addition to stratifying by the race/ethnicity groups mentioned above, these three strata will be sorted by a poverty variable and the sample will be selected systematically from the sorted list to maintain the true poverty level proportions in the sample. The Census tracts will be used to define the two categories of poverty level as households in:

- 1) Tracts with 20% or more below the poverty line
- 2) Tracts with less than 20% below the poverty line

Households will be randomly assigned to one of five data collection treatment conditions (baseline, targeted mailing, updated sequential mixed-mode, choice-plus, and modeled mode conditions).

- The baseline sample will mimic the 2016 sequential mixed mode design protocols which were developed for the web experiment in that administration. These cases will allow us to compare 2019 response rates to 2016 response rates.
- The sample allocated for targeted mailings will be reserved for experimental testing of a set of survey materials designed specifically for Spanish-speakers. Addresses in this group that are predicted to include a Spanish speaker will be sent materials designed to attract response specifically from them. Other addresses in this group will receive the same English language materials as the baseline sample.
- The updated sequential mixed-mode condition utilizes the data about response from previous collections to create a survey protocol that represents what we predict will be the optimal set of survey protocols for 2019. Multiple experiments are embedded in this group to further tease out optimal cost/response tradeoffs.
- The choice-plus treatment replicates a design tested in the U.S. Energy Information Administration's Residential Energy Consumption Survey. Households are provided web and paper options but incentivized to respond by web.
- The modeled mode condition will allocate sample for the purpose of testing a model that predicts paper response; addresses in this group will be assigned to receive either all paper throughout the survey administration or will be administered the baseline survey protocol.

A total of 40,000 households will be chosen for the baseline condition, 15,000 for the targeted mailing condition, 80,000 for the updated sequential mixed mode condition, 30,000 for the choice-plus condition, and 40,000 households for the modeled mode condition.

### **B.1.2 Within-Household Sampling**

Among households that complete the screener and report children eligible for one or more topical surveys, a three-phase procedure will be used to select a single household member to receive a topical questionnaire. In order to minimize household burden, only one eligible child from each household will be sampled; therefore, each household will receive only one of the two topical surveys. In households with multiple children, homeschooled children are chosen at higher rates than enrolled children, and young children who are eligible to be sampled for the ECPP survey are chosen at higher rates than K-12 enrolled students. The topical sampling procedure works by randomly assigning two pre-designations to each household. These pre-designations include homeschool status and the domain which is the type of topical survey (PFI or ECPP). Depending on the composition of the household, some or all of these pre-designations are used to assign the household to one of the two topical surveys. If the household has more than one member eligible for the survey for which it is selected, the final phase of sampling randomly selects one of these members for that survey. The remainder of this section provides additional detail on each of the three phases, and provides examples of how this procedure would work for three hypothetical households.

### First-phase sampling: Homeschooling Status (Homeschool child or not)<sup>3</sup>

In the first phase of sampling, each address will be randomly pre-designated as either a “homeschool household” or an “other household.” This pre-designation will be referred to as the homeschooling pre-designation. A household will be pre-designated as a “homeschool household” with a 0.8 probability and as an “other household” with a 0.2 probability. This pre-designation will only be used when a household has at least one child eligible for the PFI and their enrollment response on the screener is “homeschool instead of attending a public or private school for some or all classes” and at least one child is eligible for one of the other topical surveys. It will be used to determine which person in the household should be sampled. If a household only has a child or children eligible for the PFI and are homeschooled, that household will receive the PFI regardless of pre-designation. Likewise, if a household only has members eligible for one of the other surveys, that household will receive a survey regardless of pre-designation. The purpose of sampling homeschool children in the first phase of topical sampling is to increase the number of homeschooled children for whom data are collected. NHES is the only nationally representative survey that collects information about homeschooled children in the United States. Because this is such a small sub-population, in NHES:2016 this sampling scheme was introduced and resulted in doubling the number of homeschool respondents compared to the NHES:2012 administration, which did not sample households separately for homeschool children.<sup>4</sup> Due to the small number of households with homeschooling children, this is expected to have a minimal impact on sample sizes and design effects for the topical surveys.

### Second-Phase Sampling for Households Selected to Receive a Child Topical Survey: Domain (ECPP or PFI)

The second phase of sampling is at the type-of-topical-survey-sent level, or the domain level. There are two topical surveys that a child could possibly be eligible for in the second phase: the ECPP or the PFI survey. A child’s age, among other factors, will determine if the child is eligible for the ECPP or the PFI topical survey. The eligibility criterion only allows a child to be eligible for one of the topical surveys, not both.

In addition to the homeschool pre-designation, each household will be pre-designated as an “ECPP household” with a 0.7 probability or a “PFI household” with a 0.3 probability. This pre-designation will be referred to as the topical survey pre-designation. Because children eligible for the ECPP survey comprise a smaller portion of the population compared to children eligible for the PFI survey, differential sampling is done to ensure a sufficient sample size for the ECPP group. This designation will only be used when a household was not selected for the homeschool designation in the first phase and has a child/children eligible for the ECPP topical survey and a child/children eligible for the PFI topical survey, to determine which person in the household should be sampled. If a household only has a child/children eligible for the ECPP topical survey and not the PFI topical survey, that household will receive an ECPP topical survey, regardless of pre-designation. Likewise, if a household only has a child/children eligible for the PFI topical survey and not the ECPP topical survey, that household will receive a PFI topical survey, regardless of pre-designation.

### Third-phase Sampling: Child Level

The third phase of sampling is at the child level. If any household has only one child that is eligible for the domain that was selected in the first two phases of sampling, that child is selected as the sample member about whom the topical survey is focused. If any household has two or more children eligible for the domain that was selected in the first two phases of sampling, one of those children will be randomly selected (with equal probability) as the sample member to be the focus of the appropriate topical survey. For example, if there are two children in the house eligible for the ECPP survey and the household was pre-designated an ECPP household, then one eligible child will be randomly selected (with equal probability). Therefore, at the end of the three phases of sampling, one eligible child within each household that has a completed screener will be sampled for the ECPP or PFI topical survey, respectively.

### Topical Sampling Examples

The use of the topical pre-designations can be illustrated by considering hypothetical households with members eligible for various combinations of the topical surveys. First, consider a household with four children: one child eligible for the ECPP, three children eligible for the PFI (one child is homeschooled). Suppose that the household is pre-designated as an “other household” in the first phase and a “PFI household” in the second phase. The first pre-designation indicates that the household should not sample the homeschool child; the second pre-designation indicates

<sup>3</sup> In prior administrations, NHES administered a PFI-Homeschool and PFI-enrolled as separate surveys. NHES:2019 has combined both surveys into one survey but the sampling of homeschooled children will be the same as it was in NHES:2016.

<sup>4</sup> In prior administrations, before NHES:2019, a household received the PFI-Homeschool survey if the household was sampled for PFI, and the individual child sampled in the final phase of sampling was identified on the screener as a homeschooler. In NHES:2019, screener identification as a homeschooler will assist with oversampling homeschoolers, but the same PFI survey will be used for all K-12 sample members.

that the household should receive the PFI. Because the household has two children eligible for the PFI who are not homeschooled, one of these non homeschooled children is randomly selected for the PFI with 0.5 probability in the final phase.

Second, consider a household with two children: one child eligible for the PFI who is homeschooled and one child eligible for ECPP. Suppose that the household is pre-designated as a “homeschool household” in the first phase and an “ECPP household” in the second phase. Because the household has a homeschool child eligible for the PFI, the first pre-designation is used to assign the household to the PFI. Therefore, the other pre-designation is not used. Because only one member is a homeschool child eligible for the PFI, this child is selected with certainty for the PFI in the final phase.

Finally, consider a household with three children: one child eligible for the ECPP and two children eligible for PFI (no children are homeschooled). Suppose that the household is pre-designated as a “homeschool household” in the first phase and a “ECPP household” in the second phase. Because the household has no homeschool children eligible for the PFI, the first pre-designation is not used. Since the household second predesignation is “ECPP household” the household is assigned to receive the ECPP. Because only one child is eligible for the ECPP, this child is selected with certainty for the ECPP in the final phase.

### B.1.3 Expected Yield

As described above, the initial sample will consist of approximately 205,000 addresses. An expected screener response rate of approximately 53.545 percent and an address ineligibility<sup>5</sup> rate of approximately 10 percent are assumed, based on results from NHES:2016 and expected differences in screener eligibility and response rates between the sampling strata and experimental treatment groups. Under this assumption, the total number of expected completed screeners is 98,790.

The ECPP or PFI topical surveys will be administered to households with completed screeners that have eligible children. For NHES:2019, we expect to achieve a total percentage of households with eligible children of approximately 28.735 percent.<sup>6</sup> Expected estimates of the percentage of households with eligible children overall and in each sampling domain are given in Table 1, as well as the expected number of screened households in the nationally representative sample of 205,000, based on the distribution of household composition and assuming 98,790 total completed screeners.

**Table 1. Expected percentage of households with eligible children or adults, by sampling domain**

Household composition	Percent of households	Expected number of screened households
<b>Total households with eligible children</b>	<b>28.735</b>	<b>28,387</b>
Households with at least one PFI eligible child enrolled in K-12 and no homeschool or ECPP eligible children	17.640	17,427
Households with at least one homeschooled child and no other PFI or ECPP eligible children	0.578	571
Households with at least one ECPP eligible child and no PFI eligible children enrolled in K-12 or homeschooled	5.074	5,013
Households with at least one PFI eligible child enrolled in K-12, at least one homeschool eligible child, and no ECPP eligible children	0.216	213
Households with at least one PFI eligible child enrolled in K-12, at least one ECPP eligible child, and no homeschool eligible children	4.968	4,908
Households with at least one homeschool eligible child, at least one ECPP eligible child, and no other PFI eligible children	0.191	189
Households with at least one PFI eligible child enrolled in K-12, at least one homeschool eligible child, and at least one ECPP eligible child	0.069	68

NOTE: The distribution in this table assumes 98,790 screened households. Detail may not sum to totals because of rounding. Estimates are based on calculations from NHES:2016.

Table 2 summarizes the expected numbers of completed interviews for NHES:2019. These numbers take into account within-household sampling. Based on results from NHES:2016, a topical response rate of approximately 82.216% is expected for the ECPP, approximately 82.909% for the PFI.<sup>7</sup> Based on an initial sample of 205,000 addresses and an

<sup>5</sup> Ineligible addresses are those that are undeliverable. Screener mailings for an address where one or more mailings are returned as a postmaster return (PMR) and no mailings are returned complete or refused will lead to an address being coded as ineligible.

<sup>6</sup> Percentages based on estimates from the NHES:2016 screener.

<sup>7</sup> Expected topical response rates represent rounded weighted averages of the response rates expected within each experimental group.

expected eligibility rate of 90 percent, the expected number of completed screener questionnaires is 98,790. Of these, we expect to have approximately 8,495 children selected for the ECPP and 19,893 children for the PFI (947 of which will be homeschooled children).

**Table 2. Expected number of sampled and completed screeners and topical surveys for households**

Survey	Expected number sampled NHES:2019	Expected number of completed interviews NHES:2019
Household screeners	205,000	98,790
ECPP	8,495	6,984
PFI	19,893	16,493
Homeschooled children	947	655

Note: Details may not sum to totals. The PFI combined response rate is expected to be 82.909% but the expected response rates differ depending on the type of respondent. For PFI homeschooled children the response rate is 69.196% and those PFI children enrolled in school is 83.602%.

For comparison purposes, Table 3 shows the number of ECPP and PFI completed interviews from previous NHES surveys administrations.

**Table 3. Numbers of completed topical interviews in previous NHES surveys administrations**

Survey	NHES surveys administration									
	1993	1995	1996	1999	2001	2003	2005	2007	2012	2016
ECPP	†	7,564	†	6,939	6,749	†	7,209	†	7,893	5,844
PFI - Enrolled	19,144	†	17,774	17,376	†	12,164	†	10,370	17,166	13,523
PFI - Homeschooled	†	†	†	285	†	262	†	311	397	552

† Not applicable; the specified topical survey was not administered in the specified year.

#### B.1.4 Reliability

The reliability of estimates produced under specific sample sizes is measured by the margins of error. The reliability of the screener sample size (98,790) is shown in Table 4. As an example, if an estimated proportion is 30 or 70 percent, the margin of error will be below 1 percent for the overall population, as well as within subgroups that constitute 50 percent, 20 percent, and 10 percent of the population. The same is true for survey estimate proportions of 10, 20, 80 or 90 percent.

**Table 4. Expected margins of error for NHES:2019 screener, by survey proportion estimate and subgroup size†**

Survey estimate proportion	Margin of error for estimate within:			
	All respondents	50% subgroup	20% subgroup	10% subgroup
10% or 90%	0.19%	0.27%	0.43%	0.61%
20% or 80%	0.26%	0.36%	0.58%	0.81%
30% or 70%	0.30%	0.42%	0.66%	0.93%
40% or 60%	0.32%	0.45%	0.71%	1.00%
50%	0.32%	0.46%	0.72%	1.02%

† Based on NHES:2016, a design effect of 1.066 was used in the calculations for this table. This represents the design effect due to unequal weighting at the screener level. The margins of error were calculated assuming a confidence level of 95 percent, using the following formula:  $1.96 \cdot \sqrt{p \cdot (1-p) / ne}$ , where p is the proportion estimate and ne is the effective sample size for the screener survey.

Table 5 shows the expected margin of error for all topical respondents and subgroups that are 50, 20, and 10 percent of the full sample size. The precision is dependent on the expected margin of error for each estimate, which is expected to be between 1 percentage point and 15.0 percentage points depending on the subgroup size.

**Table 5. Expected margins of error for NHES:2019 topical surveys percentage estimates, by subgroup size, topical survey, and survey estimate proportion†**

Topical survey	Survey estimate proportion	Margin of error for estimate within:			
		All respondents	50% subgroup	20% subgroup	10% subgroup
PFI	10% or 90%	0.65%	0.91%	1.44%	2.04%
	20% or 80%	0.86%	1.22%	1.92%	2.72%
	30% or 70%	0.99%	1.39%	2.20%	3.12%

	40% or 60%	1.05%	1.49%	2.36%	3.33%
	50%	1.08%	1.52%	2.41%	3.40%
ECPP	10% or 90%	0.93%	1.32%	2.09%	2.95%
	20% or 80%	1.24%	1.76%	2.78%	3.94%
	30% or 70%	1.43%	2.02%	3.19%	4.51%
	40% or 60%	1.52%	2.16%	3.41%	4.82%
	50%	1.56%	2.20%	3.48%	4.92%
PFI homeschooled children	10% or 90%	2.84%	4.01%	6.35%	8.98%
	20% or 80%	3.79%	5.35%	8.46%	11.97%
	30% or 70%	4.34%	6.13%	9.70%	13.71%
	40% or 60%	4.64%	6.56%	10.37%	14.66%
	50%	4.73%	6.69%	10.58%	14.96%

† Based on preliminary analysis, the following design effects were used in the calculations for this table: 2.039 for the PFI, 1.761 for the ECPP, and 1.524 for the PFI homeschooled children. These represent the design effects due to unequal weighting at the screener and topical levels. The margins of error were calculated assuming a confidence level of 95 percent, using the following formula:  $1.96 \cdot \sqrt{p \cdot (1-p) / ne}$ , where p is the proportion estimate and ne is the effective sample size for the topical survey.

### B.1.5. Estimation Procedures

The data sets from NHES:2019 will have weights assigned to facilitate estimation of nationally representative statistics. All households responding to the screener will be assigned weights based on their probability of selection and a non-response adjustment, making them representative of the household population. All individuals in the national sample responding to the topical questionnaires will have a record with a person weight designed such that the complete data set represents the target population.

The estimation weights for NHES:2019 surveys will be formed in stages. The first stage is the creation of a base weight for the household, which is the inverse of the probability of selection of the address. The second stage is a screener nonresponse adjustment to be performed based on characteristics available on the frame.

The household-level weights are the base weights for the person-level weights. For each completed topical questionnaire, the person-level weights also undergo a series of adjustments. The first stage is the adjustment of these weights for the probability of selecting the child within the household. The second stage is the adjustment of the weights for topical survey nonresponse to be performed based on characteristics available on the frame and screener. The third stage is the raking adjustment of the weights to Census Bureau estimates of the target population. The variables that may be used for raking at the person level include race and ethnicity of the sampled child, household income, home tenure (own/rent/other), region, age, grade of enrollment, gender, family structure (one parent or two parent), and highest educational attainment in household. These variables (e.g., family structure) have been shown to be associated with response rates. The final raked person-level weights include undercoverage adjustments as well as adjustments for nonresponse.

Standard errors of the estimates will be computed using a jackknife replication method. The replication process repeats each stage of estimation separately for each replicate. The replication method is especially useful for obtaining standard errors for the statistics that NHES:2019 produces.

### B.1.6 Nonresponse Bias Analysis

To the extent that those who respond to surveys and those who do not differ in important ways, there is a potential for nonresponse biases in estimates from survey data. The estimates from NHES:2019 are subject to bias because of unit nonresponse to both the screener and the extended topical surveys, as well as nonresponse to specific items. Per NCES statistical standards, a unit-level nonresponse bias analysis will be conducted if the NHES:2019 overall unit response rate (the screener response rate multiplied by the topical response rate) falls below 85 percent (as is expected). Additionally, any item with an item-level response rate below 85 percent will be subject to an examination of bias due to item nonresponse.

#### Unit nonresponse

To identify characteristics associated with unit nonresponse, a multivariate analysis will be conducted using a categorical search algorithm called Chi-Square Automatic Interaction Detection (CHAID). CHAID begins by identifying the characteristic of the data that is the best predictor of response. Then, within the levels of that characteristic, CHAID identifies the next best predictor(s) of response, and so forth, until a tree is formed with all of the response predictors that were identified at each step. The result is a division of the entire data set into cells by attempting to determine sequentially the cells that have the greatest discrimination with respect to the unit response



rates. In other words, it divides the data set into groups so that the unit response rate within cells is as constant as possible, and the unit response rate between cells is as different as possible. Since the variables considered for use as predictors of response must be available for both respondents and nonrespondents, demographic variables from the sampling frame provided by the vendor (including household education level, household race/ethnicity, household income, age of head of household, whether the household owns or rents the dwelling, and whether there is a surname and/or phone number present on the sampling frame) and data from the screener (including number of children in the household) will be included in the CHAID analysis at the topical person-level nonresponse. The results of the CHAID analysis will be used to statistically adjust estimates for nonresponse. Specifically, nonresponse-adjusted weights will be generated by multiplying each household's and person's unadjusted weight (the reciprocal of the probability of selection, reflecting all stages of selection) by the inverse response rate within its CHAID cell.

In NHES:2016 an alternative method for identifying characteristics associated with unit nonresponse was tested that used regression trees to model all the covariate relationships in the data. Each covariate was put into one regression model to predict unit nonresponse. We calculated nonresponse adjustments using both the CHAID method and regression tree method in NHES:2016 and found no differences in estimates between the methods. For NHES:2019 an additional frame data source will be appended to the file to determine if additional auxiliary variables will improve the nonresponse weighting adjustment. We plan to calculate nonresponse adjustments using the CHAID method used in 2016 and include these additional variables to see if differences in estimates would result.

The extent of potential unit nonresponse bias in the estimates will be analyzed in several ways. First, the percentage distribution of variables available on the sampling frame will be compared between the entire eligible sample and the subset of the sample that responded to the screener. While the frame variables are of unknown accuracy and may not be strongly correlated with key estimates, significant differences between the characteristics of the respondent pool and the characteristics of the eligible sample may indicate a risk of nonresponse bias. Respondent characteristics will be estimated using both unadjusted and nonresponse-adjusted weights (with and without the raking adjustment), to assess the potential reduction in bias attributable to statistical adjustment for nonresponse. A similar analysis will be used to analyze unit nonresponse bias at the topical level, using both sampling frame variables and information reported by the household on the screener.

In addition to the above, the magnitude of unit nonresponse bias and the likely effectiveness of statistical adjustments in reducing that bias will be examined by comparing estimates computed using adjusted weights to those computed using unadjusted weights. The unadjusted weight is the reciprocal of the probability of selection, reflecting all stages of selection. The adjusted weight is the extended interview weight adjusted for unit nonresponse (with and without the raking adjustment). In this analysis, the statistical significance of differences in estimates will be investigated only for key survey estimates including, but not limited to, the following:

#### All surveys

- Age/grade of child
- Census region
- Race/ethnicity
- Parent 1 employment status
- Parent 1 home language
- Educational attainment of parent 1
- Family type
- Household income
- Home ownership

#### Early Childhood Program Participation (ECPP)

- Child receiving relative care
- Child receiving non-relative care
- Child receiving center-based care
- Number of times child was read to in past week
- Someone in family taught child letters, words, or numbers
- Child recognizes letters of alphabet
- Child can write own name
- Child is developmentally delayed
- Child has health impairment
- Child has good choices for child care/early childhood programs

#### Parent and Family Involvement in Education (PFI)

- School type: public, private, homeschool, virtual school
- Whether school assigned or chosen
- Child’s overall grades
- Contact from school about child’s behavior
- Contact from school about child’s school work
- Parents participate in 5 or more activities in the child’s school
- Parents report school provides information very well
  - About how child is doing in school
  - About how to help child with his/her homework
  - About why child is placed in particular groups or classes
  - About the family’s expected role at child’s school
  - About how to help child plan for college or vocational school
- Parents attended a general school meeting (open house), back-to-school night, meeting of parent-teacher organization
- Parents went to a regularly scheduled parent-teacher conference with child’s teacher
- Parents attended a school or class event (e.g., play, sports event, science fair) because of child
- Parents acted as a volunteer at the school or served on a committee
- Parents check to see that child’s homework gets done

The next component of the bias analysis will include comparisons between respondent characteristics and known population characteristics from extant sources including the Current Population Survey (CPS) and the American Community Survey (ACS). Additionally, for substantive variables, weighted estimates will be compared to prior NHES administrations if available. While differences between NHES:2019 estimates and those from external sources as well as prior NHES administrations could be attributable to factors other than bias, differences will be examined in order to confirm the reasonableness of the 2019 estimates. The final component of the bias analysis will include a comparison of base-weighted estimates between early and late respondents of the topical surveys to assess bias risk in the estimates.

**Item nonresponse**

In order to examine item nonresponse, all items with response rates below 85 percent will be listed. Alternative sets of imputed values will be generated by imposing extreme assumptions on the item nonrespondents. For most items, two new sets of imputed values—one based on a “low” assumption and one based on a “high” assumption—will be created. For most continuous variables, a “low” imputed value variable will be created by resetting imputed values to the value at the fifth percentile of the original distribution; a “high” imputed value variable will be created by resetting imputed values to the value at the 95<sup>th</sup> percentile of the original distribution. For dichotomous and most polytomous variables, a “low” imputed value variable will be created by resetting imputed values to the lowest value in the original distribution, and a “high” imputed value variable will be created by resetting imputed values to the highest value in the original distribution. Both the “low” imputed value variable distributions and the “high” imputed value variable distributions will be compared to the unimputed distributions. This analysis helps to place bounds on the potential for item nonresponse bias using “worst case” scenarios.

**B.2 Statistical Procedures for Collection of Information**

This section describes the data collection procedures to be used in NHES:2019. These procedures represent a combination of best practices to maximize response rates based on findings from NHES:2012, NHES:2016, and NHES:2017 Web Test within NCES’s cost constraints. NHES is a two-phase self-administered survey. In the first phase, households are screened to determine if they have eligible children. In the second phase, a detailed topical questionnaire is sent to households with eligible children. NHES employs multiple contacts with households to maximize response. These include an advance letter and up to four questionnaire mailings for both the screener and the topical surveys. In addition, households will receive one reminder postcard or pressure sealed envelope after the initial mailing of a screener or topical questionnaire. Respondent contact materials are provided in Appendix 1, paper survey instruments in Appendix 2, and web screener and topical instruments in Appendix 3.

Additionally, in conjunction with NHES:2019, NCES plans to conduct an In-Person Study of NHES:2019 Nonresponding Households, designed to provide insight about nonresponse that can help plan future survey administrations. This in-person study is described in detail in Appendix 4 of this submission; its contact materials, final interview protocols, and final sample selection details will be submitted to OMB for review with a 30-day public comment period in fall 2018.

***Screener Procedures***

*Advanced mailer campaign.* For a select number of cases (23,333 randomly assigned cases) the NHES:2019 data collection will begin with two separate mailings of over-sized postcards in mid December 2018. The mailings will not instruct the cases to do anything; instead, their purpose is to build awareness of NCES and NHES by providing brief statistics about early childhood care and children’s K-12 education from prior administrations of NHES. One purpose of the mailings is to make the people who live at the address more familiar with NCES and NHES so that when they receive the NHES screener questionnaire they will be more likely to recognize from where it is coming. The advance mailer campaign also may make it more likely that the people at the sampled address would be favorably disposed to the value of NHES, which would make it more likely that they would complete the questionnaire when they receive it. These mailings will be two glossy over-sized postcards that are professionally designed with attractive and eye-catching graphics. These two over-sized postcards are provided in Appendix 1.

*Advance letter.* Outside of those cases that receive the advanced mailer campaign, the NHES:2019 data collection will begin with the mailing of an advance notification letter in early January 2019. A select number of cases (23,333 randomly assigned cases) in the sequential mixed mode contact strategies experimental group, described in Part A, section A.2 of this package, will not receive an advance letter. This is to test whether or not an advance letter increases response rates in a mixed-mode design that invites the household to respond initially by web. Advance letters have increased response in the landline phone and paper modes in previous NHES administrations, but the NHES program has not yet tested their effects in a web-first mode.

*Screener package.* Most households will receive an invitation to respond to the web survey in their first screener package. A select number of cases in the modeled mode experimental condition will receive the mail treatment, which involves receiving a paper screener questionnaire in the first screener mailing package. Figures 1-3 present the data collection plan for the web mode households. The initial mailing will include an invitation letter by mail (which includes the web survey URL and log-in credentials) and a cash incentive. As shown in Figures 1-3, mail modelled households and choice-plus households will be mailed a questionnaire package one week after the advance letter. The packages will contain a cover letter, household screener, business reply envelope, and a cash incentive. In addition, households in the Choice-Plus incentive treatment will receive all materials listed above in the first screener mailing package. The cover letter will explain to the household that web and paper are options, but that choosing the web survey will save the government money and thus will be further incentivized.

*Incentive.* Many years of testing in NHES have shown the effectiveness of a small cash incentive on increasing response (see Part A, section A.9 of this submission). NHES:2019 will continue to use a \$5 screener incentive.

An additional incentive experiment will be conducted in NHES:2019 for a select number of cases to encourage sample members to respond via web or the helpdesk phone (Census Telephone Questionnaire Assistance line) instead of by paper. A \$10 or \$20 contingent incentive will be offered if the sample member responds via web or through the Census telephone questionnaire assistance (TQA) line to all the surveys they are eligible for. For example, if the household does not have an eligible child they will receive the incentive for completing the screener. A household that does have an eligible child will receive the promised incentive for completion of the screener and topical surveys. The incentive will be mailed to the respondent 3-5 weeks after they respond through the appropriate mode. Part A, section A.9 of this submission provides a detailed description of this Choice-Plus incentive experiment. Sample members allocated to this experiment will be randomly assigned to either the \$10 contingent incentive group or the \$20 contingent incentive group.

*Reminders and nonresponse follow-up.* The reminder and nonresponse follow-up varies depending on the experimental treatment to which the case is assigned. The majority of cases are assigned to the web mode to respond. A thank you/reminder postcard or pressure-sealed envelope will be sent to web sampled addresses approximately one week after the first mailing. Unlike the postcard, the thank you/reminder pressure-sealed envelope will include the URL and log-in credentials to complete the survey. Three nonresponse follow-up mailings follow this postcard/pressure-sealed envelope. For the web mode, the mailings will contain reminder letters with URL and log-in credentials. A second reminder mailing will be sent to nonresponding households approximately two weeks after the postcard/pressure-sealed envelope. For some cases the second reminder mailing will be sent using rush delivery (FedEx or USPS Priority Mail<sup>®</sup>). Cases receiving FedEx at the second mailing will be among those that are allocated to the sample within the sequential mixed mode contact strategies FedEx experiment. Approximately three weeks after the second mailing, a third reminder mailing will be sent to non-responding households using rush delivery if they did not receive rush delivery in the second mailing. Most sampled cases will receive the third survey package by FedEx. This mailing will include a paper screener questionnaire with a business reply envelope. A small number of cases that are part of the

---

<sup>®</sup> Because FedEx cannot deliver to P.O. Boxes, P.O. Box-only addresses will be delivered using USPS Priority Mail.

FedEx mailing experiment will receive this third mailing through USPS. Finally, the fourth questionnaire package will be sent to nonresponding households approximately three weeks after the third questionnaire mailing. For a select number of cases that did not receive a rush delivery package in either the second or third mailing, this final mailing will be sent using rush delivery.

Sampled addresses in the modeled mode experiment that were determined to have a high propensity to respond to a paper survey and low propensity to respond to the web will be part of the mail mode. For these cases, a thank you/reminder postcard will be sent to sampled addresses approximately one week after the first mailing. Three nonresponse follow-up questionnaire mailings follow this postcard. All nonresponse follow-up questionnaire mailings for mail treatment households will contain a cover letter, replacement screener questionnaire, and business reply envelope. A second questionnaire mailing will be sent to nonresponding households approximately two weeks after the postcard. Approximately three weeks after the second mailing, a third questionnaire mailing will be sent to nonresponding households using rush delivery. The fourth questionnaire package will be sent to nonresponding households approximately three weeks after the third questionnaire mailing.

*Language.* The following criteria were used in NHES:2012 and NHES:2016 and are proposed again for NHES:2019 to determine which addresses receive a bilingual screener package:

- First mailing criteria: An address is in the Hispanic stratum, or a Hispanic surname is associated with the address, or the address is in a Census tract where 10 percent or more of households meet the criterion of being linguistically isolated Spanish-speaking.
- Second mailing criteria: An address is in the Hispanic stratum, or a Hispanic surname is associated with the address, or the address is in a Census tract where 3 percent or more of households meet the criterion of being linguistically isolated Spanish-speaking.
- Third and fourth mailing criteria: An address is in the Hispanic stratum, or a Hispanic surname is associated with the address, or the address is in a Census tract where 2 percent or more of households meet the criterion of being linguistically isolated Spanish-speaking, or the address is in a Census tract where 2 percent or more of the population spoke Spanish.

*Targeted mailings.* As an experiment, some households sampled for NHES:2019 will receive a set of materials that are targeted to Spanish-speakers. Households in this group will follow the same sequence of contacts as those described above for addresses receiving a sequential multi-mode protocol of two web invitations followed by two paper survey invitations at both the screener and topical phases. Materials sent to the treatment group for this experiment may include additional images of Hispanics/ Latinos. They may also include messaging that may appeal to Hispanics/Latinos. Design details are under development at the time of this submission.

The sample for the targeted mailings experiment will be randomly selected. However, among the cases allocated to the experiment, the following criteria will be applied to determine which of these cases receives mailings targeted for Spanish speakers:

- The household is flagged on the frame as having a Hispanic surname,
- The household is in a tract with 40% or more Hispanic persons (and not 25% or more Black persons), and
- The household is in a tract where 10% or more of households speak limited English.

About 22% of sampled cases (3,300) in this condition will receive tailored materials and the other 78% (11,700) will receive the same mailing protocol as the majority of cases. Figure 1 shows the screener contact procedures in a flow chart.

### **Topical Procedures**

Topical survey mailings will follow procedures similar to the screener procedures. As described earlier, households with at least one eligible person will be assigned to one of two topical survey mailing groups:

- (1) Households that have been selected to receive an ECPP questionnaire and
- (2) Households that have been selected to receive a PFI questionnaire.

Within each group, for cases in the mail mode group, questionnaires will be mailed in waves to minimize the time between the receipt of the screener and the mailing of the topical. Regardless of wave, households with a selected sample member will have the same topical contact strategy within the mail treatment group and the web treatment group. The initial topical mailing for sampled households in the mail mode group will include a cover letter,

questionnaire, business reply envelope, and a \$5 or \$15 cash incentive. A reminder/thank you postcard is then sent to all topical sample members. A total of three nonresponse follow-up mailings follow. All nonresponse follow-up mailings will contain a cover letter, replacement questionnaire, and business reply envelope. The third mailing will be sent rush delivery.

Web mode cases that break off prior to completing any topical questionnaire items will receive an initial topical mailing (with the web survey URL and log-in credentials) and a \$5 or \$15 cash incentive. Web nonrespondents will receive a 2<sup>nd</sup> followup mailing with the URL and log-in credentials. In the 3<sup>rd</sup> and 4<sup>th</sup> mailings, web nonrespondents will be moved to the mail mode process, receiving a paper-and-pencil instrument and business reply envelope each time.

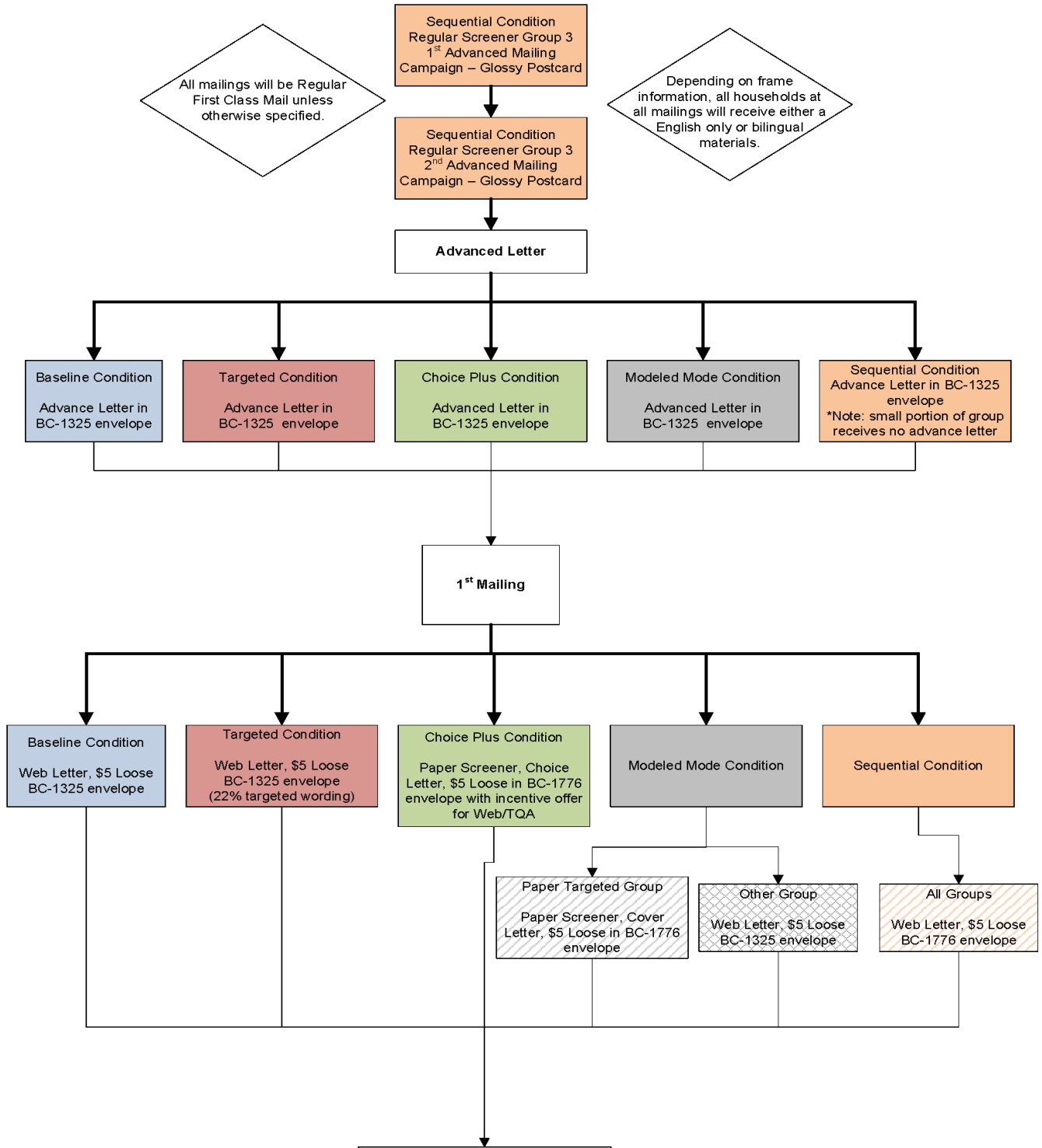
*Incentive.* The NHES:2011 Field Test, NHES:2012, NHES-FS, and NHES:2016 showed that a \$5 incentive was effective with most respondents at the topical level. The results from NHES:2011 Field Test indicate that, among later screener responders, the \$15 incentive was associated with higher topical response rates compared to the \$5 incentive. Based on these findings, we used an incentive model that provides a \$5 topical incentive for early-screener responders and a \$15 topical incentive for late-screener respondents for NHES:2016. As a result, for NHES:2019 we will offer these late screener responders a \$15 prepaid cash incentive with their first topical questionnaire mailing. All other respondents in the main sample will receive \$5 cash with their initial mailing.

An additional incentive experiment will be conducted in NHES:2019 for a select number of cases to encourage respondents to respond via web or the TQA helpdesk phone instead of by paper. A \$10 or \$20 contingent incentive will be offered if the respondent responds via web or the TQA to all the surveys they are eligible for. For example, if the household does have an eligible child they will need to complete both the screener and topical survey to receive the incentive. The incentive will be mailed to the respondent 3-5 weeks after they respond through the appropriate mode. Part A, section A.9 of this submission provides a detailed description of this Choice-Plus incentive experiment.

*Language.* Households will receive a topical package in English or Spanish depending on the language they used to complete the screener. Households that prefer to complete the topical questionnaire in a different language from that used to complete the screener will be able to contact the Census Bureau to request the appropriate questionnaire. Households in the web mode group will be able to toggle between Spanish and English at any time in the web topical instruments.

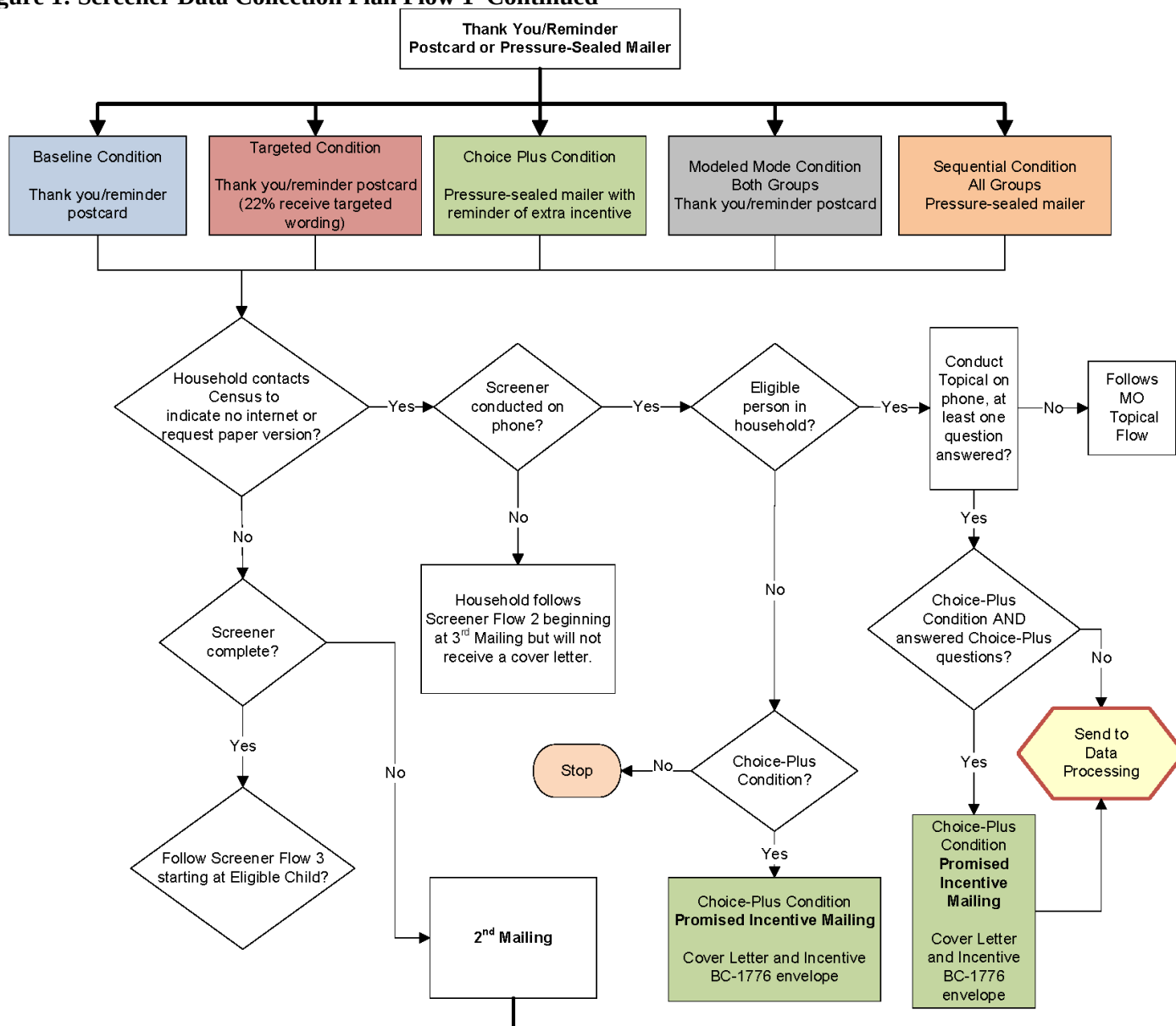
*Reminders and nonresponse follow-up.* Nonresponding households in the mail mode group will be mailed a second topical questionnaire approximately two weeks after the reminder postcard. If households that have been mailed a second topical questionnaire do not respond, a third package will be mailed by rush delivery (FedEx or USPS Priority Mail) approximately three weeks after the second mailing. Most households will be sent packages via FedEx delivery; only those households with a P.O. Box-only address will be sent USPS Priority Mail. If the household does not respond to the third mailing, a fourth mailing will be sent via USPS First Class mail. Figure 4 show the topical data collection plan for mail treatment households for the ECPP and PFI. Nonresponding households in the web mode groups will receive a second letter with the survey URL and log-in credentials two weeks after the reminder pressure sealed envelope. If they do not respond by web after the second mailing, they will be merged into the mail treatment data collection stream receiving a cover letter, paper questionnaire, and business reply envelope in the third (FedEx or USPS Priority Mail) and fourth mailings. Figure 5 shows the topical data collection plan for web mode households.

**Figure 1: Screener Data Collection Plan Flow 1**



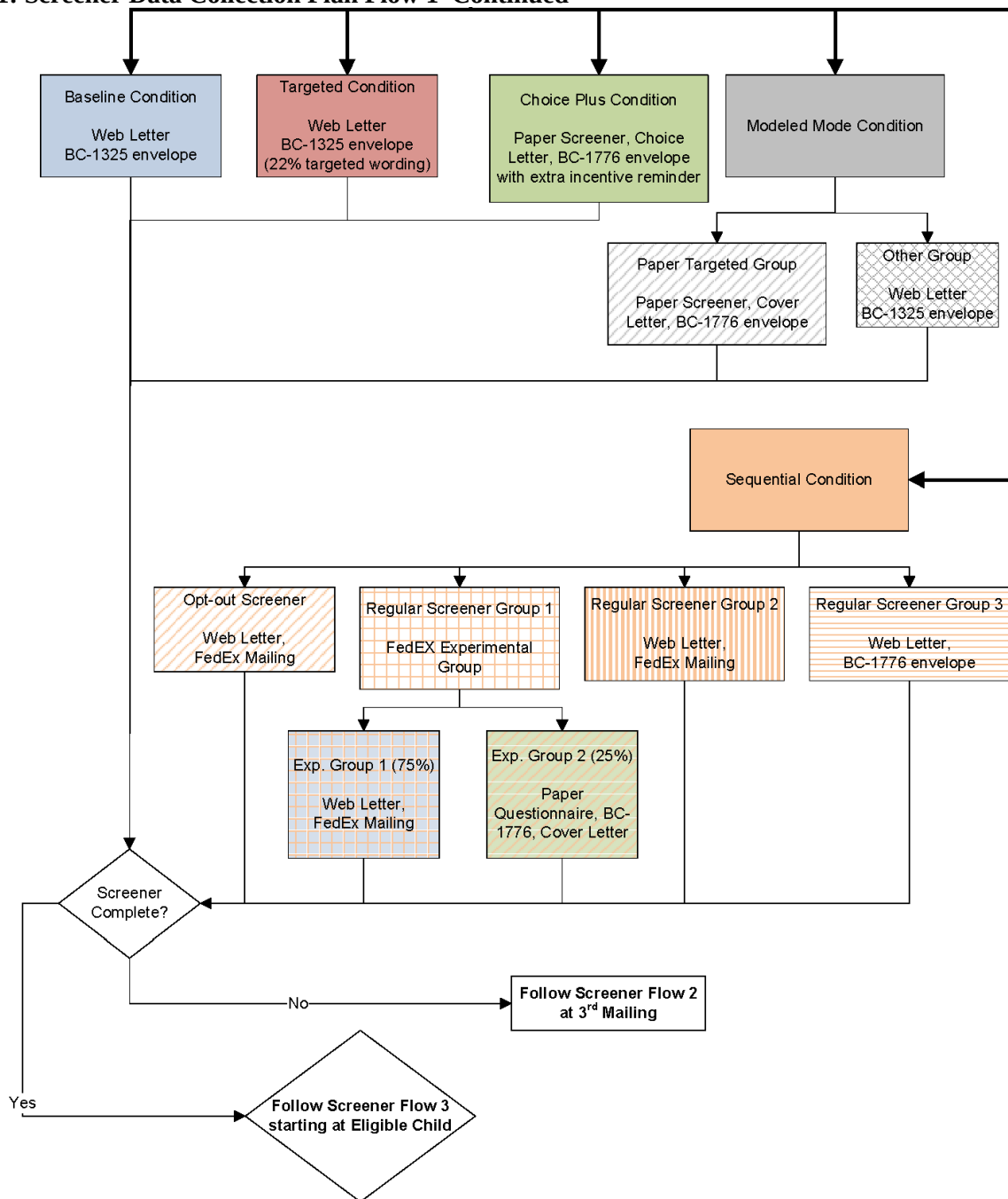
See continued flow chart on the next page.

**Figure 1: Screener Data Collection Plan Flow 1–Continued**



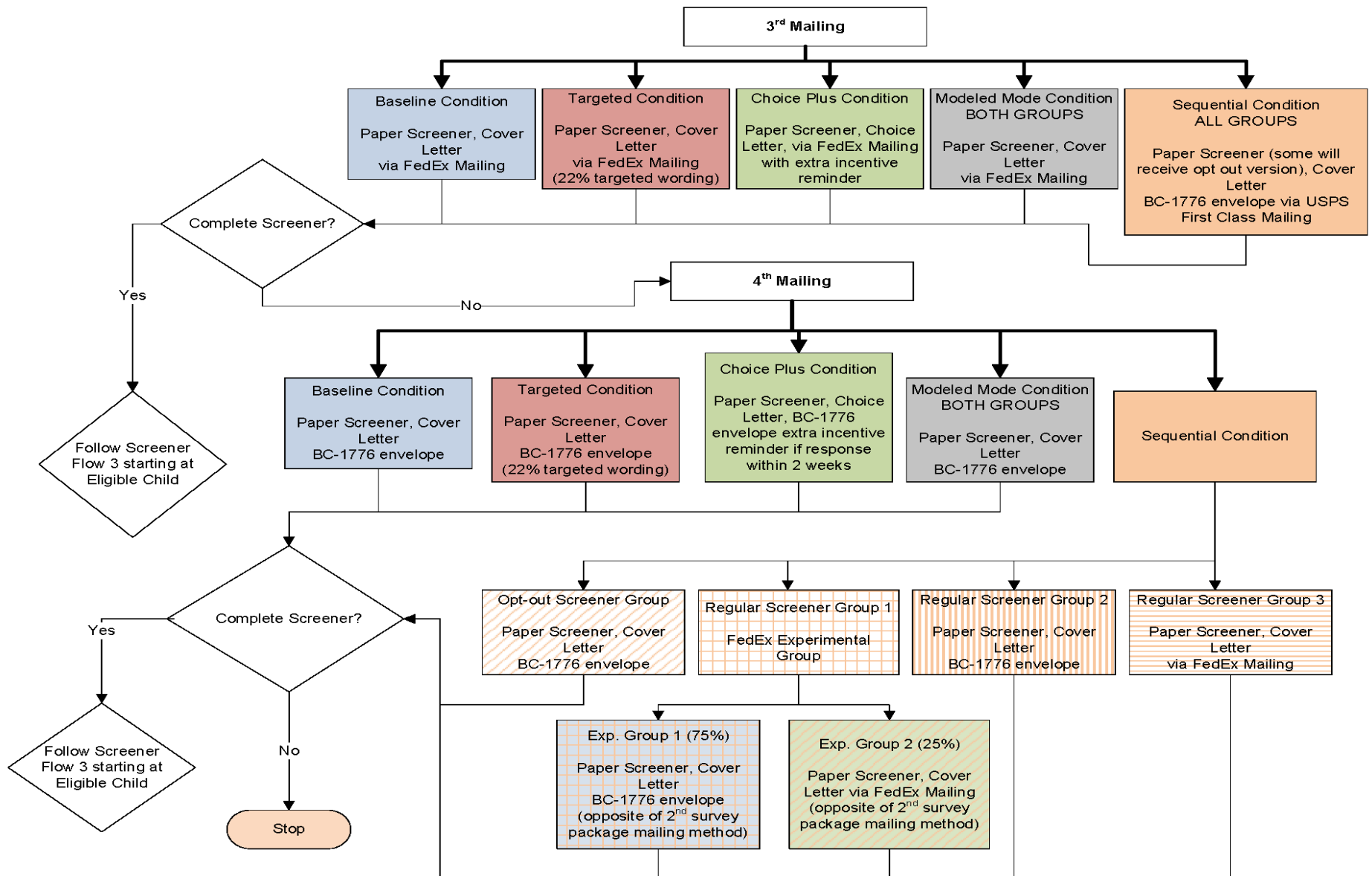
See continued flow chart on the next page.

**Figure 1: Screener Data Collection Plan Flow 1–Continued**

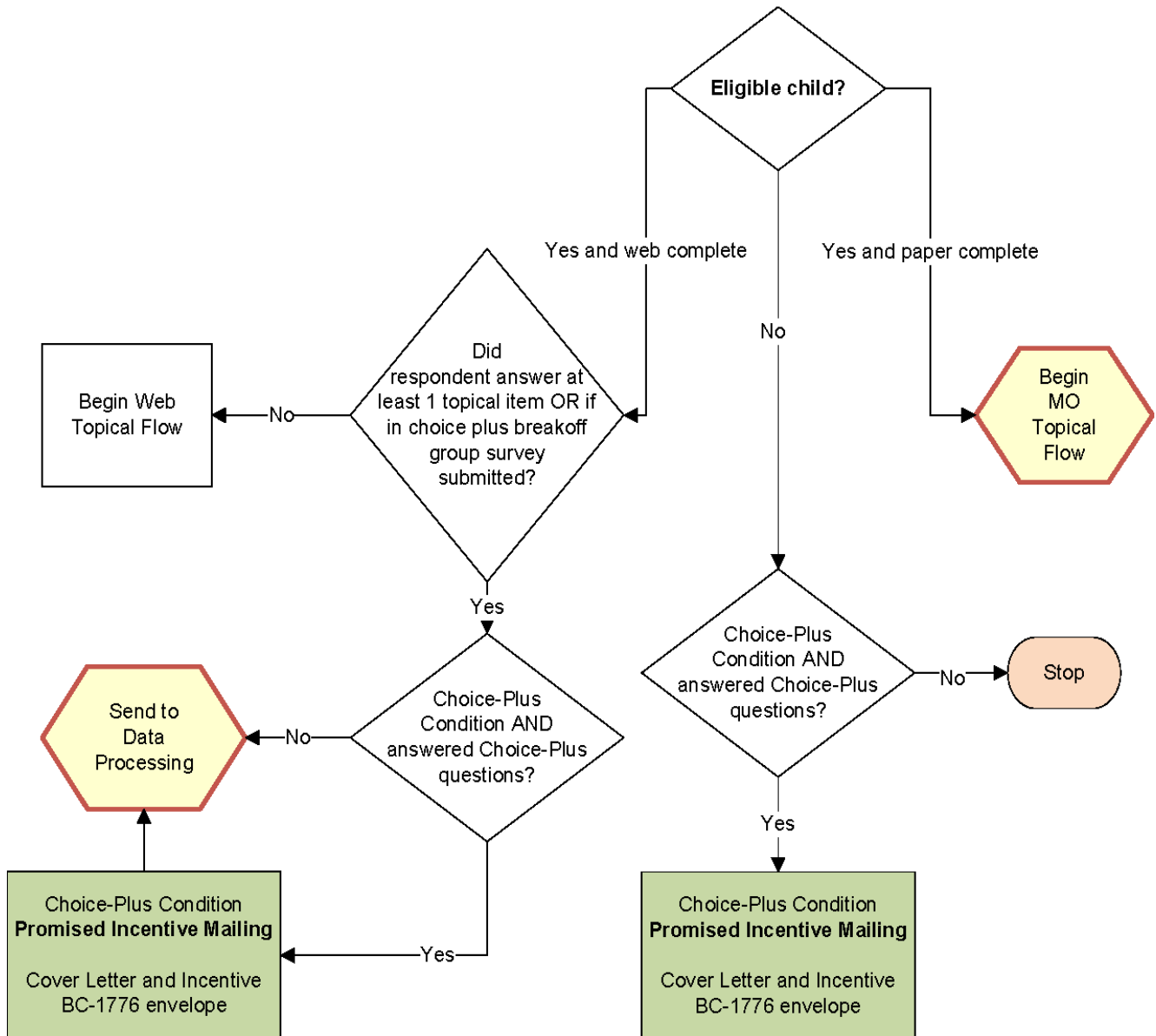




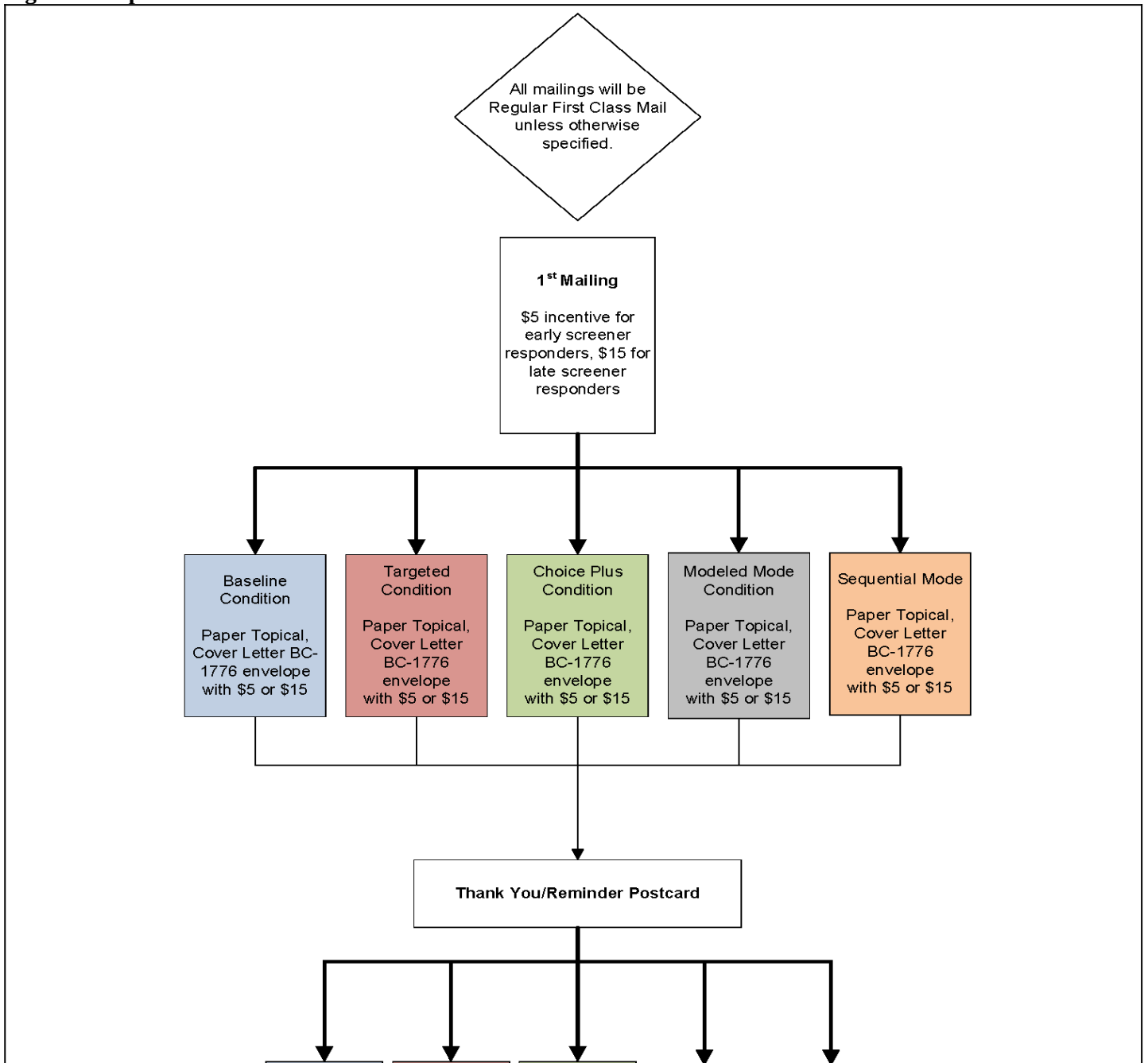
**Figure 2: Screener Data Collection Plan Flow 2**



**Figure 3: Screener Data Collection Plan Flow 3**

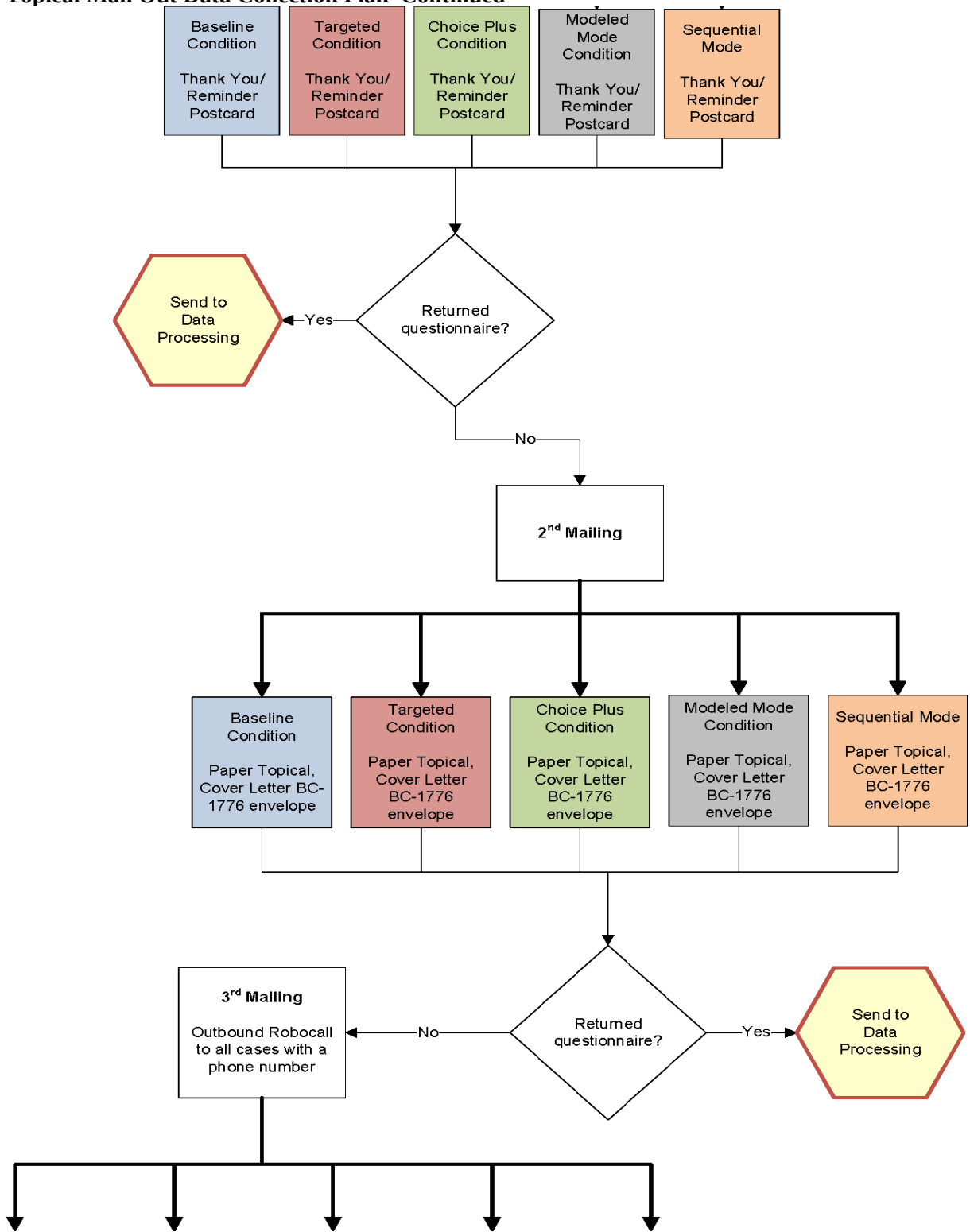


**Figure 4: Topical Mail Out Data Collection Plan**



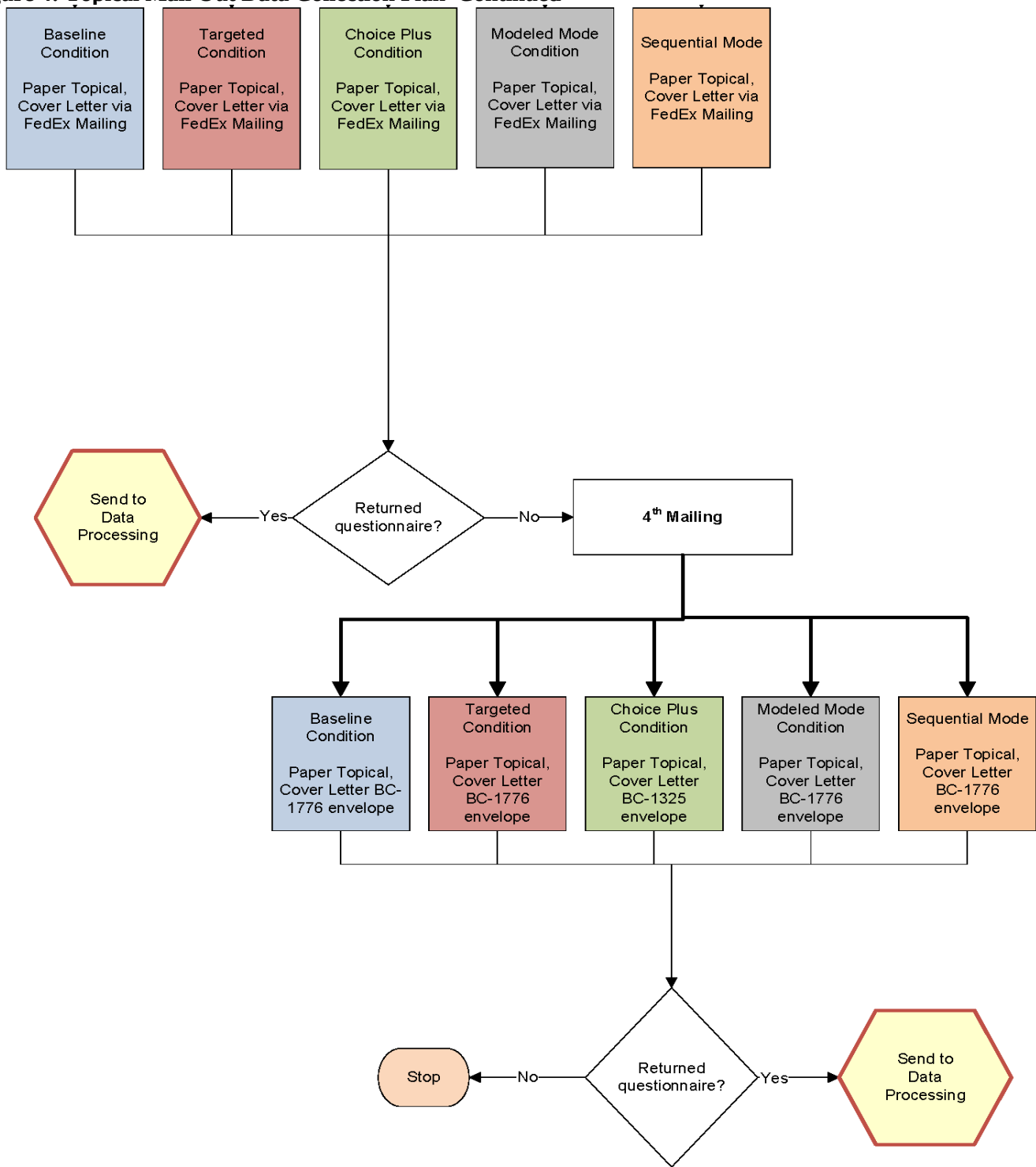
See continued flow chart on the next page.

**Figure 4: Topical Mail Out Data Collection Plan—Continued**

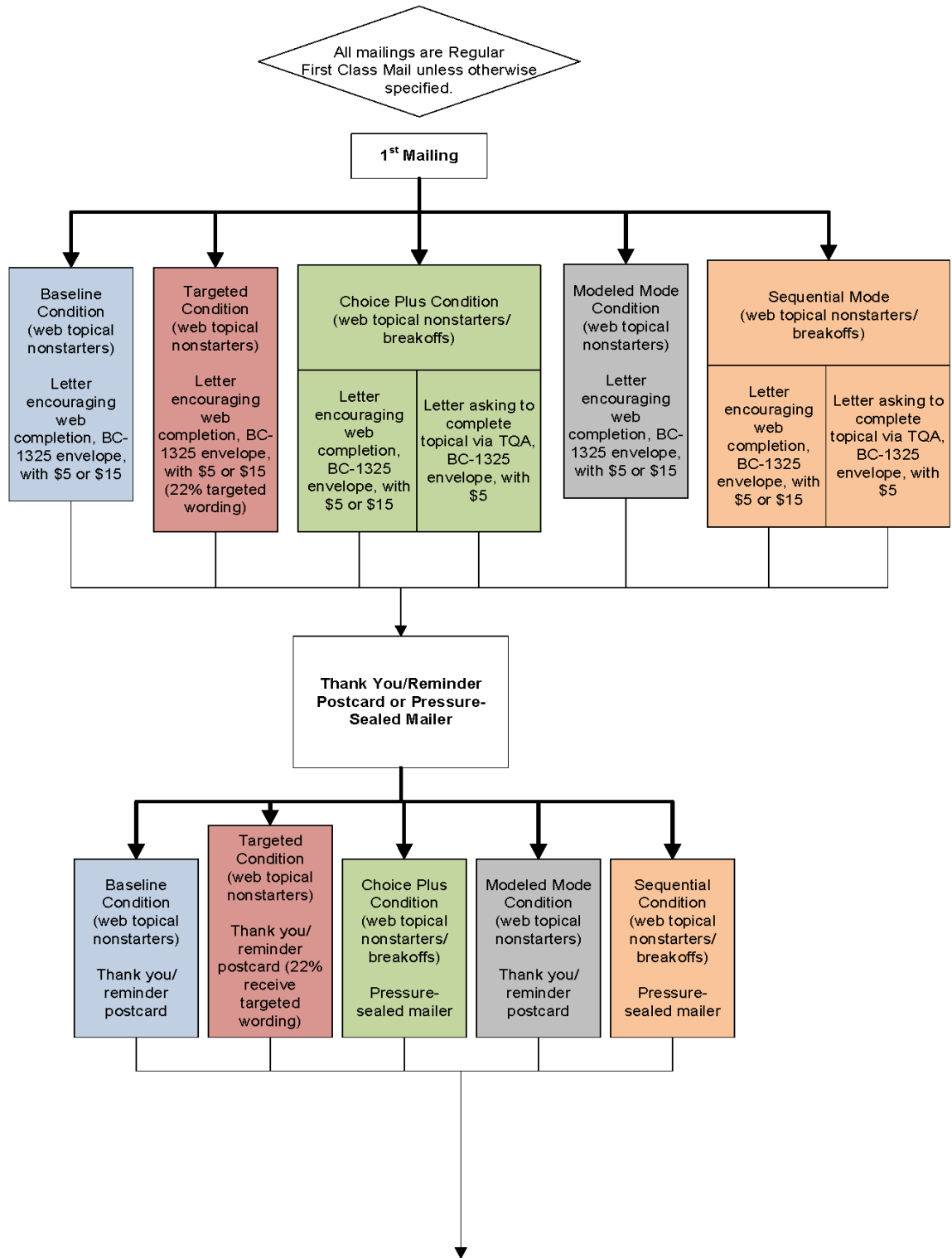


See continued flow chart on the next page.

**Figure 4: Topical Mail Out Data Collection Plan—Continued**

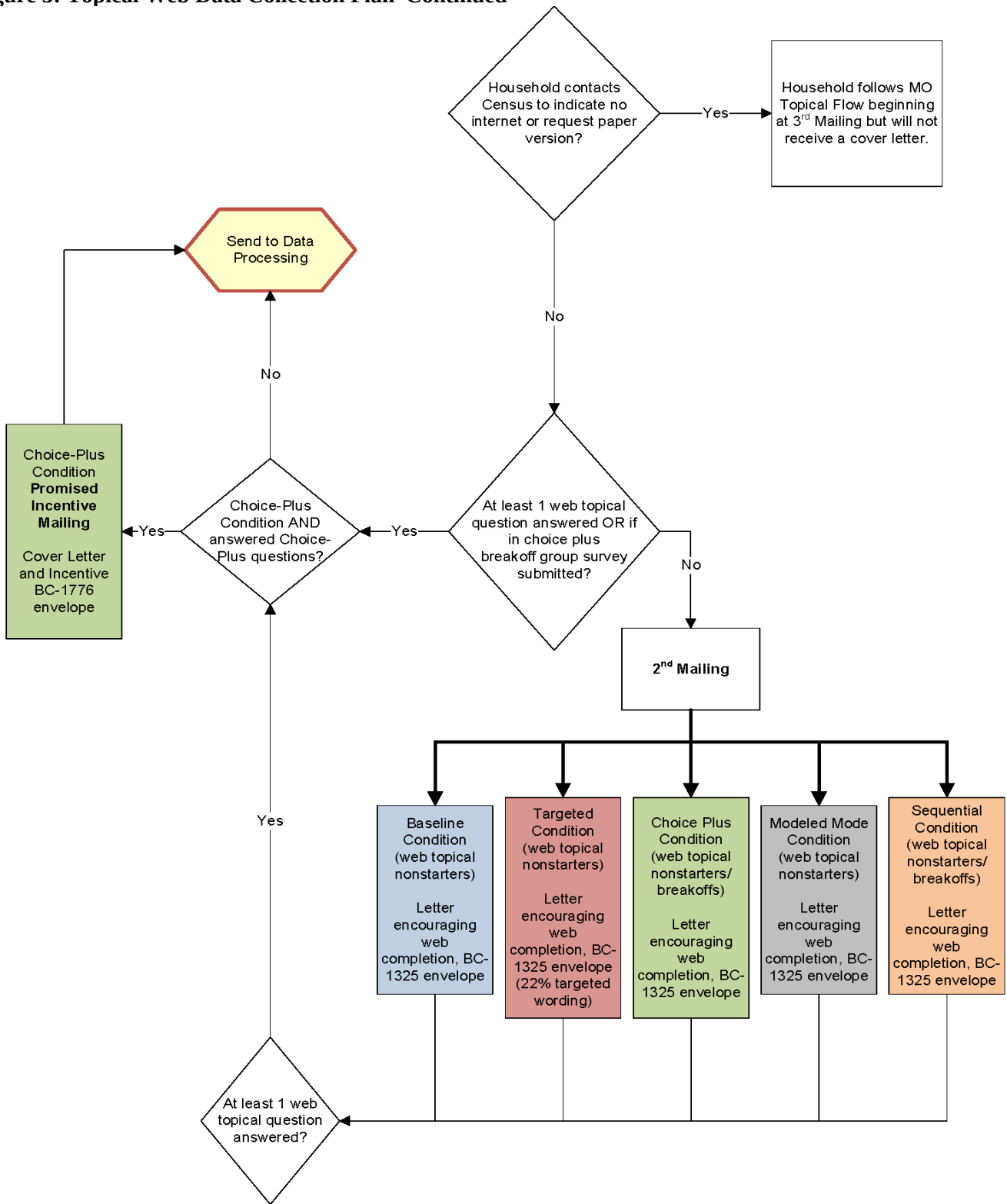


**Figure 5: Topical Web Data Collection Plan**



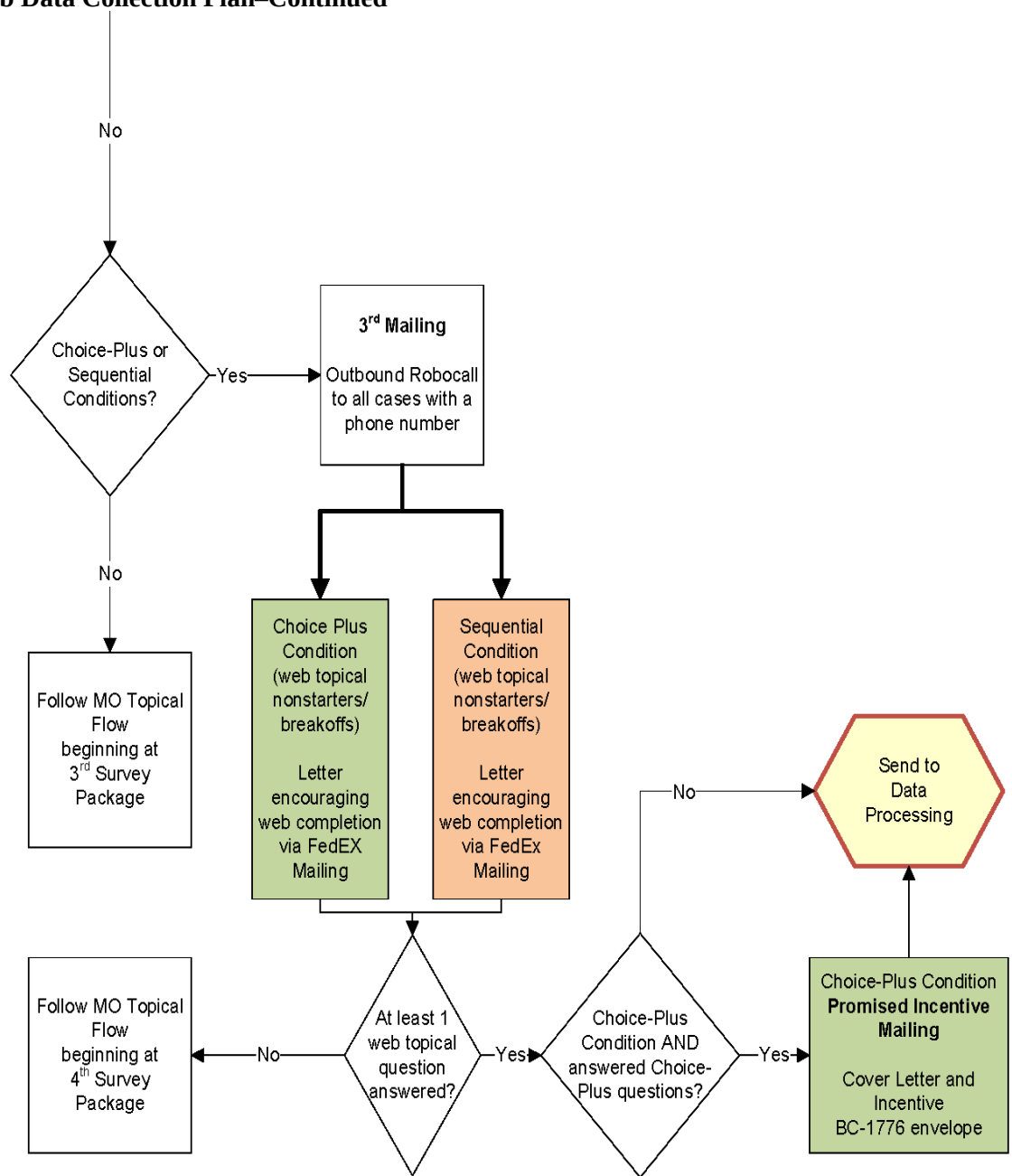
See continued flow chart on the next page.

**Figure 5: Topical Web Data Collection Plan–Continued**



See continued flow chart on the next page.

**Figure 5: Topical Web Data Collection Plan—Continued**





## Survey Monitoring

Mail survey returns will be processed upon receipt, and reports from the survey management system will be prepared at least weekly. The reports will be used to continually assess the progress of data collection.

### *In-person Study of NHES:2019 Nonresponding Households Sampling and Procedures*

The in-person study of nonresponding households will use structured and qualitative in-person interviews, as well as address and household observations to assess whether the existing theories surrounding household survey nonresponse are applicable to NHES. The study is designed to provide additional actionable information about how to combat the growing nonresponse problem in NHES and other federal government household surveys that use mail to contact sample members. It is expected that the results of this study will be used to improve the design of NHES:2022, with the goal of increasing the response rate and the representativeness of the respondents. Potential changes could include modifications to the nonresponse follow-up protocol, the type of incentives offered, and the presentation of contact materials.

This in-person study is described in detail in Appendix 4 of this submission. The contact materials, final interview protocols, and final sample selection details for this study will be submitted to OMB for review with a 30-day public comment period in fall 2018.

## B.3 Methods for Maximizing Response Rates

The NHES:2019 design incorporates several features to maximize response rates. This section discusses those features.

**Total Design Method/Respondent-Friendly Design.** Surveys that take advantage of respondent-friendly design have demonstrated increases in survey response (Dillman, Smyth, and Christian 2008; Dillman, Sinclair, and Clark 1993). We have honed the design of the NHES forms through multiple iterations of cognitive interviewing and field testing. These efforts have focused on the design and content of all respondent contact materials. In addition, in January 2018, an expert panel was convened at NCES to review contact materials center-wide. The NHES materials were reviewed, and the panel provided guidance for improvement to the letters that was applied to all NHES:2019 contact materials. Specifically, the panel advised that each letter convey one primary message, and that the sum total of all letters should constitute a campaign to encourage the sample member's response. This is in contrast to previous contact strategies which tended to reiterate the same information in the letter for each stage of nonresponse follow-up.

As noted previously, we will include a respondent incentive in the initial screener invitation mailing. Respondent incentives will also be used in the initial topical mailing where applicable. Many years of testing in NHES have shown the effectiveness of a small prepaid cash incentive on increasing response. The Census Bureau will maintain an email address and a TQA line to answer respondent questions or concerns. If a respondent chooses to provide their information to the TQA staff, staff will be able to collect the respondent's information on Internet-based screener and topical instruments. Additionally, the web respondent contact materials, questionnaires, and web data collection instrument contain frequently asked questions (FAQs) and contact information for the Project Officer.

**Engaging Respondent Interest and Cooperation.** The content of respondent letters and FAQs is focused on communicating the legitimacy and importance of the study. Experience has shown that the NHES child survey topics are salient to most parents. In NHES:2012, Census Bureau "branding" was experimentally tested against Department of Education branding. Response rates to Census Bureau branded questionnaires were higher compared to Department of Education branded questionnaires and was used in the NHES:2016 collection. NHES:2019 will continue to use Census Bureau-branded materials in the data collection.

Additionally, NCES is experimenting with materials targeted to Spanish speaking households in 2019. These materials will include images, language, and organizational endorsements designed to appeal to Spanish speakers. The endorsements that NCES will seek will be a subset of the following organizations: The Aspira Association; AVANCE; Congressional Hispanic Caucus Institute (CHCI); Hispanas Organized for Political Equality (HOPE); Hispanic Access Foundation; Hispanic Association of Colleges and Universities; Hispanic Federation; Hispanic Heritage Foundation; National Latino Education Institute; Hispanic Scholarship Fund; The Latino Congreso (of the William C. Velasquez Institute); The League of United Latin American Citizens (LULAC); ListoAmerica Inc.; MAES – Latinos in Science and Engineering; MANA – A National Latina Organization; National Alliance for Hispanic Health; National Association of Latino Elected and Appointed Officials (NALEO); National Hispanic Leadership Agenda; National Women, Infants and Children's program (WIC); Society for Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS); Southwest Voter Registration Education Project (SVREP); UnidosUS; and U.S. Hispanic Chamber of Commerce. Being able to achieve a higher response from Spanish-speaking households is

expected to reduce bias and increase the precision of NHES estimates.

**Nonresponse Follow-up.** The data collection protocol includes several stages of nonresponse follow-up at each phase. In addition to the number of contacts, changes in method (USPS First Class mail, FedEx, and automated reminder phone calls) and materials (letters, postcards, and pressure sealed envelopes) are designed to capture the attention of potential respondents.

## **B.4 Tests of Procedures and Methods**

NHES has a long history of testing materials, methods, and procedures to improve the quality of its data. Section B.4.1 describes the test that have most influenced the NHES design. The NHES:2016 and NHES:2017 web test experiments are described below. Section B.4.2 describes the proposed experiments for NHES:2019.

The following tests are planned for the NHES:2019 data collection:

- Targeted screener mailing experiment
- Sequential mixed mode contact strategies experiments
- Choice-plus incentive experiment
- Modeled mode experiment

### **B.4.1 Tests Influencing the Design of NHES:2019**

#### **NHES:2016 Full-Scale Collection Experiments**

Two experiments were designed in the NHES:2016 to determine methods to improve response rates and data quality: 1) the web test experiment and 2) incentive experiment. Each of these experiments is briefly described below, along with its results and implications for future NHES data collections.

- 1. Web test.** This experiment was designed to evaluate response rates for a subsample of respondents requested to complete the screener and topical instruments over the internet. From the original sample of 206,000 households, 35,000 were allocated to the web experiment. Additionally, a random half of the web respondents were asked for an email address for nonresponse followup to evaluate whether that request leads to breakoffs and invalid responses.

In the web test, the results showed a slight decrease in the overall screener response rate in the web treatment as compared to the paper version (62.1 percent web vs. 67.2 percent paper) but the overall topical response rate for those who completed on the web was much higher than those who completed on paper (PFI 93.3 percent web vs. 72.6 percent paper; ECPP 91.9 percent web vs. 71.5 percent paper). Also, a mode effects analysis was conducted to assess the prevalence and the extent of selection effects (whether the respondents in each mode differed on the characteristics measured by key topical survey questions) and measurement effects (whether the response mode affected how individuals responded to those key topical survey questions). Some evidence of both types of effects was found in the NHES:2016 data. However, the affected items were scattered throughout the topical questionnaires, the magnitude of most effects was small, and no clear patterns to the effects were found.

For the email experiment, results indicated that the majority of e-mail addresses were valid (about 1.6 percent of e-mails bounced back) but item response rates for providing an email address were lower than expected; about 76 percent of Adult Training and Education Survey (ATES) respondents and about 5 percent of child survey respondents broke off from completing the survey at the email item. As a result of these findings and the fact that the NHES:2017 web test indicated that zero completes were garnered by following up via email with web breakoffs, we determined that it is feasible to use web as the primary mode of completion in NHES but that asking for emails is not necessary in data collection. In NHES:2019, we plan to ask the majority of households by mailed correspondence to respond via web, then switching to a paper questionnaire at the third mailing. We will not request email addresses in NHES:2019 given that in the web test they did not help to improve response rates nor data quality.

- 2. Incentive experiment.** NHES:2016 also included a tailored incentive experiment designed to examine the effectiveness of leveraging auxiliary frame data to target lower screener incentives to households expected to be most likely to respond regardless of incentive amount, and higher screener incentives to households expected to be less likely to respond without an additional incentive. Households that were expected to be most likely to respond received \$0 with the NHES screener. Some households that were likely to respond received a \$2 screener incentive, and households that were predicted to be least likely to respond received \$10. The remaining households received the same \$5 screener incentive as the main NHES:2016 sample. The results related to the

actual effects of the specific tailored incentive were mixed. In general, the primary value of a tailored incentive protocol appears to be as a means of reducing incentive costs, while minimizing the risk of reducing response rates or increasing nonresponse bias. However, higher mailing costs then reverse the incentive cost savings. For example, the group that was modeled to receive \$0 responded at a rate of 82 percent. The comparable propensity group that was part of the \$2 control or the \$5 control responded at 87 percent, suggesting that NHES would lose 5 percentage points in response from NHES sample members modeled as “most likely to respond” by not sending any incentive to these cases (about 5 percent of the sample). Similarly, the group that was modeled to receive \$2 responded at a rate of 77 percent. The comparable propensity group that was part of the \$5 control group responded at a rate of 79 percent, suggesting that NHES would lose 2 percentage points in response from NHES sample members modeled as “likely to respond” by sending \$2 instead of \$5 to these cases (about 20 percent of the sample). Though these losses in response may be considered modest in comparison to the cost savings, further analyses demonstrated that higher incentives shifted the distribution of responses to earlier in the data collection, while lower incentives shifted the distribution to later in the data collection, so that the NHES program spent more money on additional nonresponse follow-up mailings for cases that received \$0 or \$2 relative to what would have been spent if these cases received \$5 with the first survey mailing. For example, among the cases modeled as “likely to respond,” the \$2 incentive reduced the return rate to the first NHES survey invitation mailing by about 3 percentage points, relative to \$5. Among the cases modeled as “most likely to respond,” the \$0 incentive reduced the return rate to the first NHES survey invitation mailing by about 8 percentage points. In neither of the groups that were modeled as “likely” or “most likely to respond” was the reduction in the return rate to the first NHES survey invitation mailing substantially offset by an increase in the return rate at the second mailing, which was the first nonresponse follow-up. Thus, the lower incentive increased the proportion of screener respondents that received the second nonresponse follow-up mailing, which was sent by FedEx (or U.S. Postal Service Priority mail for P.O. Box addresses) and was substantially more expensive. On the polar opposite end of the experiment, in which we hypothesized that sending \$10 to cases modeled as “least likely to respond” would raise response among those sample members, our hypothesis was only modestly supported. There was no statistically significant difference in response between the group that received \$10 and the comparable response propensity group that received \$5 (42 percent and 40 percent response, respectively). The difference in response rate between the group that received \$10 and the comparable group that received \$2 was only 4 percentage points (42 percent and 38 percent) despite an outlay of an additional \$8 per sample member. The results provided little evidence that tailoring noncontingent incentives can substantially increase response rates, or substantially reduce the risk of nonresponse bias. In particular, there was little evidence that key survey estimates were notably impacted by the noncontingent incentive protocol that was tested in this experiment. As a result of this experiment, in NHES:2019, we will not utilize a tailored incentive approach, and all households will receive a \$5 screener incentive.

An additional experiment was conducted in NHES:2016, where we used a seeded sample of certificate holders to evaluate new measures of non-degree credentials for the ATES topical survey. Since ATES will not be fielded in NHES:2019, results of this experiment are not included in this document.

### **NHES:2017 Web Test Experiments**

Four experiments were designed in the NHES:2017 Web Test to determine methods to increase the number of completed topical surveys and improve response rates and data quality: 1) the dual household sampling experiment; 2) incentive experiment; 3) envelope experiment; and 4) screener split-panel experiment. Each of these experiments is briefly described below, along with its results and implications for future NHES data collections. Detailed results of these experiments are provided in the Appendix 6 that is part of this submission.

- 1. Dual Household Sampling Experiment.** This experiment tested whether households with members eligible for two or more topical surveys can be asked to complete two topical surveys without negatively impacting topical response rates and data quality. Of the 97,500-household screener sample, 2/3 (65,000) were assigned no dual sampling; these households were sampled for no more than one topical survey, regardless of the number of surveys for which they have eligible members. The remaining 1/3 of households were assigned to dual sampling; these households were sampled for two topical surveys if they had members eligible for two or more. This experiment found that the dual-topical condition led to a significant decrease in the response rate for all topicals except PFI-Homeschool. The ECPP response rate was 86 percent in the dual-topical condition compared to the 89 percent response rate in the single-topical condition. The PFI-Enrolled response rate was 87 percent in the dual-topical condition compared to the 92 percent response rate in the single-topical condition. Though ATES will not be fielded in 2019, the results of the dual-topical experiment were most dramatic for that survey, wherein the

response rate was 60 percent in the dual-topical condition compared to the 68 percent response rate in the single-topical condition. Also, the topical survey being presented second in the dual-topical condition had a negative effect on its response rate (as compared to being presented first) for all topicals. The differences in response rates by order were as follows: for the ECPP, the response rate was 88 percent when offered first compared to 81 percent when offered second; for the PFI-Enrolled, the response rate was 90 percent when offered first compared to 80 percent when offered second; for the PFI-Homeschool survey, the response rate was 84 percent when offered first compared to 75 percent when offered second; and for the ATES, the response rate was 89 percent when offered first and the screener respondent was also the topical sample member and 76 percent when offered after a child survey to the screener respondent. Finally, the dual-topical condition led to significantly more breakoffs for PFI-Enrolled survey—13 percent compared to 11 percent—(but not for any of the other topicals). NHES:2019 will not have a dual household component because the NHES will only collect data on children in 2019 and the biggest gain from the dual household sampling was to receive information on both an adult and a child in the household. Due to operational and data processing complexities, such as creating a web tool that minimizes redundancy in household questions for two sampled children, the gain in achieving two child topical surveys does not outweigh the cost to collect and process these data.

2. **Incentive Experiment.** This experiment was designed to provide NCES with data on how pre-paid, non-contingent incentives affect response rates to web surveys. A random sample of 14,625 addresses received \$2 in the initial invitation letter. The remaining addresses received \$5 in the initial invitation letter. Results of this experiment showed that using a \$2 screener incentive resulted in a significant reduction in the screener response rate (41 percent compared to 44 percent) but had no statistically significant effect on the topical response rate. Also, the positive effect of the \$5 incentive over the \$2 incentive was greatest at the beginning of the administration – it was particularly effective at getting sample members to respond to one of the first two screener contacts. Given that the \$2 incentive had a negative impact on response rates in the NHES:2017 Web Test, NHES:2019 will use the \$5 incentive in the initial invitation letter.
3. **Envelope Experiment.** This experiment was designed to determine whether initial login rates to the web instrument are affected by whether the web letters are mailed using large (i.e. full size) or small (i.e. letter-size) envelopes. Of the 97,500-household screener sample, 92,625 were assigned to the large envelope and 4,875 to the small envelope. Also, to determine if another method other than FedEx would have a similar effect in increasing response rates, half of the sample was assigned to receive the third screener mailing in a FedEx envelope (standard NHES approach), while the other half was assigned to receive the mailing in a First Class cardboard priority mail envelope. Using a letter-size envelope instead of a full-size envelope for two of the screener mailings did not have a significant effect on the final screener response rate (43 percent response among sample members who received the large, full-size envelope and 43 percent response among sample members who received the small, letter-sized envelopes), or on the response rate after any specific mailing, and no statistically significant decrease on topical response rates. For the FedEx/First Class experiment, sending the final screener mailing using First Class instead of by FedEx resulted in a significant reduction in the screener response rate. The sample that received FedEx responded to the survey at a 45 percent response rate compared to the 42 percent response rate from the recipients of the First Class mailing. It did not have a significant effect on the topical response rates. As a result of this experiment, given the lower postage cost and lack of effect on the response rate, we plan to use for the majority of cases the small envelope (except when paper surveys are included in the mailing, in which cases we will continue to use the large envelope). Finally, NHES:2019 will continue to use FedEx instead of First Class priority envelopes.
4. **Screener Split-Sample Experiment.** This experiment was designed to determine which screener format would provide higher data quality than the other. Half of sampled households received the screener used in NHES:2016, which asks respondents to first indicate the number of people living in the household and then provide more detailed information person-by-person (e.g., all items for Person 1, all items for Person 2, and so on). The other half of the sampled households received a redesigned screener, which asks respondents to first list the names of all the individuals living in the household and then provide more detailed information item-by-item (e.g., date of birth for Person 1, date of birth for Person 2, sex for Person 1, sex for Person 2, and so on). Results from the experiment showed that the screener version did not have an impact on the screener response rate (43 percent in both conditions) or screener breakoff rate (3 percent in both conditions). It also did not have a significant effect on the topical response rates, with no statistically significant differences between topical response rates by topical and screener version. Screener version did not have a significant impact on the percentage of web screener respondent households that reported at least one household member eligible for at least one topical survey, but

households in the redesigned version were significantly more likely to report at least one household member eligible for ECPP or PFI-E, although the magnitude of these differences is again quite small and may not be of practical concern. About 10 percent of screener respondents to the 2016 version of the questionnaire reported an ECPP-eligible household member compared to 11 percent of the redesigned version. About 22 percent of screener respondents to the 2016 version of the questionnaire reported a PFI-eligible household member compared to 23 percent of respondents to the redesigned screener. For several other outcomes, the redesigned screener performed worse than the 2016 version among web screener respondents. Respondent households in the redesigned version:

- were more likely to have missing data for at least one household member for most screener items (0.5 percent missing name in 2016 version compared to 1.1 percent missing name in the redesigned version; 1.0 percent missing sex in the 2016 version compared to 1.6 percent missing sex in the redesigned version; 0.9 percent missing enrollment status in the 2016 version compared to 1.2 percent missing enrollment status in the redesigned version; and 0.4 missing grade in the 2016 version compared to 0.8 percent missing grade in the redesigned version);
- were more likely to provide an inconsistent response for at least one household member (2 percent in the 2016 version compared to 3 percent in the redesigned version);
- were more likely to have at least one household member receive an “unknown eligibility” sampling status (1 percent in the 2016 version and 2 percent in the redesigned version); and
- took longer on average to complete the screener (3.9 minutes to complete the 2016 version compared to 4.4 minutes to complete the redesigned version)

As a result of this experiment, for NHES:2019, we plan to combine the best functioning parts of both screener versions through asking for a list of the names of children that live in the household and then ask in a child-by-child format the characteristics of each child.

An additional experiment was conducted in the NHES:2017 Web Test only for the ATES. This was a split panel experiment to test revisions to wording and structure of some of the key ATES items. Since ATES will not be fielded in NHES:2019, results of that experiment are not included in this document.

#### **B.4.2 Tests Included in the Design of NHES:2019**

To address declining response rates among households in NHES:2016, NCES is proposing to test different contact strategies, modes of collection, and incentive protocols to increase response in the 2019 collection.

- 1. Targeted Screener Mailing Experiment.** As described in Part A, section A.2 of this submission and section B.2 of this document, NHES:2019 includes targeted screener mailings for likely Spanish-speaking households (based on information available on the sampling frame), which is expected to increase the response rate for those cases. If an acceptable response rate is achieved for those cases, continued targeting of materials for likely Spanish-speaking households will continue for data collection in the future. A total of 15,000 of the 205,000 screened households will be selected for the targeted screener mailings experiment, of which 3,300 cases will be predicted as likely Spanish-speaking and 11,700 cases will receive the baseline data collection protocol. With 80 percent power and a significance level of 0.05, the treatment group size of 15,000 will allow a minimum detectable difference in screener response rates between the targeted screener mailing treatment group and the baseline condition of approximately 5 percentage points. At the screener level, a minimum detectable difference in screener response rates between the number of respondents falling into the targeted group of likely Spanish-speaking households and the baseline condition will be 3 percentage points. The next section describes planned analyses for the results of this experiment.

#### **Planned analyses of the targeted screener mailing experiment**

- Does using targeted letter wording increase the final screener response rate for the groups that are targeted? This compares the overall screener response rates among Spanish-speaking households between two groups—that is, the percentage of Spanish-speaking households receiving targeted wording that return a screener, versus the percentage of Spanish-speaking households receiving regular wording that return a screener.
- Does using targeted letter wording increase the representation of Spanish-speaking households among screener respondents? This compares the percentage of screener respondents identified as Spanish-speaking between the two groups. Since targeted wording would be sent to population identified as likely Spanish-speaking households in the frame, and it is expected to increase Spanish-speaking participation, this

comparison examines any difference in the percentage of screener respondents who are identified as Spanish-speaking in the frame in the two groups.

- 2. Sequential Mixed Mode Contact Strategies Experiments:** As described in Part A, sections A.2 and section B.2 of this document, NHES:2019 will also include experiments designed to examine the effectiveness of various operations that may hold promise for increasing response to NHES:2019 and future collections. Two experimental designs will be included 1) an “opt-out” screener and cover letters; and 2) a 3x3 mailing experiment that will vary the type of advance contact received (advance mailer campaign vs. advance letter vs. no advance letter) and when a FedEx package is sent. The FedEx experiment includes a modeled FedEx timing treatment to determine which households should receive the FedEx for a second survey package and which for the fourth. A total of 80,000 households will be allocated to these experiments: 10,000 to the “opt-out” screener group, and 70,000 to the 3x3 mailing experiments.

The next section describes planned analyses for these experiments. Below each summary of the analysis we plan to perform, given its corresponding sample sizes, and with 80 percent power and a significance level of 0.05, we provide the minimum detectable difference in screener response rates and topical response rates as compared to the baseline condition or other conditions noted for the given analysis.

### **Planned analyses of the sequential mixed mode contact strategies experiment**

- Overall:
  - Does an “improved” sequential mixed mode design that takes into account lessons learned in NHES:2016 and NHES:2017 web test and utilizes new contact strategies increase the final screener response rate? This analysis will compare the overall screener response rates between the group receiving the advance mailing campaign and modeled FedEx, and the baseline group. The baseline group will receive a regular advance letter and a FedEx letter at the third mailing. This comparison examines whether the “improved” contact strategies of sending advance mailing campaign and targeting people who are more likely to respond after receiving FedEx at an earlier mailing will increase participation.
    - Effect of “improved” sequential mixed mode design on screener response rate will be detected at a minimum 3 percentage point change relative to the baseline condition.
- Opt-out screener group:
  - Does using opt-out screener increase final screener response rate? This will compare the overall screener response rates between the opt-out screener group and those who receive advance letter and FedEx at second mailing in the updated sequential mixed-mode experiment to evaluate the general impact of providing an opt-out option on response rates.
    - Effect of opt-out screener on screener response rate will be detected at a minimum 3 percentage point change relative to regular screener cases in updated sequential mixed mode condition that receive advance mailing and FedEx second.
  - Does using the opt-out screener increase final screener response rate among the households flagged as not having children? This analysis will compare the overall screener response rates among households without children between the opt-out screener group and those who received the advance letter and FedEx at second mailing.
    - Effect of opt-out screener on screener response rate for households flagged as having children on the frame will be detected at a minimum 5 percentage point change relative to equivalent regular screener cases in the updated sequential mixed mode condition that receive the advance mailing and FedEx second. The analysis will also be done for households not flagged as having children on the frame.
- Advance mailing experiment:
  - Does removing the advance letter lower the final screener response rate? We will analyze the effect of sending an advance letter relative to not sending the letter by examining the overall response rates between the two groups.

- Effect of adding advance mailing or removing advance letter on screener response rate, within the context of the updated sequential mixed mode protocol, will be detected at a minimum 3 percentage point change relative to cases in updated sequential mixed mode condition that just receive an advance letter. For this analysis the FedEx experiment conditions will be collapsed.
- Does adding an informational advance mailing campaign increase the final screener response rate? We will analyze the effect of sending an advance letter relative to sending out an advance mailing campaign by examining the overall response rates between the two groups.
  - Effect of adding advance mailing campaign on screener response rate, within the context of the updated sequential mixed mode protocol, will be detected at a minimum 3 percentage point change relative to cases in updated sequential mixed mode condition that just receive an advance letter. For this analysis the FedEx experiment conditions will be collapsed.

- FedEx timing/modeling experiment:

- Does sending the second survey package via FedEx increase the final screener response rate (versus sending it as the fourth survey package)? We will evaluate the overall screener response rates between sending FedEx at the second mailing and sending FedEx at the fourth mailing, excluding the cases in the modeled FedEx experiment.
  - Effect of sending FedEx second to all cases on screener response rate, within the context of the updated sequential mixed mode protocol, will be detected at a minimum 3 percentage point change relative to cases that receive FedEx fourth. For this analysis the advance mailing experiment conditions will be collapsed.
- Does waiting longer to send the FedEx survey package to the cases that advance modeling deems as being “FedEx-least-insensitive” decrease the final screener response rate for these kinds of cases compared to sending FedEx to these households as the second survey package? This question evaluates the overall screener response rate among households deemed “FedEx-least-insensitive” between those receiving FedEx at fourth package in the modeled FedEx condition, and those receiving FedEx at the second package. For the modeled FedEx condition, a model will be run to predict the increase in response propensity attributable to the use of FedEx. Each household will be put into quartiles based on the “sensitivity score” of responding to FedEx. The group of cases deemed as less sensitive to FedEx treatment, the “FedEx-least-insensitive” group, will be assigned to receive FedEx at the fourth package. This comparison examines whether sending this group at the fourth versus second mailing makes a difference for this specific subset of households.
  - Effect on screener response rate of waiting to send FedEx mailing fourth for those cases that modeling deems “FedEx-least-insensitive”, within the context of the updated sequential mixed mode protocol, will be detected at a minimum of 3 percentage point change relative to equivalent cases in “FedEx second” group. For this analysis the advance mailing experiment conditions will be collapsed.
- Does sending the second survey package via FedEx increase the percentage of screener respondents that respond online or using the helpdesk phone (the Census TQA) (versus sending it as the fourth survey package)? This question compares the percentage of screener respondent cases responding online or via TQA between those receiving FedEx at the second package and those receiving FedEx at the fourth package, to understand whether sending out FedEx sooner than later would encourage more online and TQA participants. The comparison excludes the cases in modeled FedEx experiment.
  - Effect of sending FedEx second to all cases on percentage of screener respondent cases that respond to the screener online/TQA, within the context of the updated sequential mixed mode protocol, will be detected at a minimum of 3 percentage point change relative to cases that receive FedEx fourth. For this analysis, the advance mailing experiment conditions will be collapsed.

**3. Choice-Plus Incentive Experiment:** As described in Part A, section A.2 and section B.2 of this document, NHES:2019 will include an experiment that will offer concurrent web and paper response options and will test two levels of promised incentives to encourage respondents to respond via the web or helpdesk phone (the Census TQA) instead of by paper. For 24,000 randomly assigned cases, respondents will be offered a \$10 promised

incentive for web or TQA completion and, for 6,000 randomly assigned cases, respondents will be offered a \$20 promised incentive for web or TQA completion.

The next section describes planned analyses for this experiment. Below each summary of the analysis we plan to perform, given the treatment group size of 24,000 - \$10 cases and 6,000 - \$20 cases, and with 80 percent power and a significance level of 0.05, we provide the minimum detectable difference in screener response rates and topical response rates as compared to the baseline condition or other conditions noted for the given analysis.

### **Planned analyses of the choice-plus incentive experiment**

- \$20 choice-plus incentive group:
    - Does the choice-plus condition increase the final screener response rate among cases offered a \$20 incentive? This analysis will compare overall screener response rates between the respondents offered a \$20 promised incentive for web completion (in addition to the \$5 prepaid incentive) and those in the baseline group (who receive the \$5 prepaid incentive but no promised incentive) or those in the \$10 group of the choice-plus condition.
      - Effect of choice-plus \$20 condition on screener response rate will be detected at a minimum of 3 percentage point change relative to the \$10 group of choice-plus condition OR relative to the baseline condition.
  - \$10 choice-plus incentive group:
    - Does the choice-plus condition increase the final screener response rate among cases offered a \$10 incentive? This compares overall response rates between the respondents offered a \$10 promised incentive for web completion and those in the baseline group.
      - Effect of choice-plus condition on screener response rate will be detected at a minimum of 3 percentage point change relative to the baseline condition.
    - Does the choice-plus condition increase the final screener response rate for each of the response propensity (RP) quartiles among cases offered a \$10 incentive? This analysis compares overall response rates within RP quartiles between the respondents offered a \$10 promised incentive for web completion and those in the baseline group. An RP model will be built based on the 2016 NHES web experiment data, and the RP score will be calculated for every household. Households will be grouped into four quartiles according to their RP score. Respondents in the same RP quartiles from the two incentive groups will be compared.
      - Effect of choice-plus condition on screener response rate for RP quartiles will be detected at a minimum of 3 percentage point change relative to equivalent cases in the baseline condition.
    - Does the choice-plus condition increase the percentage of respondent cases that complete the screener online/TQA? This will compare the percentage of screener respondent cases responding on the web or through TQA between the respondents offered a \$10 promised incentive for web completion and those in the baseline group.
      - Effect of choice-plus condition on percentage of screener respondent cases that complete the screener online/TQA will be detected at a minimum of 3 percentage point change relative to the baseline condition.
    - Does the choice-plus condition have an effect on the topical response rate? This analysis compares the final topical response rate between the respondents offered a \$10 promised incentive for web completion and those in the baseline group. Because the promised incentive will be provided after respondents complete the topical questionnaire, it is expected that the use of a promised incentive will motivate more completion.
      - Effect of choice-plus condition on topical response rate, across all topicals combined will be detected at a minimum of 5 percentage point change relative to the baseline condition.
- 4. Modeled Mode Experiment:** As described in Part A, section A.2 and section B.2 of this document, NHES:2019 will include an experiment that will test whether targeting some cases for a paper-only protocol increases the response rate among those cases. For a total of 40,000 of the 205,000 households, we will use a model predicting



response mode preference and data about the characteristics of the households from NHES:2016 to identify those that are most likely to respond by paper/least likely to respond by web. This model will be used to identify NHES:2019 households with high propensity for paper response and low propensity for web response. These households will be sent the paper questionnaire only and will not be offered a web version.

The next section describes planned analyses for this experiment. Below each summary of the analysis we plan to perform, given the treatment group size of 40,000, and with 80 percent power and a significance level of 0.05, we provide the minimum detectable difference in screener response rates and topical response rates as compared to the baseline condition or other conditions noted for the given analysis.

### **Planned analyses of the modeled mode experiment**

- Does sending paper-only to cases identified as being better for a paper-only protocol (vs. baseline mixed-mode) increase the final screener response rate for these cases? This analysis compares the overall screener response rates among paper-sensitive cases between the modeled mode group and the baseline group. The paper-sensitive cases are the respondents who are identified by the model to be the most likely to respond to a paper than a mixed-mode protocol.
  - Effect of modeled mode protocol on screener response rate among cases assigned to paper-only protocol will be detected at a minimum of 3 percentage point change relative to equivalent cases in baseline condition.
- Does sending paper-only to cases identified as being better for a paper-only protocol (vs. baseline mixed-mode) increase the final screener response rate? This will compare the overall screener response rates between modeled mode group and baseline group. The use of paper questionnaire and materials for a portion of the modeled mode group is expected to increase the screener response rate for people assigned to the modeled mode overall.
  - Effect of modeled mode protocol on screener response rate among cases assigned to paper-only protocol will be detected at a minimum of 3 percentage point change relative to equivalent cases in baseline condition.
- Does sending paper-only to cases identified as being better for a paper-only protocol (vs. baseline mixed-mode) increase the final topical response rate? This analysis compares the final topical response rates between modeled mode group and baseline group. The preliminary modeling work has suggested a lower topical response rate for paper-only protocol. Since paper questionnaire and materials will be used throughout the data collection period for a portion of the modeled mode group, it is expected that using a paper-only protocol would have a slightly lower response rate for people assigned to the modeled mode overall.
  - Effect of modeled mode protocol on topical response rate, for all topicals combined will be detected at a minimum of 5 percentage point change relative to baseline condition.
- Does sending paper-only to cases identified as being better for a paper-only protocol (vs. baseline mixed-mode) increase the final topical response rate for these cases? We will evaluate specifically the final topical response rates among paper-sensitive cases in the modeled mode and baseline group. We will evaluate the effect of sending paper questionnaires solely to this type of household, relative to sending them web survey materials.
  - Effect of modeled mode protocol on topical response rate among cases assigned to paper-only protocol, for all topicals combined will be detected at a minimum of 7.5 percentage point change relative to equivalent cases in baseline condition.

**5. In-Person Study of NHES:2019 Nonresponding Households:** As described in Part A, section A.2, and section B.2, the in-person study of nonresponding households will use structured and qualitative in-person interviews, as well as address and household observations to assess whether the existing theories surrounding household survey nonresponse are applicable to NHES. 230 completed structured interviews and 80 semi-structured, qualitative interviews will be completed with NHES nonrespondents. Additional detail is provided in Appendix 4. The final sample selection details and planned analyses for this study will be submitted to OMB for review with a 30-day public comment period in fall 2018.

## **B.5 Individuals Responsible for Study Design and Performance**

The following persons participated in the study design and are responsible for the collection and analysis of the NHES:2019 data: Sarah Grady and Andrew Zukerberg of NCES, and Shawna Cox and Christina Cox of Census.

### **REFERENCES**

- Dillman, D.A., Sinclair, M.D., and Clark, J.R. (1993). Effects of questionnaire length, respondent-friendly design, and difficult questions on response rates for occupant-addressed Census mail surveys. *Public Opinion Quarterly*, 57, 289-304.
- Dillman, D.A., Smyth, J.D., and Christian, L.M. (2008). *Internet, mail, and mixed mode surveys: The Tailored Design Method*. New York: Wiley.