

ICPSR 36498

## **Population Assessment of Tobacco and Health (PATH) Study [United States] Public-Use Files**

*United States Department of Health and  
Human Services. National Institutes of  
Health. National Institute on Drug Abuse*

*United States Department of Health and  
Human Services. Food and Drug  
Administration. Center for Tobacco  
Products*

Public-Use Files User Guide

## **Terms of Use**

The terms of use for this study can be found at:  
<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36498/terms>

## **Information about Copyrighted Content**

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

**NOTICE**  
**WARNING CONCERNING COPYRIGHT RESTRICTIONS**

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

**NOTE**

This document does not currently comply with 508 standards. If you need an alternative means of access to any information, please contact NAHDAP at [nahdap@icpsr.umich.edu](mailto:nahdap@icpsr.umich.edu). Let us know the nature of your accessibility problem, the Web address of the required information, and your contact information.



*A collaboration between the NIH and FDA*

## PATH Study Public Use Files User Guide

**August 2, 2024**  
**Updated for Wave 7**

Contract Number: HHSN271201100027C  
Reference Number: N01DA-11-556

Contract Number: HHSN271201600001C  
Reference Number: N01DA-16-5582

**Submitted to:**

Heather Kimmel, PhD, PATH Study COR  
Joshua Lazarus, NIDA Contracting Officer  
National Institute on Drug Abuse  
National Institutes of Health  
Office of Acquisitions  
3WFN MSC 6012  
301 N Stonestreet Ave.  
Bethesda, Maryland 20892

**Submitted by:**

Westat  
An Employee-Owned Research Corporation®  
1600 Research Boulevard  
Rockville, Maryland 20850-3129  
(301) 251-1500

## **Table of Contents**

<b><u>Chapter</u></b>		<b><u>Page</u></b>
1	Introduction .....	1
	1.1    Background of the PATH Study .....	1
	1.2    Public Use Data Contents.....	7
2	Sample Design.....	9
	2.1    Wave 1 Sample Design.....	10
	2.1.1    Target Population .....	10
	2.1.2    Multi-Stage Sampling .....	11
	2.2    Sample Design for Waves 2 and 3.....	16
	2.3    Wave 4 Sample Design.....	17
	2.3.1    Target Populations .....	17
	2.3.2    Replenishment Sampling.....	18
	2.4    Wave 4.5 Sample Design.....	19
	2.5    Wave 5 Sample Design.....	19
	2.6    Wave 5.5 Sample Design.....	20
	2.7    Adult Telephone Survey Sample Design.....	22
	2.8    Wave 6 Sample Design.....	23
	2.9    Wave 7 Sample Design.....	23
	2.9.1    Target Populations .....	24
	2.9.2    Replenishment Sampling.....	24
3	Data Collection .....	27
	3.1    Wave 1 Data Collection .....	30
	3.1.1    Advance Mail and Prepaid Incentives .....	30
	3.1.2    Household Screener .....	30
	3.1.3    Interview .....	31
	3.1.4    Biospecimen Collection .....	33
	3.2    Wave 2 Data Collection .....	34
	3.2.1    Advance Mail and Telephone Contact .....	35
	3.2.2    Interview .....	35

## **Contents (continued)**

<b><u>Chapter</u></b>		<b><u>Page</u></b>
3.3	Wave 3 Data Collection .....	38
3.3.1	Advance Mail and Telephone Contact .....	38
3.3.2	Interview .....	39
3.4	Wave 4 Data Collection .....	39
3.4.1	Advance Mail and Telephone Contact .....	40
3.4.2	Household Screener (Replenishment Sample) .....	40
3.4.3	Interview .....	41
3.5	Wave 4.5 Data Collection .....	41
3.5.1	Advance Mail and Telephone Contact .....	42
3.5.2	Interview .....	42
3.6	Wave 5 Data Collection .....	43
3.6.1	Advance Mail and Telephone Contact .....	43
3.6.2	Interview .....	44
3.7	Wave 5.5 Data Collection .....	44
3.7.1	Advance Mail and Telephone Contact .....	45
3.7.2	Interview .....	45
3.8	Adult Telephone Survey Data Collection.....	47
3.8.1	Advance Mail and Telephone Contact .....	47
3.8.2	Interview .....	47
3.9	Wave 6 Data Collection .....	48
3.9.1	Advance Mail and Telephone Contact .....	48
3.9.2	Interview .....	48
3.10	Wave 7 Data Collection .....	49
3.10.1	Advance Mail and Telephone Contact .....	50
3.10.2	Household Screener (Replenishment Sample) .....	50
3.10.3	Interview .....	51
3.10.4	Web Pilot Test.....	52

## **Contents (continued)**

<b><u>Chapter</u></b>		<b><u>Page</u></b>
4	Response Rates .....	53
	4.1    Wave 1 Response Rates .....	54
	4.2    Response Rates for Wave 2 Through Wave 7 for the Wave 1 Cohort .....	57
	4.3    Response Rates for the Wave 4 Replenishment Sample.....	62
	4.4    Response Rates for Wave 4.5 Through Wave 7 for the Wave 4 Cohort .....	64
	4.5    Response Rates for the Wave 7 Replenishment Sample.....	67
5	Weights and Imputation .....	70
	5.1    Adult and Youth Weights .....	70
	5.2    Variance Estimation .....	72
	5.2.1    Creation of Variables for Variance Estimation .....	77
	5.2.2    Software Options.....	78
	5.3    Imputation.....	79
	5.3.1    Wave 1 Imputation.....	79
	5.3.2    Wave 4 Imputation.....	79
	5.3.3    Wave 7 Imputation.....	80
6	Analytic Considerations.....	81
	6.1    Background .....	81
	6.2    Target Populations.....	83
	6.3    Selecting the Appropriate Weights.....	84
	6.3.1    Longitudinal Analyses Using Data from the First Wave of a Cohort in Estimating a Later Wave Outcome .....	89
	6.3.2    Longitudinal Analyses Using Data from Preceding Waves in Estimating a Later Wave Outcome .....	90
	6.3.3    Cross-Sectional Analyses Comparing Different (or Partially Overlapping) Sets of Persons Between Waves .....	91
	6.3.4    Cross-Sectional Analyses Describing Findings for a Single Wave Without Comparison to Other Waves.....	93
	6.3.5    Analyses Using PATH-ATS Data .....	93

## **Contents (continued)**

<b><u>Chapter</u></b>		<b><u>Page</u></b>
6.4	Impact of Data Collection Methods .....	95
6.5	Age Collection, Calculation, and Inconsistencies Across Waves.....	96
6.5.1	Collection of Age Information .....	97
6.5.2	Calculating Interview Ages.....	98
6.5.3	Potential for Age Inconsistencies Across Waves.....	99
6.5.4	Impact of Age Inconsistencies Across Waves .....	100
6.6	Considerations for Pooling Data .....	101
6.7	Considerations for Small Domains .....	102
7	Data Files .....	104
7.1	Data Structure.....	104
7.2	Record Identifier .....	105
7.3	Excluded Variables .....	105
7.4	Variable Names .....	106
7.5	Variable Labels .....	111
7.6	Value Labels.....	112
7.6.1	Interview Variables.....	112
7.6.2	Recoded Values.....	112
7.6.3	Values for Derived, Imputed, and External Variables.....	112
7.7	Missing Values.....	112
7.8	Outlier Values.....	115
8	Linking Files .....	116
8.1	Master Linkage Data File Structure.....	116
8.2	Linking Data Files Across Waves.....	119
8.3	File Merging Notes and Recommendations.....	120
8.4	Case Identification Number .....	120
	References.....	121

## **Contents (continued)**

<b><u>Appendices</u></b>		<b><u>Page</u></b>
A	PATH Study Protocol by Wave .....	123
B	Weighting Procedures .....	126
C	Example Program Code for Analysis .....	196
D	Imputation .....	223
E	Excluded Questionnaire Variables.....	238
F	Variables with Coded Outlier Values .....	413

## **Tables**

1-1	Expected ages of interviewed PATH Study participants by cohort and wave.....	5
1-2	PATH Study data collection dates and numbers of interviews, by wave .....	6
2-1	Wave 1 Cohort and Wave 4 Cohort sample sizes at the time the Wave 4 Cohort was formed .....	19
2-2	Wave 4 Cohort and Wave 7 Cohort sample sizes at the time the Wave 7 Cohort was formed .....	26
3-1	PATH Study data and collection methods by wave and whether biospecimens were collected.....	29
4-1	Wave 1 unweighted and weighted household screener response rates.....	57
4-2	Wave 1 unweighted and weighted adult interview response rates.....	57
4-3	Wave 1 unweighted and weighted youth interview and shadow youth response rates.....	57
4-4	Wave 2 unweighted and weighted interview response rates for the Wave 1 Cohort.....	59
4-5	Wave 3 unweighted and weighted interview response rates for the Wave 1 Cohort.....	59

## **Contents (continued)**

<b>Tables</b>		<b>Page</b>
4-6	Wave 4 unweighted and weighted interview response rates for the Wave 1 Cohort .....	60
4-7	Wave 4.5 unweighted and weighted interview response rates for the Wave 1 Cohort .....	60
4-8	Wave 5 unweighted and weighted interview response rates for the Wave 1 Cohort .....	61
4-9	Wave 6 unweighted and weighted interview response rates for the Wave 1 Cohort .....	62
4-10	Wave 7 unweighted and weighted interview response rates for the Wave 1 Cohort .....	62
4-11	Unweighted and weighted household screener response rates for the Wave 4 replenishment samples .....	64
4-12	Wave 4 unweighted and weighted adult interview response rates for the Wave 4 AYS replenishment subsample.....	64
4-13	Wave 4 unweighted and weighted youth interview response rates for the Wave 4 AYS replenishment subsample.....	64
4-14	Wave 4 unweighted and weighted shadow youth response rates for the entire Wave 4 replenishment sample .....	64
4-15	Wave 4.5 unweighted and weighted interview response rates for the Wave 4 Cohort .....	65
4-16	Wave 5 unweighted and weighted interview response rates for the Wave 4 Cohort .....	65
4-17	Wave 5.5 unweighted and weighted interview response rates for the Wave 4 Cohort .....	66
4-18	PATH-ATS unweighted and weighted response rates for the Wave 4 Cohort .....	66
4-19	Wave 6 unweighted and weighted interview response rates for the Wave 4 Cohort .....	67
4-20	Wave 7 unweighted and weighted interview response rates for the Wave 4 Cohort .....	67

## **Contents (continued)**

<b><u>Tables</u></b>		<b><u>Page</u></b>
4-21	Unweighted and weighted household screener response rates for the Wave 7 replenishment samples.....	68
4-22	Wave 7 unweighted and weighted adult interview response rates for the Wave 7 AYS replenishment subsample.....	69
4-23	Wave 7 unweighted and weighted youth interview and shadow youth response rates for the entire Wave 7 replenishment sample.....	69
5-1	Full sample and replicate weight variables for each wave, cohort, weight type, and interview type .....	72
6-1	Descriptions of weights available for each wave and cohort.....	85
7-1	Weight variable name mnemonic components.....	109
7-2	Tobacco-use variable name mnemonic components .....	110
7-3	PATH Study missing value codes and their descriptions .....	113
8-1	Possible response scenarios when the variable WAVE1_INTERVIEW is compared to variables added for subsequent waves.....	117
8-2	Example response scenarios when the variables RECRUITMENT_WAVE, WAVE1_COHORT, WAVE4_COHORT, and WAVE4_INTERVIEW for the PATH Study sample at Wave 4 are compared.....	118
8-3	Example response scenarios when the variables RECRUITMENT_WAVE, WAVE1_COHORT, WAVE4_COHORT, WAVE7_COHORT, and WAVE7_INTERVIEW for the PATH Study sample at Wave 7 are compared.....	119
A-1	PATH Study Protocol by Wave for Waves 1 – 7 .....	123
B-1	Trimming cells used in creating the Wave 4 final weights for the Wave 4 Cohort.....	155
E-1	Wave 1 questionnaire variables entirely excluded from the PUFs .....	238
E-2	Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable.....	243

## **Contents (continued)**

<b><u>Tables</u></b>		<b><u>Page</u></b>
E-3	Wave 2 questionnaire variables entirely excluded from the PUFs .....	256
E-4	Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable.....	263
E-5	Wave 3 questionnaire variables entirely excluded from the PUFs .....	271
E-6	Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable.....	278
E-7	Wave 4 questionnaire variables entirely excluded from the PUFs .....	287
E-8	Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable.....	296
E-9	Wave 4.5 questionnaire variables entirely excluded from the PUFs.....	309
E-10	Wave 4.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable .....	314
E-11	Wave 5 questionnaire variables entirely excluded from the PUFs .....	318
E-12	Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable.....	328
E-13	Wave 5.5 questionnaire variables entirely excluded from the PUFs.....	338
E-14	Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable .....	348
E-15	PATH-ATS questionnaire variables entirely excluded from the PUFs.....	358
E-16	PATH-ATS questionnaire variables excluded from the PUFs that are replaced with a derived variable .....	361
E-17	Wave 6 questionnaire variables entirely excluded from the PUFs .....	362
E-18	Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable.....	375
E-19	Wave 7 questionnaire variables entirely excluded from the PUFs .....	386

## **Contents (continued)**

<b><u>Tables</u></b>		<b><u>Page</u></b>
E-20	Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable.....	401
F-1	Wave 1 questionnaire variables with coded outlier values.....	413
F-2	Wave 2 questionnaire variables with coded outlier values.....	418
F-3	Wave 3 questionnaire variables with coded outlier values.....	421
F-4	Wave 4 questionnaire variables with coded outlier values.....	424
F-5	Wave 4.5 questionnaire variables with coded outlier values .....	429
F-6	Wave 5 questionnaire variables with coded outlier values.....	430
F-7	Wave 5.5 questionnaire variables with coded outlier values .....	434
F-8	PATH-ATS questionnaire variables with coded outlier values .....	436
F-9	Wave 6 questionnaire variables with coded outlier values.....	438
F-10	Wave 7 questionnaire variables with coded outlier values.....	442

<b><u>Figures</u></b>		
1	Illustration of the relationship between two successive PATH Study Cohorts.....	4
2	Illustration of the sample sizes for the Wave 1 Cohort, Wave 4 Cohort, and Wave 7 Cohort.....	10
3	Flowchart providing guidance on which weights to use for analyses of PATH Study data .....	88
4	PATH Study variable name components.....	106
5	Example of question and response options.....	111
6	PATH Study variable labeling convention.....	111

# 1. Introduction

---

This user guide describes the data files included in the Public Use Files (PUFs) for the Population Assessment of Tobacco and Health (PATH) Study. It includes information about the sample design, data collection methods, response rates, and weighting and imputation methods; it also guides data users on the variance estimation of complex survey data. The PUFs contain extensive data from the adult interview, youth and parent interviews. Every effort is made to protect the identity of individual respondents. The process of preparing the files for release includes a formal disclosure risk analysis. Any information from the data files that might be used to directly identify respondents is removed and some edits have been made for purposes of confidentiality. Additionally, for a small percentage of cases, values have been swapped across cases with similar characteristics to make it very difficult to identify a respondent with certainty. The modifications used to reduce the likelihood that any respondent could be identified in the data do not affect the overall data quality. However, they may result in estimates that differ slightly from those created using the PATH Study's Restricted Use Files.

In this document, the term "participants" is used to indicate people who agreed to be part of the PATH Study, whether or not they completed an interview in a particular wave. The term "respondents" is used to indicate the subset of people who completed an interview or, in the case of shadow youth, verified their information with the study in a particular wave; where needed for clarity, this term may be qualified to indicate the type of respondent (e.g., "shadow youth respondent," "adult interview respondent").

## 1.1 Background of the PATH Study

On June 22, 2009, the Family Smoking Prevention and Tobacco Control Act (TCA) was signed into law. The TCA gave the Food and Drug Administration's (FDA's) Center for Tobacco Products (CTP) regulatory authority over the manufacture, marketing, and distribution of tobacco products with the goal of protecting the nation's health. This authority allows the CTP to:

- Develop product standards, e.g., regulate levels of nicotine and other ingredients in tobacco products;
- Restrict tobacco marketing and sales to youth;

- Mandate labeling of tobacco products including health warnings on packaging and in advertisements;
- Mandate warning labeling on smokeless tobacco products;
- Ensure “modified risk” claims are supported by scientific evidence;
- Require disclosure of ingredients in tobacco products; and
- Preserve state, local, and tribal authority over tobacco regulation.

In contrast to other centers within the FDA that review therapeutic products using a “safe and effective” standard, the CTP regulates tobacco products using a standard based on population health. This standard allows the CTP to weigh the potential benefits and harm to persons who currently, formerly, or never used tobacco products when instituting regulations.

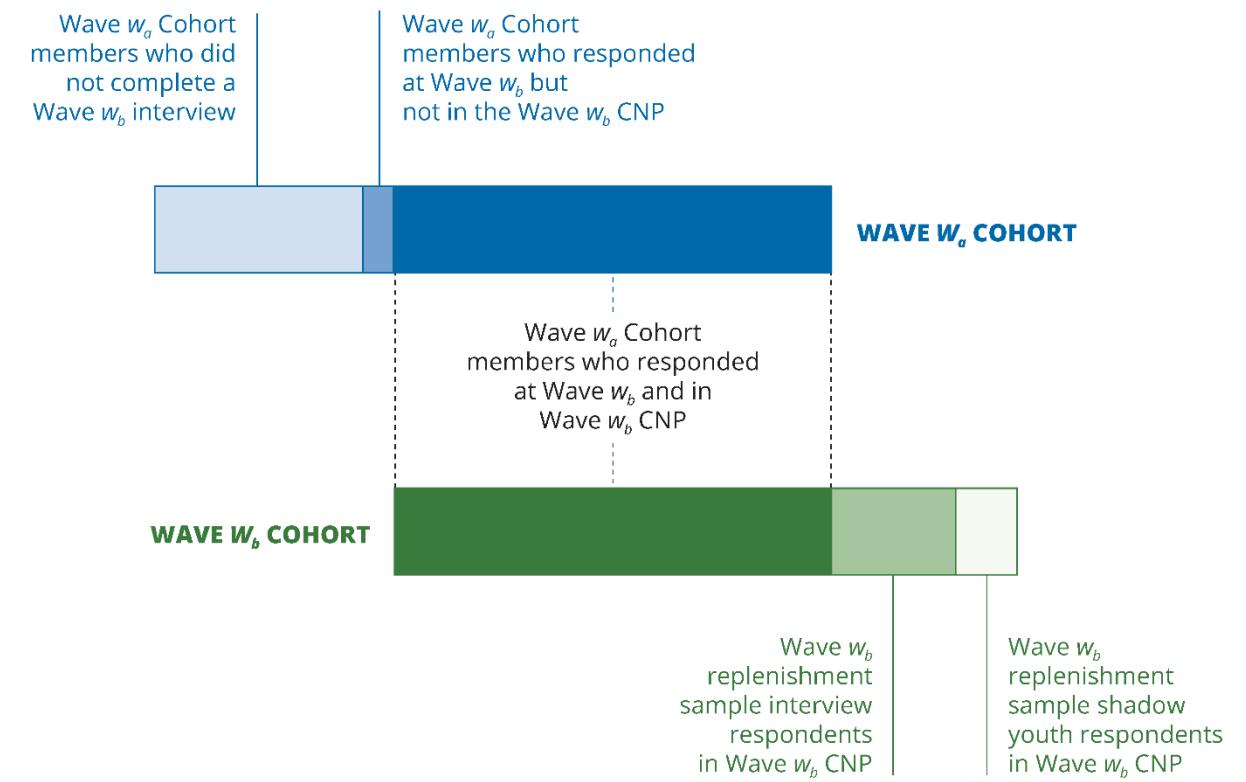
The PATH Study is a nationally representative, longitudinal cohort study of tobacco use and how it affects the health of people in the United States. The PATH Study is the first large joint research effort on this topic by the National Institutes of Health (NIH) and FDA since the 2009 TCA authorized the FDA to regulate tobacco products. The PATH Study will provide an empirical evidence base for developing, implementing, and evaluating regulations governing tobacco products by measuring the behavioral and health effects associated with changes in such regulations. The PATH Study interviews persons (whether they use or do not use tobacco) about multiple tobacco products, including cigarettes, e-cigarettes and other electronic nicotine delivery systems, traditional cigars, cigarillos, filtered cigars, smokeless tobacco, snus pouches, pipe tobacco, hookah, dissolvable tobacco (and for youth, bidis, kreteks), and other new/emerging products when applicable such as nicotine pouches and other oral nicotine products. Biomarkers of exposure and potential harm are measured in the blood and urine specimens of a subsample of respondents; these biomarker data can be analyzed with associated data on any tobacco product use, including product use for the 3 days before biospecimen collection.

A four-stage stratified area probability sample design was used in Wave 1 of the PATH Study to select adults ages 18 and older and youth ages 12 to 17 from the U.S. civilian noninstitutionalized population (CNP); an additional “shadow sample” of youth ages 9 to 11 was selected to be interviewed at later waves. At the first stage, a stratified sample of geographical primary sampling units (PSUs) was selected, in which a PSU was a county or group of counties. For the second stage,

within each selected PSU, smaller geographical segments were formed and then a sample of these segments was drawn. At the third stage, the sampling frame consisted of the residential addresses located in these segments. The fourth stage selected adults and youth from the sampled households identified at these addresses, with varying sampling rates for adults by age, race, and tobacco-use status. Adults were sampled in two phases: Phase 1 sampling used the information provided by one adult household member in the household screener and Phase 2 sampling used information that the sampled adult provided in the Phase 2 screener at the beginning of the adult interview. Parents of youth respondents do not constitute a separate sample. Parents who provide permission for their child to complete a youth interview are asked to complete a brief parent interview about their youth selected for the PATH Study. All Wave 1 sample participants form the Wave 1 Cohort.

At Wave 4, the Wave 1 Cohort was replenished with a probability sample of adults, youth, and shadow youth ages 10 and 11 selected from the CNP at the time of Wave 4. This sample was selected from residential addresses not selected for Wave 1 in the same sampled PSUs and segments using similar within-household sampling procedures. This “replenishment sample” was designed to supplement the Wave 1 sample. It was not intended to be used for estimation purposes on its own; rather, it was intended to be combined for estimation and analysis purposes with Wave 4 adult and youth respondents from the Wave 1 Cohort who were in the CNP at the time of Wave 4. This combined set of Wave 4 participants forms the Wave 4 Cohort. A more detailed description of the Wave 4 Cohort is in Chapter 2 (see Figure 1). Waves in which the PATH Study administered household screeners and sampled adults, youth, and shadow youth are herein referred to as recruitment waves.

At Wave 7, the sample was replenished a second time with a probability sample of adults, youth, and shadow youth ages 9 to 11 selected from the CNP at the time of Wave 7. This sample was selected from residential addresses not selected for Wave 1 or Wave 4 in the same sampled PSUs and segments, using similar within-household sampling procedures used at Wave 4. A new cohort was formed by combining this new sample with Wave 7 respondents in the Wave 4 Cohort who were in the CNP at the time of Wave 7. Waves in which the PATH Study administered household screeners and sampled adults, youth, and shadow youth are herein referred to as recruitment waves. Figure 1 is an illustration of the relationship between two successive PATH Study cohorts;  $w_a$  and  $w_b$  denote the two successive recruitment waves (where  $a=1, b=4$  or  $a=4, b=7$ ). More detailed descriptions of the PATH Study cohorts are in Chapter 2.

**Figure 1.** Illustration of the relationship between two successive PATH Study Cohorts

Computer-assisted personal interviewing (CAPI) was used for the household screener in Wave 1, Wave 4, and Wave 7, and for parent interviews in each wave. Audio computer-assisted self-interviewing (ACASI) was used for adult and youth interviews.

Data collections were conducted each year following the creation of the Wave 1 Cohort. Table 1-1 provides the expected ages of interviewed PATH Study participants by cohort and wave. See Section 6.5 for the calculation of interview ages. Wave 2 and Wave 3 were the first two annual follow-up waves for the Wave 1 Cohort. Wave 4 was the third annual follow-up wave for the Wave 1 Cohort, and the first wave for the Wave 4 Cohort. These data collections, which included all available sample youth and adults, are referred to as primary waves. Wave 4.5 was a special data collection for youth ages 12 to 17 at the time of the interview; Wave 5 was a primary wave including participants ages 12 and older at the time of the interview. Wave 5.5, conducted in 2020, was a special data collection for Wave 4 Cohort youth and young adults ages 13 to 19 at the time of the Wave 5.5 interview. Due to the amount of overlap between the Wave 1 and Wave 4 Cohorts, data for Wave 1 Cohort youth and young adults ages 15 to 19 also are available for Wave 5.5.

**Table 1-1. Expected ages of interviewed PATH Study participants by cohort and wave**

Wave	Data collection year <sup>a</sup>	Expected ages of Interviewed participants <sup>b</sup>		
		Wave 1 Cohort	Wave 4 Cohort	Wave 7 Cohort
1	2014	12+		
2	2015	12+	N/A	
3	2016	12+		
4	2017	12+	12+	
4.5	2018	13-17	12-17	
5	2019	14+	12+	
5.5	2020 <sup>c</sup>	15-19	13-19	
PATH-ATS		20+	20+	
6	2021	16+	14+	
7	2022	17+	15+	12+

<sup>a</sup> Year in which all or most of the data collection occurred. See Table 1-2 for the specific data collection start and end dates.

<sup>b</sup> The age progression illustrated in the table assumes respondents have one birthday between annual interviews. Logically, a respondent with age z at Wave w is expected to have an age of z, z+1, or z+2 in an annual follow-up wave depending on the dates of the two interviews in relation to the respondent's date of birth. See Section 6.5 for more information on age collection, calculation, and inconsistencies across waves.

<sup>c</sup> Only Wave 4 Cohort members were asked to participate in Wave 5.5 and PATH-ATS; those in the Wave 1 Cohort, but not in the Wave 4 Cohort, were not approached for an interview. Data collection occurred in the latter portion of 2020.

Also in 2020, several Federal tobacco policies were enacted. Examples include the nationwide Tobacco 21 policy, flavor restrictions, and increased enforcement of youth marketing restrictions. At the same time, the COVID-19 pandemic changed the way people used and purchased tobacco, interrupted data collection for the PATH Study and other national surveys used to understand tobacco use and resulted in less tobacco-use data being collected in 2020 than in most years. To address the challenges posed by the COVID-19 pandemic and because some of the Federal policy activities potentially impacted adult tobacco use, a subsample of Wave 4 Cohort adults ages 20 and older on August 31, 2020 was surveyed via the PATH Study Adult Telephone Survey (PATH-ATS) to collect data for adults critical in understanding tobacco use in this population. (See Chapter 2 for details about the PATH-ATS sample design.)

Wave 6 was a primary data collection wave including all available sample youth and adults. As described earlier, Wave 7 was a recruitment wave and the first wave for the Wave 7 Cohort. Table 1-2 provides the data collection dates and numbers of interviews completed in each PATH Study wave and the PATH-ATS for which data are available.

**Table 1-2. PATH Study data collection dates and numbers of interviews, by wave**

Wave	Data collection		Interviews conducted <sup>a</sup>		
	Start date	End date	Adults	Youth	Parents of youth
1	September 12, 2013	December 14, 2014	32,320	13,651	13,588
2	October 23, 2014	October 30, 2015	28,362	12,172	12,129
3	October 19, 2015	October 23, 2016	28,148	11,814	11,807
4	December 1, 2016	January 3, 2018	33,822	14,798	14,709
4.5	December 1, 2017	December 1, 2018	NA	13,131	12,965
5	December 1, 2018	November 30, 2019	34,309	12,098	12,027
5.5	July 3, 2020 <sup>b</sup>	December 31, 2020	3,628	7,129	7,051
PATH-ATS	September 10, 2020	December 20, 2020	8,874	NA	NA
6	March 1, 2021	November 30, 2021	30,516	5,652	5,618
7	January 6, 2022	April 2, 2023	30,801	10,834	10,790

<sup>a</sup> For Wave 4 and subsequent waves, the number of interviews available for analysis in the Wave 1 and Wave 4 Cohorts will be smaller than these totals.

<sup>b</sup> Data collection for Wave 5.5 originally began in person on December 1, 2019, but was suspended on March 17, 2020 due to the COVID-19 pandemic. The released data for this wave were collected via telephone starting on July 3, 2020. See Chapter 2 for details.

In Wave 1, every respondent who completed an adult interview was asked to provide urine and blood specimens. A subset of these adults was asked to provide urine specimens in follow-up waves through Wave 5. In Wave 2 through Wave 5, participants completing an adult interview for the first time, including newly sampled adults in Wave 4, were also asked to provide urine and blood specimens. Additionally, participants completing a youth interview in Wave 4 were asked to provide a urine specimen. There was no biospecimen collection in Wave 4.5. In Wave 5, youth who provided urine at Wave 4 and participants completing the youth interview for the first time in Wave 4.5 or Wave 5 were asked to provide a urine specimen. There was no biospecimen collection in Wave 5.5, for the PATH-ATS, or in Wave 6. Biospecimen collection was resumed at Wave 7. Subsamples of adults interview respondents (including newly sampled adults) at Wave 7 were asked to provide urine and blood specimens; subsamples of youth interview respondents (excluding newly sampled youth) at Wave 7 were asked to provide a urine specimen.

The multi-wave design allows for the longitudinal assessment of patterns of use of tobacco products, tobacco exposures, health, and risks for disease. Other aspects being examined in the PATH Study include changes in awareness, knowledge, risk perceptions, and attitudes about current and newly emerging tobacco products.

## 1.2 Public Use Data Contents

The public use data consist of the following files that will be included for each wave and the PATH-ATS:

- **Adult data file:** A data file containing adult questionnaire responses, derived variables, corresponding weight variables (for Waves 1 and 2 only), and related external data (i.e., 2010 census data). (Note that no adult data are available for Wave 4.5.)
- **Youth/parent data file:** A data file containing youth questionnaire responses, parent questionnaire responses about the youth and the parent, derived variables, corresponding weight variables (for Waves 1 and 2 only), and related external data (i.e., 2010 census data).
- **Weight files:** Starting with Wave 3, a set of files containing weight variables, as there are multiple sets of weights available for analysis of the respective adult and youth data.
- **Ever/Never reference data file:** Also starting with Wave 3, a reference data file containing derived variables that simplify the definitions of tobacco-use variables in follow-up waves.
- **Master linkage data file:** A master linkage data file provides each participant's unique identifier (PERSONID) and indicators as to which file(s) in the list above contains data specific to that respondent within a wave and across waves.

The adult data file and the youth/parent data file have similar content for a given wave. They both include the questionnaire items from the interview and some derived variables such as tobacco-use definitions. The data files are accompanied by annotated instruments for each interview. The annotated instruments provide a link between the variables and the corresponding questions asked in each interview. These documents can serve as references for data users to interpret variables for analyses. (See Chapter 7 for details about the adult data files and youth/parent data files.) In addition to question wording and routing, users are encouraged to read the “Explanatory Notes” section of the appropriate instrument, as that section provides information critical to understanding the complex PATH Study instruments.

Chapter 2 of this user guide contains information about the PATH Study’s multi-stage sample design. Chapter 3 describes the data collection methods, and Chapter 4 presents the response rates. Chapter 5 discusses weighting procedures, imputation, and variance estimation. Chapter 6 discusses analytic considerations, including selecting the appropriate weights for analysis. Chapter 7 describes the structure of the adult data files and youth/parent data files, including the record identifier,

## PATH Study Public Use Files User Guide

variable naming convention, and variable values and labels. Chapter 8 provides information about linking individual data files included in the release.

## 2. Sample Design

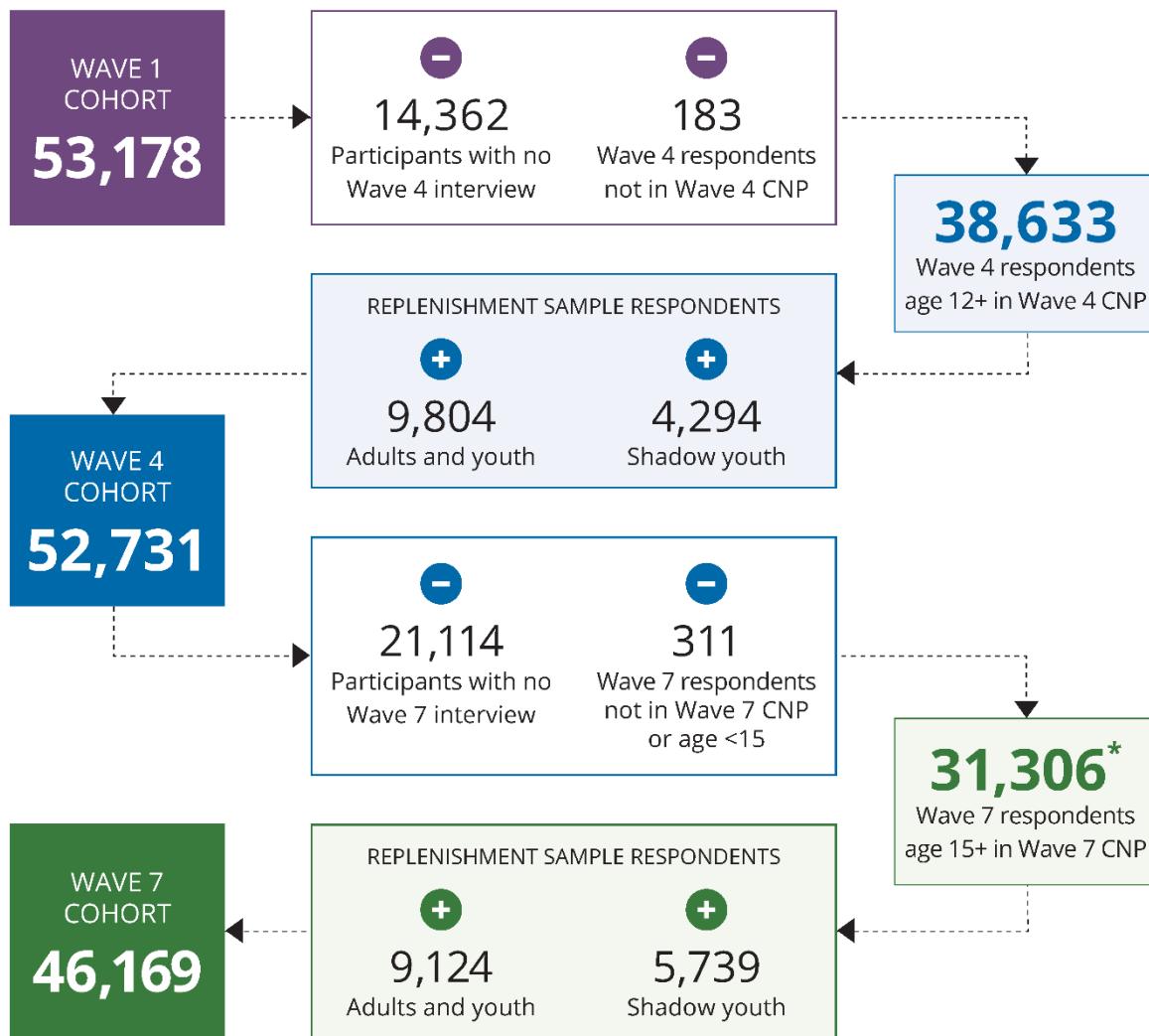
---

The PATH Study is a nationally representative, longitudinal cohort study of adults and youth. The PATH Study's design allows for the longitudinal assessment of patterns of use for a spectrum of tobacco products, including initiation, cessation, relapse, and transitions between products, as well as factors associated with use patterns. The study currently follows three cohorts:

- **The Wave 1 Cohort:** Participants selected using a probability sample at Wave 1 (conducted September 12, 2013, to December 14, 2014) when they were ages 9 and older. See Section 2.1 for information regarding the selection of the Wave 1 Cohort.
- **The Wave 4 Cohort:** Participants selected using a probability replenishment sample at Wave 4 (conducted December 1, 2016, to January 3, 2018) when they were ages 10 and older, combined with members of the Wave 1 Cohort who completed an interview at Wave 4 and were in the CNP at that time. See Section 2.3 for information regarding the selection of the additional participants contributing to the Wave 4 Cohort.
- **The Wave 7 Cohort:** Participants selected using a probability replenishment sample at Wave 7 (conducted January 6, 2022, to April 2, 2023) when they were ages 9 and older, combined with members of the Wave 4 Cohort who completed an interview at Wave 7, were ages 15 and older, and were in the CNP at that time. See Section 2.9 for information regarding the selection of the additional participants contributing to the Wave 7 Cohort.

Figure 2 illustrates the sample sizes for the Wave 1 Cohort, Wave 4 Cohort, and Wave 7 Cohort.

**Figure 2.** Illustration of the sample sizes for the Wave 1 Cohort, Wave 4 Cohort, and Wave 7 Cohort



\* 77.5% of the Wave 4 Cohort Wave 7 respondents age 15+ in the Wave 7 CNP were sampled in Wave 1.

## 2.1 Wave 1 Sample Design

### 2.1.1 Target Population

The PATH Study's Wave 1 target population is the civilian household population ages 9 and older in the United States (all 50 States and the District of Columbia). College students were sampled through their permanent residence rather than at their dormitory. Active-duty members of the military (Army, Navy, Marines, Air Force, and Coast Guard) were ineligible, as were all persons

living in institutional and noninstitutional group quarters other than college dormitories. Spouses and children of active-duty military living in the United States were eligible for the study.

## **2.1.2 Multi-Stage Sampling**

The Wave 1 sample was selected using a four-stage, stratified probability sample design involving the selection of (1) 156 PSUs consisting of counties or groups of contiguous counties, (2) 6,049 second-stage sampling units (referred to as segments), (3) 166,088 mailing addresses, and (4) 76,539 sampled persons within households occupying dwelling units at sampled addresses. The sampling rates for adults varied by age, race, and tobacco-use status. Two-phase sampling was used for adult selection within households to correct for potential misreporting by the household screener respondent of the tobacco-use status of adult members of the household. The sampling rates for the two phases were designed to achieve sufficiently large sample sizes for young adults (ages 18 to 24) and adults of all ages who use tobacco. Generally, up to two youths ages 12 to 17 were sampled within each household (see Section 2.1.2.4 for details). In addition, a shadow sample of youth ages 9 to 11 was selected to be interviewed as youth in later waves.

### **2.1.2.1 PSU Sampling**

At the first stage, a stratified sample of 156 PSUs was selected using probability proportional to size (PPS) sampling. The measure of size (MOS) was defined as a weighted sum of estimated PSU adult population counts of the subgroups that would be sampled within households at different rates, where the weights used to construct the MOS were proportional to the expected overall sampling rates to be applied for each subgroup. The use of this composite MOS was designed to (1) give relatively higher probabilities of selection to PSUs with higher proportions of the key subgroups of young adults, Black or African American adults,<sup>1</sup> and adults who use tobacco; (2) to improve the chances that sufficient numbers of the various sampling subgroups would be included in the sample; and (3) to produce more balanced workloads for field interviewers.

---

<sup>1</sup> Black or African Americans were oversampled because of their higher rate of use of menthol cigarettes compared to other racial/ethnic groups. For this purpose, Black or African American was defined as Black or African American alone or in combination with other race(s), whether Hispanic or non-Hispanic.

Thirty-five PSUs were selected with certainty. These PSUs in effect serve as their own strata for variance estimation. The remaining 121 PSUs were grouped into 57 strata based on known PSU characteristics that were related to the PATH Study variables of interest. This stratification improves the precision of estimates of variables of interest (e.g., tobacco use, perceptions, health, and possible changes over time in those characteristics) because PSUs within the same strata are in general more homogeneous than PSUs in the sampling frame. The variables used for stratification included census region and division, urban/rural designation, Core Based Statistical Area status and size, percent of adults ages 25 and older with at least a bachelor's degree, percent of population with family income below 200 percent of the poverty level, percent of population who were Black or African American, and percent of population who were Hispanic. These variables were available at the county level from the 2010 decennial census and American Community Survey (ACS) 5-year (2006-2010) data. Fifty-seven strata were formed using these variables. From each of these 57 strata, two, or occasionally three, PSUs were sampled systematically with probability proportional to the MOS, resulting in a total of 121 non-self-representing PSUs.

### **2.1.2.2 Segment Sampling**

Sampled PSUs were divided into segments based on census-defined blocks and had, in general, a minimum of 100 occupied housing units. This formation of segments sought to achieve desired sample sizes while minimizing within-segment travel time for field staff. A sample of segments was subsequently selected within each of the sampled PSUs.

Each segment was assigned a MOS to be used in drawing a PPS sample of segments. The segment composite MOS was computed in the same manner as the PSU composite MOS, with the expected counts of adults in each of eight key sampling subgroups defined by age, race, and tobacco use being computed at the segment level, rather than at the PSU level.

Socio-demographic data from the 2010 census were used to group and sort the segments on the frame before drawing a PPS systematic sample from each PSU. The demographic factors used were percent Black or African American,<sup>2</sup> percent Hispanic, and percent of occupied housing units that were owner-occupied. Segments similar to each other on these three factors were grouped using a

---

<sup>2</sup> Black or African American was defined as Black or African American alone or in combination with other race(s), whether Hispanic or non-Hispanic.

clustering procedure (Golder and Yeomans, 1973; Judkins and Singh, 1981). Consequently, this procedure helped to ensure that a well-balanced sample of segments was taken from each PSU such that the resulting sample of segments would produce estimates with smaller variance. A systematic PPS sample of about 40 segments was drawn within each noncertainty PSU,<sup>3</sup> with more segments drawn in the larger certainty PSUs, for a total of 6,049 segments.

### **2.1.2.3 Address Sampling**

At the third stage of sampling, addresses within each segment were ordered by census block, and a sample was selected by systematic sampling. The goal in assigning address selection probabilities was to minimize the variation in weights across the first three stages of sampling and to produce a relatively even workload for field interviewers in each segment. The sampling rate within each segment was determined so that each address selected for the sample would have approximately the same unconditional probability of selection over all three stages of sampling.<sup>4</sup> Due to the form of the composite MOS used for PPS sampling at the first two stages of the design, the allocated number of sampled addresses varied by segment. However, the sampling scheme was intended to produce approximately equal segment workloads in terms of the number of adult interviews conducted. A systematic sample of addresses with the desired sampling rate was chosen from each segment.

### **2.1.2.4 Within-Household Sampling**

The fourth stage of sampling selected persons from the sampled households. During the household screener, one adult household member (referred to as the screener respondent) was asked to list members of the household and provide demographic, active military service status (for ages 17 and

---

<sup>3</sup> A few small PSUs had fewer than 40 segments; in these, all segments were selected with certainty.

<sup>4</sup> The sampling rate was adjusted for some segments to yield an expected number of addresses between 10 and 40. In addition, in some cases, a single address represented a multi-unit structure that contained many dwelling units, all receiving their mail at a single mail “drop point.” These drop points were sampled at a higher rate and then a subsample of the dwelling units was drawn from the selected drop point.

older only), and tobacco-use information (for adults only) about each person for use in sampling three main groups of interest:

- Adults (up to two adults per household were sampled);
- Children ages 12 to 17 (referred to as “youth,” generally up to two per household were sampled); and
- Children ages 9 to 11 (referred to as “shadow youth,” generally up to two per household were sampled) to be interviewed as youth after reaching age 12.

The sections below provide more details regarding the selection of adults, youth, and shadow youth. All adults and youth completing an interview at Wave 1 (45,971 participants), along with all shadow youth permitted by a parent or guardian to participate in the study (7,207 participants), form the Wave 1 Cohort (53,178 participants).

### ***Adults***

The sampling procedure for selecting adults within a household had two phases. Phase 1 sampling depended on the information provided by the household screener respondent. For the Phase 1 sampling, adults in the household were classified into one of eight subgroups defined by the cross-classification of age (18 to 24, 25 or older), race (Black or African American,<sup>5</sup> all other), and use of tobacco<sup>6</sup> (meets the oversampling criteria, does not meet the oversampling criteria), as reported by the screener respondent. Predetermined rates were used to sample adults in Phase 2 of the screening process. This led to the final sampling for the adult interview, subject to the constraint that, at most, two adults could be sampled from each household. Phase 2 sampling allowed for the correction of classification errors by household screener respondents, in particular when the screener respondent’s information erroneously classified the sampled adult as not meeting the tobacco-use oversampling criteria when in fact they did. Phase 2 sampling was based on the sampled individual’s self-reported

---

<sup>5</sup> Black or African American was defined as Black or African American alone or in combination with other race(s), whether Hispanic or non-Hispanic.

<sup>6</sup> Because of the PATH Study’s interest in persons who are experimenting with tobacco products or are likely to use tobacco products in the future, a broad set of tobacco-use criteria was used to classify adults into Phase 1 sampling domains. Adults met the tobacco-use oversampling criteria if it was reported that they smoked a cigarette, cigar, or pipe, or used smokeless tobacco every day or some days; and/or had ever used an e-cigarette, snus, or dissolvable tobacco, or had ever smoked tobacco in a hookah.

information and thus considered more accurate.<sup>7</sup> The sampling rates for the two phases were designed to achieve predetermined minimum sample sizes for key subgroups of people, such as young adults (ages 18 to 24) and adults of all ages who met the tobacco-use oversampling criteria.

At Phase 1, sampling rates for adults who did not meet the tobacco-use oversampling criteria were kept within reasonable bounds, compared to the sampling rates for adults who did to ensure that the weights of any adults sampled at Phase 1 as not meeting the tobacco-use oversampling criteria who then reported information at Phase 2 which did meet the criteria would be similar to the weights of those who were correctly classified at Phase 1. Misclassification in the other direction—with the screener respondent reporting information which classified the adult as meeting the tobacco-use oversampling criteria when the person self-reported information which did not—was handled by deselecting some members of this group so that those retained would have sampling rates similar to those of other adults who did not meet the tobacco-use oversampling criteria.

The population proportions of adults within the eight age/race/tobacco-use subgroups were unknown at the initial design stage. The within-household adult sampling rates were adjusted during Wave 1 data collection as more accurate information was accrued to achieve the desired sample sizes in the eight domains.

The two-phase sampling procedure and disproportionate sampling of younger adults, Black or African American adults, and adults who met the tobacco-use oversampling criteria had two main effects on the sampling errors. First, the procedure resulted in larger sample sizes in the oversampled subgroups than would have occurred using equal probability sampling, which reduced the sampling errors for estimates calculated for those subgroups. However, the increased precision for those subgroups increased weight variation for the adult sample as a whole. The average selection probabilities for subgroups ranged from approximately 0.10 (for non-Black or African American adults ages 25 and older who did not meet the tobacco-use oversampling criteria) to 0.80 (for Black or African American adults ages 18 to 24 who met the tobacco-use oversampling criteria), which leads to weight factors between 10 and 1.25. Consequently, under this design, estimates calculated for all adults have larger standard errors than they would if all adults were sampled at the same rate. However, sampling all adults at the same rate would have required an enormous sample size to

---

<sup>7</sup> Because of the PATH Study's interest in persons who are experimenting with tobacco products or are likely to use tobacco products in the future, a "wide net" definition of tobacco use was implemented when classifying adults into Phase 2 sampling domains. Adults who met the tobacco-use oversampling criteria reported that they had smoked a cigarette, cigar, or pipe, or had used smokeless tobacco in the past 30 days; and/or had ever used an e-cigarette, snus, or dissolvable tobacco, or had ever smoked tobacco in a hookah.

achieve the number of adults who use tobacco needed for the scientific aims of the PATH Study and would have resulted in a large group of adults who do not use tobacco.

### ***Youth and Shadow Youth***

The PATH Study youth sample has two components: selection of youth ages 12 to 17 for the Wave 1 and follow-up interviews, and selection of a shadow sample of youth ages 9 to 11 for inclusion in future waves. Youth in the shadow sample were selected at Wave 1 to replenish the 12-to 17-year-old youth sample in later waves; they were not interviewed at Wave 1. The sampling of youth within sampled households was independent of adult sampling and did not involve any oversampling by race, ethnicity, sex, or tobacco use. The number of persons ages 12 to 17 was tallied for the household. If only one or two children in that age range were in the household, those children were included in the youth sample. If there were more than two children in that age range, two were randomly selected for the youth sample.

The shadow youth selection procedure was identical to that for the youth ages 12 to 17 and was carried out independently of the youth and adult sampling within sampled households. The number of 9- to 11-year-old children in the household was tallied. If only one or two children in that age range were in the household, those children were included in the shadow sample. If there were more than two children in that age range, two were randomly selected for the shadow sample.

Given a special analytic interest in multiple-birth youth (e.g., twins), the shadow and youth sampling procedures were modified when households containing multiple-birth youth were encountered so that the multiple-birth youth would have relatively higher probabilities of selection. This resulted in some households with more than two children selected for the youth and/or shadow youth samples.

## **2.2      Sample Design for Waves 2 and 3**

There was no additional sampling for Wave 2 or 3. All Wave 1 participants were eligible for Wave 2 as long as they continued to live in the United States and were not incarcerated. For longitudinal studies, issues of eligibility and target populations of inference over time require careful consideration. As stated in Section 2.1, the PATH Study target population at Wave 1 excluded all active-duty members of the military and all persons living in group quarters other than college dormitories. The exclusion applied to both institutional and noninstitutional group quarters. Thus,

the target population for the PATH Study in Wave 2 is the resident U.S. population at the time of Wave 2 (other than those who were incarcerated) who were in the CNP and ages 9 and older at Wave 1. Similarly, the target population for the PATH Study in Wave 3 is the resident U.S. population the time of Wave 3 (other than those who were incarcerated) who were in the CNP and ages 9 and older at Wave 1.

## **2.3      Wave 4 Sample Design**

The objectives of the Wave 4 sample design were to replenish the Wave 1 sample to account for the aging of Wave 1 participants and sample attrition, providing an opportunity to replenish the shadow sample as a source of 12- and 13-year-olds in future waves. The replenishment also provided the opportunity for a panel conditioning assessment to monitor the potential effects of participating in the PATH Study over time.<sup>8</sup> The replenishment resulted in the formation of a new cohort.

Section 2.3.1 describes the target populations at the time of Wave 4 for the Wave 1 Cohort and the Wave 4 Cohort, respectively. Section 2.3.2 describes the sample selection of the replenishment portion of the Wave 4 Cohort.

### **2.3.1    Target Populations**

For the Wave 1 Cohort, Wave 4 of the PATH Study was the third follow-up wave. These participants were eligible for Wave 4 if they were still U.S. residents and not incarcerated. The target population for the Wave 1 Cohort in Wave 4 is the resident U.S. population at the time of Wave 4 (other than those who were incarcerated) who were in the CNP and ages 9 and older at Wave 1.

At Wave 4, a probability replenishment sample of adults, youth, and shadow youth (ages 10 and 11) was selected from the CNP at the time of Wave 4, including persons who were not in the CNP at the time of Wave 1 (such as recent immigrants or those returning from the military). Only members of the Wave 1 Cohort who were also in the CNP at the time of Wave 4 were combined with the

---

<sup>8</sup> Panel conditioning is a term used to describe the situation where the reported attitudes and/or behaviors of survey respondents who have been retained in a study for an extended period of time begin to be influenced by study participation itself. After the Wave 4 replenishment, such an assessment was conducted comparing PATH Study estimates between two portions of the Wave 4 Cohort: those sampled at Wave 4, and those sampled at Wave 1. The results of this analysis are presented in Opsomer et al. (2023).

Wave 4 replenishment sample to form the new Wave 4 Cohort. As a result, the target population for the Wave 4 Cohort is the CNP who were ages 10 and older at the time of Wave 4.

### **2.3.2 Replenishment Sampling**

As part of the Wave 4 replenishment effort, adults, youth, and shadow youth were sampled within the existing PATH Study sample segments from among the addresses not selected for Wave 1. The goals for the Wave 4 address selection were to achieve desired sample sizes, to minimize variation in weights across the first three stages of sampling, and to produce a relatively even workload for field interviewers in each segment. A total of 174,273 mailing addresses were selected. To meet the needs for the Wave 4 Cohort shadow sample (i.e., to satisfy Wave 5 youth sample size requirements for 12- and 13-year-olds), a randomly selected subset of the sampled addresses (115,536 or close to two-thirds of the addresses) were screened solely to identify shadow youth ages 10 and 11. The remaining addresses (58,737) were screened for adults, youth, and shadow youth ages 10 and 11. These are referred to as the “SO” (shadow youth only) and “AYS” (adults, youth, and shadow youth) replenishment samples, respectively.

The Wave 4 within-household sampling procedures mirrored those used at Wave 1, with sampling rates varying for adults by age, race, and tobacco-use status;<sup>9</sup> however, the within-household sampling rates were designed to bring the Wave 4 adult and youth sample sizes up to a level commensurate with Wave 1 sample sizes by sampling domain. This required consideration for the expected combined effect of the aging of Wave 1 participants and the loss of sample size due to attrition by the time of Wave 4. In particular, it was necessary to oversample adults based on their tobacco use because the Wave 1 respondents who became adults after Wave 1 were not oversampled with respect to tobacco-use status.

A total of 18,360 adults ages 18 and above and youth ages 12 to 17 were selected as part of the replenishment sample. Those who completed an interview ( $n=9804$ ) were combined with Wave 1 Cohort interview respondents at Wave 4 in the CNP ( $n=38,633$ ) to form a sample of 48,437 participants total. These interview respondents together with the newly sampled shadow youth

---

<sup>9</sup> The same two-phase sampling procedure for adults used at Wave 1 (and described in Section 2.1.2.4) was used for the replenishment sample, but the definitions of tobacco use were expanded to include use of all e-products and not only use of e-cigarettes.

permitted by a parent or guardian to participate in the study (4,294 participants) formed the Wave 4 Cohort (52,731 participants). Table 2-1 presents the Wave 1 Cohort and Wave 4 Cohort sample sizes at the time the Wave 4 Cohort was formed.

**Table 2-1. Wave 1 Cohort and Wave 4 Cohort sample sizes at the time the Wave 4 Cohort was formed**

	<b>Wave 1 Cohort</b>	<b>Wave 4 Cohort</b>
Wave 1 Cohort members who did not complete a Wave 4 interview	<b>14,362</b>	
Wave 4 interview respondents in the Wave 1 Cohort and in the Wave 4 CNP	<b>38,633</b>	<b>38,633</b>
Wave 4 interview respondents in the Wave 1 Cohort, but not in the Wave 4 CNP	<b>183</b>	
Wave 4 replenishment sample interview respondents		<b>9,804</b>
Wave 4 replenishment sample shadow youth respondents		<b>4,294</b>
Total (shadow youth, youth, and adults)	<b>53,178</b>	<b>52,731</b>

## 2.4 Wave 4.5 Sample Design

Wave 4.5 was a special data collection for which only youth ages 12 to 17 at the time of Wave 4.5 were interviewed. There was no additional sampling for Wave 4.5, and no biospecimens were collected. The target population for the Wave 1 Cohort in Wave 4.5 is the resident U.S. population who were ages 13 to 17 at the time of Wave 4.5 (other than those who were incarcerated) and part of the CNP at the time of Wave 1. The target population for the Wave 4 Cohort in Wave 4.5 is the resident U.S. population who were ages 12 to 17 at the time of Wave 4.5 (other than those who were incarcerated) and part of the CNP at the time of Wave 4.

## 2.5 Wave 5 Sample Design

There was no additional sampling for Wave 5. All study members were eligible for Wave 5 as long as they continued to live in the United States and were not incarcerated. The target population for the Wave 1 Cohort in Wave 5 is the resident U.S. population at the time of Wave 5 (other than those who were incarcerated) who were in the CNP and ages 9 and older at Wave 1. The target population for the Wave 4 Cohort in Wave 5 is the resident U.S. population at the time of Wave 5 (other than those who were incarcerated) who were in the CNP and ages 10 and older at Wave 4.

## 2.6 Wave 5.5 Sample Design

Wave 5.5 was a special data collection for which only youth ages 13 to 17 and young adults ages 18 to 19 at the time of Wave 5.5 were interviewed. As shown in Table 1-1, youth age 13 at the Wave 5.5 interview are the youngest members of the Wave 4 Cohort. There was no additional sampling for Wave 5.5, and no biospecimens were collected. Wave 5.5 data collection began on December 1, 2019, with the same in-person methods used in prior non-replenishment waves of the PATH Study and was originally scheduled to continue until November 30, 2020. However, in-person data collection was suspended on March 17, 2020, due to the COVID-19 pandemic. Data collection resumed via telephone on July 3, 2020, and continued until December 31, 2020.

If the data collected in person, before the COVID-19 pandemic and related restrictions were widespread in the United States, had been simply pooled with the data obtained in the latter portion of 2020 via telephone, the resulting Wave 5.5 estimates would have the potential for bias. This is because even though the combined data would represent tobacco-use behavior over most of the year, it would not include the 15-week period during which U.S. COVID-19 cases were rising and most state populations were under stay-at-home orders; such factors may have influenced tobacco-use behaviors. Because of potential changes in tobacco-use behaviors due to the COVID-19 pandemic and absence of data collection for an extended period, the combined data are not representative of the full year. For this reason, it was decided that only data collected via telephone would be used to produce Wave 5.5 estimates, and the target population would be re-defined to cover the latter portion of 2020. However, the PATH Study participants eligible for Wave 5.5 were scheduled for interviews approximately 1 year from their Wave 5 interview, and not in random subsamples. (See Section 3.7 for more information about PATH Study data collection procedures.) This means that the participants scheduled for interviews in the latter portion of the year were not a random subsample of those eligible for Wave 5.5, so analyses of data collected from only these participants could also result in biased estimates. To overcome this concern, all participants who completed a Wave 5.5 interview in person on or before March 17, 2020, and were still age-eligible

for Wave 5.5 (i.e., ages 13 to 19) were re-contacted for an interview by telephone starting August 3, 2020.<sup>10</sup>

During the data collection hiatus, some of the 19-year-olds interviewed in person turned age 20 and others would turn age 20 after the resumption of data collection. To ensure that the in-person respondents re-contacted by telephone would still be age eligible for Wave 5.5, August 31, 2020, was used as a date of demarcation to allow some older 19-year-olds (e.g., those with 20<sup>th</sup> birthdays in September) to be part of Wave 5.5. Those ages 13 to 19 on August 31, 2020, who previously completed the Wave 5.5 interview in person were eligible for the Wave 5.5 telephone re-interview. More generally, participants who were ages 19 and younger<sup>11</sup> on August 31, 2020, were eligible for Wave 5.5. This date of demarcation was also applied to determine the eligibility for PATH-ATS, which is discussed in Section 2.7.

Only data collected via telephone interviews are included in the Wave 5.5 data files. Although data collection at Wave 5.5 focused on members of the Wave 4 Cohort, data for the Wave 1 Cohort are available for Wave 5.5 analysis given the amount of overlap between the two cohorts. The target population for the Wave 1 Cohort in Wave 5.5 is the resident U.S. population who were ages 15-19 in the latter portion of 2020 (other than those who were incarcerated) and part of the CNP at the time of Wave 1. The target population for the Wave 4 Cohort in Wave 5.5 is the resident U.S. population who were ages 13 to 19 in the latter portion of 2020 (other than those who were incarcerated) and part of the CNP at the time of Wave 4.

Because the same date of demarcation was applied to determine age eligibility for the Wave 5.5 telephone interview and the PATH-ATS, the two samples are mutually exclusive in terms of their coverage of adults and jointly exhaustive for the adult population. Thus, the Wave 5.5 and the PATH-ATS data may be combined for analysis. See Section 2.7 for a description of the target

---

<sup>10</sup>Most participants whose in-person interview status was finalized as “nonrespondent” prior to the resumption of data collection on July 3, 2020, were participants who refused to respond to the interview. Because it was presumed that these participants would also refuse a request to complete a telephone interview, it was decided that no participants finalized as nonrespondent prior to the resumption of data collection would be recontacted for a Wave 5.5 telephone interview. These cases were accounted for in the weighting process as described in Appendix B.7.

<sup>11</sup>This includes some participants who were age 12 on August 31, 2020; these participants were contacted for an interview on or after their 13<sup>th</sup> birthday.

population for the combined Wave 5.5 and PATH-ATS, and Section 6.3.5 for more information on the appropriate use of weights for such analyses.

## **2.7      Adult Telephone Survey Sample Design**

The PATH Study Adult Telephone Survey (PATH-ATS) was a special data collection for which a subsample of adults ages 20 and older were interviewed. The survey was developed in 2020 in light of new Federal tobacco policies, such as the nationwide Tobacco 21 policy and flavor restrictions, and the COVID-19 pandemic and its impact on the way people used and purchased tobacco. The objectives of the PATH-ATS were to (1) provide data that can be used to understand the impact of the Federal policy and regulatory changes in 2020; (2) provide data that can be used to assess the relationship between tobacco use and the COVID-19 pandemic; and (3) fill an unanticipated gap in nationally representative data on adult tobacco use, prevalence, cessation, initiation, and relapse.

The PATH-ATS was conducted from September 10, 2020, to December 20, 2020. Because the planned data collection period was shorter than most PATH Study data collection efforts, only a subsample of participants were interviewed. Participants eligible for the PATH-ATS sample were ages 20 and older on August 31, 2020, members of the Wave 4 Cohort, and respondents to the Wave 5 adult interview. A stratified random sample of 18,601 PATH Study participants was selected for the PATH-ATS from a total of 31,343 eligible participants, with oversampling and under sampling based on age (ages 20 to 24, ages 25 and older), tobacco product use (electronic nicotine delivery systems (ENDS), cigarettes), and frequency of use (ever, past 12 months, past 30 days).

The target population for PATH-ATS estimates based on the Wave 4 Cohort is the resident (not incarcerated) U.S. population ages 20 and older in the latter portion of 2020 who were in the CNP at the time of Wave 4 and the resident U.S. population at Wave 5 (other than those who were incarcerated). However, a large portion of the Wave 4 Cohort is made up of members from the Wave 1 Cohort, specifically, those who were both Wave 4 interview respondents and members of the CNP at the time of Wave 4. An additional target population for PATH-ATS analyses restricted to the Wave 1 Cohort is the resident (not incarcerated) U.S. population ages 20 and older at the time of the PATH-ATS who were in the resident U.S. population at Waves 2, 3, 4, and 5 and were not incarcerated at any of those waves and part of the CNP at the time of Wave 1.

In light of the pandemic, it was of analytic interest to learn how the public health crisis affected the full target population(s) of the PATH Study, adults as well as youth. The PATH-ATS was designed to complement the Wave 5.5 sample consisting of those ages 13 to 19 who were interviewed by telephone in the latter portion of 2020. As indicated in Section 2.6, the Wave 5.5 and the PATH-ATS data may be combined for analysis. The target population for the combined Wave 4 Cohort data is the resident (not incarcerated) U.S. population ages 13 and older in the latter portion of 2020 and part of the CNP at the time of Wave 4. The target population for the combined Wave 1 Cohort data is the resident U.S. population who were ages 15 and older in the latter portion of 2020 (other than those who were incarcerated), and part of the CNP at the time of Wave 1. See Section 6.3.5 for more information on the appropriate use of weights for such analyses.

## **2.8      Wave 6 Sample Design**

There was no additional sampling for Wave 6. All study members were eligible for Wave 6 as long as they continued to live in the United States and were not incarcerated. The target population for the Wave 1 Cohort in Wave 6 is the resident U.S. population at the time of Wave 6 (other than those who were incarcerated) who were in the CNP and ages 9 and older at Wave 1. The target population for the Wave 4 Cohort in Wave 6 is the resident U.S. population at the time of Wave 6 (other than those who were incarcerated) who were in the CNP and ages 10 and older at Wave 4. Wave 6 is the first wave of the PATH Study that includes data from interviews conducted in both in-person and telephone modes.

## **2.9      Wave 7 Sample Design**

At Wave 7 the sample was replenished again, and a new cohort was formed. The main objectives of the Wave 7 sample design were to replenish the sample to account for attrition and the aging of participants. Particularly, at the start of Wave 7, the youngest participants in the Wave 4 Cohort were primarily 15 years old, so the replenishment effort provided a means of selecting a new sample of youth ages 12 to 14, in addition to a new shadow sample of youth ages 9 to 11 to be interviewed after they reach age 12 at later waves. The replenishment also provided the opportunity for a second panel conditioning assessment.

Section 2.9.1 describes the target populations at the time of Wave 7 for the Wave 1, Wave 4, and Wave 7 Cohorts. Section 2.9.2 describes the sample selection of the replenishment portion of the Wave 7 Cohort.

### **2.9.1 Target Populations**

Wave 1 Cohort participants were eligible for Wave 7 if they were still U.S. residents and not incarcerated. The target population for the Wave 1 Cohort at Wave 7 is the resident U.S. population at the time of Wave 7 (other than those who were incarcerated) who were in the CNP and ages 9 and older at Wave 1.

Similarly, Wave 4 Cohort participants were eligible for Wave 7 if they were still U.S. residents and not incarcerated. The target population for the Wave 4 Cohort at Wave 7 is the resident U.S. population at the time of Wave 7 (other than those who were incarcerated) who were in the CNP and ages 10 and older at Wave 4.

At Wave 7, a probability replenishment sample of adults, youth, and shadow youth (ages 9 to 11) was selected from the CNP at the time of Wave 7, including persons who were not in the CNP at the time of Wave 4 (such as recent immigrants or those returning from the military). Only members of the Wave 4 Cohort who were ages 15 and older and were also in the CNP at the time of Wave 7 were combined with the Wave 7 replenishment sample to form the new Wave 7 Cohort. As a result, the target population for the Wave 7 Cohort is the CNP who were ages 9 and older at the time of Wave 7.

### **2.9.2 Replenishment Sampling**

As part of the Wave 7 replenishment effort, adults, youth, and shadow youth were sampled within the existing PATH Study sample segments from among the addresses not selected for Wave 1 or Wave 4. The goals for the Wave 7 address selection were to achieve desired sample sizes and to minimize variation in weights across the first three stages of sampling. A total of 244,338 mailing addresses were selected. To meet the needs for the Wave 7 youth sample (given that the youngest of the Wave 4 Cohort were expected to be 15 years old at Wave 7) and the need for a Wave 7 Cohort shadow sample (i.e., to satisfy future youth sample size requirements), the address sample was randomly divided into three subsamples. A subset of 111,423 addresses (or close to 45 percent) were

screened solely to identify youth ages 9 to 14; another subset of 96,752 addresses (or close to 40 percent) were screened to identify youth ages 9 to 17. The remaining 36,163 addresses were screened for adults, youth, and shadow youth ages 9 to 11. These subsamples are referred to as the “YYO” (young youth only ages 9 to 14), “YO” (youth only ages 9 to 17) and “AYS” (adults ages 18 and above, youth ages 12 to 17, and shadow youth ages 9 to 11) replenishment samples, respectively.

The Wave 7 within-household sampling procedures were similar to those used at Wave 1 and Wave 4, with sampling rates varying for adults by age, race, and tobacco-use status.<sup>12</sup> The within-household sampling rates were designed to bring the Wave 7 adult and youth sample sizes up to a level commensurate with Wave 4 sample sizes. This required consideration for the expected combined effect of the aging of continuing participants and the loss of sample size due to attrition by the time of Wave 7. The number of responding adults from the continuing sample at Wave 7 was projected to be very close to Wave 4 levels without replenishment. However, small samples of adults were selected from each of the sampling domains to allow representation of immigrants and other adults who have joined the CNP since Wave 4, and to support a panel conditioning assessment. The numbers of targeted adults made it feasible to sample adults at Phase 1 based only on race and age, and to subsample at Phase 2 based on self-reported tobacco use, in addition to self-reported age and race.

A total of 15,684 adults ages 18 and above and youth ages 12 to 17 were selected as part of the replenishment sample. Those who completed an interview were combined with Wave 4 Cohort interview respondents at Wave 7 who were in the CNP at that time and at least age 15 (31,306 participants total; 77.5% of these participants were sampled at Wave 1). These interview respondents together with the newly sampled shadow youth permitted by a parent or guardian to participate in the study (5,739 participants) form the Wave 7 Cohort (46,169 participants). Table 2-2 presents the Wave 4 Cohort and Wave 7 Cohort sample sizes at the time the Wave 7 Cohort was formed.

---

<sup>12</sup>The main difference between the definition of tobacco use for sampling adults at Waves 1, 4, and 7 is that “e-cigarettes” were replaced by “electronic nicotine products” starting at Wave 4, and ever use of IQOS, nicotine pouches, and other types of oral nicotine products were added at Wave 7.

**Table 2-2. Wave 4 Cohort and Wave 7 Cohort sample sizes at the time the Wave 7 Cohort was formed**

	<b>Wave 4 Cohort</b>	<b>Wave 7 Cohort</b>
Wave 4 Cohort members who did not complete a Wave 7 interview	<b>21,114</b>	
Wave 7 interview respondents ages 15+ in the Wave 4 Cohort and in the Wave 7 CNP	<b>31,306</b>	<b>31,306</b>
Wave 7 interview respondents ages <15 in the Wave 4 Cohort	<b>117</b>	
Wave 7 interview respondents in the Wave 4 Cohort, but not in the Wave 7 CNP	<b>194</b>	
Wave 7 replenishment sample interview respondents		<b>9,124</b>
Wave 7 replenishment sample shadow youth respondents		<b>5,739</b>
<b>Total (shadow youth, youth, and adults)</b>	<b>52,731</b>	<b>46,169</b>

## 3. Data Collection

---

The PATH Study collects baseline and follow-up information on tobacco-use patterns, trends in risk perceptions and attitudes regarding harmful constituents, new and emerging tobacco products, and tobacco initiation, cessation, and relapse behaviors among youth ages 12 to 17 and adults ages 18 and older. Parents who provide permission for their child to complete a youth interview are asked to complete a brief parent interview about themselves, including their tobacco use and the selected youth's parental supervision, school performance, and tobacco use.

The PATH Study also collects biospecimens from consenting respondents to assess markers of nicotine exposure, to detect and compare intermediate endpoints and incident health outcomes associated with the use of tobacco products and related disease processes, and to validate self-reported behavioral and health data. At Wave 1 through Wave 5, respondents completing the adult interview for the first time were asked to provide samples of urine and blood (and buccal cells, until this collection was discontinued in May 2014). Additionally, some adults who provided urine in a previous wave were asked to continue providing a urine specimen in subsequent waves, with the goal of collecting a specimen from the same individuals over several waves. All youth completing the interview at Wave 4 were also asked to provide a urine specimen. All participants who provided a urine specimen when they were youth at Wave 4 were also asked to provide one at Wave 5 upon interview completion; participants completing the youth interview for the first time at Wave 5 were also asked to provide a urine specimen. Biospecimens were not collected in the special data collections (Wave 4.5, Wave 5.5, and PATH-ATS) or in Wave 6. At Wave 7, subsamples of adult interview respondents were asked to provide a urine specimen, or a blood specimen, or both specimens until blood collection was discontinued in March 2022<sup>13</sup>. While youth in the replenishment sample were not asked to provide a urine specimen at Wave 7, subsamples of continuing youth interview respondents were asked to provide a urine specimen.

Data collection at Wave 1 involved four main components: (1) a CAPI household screening instrument, (2) ACASI instruments (separate instruments for youth and adults), (3) a CAPI parent

---

<sup>13</sup>Blood specimens were collected for a limited time, from January 6, 2022 to March 4, 2022.

instrument, and (4) collection of urine, blood, and buccal cell specimens from consenting adults.<sup>14</sup> The collection of biospecimens is not a requirement for adult participation; however, completion of an extended interview is required. At Wave 4, all associated procedures in data and biospecimen collection from Wave 1, including the household screener, were updated to reflect Wave 4 requirements such as the age range of 10 and 11 only for shadow youth. The other data collection components were similar to previous waves, as described above.

For Wave 4.5, only the youth ACASI instrument and the CAPI parent instrument are applicable. For Wave 5.5, only the ACASI instruments for youth and adults and the CAPI parent instrument were used before data collection was suspended on March 17, 2020; a modified version of the ACASI instruments and the CAPI parent instrument were created to facilitate the resumption of interviewing by telephone on July 3, 2020. For the PATH-ATS, a shortened version of the adult instrument was developed for Computed-Assisted Telephone Interviewing (CATI) administration. The Wave 6 ACASI and CAPI instruments were designed to support both telephone and in-person collection modes.

At Wave 7, procedures in data and biospecimen collections from Wave 4 were updated to reflect Wave 7 requirements, such as the age range of 9 to 11 for shadow youth, with the ACASI and parent CAPI instruments supporting both telephone and in-person collection modes. Additionally, a Web pilot test was conducted with selected participants, introducing a new data collection mode for the PATH Study. Table 3-1 provides a summary of the PATH Study data and collection methods by wave, whether biospecimens were collected, and incentives for interviews and biospecimen collections; see Appendix A for an expanded version of this table summarizing the PATH Study protocol by wave.

---

<sup>14</sup>Buccal cells were collected for a limited time, from September 12, 2013, to May 18, 2014.

**Table 3-1.** PATH Study data and collection methods by wave and whether biospecimens were collected

Wave	Instrument	Method	Mode	Biospecimens <sup>a</sup>	Interview/Biospecimen Collection Incentive
1	Household screener <sup>b</sup>	CAPI	In-person	NA	\$2 mailed with advance letter
1	Adult	ACASI	In-person	Urine, blood, buccal cells	\$35/\$25 per specimen
1	Youth	ACASI	In-person	NA	\$25
1	Parent	CAPI	In-person	NA	\$10
2	Adult	ACASI	In-person	Urine, blood <sup>c</sup>	\$35/\$25 per specimen
2	Youth	ACASI	In-person	NA	\$25
2	Parent	CAPI	In-person	NA	\$10
3	Adult	ACASI	In-person	Urine, blood <sup>c</sup>	\$35/\$25 per specimen
3	Youth	ACASI	In-person	NA	\$25
3	Parent	CAPI	In-person	NA	\$10
4	Household screener <sup>b</sup>	CAPI	In-person	NA	\$2 mailed with advance letter
4	Adult	ACASI	In-person	Urine, blood <sup>c</sup>	\$35/\$25 per specimen
4	Youth	ACASI	In-person	Urine	\$25/\$25
4	Parent	CAPI	In-person	NA	\$10
4.5	Youth	ACASI	In-person	NA	\$25
4.5	Parent	CAPI	In-person	NA	\$10
5	Adult	ACASI	In-person	Urine, blood <sup>c</sup>	\$50/\$25 per specimen
5	Youth	ACASI	In-person	Urine	\$35/\$25
5	Parent	CAPI	In-person	NA	\$15
5.5	Adult <sup>d</sup>	ACASI	Telephone	NA	\$50
5.5	Youth <sup>d</sup>	ACASI	Telephone	NA	\$35
5.5	Parent <sup>d</sup>	CAPI	Telephone	NA	\$15
PATH-ATS	Adult <sup>e</sup>	CATI	Telephone	NA	\$50
6	Adult	ACASI	In-person, Telephone	NA	\$50
6	Youth	ACASI	In-person, Telephone	NA	\$50
6	Parent	CAPI	In-person, Telephone	NA	\$15
7	Household screener <sup>b</sup>	CAPI	In-person	NA	\$2 or \$5 mailed with advance letter <sup>g</sup>
7	Adult	ACASI, Web interview	In-person, Telephone, Web	Urine, blood <sup>f</sup>	\$50/\$25 per specimen <sup>h</sup>
7	Youth	ACASI, Web interview	In-person, Telephone, Web	Urine	\$50/\$25 <sup>h</sup>
7	Parent	CAPI, Web interview	In-person, Telephone, Web	NA	\$15

<sup>a</sup> Biomarker data are available only in restricted-use format. See the PATH Study Biomarker Restricted Use Files User Guide for information about the related biomarker data available for each wave.

<sup>b</sup> Household screener data are not available in the released data files.

<sup>c</sup> In Waves 2, 3, and 5 blood was collected from aged-up adults only; in Wave 4 blood was collected from aged-up adults and adults in the replenishment sample.

<sup>d</sup> Modified for telephone administration as discussed in Section 3.7.

<sup>e</sup> Shortened version of the PATH Study adult instrument based on the existing Waves 5 and 5.5 adult instruments, modified for telephone administration, as discussed in Section 3.8.

<sup>f</sup> Blood collection discontinued in March 2022.

<sup>g</sup> Addresses in the AYS sample were mailed \$2; YYO and YO portions of the replenishment sample were mailed \$5 in advance. See Section 3.10 for more information.

<sup>h</sup> Adults and youth who completed the interview by telephone and self-collected and shipped a urine sample received a total of \$35 (\$25 for the urine sample and \$10 for shipping the sample) after it was received at the repository as discussed in Section 3.10.

Before the start of data collection for each wave, including the PATH-ATS and the resumption of Wave 5.5 interviewing by telephone, the PATH Study received approval from the Westat Institutional Review Board. The PATH Study also obtained Certificates of Confidentiality from the National Institutes of Health covering each data collection period.

The following sections provide more information regarding data collection in each completed wave for which data are available. Table 1-2 in Chapter 1 details the data collection dates and the number of interviews conducted. Information about biospecimen collection is provided to give a full picture of data collection in the PATH Study. Biomarker data are not currently available on the PUFs. However, indicators of biospecimen collection by type of biospecimen are included on the adult and youth/parent data files pertaining to the waves in which biospecimens were requested.

## **3.1 Wave 1 Data Collection**

### **3.1.1 Advance Mail and Prepaid Incentives**

Advance letters and brochures were mailed to all sampled addresses 1 to 2 weeks before the field interviewer's Wave 1 visit to inform the target respondents about the PATH Study sponsor, the nature and uses of the data collected, legal authorities, the voluntary nature of participation, and protection of the information. The advance letter contained a \$2 bill as an attention-getter. Before the administration of the household screener, the field interviewer asked if the household received the introductory letter and brochure. Those respondents who did not recall receiving or reading them were provided with copies and were given time to read the information and ask any questions.

### **3.1.2 Household Screener**

Up to two adults and two youths per eligible household were randomly selected using a CAPI screening instrument. One or more additional youth could also be selected among households with multiple-birth youth. The screener respondent was an adult household member age 18 or older; he/she provided oral consent for the screener. The screener used a full household enumeration process to collect information on age and race for each reported household member, as well as information on active military service status for each household member ages 17 and older, and tobacco usage for each adult household member. The relationship of all household members to the

screener respondent was also collected. In addition to household enumeration information, the household screener respondent's and each sampled person's telephone numbers were collected to allow re-contact of the household for quality control purposes or to set appointments for interviews if any of the sampled persons or parents were unavailable at the time of the screening. The household's mailing address was also collected for purposes of re-contacting the sampled person(s).

The sampling algorithm for selecting up to two adults and two youths (except in the case of multiple-birth youth) per household was programmed within the CAPI screener software and tested extensively before data collection began.

### **3.1.3 Interview**

The data collection procedures differed for adults and youth. For adults, some questions were asked at the beginning of the interview to serve the purpose of Phase 2 sampling, as discussed in Section 2.1.2.4.

#### ***Adult Data Collection***

Among other purposes, the household screener (which is also the Phase 1 screener for adults) collected a minimum amount of high-level information about each adult's tobacco usage to classify him/her sufficiently for potential selection based on the PATH Study sampling algorithm. The household screener obtained tobacco-use information about all adults from the household screener respondent, and this information could be inaccurate. To obtain more complete and accurate information from an adult sampled through the household screener, a Phase 2 screener was administered to the sampled adult directly to ask a more extensive panel of questions about tobacco usage. Phase 2 questions were also asked in a private setting using ACASI rather than CAPI.

Following the administration of the Phase 1 screener, if the sampled adult was available and had an adequate amount of time to complete the interview, the field interviewer (1) requested informed consent; (2) administered the Phase 2 screener and adult interview, which included gathering additional contact information about the adult; (3) requested consent for the biospecimen collection; (4) collected the urine specimen (and buccal cell specimen through May 2014); (5) arranged a follow-up appointment for a phlebotomist to collect a blood specimen; and (6) at the completion of the first home visit, paid an incentive to the respondent. If the sampled adult was unavailable or unable

to complete the interview at that time, the field interviewer attempted to schedule an appointment for a return visit or at a minimum, determined the best time for a return visit.

To begin the interview, the field interviewer provided a brief automated tutorial to the adult on using ACASI and launched the ACASI interview. The first part of the interview process was the Phase 2 screener, which may have confirmed or contradicted the information provided in the Phase 1 (household) screener by the household screener respondent. Depending on the individual's self-reports (e.g., on tobacco usage), the computerized sampling algorithm may have de-selected the respondent and he or she was not asked to complete the remainder of the extended interview. Throughout the interview, the field interviewer provided aid to the sampled person if needed on the use of ACASI or related questions. At the end of the extended interview, the field interviewer collected additional contact information for that person and asked the respondent to consent to provide biospecimens.

Adults who completed the interview or were de-selected based on their responses to the Phase 2 screener received \$35 (the adult interview incentive) in appreciation for his/her time for completing the interview, as well as a thank-you letter. Sampled adults who initially declined to participate or were difficult to contact were sent a follow-up refusal conversion letter approximately 3 weeks later.

### ***Youth Data Collection***

Following the administration of the household screener, if a parent or guardian of the selected youth was available and had time, the field interviewer (1) requested parent permission for the youth to participate, (2) requested consent for the short parent interview, and (3) administered the CAPI parent interview, which included collecting additional contact information from the parent. If a parent of the sampled youth was unavailable or unable to participate at that time, the field interviewer attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit. Except for emancipated youth, parental permission was required before conducting the youth interview. The parent who completed a parent interview about the youth received \$10 as a token of appreciation for completing the interview.

For a selected youth with parental permission, if the youth was available and had an adequate amount of time to complete the interview, the field interviewer requested youth assent. If a sampled youth was unavailable or unable to complete the interview at that time, the field interviewer

attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit.

After obtaining assent from the selected youth, the field interviewer provided a brief automated tutorial on using ACASI and launched the ACASI interview. Throughout the interview, the field interviewer provided aid to the sampled youth in using the ACASI instrument if necessary.

Youth who completed the extended interview received \$25 (the youth interview incentive) as a token of appreciation for completing the interview. The parent of the youth respondent received a thank-you letter. The parents of sampled youth who were difficult to contact or initially declined permission for their youth to participate were sent a refusal conversion letter approximately 3 weeks later.

### **3.1.4 Biospecimen Collection**

The field interviewer asked each adult interview respondent to consent to provide biospecimens as part of the PATH Study. However, providing biospecimens was voluntary and not a condition of participation. Completion of the Wave 1 adult interview was required from all respondents in order for them to provide biospecimens.

#### **3.1.4.1 Buccal Cells and Urine Collection**

For adults who consented to provide buccal cells and/or urine, the field interviewer collected the specimens following the completion of the interview. The field interviewer provided written and oral instructions to the respondent for the collection of buccal cells and/or the urine specimen. The field interviewer packed the specimen(s) and shipped the package to the PATH Study biorepository.

If the adult was unavailable or unable to continue with specimen collection immediately following completion of the interview, the field interviewer attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit.

Adults who provided a biospecimen during a field interviewer's visit received \$25 as a token of appreciation for participating in the buccal cell and/or urine collection component of the PATH Study.

Midway through Wave 1 in May 2014, the PATH Study ended buccal cell specimen collection.

### **3.1.4.2 Blood Collection**

For adults who consented to provide blood, the field interviewer scheduled the appointment for the visit by a phlebotomist to obtain the blood specimen. Upon visiting the respondent's home, the phlebotomist administered blood suitability exclusion questions for blood collection (CAPI instrument) and asked the respondent to answer items about his/her recent use of tobacco products (ACASI instrument). The phlebotomist then collected the blood specimen and packed and shipped it to the PATH Study biorepository.

Adults who provided a blood specimen during this later home visit received \$25 as a token of appreciation for participating in the blood collection component of the PATH Study.

## **3.2 Wave 2 Data Collection**

Wave 2 was the first annual follow-up wave of the PATH Study. All Wave 1 participants were eligible for Wave 2 as long as they continued to live in the United States and were not incarcerated. In order to define a reasonable target period for completing the Wave 2 interview, each participant was assigned an "anniversary month" for Wave 2 (approximately 1 year after the Wave 1 interview was completed) and a target data collection period that would improve the likelihood of the interviews taking place 1 year after each individual's Wave 1 interview and give the field interviewers a target for their data collection efforts.

The Wave 2 anniversary month was defined as the calendar month containing the date of the earliest Wave 1 interview completed by a member of the study participant's household, 1 year after the Wave 1 interview. Given the challenges of contacting and scheduling the interviews, the target period encompassed 4 months, starting with the month before the anniversary month and ending 2 months after the anniversary month; however, if necessary, efforts to complete the interview could

continue past this period. Data collection efforts for nonrespondents could have continued up to the last day of the data collection period.

### **3.2.1 Advance Mail and Telephone Contact**

As noted in Section 3.1.3, after the Wave 1 interview, adult respondents and parents of responding youth were mailed thank-you letters. They were also mailed cards on the adult or youth participant's birthday. These mailings were conducted to maintain current contact information and to keep participants and parents engaged in the study.

In addition to these contacts, approximately 3 months in advance of the anniversary month, a reminder letter was mailed to adult participants, parents of youth participants, and parents of shadow youth to remind them of the PATH Study and the upcoming contact from a field interviewer. Also included in the mailing was a refrigerator magnet, a form for updating contact information, and a postage-paid envelope to encourage participants to confirm or update their contact information. Approximately 1 month before the anniversary month, the field interviewer telephoned the adult participant and the parent of the youth participant to arrange a convenient time for an in-person visit by the interviewer to the person's home.

There is one exception to the timing described above. Parents in households with only shadow youth who were not expected to be age 12 at Wave 2 were mailed the package 6 months in advance of the anniversary month. This timing helped the PATH Study increase contact and engagement with such parents because they were not as involved in the study as other participants. Again, the package included a refrigerator magnet, a form for updating contact information, and a postage-paid envelope. During the target period, the household was telephoned to obtain updated contact information.

### **3.2.2 Interview**

The data collection procedures differed for: (1) adults at Wave 1 (continuing adults), (2) youth at Wave 1 who were age 17 or younger at Wave 2 (continuing youth) and their parents, (3) youth who turned age 18 and were eligible for the adult interview at Wave 2 (aged-up adults), and (4) shadow youth who turned age 12 and were eligible for the youth interview at Wave 2 (aged-up youth) and their parents. As in Wave 1, the field interviewer provided a brief automated tutorial on using

ACASI to all adult and youth participants, and launched the appropriate automated ACASI interview. The field interviewer was available throughout the interview to assist the participant if needed on the use of ACASI or related questions.

The Wave 2 interviews build on the information collected from adults and youth who completed interviews at Wave 1, i.e., at their baseline. Stable information such as demographic characteristics (e.g., sex and race) was collected only at baseline. Similarly, information on lifetime use of tobacco products up to the time of the baseline interview was not requested again. The Wave 2 interview updated information on the use of the products since the Wave 1 interview.

### **3.2.2.1 Data Collection for Continuing Adults**

At the in-person visit, the field interviewer: (1) reviewed the main elements of the informed consent for interview provided by the adult upon recruitment; (2) administered the adult interview, which included updating contact information about the adult; (3) as appropriate, reviewed the main elements of consent for the biospecimen collection obtained at recruitment; (4) collected the urine specimen (urine was requested from some adults who provided urine specimens after their first adult interview); and (5) paid the incentive to the respondent at the completion of the first home visit. (The urine collection procedures were the same as described in Section 3.1.4.) If an adult was unavailable or unable to complete the interview at the scheduled time, the field interviewer attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit.

Adults who completed the interview received \$35 (the adult interview incentive) as a thank-you for completing the interview. A refusal conversion letter was sent to adults who initially declined to participate or were difficult to contact.

### **3.2.2.2 Data Collection for Continuing Youth**

At the in-person visit, the field interviewer: (1) reviewed with the parent the main elements of parent permission for the youth to participate obtained upon recruitment; (2) reviewed with the parent the main elements of consent for the short parent interview obtained at recruitment; and (3) administered the CAPI parent interview, which included updating contact information for the parent. If a parent of the youth was unavailable or unable to participate at that time, the field

interviewer attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit. The youth interview was not conducted until parental informed consent was reviewed. Parents who completed a parent interview for the youth received \$10 as a thank-you for completing the interview.

If the youth had parental permission, was available, and had an adequate amount of time to complete the interview, the field interviewer reviewed with the youth the main elements of assent for the interview obtained at baseline and requested the youth to complete the automated ACASI instrument. If the youth was unavailable or unable to complete the interview at that time, the field interviewer attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit.

Youth who completed the interview received \$25 (the youth interview incentive) as a thank-you for completing the interview. Parents of youth respondents received a thank-you letter. A refusal conversion letter also was sent to the parents of youth who were difficult to contact.

### **3.2.2.3 Data Collection for Aged-Up Adults**

At the in-person visit, the field interviewer: (1) requested informed consent; (2) administered the adult interview, which included gathering additional contact information about the adult; (3) requested consent for the biospecimen collection; (4) collected the urine specimen if consent was given; (5) arranged a follow-up appointment for a phlebotomist to collect a blood specimen if consent was given; and (6) paid the incentive to the respondent at the completion of the first home visit. The biospecimen collection procedures for these adults were the same as described in Section 3.1.4. The incentive amounts were the same as described for continuing adults. If an aged-up adult was unavailable or unable to complete the interview at that time, the field interviewer attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit.

### **3.2.2.4 Data Collection for Aged-Up Youth**

At the in-person visit, the field interviewer: (1) requested parent permission for the youth to participate; (2) requested consent for the short parent interview; and (3) administered the CAPI parent interview, which included updating contact information for the parent. If a parent of an aged-

up youth was unavailable or unable to participate at that time, the field interviewer attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit. The youth interview was not conducted until parental informed consent was obtained. The parent and youth incentive amounts were the same as described for continuing youth.

If the youth had parental permission, was available, and had an adequate amount of time to complete the interview, the field interviewer requested youth assent and, if provided, requested the youth to complete the automated ACASI instrument. If an aged-up youth was unavailable or unable to complete the interview at that time, the field interviewer attempted to schedule an appointment for a return visit, or at a minimum, determined the best time for a return visit.

### **3.3 Wave 3 Data Collection**

Wave 3 was the second annual follow-up wave for the PATH Study. All Wave 1 Cohort members were eligible for the Wave 3 interview as long as they continued to live in the United States and were not incarcerated. To define a reasonable target period for completing the Wave 3 interview, each participant was again assigned an anniversary month for Wave 3 and a target data collection period that would improve the likelihood of the interviews taking place at 1-year intervals. The anniversary month was approximately 1 year after the Wave 2 interview was completed or, if there was no Wave 2 interview, 2 years after the completion of the Wave 1 interview.

The target period for contacting and scheduling the interviews encompassed 3 months, starting with the month before the anniversary month and ending 1 month after the anniversary month; however, if necessary, efforts to complete the interview could continue past this period. For some nonrespondents, data collection efforts may have continued up to the last day of the data collection period.

#### **3.3.1 Advance Mail and Telephone Contact**

The advance mailing and telephone contact leading up to Wave 3 followed the same protocol as that preceding Wave 2 and described in Section 3.2.1.

### 3.3.2 Interview

The data collection procedures differed for: (1) participants who had completed an adult interview in any previous wave (continuing adults), (2) participants who had completed a youth interview in any previous wave and who were still ages 12 to 17 (continuing youth) and their parents, (3) youth who turned age 18 and were eligible for the adult interview at Wave 3 (aged-up adults), and (4) shadow youth who turned age 12 and were eligible for the youth interview at Wave 3 (aged-up youth) and their parents. Unlike Waves 1 and 2 in which the ACASI tutorial was required of everyone, a brief automated tutorial on using ACASI in Wave 3 was required of all aged-up youth, as they had no prior ACASI experience, but was optional for all other adults and youth. The field interviewer set up the ACASI interview and was available throughout the interview to assist the respondent if needed on the use of ACASI or related questions. The interviewer administered the parent interview using CAPI.

The Wave 3 interview builds on the information collected from the study participant in a prior wave. Questions regarding stable information such as demographic characteristics (e.g., sex and race) and information on lifetime use of tobacco products asked in a previous interview were not asked again.

At the in-person visit, the field interviewer followed the same procedures described in Section 3.2.2 with a few exceptions. A refusal conversion letter was sent to adults (continuing or aged-up) and parents who initially refused to participate in Wave 3 with one exception. In Wave 3, if an adult or parent had already refused participation in Wave 2 and refused again in Wave 3, that refusal was considered final. Refusal conversion was not attempted in these cases. If a youth (continuing or aged-up) refused the interview after the parent provided consent, the field interviewer did not attempt to convert that refusal.

## 3.4 Wave 4 Data Collection

Wave 4 data collection procedures mirrored both those of Wave 1 for the replenishment sample and Waves 2 and 3 for the Wave 1 Cohort. Like Wave 1, Wave 4 data collection for the replenishment sample included newly sampled addresses and sampled adults, sampled youth and their parents, and shadow youth and their parents.

Wave 4 was the third annual follow-up wave for the Wave 1 Cohort. All Wave 1 participants were eligible for Wave 4 as long as they continued to live in the United States and were not incarcerated. However, Wave 1 Cohort members who refused to participate in both Wave 2 and Wave 3 were not approached for an interview at Wave 4.

To ensure operational efficiency, each Wave 1 Cohort member was again assigned an anniversary month as described in Section 3.3. As in previous waves, data were not collected from shadow youth in Wave 4. Any remaining shadow youth from the Wave 1 Cohort were anticipated to be youth (age 12) by Wave 4; they were interviewed only if the parent confirmed their age as 12 or older.

### **3.4.1 Advance Mail and Telephone Contact**

For the Wave 1 Cohort, the advance mailing and telephone contact leading up to Wave 4 followed the same protocol as that preceding Waves 2 and 3 and described in Section 3.2.1.

For the AYS portion of the replenishment sample, advance letters and brochures were mailed to each sampled address several weeks before the field interviewer's first contact. The advance letter contained a \$2 bill to gain the recipients' attention.

Sampled addresses for the SO replenishment sample were mailed a paper screener designed to determine the presence of any 10- or 11-year-olds in the household as the first step in data collection. A postage-paid envelope, letter of introduction, and a \$2 bill were included with the paper screener. Addresses for which the paper screener was not returned were released for in-person data collection at the conclusion of the mail protocol. Households at sampled addresses that returned the paper screener and reported a 10- or 11-year-old household member were also released for in-person data collection within 2 weeks of receiving the returned paper screener. Advanced letters were mailed to thank the household for returning the paper screener and to notify them of the upcoming field interviewer visit. If a paper screener was returned reporting the absence of 10- or 11-year-olds, the household at that address was not contacted further.

### **3.4.2 Household Screener (Replenishment Sample)**

The household screener for Wave 1 (see Section 3.1.2) was reinstated for the replenishment sample. The purpose of the screener was to identify eligible household members to participate in the PATH

Study and to select specific participants. The same household screener was used for the AYS and SO replenishment samples.

### **3.4.3 Interview**

The data collection procedures differed for: (1) Wave 1 Cohort adults who had previously completed an adult interview (continuing adults), (2) Wave 1 Cohort youth who had previously completed a youth interview and who were still ages 12 to 17 (continuing youth) and their parents, (3) Wave 1 Cohort youth who turned age 18 and were eligible for the adult interview at Wave 4 (aged-up adults), (4) Wave 1 Cohort shadow youth who turned age 12 and were eligible for the youth interview at Wave 4 (aged-up youth) and their parents, (5) adults selected from the AYS replenishment sample at Wave 4 (new adults), and (6) youth selected from the AYS replenishment sample at Wave 4 (new youth) and their parents. For new adults and youth, the data collection procedures mirrored those described in Section 3.1.3 for Wave 1. The data collection procedures for the Wave 1 Cohort mirrored those described in Section 3.3.2 for Wave 3.

One difference between the data collection procedures for youth at Wave 4 was that for the first time in the PATH Study, all youth (continuing, new, and aged-up) who completed a Wave 4 youth interview were asked to provide a urine specimen in Wave 4. Field interviewers requested permission from parents of youth to collect a urine specimen. If permission was given, field interviewers administered the assent for urine collection and collected a urine specimen from youth. Youth who provided a urine specimen received a \$25 incentive.

## **3.5 Wave 4.5 Data Collection**

All participants ages 12 to 17 at the time of Wave 4.5 were eligible for the Wave 4.5 youth interview as long as they continued to live in the United States and were not incarcerated. However, Wave 1 Cohort youth who previously refused to participate in two consecutive waves were not contacted for an interview at Wave 4.5. Wave 4.5 was the fourth annual follow-up wave for those who were members of the Wave 1 Cohort. For those who were sampled at Wave 4, Wave 4.5 was the first annual follow-up wave.

To define a reasonable target period for completing the Wave 4.5 interview, each youth participant was again assigned an anniversary month and a target data collection period that would improve the likelihood of the interviews taking place at 1-year intervals. For participants who completed a Wave 4 interview, the Wave 4.5 anniversary month was approximately 1 year after the Wave 4 interview was completed. For participants in the Wave 1 Cohort who did not complete a Wave 4 interview but completed a Wave 3 interview, the anniversary month was approximately 2 years after the completion of the Wave 3 interview. The target period for Wave 4.5 data collection encompassed 3 months as described in Section 3.3.

Because Wave 4.5 included only youth data collection, there is one exception to the above standard fielding plan. If a Wave 4 household had only one shadow youth who was age 11 at the Wave 4 parent contact and the shadow youth was expected to age up to age 12 after the start of their Wave 4.5 data collection target window, then the anniversary month was set to the shadow youth's birth month plus 2 months (but no later than November 30, 2018) so that the youth could be eligible for the Wave 4.5 interview.

### **3.5.1 Advance Mail and Telephone Contact**

Approximately 3 months in advance of the anniversary month, a reminder letter was mailed to parents of youth participants to remind them of the PATH Study and the upcoming contact from a field interviewer. The advance mailing and telephone contact leading up to Wave 4.5 followed the same protocol as described in Section 3.2.1.

There is one exception to the protocol used in previous waves. The field interviewer contacted a parent or guardian preferably by telephone to verify contact information and confirm the age for youth who were age 16 at Wave 4 to avoid field visits for youth who would turn age 18 or older at Wave 4.5.

### **3.5.2 Interview**

The Wave 4.5 data collection procedures differed for: (1) youth who completed a youth interview in a prior wave and who were still ages 12 to 17 on the date of the Wave 4.5 interview (continuing youth) and their parents and (2) shadow youth whose parent completed a brief parent interview in a

prior wave, but who has not yet completed a youth interview in any wave, and who turned age 12 by the date of their Wave 4.5 interview (aged-up youth) and their parents.

The consent portion of the Wave 4.5 parent interview confirmed that a shadow youth had aged up to a youth by the date of their Wave 4.5 youth interview. Youth were interviewed only if their parent confirmed they were ages 12 to 17. If the assent portion of the youth interview indicated that the youth was age 11, the youth interview was terminated.

At the in-person visit, the field interviewer followed the same procedures described in Section 3.3.2.

## **3.6      Wave 5 Data Collection**

Wave 5 was the fourth or fifth follow-up for those in the Wave 1 Cohort and the first or second follow-up for those in the Wave 4 replenishment sample (depending on whether or not the study participant was part of the Wave 4.5 special data collection). There was no additional sampling for Wave 5, and all study members were eligible if they were still residents of the United States and not incarcerated. Nonrespondents from previous waves were asked to participate in Wave 5 unless they were deceased, had moved out of the United States permanently, were incarcerated long-term, had specifically requested withdrawal from the study, were firm or hostile refusals at a previous wave, were unable to complete a previous interview in English or Spanish, or had a physical or mental disability or chronic illness that prevented participation in the study. In addition, those who had refused participation at any two consecutive waves were not fielded for data collection.<sup>15</sup> The target period for Wave 5 data collection encompassed 3 months as described in Section 3.3.

### **3.6.1    Advance Mail and Telephone Contact**

The advance mailing and telephone contact leading up to Wave 5 followed the same protocol as described in Section 3.2.1.

---

<sup>15</sup>For youth, refusal could have come from the youth or their parent/guardian. A refusal at Wave 4.5 was counted in the calculation of “two consecutive waves” of refusal for participants who were part of the special data collection. Otherwise, Waves 3 and 4 were the last two consecutive waves that were considered in the calculation for Wave 5.

### 3.6.2 Interview

The data collection procedures differed for: (1) adults who had completed an adult interview in any previous wave (continuing adults), (2) youth who had completed a youth interview in any previous wave and who were still ages 12 to 17 (continuing youth) and their parents, (3) youth who turned age 18 and were eligible for the adult interview at Wave 5 (aged-up adults), and (4) shadow youth who turned age 12 and were eligible for the youth interview at Wave 5 (aged-up youth) and their parents. At the in-person visit, the field interviewer followed the same procedures described in Section 3.2.2. Incentives for adult, youth, and parent interviews increased at Wave 5. Adults who completed an interview received \$50, youth who completed an interview received \$35, and parents who completed an interview about the youth received \$15. As in previous waves, urine was requested from some adults who provided urine specimens after their first adult interview in Wave 1; urine and blood specimens were requested from aged-up adults completing their first adult interview at Wave 5. Adults who provided biospecimens received a \$25 incentive for each.

At Wave 5, participants who provided a urine specimen at Wave 4 as youth and all youth who aged up in Wave 4.5 or Wave 5 were requested to provide a urine specimen with parent permission. Field interviewers requested permission from parents of youth to collect a urine specimen. If permission was given, field interviewers administered the assent for urine collection and collected a urine specimen from youth. Youth who provided a urine specimen received a \$25 incentive.

### 3.7 Wave 5.5 Data Collection

Wave 5.5 was a special data collection for Wave 4 Cohort participants who were younger than 18 years at the time of the Wave 4.5 special data collection and were 13 to 19 years at the time of the Wave 5.5 interview. Participants who were also members of the Wave 1 Cohort were 15 to 19 years at the time of the Wave 5.5 interview. Most nonrespondents from previous waves were asked to participate in Wave 5.5; the same exceptions noted in Section 3.6 for Wave 5 applied for Wave 5.5. The PATH Study participants eligible for Wave 5.5 were originally scheduled for interviews approximately 1 year from their Wave 5 interview.

Wave 5.5 data collection began on December 1, 2019, with the same in-person methods used in prior non-replenishment waves of the PATH Study and was originally scheduled to continue until November 30, 2020. However, in-person data collection was suspended on March 17, 2020, due to

the COVID-19 pandemic, the declaration of a national emergency, and social distancing guidelines. Westat field interviewers contacted participants with appointments to cancel those appointments. A letter was sent to all adults, youth, and parents who had been contacted with an initial letter regarding upcoming data collection, or who had been contacted by a field interviewer, describing the suspension of data collection. Data collection resumed via telephone on July 3, 2020, and continued until December 31, 2020. Wave 5.5 in-person data collection procedures were modified to accommodate telephone interviewing.

All participants who completed a Wave 5.5 interview in person on or before March 17, 2020, who were still age-eligible for Wave 5.5 (i.e., ages 13-19 on August 31, 2020, as discussed in Section 2.6) were also re-contacted for interview by telephone starting August 3, 2020.

### **3.7.1 Advance Mail and Telephone Contact**

The advance mailing and telephone contact leading up to Wave 5.5 followed the same protocol as described in Section 3.2.1.

Prior to the re-launch of data collection via telephone, reminder letters were mailed to inform participants that the PATH Study would be contacting them for a telephone interview. Letters were also mailed to eligible Wave 5.5 in-person interview respondents to inform them that field interviewers would be in touch for the re-interviews by telephone.

### **3.7.2 Interview**

The data collection procedures differed for: (1) adults who had completed an adult interview in any previous wave (continuing adults), (2) youth who had completed a youth interview in any previous wave and who were still ages 13 to 17 (continuing youth) and their parents, and (3) youth who turned age 18 and were eligible for the adult interview at Wave 5.5 (aged-up adults), and (4) participants selected as shadow youth who had not yet completed a youth interview and who turned age 13 and were eligible for the youth interview at Wave 5.5 (aged-up youth) and their parents. At the in-person visit, the field interviewer followed the same procedures described in Section 3.2.2. Adults who completed an interview received \$50, youth who completed an interview received \$35, and parents who completed an interview about the youth received \$15.

To transition from in-person administration to telephone data collection and to capture timely and critical information related to the COVID-19 pandemic, several modifications were made to the instruments. Show cards were developed to support telephone administration of the adult, youth, and parent interviews for items that contained visual imagery and long or complex response options. The show cards were made available electronically or were mailed in paper form to participants who were unable to access the Internet.

In response to the COVID-19 pandemic, the PATH Study leveraged the unique opportunity of the Wave 5.5 re-launch to collect timely information on participant experiences during the public health crisis, including a set of questions related to social distancing practices, impacts on tobacco use and stress, perceptions of the COVID-19 pandemic, receiving a COVID-19 diagnosis, and experiencing COVID-19-related symptoms.

Because all Wave 5.5 participants had a history with the PATH Study, telephone numbers (in some instances multiple telephone numbers) were available for contacting adults and parents of youth. The data collection procedures continued to differ by type of participant as noted at the beginning of this section. The field interviewers contacted the parent or adult, introduced the telephone interview, scheduled appointments as needed, and verified Internet access.

The field interviewers reviewed consent/assent with continuing participants. If the participant had aged up to a new category or if the parent was not experienced with the PATH Study, the field interviewer asked for stated verbal consent/assent and recorded the consent/assent. Participants could also request to receive the consent forms and show cards in the mail prior to the telephone interview. After the consent/assent process, the field interviewer administered the interview using the modified instrument (parent, youth, or adult as appropriate), directing the participant to the show cards described above as necessary. Adults who completed a telephone interview received \$50, youth who completed a telephone interview received \$35; and parents who completed an interview about the youth received \$15, this includes respondents who also completed the Wave 5.5 interview in person prior to the suspension of data collection in March 2020.

## **3.8 Adult Telephone Survey Data Collection**

The PATH-ATS was a special data collection for a subsample of Wave 5 interview respondents in the Wave 4 Cohort who were ages 20 and older on August 31, 2020 (as discussed in Section 2.6).

Trained telephone interviewers began collecting data on September 10, 2020, using specialized scheduling, management, and computer-assisted telephone interviewing (CATI) systems. Because all participants selected for PATH-ATS had a history with the PATH Study, telephone numbers (in some instances multiple telephone numbers) were available. The telephone protocol included an algorithm that scheduled calls across multiple combinations of times and days of the week to maximize contact likelihood. The CATI system cycled through the available phone numbers after efforts with the preferred number were unsuccessful. The system also collected new phone numbers if a participant requested a call at a number they had not previously provided.

### **3.8.1 Advance Mail and Telephone Contact**

Advance letters were mailed to the sampled participants notifying them of the telephone survey and that an interviewer would be calling them. The advance letter included a toll-free number to call to schedule an appointment or to ask questions.

### **3.8.2 Interview**

The PATH-ATS used an interviewer-administered instrument. To maintain comparability with existing PATH Study data, the PATH-ATS instrument was a shortened version of the PATH Study adult instrument based on the existing Wave 5 and 5.5 adult instruments and included information related to the COVID-19 pandemic as discussed in Section 3.7.2.

To support telephone administration of the PATH-ATS instrument, electronic show cards were used to convey information to respondents for select items, similar to the approach used for Wave 5.5 telephone interviewing. Questions containing images or lengthy/complex response options used show cards to present the information to the participant as the questions were administered. For participants with Internet access, the show cards were provided electronically. For participants who were unable to access the Internet, a paper version of the show cards was provided before the interview was administered.

All of the PATH-ATS participants were continuing adult participants in the PATH Study and had previously signed a consent form. The elements of consent were reviewed with each participant prior to the interview. PATH-ATS participants could access the consent form electronically or request a paper version be mailed to them. Adults completing the PATH-ATS interview received a \$50 incentive.

## **3.9 Wave 6 Data Collection**

There was no additional sampling for Wave 6, and all study members were eligible if they were still residents of the United States and not incarcerated. Nonrespondents from previous waves were asked to participate at Wave 6 unless they were deceased, had moved out of the United States permanently, were incarcerated long-term, had specifically requested withdrawal from the study, were firm or hostile refusals at a previous wave, were unable to complete a previous interview in English or Spanish, or had a physical or mental disability or chronic illness that prevented participation in the study. In addition, those who had refused participation at any two consecutive waves or otherwise did not respond at three consecutive waves were not fielded for data collection.<sup>16</sup>

### **3.9.1 Advance Mail and Telephone Contact**

The advance mailing and telephone contact leading up to Wave 6 followed the same protocol as described in Section 3.2.1.

### **3.9.2 Interview**

In-person data collection and telephone data collection procedures were used for Wave 6. Due to the COVID-19 pandemic, and for the continued safety of PATH Study participants and field interviewers, Wave 6 data collection began with telephone interviewing. As conditions improved in

---

<sup>16</sup>For youth, refusal could have come from the youth or their parent/guardian. A refusal at Wave 4.5 or Wave 5.5 was counted in the calculation of “two consecutive waves” of refusal for participants who were part of those special data collections. A nonresponse at Wave 4.5 or Wave 5.5 was counted in the calculation of “three consecutive waves” of nonresponse for participants who were part of those special data collections. Otherwise, Waves 4 and 5 were the last two consecutive waves that were considered when determining whom to field for Wave 6 data collection. For youth and adults ages 18 to 19 approached for re-interview by telephone at Wave 5.5 and for adults selected for the PATH-ATS, refusing or nonresponding to the interview was not counted in the calculation of “two consecutive waves” nor “three consecutive waves” of refusal or nonresponse.

certain parts of the country, in-person interviews began on May 7, 2021. Where and when in-person visits were permitted by Westat, NIDA and FDA, and local jurisdictions, in-person data collection was given priority, though each participant made the final decision regarding their interview mode. All in-person contacts with participants were conducted in compliance with local and state restrictions for COVID-19 mitigation.

The data collection procedures differed for: (1) adults who had completed an adult interview in any previous wave (continuing adults), (2) youth who had completed a youth interview in any previous wave and who were still under 18 (continuing youth) and their parents, and (3) youth who turned age 18 and were eligible for the adult interview at Wave 6 (aged-up adults). At the in-person visit, the field interviewer followed the same procedures described in Section 3.2.2. To support telephone administration of the adult, youth, and parent interviews, show cards were similar to those developed for Wave 5.5 telephone interviews. Adults who completed an interview received \$50, youth who completed an interview received \$35, and parents who completed an interview about the youth received \$15.

## **3.10     Wave 7 Data Collection**

In Wave 7, the PATH Study recruited new samples of adults, youth, and shadow youth (ages 9-11), with the main focus on recruiting youth. All Wave 1 Cohort and Wave 4 Cohort participants were eligible for Wave 7 as long as they continued to live in the United States and were not incarcerated. Nonrespondents from previous waves were asked to participate at Wave 7 with the same exceptions noted in Section 3.9 for Wave 6. Wave 7 interviews were conducted for adults, parents of youth, youth, and parents of newly recruited shadow youth, and biospecimens collected from selected adults and youth.

In-person was the preferred data collection mode for Wave 7, with the option to conduct interviews by telephone as needed due to any public health considerations. Selected participants completed a Web interview, as described in Section 3.10.4.

To ensure operational efficiency, each of the Wave 1 Cohort and Wave 4 Cohort was again assigned an anniversary month as described in Section 3.3. As in previous waves, data were not collected from shadow youth in Wave 7.

### **3.10.1 Advance Mail and Telephone Contact**

For Wave 1 Cohort and Wave 4 Cohort participants, the advance mailing and telephone contact leading up to Wave 7 followed the same protocol described in Section 3.2.1. For the AYS portion of the replenishment sample, advance letters and brochures were mailed to each sampled address several weeks before the field interviewer's first contact. The advance letter contained a \$2 incentive to gain the recipients' attention.

Addresses in the YYO and YO portions of the replenishment sample were pre-screened for youth household membership. All addresses in these subsamples were mailed a letter of introduction, a prepaid token cash incentive of \$5, an initial invitation to complete the pre-screener on the Web, and an offer of \$5 for completing the online survey. After this initial mailing and a reminder, a third package was mailed to addresses from which no response was received, including a paper questionnaire that could be completed and returned via mail. All households that completed the pre-screener received the completion incentive: \$5 if the pre-screener was returned via mail or if it was returned via the Web before the nonresponse follow-up mailing; \$10 if the pre-screener was returned via the Web after the nonresponse follow-up mailing. Addresses for which a pre-screener was not returned were released for in-person data collection at the conclusion of the mail protocol. All households that responded to the pre-screener but did not report a youth household member received a letter thanking them for their participation with their incentive and were not contacted further. All households that responded and reported a youth received a letter thanking them for their participation and heralding the upcoming in-person visit along with the incentive for completing the pre-screener.

### **3.10.2 Household Screener (Replenishment Sample)**

The household screener (see Section 3.1.2) was reinstated for the replenishment sample. The purpose of the screener was to identify eligible household members to participate in the PATH Study and to select specific participants. The same household screener was used for the AYS, YYO, and YO replenishment samples.

### 3.10.3 Interview

The data collection procedures differed for: (1) Wave 1 Cohort and Wave 4 Cohort adults who had previously completed an adult interview (continuing adults), (2) Wave 1 Cohort and Wave 4 Cohort youth who had previously completed a youth interview and who were still ages 12 to 17 (continuing youth) and their parents, (3) Wave 1 Cohort and Wave 4 Cohort youth who turned age 18 and were eligible for the adult interview at Wave 7 (aged-up adults), (4) adults selected from the AYS replenishment sample at Wave 7 (new adults), (5) youth selected from the replenishment sample at Wave 7 (new youth) and their parents, and (6) new shadow youth parents. For new adults and youth, the data collection procedures mirrored those described in Section 3.1.3 for Wave 1, with the exception of procedures for telephone interviews which were similar to those described in Section 3.7.2. The data collection procedures for the Wave 1 Cohort and the Wave 4 Cohort participants mirrored those described in Section 3.3.2 for Wave 3, again with the exception of procedures for telephone interviews. Adults who completed an interview received \$50, youth who completed an interview received \$35, and parents who completed an interview about the youth received \$15.

Biospecimen collection was resumed at Wave 7 with procedures mirroring those used in Wave 1 (described in Section 3.1.4). Subsamples of new and continuing adults were asked for urine, blood, or both specimens until blood collection was discontinued on March 4, 2022. A subsample of continuing youth was asked for urine specimens. New youth were not asked for a urine specimen in Wave 7. Adults and youth who provided a urine specimen received \$25. Adults who provided a blood specimen received \$25.

Biospecimen collection procedures differed for participants completing the interview by telephone. If the participant was selected for urine collection, the consent and collection processes was reviewed with the adult or parent of the selected youth over the telephone. The participant was then mailed a urine self-collection package with instructions for collecting and shipping the specimen to the PATH Study biorepository. Adults and youth who completed the process received a total of \$35 (\$25 for the urine sample and \$10 for shipping the sample) after it was received at the repository. If the adult was selected for blood collection, the interviewer obtained consent over the telephone before scheduling an appointment for a phlebotomist to visit the home for specimen collection.

### **3.10.4 Web Pilot Test**

For Wave 7 a pilot test was conducted to evaluate the feasibility of introducing Web versions of the PATH Study questionnaires. A sample of continuing adults, parents, and youth was selected to complete the Wave 7 data collection via the Web with the goal of obtaining 300 adult and 300 youth completed Web interviews. Various personal characteristics were used for selection to ensure inclusion of adult and youth participants who would be exposed to different types of questions (e.g., tobacco use and health conditions as of Wave 5) and who might differ in facility with the Web or in propensity to use the Web (e.g., age and education). In addition, there were some operational criteria that took into consideration various data collection activities that could improve participation in the pilot study while minimizing disruption to the Wave 7 fielding schedule. Sampled adults and parents of sampled youth were mailed an invitation package with instructions including the URL for the Web interview, and a PIN to access their interview. Within the mailing to the parent was an envelope for the youth that contained a separate letter with their instructions. Adults, youth and parents who completed a Web interview received \$50, \$35, or \$15, respectively (the same incentive received for completing the respective in-person interview). Regardless of whether the selected adult, youth, or parent completed the Web interview, they were contacted by a field interviewer to schedule an in-person follow-up interview; if the in-person interview was completed, the participant received the incentive amount noted above. If the participant completed the interview via the Web, the data from that interview are included in the respective adult or youth/parent data file; if they did not complete the Web interview, but completed the ACASI interview, data from the ACASI interview are included.

## 4. Response Rates

---

Some addresses sampled for the PATH Study could not be located or accessed, others were found to be ineligible (e.g., vacant lots and group quarters), and some eligible households did not complete the household screener. Further, not all sampled persons within eligible households agreed to participate in the PATH Study, and those who were recruited into the study at Wave 1, Wave 4, or Wave 7 may not have responded at some or all of the follow-up waves. Section 4.1 presents the calculations of the Wave 1 response rates for the household screener, the adult interview, the youth interview, and parent consent for shadow youth. Section 4.2 presents the adult interview and youth interview response rates for Wave 2 through Wave 7 for the Wave 1 Cohort.

Section 4.3 presents the household screener response rates for both the entire replenishment sample and the AYS replenishment subsample (from which all Wave 4 replenishment sample adults and youth were selected) in Wave 4. The Wave 4 AYS replenishment subsample household response rate is more relevant for analyses that do not involve shadow youth recruited at Wave 4. This section also contains the adult interview and youth interview response rates for the Wave 4 AYS replenishment subsample, and parent consent for shadow youth response rates for the entire Wave 4 replenishment sample. Section 4.4 presents the Wave 4.5 youth interview response rates, as well as the Wave 5, Wave 5.5, Wave 6, and Wave 7 adult interview and youth interview response rates for the Wave 4 Cohort. Section 4.4 also presents the PATH-ATS interview response rates.

Section 4.5 contains the household screener response rates for both the entire replenishment sample and the AYS replenishment subsample (from which all Wave 7 replenishment sample adults were selected) at Wave 7. The Wave 7 AYS replenishment subsample household response rates are more pertinent for analyses that only include adults recruited at Wave 7. Section 4.5 also presents the adult response rates for the Wave 7 AYS replenishment subsample, as well as the youth interview response rates and parent consent for shadow youth response rates for the entire Wave 7 replenishment sample.

As described in Chapter 2, the Wave 4 Cohort is comprised of two groups of study members recruited approximately 3 years apart. Wave 4 was the third follow-up attempt for those sampled at Wave 1, whereas members of the replenishment sample were asked to participate in the PATH Study for the first time. Follow-up wave response rates for the Wave 1 Cohort presented in

Section 4.2 were conditioned on Wave 1 response for those sampled at Wave 1, whereas Wave 4 response rates for those from the replenishment sample presented in Section 4.3 were conditioned on completion of the Wave 4 household screener. It was clear that the Wave 4 response rates for the two groups of study members comprising the Wave 4 Cohort were conceptually different. For this reason, no attempt was made to compute “blended” response rates for the Wave 4 Cohort at Wave 4. Starting in Wave 4.5, the response rates of the two groups of Wave 4 Cohort members who were recruited at different waves were conditioned on the same type of response, i.e., the Wave 4 youth interview or parent consent for shadow youth response. In addition, the Wave 4 Cohort data are to be analyzed as a whole, not by recruitment wave (i.e., not analyzing Wave 4 Cohort members who were recruited in Wave 1 separately from those recruited as part of the Wave 4 replenishment sample). Thus, from Wave 4.5 onward, “blended” response rates are reported for the Wave 4 Cohort.

The Wave 7 Cohort consists of groups of study members recruited at different times. Wave 4 Cohort members who were in the Wave 7 CNP had been asked to participate multiple times since their initial recruitment wave, whereas members of the Wave 7 replenishment sample were asked to participate in the PATH Study for the first time at Wave 7. For the same reason described above for the Wave 4 Cohort, “blended” Wave 7 response rates were not computed for the Wave 7 Cohort. Starting with Wave 7.5, “blended” response rates will be reported for the Wave 7 Cohort.

The response rate calculations were based on the RR3 response rate formula provided by the American Association for Public Opinion Research (AAPOR, 2023). This formula calls for calculating response rates as the ratio of the number of completed cases to the number of eligible sample cases.

## 4.1 Wave 1 Response Rates

This section summarizes the Wave 1 response rates for the household screener, the adult interview, the youth interview, and parent consent for shadow youth.

The household screener response rate, denoted as  $RR_{HH}$ , was calculated using equation 4.1.1:

$$RR_{HH} = C_{HH}/(C_{HH} + N_{HH} + e_{HH} \times U_{HH}) \quad (4.1.1)$$

where

$C_{HH}$  = number of completed cases;

$N_{HH}$  = number of nonresponding cases known to be eligible;

$e_{HH}$  = estimated proportion of nonresponding cases with unknown eligibility that were eligible; and

$U_{HH}$  = number of nonresponding cases with unknown eligibility.

Wave 1 response rates for adults and youth depended on completion of the Phase 1 household screener.<sup>17</sup> The adult interview response rate (conditioned on completion of the Phase 1 household screener), denoted as  $RR_A$ , was calculated as the product of (1) the Phase 2 screener response rate, and (2) the proportion of adults who completed the adult interview among those who completed the Phase 2 screener and were selected for the adult interview, as shown in equation 4.1.2:

$$RR_A = (C_{P2}/(C_{P2} + N_{P2})) \times (C_A/(C_A + N_A)) \quad (4.1.2)$$

where

$C_{P2}$  = number of completed cases for the Phase 2 screener;

$N_{P2}$  = number of nonrespondents to the Phase 2 screener;

$C_A$  = number of completed cases for the adult interview; and

$N_A$  = number of nonrespondents to the adult interview.

---

<sup>17</sup>At the adult and youth levels, all cases involved in the response rate calculations were eligible. In some instances, selected household residents thought to be eligible for the PATH Study during screening were discovered to be ineligible prior to the interview and were excluded from interview response rate calculations. Examples include selected household residents discovered to be on active military duty after screening but prior to the interview and those whose age was incorrectly recorded during screening and whose corrected age was not in the eligible age range for the address subsample to which their household had been randomly assigned.

The youth interview response rate (conditioned on completion of the household screener), denoted as  $RR_Y$ , was calculated using equation 4.1.3:

$$RR_Y = C_Y / (C_Y + N_Y) \quad (4.1.3)$$

where

$C_Y$  = number of completed cases for the youth interview; and

$N_Y$  = number of nonrespondents to the youth interview.

The parent consent for shadow youth response rate (conditioned on completion of the household screener), denoted as  $RR_{SY}$ , was calculated as:

$$RR_{SY} = C_{SY} / (C_{SY} + N_{SY}) \quad (4.1.4)$$

where

$C_{SY}$  = number of shadow youth whose parents provided parent consent; and

$N_{SY}$  = number of shadow youth nonrespondents.

Unweighted response rates were calculated based on actual case counts, whereas weighted response rates were based on the sums of the Wave 1 inverse probability of selection (IPS) weights. The household-level IPS weight is described in equation B.1 in Appendix B. At the adult level, youth level, and shadow youth level, the IPS weight is the product of the household IPS weight and the inverse of the within-household probability of selection. The unweighted response rate measures the success of field operations in obtaining responses from the selected sample. The weighted response rate estimates the proportion of the population represented by the sample group that would have responded if they all had been asked to participate in the study, and it provides a measure of the potential impact of nonresponse on study estimates.

Tables 4-1 through 4-3 show the unweighted and weighted response rates for the Wave 1 household screener, adult interview, youth interview, and parent consent for shadow youth participation. The weighted response rates were 54.0 percent, 74.0 percent, 78.4 percent, and 80.2 percent for the household screener, adult interview, youth interview, and parent consent for shadow youth, respectively. Response rates differed across demographic groups such as age, sex, race/ethnicity, and

education. Differential response rates were compensated for by using nonresponse adjustments in the weighting methods, as described in Chapter 5 and Appendix B.

**Table 4-1. Wave 1 unweighted and weighted household screener response rates**

Type of interview	$C_{HH}$ : Completed (n)	$N_{HH}$ : Nonresponse known to be eligible (n)	$e_{HH} \times U_{HH}$ : Unknown eligibility estimated to be eligible (n)	$RR_{HH}$ : Unweighted response rate (%)	$RR_{HH}$ : Weighted response rate (%)
Household screener	79,198	62,332	4,760	54.1	54.0

**Table 4-2. Wave 1 unweighted and weighted adult interview response rates**

Type of interview	Phase 2 screener		Adult interview		$RR_A$ : Unweighted response rate (%)	$RR_A$ : Weighted response rate (%)
	$C_{P2}$ : P2 screener, completed (n)	$N_{P2}$ : P2 screener, nonresponse (n)	$C_A$ : Adult interview, completed (n)	$N_A$ : Adult interview, nonresponse (n)		
Adult interview	44,303	14,785	32,320	80	74.8	74.0

**Table 4-3. Wave 1 unweighted and weighted youth interview and shadow youth response rates**

Type of Interview <sup>a</sup>	$C_Y$ : Completed (n)	$N_Y$ : Nonresponse (n)	$RR_Y$ : Unweighted response rate (%)	$RR_Y$ : Weighted response rate (%)
Youth interview	13,651	3,800	78.2	78.4
Shadow youth (parent consent)	7,207	1,735	80.6	80.2

<sup>a</sup> For shadow youth calculations, please refer to the components of equation 4.1.4 with subscript SY rather than Y as indicated in the table headers.

## 4.2 Response Rates for Wave 2 Through Wave 7 for the Wave 1 Cohort

The response rates for Wave 2 through Wave 7 were calculated as the ratio of the number of completed cases to the number of cases eligible for the interview at each wave among Wave 1 participants. Wave 5.5 and PATH-ATS response rates for the Wave 1 Cohort are not presented because only members of the Wave 4 Cohort were fielded for Wave 5.5 and PATH-ATS. Wave 1 Cohort members not in the Wave 4 Cohort were not eligible for Wave 5.5 or PATH-ATS.

The response rates were calculated separately for the adult interview and the youth interview. Respondents were classified as adults or youth according to their age on the date of the interview:

persons 18 and older were asked to complete an adult interview, and persons ages 12 to 17 were asked to complete a youth interview.

However, a Wave 1 participant who was a nonrespondent in any follow-up wave did not have an interview date for that wave, so the latest available date of birth or age information was used to determine their age. The age classification date for a nonrespondent in the Wave 1 Cohort was set to the last day of the last month of the target data collection period (see Sections 3.2 through 3.4) or the final date of data collection for the wave, whichever came earlier.<sup>18</sup> Nonrespondents determined to be ages 12 to 17 on the classification date were classified as youth at the wave; those determined to be ages 18 and older on the classification date were classified as adults.

Both unweighted and weighted response rates were calculated. The numbers of cases used in the calculations for the unweighted rates were the actual case counts. The numbers of cases used in the calculations for the weighted rates were the sums of the Wave 1 IPS weights. The person-level IPS weight is the product of the household IPS weight (described in Section B.1 in Appendix B) and the inverse of the within-household probability of selection.

Even though the response rates for adults and youth, denoted as RR, were calculated separately, they were calculated using the same equations (4.2.1 and 4.2.2) shown below:

$$RR = C/(C + N + e \times U) \quad (4.2.1)$$

$$e = (C + N)/(C + N + I) \quad (4.2.2)$$

where

$C$  = number of completed cases;

$N$  = number of nonrespondents known to be eligible;

$e$  = estimated proportion of nonresponding cases with unknown eligibility who were eligible.

---

<sup>18</sup>This approach was modified slightly for Wave 7 due to the extended field period. A modified approach was also necessary for Wave 5.5 due to pandemic-related interruptions to data collection.

$U$  = number of nonrespondents with unknown eligibility; and

$I$  = number of ineligible cases.<sup>19</sup>

Table 4-4 provides the unweighted and weighted response rates for the Wave 2 adult and youth interviews among Wave 1 participants. The weighted response rates were 83.2 percent for adults and 87.3 percent for youth.

**Table 4-4. Wave 2 unweighted and weighted interview response rates for the Wave 1 Cohort**

Type of interview	C: Completed (n)	I: Ineligible (n)	N: Nonresponse known to be eligible (n)	U: Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	RR: Weighted response rate (%)
Adult interview	28,362	284	4,266	1,647	82.8	83.2
Youth interview	12,172	20	1,331	444	87.3	87.3

Table 4-5 provides unweighted and weighted response rates for the Wave 3 adult and youth interviews among Wave 1 participants. Only those participants temporarily, but not permanently, ineligible at Wave 2 (e.g., incarcerated) were included in the calculation of equation 4.2.2. There were a small number of participants discovered to be permanently ineligible at Wave 2 (e.g., deceased, incarcerated long-term, or permanently moved outside the United States). This set of cases, denoted as  $IP$ , was not included in the calculation of equation 4.2.2 but is shown in Table 4-5 for reference. The Wave 3 weighted response rates for the Wave 1 Cohort were 78.4 percent for adults and 83.3 percent for youth.

**Table 4-5. Wave 3 unweighted and weighted interview response rates for the Wave 1 Cohort**

Type of interview	C: Completed (n)	IP: Permanently Ineligible prior to Wave 3 <sup>a</sup> (n)	I: Other Ineligible (n)	N: Nonresponse known to be eligible (n)	U: Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	RR: Weighted response rate (%)
Adult interview	28,148	187	436	5,192	2,781	78.0	78.4
Youth interview	11,814	4	27	1,742	636	83.3	83.3

<sup>a</sup> These cases were not included in equation 4.2.2.

<sup>19</sup>Excludes participants discovered to be permanently ineligible (i.e., deceased, moved out of the United States permanently, or incarcerated long term) at a prior wave. At Wave 4.5, only youth fielded and found to be ineligible are included in this set of cases for the response rate calculation. Study participants not fielded based on their age, as are the twenty participants who became permanently ineligible prior to Wave 4.5.

Table 4-6 provides unweighted and weighted response rates for the Wave 4 adult and youth interviews among Wave 1 participants. Only those participants who had not become permanently ineligible before Wave 4 were included in the calculation of equation 4.2.2. The numbers of participants discovered to be permanently ineligible before Wave 4 are shown in Table 4-6 for reference. The Wave 4 weighted response rates for the Wave 1 Cohort were 73.5 percent for adults and 79.5 percent for youth.

**Table 4-6. Wave 4 unweighted and weighted interview response rates for the Wave 1 Cohort**

Type of Interview	C: Completed (n)	I/P: Permanently Ineligible prior to Wave 4 <sup>a</sup> (n)	A: Other Ineligible (n)	N: Nonresponse known to be eligible (n)	U: Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	WR: Weighted response rate (%)
Adult interview	27,757	400	519	4,133	6,292	72.9	73.5
Youth interview	11,059	8	36	1,365	1,494	79.5	79.5

<sup>a</sup> These cases were not included in equation 4.2.2.

Table 4-7 provides unweighted and weighted response rates for the Wave 4.5 youth interview among Wave 1 respondents. Only Wave 1 respondents who had not become permanently ineligible before Wave 4.5 were included in the calculation of equation 4.2.2. The number of participants discovered to be permanently ineligible before Wave 4.5 is shown in Table 4-7 for reference. The Wave 4.5 weighted youth interview response rate for the Wave 1 Cohort was 74.6 percent.

**Table 4-7. Wave 4.5 unweighted and weighted interview response rates for the Wave 1 Cohort**

Type of Interview	C: Completed (n)	I/P: Permanently Ineligible prior to Wave 4.5 <sup>a</sup> (n)	A: Other Ineligible (n)	N: Nonresponse known to be eligible (n)	U: Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	WR: Weighted response rate (%)
Youth interview	8,761	20	91	1,353	1,665	74.5	74.6

<sup>a</sup> These cases were not included in equation 4.2.2.

Table 4-8 provides unweighted and weighted response rates for the Wave 5 adult and youth interviews among Wave 1 participants. Only those participants who had not become permanently

ineligible before Wave 4.5<sup>20</sup> were included in the calculation of equation 4.2.2. The numbers of participants discovered to be permanently ineligible before Wave 4.5 are shown in Table 4-8 for reference. The Wave 5 weighted response rates for the Wave 1 Cohort were 69.4 percent for adults and 72.3 percent for youth.

**Table 4-8. Wave 5 unweighted and weighted interview response rates for the Wave 1 Cohort**

Type of Interview	C. Completed (n)	I/P. Permanently Ineligible prior to Wave 4.5 <sup>a</sup> (n)	A. Other Ineligible (n)	N. Nonresponse known to be eligible (n)	U. Nonresponse with unknown eligibility (n)	RR. Unweighted response rate (%)	RR. Weighted response rate (%)
Adult interview	28,970	714	840	4,147	9,063	69.0	69.4
Youth interview	6,760	15	41	803	1,825	72.1	72.3

<sup>a</sup> These cases were not included in equation 4.2.2.

Table 4-9 provides unweighted and weighted response rates for the Wave 6 adult and youth interviews among Wave 1 participants. Only Wave 4 Cohort members who had not become permanently ineligible before Wave 6<sup>21</sup> were included in the calculation of equation 4.2.2. The numbers of participants discovered to be permanently ineligible before Wave 5.5 and PATH-ATS are shown in Table 4-9 for reference. The Wave 6 weighted response rates for the Wave 1 Cohort were 57.5 percent for adults and 56.6 percent for youth.

---

<sup>20</sup>At Wave 4.5 only households with age-eligible participants (i.e., ages 12 to 17 prior to the start of Wave 4.5 data collection) were contacted. No attempt was made to determine the eligibility status of PATH Study participants in other households. Thus, cases determined to be permanently ineligible at Wave 4.5 were treated for the Wave 5 response rate calculations for the Wave 1 Cohort and in Table 4-8 as though they were first discovered to be permanently ineligible at Wave 5 and assigned to the “Other ineligible” response rate category.

<sup>21</sup>At Wave 5.5 and PATH-ATS, only households with specific participants (for Wave 5.5, Wave 4 Cohort participants ages 13 to 19 as of August 31, 2020; for PATH-ATS, selected Wave 4 Cohort participants ages 20 and older as of August 31, 2020) were contacted. No attempt was made to determine the eligibility status of PATH Study participants in other households. Thus, cases first discovered to be permanently ineligible at Wave 5.5 or PATH-ATS were treated for the Wave 6 response rate calculations for both cohorts (Wave 1 Cohort in Table 4-9 and Wave 4 Cohort in Table 4-19) as though they were first discovered to be permanently ineligible at Wave 6 and assigned to the “Other ineligible” response rate category.

**Table 4-9. Wave 6 unweighted and weighted interview response rates for the Wave 1 Cohort**

Type of Interview	C Completed (n)	I/P: Permanently Ineligible prior to Wave 6 <sup>a</sup> (n)	A Other Ineligible (n)	N: Nonresponse known to be eligible (n)	U: Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	RR: Weighted response rate (%)
Adult interview	25,604	1,266	900	5,801	15,382	55.2	57.5
Youth interview	2,377	15	12	426	1,395	56.7	56.6

<sup>a</sup> These cases were not included in equation 4.2.2.

Table 4-10 provides unweighted and weighted response rates for the Wave 7 adult and youth interviews among Wave 1 participants. Only those participants who had not become permanently ineligible before Wave 7 were included in the calculation of equation 4.2.2. The numbers of participants discovered to be permanently ineligible before Wave 7 are shown in Table 4-10 for reference. The Wave 7 weighted response rates for the Wave 1 Cohort were 52.4 percent for adults and 54.3 percent for youth.

**Table 4-10. Wave 7 unweighted and weighted interview response rates for the Wave 1 Cohort**

Type of Interview	C Completed (n)	I/P: Permanently Ineligible prior to Wave 7 <sup>a</sup> (n)	A Other Ineligible (n)	N: Nonresponse known to be eligible (n)	U: Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	RR: Weighted response rate (%)
Adult interview	24,408	2,028	424	5,529	19,121	50.0	52.4
Youth interview	899	8	3	170	588	54.3	54.3

<sup>a</sup> These cases were not included in equation 4.2.2.

Response rates for Wave 2 through Wave 7 differed across groups based on data from the Wave 1 interviews (e.g., age, sex, race/ethnicity, and tobacco use). The weighting adjustments attempted to compensate for differential response rates across groups, as described in Chapter 5 and Appendix B.

## 4.3 Response Rates for the Wave 4 Replenishment Sample

This section summarizes the household screener response rates for the AYS replenishment subsample and the entire replenishment sample in Wave 4. Furthermore, it reviews the adult

interview and youth interview response rates for the Wave 4 AYS replenishment subsample, and parent consent for shadow youth response rates for the entire Wave 4 replenishment sample.

Both household screener response rates were calculated using equation 4.1.1, whereas the adult interview response rate was calculated using equation 4.1.2. The youth interview and parent consent for shadow youth response rates were calculated using equations 4.1.3 and 4.1.4, respectively.

Unweighted response rates were calculated based on actual case counts, whereas weighted response rates were based on the sums of the Wave 4 IPS weights described in Section B.4.3.1 in Appendix B. At the adult level, youth level, and shadow youth level, the IPS weight is the product of the household IPS weight and the inverse of the within-household probability of selection. Although it was not designed to generate estimates on its own, the Wave 4 replenishment sample was a probability-based nationally representative sample, so the weighted response may be interpreted as described in Section 4.1.

Tables 4-11 through 4-14 show the unweighted and weighted household screener, adult interview, and youth interview response rates for the Wave 4 AYS replenishment subsample, as well as the unweighted and weighted response rates for the household screener and parent consent for shadow youth for the entire Wave 4 replenishment sample.<sup>22</sup> The weighted response rates for the Wave 4 AYS replenishment sample were 52.8 percent, 68.0 percent, and 70.6 percent for the household screener, adult interview, and youth interview, respectively. The weighted response rates for the entire Wave 4 replenishment subsample were 51.2 percent and 78.7 percent for the household screener and parent consent for shadow youth, respectively. Response rates differed across subgroups. Differential response rates were compensated for by using nonresponse adjustments in the weighting methods, as described in Chapter 5 and Appendix B.

---

<sup>22</sup>Similar to Wave 1, at the adult and youth levels, all cases involved in the Wave 4 response rate calculations were eligible.

**Table 4-11.** Unweighted and weighted household screener response rates for the Wave 4 replenishment samples

Wave 4 replenishment sample group	$C_{HH}$ : Completed (n)	$N_{HH}$ : Nonresponse known to be eligible (n)	$e_{HH} \times U_{HH}$ : Unknown eligibility estimated to be eligible (n)	$RR_{HH}$ : Unweighted response rate (%)	$RR_{HH}$ : Weighted response rate (%)
AYS replenishment subsample	27,364	22,010	2,230	53.0	52.8
Entire replenishment sample	79,540	66,437	8,964	51.3	51.2

**Table 4-12.** Wave 4 unweighted and weighted adult interview response rates for the Wave 4 AYS replenishment subsample

Type of interview	Phase 2 screener		Adult interview		$RR_A$ : Unweighted response rate (%)	$RR_A$ : Weighted response rate (%)
	$C_{P2}$ : P2 screener, completed (n)	$N_{P2}$ : P2 screener, nonresponse (n)	$C_A$ : Adult interview, completed (n)	$N_A$ : Adult interview, nonresponse (n)		
Adult interview	8,927	4,120	6,065	12	68.3	68.0

**Table 4-13.** Wave 4 unweighted and weighted youth interview response rates for the Wave 4 AYS replenishment subsample

Type of interview	$C_Y$ : Completed (n)	$N_Y$ : Nonresponse (n)	$RR_Y$ : Unweighted response rate (%)	$RR_Y$ : Weighted response rate (%)
Youth interview	3,739	1,574	70.4	70.6

**Table 4-14.** Wave 4 unweighted and weighted shadow youth response rates for the entire Wave 4 replenishment sample

Type of Interview	$C_{SY}$ : Completed (n)	$N_{SY}$ : Nonresponse (n)	$RR_{SY}$ : Unweighted response rate (%)	$RR_{SY}$ : Weighted response rate (%)
Shadow youth (parent consent)	4,294	1,049	80.4	78.7

## 4.4 Response Rates for Wave 4.5 Through Wave 7 for the Wave 4 Cohort

This section reviews the Wave 4 Cohort response rates for Wave 4.5 through Wave 7. The response rates for each of these data collections were calculated as the ratio of the number of completed cases to the number of cases eligible for the interview at each wave among Wave 4 Cohort participants.

The response rates were calculated separately for the adult interview and the youth interview. The same protocols used for the Wave 1 Cohort to classify respondents and nonrespondents as adults or youth at each wave were applied to the Wave 4 Cohort (see Section 4.2).

Both unweighted and weighted response rates were calculated using equations 4.2.1 and 4.2.2. The numbers of cases used in the calculations for the unweighted rates were the actual case counts. Except for the PATH-ATS, the numbers of cases used in the calculations for the weighted rates were the sums of the Wave 4 cross-sectional weights. For the PATH-ATS, the weighted response rate was based on the sums of the product of the Wave 4 cross-sectional weight and the inverse of the PATH-ATS probability of selection,<sup>23</sup> which serves as proxy for a PATH-ATS IPS weight.

Table 4-15 provides unweighted and weighted response rates for the Wave 4.5 youth interview among the Wave 4 Cohort. The weighted Wave 4.5 youth interview response rate for the Wave 4 Cohort was 89.1 percent.

**Table 4-15. Wave 4.5 unweighted and weighted interview response rates for the Wave 4 Cohort**

Type of Interview	C Completed (n)	I Ineligible (n)	N Nonresponse known to be eligible (n)	U Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	RR: Weighted response rate (%)
Youth interview	12,918	107	1,252	246	89.6	89.1

Table 4-16 provides unweighted and weighted response rates for the Wave 5 adult and youth interviews among the Wave 4 Cohort. The Wave 5 weighted response rates for the Wave 4 Cohort were 88.0 percent for adults and 83.5 percent for youth.

**Table 4-16. Wave 5 unweighted and weighted interview response rates for the Wave 4 Cohort**

Type of Interview	C Completed (n)	I Ineligible (n)	N Nonresponse known to be eligible (n)	U Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	RR: Weighted response rate (%)
Adult interview	32,687	664	3,274	1,925	86.4	88.0
Youth interview	11,976	48	1,508	637	84.8	83.5

<sup>23</sup>The PATH-ATS probability of selection is described in equation B.34 in Appendix B. As described in Section 2.7, participants eligible for the PATH-ATS sample were ages 20 and older on August 31, 2020, part of the Wave 4 Cohort, and respondents to the Wave 5 adult interview.

Table 4-17 shows the unweighted and weighted response rates for the Wave 5.5 adult and youth interviews among the Wave 4 Cohort. Only Wave 4 Cohort members who were ages 19 and younger on August 31, 2020 and had not become permanently ineligible at Wave 5 or earlier were included in the calculation of equation 4.2.2. The numbers of participants discovered to be permanently ineligible before resumption of Wave 5.5 are shown in Table 4-16 for reference. The Wave 5.5 weighted response rates for the Wave 4 Cohort were 69.9 percent for adults and 66.8 percent for youth.

**Table 4-17. Wave 5.5 unweighted and weighted interview response rates for the Wave 4 Cohort**

Type of Interview	C Completed (n)	I/P: Permanently ineligible prior to Wave 5.5 <sup>a</sup> (n)	A Other ineligible (n)	N: Nonresponse known to be eligible (n)	U: Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	RR: Weighted response rate (%)
Adult interview	3,628	11	23	634	922	70.1	69.9
Youth interview	7,129	18	33	1,302	2,122	67.6	66.8

<sup>a</sup> These cases were not included in equation 4.2.2.

Table 4-18 shows the unweighted and weighted response rates for the PATH-ATS interviews among the Wave 4 Cohort. As described in Section 2.7, participants eligible for the PATH-ATS sample were ages 20 and older on August 31, 2020, part of the Wave 4 Cohort, and respondents to the Wave 5 adult interview. The weighted PATH-ATS response rate was 55.6 percent.

**Table 4-18. PATH-ATS unweighted and weighted response rates for the Wave 4 Cohort**

Type of Interview	C Completed (n)	A Ineligible (n)	N: Nonresponse known to be eligible (n)	U: Nonresponse with unknown eligibility (n)	RR: Unweighted response rate (%)	RR: Weighted response rate (%)
PATH-ATS	8,874	109	3,313	6,305	48.1	55.6

Table 4-19 shows the unweighted and weighted response rates for the Wave 6 adult and youth interviews among the Wave 4 Cohort. Only Wave 4 Cohort members who had not become permanently ineligible before Wave 6 were included in the calculation of equation 4.2.2. The numbers of participants discovered to be permanently ineligible before Wave 6 are shown in Table 4-19 for reference. The Wave 6 weighted response rates for the Wave 4 Cohort were 73.5 percent for adults and 63.6 percent for youth.

**Table 4-19. Wave 6 unweighted and weighted interview response rates for the Wave 4 Cohort**

Type of Interview	C Completed (n)	I/P Permanently Ineligible prior to Wave 6 <sup>a</sup> (n)	A Other Ineligible (n)	N Nonresponse known to be eligible (n)	U Nonresponse with unknown eligibility (n)	RR Unweighted response rate (%)	RR Weighted response rate (%)
Adult interview	29,516	469	881	6,822	6,367	69.4	73.5
Youth interview	5,585	14	34	1,242	1,800	64.8	63.6

<sup>a</sup> These cases were not included in equation 4.2.2.

Table 4-20 shows the unweighted and weighted response rates for the Wave 7 adult and youth interviews among the Wave 4 Cohort. Only Wave 4 Cohort members who had not become permanently ineligible before Wave 7 were included in the calculation of equation 4.2.2. The numbers of participants discovered to be permanently ineligible before Wave 7 are shown in Table 4-20 for reference. The Wave 7 weighted response rates for the Wave 4 Cohort were 66.9 percent for adults and 55.9 percent for youth.

**Table 4-20. Wave 7 unweighted and weighted interview response rates for the Wave 4 Cohort**

Type of Interview	C Completed (n)	I/P Permanently Ineligible prior to Wave 7 <sup>a</sup> (n)	A Other Ineligible (n)	N Nonresponse known to be eligible (n)	U Nonresponse with unknown eligibility (n)	RR Unweighted response rate (%)	RR Weighted response rate (%)
Adult interview	28,165	1,191	463	6,680	10,165	62.8	66.9
Youth interview	3,452	32	8	778	1,796	57.3	55.9

<sup>a</sup> These cases were not included in equation 4.2.2.

Response rates for Wave 4.5 through 7 differed across groups based on data from the Wave 4 interviews (e.g., age, sex, race/ethnicity, and tobacco use). Response rates also differed by recruitment wave. The weighting adjustments attempted to compensate for differential response rates across groups, as described in Chapter 5 and Appendix B.

## 4.5 Response Rates for the Wave 7 Replenishment Sample

This section summarizes the household screener response rates for the AYS replenishment subsample and the entire replenishment sample at Wave 7. Moreover, it reviews the adult interview

response rate for the Wave 7 AYS replenishment subsample, as well as the youth interview response rate and parent consent for shadow youth response rate for the entire Wave 7 replenishment sample.

Both household screener response rates were calculated using equation 4.1.1, whereas the adult interview response rate was calculated using equation 4.1.2. The youth interview and parent consent for shadow youth response rates were calculated using equations 4.1.3 and 4.1.4, respectively.

Unweighted response rates were calculated based on actual case counts, whereas weighted response rates were based on the sums of the Wave 7 IPS weights, which are akin to the Wave 4 IPS weights described in Section B.4.3.1 in Appendix B. At the adult level, youth level, and shadow youth level, the IPS weight is the product of the household IPS weight and the inverse of the within-household probability of selection. Although it was not designed to generate estimates on its own, the Wave 7 replenishment sample was a probability-based nationally representative sample, so the weighted response may be interpreted as described in Section 4.1.

Tables 4-21 through 4-23 show the unweighted and weighted household screener and adult interview response rates for the Wave 7 AYS replenishment subsample, as well as the unweighted and weighted response rates for the household screener, youth interview, and parent consent for shadow youth for the entire Wave 7 replenishment sample. The weighted response rates for the Wave 7 AYS replenishment subsample were 30.3 percent and 55.1 percent for the household screener and adult interview, respectively. The weighted response rates for the entire Wave 7 replenishment sample were 55.6 percent, 64.0 percent, and 77.9 percent for the household screener, youth interview, and parent consent for shadow youth, respectively. Response rates differed across subgroups. Differential response rates were compensated for by using nonresponse adjustments in the weighting methods, as described in Chapter 5 and Appendix B.

**Table 4-21. Unweighted and weighted household screener response rates for the Wave 7 replenishment samples**

Wave 7 replenishment sample group	$C_{HH}$ : Completed (n)	$N_{HH}$ : Nonresponse known to be eligible (n)	$e_{HH} \times U_{HH}$ : Unknown eligibility estimated to be eligible (n)	$RR_{HH}$ : Unweighted response rate (%)	$RR_{HH}$ : Weighted response rate (%)
AYS replenishment subsample	10,013	20,028	2,741	30.5	30.3
Entire replenishment sample	125,706	84,953	14,891	55.7	55.6

**Table 4-22.** Wave 7 unweighted and weighted adult interview response rates for the Wave 7 AYS replenishment subsample

Type of interview	Phase 2 screener		Adult interview		RR <sub>A</sub> : Unweighted response rate (%)	RR <sub>A</sub> : Weighted response rate (%)
	C <sub>P2</sub> : P2 screener, completed (n)	N <sub>P2</sub> : P2 screener, nonresponse (n)	C <sub>A</sub> : Adult interview, completed (n)	N <sub>A</sub> : Adult interview, nonresponse (n)		
Adult interview	2,404	1,953	1,803	0	55.2	55.1

**Table 4-23.** Wave 7 unweighted and weighted youth interview and shadow youth response rates for the entire Wave 7 replenishment sample

Type of interview <sup>a</sup>	C <sub>Y</sub> : Completed (n)	N <sub>Y</sub> : Nonresponse (n)	RR <sub>Y</sub> : Unweighted response rate (%)	RR <sub>Y</sub> : Weighted response rate (%)
Youth interview	7,321	4,006	64.6	64.0
Shadow youth (parent consent)	5,739	1,538	78.9	77.9

<sup>a</sup> For shadow youth calculations, please refer to the components of equation 4.1.4 with subscript SY rather than Y as indicated in the table headers.

## 5. Weights and Imputation

---

Analysis of data from complex sample designs, such as the PATH Study design, requires the use of weights to compensate for variable probabilities of selection, differential nonresponse rates, and possible deficiencies in the sampling frame (e.g., undercoverage of certain population groups). It is also necessary to implement variance estimation procedures that appropriately account for sampling design factors (such as the stratification and sampling of PSUs and area segments and the use of oversampling) and nonresponse adjustment factors. Section 5.1 describes the adult and youth weights available for analysis; the procedures used to calculate the adult and youth weights are described in Appendix B. Section 5.2 describes methods that should be used to estimate variances. The weights and variance estimation methods permit correct inferences for analyses of the PATH Study adult data and youth/parent data. Section 5.3 and Appendix D describe the imputed variables included in the adult and youth/parent data files.

### 5.1 Adult and Youth Weights

Appendix B describes the procedures used to calculate the adult and youth weights for each wave and cohort. Cross-sectional and longitudinal (single-wave or all-waves) adult and youth weights were created to support analyses of each cohort of the PATH Study. Weights for shadow youth not old enough to complete an interview in a given wave were created (but not released with the data files) to support the formation of weights in subsequent waves. For some waves, there are multiple weight types available. The different weight types are described below.

- **Cross-sectional weights:** These were created for each wave at which a cohort was formed and were assigned to each respondent at that wave in the cohort, (e.g., all Wave 1 respondents in the Wave 1 Cohort).
- **Single-wave weights:** These were created for each wave after the formation of a cohort, and until a newer cohort is formed the single-wave weights are assigned to each respondent at that wave in that cohort, regardless of their participation in previous waves (e.g., all Wave 3 interview respondents, including those in the shadow sample at Wave 1, received a Wave 3 single-wave weight for the Wave 1 Cohort). After the establishment of a newer cohort, the single-wave weights for an older cohort are only assigned to respondents who completed an interview at the first wave of the older cohort (e.g., only Wave 5 respondents who completed an interview at Wave 1 received a Wave 5 single-wave weight for the Wave 1 Cohort).

- **All-waves weights:** Starting with the second wave after the formation of a cohort (e.g., Wave 3 for the Wave 1 Cohort), these were created for respondents who also responded at all previous waves relevant to the cohort.

For primary waves, this includes respondents at that wave who either:

- Completed an interview (adult or youth) at all previous primary waves relevant to the cohort; or
- Were in the shadow sample at the start of the cohort, were at least age 12 at the time of the wave, and completed an interview for all waves relevant to the cohort in which they were old enough to do so, or verified their information with the study for waves in which they were not old enough to be interviewed, excluding special collections.

For special collections (Wave 4.5, Wave 5.5, and PATH-ATS), this includes respondents at that wave who either:

- Completed an interview (adult or youth) at all previous waves relevant to the cohort and special collection as appropriate (e.g., including Wave 4.5 for the Wave 5.5 all-waves weights, and excluding Wave 4.5 for the PATH-ATS all-waves weights); or
- Were in the shadow sample at the start of the cohort, were at least age 12 at the time of the wave, and completed an interview for all waves relevant to the cohort and special collection in which they were old enough to do so, or verified their information with the study for waves in which they were not old enough to be interviewed.<sup>24</sup>

- **Special collection all-waves weights:** Starting with Wave 5, special collection all-waves weights were created in each primary wave for respondents at that wave who also responded at all previous waves relevant to the cohort and in all previous special data collections as appropriate. This includes respondents who either:

- Completed an interview at all previous waves relevant to the cohort, including previous special collections as appropriate (e.g., Wave 4 and Wave 4.5 for the Wave 4 Cohort special collection all-waves weights for Wave 5); or
- Were in the shadow sample at the start of the cohort, were at least age 12 at the time of the wave, and completed an interview for all waves relevant to the cohort in which they were old enough to do so, or verified their information with the study for waves in which they were not old enough to be interviewed including previous special collections as appropriate.<sup>24</sup>

---

<sup>24</sup>By design, Wave 4 Cohort shadow youth who were age 10 at Wave 4 were not contacted for the Wave 4.5 special data collection. However, if they completed an interview (when old enough to do so) or verified their information with the study (when not old enough to be interviewed) for all other waves relevant to the Wave 4 Cohort, they were assigned this weight.

The weight variables for each wave, cohort, weight type, and interview type are summarized in Table 5-1. Researchers are encouraged to review Chapter 6 for guidelines on selecting the appropriate weight for their analysis.

## 5.2 Variance Estimation

Accompanying the adult data files and youth/parent data files are replicate weights that can be used to estimate variances. The replication variance estimation approach is the preferred method for estimating variances from PATH Study data. The replicate weight variables for each wave, cohort, weight type, and interview type are summarized in Table 5-1.

**Table 5-1. Full sample and replicate weight variables for each wave, cohort, weight type, and interview type**

Wave	Cohort	Weight Type <sup>a</sup>	Interview type	Full sample weight variable name <sup>b</sup>	Replicate weight variable names <sup>b</sup>
1	Wave 1	Cross-sectional	Adult	R01_A_PWGT	R01_A_PWGT1 – R01_A_PWGT100
			Youth	R01_Y_PWGT	R01_Y_PWGT1 – R01_Y_PWGT100
2	Wave 1	Single-wave/All-waves	Adult	R02_A_PWGT	R02_A_PWGT1 – R02_A_PWGT100
			Youth	R02_Y_PWGT	R02_Y_PWGT1 – R02_Y_PWGT100
3	Wave 1	All-waves	Adult	R03_A_AWGT	R03_A_AWGT1 – R03_A_AWGT100
			Youth	R03_Y_AWGT	R03_Y_AWGT1 – R03_Y_AWGT100
3	Wave 1	Single-wave	Adult	R03_A_SWGT	R03_A_SWGT1 – R03_A_SWGT100
			Youth	R03_Y_SWGT	R03_Y_SWGT1 – R03_Y_SWGT100
4	Wave 1	All-waves	Adult	R04_A_A01WGT	R04_A_A01WGT1 – R04_A_A01WGT100
			Youth	R04_Y_A01WGT	R04_Y_A01WGT1 – R04_Y_A01WGT100
4	Wave 1	Single-wave	Adult	R04_A_S01WGT	R04_A_S01WGT1 – R04_A_S01WGT100
			Youth	R04_Y_S01WGT	R04_Y_S01WGT1 – R04_Y_S01WGT100
4	Wave 4	Cross-sectional	Adult	R04_A_C04WGT	R04_A_C04WGT1 – R04_A_C04WGT100
			Youth	R04_Y_C04WGT	R04_Y_C04WGT1 – R04_Y_C04WGT100

**Table 5-1.** Full sample and replicate weight variables for each wave, cohort, weight type, and interview type (continued)

<b>Wave</b>	<b>Cohort</b>	<b>Weight Type<sup>a</sup></b>	<b>Interview type</b>	<b>Full sample weight variable name<sup>b</sup></b>	<b>Replicate weight variable names<sup>b</sup></b>
4.5	Wave 1	All-waves	Youth	X04_Y_A01WGT	X04_Y_A01WGT1 - X04_Y_A01WGT100
4.5	Wave 1	Single-wave	Youth	X04_Y_S01WGT	X04_Y_S01WGT1 - X04_Y_S01WGT100
4.5	Wave 4	Single-wave/ All-waves	Youth	X04_Y_S04WGT	X04_Y_S04WGT1 - X04_Y_S04WGT100
5	Wave 1	All-waves	Adult	R05_A_A01WGT	R05_A_A01WGT1 - R05_A_A01WGT100
			Youth	R05_Y_A01WGT	R05_Y_A01WGT1 - R05_Y_A01WGT100
5	Wave 1	Special collection all-waves	Adult <sup>c</sup>	R05_A_AX01WGT	R05_A_AX01WGT1 - R05_A_AX01WGT100
			Youth	R05_Y_AX01WGT	R05_Y_AX01WGT1 - R05_Y_AX01WGT100
5	Wave 1	Single-wave	Adult	R05_A_S01WGT	R05_A_S01WGT1 - R05_A_S01WGT100
			Youth	R05_Y_S01WGT	R05_Y_S01WGT1 - R05_Y_S01WGT100
5	Wave 4	Special collection all-waves	Adult <sup>c</sup>	R05_A_AX04WGT	R05_A_AX04WGT1 - R05_A_AX04WGT100
			Youth	R05_Y_AX04WGT	R05_Y_AX04WGT1 - R05_Y_AX04WGT100
5	Wave 4	Single-wave/ All-waves	Adult	R05_A_S04WGT	R05_A_S04WGT1 - R05_A_S04WGT100
			Youth	R05_Y_S04WGT	R05_Y_S04WGT1 - R05_Y_S04WGT100
5.5	Wave 1	All-waves	Adult	X05_A_A01WGT	X05_A_A01WGT1 - X05_A_A01WGT100
			Youth	X05_Y_A01WGT	X05_Y_A01WGT1 - X05_Y_A01WGT100
5.5	Wave 4	All-waves	Adult	X05_A_A04WGT	X05_A_A04WGT1 - X05_A_A04WGT100
			Youth	X05_Y_A04WGT	X05_Y_A04WGT1 - X05_Y_A04WGT100
5.5	Wave 4	Single-wave	Adult	X05_A_S04WGT	X05_A_S04WGT1 - X05_A_S04WGT100
			Youth	X05_Y_S04WGT	X05_Y_S04WGT1 - X05_Y_S04WGT100
PATH-ATS	Wave 1	All-waves	Adult	T05_A_A01WGT	T05_A_A01WGT1 - T05_A_A01WGT100
PATH-ATS	Wave 4	All-waves	Adult	T05_A_A04WGT	T05_A_A04WGT1 - T05_A_A04WGT100
6	Wave 1	All-waves	Adult	R06_A_A01WGT	R06_A_A01WGT1 - R06_A_A01WGT100
			Youth	R06_Y_A01WGT	R06_Y_A01WGT1 - R06_Y_A01WGT100
6	Wave 1	Special collection all-waves	Adult	R06_A_AX01WGT	R06_A_AX01WGT1 - R06_A_AX01WGT100
			Youth	R06_Y_AX01WGT	R06_Y_AX01WGT1 - R06_Y_AX01WGT100

**Table 5-1.** Full sample and replicate weight variables for each wave, cohort, weight type, and interview type (continued)

<b>Wave</b>	<b>Cohort</b>	<b>Weight Type<sup>a</sup></b>	<b>Interview type</b>	<b>Full sample weight variable name<sup>b</sup></b>	<b>Replicate weight variable names<sup>b</sup></b>
6	Wave 1	Single-wave	Adult	R06_A_S01WGT	R06_A_S01WGT1 - R06_A_S01WGT100
			Youth	R06_Y_S01WGT	R06_Y_S01WGT1 - R06_Y_S01WGT100
6	Wave 4	All-waves	Adult	R06_A_A04WGT	R06_A_A04WGT1 - R06_A_A04WGT100
			Youth	R06_Y_A04WGT	R06_Y_A04WGT1 - R06_Y_A04WGT100
6	Wave 4	Special collection all-waves	Adult	R06_A_AX04WGT	R06_A_AX04WGT1 - R06_A_AX04WGT100
			Youth	R06_Y_AX04WGT	R06_Y_AX04WGT1 - R06_Y_AX04WGT100
6	Wave 4	Single-wave	Adult	R06_A_S04WGT	R06_A_S04WGT1 - R06_A_S04WGT100
			Youth	R06_Y_S04WGT	R06_Y_S04WGT1 - R06_Y_S04WGT100
7	Wave 1	All-waves	Adult	R07_A_A01WGT	R07_A_A01WGT1 - R07_A_A01WGT100
			Youth	R07_Y_A01WGT	R07_Y_A01WGT1 - R07_Y_A01WGT100
7	Wave 1	Special collection all-waves	Adult	R07_A_AX01WGT	R07_A_AX01WGT1 - R07_A_AX01WGT100
			Youth	R07_Y_AX01WGT	R07_Y_AX01WGT1 - R07_Y_AX01WGT100
7	Wave 1	Single-wave	Adult	R07_A_S01WGT	R07_A_S01WGT1 - R07_A_S01WGT100
			Youth	R07_Y_S01WGT	R07_Y_S01WGT1 - R07_Y_S01WGT100
7	Wave 4	All-waves	Adult	R07_A_A04WGT	R07_A_A04WGT1 - R07_A_A04WGT100
			Youth	R07_Y_A04WGT	R07_Y_A04WGT1 - R07_Y_A04WGT100
7	Wave 4	Special collection all-waves	Adult	R07_A_AX04WGT	R07_A_AX04WGT1 - R07_A_AX04WGT100
			Youth	R07_Y_AX04WGT	R07_Y_AX04WGT1 - R07_Y_AX04WGT100
7	Wave 4	Single-wave	Adult	R07_A_S04WGT	R07_A_S04WGT1 - R07_A_S04WGT100
			Youth	R07_Y_S04WGT	R07_Y_S04WGT1 - R07_Y_S04WGT100
7	Wave 7	Cross-sectional	Adult	R07_A_C07WGT	R07_A_C07WGT1 - R07_A_C07WGT100
			Youth	R07_Y_C07WGT	R07_Y_C07WGT1 - R07_Y_C07WGT100

<sup>a</sup> See the introduction to Section 5.1 for descriptions of the various weight types.

<sup>b</sup> Full sample weight, replicate weight, pseudo-strata, and pseudo-PSU variables are included on the files with the questionnaire data for Wave 1 and Wave 2. For all other waves, these variables are on separate files corresponding to the respective wave, cohort, weight type, and interview type.

<sup>c</sup> Adults with Wave 5 special collection all-waves weights completed youth interviews at Wave 4 and Wave 4.5.

Replication variance estimation methods are used to provide analysts with a method for calculating standard errors of statistics. They produce consistent estimators of the variance for statistics that are smooth functions of estimated totals (Krewski and Rao, 1981); these include most commonly used statistics such as means, ratios, linear/logistic/Poisson regression coefficients, correlation coefficients, and many measures of association for categorical data. In most complex designs, such as the multi-stage sample design used in the PATH Study, the variance is estimated by assuming that the first-stage sampling is performed with replacement (Wolter, 2007).

The basic idea behind replication is to select subsamples repeatedly from the whole sample, calculate the statistic of interest for each subsample, and then use these subsamples or replicate statistics to estimate the variance of the full-sample statistic. Different ways of creating subsamples from the full sample result in different replication methods. The subsamples are called replicates and the statistics calculated from these replicates are called replicate estimates.

One major advantage of replication methods is that they can produce variance estimates for statistics that might not be available in standard software. Another advantage is that the replication variance estimation provides a simple way to account for adjustments that are made in weighting. As described above, the PATH Study full-sample weights are adjusted for nonresponse, trimming, and raking to control totals. By separately computing the weighting adjustments for each replicate, it is possible to reflect the effects of weight adjustments in the estimates of variance, which frequently results in a smaller variance because certain estimates have been calibrated to known control totals (Valliant, 1993). Taylor series (linearization) methods for estimating the variance do not account for the variance reduction resulting from raking, and consequently often (but not always) provide variance estimates that are too large (Valliant, 2004; Chowdhury, 2013).

There are several ways of forming replicate weights, including balanced repeated replication (BRR), jackknife (JK-1, JK-2, and JK-n), and bootstrap. The choice of what kind of replicate weights to create is determined by the type of sampling design that was used to collect the data; in particular, whether or not stratification was used and how many PSUs were selected in each stratum. The replication method selected for the PATH Study is BRR (McCarthy, 1969).

The BRR method was selected for the PATH Study because (1) it allows calculation of the variance with fewer replicate weight variables than would be needed for the bootstrap (this is possible

because the subsets of PSUs for the replicates are carefully selected according to an orthogonal experimental design) and (2) BRR, unlike jackknife, produces consistent variance estimates for non-smooth statistics such as quantiles (Rao and Shao, 1993, 1999). The BRR method also allows correct estimation of standard errors for analyses involving quantiles and quantile regressions. This may be of interest for some of the biomarkers assessed by the PATH Study, such as the 10<sup>th</sup> percentile or the median of cotinine level. The PATH Study uses a variant of BRR known as Fay's method (Judkins, 1990). Fay's method produces more stability for the variance estimates for quantities in domains with small sample sizes. Using Fay's method, one half of the sample is weighted down by a factor  $\epsilon$  set to 0.3 for the PATH Study) and the remaining half is weighted up by a compensating factor  $2 - \epsilon$ .

Suppose that  $\hat{\theta}$  is the full-sample estimate of a population parameter  $\theta$ , which can be a smooth or non-smooth function of a linear estimator. Then the variance estimator  $v(\hat{\theta})$  takes the form as shown in equation 5.2.1:

$$v(\hat{\theta}) = c \sum_{g=1}^G (\hat{\theta}_{(g)} - \hat{\theta})^2 \quad (5.2.1)$$

where

$\hat{\theta}$  is the estimate of  $\theta$  calculated using the full-sample weight,

$\hat{\theta}_{(g)}$  is the estimate of  $\theta$  calculated using the  $g$ -th replicate weight,

$G$  is the total number of replicates formed, and

$c$  is a constant that depends on the replication method.

For standard BRR,  $c = \frac{1}{G}$ .

For Fay's method,  $c = \frac{1}{G(1-\epsilon)^2}$ , where  $\epsilon$  is the factor used in Fay's method.

For the PATH Study, the value of the constant to be used when calculating the variance is

$$c = \frac{1}{100(1-0.3)^2} = 0.020408. \quad (5.2.2)$$

### 5.2.1 Creation of Variables for Variance Estimation

The first step in constructing the replicate weights was to create variables for pseudo-strata and pseudo-PSUs that reflect the variance structure. These were created using methods described in Korn and Graubard (2011, p. 206) and are called VARSTRAT and VARPSU on the data files.

There is a total of 92 strata and 156 PSUs in the PATH Study sampling design. Fifty of the strata have two PSUs sampled; these were left as is for the variance estimation. Seven strata have three PSUs sampled. Because the BRR method assumes that two PSUs are selected from each stratum, two pseudo-PSUs were created for each of these seven strata by randomly selecting two of the three PSUs in each stratum and combining them into one pseudo-PSU. The remaining PSU then became the second pseudo-PSU in the stratum.

Thirty-five of the strata have one PSU selected with certainty and called self-representing (SR). In these strata, the variability comes from the secondary sampling units (segments), which were selected from a list of segments ordered such that similar segments were close together (see Section 2.1.2.2). The following procedure was used to create pseudo-strata and pseudo-PSUs within the SR PSUs for variance estimation purposes.

- Each SR PSU with fewer than 60 segments was treated as one pseudo-stratum with two pseudo-PSUs. The segments in the SR PSU were assigned to the two pseudo-PSUs so that the odd-numbered segments in the sorted list were assigned to one pseudo-PSU and the even-numbered segments in the sorted list were assigned to the other pseudo-PSU. In this way, adjacent segments in the sorted list were assigned to different pseudo-PSUs.
- Four of the SR PSUs had large numbers of segments selected, and these were divided into pseudo-strata for variance estimation purposes. The pseudo-strata were formed by cutting the ordered list of segments into the desired number of pseudo-strata. Then, the segments within the pseudo-strata were assigned to pseudo-PSUs as described above.

Applying the BRR method to the PATH Study's pseudo-strata and the initial household weights yields 100 initial replicate weights. These 100 initial replicate weights were adjusted using the steps described in Appendix B for the full-sample weights to arrive at the final set of corresponding replicate weights for variance estimation.

## 5.2.2 Software Options

The data files are provided in several formats. When the data analysis software package allows, the replicate weights should be used for all variance calculations to reflect the complex sample design and the various weighting adjustments on standard errors. Note that, although variables for pseudo-strata (VARSTRAT) and pseudo-PSUs (VARPSU) are included with the data files, variance estimates calculated using these variables with linearization (for example, by using the STRATA and CLUSTER statements in SAS<sup>®</sup>) **do not** reflect the impact of the weighting adjustments and may result in incorrect inferences. Some example SAS, SUDAAN<sup>®</sup>, Stata<sup>®</sup>, R, and SPSS<sup>®</sup> program code for generating popular statistics is provided in Appendix C.1.

The BRR replication method of variance estimation is available in both the SAS and Stata software packages. SPSS Complex Samples™ currently does not offer BRR or other forms of replication-based estimation. Note that the open-source R statistical software language does not have core programs that do survey data analysis. There are various contributed packages to R (such as the survey package) that analyze data from complex surveys and handle replication methods for variance estimation; however, these contributed packages are not peer-reviewed or subject to quality standards, and their features are subject to change in future versions of R.

Although software packages do not universally accommodate replicate weights for all analytic methods, the replication method can be applied by repetition to any analytic routine. That is, the desired analysis would first be run using the full-sample weight. Then, it would be repeated replacing the full-sample weight by each replicate weight, in turn (i.e., 100 times for the PATH Study data). The formula for BRR variance estimates (provided in equation 5.2.1) would then be used to estimate the variance of any parameter (e.g., regression coefficient) of interest.

Variance estimates for small domains may be unstable because some PSUs may contain no observations belonging to the domain. For small domains, the variance estimates produced by statistical software packages may differ because they use different methods to adjust for strata in which only one PSU contains domain members. For more information, see Graubard and Korn (1996), Lohr (2010, p. 570), or Lewis (2013).

## 5.3 Imputation

Demographic characteristics of adults and youth selected for the PATH Study were used in the creation of the weights. These included variables indicating sex, age, highest education (adults only), race, and ethnicity of the sampled persons. However, because some of this information may be missing for some sampled adults and youth, imputation methods were used to assign values when self-reported information was not available.

### 5.3.1 Wave 1 Imputation

For both adults and youth, imputation was performed by first considering information provided in the household screener and then by using statistical imputation methods. The imputation methods were performed for respondents and nonrespondents<sup>25</sup> because data for both were needed for weighting, but only the values for respondents are included on the data files. There was no imputation performed for Waves 2 and 3 because the weighting process relied on demographic and other characteristics from Wave 1.

Methods used for imputing sex, age, education, race, and ethnicity for Wave 1 are described in details in Appendix D.1.

### 5.3.2 Wave 4 Imputation

Respondents to Wave 4 of the PATH Study were from two separate samples selected at different times: the Wave 4 replenishment sample and the Wave 1 sample. The imputation approach differed for these two groups. For respondents from the replenishment sample, imputation was performed by first considering information provided in the Wave 4 household screener and then by using statistical imputation methods. For respondents from the Wave 1 Cohort, imputed values from Wave 1 were used in addition to new statistical imputation after considering information provided in waves subsequent to Wave 1. Even though Wave 1 Cohort members who are not in the Wave 4 Cohort were excluded from the Wave 4 Cohort cross-sectional weighting process, they were

---

<sup>25</sup>These are adults and youth belonging to responding households who did not respond to the interview.

included in the imputation process so that the imputed variables could be made available for all Wave 4 respondents.

Non-missing values for sex, age, highest education (adults only), race, and ethnicity were needed for raking the Wave 4 Cohort cross-sectional weights to control totals. Imputation was performed only for Wave 4 respondents because nonrespondents are not included in the raking procedure.

All Wave 4 respondents had non-missing values for age. Methods used for imputing sex, education, race, and ethnicity for Wave 4 are described in details in Appendix D.2.

### **5.3.3     Wave 7 Imputation**

Respondents to Wave 7 of the PATH Study were from three separate groups selected at different times. Those from the Wave 7 replenishment sample were selected for the study at Wave 7 and those from the “Wave 4 Cohort” were selected for the study either at Wave 1 or Wave 4.

Additionally there are Wave 7 respondents who are members of the Wave 1 Cohort only, all selected at Wave 1. The imputation approach differed for these three groups. For respondents from the Wave 7 replenishment sample, imputation was performed by first considering information provided in the Wave 7 household screener and then by using statistical imputation methods. For other respondents, imputed values from Wave 4 and Wave 1 were used in addition to new statistical imputation after considering information provided in other waves. Even though respondents who are not in the Wave 7 Cohort were excluded from the Wave 7 Cohort cross-sectional weighting process, they were included in the imputation process so that the imputed variables could be made available for all Wave 7 respondents.

Non-missing values for sex, age, highest education (adults only), race, and ethnicity were needed for raking the Wave 7 Cohort cross-sectional weights to control totals. Imputation was performed only for Wave 7 respondents because nonrespondents are not included in the raking procedure.

Note that there were no missing values for age among Wave 7 respondents, so no age imputation was needed. Methods used for imputing sex, education, race, and ethnicity for Wave 7 are described in detail in Appendix D.3.

## 6. Analytic Considerations

---

This chapter provides information to assist users in determining appropriate approaches for the analysis of the PATH Study data and the weights to use with each approach. For example, estimates can be made comparing tobacco use reported in the latest wave to that used in Wave 1 or comparing all prior waves for the same set of individuals. These are longitudinal comparisons, and the weights used would be associated with the latest wave.

As another example, estimates can be obtained for the CNP at Wave 4, and these can be compared to the same estimates representing the CNP at a prior wave, based on those who responded at that wave. This is a cross-sectional comparison and some participants may have responded in one wave but not the other. The PATH Study was not designed with a focus on cross-sectional estimation and comparisons. However, if variance calculations are correctly specified, they will correctly reflect the correlation between the overlapping sets of respondents and the estimates will be appropriate but approximate. Cross-sectional comparisons between waves should be undertaken with care, and are not appropriate in all situations. See Section 6.3.3 for more information.

Section 6.1 considers different designs that observational studies may use to measure change over time. Sections 6.2 and 6.3 present the target populations of inference and guidance for selecting the appropriate weights for the PATH Study. Regardless of the analysis, it is important that analysts use the appropriate framework and methods of estimating standard errors to appropriately reflect the complex sample design employed for the PATH Study. See Appendix C.1 for example program code illustrating the correct specifications for creating appropriate variance estimates that reflect this complex sample design. Section 6.4 describes the impact of data collection methods. Section 6.5 provides guidance for analyses involving age. Section 6.6 provides considerations for pooling data. Section 6.7 discusses considerations for significance testing for small domains.

### 6.1 Background

The PATH Study is a longitudinal cohort study, not a repeated cross-sectional study. In a repeated cross-sectional study, such as the National Health Interview Survey,<sup>26</sup> a sample of persons is selected

---

<sup>26</sup>U.S. DHHS, CDC/National Center for Health Statistics.

at time 1, and that sample is used to produce estimates for variables of interest at time 1. For example, an estimate of interest might be the percentage of 18-year-olds who use ENDS. Then, at time 2, a new sample is selected, and the percentage of 18-year-olds using ENDS is estimated from that new sample. The sampling may be repeated as many times as desired. Each sample provides a snapshot of what is happening in the target population or subpopulation at the time the sample is taken, but it does not provide information about a particular sample at other times. Change is estimated by comparing the time-1 sample with the time-2 sample. Thus, repeated cross-sectional samples may be able to provide information on whether ENDS usage among the subpopulation of 18-year-olds increased or decreased at time 2 compared to time 1. However, they do not permit analysis of change for the same set of persons over time, for instance, as to whether 18-year-olds who use ENDS at time 1 are still using the product at time 2.

In a longitudinal cohort study, such as the PATH Study, a sample of persons is selected at time 1 (as in a repeated cross-sectional study). At time 2 and subsequent points in time, however, the same persons are re-interviewed. They may be asked the same questions as at time 1, and they may also be asked different questions at later interviews. Thus, a longitudinal study provides information on how individual participants in the study change over time, and allows the researcher to investigate the individual trajectories experienced by different people over time. A longitudinal study allows analysis of “within-person” change over time and can help researchers understand persistence of behavior or factors associated with changing behaviors and outcomes.

A longitudinal study can provide important precision gains for assessing change when compared with a repeated cross-sectional study. Assessing change over time “within person” takes maximum advantage of the correlation between measures on the same respondents over time to increase precision and power. An issue with a longitudinal study is that it does not include new members of the population such as recent immigrants. Rather, it includes only persons who were eligible for the study at the time of recruitment. To help address this issue, the PATH Study design includes periodic “refreshment” of the sample, randomly selecting new study participants from the CNP. The first refreshment (or replenishment) sample was selected at Wave 4 and the second was selected at Wave 7.

## 6.2 Target Populations

The target population for a study is the population for which the sample weights can be used for inferential purposes. The sample that represents the target population for a longitudinal study is generally referred to as a cohort and the target population for that sample at a given point in time is the population from which the sample was selected that remains eligible at the time in question. The PATH Study currently follows three cohorts, the Wave 1 Cohort, Wave 4 Cohort, and the Wave 7 Cohort. Descriptions of these cohorts and their target populations at each wave are in Chapter 2.

Generally, for the PATH Study, the target population for a given cohort is the CNP of a particular age range at the time of sampling. For the waves that follow, the target population is the same population who at that wave still resides in the United States and is not incarcerated. However, the participants at later waves do not represent some members of the CNP at the time of that wave. For example, at Wave 2, the Wave 1 Cohort does not represent immigrants who were not residents at Wave 1 or persons who were in the military or an institution (e.g., health care or incarcerated) at the time of Wave 1 who re-joined the CNP at Wave 2. This means that in follow-up waves, the PATH Study cohorts do not perfectly align with the CNP at the time of follow-up.

However, the subpopulations that form this “over” or “under” coverage are relatively small for follow-up waves close in time to the formation of the cohort. So, the weights developed for the cohort for some follow-up waves are suitable for producing useful cross-sectional estimates of the CNP at the time of the respective wave. For example, the weights developed for Wave 2 and Wave 3 are suitable for cross-sectional estimation for those waves. See Table 6-1 for the weights suitable for cross-sectional estimation for follow-up waves.

With the establishment of each new cohort, weights that truly represent the CNP at that wave are developed so those weights should be used for cross-sectional estimation for that wave and for subsequent waves until a newer cohort is established. For example, the Wave 4 weights for the Wave 4 Cohort should be used for cross-sectional estimation for the time of Wave 4. Because weights developed for the Wave 4 Cohort will more closely represent the population at the time of subsequent waves than weights for the Wave 1 Cohort, weights developed for the Wave 4 Cohort should be used for cross-sectional estimation for follow-up waves until a newer cohort is established, i.e., up until Wave 7. Similarly, the Wave 7 weights for the Wave 7 Cohort should be used for cross-sectional estimation for the time of Wave 7.

## 6.3 Selecting the Appropriate Weights

As the PATH Study is a longitudinal study, there will be many analyses examining within-person transitions over time with respect to the use of tobacco products, and changes in health, tobacco-related attitudes, and other factors. Thus, for most analyses, the focus will be on the characteristics of the respondents in the latest wave, how those may have changed from previous waves, and what factors may have been associated with such changes. For each wave after Wave 1, longitudinal weights are available to permit such analyses. These types of analyses are discussed more in Sections 6.3.1 and 6.3.2.

Cross-sectional weights are available for estimation of the Wave 1, Wave 4, and Wave 7 target populations. As discussed in Section 6.2, some weights, although longitudinal in nature, can be used to approximate the cross-sectional population for the respective wave. These types of analyses, and when they are appropriate, are discussed more in Sections 6.3.3 and 6.3.4.

Two longitudinal all-waves weights are publicly available for estimation using data from the PATH-ATS: one for the Wave 1 Cohort and another for the Wave 4 Cohort. Although designed for longitudinal analyses, the weights created for the Wave 4 Cohort can be used to approximate the cross-sectional population in the latter portion of 2020. The PATH-ATS weights also can be combined with the respective Wave 5.5 weights for longitudinal and cross-sectional analyses for the Wave 4 Cohort. These types of analyses are discussed more in Section 6.3.5.

For the PATH Study analyses, the recommended approach for computing sampling errors is to employ the replicate weights provided for each weight type. If an analyst chooses to use Taylor Series methods and estimates are being made for subdomains of interest, then all records in the full sample should be retained in the analysis data file and a domain analysis conducted so that the software package employed can appropriately compute variance estimates.

Table 6-1 summarizes the weights available for each wave and cohort, for different analysis types, including which PATH Study participants were assigned each weight. (Note that if there is only one set of longitudinal weights available for a wave that is not the first wave of a new cohort, such as with Wave 2, that set of weights serves as all-waves and single-wave weights.) The appropriate weight for any analysis depends on the PATH Study participants included in the analysis.

(e.g., Wave 2 adult interview respondents) and not any particular analysis tool (e.g., logistic regression).

**Table 6-1.** Descriptions of weights available for each wave and cohort

Wave	Cohort	Weight type <sup>a</sup>	Analysis type	Assigned to...
1	Wave 1	Cross-sectional	Cross-sectional	Wave 1 interview respondents in the Wave 1 Cohort; no youth selected in the shadow sample at Wave 1 have this weight
2	Wave 1	Single-wave/ All-waves	Longitudinal, cross-sectional <sup>c</sup>	Wave 2 interview respondents in the Wave 1 Cohort
3	Wave 1	All-waves	Longitudinal	Wave 3 interview respondents in the Wave 1 Cohort who also responded at Wave 2
3	Wave 1	Single-wave	Longitudinal, cross-sectional <sup>c</sup>	Wave 3 interview respondents in the Wave 1 Cohort
4	Wave 1	All-waves	Longitudinal	Wave 4 interview respondents in the Wave 1 Cohort who also responded at Wave 2 and Wave 3
4	Wave 1	Single-wave <sup>b</sup>	Longitudinal	Wave 4 interview respondents in the Wave 1 Cohort who completed a Wave 1 interview
4	Wave 4	Cross-sectional	Cross-sectional	Wave 4 interview respondents in the Wave 4 Cohort; no youth selected in the shadow sample at Wave 4 have this weight
4.5	Wave 1	All-waves	Longitudinal	Wave 4.5 youth interview respondents in the Wave 1 Cohort who also responded at Wave 2, Wave 3, and Wave 4
4.5	Wave 1	Single-wave <sup>b</sup>	Longitudinal	Wave 4.5 youth interview respondents in the Wave 1 Cohort who completed a Wave 1 interview
4.5	Wave 4	Single-wave/ All-waves	Longitudinal, cross-sectional <sup>c</sup>	Wave 4.5 youth interview respondents in the Wave 4 Cohort
5	Wave 1	All-waves	Longitudinal	Wave 5 interview respondents in the Wave 1 Cohort who also responded at Wave 2, Wave 3, and Wave 4
5	Wave 1	Special collection all-waves	Longitudinal	Wave 5 interview respondents in the Wave 1 Cohort who also participated in Wave 2, Wave 3, Wave 4, and Wave 4.5
5	Wave 1	Single-wave <sup>b</sup>	Longitudinal	Wave 5 interview respondents in the Wave 1 Cohort who completed a Wave 1 interview
5	Wave 4	Special collection all-waves	Longitudinal	Wave 5 interview respondents in the Wave 4 Cohort who either responded or were < age 12 at Wave 4.5
5	Wave 4	Single-wave/ All-waves	Longitudinal, cross-sectional <sup>c</sup>	Wave 5 interview respondents in the Wave 4 Cohort

**Table 6-1. Descriptions of weights available for each wave and cohort (continued)**

<b>Wave</b>	<b>Cohort</b>	<b>Weight type<sup>a</sup></b>	<b>Analysis type</b>	<b>Assigned to...</b>
5.5	Wave 1	All-waves	Longitudinal	Wave 5.5 telephone interview respondents in the Wave 1 Cohort who also responded at Wave 2, Wave 3, Wave 4, Wave 4.5, and Wave 5
5.5	Wave 4	All-waves	Longitudinal, cross-sectional <sup>d</sup>	Wave 5.5 telephone interview respondents in the Wave 4 Cohort who also responded at Wave 4.5 and Wave 5
5.5	Wave 4	Single-wave	Longitudinal, cross-sectional <sup>c</sup>	Wave 5.5 telephone interview respondents in the Wave 4 Cohort
PATH-ATS	Wave 1	All-waves	Longitudinal <sup>d</sup>	PATH-ATS interview respondents in the Wave 1 Cohort who also responded at Wave 2, Wave 3, Wave 4, and Wave 5
PATH-ATS	Wave 4	All-waves	Longitudinal, cross-sectional <sup>c,d</sup>	All PATH-ATS interview respondents
6	Wave 1	All-waves	Longitudinal	Wave 6 interview respondents in the Wave 1 Cohort who also participated in Wave 2, Wave 3, Wave 4, and Wave 5
6	Wave 1	Special collection all-waves	Longitudinal	Wave 6 interview respondents in the Wave 1 Cohort who also responded at Wave 2, Wave 3, Wave 4, Wave 4.5 (if applicable), Wave 5, and Wave 5.5/PATH-ATS
6	Wave 1	Single-wave <sup>b</sup>	Longitudinal	Wave 6 interview respondents in the Wave 1 Cohort who completed a Wave 1 interview
6	Wave 4	All-waves	Longitudinal	Wave 6 interview respondents in the Wave 4 Cohort who also responded at Wave 5
6	Wave 4	Special collection all-waves	Longitudinal	Wave 6 interview respondents in the Wave 4 Cohort who also responded at Wave 4.5 (if applicable), Wave 5, and Wave 5.5/PATH-ATS
6	Wave 4	Single-wave	Longitudinal, cross-sectional <sup>c</sup>	Wave 6 interview respondents in the Wave 4 Cohort
7	Wave 1	All-waves	Longitudinal	Wave 7 interview respondents in the Wave 1 Cohort who also responded at Wave 2, Wave 3, Wave 4, Wave 5, and Wave 6
7	Wave 1	Special collection all-waves	Longitudinal	Wave 7 interview respondents in the Wave 1 Cohort who also responded at Wave 2, Wave 3, Wave 4, Wave 4.5 (if applicable), Wave 5, Wave 5.5/PATH-ATS, and Wave 6
7	Wave 1	Single-wave <sup>b</sup>	Longitudinal	Wave 7 interview respondents in the Wave 1 Cohort who completed a Wave 1 interview
7	Wave 4	All-waves	Longitudinal	Wave 7 interview respondents in the Wave 4 Cohort who also responded at Wave 5 and Wave 6

**Table 6-1. Descriptions of weights available for each wave and cohort (continued)**

<b>Wave</b>	<b>Cohort</b>	<b>Weight type<sup>a</sup></b>	<b>Analysis type</b>	<b>Assigned to...</b>
7	Wave 4	Special collection all-waves	Longitudinal	Wave 7 interview respondents in the Wave 4 Cohort who also responded at Wave 4.5 (if applicable), Wave 5, Wave 5.5/PATH-ATS, and Wave 6
7	Wave 4	Single-wave	Longitudinal	Wave 7 interview respondents in the Wave 4 Cohort who completed a Wave 4 interview
7	Wave 7	Cross-sectional	Cross-sectional	Wave 7 interview respondents in the Wave 7 Cohort; no youth selected in the shadow sample at Wave 7 have this weight

<sup>a</sup> See the introduction to Section 5.1 for descriptions of the various weight types.

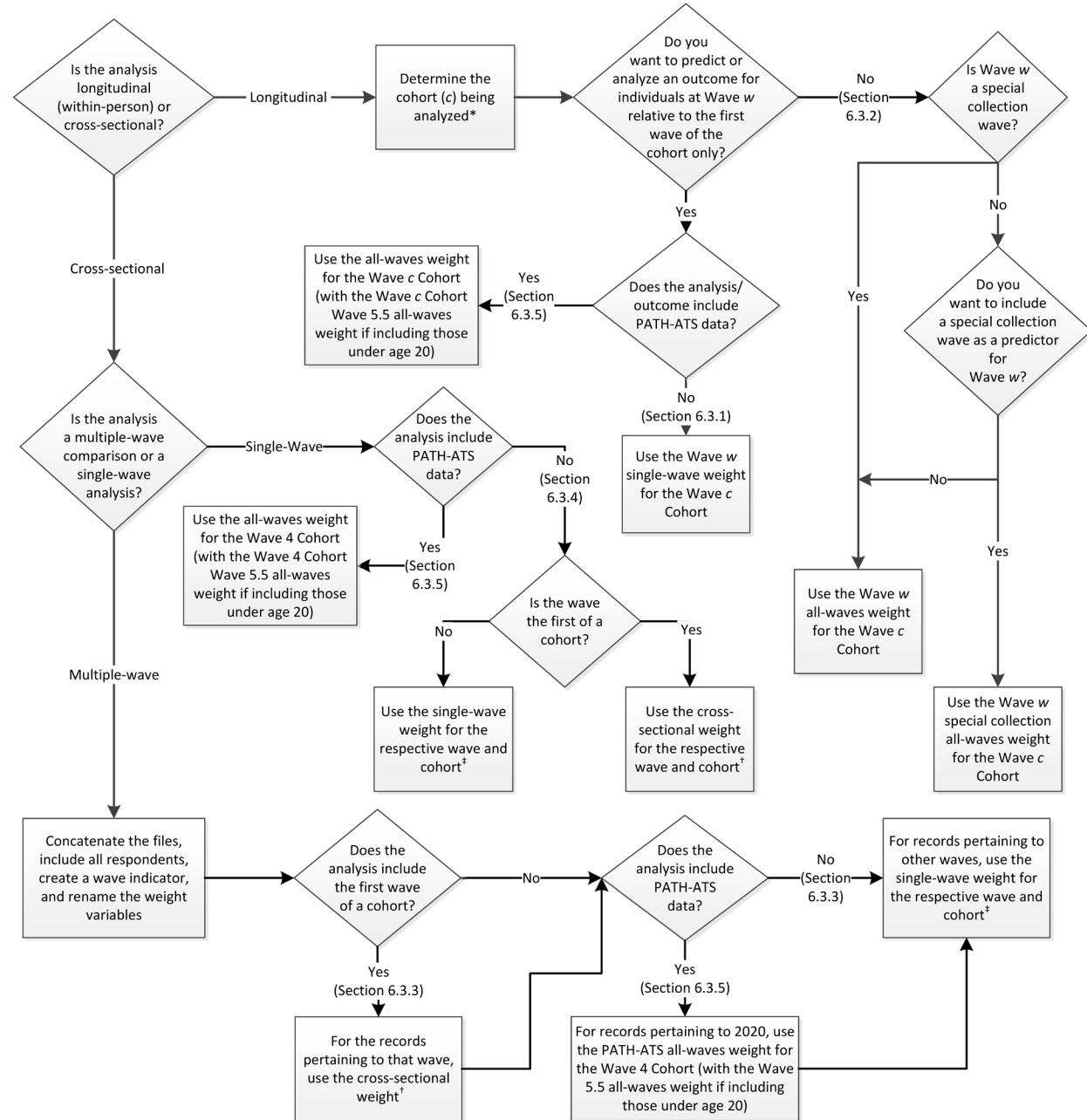
<sup>b</sup> Youth selected in the shadow sample at Wave 1 do not have this weight available with the data files because they did not complete a Wave 1 interview and have no data available for longitudinal analyses involving Wave 1 and any later wave.

<sup>c</sup> The single-wave weight serves as a pseudo cross-sectional weight for the interim waves prior to the formation of a newer cohort (see Section 6.2). If separate single-wave weights are not available (e.g., Wave 2, Wave 5, PATH-ATS) the available set of weights serves as an all-waves and a single-wave weight.

<sup>d</sup> For analyses combining Wave 5.5 and PATH-ATS including those younger than 20 years old, the Wave 5.5 all-waves weight should be used in combination with the PATH-ATS all-waves weight for the cohort of interest. To approximate the cross-sectional population, the Wave 4 Cohort all-waves weights for Wave 5.5 and PATH-ATS should be used (see Section 6.3.5).

The flowchart in Figure 3 summarizes the types of decisions researchers will make in analyzing PATH Study data. This is further discussed, with examples of analyses of each type, in Sections 6.3.1 through 6.3.4. Figure 3 also illustrates analyses using the PATH-ATS data; for more information on analyses using the PATH-ATS data, see Section 6.3.5.

**Figure 3.** Flowchart providing guidance on which weights to use for analyses of PATH Study data



### 6.3.1 Longitudinal Analyses Using Data from the First Wave of a Cohort in Estimating a Later Wave Outcome

This section provides examples where the desired analysis is to relate a behavior or outcome at a later wave (Wave  $w$ ) to characteristics of the same person at the first wave of a cohort (i.e., Wave 1 for the Wave 1 Cohort, Wave 4 for the Wave 4 Cohort, Wave 7 for the Wave 7 Cohort). In this instance, the later wave's single-wave weight is to be used. If the later "wave" includes the PATH-ATS data, the all-waves weight is to be used (see Section 6.3.5).

The defining characteristic for this type of analysis is that the interest is in the trajectory of the same individuals traced from the first wave of the cohort to Wave  $w$ . The goal is to estimate at Wave  $w$  the extent to which individuals have changed since the first wave of the cohort based on characteristics from that wave and Wave  $w$  only. These are within-person analyses for which the PATH Study was designed.

Because no data are collected for shadow youth at their recruitment wave, Wave  $c$ , they may not be included in such an analysis of Wave  $c$  cohort data. For example, Wave 1 Cohort participants selected as shadow youth at Wave 1, cannot be included in an analysis of data using characteristics from Wave 1 and Wave  $w$  because they have no Wave 1 data available for analysis.

Examples of analyses using Wave 1 information to estimate a Wave  $w$  outcome for the Wave 1 Cohort:

- Estimating the proportion of persons using e-cigarettes at Wave  $w$  who used e-cigarettes at Wave 1. This is a longitudinal analysis using the persons who have data at both Wave 1 and Wave  $w$ .
- Calculating an odds ratio such as the ratio comparing (a) the odds of a person using e-cigarettes at Wave  $w$  if the person used e-cigarettes at Wave 1 to (b) the odds of a person using e-cigarettes at Wave  $w$  if the person did not use e-cigarettes at Wave 1. The target population is that at Wave  $w$ . The odds ratio can be calculated from estimated proportions obtained from the two-by-two contingency table below. (Note that the cells of the contingency table are based on the characteristics of persons who responded at both waves.)

E-cigarette use	Did not use at Wave $w$	Used at Wave $w$	Total
Did not use at Wave 1	Estimated number of persons who did not use at both Wave $w$ and Wave 1	Estimated number of persons who used at Wave $w$ although did not use at Wave 1	
Used at Wave 1	Estimated number of persons who did not use at Wave $w$ although users at Wave 1	Estimated number of persons who used at both Wave $w$ and Wave 1	
Total			

- Logistic regression estimating Wave  $w$  e-cigarette use from both Wave 1 and Wave  $w$  characteristics. This might include Wave 1 e-cigarette use status (e.g., experimenting with e-cigarettes), opinion on tobacco use (as identified at Wave 1) of people important to the respondent, education level at Wave  $w$ , and demographic information. This analysis can be thought of as estimating the probability that a person uses e-cigarettes at Wave  $w$  from that person's Wave 1 and Wave  $w$  covariates.

To perform such longitudinal analyses, merge the wave-specific files by PERSONID and then limit the resulting file to records that are common to both files. Use the Wave  $w$  full-sample single-wave weight and replicate weights if replication methods are to be used for variance estimation purposes. If the PATH-ATS data are included, the full-sample all-waves weight and replicate weights should be used.

### 6.3.2 Longitudinal Analyses Using Data from Preceding Waves in Estimating a Later Wave Outcome

Analyses estimating Wave  $w$  outcomes/behaviors/characteristics from data available from preceding waves are also longitudinal analyses. However, they involve persons who participated in all these waves (as opposed to the analyses discussed in Section 6.3.1, which do not use information from the waves between the first wave of a cohort and Wave  $w$ ). For these analyses, use the Wave  $w$  all-waves weight. Also, use this weight if it is desired to estimate Wave  $w$  outcomes from characteristics from an earlier wave other than that from the first wave of a cohort (for example, to estimate Wave 3 outcomes using Wave 2 characteristics). This weight is the appropriate weight to use in the analysis of Wave  $c$  cohort data including shadow youth recruited at Wave  $c$  since they will only have data for waves after the first wave of the cohort. Note that if Wave  $w$  is not a special collection wave, the Wave  $w$  special collection all-waves weight is needed if the analysis includes information from a special collection wave as a predictor for Wave  $w$ .

Examples of such analyses are as follows:

- Logistic regression estimating a person's ENDS use at Wave 4.5 from the person's usage status and demographic characteristics at Waves 1 through 4.
- Mediation analysis estimating a person's ENDS use at Wave 4 from the person's usage and demographic characteristics at Wave 1, mediated by peer opinions at the intermediate waves (Wave 2 and Wave 3).
- Logistic regression estimating cigarette usage at Wave 4 among persons who used cigarettes daily at Wave 1, relating to the history of e-cigarette use (at Wave 1) and ENDS use at Wave 2 and Wave 3.
- Survival analysis estimating time to first trying ENDS at Wave 4.5 for a person who has never tried ENDS before, relating the occurrence to peer pressure and social media exposure at Wave 1 through Wave 4.
- Generalized estimating equation analysis using wave pairs (e.g., Wave 1-Wave 2, Wave 2-Wave 3, Wave 3-Wave 4) to estimate the association between demographic and tobacco-use characteristics at the 'baseline' wave and initiation of ENDS use at the 'follow-up' wave of the wave pairs.

To perform such longitudinal analyses, merge the wave-specific files by PERSONID and then keep the records that are common to all these wave-specific files (i.e., keep the persons who responded at the waves under analysis). Use the Wave *w* full-sample all-waves weight and replicate weights if replication methods are to be used for variance estimation purposes.

### **6.3.3 Cross-Sectional Analyses Comparing Different (or Partially Overlapping) Sets of Persons Between Waves**

Some researchers may want to use the PATH Study data to answer questions that relate to the cross-sectional populations at the different waves. The cross-sectional weights created at recruitment waves upon cohort formation (e.g., the Wave 4 weights created for the Wave 4 Cohort) allow representation of the CNP at those waves. However, as explained in Section 6.2, there are subtle differences between the PATH Study target population for a cohort at subsequent waves and the CNP at those points in time. Therefore, any comparisons of these cross-sectional populations using the PATH Study data are approximations.

The PATH Study was not designed to answer questions that relate to the cross-sectional population at different waves. That said, there are situations where cross-sectional estimation may be called for. The following are some examples where cross-sectional comparisons between waves may be desired.

- Comparing the estimated percentage of 8<sup>th</sup> graders who have tried e-cigarettes at Wave 1 with the estimated percentage of 8<sup>th</sup> graders who have tried e-cigarettes at a later wave. Most of the 8<sup>th</sup> graders at Wave 1 are different persons than the 8<sup>th</sup> graders at a later wave because the targeted interval between waves is at least 1 year, so a longitudinal analysis is not appropriate (longitudinal analyses are based on data collected from the same persons followed over time).
- Comparing the estimated percentage of 18- to 24-year-olds who use e-cigarettes at Wave 1 with the estimated percentage of 18- to 24-year-olds who use e-cigarettes at a later wave, Wave  $w$ . There may be some persons who are ages 18 to 23 at Wave 1 and still in the 18 to 24 age group at Wave  $w$ . However, generally, the persons who are 24 at Wave 1 will not be in the age group at Wave  $w$ , and the persons age 18 at Wave  $w$  will not be in the comparison group for Wave 1. As with the first example, the analysis is not based on data collected from the same set of persons followed over time, so it is not longitudinal in nature.

To create cross-sectional estimates for comparing waves, take the following steps (see Section 6.3.5 for steps needed to incorporate the PATH-ATS data):

- Identify the appropriate wave-specific weight variables for the comparison by consulting Table 6-1. Note that for each type of weight there are separate weight variables for youth and adults, but this distinction is ignored here for simplicity;
- Rename all wave-specific variables, including the weight variables, to obtain a single common name for each set of comparable variables;
- Create a wave indicator variable;
- Concatenate or “stack” data files from each wave to form a single file with one record per respondent per wave in which they provided data.

The subsequent analyses must include the newly created wave indicator variable and the design correctly specified in a software package that can capture sample variability, as described in Appendix C.1. Even though there may not be complete overlap between the two sets of respondents, there are still correlations between the two groups that should be reflected due to potential partial overlap and because some persons may be in the same PSUs. This correlation serves to reduce the estimated variance of the comparison. Manipulating the variables, records, and files as described above and using the appropriate variance estimation methods will correctly reflect these correlations.

Researchers are advised not to compare groups at two different waves (or indeed, two subgroups in the same wave) by computing separate confidence intervals for the two estimates and determining if they overlap. Such an analysis would be incorrect even for simple random samples<sup>27</sup> and would result in misleading inferences when applied to a complex design such as used for the PATH Study. This is because such an approach would not account for the correlation between the two groups caused by the complex sample design.

#### **6.3.4 Cross-Sectional Analyses Describing Findings for a Single Wave Without Comparison to Other Waves**

Before the release of the Wave 2 PATH Study data, the only analyses possible were cross-sectional analyses of the Wave 1 data. With the release of later waves of data, most of the analyses conducted are anticipated to be longitudinal, developing estimates related to a Wave  $n$  outcome from characteristics from previous waves (as described in Sections 6.3.1 and 6.3.2). Cross-sectional analyses comparing different (or partially overlapping) sets of persons between waves may also be undertaken (as described in Section 6.3.3). However, there may be some situations in which analysis of later-wave data alone is of interest, for example, to analyze new content introduced in that wave.

When performing cross-sectional analyses describing findings for a single wave without comparison to other waves, use the data files and weights identified for cross-sectional use in Table 6-1. For such analyses, the populations represented at Waves 1, 4, and 7 will be the CNP at the time of those waves. The populations represented at waves other than Waves 1, 4, and 7 would be, strictly speaking, slightly different than the CNP at those waves, as discussed earlier (e.g., at Wave 2, the estimates represent those in the CNP at Wave 1 who were residents of the United States and not incarcerated at Wave 2). However, the estimates would be suitable to serve as reasonable approximations of the CNP at those times and could be described as such.

#### **6.3.5 Analyses Using PATH-ATS Data**

As described in Section 2.7, The PATH-ATS was selected from Wave 4 Cohort participants who responded to the Wave 5 adult interview and were ages 20 and older on August 31, 2020. This

---

<sup>27</sup>See Schenker and Gentleman (2001).

means that all PATH-ATS respondents are by definition Wave 4 Cohort “all-wave respondents,” and so the weight assigned to all PATH-ATS respondents is a Wave 4 Cohort all-waves weight. This weight may be used for analyses of PATH-ATS data with data from Wave 4 and/or Wave 5. A Wave 1 Cohort longitudinal all-waves weight is also available for analysis of PATH-ATS data, given the amount of overlap between the two cohorts; this weight is assigned to PATH-ATS respondents who also responded at Waves 1, 2, 3, 4, and 5. This weight may be used for analyses of PATH-ATS data with data from Waves 1, 2, 3, 4, and 5, or any combination of these waves.

The all-waves weight is the appropriate weight to use for longitudinal analyses (see Sections 6.3.1 and 6.3.2). An example of a PATH-ATS longitudinal analysis is estimating the percentage of the adult population ages 20 and older that uses ENDS in the latter portion of 2020 from usage and demographic characteristics at Wave 3; the appropriate weight for this analysis is the Wave 1 Cohort all-waves weight.

Because the Wave 4 Cohort all-waves weights is the only PATH-ATS weight available for the cohort, this weight may also be used for cross-sectional estimation. An example of such a cross-sectional comparison between waves is comparing the estimated percentage of 20- to 24-year-olds who use e-cigarettes at Wave 1 with the estimated percentage of 20- to 24-year-olds who use e-cigarettes in the latter portion of 2020 (see Section 6.3.3). Another cross-sectional application is analyzing new content introduced in the latter portion of 2020 (see Section 6.3.4) related to COVID-19 (e.g., social distancing practices, impacts on tobacco use and stress, perceptions of COVID-19 pandemic, receiving a COVID-19 diagnosis, experiencing COVID-19-related symptoms). Some researchers may want to analyze the PATH Study data for the full adult population (ages 18-19 at Wave 5.5 and ages 20 and older at PATH-ATS interview) or some subset of the adult population (e.g., ages 18-24) in the latter portion of 2020. Because the respondents to the Wave 5.5 telephone interview and the respondents to the PATH-ATS are mutually exclusive, and the two samples were combined for the last step in the all-waves weighting process (see Section B.8), the two samples may be pooled for analyses.

To create a combined data file for analyzing adult interview data from Wave 5.5 and PATH-ATS representing the latter portion of 2020, take the following steps:

- Identify the specific cohort and the appropriate all-waves weight variables for Wave 5.5 and PATH-ATS by consulting Table 6-1. Rename the weight variables to obtain a single

common name for variables in each file, and concatenate or “stack” the Wave 5.5 and PATH-ATS files to form a single weight file;

- Identify or create comparable analysis variables using the Wave 5.5 adult and PATH-ATS interview data files, renaming them as necessary so that each analysis variable has a single name across the two files.
- Concatenate or “stack” the Wave 5.5 adult and PATH-ATS files to form a single analysis file.

Merge the analysis and weight files by PERSONID to form a single file with all adult Wave 5.5 and PATH-ATS respondents.

## 6.4 Impact of Data Collection Methods

Due to the COVID-19 pandemic, data collection methods used in 2020 differed from previous waves in important ways. Some of these are noted below; however, analysts are encouraged to review the Wave 5.5 and PATH-ATS sample designs presented in Sections 2.6 and 2.7 and the data collection methods which are detailed in Sections 3.7 and 3.8 and summarized in Table 3-1 and Appendix A.

- The data in the Wave 5.5 and the PATH-ATS released data files were collected solely via telephone, unlike the previous waves where households were visited in-person and interviews were completed solely by ACASI. The major difference between telephone interviewing and ACASI is the involvement of interviewers in the data collection. Telephone interviewers read questions to the participants and listen to and record their answers to the questions; with ACASI, interviewers are present to assist if requested, but do not know answers the participants provide. This difference in interviewer involvement in data collection may have a greater impact on answers to sensitive questions (Tourangeau and Yan, 2007; Yan and Cantor, 2019; Yan, 2021).
- In order to avoid the potential for bias resulting from pooling the Wave 5.5 data collected prior to the suspension of in-person data collection with the data collected in the latter portion of 2020, all Wave 5.5 in-person interview respondents who were still age-eligible were re-contacted for an interview by telephone.
- The data collection periods for Wave 5.5 and the PATH-ATS were shorter than previous waves and cover the latter portion of 2020, after the onset of the COVID-19 pandemic. (See Table 1-2 for the specific data collection start and end dates.) As such, the data collected are not representative of the full year.

Because of the differences in data collection mode, period of data collection, and the impact of the COVID-19 pandemic, data users should exercise caution when comparing estimates from 2020 to estimates from other years.

As described in Section 3.6.2, both in-person data collection and telephone data collection were conducted for Wave 6. Wave 6 was initially conducted only by telephone until May 7, 2021. After that date, Westat assessed the risk levels for each PSU using COVID-19 numbers released by CDC and in-person visits were resumed in PSUs where it was determined safe to conduct in-person interviews. Starting in August 2021, in-person interviews administered by fully vaccinated interviewers were extended to all PSUs. Participants visited by field interviewers during Wave 6 completed the PATH Study questionnaires by ACASI. Participants who did not want an in-person visit were offered the opportunity to complete the questionnaire by telephone. The option to complete the interview via telephone for participants with health concerns continued in Wave 7. Also in Wave 7, a limited number of adults and youth were asked to complete their interview via the Web as part of a pilot study (see Section 3.10.4).

Because of the limited interview involvement with ACASI (noted above), there are fewer differences between ACASI and Web self-interviewing than between these modes and telephone interviewing. Regardless, the mode of data collection may affect the demographic composition of people who responded to a particular mode, the interview setting, as well as the answers provided by respondents. When using data collected via multiple modes for cross-sectional or longitudinal purposes, data users are advised to first examine differences in demographic composition and key variables of analytic interest (e.g., tobacco use) by mode of data collection, to conduct sensitivity analyses, if needed, and to consider how these might influence interpretation of findings.

## **6.5 Age Collection, Calculation, and Inconsistencies Across Waves**

The sections below describe how age information is collected in the PATH Study (Section 6.5.1), how the age variables are calculated (Section 6.5.2), and how inconsistencies in ages across waves can result (Section 6.5.3). Analytic considerations in light of such age inconsistencies are provided in Section 6.5.4.

### **6.5.1 Collection of Age Information**

All adults sampled in recruitment waves were asked to provide their date of birth as part of the consent process. If the adult refused or reported that they did not know their date of birth, they were asked for an exact age (at the time of the consent). If the adult also refused or reported that they did not know their exact age, they were asked to choose a response from among specified age ranges. If the adult refused to provide or did not know any of this information, the consent process was ended and the adult was not included in the PATH Study.

Parents of youth sampled in recruitment waves were asked to provide the date of birth for their child. If a parent refused or reported that they did not know the sampled youth's date of birth, the youth was not included in the PATH Study. The reason for the more stringent requirement for youth respondents was to ensure that the appropriate consent and assent procedures were followed for study participants under age 18.

Because the PATH Study includes youth under age 18 and asks about illegal and/or sensitive behaviors, ensuring that the study correctly identifies the age of each respondent before a follow-up interview is of high priority. This requires confirming or asking again about age at follow-up waves for all respondents, as described below.

The process for confirming or collecting age at follow-up waves for adults depended on respondent type:

- If continuing adult respondents provided their date of birth at a previous wave, they were asked to verify the date of birth.
- If continuing adult respondents did not provide their date of birth at a previous wave, they were asked for their date of birth.
- Aged-up adults were asked to verify the date of birth previously provided by their parent.

In any of the above scenarios, if an adult respondent indicated that their date of birth provided at a previous wave was incorrect, they were asked for the correct date of birth, with follow-up questions similar to those used at the recruitment wave if the respondent refused or reported that they did not know the information. If an adult respondent did not provide updated age information, the interview for the respective wave proceeded.

For youth respondents, the age information was confirmed in follow-up waves by asking the parent to verify the date of birth that was provided for their child at a previous wave. If the parent reported that the information was incorrect, they were asked to provide the correct date of birth. If a parent did not provide updated information, the youth interview was coded as nonresponse for that wave.

## 6.5.2 Calculating Interview Ages

For each wave, the calculated interview is used to create an age range variable R0nR\_x\_AGECAT# (or X0nR\_x\_AGECAT# for special collection waves that occur between primary waves, for example, Wave 4.5; or T05R\_A\_AGECAT# for PATH-ATS), where *n* indicates the wave number (or prior wave number for special collection waves) and *x* indicates “A” for adults or “Y” for youth, and # indicates the number of categories which is determined by the disclosure risk analysis (for example, # = 7 for Wave 1 to Wave 3 adults, # = 6 for Wave 4 to Wave 6 adults, including Wave 5.5 and PATH-ATS, and # = 2 for youth in all waves).

For recruitment waves, ages for all respondents who provided dates of birth were calculated by subtracting the date of birth from the interview date for the recruitment wave. If an exact age was provided, that value was used as the age for the recruitment wave. If only an age range was provided, the calculated age for the recruitment wave was missing. An age range variable was then created using the calculated age.

For follow-up waves, age at each wave was calculated as follows:

- For respondents who verified the date of birth provided at a previous wave, their age as of the date of the respective follow-up interview was calculated by subtracting the confirmed date of birth from the date of the follow-up interview.
- If a respondent revised their date of birth at a follow-up wave, their age as of the date of the respective follow-up interview was calculated using the revised date of birth.
- If a respondent provided an exact age at a follow-up wave, that value was used.
- If a respondent only provided an age range at a follow-up wave, the exact interview age is missing. However, respondents were routed appropriately through the instrument for the given age group at each respective wave.

### 6.5.3 Potential for Age Inconsistencies Across Waves

Logically, a respondent with age  $z$  at Wave  $w$  is expected to have an age of  $z$ ,  $z+1$ , or  $z+2$  in an annual follow-up wave, depending on the dates of the two interviews in relation to the respondent's date of birth. Although every effort is made to complete follow-up interviews at approximately 1-year intervals in annual follow-up waves, participants have the opportunity to respond until the end of each wave. Therefore, a respondent could have two birthdays between interviews, one occurring soon after their Wave  $w$  interview and another soon before their annual follow-up interview. As such, a difference of 2 years between interview ages in two consecutive annual waves is considered acceptable.<sup>28</sup>

In a biennial follow-up wave (when it occurs), a respondent with age  $z$  at Wave  $w$  is expected to have an age of  $z+1$ ,  $z+2$ , or  $z+3$ , depending on the dates of the two interviews in relation to the respondent's date of birth. Although every effort is made to complete follow-up interviews at approximately 2-year intervals in biennial follow-up waves, participants have the opportunity to respond until the end of each wave. Therefore, a respondent could have three birthdays between interviews, one occurring soon after their Wave  $w$  interview and another soon before their biennial follow-up interview. As such, a difference of 3 years between interview ages in two consecutive biennial follow-up waves is considered acceptable.

However, if in an annual or a biennial follow-up wave a respondent provided a revised date of birth or age that constitutes a logical inconsistency from previously recorded age(s), the longitudinal age progression at any follow-up wave may fall outside of the expected range.

A change to information that results in inconsistent interview ages between waves is problematic for analysis of longitudinal data. Because respondents are expected to increase in age by no more than 2 or 3 years between consecutive interviews depending on the timing of their birth dates and their

---

<sup>28</sup>Some study members respond to the PATH Study in non-consecutive waves. In such instances, a 2-year difference for each wave is included in the total acceptable interval between the annual waves completed by respondents. For example, if a participant responds in Wave 1, does not respond in Wave 2 and responds in Wave 3, a difference of 4 years (between Wave 3 and Wave 1) is considered acceptable.

interview dates, the problematic cases are those where a respondent revised their information such that:

- a) at their annual follow-up interview they were younger or more than 2 years older than they appeared to be at Wave  $w$ ; or
- b) at their biennial follow-up interview they were younger or more than 3 years older than they appeared to be at Wave  $w$ .

So that respondents with inconsistent ages across waves are readily apparent, the dichotomous variable R0 $wR_x$ \_AGE\_CHECK is included on the adult and youth data files after Wave 1 (where  $w$  indicates any wave after Wave 1 and  $x$  indicates “A” or “Y” as appropriate) with the following categories:

- 1 = Age is within expected range
- 2 = Age is not within expected range.

A missing value (-99933) is assigned to the Wave 4 version of these variables for respondents added to the sample through that wave’s replenishment effort. The Wave 7 version of these variables are set to system missing for respondents added to the sample through that wave’s replenishment effort. Respondents added to the sample through replenishment efforts are assigned a non-missing value (1 or 2) for these variables in respective waves that follow. In special collection waves that occur between primary waves, the dichotomous variable X0 $wR_x$ \_AGE\_CHECK is included on the respective file(s), following the variable nomenclature described in Section 7.4. For PATH-ATS, the dichotomous variable is named T05R\_A\_AGE\_CHECK.

Some respondents with inconsistent ages across waves will not appear to have inconsistent age ranges; because the ages are within the same age range or result in a change from one range to the next higher range. The age-check variable is included on the files, even though single-year of age is not provided, to identify all suspect cases that analysts may want to exclude or categorize separately as discussed in the next section.

#### **6.5.4 Impact of Age Inconsistencies Across Waves**

As noted above, logically a respondent with age  $z$  at Wave  $w$  may still be age  $z$  in an annual follow-up wave, or age  $z+1$  in a biennial follow-up wave, depending on the dates of the two interviews in relation to the respondent’s date of birth. In most analyses, cases with this age progression may not

be apparent. However, in analyses including the youngest of a cohort, these cases may appear unusual. For example, although we would expect all participants in the Wave 1 Cohort to be at least 16 years old by Wave 6, there are some 15-year-olds in this cohort who completed the interview. Table 1-1 indicates the expected ages of interviewed PATH Study participants by cohort and wave.

Although the number of cases appearing to be too young for their cohort or with inconsistent ages across waves is small, users may want to consider the influence these cases have on their analyses:

- For analyses that predict an outcome at a follow-up wave based on data from an earlier wave, the age at the earlier wave is probably the more appropriate choice (for example, when predicting first use of e-cigarettes among Wave 1 participants who did not use e-cigarettes, the Wave 1 age would be employed).
- For analyses that include only a single wave of data (for example, an analysis of Wave 6 ENDS use), the age at the wave being analyzed is probably the more appropriate choice. For cases appearing to be too young for their cohort, analysts may decide to impute the respondents' age to be the expected youngest age for the cohort indicated in Table 1-1. Care is needed in this instance to review any age-related skip patterns in the respective wave's questionnaire to determine the appropriateness of this approach for the specific analysis.

For these and other types of analysis, it is suggested that a sensitivity analysis be conducted to determine if including respondents who have an age that is not within the expected range makes a difference in the results.

## 6.6 Considerations for Pooling Data

Users interested in pooling PATH Study data with data from other sources should first consider the advantages and disadvantages of any such analysis. Pooling data may increase statistical power, particularly for subgroups; however, independent studies can differ in their measurement of key constructs, timing of the study, and target population. Measures of certain constructs may even differ across assessment periods within the same study. Analysts should use weights and/or design variables appropriately to ensure that the complex design features of the PATH Study (and potentially other data sources) are reflected in estimates based on pooled data.

## 6.7 Considerations for Small Domains

As noted in Section 5.2, variance estimates for small domains may be unstable because some PSUs may contain no observations belonging to the domain. Consider the example of an analysis of adults in a limited age range with and without a particular health condition. A reasonable approach is to create estimates separately for adults with and without the condition. If the health condition is rare, it is likely that the estimates for the group with the health condition will exclude some PSUs. This would not be the case for estimates made separately for more common characteristics, such as for males and females; in that case, both estimates will include respondents from all PSUs, because all PSUs include males and females.

The smaller the number of PSUs included in an estimate, the greater the instability of the variance estimate. For complex samples, such as the PATH Study sample, the number of “degrees of freedom” ( $df$ ) reflects the stability of variance estimates and is generally associated with the number of PSUs available for variance estimation (Valliant and Rust, 2010). However, software packages designed for analysis of complex sample data using replication generally assume by default that all PSUs are included in an estimate when creating confidence intervals and in significance testing.

The customary  $df$  attributed to a test statistic is the total number of PSUs minus the total number of strata (Korn and Graubard, 2011). When using the full set of PSUs in the analysis of PATH Study data, this calculation results in 99  $df$ .<sup>29</sup> For estimates based on small domains or specific to a census region, however, the  $df$  available will be smaller. Researchers should consider calculating approximate  $df$  for each estimate that is based on a small domain. The resulting  $df$  may then be appropriately specified in the software package. Specifically, researchers should consider the following:

- Manually calculate the  $df$  separately for each subgroup as the number of PSUs in that subgroup minus the number of strata in that subgroup using the pseudo-strata and pseudo-PSU variables included on the respective weight file (VARSTRAT and

---

<sup>29</sup>Note that when using replication, the number of degrees of freedom assumed by most software packages is the number of replicate weights, which for the PATH Study is 100.

VARPSU, respectively).<sup>30</sup> Appendix C.2 provides example program code for this calculation.

- Analyze subgroups separately (i.e., not using a “by” statement in SAS or equivalent statement in other software packages) and specify the calculated *df* for each subgroup in the software package. Appendix C.2 provides example code for the specification of the *df* in analysis using various software packages.
- When testing to determine if there is a significant difference between groups of participants with different characteristics, use the smallest *df* calculated across the groups in the comparison.

As noted in Section 6.3, analysts using Taylor Series methods to create estimates for subdomains of interest may retain all records in the analysis data file for their analysis; in such an instance the software package appropriately computes the *df* and variance estimates. However, the recommended approach for computing sampling errors for the PATH Study is to employ the replicate weights provided for each weight type.

Decreasing the *df* used in analyses will increase the width of confidence intervals around the corresponding estimate, thus reflecting the instability of the variance to the extent possible. The increased widths will provide some protection against misleading findings. Estimates based on fewer than 8 *df* have a relative standard error of the variance approximately 50% or higher. For this reason, researchers should proceed with caution if any of the domains result in fewer than 8 *df*.<sup>31</sup>

---

<sup>30</sup>Full sample weight, replicate weight, pseudo-strata, and pseudo-PSU variables are included on the files with the questionnaire data for Wave 1 and Wave 2. For all other waves, these variables are on separate files corresponding to the respective wave, cohort, weight type, and interview type.

<sup>31</sup>This is calculated based on the approximation given in Valliant and Rust (2010)  $RSE(var(\hat{\theta})) = 100\sqrt{2/df}$ .

## 7. Data Files

---

### 7.1 Data Structure

As discussed in Section 1.2, the adult data file and youth/parent data file for a given wave have similar content. They both include the questionnaire items from the interviews and some derived variables such as those pertaining to tobacco-use definitions. The Wave 1, Wave 4 and Wave 7 data files include census region

Starting with Wave 3, an auxiliary data file was developed to simplify the derivation of certain variables in the adult and youth data files at follow-up waves. This new data file is called the ever/never reference data file.

**Adult:** The adult data file contains one record for every adult interview completed in the given wave or special collection.

**Youth/Parent:** The youth/parent data file contains one record for every youth interview completed in the given wave or special collection and contains responses to the brief parent interview about the youth who completed the youth interview. Not every youth interview has a corresponding parent interview because a parent/guardian can refuse to complete the parent interview but give permission for their youth to complete the youth interview. A youth may also have a completed parent interview in one wave but not the next. In addition, different parents may complete the interviews across waves. A single parent could have completed the parent interview for multiple youths within a household; in this instance, there are certain questions in the parent interview that are asked only once. The data for questions that were not asked in additional interviews with the same parent have been copied from the record of the first interview completed by that parent. For reference, questions that are not repeated in these instances in the parent interview are identified in the parent annotated instrument (combined in the same file as the youth annotated instrument). For example, in Wave 1, these are identified in Boxes R01\_PTR01 and R01\_PNR01.

**Ever/Never Reference:** The ever/never reference data file is an auxiliary file that provides respondents' ever/never statuses for tobacco use as of each wave. The objective of this data file is to consolidate information across waves and to ascertain whether respondents ever reported using any tobacco products. Consolidating this information into a single data file helps simplify the algorithms

to derive tobacco use variables in subsequent waves. Without the consolidation of this information, users would need to merge data across multiple waves to determine if respondents reported ever using any tobacco products.

The Wave 3 ever/never reference data file contains one record for each Wave 1 adult, youth, and shadow youth respondent. It includes a set of marker derived variables (MDVs), as well as flag variables for shadow youth and respondents who skipped one or more waves of the study. Each MDV indicates respondents' ever/never status of tobacco product use, as of the most recent wave in which they responded. The MDV variables in the Wave 3 ever/never reference data file are used to derive ever/never status variables in the Wave 4 adult and youth data files. Ever/never reference files are created for all waves after Wave 3; ever/never status variables in each wave after Wave 4 are derived using the MDV variables in the prior waves' ever/never reference data file.

The intent of the MDVs in the ever/never reference data file is to simplify derivation of tobacco use variables. As such, neither the ever/never reference data file nor the MDVs it contains are recommended for any analyses.

## **7.2 Record Identifier**

The variable named PERSONID serves as the unique identifier for each record in the PATH Study data files. Each PERSONID represents a participant in the PATH Study: an adult, a youth, or a shadow youth. The PERSONID will remain constant throughout the PATH Study even as a shadow youth ages to a youth or a youth respondent ages into adulthood. Each PERSONID begins with a "P" and is followed by a randomly assigned nine-digit number that does not include any direct or indirect references to personally identifiable information or geographic location.

## **7.3 Excluded Variables**

Some specific questionnaire items are not provided in the data files. This may be for one of two reasons:

- Some items are excluded to reduce the risk of disclosing a respondent's identity. All direct identifiers such as names, addresses, and phone numbers are excluded from the files, but other specific questionnaire items are also excluded for this reason.

- Some items were collected or created in the instrument for operational purposes only. These include items derived throughout the instrument for routing purposes, including items that indicate types and levels of tobacco usage. (See the appendixes of the annotated instruments for the full list of tobacco-use variables used for routing purposes.) Variables indicating types and levels of tobacco use that are better suited for analysis were derived and provided with the data files.

Appendix E contains the full list of questionnaire items excluded from the data files.

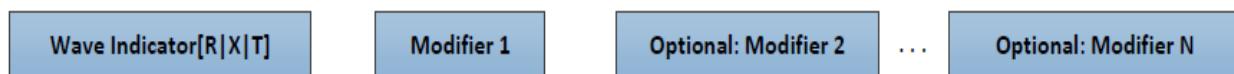
## 7.4 Variable Names

The PATH Study variable naming convention was designed to provide context for each variable based on the content being represented by the item and thus to facilitate the identification and classification of items for analyses. The variable naming convention also includes an indicator that specifies the longitudinal data collection wave.

The naming convention applies to all variables on the data files, with the exception of record identifiers and the variables that reflect the variance structure. It is anticipated that these items will retain the same variable name throughout the course of the PATH Study.

Variables on the PATH Study data files are named using multiple components, each of which is separated by an underscore “\_” character for readability. The naming convention is based on a syntax methodology that begins with broader modifiers, moving to more specific modifiers. The basic structure of the naming convention is given in Figure 4. This structure creates variables that are platform independent (no unusual characters) and easy to read and allows variables to be visually identifiable as related. The general structure applied across all PATH Study file types aims to prevent the situation where a variable of the same name has different contents in different files. This helps users to avoid unintentionally overwriting data when merging records from two or more files (e.g., when merging adult and youth/parent interview data files). Users may stack (append) datasets for analysis including both adults and youth, after renaming the variables of interest including the weight and replicate weight variables.

**Figure 4. PATH Study variable name components**



The components of the variable names are:

- **Wave indicator:** Each variable name begins with a wave indicator, represented by three characters. In Waves 1 through 4 and the primary waves that follow, the first character is “R”, which is followed by a two-digit zero-filled number indicating the data collection wave, beginning with “01” for Wave 1. As such, all variables in the Wave 1 data files other than PERSONID, VARSTRAT and VARPSU have the prefix “R01”. Variables in the Wave 2 data files have the prefix “R02,” and so on for each wave in which data collection is attempted for all PATH Study participants. In special collection waves between the primary waves (i.e., the “point 5” waves), the first character is “X,” followed by a two-digit zero-filled number indicating the previous data collection wave. The first “point 5” wave in the PATH Study occurred between primary Waves 4 and 5. Because the “point 5” waves are considered follow-up waves to the prior waves (e.g., Wave 4.5 was a follow-up to Wave 4, with a subset of all PATH Study participants), the number indicating the wave before the “.5” is used as the wave indicator. As such, variables in the Wave 4.5 data files have the prefix “X04.” Similarly, variables in the Wave 5.5 data files have the prefix “X05.” For PATH-ATS, the first character is “T,” followed by a two-digit zero filled number indicating the data collection wave. PATH-ATS occurred between primary Waves 5 and 6. Therefore, following the convention used for the “point 5” waves, the variables for PATH-ATS have the prefix “T05.”

All derived and imputed variables include the additional character “R” (indicating that the variable is a recode) after the wave indicator. As such, all derived and imputed variables on the Wave 1 data files have the prefix “R01R”. All derived and imputed variables in the Wave 2 data files have the prefix “R02R,” and so on for each wave in which data collection is attempted for all PATH Study participants. Similarly, all variables created from an external source include the additional character “X” so that all such variables on the Wave 1 data files have the prefix “R01X”. In special collection waves between the primary waves, derived and imputed variables also have the additional character “R” after the wave indicator. So, all derived and imputed variables in the Wave 4.5 data files have the prefix “X04R.” Similarly, the prefix in the Wave 5.5 data files for derived and imputed variables is “X05R.” All derived and imputed variables in the PATH-ATS data file have the prefix “T05R.”

All marker derived variables in the ever/never reference files include the additional character “M” after the wave indicator. As such, all marker derived variables in the Wave 3 ever/never reference file have the prefix “R03M,” and so on for each wave in which data collection is attempted for all PATH Study participants. In special collection waves between primary waves, all marker derived variables in the ever/never reference files also have the additional character “M” after the wave indicator. So, all marker derived variables in the Wave 4.5 ever/never reference file have the prefix “X04M” and all marker derived variables in the Wave 5.5 ever/never reference file have the prefix “X05M.”

- **Modifiers**
  - **PATH ID:** A PATH ID represents a unique question in an instrument. All interview variables are named using the PATH IDs that were assigned to questions in the instrument specifications. PATH IDs are represented in the format: @\$\$\$\$\$, where @ represents the instrument type (“A” for Adult, “P” for Parent, and “Y” for Youth), \$ represents the item content (e.g., “C” for cigarettes), and ##### is a numeric key that is unique within each instrument. The numbers in the PATH IDs do not correspond to any specific ordering of questions; they were simply assigned during the instrument design process based on availability. Some variable names include suffixes to the PATH ID to differentiate between tobacco product types when the same question was repeated for different tobacco products. For example, variable R01\_AG1005TC corresponds to the question with PATH ID AG1005 when it was asked about traditional cigars in the Wave 1 adult interview. In some instances, there are questions with similar content that are asked in different sections or in different instruments. In such cases, the number identifier in the PATH ID is the same across both sections/instruments whereas the instrument/content indicator is different. For example, in Wave 1, the question with PATH ID AC1005 (lifetime use of cigarettes) in the cigarette section is also asked about e-cigarettes in the question with PATH ID AE1005 (lifetime use of e-cigarettes).
  - **Interview identifier:** Where appropriate, administrative, management, processing, derived, and imputed variables include an indicator of the interview to which the information pertains: “A” (adult interview), “Y” (youth interview), or “P” (parent interview). For example, the full-sample weight for analyzing data from the Wave 1 adult interview is in the variable named R01\_A\_PWT; the corresponding weight for analyzing data from the Wave 1 youth or parent interview is in the variable named R01\_Y\_PWT.
  - **Mnemonic:** All administrative, management, processing, derived, external, weight, and imputed variables are named using mnemonics. Mnemonics can be full words, phrases, or abbreviations to represent the content of the variable, and can range from the very simple to the complex. For example, the variable indicating the youth’s body mass index at Wave 1 is R01R\_Y\_BMI. In other cases, the mnemonic is actually the PATH ID. For example, the adult variable indicating the age range when first alcoholic drink was consumed is recoded from the questions AX0074 (age in years) and AX0270 (age range, if age in years not given) because all age in year variables were excluded from the PUFs; the recoded variable is named R01R\_A\_AX0074. Variables created from combining parent and emancipated youth responses to the same question are given a hybrid mnemonic. For example, the variable indicating the youth’s overall health status is recoded from PT0035 (parent response) and YT0035 (emancipated youth response); the recoded variable is named R01R\_Y\_PY0035.

The weight variables have components in the mnemonic that indicate the type of weight and cohort identification (starting with Wave 4). For example, the all-

waves weight variable for Wave 4 adults in the Wave 1 Cohort is R04\_A\_A01WGT. Table 7-1 includes a list of mnemonic components in weight variables.

The tobacco-use variables have components in the mnemonic that indicate the type of use, intensity, timeframe, and product. For example, the Wave 1 variable indicating whether an adult has used cigarettes (CIGS) but now does not (former or FMR) and in the past used cigarettes regularly or in an established manner (ESTD) and has used cigarettes within the past 12 months (P12M) is R01R\_A\_FMR\_ESTD\_P12M\_CIGS. Table 7-2 includes a list of the tobacco-use mnemonic components.

**Table 7-1. Weight variable name mnemonic components**

Attribute	Abbreviation	Description
Interview identifier	A Y	Adult Youth
Weight type (Waves 1, 2, 3)	P A S	Person-level All-waves Single-wave
Cohort and weight type (Starting with Wave 4)	A01 S01 C04 S04 AX01 AX04 C07	Wave 1 Cohort, all-waves Wave 1 Cohort, single-wave Wave 4 Cohort, cross-sectional Wave 4 Cohort, single-wave Wave 1 Cohort, special collection all-waves Wave 4 Cohort, special collection all-waves Wave 7 Cohort, cross-sectional
Content	WGT	Weight values

**Table 7-2. Tobacco-use variable name mnemonic components**

Type/Attribute	Abbreviation <sup>a</sup>	Description
Status	CUR	Current
	EDY	Every day
	SDY	Some days
	EDSD	Every day or some days
	NVR	Never
	FMR	Former
	NEW	New
Intensity	ESTD	Established use
	EXPR	Experimental use
	THRSHLD, THRSH <sup>b</sup>	Lifetime threshold of use
Timeframe	12MA	12 months ago (1 year ago)
	P12M	In past 12 months
	P30D	In past 30 days
Product	BIDI	Bidis
	CIGS	Cigarettes
	CIGSMFG	Cigarettes, manufactured
	CIGSRYO	Cigarettes, roll your own
	DISSBL	Dissolvables
	ECIG	E-cigarettes
	EPRODS	E-products
	ECIGAR	E-cigars
	EPIPE	E-pipe
	EHOOK	E-hookah
	GFILTR	Filtered cigars
	GRILLO	Cigarillos
	GTRAD	Traditional cigars
	CIGAR	All cigars
	HOOK	Hookah
	IQOS	IQOS (heated tobacco product)
	KRETEK	Kreteks
	NRT	Nicotine replacement therapy
	NP	Nicotine pouches
	OT	Oral tobacco
	PIPE	Pipe
	RX	Prescription therapy
	SMKLS	Smokeless
	SNUS	Snus pouches
	TOB	All tobacco products

<sup>a</sup> The maximum length for a variable name in the PATH Study data files is 32 characters. Current and recent versions of most statistical software packages are able to accommodate variable names with 32 (or more) characters.

<sup>b</sup> The abbreviation THRSHLD was only used in tobacco-use variable names in the Wave 3 adult data file. The abbreviation THRSH is used in tobacco-use variable names in all other data files.

- **Source:** As additional sources of external data are added, a modifier for the source will be included. For example, consider item R01X\_CB\_REGION, which represents census region in which the respondent was sampled; CB identifies the information source as the Census Bureau.

- **Sub-Item identifier:** Some interview questions have multiple components for a single PATH ID, such as choose-all-that-apply items or dates where responses are recorded in multiple variables. In such cases, the variable name includes a sub-item identifier to differentiate components. For example, consider item R01\_AG9009, about cigarillos and filtered cigars. The question and response options are shown in Figure 5 for reference.

**Figure 5. Example of question and response options**

Which of the following kinds of cigarillos or filtered cigars have you smoked[ as blunts]? Choose all that apply. The kind...

<b>R01_AG9009_01</b>	1 With a plastic or wooden tip
<b>R01_AG9009_02</b>	2 With a filter (like a cigarette filter)
<b>R01_AG9009_03</b>	3 Without a tip or filter
	-8 DON'T KNOW
	-7 REFUSED

The respondent could choose multiple responses for this item. Because there are three sub-items for R01\_AG9009, there are three variables (R01\_AG9009\_01 through R01\_AG9009\_03), each of which corresponds to one of these sub-items. The variables for the sub-items correspond to the order of the responses presented in the question. Therefore, if a respondent chose “With a plastic or wooden tip” and “Without a tip or filter” as answers to R01\_AG9009, only the variables R01\_AG9009\_01 and R01\_AG9009\_03 contain the value indicating the respondent chose that response option (typically this value is 1).

## 7.5 Variable Labels

All variables on the data files include a label that briefly explains the content of the item. For clarity, consistency, and usability, labels are assigned to each variable using a standard convention. The components of the variable labeling convention are shown in Figure 6.

**Figure 6. PATH Study variable labeling convention**

Variable Name	:	[DERIVED   IMPUTED   IMPUTATION FLAG   MARKER - ]Brief Description
---------------	---	--

For all interview, administrative, and management variables, the labels begin with the variable name followed by a colon and a brief description of the content. For derived, imputed, or imputation flag variables, the labels begin with the variable name followed by a colon, the respective phrase “DERIVED”, “IMPUTED”, or “IMPUTATION FLAG”, a spaced hyphen, and then a brief description of the content. For marker derived variables in the ever/never reference data files, the

labels begin with the variable name followed by a colon, the key word “MARKER,” a spaced hyphen, and then a brief description of the content.

## **7.6 Value Labels**

The labels associated with variable values are provided. All value labels include both the data value as well as the description of the value. For example, a variable for an item with yes/no responses has “1 = Yes” and “2 = No” as value labels.

### **7.6.1 Interview Variables**

The value labels for all interview variables match the text of the response option as seen by the respondent.

### **7.6.2 Recoded Values**

Certain variables have new recoded values that were not presented as response options during the interview. Recoded values are assigned to items with Other-Specify response options. The label for any recoded value is prefixed with the term “RECODE:” for explicit identification.

### **7.6.3 Values for Derived, Imputed, and External Variables**

The labels for all created values for derived, imputed, and external variables include the coded value and a description of the value. If a value could not be assigned, it is coded as a missing value. See Section 7.7 for a description of the missing value codes.

## **7.7 Missing Values**

The data files include missing values, as currently defined in Table 7-3. The data user is advised to refer to the descriptive frequencies in the documentation accompanying the data files to ascertain the levels of missing data in variables of interest.

**Table 7-3. PATH Study missing value codes and their descriptions**

<b>Value</b>	<b>Definition<sup>a</sup></b>	<b>Applicable data types</b>	<b>Applicable variable types</b>	<b>Description<sup>a</sup></b>
-1	Inapplicable	Numeric, Character	Interview	This value is coded to identify items that were validly skipped on a route specific to a respondent.
-5	Improbable response removed	Numeric	Interview	This value is coded to represent an improbable numeric answer provided by a respondent.
-6	Unassigned PERSONID	Character	Interview	This value is coded when a new person is added as a spouse or other guardian in parent interviews for Wave 2 and beyond.
-7	Refused	Numeric, Character	Interview	This value is coded when a respondent answered a question as "Refused."
-8	Don't know	Numeric, Character	Interview	This value is coded when a respondent answered a question as "Don't know."
-9	Missing – Not ascertained	Numeric, Character	Interview	This value is coded when data were supposed to be collected on a route specific to a respondent but the collection did not occur due to a skip error in the instrument.
-99933	Missing due to respondent not being in previous wave	Numeric, Character	Derived	This value is coded when a value in a derived variable is missing because a respondent was a shadow youth in a previous wave. It is also coded when a value in a derived variable is missing for respondents added in a replenishment wave. Coded in Waves 2 through 5 only.
-99999	Missing due to data not ascertained on one or more component variables	Numeric, Character	Derived	This value is coded when a value in a derived variable is missing because data were "Not ascertained" (as described above) in one or more component variables. Coded in Waves 1 through 5 only.
-99988	Missing due to a don't know response on one or more component variables	Numeric, Character	Derived	This value is coded when a value in a derived variable is missing due to a "Don't know" response in one or more component variables and none of those variables were "Not ascertained" (as described above). Coded in Waves 1 through 5 only.
-99977	Missing due to a refused response on one or more component variables	Numeric, Character	Derived	This value is coded when a value in a derived variable is missing due to a "Refused" response in one or more component variables and none of those variables were missing due to a "Don't know" response or because they were "Not ascertained" (as described above). Coded in Waves 1 through 5 only.
-99955	Missing due to an improbable response on one or more component variables	Numeric	Derived	This value is coded when a value in a derived variable is missing due to an improbable response in one or more component variables and none of those variables were missing due to "Don't know" or "Refused" responses or because they were "Not ascertained" (as described above). Coded in Waves 1 through 5 only.

**Table 7-3. PATH Study missing value codes and their descriptions (continued)**

<b>Value</b>	<b>Definition<sup>a</sup></b>	<b>Applicable data types</b>	<b>Applicable variable types</b>	<b>Description<sup>a</sup></b>
-99966	Missing due to missing data in a prior wave for one or more component variables	Numeric, Character	Derived	This value is coded when a value in a derived variable is missing because one or more component variables from any previous wave is coded as a missing value. It is also coded when a respondent did not participate in one or more previous waves. Coded in Waves 1 through 5 only.
-99911	Missing due to an instrument skip pattern for one or more component variables	Numeric, Character	Derived	This value is coded when a value in a derived variable is missing due to items that are not applicable/validly skipped for a respondent in one or more component variables and none of those variables were missing due to the other reasons listed above. Coded in Waves 1 through 5 only.
-97777	Missing due to data removed per respondent request	Numeric, Character	Interview, Derived	This value is coded when a respondent requested data removal.
.	System missing	Numeric	Derived	This value is coded when a derived numeric variable is missing because a value was not assigned by the algorithm. Coded in derived variables starting with Wave 5.5 and PATH-ATS.
NULL/ BLANK	Empty	Character	Derived	This value is coded when a derived character variable is missing because a value was not assigned by the algorithm. Coded in derived variables starting with Wave 5.5 and PATH-ATS.

<sup>a</sup> A component variable is any variable used in the creation of a derived variable. Unless otherwise stated, “component variables” are from the current wave.

An example of the implementation of these missing values codes can be seen in the “ever” status variables. An “ever” status identifies whether a respondent has ever used a tobacco product (e.g., derived variables named R0wR\_x\_EVR\_aaaaa, where *w* indicates the wave (1 through 5), *x* indicates “A” or “Y” respectively, and *aaaaa* is an abbreviation for a tobacco product). When participants did not respond in a prior wave but indicated a “never” status at their baseline wave and a “not in the past 12 months” status for all other waves in which they responded, an “ever” status was not conclusively identified. In such instances, the derived “ever” status variables in Waves 1 through 5 have been coded as missing, with a value -99966.

Starting with Wave 5.5 and PATH-ATS, system missing is assigned in all derived variables when a value cannot be logically assigned by the algorithm. With the exception of value -97777, no other 5-digit missing values are coded in derived variables, starting with Wave 5.5 and PATH-ATS.

## 7.8 Outlier Values

In all waves and modes of the PATH Study interviews, there were checks throughout the instruments to query potentially illogical or unlikely responses. However, the number of such checks was limited in order to minimize frustration on the part of the respondent. Moreover, these checks were designed to warn the respondent if a response was unlikely, but the respondent could continue through the instrument without changing the response. In order to maintain the confidentiality afforded by the ACASI instrument, it was important to minimize the number of instances where the respondent might need to ask the field interviewer for assistance in overcoming a check that required correction of data before continuing. As a result, the PATH Study questionnaire data include outliers, i.e., values that lie outside the expected range of data. There are two types of outliers, which were handled in two different ways.

First, there are outliers that are logically impossible. For example, a few respondents reported doing something at an age that was older than their current age (such as a respondent who was age 37 reporting initiation of cigarillo use at age 39). For these clear instances where a response was logically impossible, the outlier was removed from the dataset, replaced with a value of -5, and labeled as “Improbable response removed.” See Appendix F for a full list of interview variables with coded outlier values.

Second, there are outliers that are unlikely but logically possible. For example, a few respondents reported sending more than 1,000 text messages a day, which is unlikely but not impossible. In the absence of established upper or lower limits for such unlikely but logically possible values, they were retained in the dataset. Therefore, analysts are advised to review carefully the distributions of the variables used in all analyses, and to consider the impact of these values on any findings.

## 8. Linking Files

---

The PATH Study public use files include a master linkage data file that indicates in which file(s) a participant has data. The linkage file can help analysts identify which files contain data for a particular participant (or set of participants), and the linkage data file is extended with each wave. This linkage data file is described in more details in Section 8.1.

Data files may be linked together across waves. This type of linkage is described in Section 8.2. Sections 8.3 and 8.4 provide some general file merging notes and recommendations, and information about the case identification variable created by ICPSR.

### 8.1 Master Linkage Data File Structure

The master linkage data file includes a unique record for each PERSONID corresponding to each PATH Study participant:

- Wave 1 adult and youth interview respondents.
- Wave 1 shadow youth, even though no data exist for shadow youth in the PATH Study Wave 1 data files.
- Wave 4 adult and youth interview respondents from the Wave 4 replenishment sample.
- Wave 4 shadow youth from the Wave 4 replenishment sample, even though no data exist for shadow youth in the PATH Study Wave 4 data files.
- Wave 7 adult and youth interview respondents from the Wave 7 replenishment sample.
- Wave 7 shadow youth from the Wave 7 replenishment sample, even though no data exist for shadow youth in the PATH Study Wave 7 data files.

The records for shadow youth are included in the master linkage data file to account for participants who aged up to youth at a follow-up wave.

The master linkage data file will be maintained as a standalone dataset that will be updated with each data file release. The number of records in the master linkage data file is defined by the adult, youth, and shadow samples established at Wave 1 and future replenishment efforts. The number of variables in the master linkage data file will increase as data for future waves become available.

The master linkage data file includes indicator variables for the availability of interview data, weights, and biomarker data. Interview indicators are named WAVE $w$ \_INTERVIEW where  $w$  denotes the primary wave (e.g., 1, 2, 3). Biomarker indicators are named WAVE $w$ \_descriptor where  $w$  denotes the wave and *descriptor* denotes the data file contents (e.g., URINE\_COLLECTED, BLOOD\_COLLECTED, ASSPEC\_LAB, HSCRP\_LAB). For special collection waves between the biennial primary waves (i.e., the “point 5” waves), interview indicators are named WAVE $w$ P5\_INTERVIEW where  $w$  denotes the wave and “P5” signifies the “point 5.” (e.g., 4P5 for 4.5 and 5P5 for 5.5). The interview indicator for the special collection PATH-ATS is named ATS\_INTERVIEW.

To identify the wave in which each PATH Study participant was sampled, the master linkage data file includes a variable named RECRUITMENT\_WAVE. The master linkage data file also includes variables to identify each cohort, as described in Chapter 2. These variables are named WAVE $c$ \_COHORT, where  $c$  indicates the wave at which the cohort was established. See the master linkage data file codebook for a description of the variables and their values.

Table 8-1 illustrates possible response scenarios indicated by the comparison of the variable WAVE1\_INTERVIEW to variables added for subsequent waves for the PATH Study sample established at Wave 1.

**Table 8-1. Possible response scenarios when the variable WAVE1\_INTERVIEW is compared to variables added for subsequent waves**

Description/Scenario	WAVE1_INTERVIEW	WAVE $w$ _INTERVIEW <sup>a</sup>
Adult across waves	1 = Adult	1 = Adult
Adult in Wave 1, no response in Wave $w$	1 = Adult	9 = Nonresponse
Youth in Wave 1, Adult in Wave $w$	2 = Youth	1 = Adult
Youth across waves	2 = Youth	2 = Youth
Youth in Wave 1, no response in Wave $w$	2 = Youth	9 = Nonresponse
Shadow youth in Wave 1, Adult in Wave $w$	3 = Shadow Youth (No data)	1 = Adult
Shadow youth in Wave 1, Youth in Wave $w$	3 = Shadow Youth (No data)	2 = Youth
Shadow youth across waves	3 = Shadow Youth (No data)	3 = Shadow Youth (No data)
Shadow youth in Wave 1, no response in Wave $w$	3 = Shadow Youth (No data)	9 = Nonresponse

<sup>a</sup> Note that  $w$  indicates any wave after Wave 1. For Wave 4.5 (WAVE4P5\_INTERVIEW), value “1 = Adult” is not applicable. For PATH-ATS (ATS\_INTERVIEW), values “2 = Youth” and “3 = Shadow Youth (No data)” are not applicable.

The response scenarios illustrated in Table 8-1 also apply to PATH Study participants sampled at Wave 4, but they do not have records in any data files released for Waves 1 through 3. As such, for records that represent participants sampled at Wave 4 in the master linkage data file

(RECRUITMENT\_WAVE = 4), the value “0 = Not in sample” is assigned to all variables that correspond to data files for Waves 1 through 3. The value “0 = Not in sample” is also assigned to participants who were not sampled for the special collections (Wave 4.5, Wave 5.5, and PATH-ATS). Similarly, PATH Study participants sampled at Wave 7 do not have records in any data files released for Waves 1 through 6. As such, for records that represent participants sampled at Wave 7 in the master linkage data file (RECRUITMENT\_WAVE = 7), the value “0 = Not in sample” is assigned to all variables that correspond to data files for Waves 1 through 6.

Table 8-2 illustrates example response scenarios indicated by the comparison of the variables RECRUITMENT\_WAVE, WAVE1\_COHORT, WAVE4\_COHORT, and WAVE4\_INTERVIEW for the full PATH Study sample at Wave 4, including the Wave 4 replenishment.

**Table 8-2. Example response scenarios when the variables RECRUITMENT\_WAVE, WAVE1\_COHORT, WAVE4\_COHORT, and WAVE4\_INTERVIEW for the PATH Study sample at Wave 4 are compared**

Description/Scenario <sup>a</sup>	RECRUITMENT_WAVE	Cohort indicators <sup>b</sup>	WAVE4_INTERVIEW
Adult sampled in Wave 1, responding in Wave 4 and in the Wave 4 Cohort	1	WAVE1_COHORT = 1 WAVE4_COHORT = 1	1 = Adult
Adult sampled in Wave 4, responding in Wave 4	4	WAVE1_COHORT = 2 WAVE4_COHORT = 1	1 = Adult
Adult sampled in Wave 1, responding in Wave 4 and not in the Wave 4 Cohort	1	WAVE1_COHORT = 1 WAVE4_COHORT = 2	1 = Adult
Adult sampled in Wave 1, no response in Wave 4	1	WAVE1_COHORT = 1 WAVE4_COHORT = 2	9 = Nonresponse
Youth sampled in Wave 1, responding Youth in Wave 4 in the Wave 4 Cohort	1	WAVE1_COHORT = 1 WAVE4_COHORT = 1	2 = Youth
Youth sampled in Wave 4, responding in Wave 4	4	WAVE1_COHORT = 2 WAVE4_COHORT = 1	2 = Youth
Youth sampled in Wave 1, no response in Wave 4	1	WAVE1_COHORT = 1 WAVE4_COHORT = 2	9 = Nonresponse
Shadow youth sampled in Wave 1, responding Youth in Wave 4 and in the Wave 4 Cohort	1	WAVE1_COHORT = 1 WAVE4_COHORT = 1	2 = Youth
Shadow youth sampled in Wave 4, responding in Wave 4	4	WAVE1_COHORT = 2 WAVE4_COHORT = 1	3 = Shadow Youth (No data)

<sup>a</sup> Scenarios in this table are examples of response patterns in the data and are not meant to be comprehensive.

<sup>b</sup> 1=Yes, 2=No.

Table 8-3 illustrates example response scenarios indicated by the comparison of the variables RECRUITMENT\_WAVE, WAVE1\_COHORT, WAVE4\_COHORT, WAVE7\_COHORT, and WAVE7\_INTERVIEW for the full PATH Study sample at Wave 7, including the Wave 7 replenishment sample.

**Table 8-3. Example response scenarios when the variables RECRUITMENT\_WAVE, WAVE1\_COHORT, WAVE4\_COHORT, WAVE7\_COHORT, and WAVE7\_INTERVIEW for the PATH Study sample at Wave 7 are compared**

Description/Scenario <sup>a</sup>	RECRUITMENT_WAVE	Cohort Indicators <sup>b</sup>	WAVE7_INTERVIEW
Adult sampled at Wave 1, responding at Wave 4 and Wave 7, and in the Wave 4 and Wave 7 Cohorts	1	WAVE1_COHORT = 1 WAVE4_COHORT = 1 WAVE7_COHORT = 1	1 = Adult
Adult sampled at Wave 7, responding at Wave 7	7	WAVE1_COHORT = 2 WAVE4_COHORT = 2 WAVE7_COHORT = 1	1 = Adult
Adult sampled at Wave 1, responding at Wave 4 and Wave 7, and not in the Wave 4 or Wave 7 Cohorts	1	WAVE1_COHORT = 1 WAVE4_COHORT = 2 WAVE7_COHORT = 2	1 = Adult
Adult sampled at Wave 1, no response at Wave 4 and Wave 7	1	WAVE1_COHORT = 1 WAVE4_COHORT = 2 WAVE7_COHORT = 2	9 = Nonresponse
Youth sampled at Wave 1, responding Adult at Wave 4 and Wave 7, and in the Wave 7 Cohort	1	WAVE1_COHORT = 1 WAVE4_COHORT = 1 WAVE7_COHORT = 1	1 = Adult
Youth sampled at Wave 7, responding at Wave 7	7	WAVE1_COHORT = 2 WAVE4_COHORT = 2 WAVE7_COHORT = 1	2 = Youth
Youth sampled at Wave 1, responding at Wave 4, no response at Wave 7	1	WAVE1_COHORT = 1 WAVE4_COHORT = 1 WAVE7_COHORT = 2	9 = Nonresponse
Shadow youth sampled at Wave 4, responding Youth at Wave 7, and in the Wave 4 and Wave 7 Cohorts	4	WAVE1_COHORT = 2 WAVE4_COHORT = 1 WAVE7_COHORT = 1	2 = Youth
Shadow youth sampled at Wave 7, responding at Wave 7	7	WAVE1_COHORT = 2 WAVE4_COHORT = 2 WAVE7_COHORT = 1	3 = Shadow Youth (No data)

<sup>a</sup> Scenarios in this table are examples of response patterns in the data and are not meant to be comprehensive.

<sup>b</sup> 1=Yes, 2=No.

## 8.2 Linking Data Files Across Waves

The unique identifier in all files is PERSONID. This allows users to easily merge the data from multiple waves for longitudinal analyses.

It is not necessary to use the master linkage data file to merge data across waves. For example, the Wave 1 adult data file can be merged directly with the Wave 2 adult data file using PERSONID. The

primary utility of the master linkage data file is to help analysts identify the different types of data available for a particular respondent or set of respondents as identified by PERSONID, both within and across data collection waves. For example, the master linkage data file can be used to identify the subset of respondents who completed a youth interview in Wave 1 and an adult interview in Wave 2. The resulting subset of PERSONIDs identified using the master linkage data file can be used to merge records from multiple data files.

## **8.3 File Merging Notes and Recommendations**

Data users are advised to avoid merges that retain all variables from multiple files because the adult and youth/parent data files each have a large number of variables. The process of merging data files with large numbers of variables may not be resource intensive, but the large size of the resulting (merged) file could significantly increase the time required to run analyses. As such, when merging data files, users are advised to keep only the variables needed from each file to run the intended analyses. This practice keeps the merged file(s) as small as possible, which helps analytic processes run faster.

The records included in the file(s) that result from merging two or more data files can vary depending on how the merge is specified. Users should carefully consider the set of cases needed for any specific analyses and define the merge accordingly.

## **8.4 Case Identification Number**

ICPSR creates a variable for each of its datasets named CASEID. This variable numbers the cases sequentially so that the file can be easily returned to its original order at any time. This variable should not be confused with any other identification variable previously mentioned. It should **not** be used to merge files or link records together.

## References

---

- American Association of Public Opinion Research. (2023). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, 10th edition*. AAPOR. <https://aapor.org/wp-content/uploads/2023/05/Standards-Definitions-10th-edition.pdf>.
- Battaglia, M.P., Hoaglin, D.C., and Frankel, M.R. (2013). Practical considerations in raking survey data. *Survey Practice*, 2(5).
- Chowdhury, S.R. (2013). Effects of poststratification and raking on precision of MEPS estimates. Proceedings of the Federal Committee on Statistical Methodology Conference, [https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/05/F1\\_Chowdhury\\_2013FCSM.pdf](https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/05/F1_Chowdhury_2013FCSM.pdf).
- Chowdhury, S., Khare, M., and Wolter, K. (2007). Weight trimming in the National Immunization Survey. *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 2651-2658.
- Golder, P.A. and Yeomans, K.A. (1973). The use of cluster analysis for stratification. *Applied Statistics*, 22, 213-219.
- Graubard, B.I. and Korn, E.L. (1996). Survey inference for subpopulations. *American Journal of Epidemiology*, 144, 102-106.
- Judkins, D.R. (1990). Fay's method for variance estimation. *Journal of Official Statistics*, 6, 223-239.
- Judkins, D.R. and Singh, R.P. (1981). Using clustering algorithms to stratify primary sampling units. *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 279-284.
- Korn, E. and Graubard, B. (2011). *Analysis of Health Surveys*. Hobart, NJ: Wiley.
- Krewski, D. and Rao, J.N.K. (1981). Inference from stratified samples: Properties of the linearization, jackknife and balanced repeated replication methods. *The Annals of Statistics*, 9, 1010-1019.
- Lewis, T. (2013). Considerations and techniques for analyzing domains of complex survey data. SAS Global Forum Proceedings, <http://support.sas.com/resources/papers/proceedings13/449-2013.pdf>.
- Lohr, S. (2010). *Sampling: Design and Analysis*, 2<sup>nd</sup> ed. Boston: Brooks/Cole.
- McCarthy, P.J. (1969). Pseudo-replication: Half samples. *Review of the International Statistical Institute*, 37, 239-264.

## PATH Study Public Use Files User Guide

Opsomer, J., Dohrmann, S., DiGaetano, R., Piesse, A., Taylor, E., Creamer, M. R., Han, D., Everard, C., Borek, N., Hubbard, F., Zandberg, I., Kanel, M., Taylor, K., Kimmel, H. L., Paredes, A., Sugawara, Y., Cheng, Y., Kwan, J., Carusi, C., & Hyland, A. (2023). Update to the design and methods of the PATH Study, Wave 4. *Tobacco Control*.  
<https://doi.org/10.1136/tc-2022-057851>

Rao, J.N.K. and Shao, J. (1993). On balanced half sample variance estimation in stratified sampling. *Journal of the American Statistical Association*, 91, 343-348.

Rao, J.N.K. and Shao, J. (1999). Modified balanced repeated replication for complex survey data. *Biometrika*, 86, 403-415.

SAS Institute Inc. (2018). *SAS/STAT® 15.1 User's Guide*. Cary, NC: SAS Institute, Inc.

Schenker, N. and Gentleman, J.F. (2001). On judging the significance of differences by examining the overlap between confidence intervals. *The American Statistician*, 55(3), 182-186.

Tourangeau, R. and Yan, T. (2007). “Sensitive questions in surveys.” *Psychological Bulletin*, 133(5), 859-883.

Valliant, R. (1993). Poststratification and conditional variance estimation. *Journal of the American Statistical Association*, 88, 89-96.

Valliant, R. (2004). The effect of multiple weighting steps on variance estimation. *Journal of Official Statistics*, 20, 1-18.

Valliant, R. and Rust, K.F. (2010). Degrees of Freedom Approximations and Rules-of-Thumb. *Journal of Official Statistics*, 26, 585-602.

Wolter, K.M. (2007). *Introduction to Variance Estimation*, 2nd ed. New York, NY: Springer.

Yan, T. and Cantor, D. (2019). Asking survey questions about criminal justice involvement. *Public Health Reports*, 134, 46s-56s, doi:10.1177/0033354919826566.

Yan, T. (2021). Consequences of asking sensitive questions. *Annual Review of Statistics and Its Application*, 8, doi:10.1146/annurev-statistics-040720-033353.

## Appendix A

### PATH Study Protocol by Wave

---

Table A-1. PATH Study Protocol by Wave for Waves 1 – 7

Wave	Data Collection Period	Interview Mode/ Data Collectors	PATH Study Participants Fielded	Instruments <sup>a</sup>	Biospecimens Collected <sup>b</sup>
Wave 1	September 12, 2013 – December 14, 2014	In-person/Field interviewers	Initial sample recruitment and data collection – adults ages 18 and older, youth ages 12 to 17, and shadow youth ages 9 to 11	ACASI <sup>c</sup> Adult ACASI Youth CAPI <sup>d</sup> Parent	<b>Urine</b> (adults only) <b>Blood</b> (adults only) <b>Buccal cells</b> (adults only) <sup>e</sup>
Wave 2	October 23, 2014 – October 30, 2015	In-person/Field interviewers	Continuing sample participants recruited in Wave 1 – adults and aged-up adults ages 18 and older, youth and aged-up youth ages 12 to 17 <sup>f</sup>	ACASI Adult ACASI Youth CAPI Parent	<b>Urine</b> (aged-up adults and some continuing adults who provided urine previously) <b>Blood</b> (aged-up adults only)
Wave 3	October 19, 2015 – October 23, 2016	In-person/Field interviewers	Continuing sample participants recruited in Wave 1 – adults and aged-up adults ages 18 and older, youth and aged-up youth ages 12 to 17	ACASI Adult ACASI Youth CAPI Parent	<b>Urine</b> (aged-up adults and some continuing adults who provided urine previously) <b>Blood</b> (aged-up adults only)
Wave 4	December 1, 2016 – January 3, 2018 <sup>g</sup>	In-person/Field interviewers	Continuing sample participants recruited in Wave 1 – adults and aged-up adults ages 18 and older, youth and any remaining aged-up youth ages 12 to 17  Replenishment sample participants recruited – adults ages 18 and older, youth ages 12 to 17, and shadow youth ages 10 to 11	ACASI Adult ACASI Youth CAPI Parent	<b>Urine</b> (new adults, aged-up adults, and some continuing adults who provided urine previously; and new youth, aged-up youth, and continuing youth) <b>Blood</b> (new and aged-up adults)

**Table A-1.** PATH Study Protocol by Wave for Waves 1 – 7 (continued)

Wave	Data Collection Period	Interview Mode/ Data Collectors	PATH Study Participants Fielded	Instruments <sup>a</sup>	Biospecimens Collected
Wave 4.5	December 1 2017 – December 1 2018 <sup>h</sup>	In-person/Field interviewers	Special collection wave of continuing sample youth and aged-up youth only, ages 12 to 17, recruited in Wave 1 or Wave 4	ACASI Youth CAPI Parent	No biospecimen collection
Wave 5	December 1, 2018 – November 30, 2019	In-person/Field interviewers	Continuing sample participants recruited in Wave 1 or Wave 4 – adults and aged-up adults ages 18 and older, youth and aged-up youth ages 12 to 17	ACASI Adult ACASI Youth CAPI Parent	<b>Urine</b> (aged-up adults and some continuing adults who provided urine previously; aged-up youth and continuing youth who provided a urine sample in Wave 4 or did not participate in Wave 4) <b>Blood</b> (aged-up adults)
Wave 5.5 <sup>i</sup>	July 3, 2020 – December 31, 2020	Telephone/Field interviewers	Special collection wave of continuing sample participants in the Wave 4 Cohort ages 19 and younger on August 31, 2020 – youth and aged-up youth ages 13 to 17 and adults and aged-up adults ages 18 and 19 on date of Wave 5.5 interview	ACASI Adult adapted for telephone ACASI Youth adapted for telephone CAPI Parent adapted for telephone	No biospecimens collected
PATH-ATS <sup>j</sup>	September 10, 2020 – December 20, 2020	Telephone/Trained telephone interviewers not previously associated with the PATH Study	Special collection effort of ~50% of Wave 5 respondents in the Wave 4 Cohort ages 20 and older on August 31, 2020	Adapted from Wave 5.5 adult instruments to be shorter and adapted for CATI <sup>k</sup>	No biospecimens collected
Wave 6	March 1, 2021–November 30, 2021	Telephone & In-person/Field interviewers	Continuing sample participants recruited in Wave 1 or Wave 4 – adults and aged-up adults ages 18 and older, youth and aged-up youth ages 14 to 17 <sup>l</sup>	ACASI Adult/ACASI Youth ACASI Adult adapted for telephone ACASI Youth adapted for telephone CAPI Parent	No biospecimens collected

**Table A-1.** PATH Study Protocol by Wave for Waves 1 – 7 (continued)

Wave	Data Collection Period	Interview Mode/ Data Collectors	PATH Study Participants Fielded	Instruments <sup>a</sup>	Biospecimens Collected
Wave 7	January 6, 2022 – April 2, 2023 <sup>m</sup>	Telephone & In-person/Field interviewers  Web/Self-administration <sup>n</sup>	Continuing sample participants recruited in Wave 1 or Wave 4 – adults and aged-up adults ages 18 and older, youth ages 14 to 17  Replenishment sample participants recruited – adults ages 18 and older, youth ages 12 to 17, and shadow youth ages 9 to 11	ACASI Adult/ACASI Youth ACASI Adult adapted for telephone ACASI Youth adapted for telephone Web Adult/Web Youth CAPI Parent Web Parent	<b>Urine</b> (selected new adults, selected aged-up and continuing adults who might or might not have provided urine previously; selected continuing youth who provided urine previously) <b>Blood</b> (selected new, continuing, and aged-up adults) <sup>o</sup>

<sup>a</sup> All instruments were provided in English and Spanish except those used for the Wave 7 Web pilot test. Administration time was stable over the waves with ACASI adult interviews averaging 60 minutes, ACASI youth interviews averaging 45 minutes, and CAPI parent interviews averaging 20 minutes. The telephone administration of the ACASI instruments during Wave 5.5, Wave 6, and Wave 7 increased the ACASI adult interviews to an average of 95 minutes and ACASI youth interviews to an average of 71 minutes. The PATH-ATS CATI adult interview averaged 30 minutes.

<sup>b</sup> Biomarker data are available only in restricted-use format.

<sup>c</sup> ACASI = audio-computer assisted self-interviewing.

<sup>d</sup> CAPI = computer-assisted personal interviewing.

<sup>e</sup> Buccal cell collection was discontinued midway through Wave 1.

<sup>f</sup> Aged-up adults are continuing sample participants who last participated as a youth and are now age 18 or older. Aged-up youth are shadow youth who are now age 12 or older and are eligible to participate in the PATH Study.

<sup>g</sup> Wave 4 was extended until January 3, 2018, due to severe weather events that disrupted data collection and to complete recruiting with the large shadow youth sample.

<sup>h</sup> Wave 4.5 was extended one day to conduct interviews with three additional youth who had aged-up to the appropriate youth group.

<sup>i</sup> The COVID-19 pandemic caused a switch to telephone for this wave.

<sup>j</sup> A special data collection during COVID-19 pandemic to collect data from those ages 20 and older.

<sup>k</sup> CATI = computer-assisted telephone interviewing.

<sup>l</sup> At Wave 6, the PATH Study did not have any youth ages 12 or 13.

<sup>m</sup> Wave 7 was extended until April 2, 2023, to complete recruiting the large shadow youth and young youth samples.

<sup>n</sup> A web pilot test was launched to interview a sample of adults, parents, and youth via the Web.

<sup>o</sup> Blood collection was discontinued early in Wave 7.

## Appendix B

# Weighting Procedures

---

### B.1 Wave 1 Weights

The sections below describe the computation of the household weights and the adult and youth weights included on the Wave 1 adult data file and youth/parent data file, respectively. Also included is a description of weights created for shadow youth respondents in preparation of their completing a youth interview in later waves. These weights are not included with the data files.

#### B.1.1 Household Weights

The initial household weights, denoted as  $W1HIPSWT_{ijk}$ , were calculated for all sampled addresses as the IPS as shown in equation B.1:

$$W1HIPSWT_{ijk} = \frac{1}{P_{ijk}} \quad (\text{B.1})$$

where  $P_{ijk}$  is the probability that household  $k$  in segment  $j$  of PSU  $i$  was selected to be in the sample. However, some sampled addresses could not be located/accessed, others were found to be ineligible (e.g., vacant lots and group quarters), and some eligible households did not complete the household screener. Adjustments were therefore made to the IPS weights of addresses with known eligibility status to compensate for those with unknown eligibility status. This eligibility adjustment was done separately for each census region. The household weight after this adjustment, represented as  $W1HNRUNK_{ijk}$ , was computed as shown in equation B.2:

$$W1HNRUNK_{ijk} = W1HIPSWT_{ijk} \times W1HN_r \quad (\text{B.2})$$

where  $W1HN_r$  is the sum of  $W1HIPSWT_{ijk}$  for all selected addresses in census region  $r$  divided by the sum of  $W1HIPSWT_{ijk}$  for addresses with known eligibility status in that region.

Further adjustments were made within weighting classes (or nonresponse adjustment cells) based on information available for both responding and nonresponding households, including census 2010 and ACS 5-year (2009-2013) data pertaining to the segments, tracts, and blocks in which they were located. Census 2010 data were used to calculate the percent of occupied housing units that were

owner-occupied, the percent of the population who were Black or African American,<sup>32</sup> the percent of the population who were Asian,<sup>33</sup> the percent of the population who were Hispanic, and the percent of the population ages 25 and older in the census block containing the address. The ACS 5-year data were used to estimate the median monthly housing unit costs in the census tract containing the address. Census region, the urbanicity of the PSU, and the urbanicity of the segment were also used when forming the weighting classes.

A second adjustment was made to account for nonresponse among addresses corresponding to eligible households. Within a weighting class, the weights from the previous step for the responding households were inflated proportionately so that they produced the same sum as the sum of these weights for the responding and nonresponding households combined. The nonresponse-adjusted household weight (denoted as  $W1HNRWT_{ijk}$ ) for responding household  $k$  of PSU  $i$  and segment  $j$  was calculated as shown in equation B.3:

$$W1HNRWT_{ijk} = W1HNRUNK_{ijk} \times W1HN_c \quad (\text{B.3})$$

where  $W1HN_c$  is the sum of  $W1HNRUNK_{ijk}$  for eligible sampled households in weighting class  $c$  of which household  $k$  is a member, divided by the sum of  $W1HNRUNK_{ijk}$  for all responding households in that weighting class.

The nonresponse-adjusted weights were raked to household counts from the 2013 ACS 1-year Public Use Microdata Sample (ACS PUMS) by census region and household composition. Household composition was defined by the number of non-adult persons in the household (0, 1, or 2+) and the number of adult household members (1, 2, 3+). For raking purposes, the household composition was imputed for households missing this information using logical imputation.<sup>34</sup> The final raked household weight was calculated as shown in equation B.4:

$$W1HRKWT_{ijk} = W1HNRWT_{ijk} \times W1HR_{ijk} \quad (\text{B.4})$$

---

<sup>32</sup>Black or African American was defined as Black or African American alone or in combination with other race(s), whether Hispanic or non-Hispanic.

<sup>33</sup>Asian was defined as Asian alone.

<sup>34</sup>See Lohr (2010) for a brief description of raking and imputation methods.

where  $W1HR_{ijk}$  is the household raking adjustment factor for household  $k$  of PSU  $i$  and segment  $j$ . The household weights are not included on the data files.

## B.1.2 Adult Weights

The raked household-level weight was used as the foundation for calculating the adult weight. The adult base weight, denoted as  $W1AP1BWT_{ijkl}$ , was computed as the product of the final household weight  $W1HRKWT_{ijk}$  and the inverse of the within-household probability of selection for adult  $l$  within household  $k$  of PSU  $i$  and segment  $j$ , as shown in equation B.5:

$$W1AP1BWT_{ijkl} = W1HRKWT_{ijk} \times \frac{1}{\text{Probability adult } l \text{ selected at Phase 1 from household (ijk)}}. \quad (\text{B.5})$$

The final weights for adults were computed in three steps.

First, a nonresponse adjustment was performed to account for nonresponse to the Phase 2 screener using a combination of census 2010 and ACS 5-year (2009-2013) data (used for the household nonresponse adjustment) and person-level data collected during the household screener. Weighting classes were formed based on census region and urbanicity of the PSU and segment; the percent of occupied housing units that were owner-occupied, the percent of the population who were Black or African American, the percent of the population who were Asian, the percent of the population who were Hispanic, and the percent of the population ages 25 and older in the census block containing the address; the median housing unit costs in the census tract containing the address; the age and sex of the household screener respondent; the number of adults in the household (capped at five); and the age, race/ethnicity, sex, and tobacco-use status of the selected adult (as reported by the household screener respondent).<sup>35</sup>

The resulting adult weight (denoted as  $W1AP1NRWT_{ijkl}$ ), adjusted for nonresponse between Phases 1 and 2 of the adult sampling procedure, was calculated for respondents to the Phase 2 Screener as shown in equation B.6:

$$W1AP1NRWT_{ijkl} = W1AP1BWT_{ijkl} \times W1AP1N_c \quad (\text{B.6})$$

---

<sup>35</sup>Block-level population percentages were created from census 2010 data. Tract-level median housing unit costs were extracted from the ACS 5-year (2009-2013) data.

where  $W1AP1N_c$  is the sum of  $W1AP1BWT_{ijkl}$  for adults sampled at Phase 1 in weighting class  $c$  of which respondent  $l$  is a member, divided by the sum of  $W1AP1BWT_{ijkl}$  for all adults responding to the Phase 2 screener in that weighting class.

Second, the probability of selection at Phase 2 was used to compute the Phase 2 weight, denoted as  $W1AP2WT_{ijkl}$ , as shown in equation B.7:

$$W1AP2WT_{ijkl} = W1AP1NRWT_{ijkl} \times \frac{1}{\text{Probability adult } l \text{ from household } (ijk) \text{ selected at Phase 2}}. \quad (\text{B.7})$$

Finally, raking and trimming were performed in an iterative process. The Phase 2 adult weights were raked to independent population totals based on data from the 2013 ACS PUMS. The raking was done using combinations of census region, age, race/ethnicity, sex, and educational attainment. These variables were imputed if they were missing. (See Section 5.3 for more information about this imputation.)

However, the raking algorithm did not place any restrictions on the highest and lowest values of the raked weights, and a few of the raked weights became extremely large in the process of matching to the population totals from the 2013 ACS PUMS. To reduce any extreme weights generated in the raking process, a trimming step was performed to bring any extreme weights down to the median weight plus four times the interquartile range<sup>36</sup> within groups defined by the eight age, race, and tobacco-use status categories used to select the adults at Phase 2 (see Section 2.1.2.4). After trimming, the weighted totals no longer matched the control totals, so the raking process was repeated. The trimming and raking steps were iterated until the resulting weights summed to the 2013 ACS PUMS totals for the raking dimensions and were within the bounds defined by the interquartile range criterion.

After the iterative raking and trimming process, the final adult weight, denoted as  $W1ARKWT_{ijkl}$ , was calculated as shown in equation B.8:

$$W1ARKWT_{ijkl} = W1AP2WT_{ijkl} \times W1ART_{ijkl} \quad (\text{B.8})$$

---

<sup>36</sup>The weight trimming procedure reduces the mean squared error by reducing the variation among the weights. See Battaglia et al. (2013) and Chowdhury et al. (2007).

where  $W1ART_{ijkl}$  is the combined raking and trimming adjustment for adult  $l$  within household  $k$  of PSU  $i$  and segment  $j$ . These final weights can be found on the Wave 1 adult data file in the variable named R01\_A\_PWGT, indicating that these are person-level, adult weights for Wave 1 of the PATH Study.

### B.1.3 Youth Weights

The raked household-level weight was also used as the foundation for calculating the youth weight. The youth base weight, denoted as  $W1YBWT_{ijkl}$ , was computed as the product of the final household weight  $W1HRKWT_{ijk}$  and the inverse of the within-household probability of selection for youth  $l$  within household  $k$  of PSU  $i$  and segment  $j$ , as shown in equation B.9:

$$W1YBWT_{ijkl} = W1HRKWT_{ijk} \times \frac{1}{\text{Probability youth } l \text{ selected from household (ijk)}}. \quad (\text{B.9})$$

Like the adjustment for household screener nonresponse, a nonresponse adjustment was performed to account for nonresponse to the youth interview using a combination of census 2010 and ACS 5-year (2009-2013) data (used for the household nonresponse adjustment) and person-level data collected during the household screener. Weighting classes were formed based on census region and urbanicity of the PSU and segment; the percent of occupied housing units that were owner-occupied, the percent of the population who were Black or African American, the percent of the population who were Asian, the percent of the population who were Hispanic, and the percent of the population ages 25 and older in the census block containing the address; the median housing unit costs in the census tract containing the address; the age and sex of the household screener respondent; the number of adults in the household (capped at five); and the age, sex, and race/ethnicity of the selected youth (as reported by the household screener respondent).

Within a weighting class, the base weights ( $W1YBWT_{ijkl}$ ) for the responding youth were inflated proportionately so that they produced the same sum as the sum of the base weights of the responding and nonresponding youth combined. The nonresponse-adjusted weight for responding youth is shown in equation B.10:

$$W1YNRWT_{ijkl} = W1YBWT_{ijkl} \times W1YN_c \quad (\text{B.10})$$

where  $W1YN_c$  is the sum of  $W1YBWT_{ijkl}$  for sampled youth in weighting class  $c$  of which respondent  $l$  is a member, divided by the sum of  $W1YBWT_{ijkl}$  for all responding youth in that weighting class.

For youth, the nonresponse-adjusted weights ( $W1YNRWT_{ijkl}$ ) were raked to population totals from the 2013 ACS PUMS using the same iterative process as used in the adult weighting described above. The raking was done using census region, single year of age, race/ethnicity, and sex as raking variables. These variables were imputed if they were missing. (See Section 5.3 for more information about this imputation.) The trimming threshold was set to the median weight plus four times the interquartile range within groups defined by whether or not the youth was selected with certainty.

After the iterative raking and trimming process, the final youth weight, denoted as  $W1YRKWT_{ijkl}$ , was calculated as shown in equation B.11:

$$W1YRKWT_{ijkl} = W1YNRWT_{ijkl} \times W1YRT_{ijkl} \quad (\text{B.11})$$

where  $W1YRT_{ijkl}$  is the combined trimming and raking adjustment for youth  $l$  within household  $k$  of PSU  $i$  and segment  $j$ . The final weight can be found on the Wave 1 youth/parent data file in the variable named R01\_Y\_PWT, indicating that it is person-level, youth weight for Wave 1 of the PATH Study.

#### B.1.4 Shadow Youth Weights

In preparation of the shadow youth completing a youth interview in later waves, Wave 1 weights were created for the shadow youth sample so that Wave 1 shadow youth would have a basis for the future-wave weights. These weights are not included with the data files.

The weighting process for the Wave 1 shadow youth was almost identical to that for the Wave 1 youth sample, as described above. There were some slight differences, however. These include the variables used to form the weighting classes for the nonresponse adjustment and the level of information available for the raking adjustment.

Nonresponse adjustment was performed to account for nonresponse to the parent consent using a combination of census 2010 and ACS 5-year (2009-2013) data (used for the household nonresponse

adjustment) and person-level data collected during the household screener. Weighting classes were formed based on census region; the percent of occupied housing units that were owner-occupied, the percent of the population who were Black or African American, the percent of the population who were Hispanic, and the percent of the population ages 25 and older in the census block containing the address; the median housing unit costs in the census tract containing the address; the age and sex of the household screener respondent; the number of adults in the household (capped at five); and the age and race/ethnicity of the selected youth.

Nonresponse-adjusted weights were raked to population totals from the 2013 ACS PUMS using census region, single year of age, race/ethnicity, and sex as raking variables. However, because no interview is conducted for shadow youth until they reach age 12, the only information available for this adjustment was from the household screener. These variables were imputed if they were missing, as described in Section 5.3, but without the benefit of self-reported information.

## B.2 Wave 2 Weights

The final weights assigned to the Wave 1 Cohort served as the initial (“base”) weights for use in developing the Wave 2 weights. These weights were then adjusted to account for nonresponse to the Wave 2 interview, or in the case of shadow youth, failure to verify their information with the study; the resulting weights were raked to Wave 1 population totals (also known as “control totals”) as described in the sections below. The weighting process consisted of partitioning the sample into groups defined by Wave 1 age (Section B.2.1), forming weighting classes and performing two nonresponse adjustments (Section B.2.2), and raking to Wave 1 control totals to form the final Wave 2 weights (Section B.2.3).

### B.2.1 Partitioning the Sample by Wave 1 Age

The nonresponse adjustment was done separately for each of eleven age groups based on Wave 1 age: 9, 10, 11, 12, 13, 14, 15, 16, 17, 18-24, and 25+. The main rationale for doing so was to appropriately handle those who age into a different age group (from shadow youth to youth or from youth to adult). For example, because of the timing of data collection it was possible that some 16-year-olds at Wave 1 (those who were close to 17 at the time of their Wave 1 youth interview) were 18 at Wave 2 because more than 1 year elapsed between interviews, while some 17-year-olds at

Wave 1 did not age up to the adult group at Wave 2 because less than a full year had elapsed between interviews. This issue also affected 10- and 11-year-olds regarding turning 12 at Wave 2.

## B.2.2 Nonresponse Adjustment

Final Wave 1 weights were adjusted for Wave 2 nonresponse in two stages, first accounting for nonresponse among nonrespondents for whom eligibility status could not be ascertained and then among those for whom it could. The first adjustment for nonrespondents whose eligibility status was unknown was undertaken for those in the Wave 1 Cohort who could not be contacted to complete the Wave 2 interview. The second adjustment was undertaken to account for nonrespondents known to be eligible for the Wave 2 interview. This group of nonrespondents largely consisted of those who were contacted but did not complete the Wave 2 interview (e.g., “refusals”).

A number of variables were considered in establishing weighting classes for nonresponse adjustment purposes, including variables from census 2010 and 5-year ACS (2009-2013) data (used for the household nonresponse adjustment described in Section B.1), person-level data collected during the Wave 1 household screener, and data from the Wave 1 adult and youth interviews. Because different information was collected in the Wave 1 adult and youth interviews, and no interview data were collected at Wave 1 for the shadow youth, different sets of variables were considered in the formation of the weighting classes for the different sets of Wave 1 participants: one set of variables was considered for adult respondents in Wave 1, another for youth respondents, and a third, smaller set, for shadow youth. The variables comprising the final weighting classes  $c_U$  (for the unknown eligibility adjustment) and  $c_K$  (for the nonresponse adjustment among those known to be eligible) were determined separately for the eleven age groups based on their ability to discriminate response propensities.

Suppose Wave 1 Cohort member  $l$  was assigned weight  $W1_l$  for Wave 1 and was assigned to class  $c_U$  for the nonresponse adjustment to account for those with unknown eligibility. Then the Wave 2 weight after this adjustment, represented as  $W2NRUNK_l$ , was computed as shown in equation B.12:

$$W2NRUNK_l = W1_l \times W2N_{c_U} \quad (\text{B.12})$$

where  $W2N_{c_U}$  is the sum of  $W1_l$  for all Wave 1 Cohort members in weighting class  $c_U$ , divided by the sum of  $W1_l$  for all Wave 1 Cohort members with known eligibility status at Wave 2 in that weighting class.

This weight was then adjusted for the nonresponse of those known to be eligible for the Wave 2 interview. Let  $c_K$  represent the weighting class to which Wave 1 Cohort member  $l$  was assigned for this second nonresponse adjustment. Then the final Wave 2 nonresponse-adjusted weight for respondent  $l$ , represented as  $W2NRWT_l$ , was computed as shown in equation B.13:

$$W2NRWT_l = W2NRUNK_l \times W2N_{c_K} \quad (\text{B.13})$$

where  $W2N_{c_K}$  is the sum of  $W2NRUNK_l$  for all those eligible for Wave 2 in weighting class  $c_K$ , divided by the sum of  $W2NRUNK_l$  for all Wave 2 respondents in that weighting class.

### B.2.3 Raking Adjustment

Finally, raking to control totals using Wave 1 characteristics and trimming were performed in an iterative process. Raking to control totals from the recruitment wave is used with longitudinal weights because the target population for longitudinal weights is the target population of initial interest (for the Wave 1 Cohort of the PATH Study, the CNP at the time of Wave 1) followed over time, including those who may enter the military or group quarters (and are not incarcerated) after Wave 1 (see Section 2.2). “Drifting” from important Wave 1 characteristics of analytic interest can arise over time due to nonresponse as well as to Wave 1 Cohort members becoming ineligible in later waves (i.e., some will have died, while others will have moved to a location outside of the United States). To maintain consistency with such Wave 1 characteristics, the Wave 2 weights for everyone other than Wave 2 nonrespondents were raked back to control totals for all Wave 1 participants using Wave 1 characteristics. Some of these control totals were simply the population-based control totals used in the Wave 1 raking. Others were sample-based control totals, reflecting PATH Study estimates related to tobacco use.

For Wave 2 adults who were also adults in Wave 1, the population-based control totals used were the same as those used at Wave 1. The sample-based control totals reflected any tobacco use and e-cigarette use reported at Wave 1 cross-classified with some of the same characteristics used to create

the population-based control totals (e.g., sex and age) at Wave 1. The tobacco-use variable denotes the type of tobacco use for each person who responded to the survey and had three types of use: current established use, ever use but not current established use, and never use. Those with missing data for tobacco use were pooled for raking purposes with those associated with the category “ever use but not current established use.” The e-cigarette variable had two types of use: never use and other. Those with missing data for e-cigarette use were pooled with those associated with the category “other.”

For Wave 2 participants who completed the youth interview at Wave 1, the population-based control totals used were the same as those used at Wave 1. The sample-based control totals were created from Wave 1 estimates of any tobacco use (for participants who were 16 or 17 at Wave 1) and e-cigarette use (for participants who were 17 at Wave 1);<sup>37</sup> the tobacco-use and e-cigarette variables were both defined using two levels: never user and other; those with missing data on either tobacco variable were pooled with those associated with the category “other” for that variable.

For Wave 2 participants who were shadow youth at Wave 1, the raking was done as described in Section B.1.4.

However, the raking algorithm did not place any restrictions on the highest and lowest values of the raked weights, and a few of the raked weights became extremely large in the process of matching to the control totals. To reduce any extreme weights generated in the raking process, a trimming step was performed to bring any extreme weights down to the median weight plus four times the interquartile range within groups used to select the PATH Study participants at Wave 1. For adults who were also adults at Wave 1, these were the eight groups defined by age, race, and tobacco-use status (see Section 2.12.4); for participants who were not adults at Wave 1, these were groups defined by whether or not they were selected with certainty.

---

<sup>37</sup>The ages for which the e-cigarette use and any tobacco use dimensions were used for raking were determined by the need to have sufficient sample size for stability of the weighting adjustments.

After trimming, the weighted totals no longer matched the control totals, so the raking process was repeated. The final Wave 2 weight, denoted as  $W2RKWT_l$ , was thus computed for each respondent  $l$  as the product shown in equation B.14:

$$W2RKWT_l = W2NRWT_l \times W2RT_l \quad (\text{B.14})$$

where  $W2RT_l$  is the combined raking and trimming adjustment for respondent  $l$ . The final weight for each Wave 1 Cohort member responding to the Wave 2 adult interview (“A”) or youth interview (“Y”) is stored in the variable named R02\_x\_PWGT (where  $x$  is either “A” or “Y” as appropriate), indicating that it is a person-level weight for Wave 2 of the PATH Study.

## B.3 Wave 3 Weights

There are two longitudinal weights available for analysis of Wave 3 data: the all-waves weight and the single-wave weight. The Wave 3 all-waves weight was assigned to Wave 3 respondents who also responded at both Wave 1 and Wave 2. The Wave 3 single-wave weight was assigned to all Wave 3 respondents whether or not they participated at Wave 2. So respondents in all three waves were assigned both the all-waves and single-wave weights. Both weights are longitudinal; however, the single-wave weight could serve as a proxy for a cross-sectional weight.

Section B.3.1 describes the construction of the Wave 3 all-waves weights and Section B.3.2 describes the Wave 3 single-wave weights.

### B.3.1 Creation of the Wave 3 All-Waves Weights

All-waves weights were created for Wave 3 respondents who also responded at both Wave 1 and Wave 2. This includes Wave 3 respondents who either:

- Completed an interview (adult or youth) at Wave 1 and Wave 2; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1, were at least age 12 at the time of the Wave 3 interview, and completed an interview for all waves in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed.

The Wave 3 all-waves weighting process began with all Wave 2 respondents and the nonresponse-adjusted weights assigned to them ( $W2NRWT_l$ ) during the Wave 2 weighting process. These initial

weights were then adjusted to account for nonresponse to the Wave 3 interview as described in the sections below. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 2 age and forming nonresponse adjustment cells (Section B.3.1.1), and adjusting for nonresponse and raking to control totals to form the final Wave 3 all-waves weights (Section B.3.1.2).

### **B.3.1.1 Partitioning the Sample by Wave 2 Age and Forming Nonresponse Adjustment Cells**

The adjustment of weights for nonresponse was done separately for each of ten age groups based on Wave 2 age, as this was the most recently confirmed age value available for the Wave 2 respondents. These groups were 9- to 10-year-olds and the seven individual ages from 11 to 17 as well as the two adult age ranges 18-24, and 25 and older. This was considered preferable to using the Wave 3 age information which was not confirmed for nonrespondents. Using separate nonresponse adjustment subgroups based on age provides a mechanism for proportionately allocating weights of nonrespondents appropriately between youth and adults or shadow youth and youth.

The initial (“base”) weights (the Wave 2 weights reflecting adjustments for nonresponse) were adjusted for Wave 3 nonresponse in two stages. The first stage accounted for nonresponse among the nonrespondents for whom eligibility status could not be ascertained, the second stage among those for whom it could.

There were many variables available for consideration in establishing weighting cells for nonresponse adjustment purposes. One set of variables was considered for those ages 18 and older (i.e., adults) at Wave 2. Another set was considered for those ages 12 to 17 (i.e., youth) at Wave 2. A final set was considered for those ages 9-11 (i.e., shadow youth, mostly 10- and 11-year-olds) at Wave 2. The variables comprising the final weighting classes  $c_U$  (for the unknown eligibility adjustment) and  $c_K$  (for the nonresponse adjustment among those known to be eligible) were determined separately for the ten age groups.

### **B.3.1.2 Computation of the Nonresponse and Raking Adjustments**

Suppose a Wave 2 respondent  $l$  was assigned weight  $W2NRWT_l$  after the Wave 2 nonresponse adjustments and was assigned to cell  $c_U$  for the first stage of nonresponse adjustment. Then the

resulting Wave 3 all-waves weight after this adjustment, represented as  $W3NRUNK\_A_l$ , was computed as shown in equation B.15:

$$W3NRUNK\_A_l = W2NRWT_l \times W3N\_A_{c_U} \quad (\text{B.15})$$

where  $W3N\_A_{c_U}$  is the sum of  $W2NRWT_l$  for all Wave 2 respondents in weighting class  $c_U$ , divided by the sum of  $W2NRWT_l$  for Wave 2 respondents with known eligibility status at Wave 3 in that weighting class.

This weight was then adjusted for the nonresponse of those known to be eligible for the Wave 3 interview. Let  $c_K$  represent the cell to which a Wave 2 respondent  $l$  known to be eligible for data collection at Wave 3 was assigned for nonresponse adjustment purposes. Then the Wave 3 all-waves nonresponse-adjusted weight for respondent  $l$ , represented as  $W3NRWT\_A_l$ , was computed as shown in equation B.16:

$$W3NRWT\_A_l = W3NRUNK\_A_l \times W3N\_A_{c_K} \quad (\text{B.16})$$

where  $W3N\_A_{c_K}$  is the sum of  $W3NRUNK\_A_l$  for all those eligible for Wave 3 in weighting class  $c_K$ , divided by the sum of  $W3NRUNK\_A_l$  for all Wave 3 respondents in that weighting class.

Finally, the Wave 3 nonresponse-adjusted weights for everyone other than Wave 3 nonrespondents were raked to control totals using Wave 1 characteristics with trimming also performed as part of an iterative process. The control totals and methods used were the same as those used to create the Wave 2 weights and are described in Section B.2.3; however, some slight modifications were made to the raking dimensions due to small sample sizes.

The final Wave 3 all-waves weight, denoted as  $W3RKWT\_A_l$ , was computed for each respondent  $l$  as shown in equation B.17:

$$W3RKWT\_A_l = W3NRWT\_A_l \times W3RT\_A_l \quad (\text{B.17})$$

where  $W3RT\_A_l$  is the combined raking and trimming adjustment for respondent  $l$ . The variable representing the final all-waves weight for each Wave 1 Cohort member responding to the Wave 3 adult interview (“A”) or youth interview (“Y”) is named R03\_x\_AWGT (where  $x$  is either “A” or

“Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 3 respondents participating at Wave 1 and Wave 2 will have a Wave 3 all-waves weight (as described at the beginning of Section B.3.1) and corresponding records on these weight files.

### B.3.2 Creation of the Wave 3 Single-Wave Weights

Wave 3 single-wave weights were assigned to all Wave 3 respondents regardless of their response status at Wave 2. To account for those Wave 1 Cohort members who became permanently ineligible (i.e., died or moved outside the United States) between Waves 1 and 2 separately from those who became ineligible between Waves 2 and 3, the Wave 3 single-wave weight was developed in a slightly different fashion from the all-waves weight.

The final Wave 1 weights served as the initial (“base”) weights for use in developing the Wave 3 single-wave weights. These initial weights were then adjusted to account for nonresponse to the Wave 3 interview as described in the sections below. The weighting process consisted of partitioning the sample into groups defined by Wave 1 age and forming nonresponse adjustment cells (Section B.3.2.1), and adjusting for nonresponse and raking to control totals to form the final Wave 3 single-wave weights (Section B.3.2.2).

#### B.3.2.1 Partitioning the Sample by Wave 1 Age and Forming Nonresponse Adjustment Cells

When developing the weighting adjustment cells, the same set of data is needed for both nonrespondents and respondents. For the Wave 3 single-wave weighting process, variables available for both Wave 3 respondents and nonrespondents are from Wave 1. As a result, Wave 1 data were used in the formation of weighting adjustment cells after partitioning the sample into eleven groups defined by Wave 1 age (as was done in the Wave 2 weighting process described in Section B.2.1).

As with the Wave 2 and Wave 3 all-waves weighting processes, the nonresponse adjustment was performed in two stages. The variables comprising the final weighting classes  $c_U$  (for the unknown eligibility adjustment) and  $c_K$  (for the nonresponse adjustment among those known to be eligible) were determined separately for the eleven age groups.

### B.3.2.2 Nonresponse and Raking Adjustments

Estimating the eligibility status of the nonrespondents whose eligibility at Wave 3 could not be ascertained (because, for example, they could not be located) is a standard part of the nonresponse adjustment process and is handled in a routine fashion for the all-waves weighting, as described in Section B.3.1.2. If the usual “routine” approach were to be used for the single-wave weighting, all persons known to be ineligible at Wave 3 would be used in determining the weighting adjustments that account for unknown eligibility. Doing so would overstate the portion of nonrespondents of unknown eligibility considered to be ineligible because it would apply an ineligibility rate accumulated across two waves of the study. This would not be a major concern in Wave 3 but could become one in a later wave. Therefore, Wave 1 Cohort members known to be permanently ineligible at Wave 2 were excluded from the Wave 3 single-wave nonresponse adjustment process.

Suppose a Wave 1 Cohort member  $l$  received final Wave 1 weight  $W1_l$  and was assigned to cell  $c_U$  for the Wave 3 single-wave weight adjustment to account for nonrespondents of unknown eligibility. Then the resulting Wave 3 single-wave weight after this adjustment, represented as  $W3NRUNK_S_l$ , was computed as shown in equation B.18:

$$W3NRUNK_S_l = W1_l \times W3N_S_{c_U} \quad (\text{B.18})$$

where  $W3N_S_{c_U}$  is the sum of  $W1_l$  for all Wave 1 Cohort members in weighting class  $c_U$  who were not permanently ineligible at Wave 2, divided by the sum of  $W1_l$  for Wave 1 Cohort members with known eligibility status at Wave 3 who were not permanently ineligible at Wave 2 in that weighting class.

This weight was then adjusted for the nonresponse of those known to be eligible for the Wave 3 interview. Let  $c_K$  represent the cell to which a Wave 1 Cohort member  $l$  known to be eligible for data collection at Wave 3 was assigned for nonresponse adjustment purposes. Then the Wave 3 single-wave nonresponse-adjusted weight for respondent  $l$ , represented as  $W3NRWT_S_l$ , was computed as shown in equation B.19:

$$W3NRWT_S_l = W3NRUNK_S_l \times W3N_S_{c_K} \quad (\text{B.19})$$

where  $W3N\_S_{c_K}$  is the sum of  $W3NRUNK\_S_l$  for all those eligible for Wave 3 in weighting class  $c_K$ , divided by the sum of  $W3NRUNK\_S_l$  for all Wave 3 respondents in that weighting class.

After completing the Wave 3 nonresponse adjustment process, the raking process (including ineligibles from Waves 2 and 3 not included in the nonresponse adjustment process) was the same as that employed for the Wave 3 all-waves weighting process described in Section B.3.1.2. The final raked weights of those who responded to a Wave 3 interview were then made available for the analysis of PATH Study data.

The variable representing the final single-wave weight for each Wave 1 Cohort member responding to the Wave 3 adult interview (“A”) or youth interview (“Y”) is named R03\_x\_SWGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2).

## B.4 Wave 4 Weights

There are two longitudinal weights available for analysis of Wave 4 data for the Wave 1 Cohort: the all-waves weight and the single-wave weight. The Wave 4 all-waves weight was assigned to Wave 4 respondents who also responded at all waves since recruitment. The Wave 4 single-wave weight was assigned to all Wave 4 respondents in the Wave 1 Cohort who completed an interview at Wave 1 whether or not they responded at Wave 2 or Wave 3.

In addition, there is a cross-sectional weight for all Wave 4 respondents in the Wave 4 Cohort. The Wave 4 cross-sectional weight was assigned to adult and youth respondents in the Wave 4 AYS replenishment sample and Wave 4 respondents from the Wave 1 Cohort who were still in the CNP at the time of Wave 4. Also created, but not included with the data files, are weights for the shadow youth respondents from the AYS and SO samples.

Sections B.4.1 and B.4.2, respectively, describe the creation of the Wave 4 all-waves and single-wave weights for the Wave 1 Cohort. Section B.4.3 describes the creation of the Wave 4 cross-sectional weights for the Wave 4 Cohort. Section B.4.4 describes the weights created for shadow youth respondents in preparation of their completing a youth interview in later waves.

### B.4.1 Creation of the Wave 4 All-Waves Weights for the Wave 1 Cohort

The all-waves weights were created for Wave 4 respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, and Wave 3. This includes Wave 4 respondents who either:

- Completed an interview (adult or youth) at Wave 1, Wave 2, and Wave 3; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1, were at least age 12 at the time of the Wave 4 interview, and completed an interview for all waves in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed.

The Wave 4 all-waves weighting process began with the “all-waves respondents” at Wave 3 and the nonresponse-adjusted weights assigned to them ( $W3NRWT_{A_l}$ ) during the Wave 3 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 4 interview. Wave 3 interview variables were used in the nonresponse adjustment process, except for shadow youth respondents at Wave 3 for whom the household screener variables were used.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 4, substituting “Wave 3” where “Wave 2” has been referenced and “Wave 4” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 3 age and forming nonresponse adjustment cells, and adjusting for nonresponse and raking to control totals to form the final Wave 4 all-waves weights. The last step of raking to control totals included some slight modifications to the raking dimensions used in previous waves due to small sample sizes.<sup>38</sup>

The variable representing the final all-waves weight for each Wave 1 Cohort member responding to the Wave 4 adult interview (“A”) or youth interview (“Y”) is named R04\_x\_A01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 4 respondents who also responded at Wave 1, Wave 2, and Wave 3 will have a Wave 4 all-waves weight (as described at the beginning of Section B.4.1) and corresponding records on these weight files.

---

<sup>38</sup>Because of the increased interest in e-cigarette use estimates, the raking dimensions including Wave 1 e-cigarette use were expanded to include 15- and 16-year-olds.

## B.4.2 Creation of the Wave 4 Single-Wave Weights for the Wave 1 Cohort

Wave 4 single-wave weights were assigned to all Wave 4 respondents in the Wave 1 Cohort who completed an interview at Wave 1 regardless of their response status at Wave 2 or Wave 3. These weights may be used for longitudinal analyses that use Wave 1 information to estimate an outcome at Wave 4. Because Wave 1 Cohort members selected at Wave 1 as shadow youth do not have any Wave 1 data, and so cannot be used in such an analysis, they were not assigned a Wave 4 single-wave weight. These participants were assigned a single-wave weight at Wave 3 because those weights could be used to create approximate cross-sectional estimates for that wave. However, genuine Wave 4 cross-sectional weights were created for the Wave 4 Cohort (as discussed in Section B.4.3), making such approximations unnecessary. See Chapter 6 for more information on selecting the appropriate weights for analysis.

Even though Wave 1 Cohort members selected as shadow youth at Wave 1 were not assigned a final Wave 4 single-wave weight for analysis, they were included in the weighting process so that those participants who are also in the Wave 4 Cohort would have an initial weight for the Wave 4 cross-sectional weighting process (see Section B.4.3). The final Wave 1 weights served as the initial (“base”) weights for use in developing the Wave 4 single-wave weights. These initial weights were then adjusted to account for nonresponse to the Wave 4 interview. PATH Study participants known to be permanently ineligible at Wave 3 (including those who became permanently ineligible at Wave 2) were excluded from the Wave 4 nonresponse adjustment process for the single-wave weight.

The Wave 3 single-wave weighting process described in Section B.3.2 is applicable to that for Wave 4, substituting “Wave 3” where “Wave 2” has been referenced and “Wave 4” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 1 age and forming nonresponse adjustment cells, and adjusting for nonresponse and raking to control totals to form the final Wave 4 single-wave weights. The last step of raking to control totals included some slight modifications to the raking dimensions used in previous waves due to small sample sizes.<sup>39</sup>

---

<sup>39</sup>Because of the increased interest in e-cigarette use estimates, the raking dimensions including Wave 1 e-cigarette use were expanded to include 15- and 16-year-olds.

The variable representing the final single-wave weight for each Wave 1 Cohort member completing an interview at Wave 1 and either the Wave 4 adult interview (“A”) or youth interview (“Y”) is named R04\_x\_S01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2).

### B.4.3 Creation of the Wave 4 Cross-Sectional Weights for the Wave 4 Cohort

The target population for the Wave 4 Cohort is the CNP at the time of Wave 4. Thus, the weights for the Wave 4 Cohort are cross-sectional at Wave 4. The approach for forming the cross-sectional weights is as follows:

1. Create weights for the Wave 4 AYS replenishment sample respondents.
2. Partition the set of Wave 4 AYS replenishment sample interview respondents into two groups: those who were members of the Wave 1 CNP and those who were not.
3. For those in the first group identified in step 2, rake their initial weights (from step 1) to population control totals, creating preliminary raked weights.
4. Identify those Wave 1 Cohort members who were both Wave 4 interview respondents and members of the Wave 4 CNP, and assign their Wave 4 single-wave weight after nonresponse adjustment as their initial weight.
5. Rake the Wave 1 Cohort initial weights (from step 4) to the same population control totals used in step 3, creating a second set of Wave 4 respondents with preliminary raked weights.
6. Develop compositing factors for the Wave 1 Cohort and AYS replenishment sample respondents who were members of the Wave 1 CNP (the first group identified in step 2).
7. Apply the compositing factors to the corresponding preliminary raked weights established in steps 3 and 5.
8. Combine (a) all those with composited weights with (b) the set of Wave 4 replenishment sample respondents not included in the compositing process (the second group identified in step 2), along with their initial weight. Together this set of respondents forms the Wave 4 Cohort.
9. Finally, rake and trim the Wave 4 Cohort weights from step 8.

Section B.4.3.1 describes step 1, Section B.4.3.2 describes steps 2-7, and Section B.4.3.3 describes steps 8-9.

### B.4.3.1 Weighting the Wave 4 AYS Replenishment Sample

The weighting process for the AYS replenishment sample at Wave 4 was similar to that used for the original sample at Wave 1, with the following two exceptions:

- First, adjustments made to sampling rates during the Wave 4 field period required the incorporation of two additional weighting factors for adults and one additional weighting factor for youth, compared to Wave 1. These adjustments resulted from the need to monitor and project combined yields for Wave 4 from both the replenishment sample and the Wave 1 Cohort. Further details can be found in the adult and youth weighting subsections below.
- Second, the purpose of the replenishment sample was to supplement the Wave 1 Cohort, not to support the development of national estimates based on the replenishment sample alone. As a result, person-level weights for the Wave 4 AYS replenishment sample interview respondents alone were not calibrated to population control totals. Calibration of person-level weights that included all members of the Wave 4 Cohort occurred only as a final step, as described in Section B.4.3.3.

The sections below describe the computation of the household weights as well as the adult and youth weights for the Wave 4 AYS replenishment sample.

#### ***Household Weights***

The process for creating household weights for the AYS replenishment sample was similar to that described in Section B.1.1 for Wave 1. Base weights were calculated for all sampled addresses as the inverse of the probability of selection. Adjustments were then made to the base weights to account for sampled addresses whose residential and occupancy status was unknown.

This eligibility adjustment was performed separately within weighting classes. The weighting classes were based on information available for all sampled addresses, including census 2010 and ACS 5-year (2009-2013) data pertaining to the segments, tracts, and blocks in which they were located. Census 2010 data were used to calculate the percent of occupied housing units that were owner-occupied, the percent of the population who were Black or African American, the percent of the population who were Asian, the percent of the population who were Hispanic, and the percent of the population ages 25 and older in the census block containing the address. The ACS 5-year data were used to estimate the median monthly housing unit costs in the census tract containing the

address. Census region, urbanicity of the PSU, urbanicity of the segment, an address type indicator,<sup>40</sup> and an indicator variable for PSUs that were affected by weather conditions<sup>41</sup> were also used when forming the weighting classes.

A second adjustment was made to account for nonresponse among addresses corresponding to eligible households. This nonresponse adjustment was performed separately within weighting classes that distinguished response propensities between subgroups of households, based on information available for both responding and nonresponding households. The same set of variables for the eligibility adjustment described above was used to develop the nonresponse adjustment cells. The nonresponse-adjusted weights were then raked to household counts from the 2016 ACS PUMS representing the four census regions.

### ***Adult Weights***

The number of adult respondents from the Wave 4 replenishment sample was monitored throughout the wave to ensure that targeted numbers of respondents would be met for various domains: age group (18-24, 25 and older), race (Black or African American, all other races), and tobacco use (yes, no). This monitoring took into account not only the original targeted numbers for the replenishment sample but also yields from the corresponding domains for the Wave 1 Cohort, the domains being replenished. If the yield for a domain from the Wave 1 Cohort was higher than expected, the sampling rate for the corresponding replenishment sample domain was adjusted to compensate.

At one point in the field period, the Wave 4 projected yields for one particular domain (non-Black adults who were at least age 25 and did not meet the tobacco-use oversampling criteria<sup>42</sup>) were high enough that no more adults were needed from the replenishment sample. Since two-phase sampling was used to select adults from the replenishment sample, there were Phase 1 domains and Phase 2 domains. Phase 1 domains were defined based on the household screener respondent's report of the adult's age, race, and tobacco-use status. Phase 2 domains were defined based on the adult's

---

<sup>40</sup>The address type indicator distinguished between the different sources of mailing addresses for sampling.

<sup>41</sup>The indicator variable identified clusters of PSUs where data collection efforts were affected by snow/ice storms, hurricanes, mudslides, wildfires, or flooding.

<sup>42</sup>All definitions of tobacco use relating to the domains discussed in this section are based on the “wide net” definition used for sampling adults at Wave 4.

self-report of these characteristics and whether or not the tobacco-use statuses reported by the household screener respondent and adult agreed. Partway through data collection, sampling rates were reduced to zero for the following: (1) Phase 1 domain “non-Black adults who were at least age 25 and did not meet the tobacco-use oversampling criteria,” (2) Phase 2 domain “non-Black adults who were at least age 25 and did not meet the tobacco-use oversampling criteria according to both the household screener respondent and self-report;” and (3) Phase 2 domain “non-Black adults who were at least age 25 and did not meet the tobacco-use oversampling criteria according to the household screener respondent but did based on self-report.” Two adjustment factors were constructed to account for adults in these domains who did not have a chance of selection for the replenishment sample.

The first adjustment factor accounted for those in the Phase 1 domain with no chance of sample selection after the rate changes were implemented. The second adjustment factor accounted for those in the Phase 2 domains with no chance of sample selection after the rate changes were implemented.<sup>43</sup> Application of these adjustment factors served the function of treating the samples that were selected from these domains as subsamples from the full sets of domain members identified over time.

The adjustment factor for adults in domain  $d_1$  based on the Phase 1 screener was computed as shown in equation B.20:

$$P1FACTOR_{d_1} = \frac{\text{number of adults (in responding households) in } d_1}{\text{number of adults in } d_1 \text{ with Phase 1 sampling rate} > 0}. \quad (\text{B.20})$$

For adults in the Phase 1 domain “non-Black adults who were at least age 25 and did not meet the tobacco-use oversampling criteria” the value of  $P1FACTOR_{d_1}$  was 1.469.  $P1FACTOR_{d_1}$  was 1 for adults in all other Phase 1 domains.

---

<sup>43</sup>No adults had a chance of being in the two Phase 2 domains with zero sampling rates if the household screener was completed after the corresponding Phase 1 domain sampling rate was set to zero. However, there was often a natural delay between the two phases of sampling, so for some adults the rate changes took place after they were selected from the Phase 1 domain and before completing the Phase 2 screener. The Phase 2 adjustment factor addresses this situation.

The adjustment factor for adults in Phase 2 domain  $d_2$  was computed as shown in equation B.21:

$$P2FACTOR_{d_2} = \frac{\text{number of adults (responding to Phase 2 screener) in } d_2}{\text{number of adults in } d_2 \text{ with Phase 2 sampling rate} > 0}. \quad (\text{B.21})$$

For adults in the Phase 2 domain “non-Black adults who were at least age 25 and did not meet the tobacco-use oversampling criteria according to both the household screener respondent and self-report,” the value of  $P2FACTOR_{d_2}$  was 1.147. For adults in the Phase 2 domain “non-Black adults who were at least age 25 and did not meet the tobacco-use oversampling criteria according to the household screener respondent but did based on self-report” the value of  $P2FACTOR_{d_2}$  was 1.270.  $P2FACTOR_{d_2}$  was 1 for adults in all other Phase 2 domains.

The raked household-level weight was used as the foundation for calculating the replenishment sample adult weight. The adult Phase 1 base weight, denoted as  $W4AP1BWT_{ijkl}$ , was computed as the product of the three quantities shown in equation B.4.3: the final household weight  $W4HRKWT_{ijk}$ ; the inverse of the within-household probability of selection for adult  $l$  within Wave 4 household  $k$  of PSU  $i$  and segment  $j$ ; and  $P1FACTOR_{d_1}$  (calculated as shown in equation B.22) for Phase 1 sampling domain  $d_1$  of which adult  $l$  is a member:

$$W4AP1BWT_{ijkl} = \left( W4HRKWT_{ijk} \times \frac{1}{\text{Probability adult } l \text{ selected at Phase 1 from household (ijk)}} \right) \times P1FACTOR_{d_1}. \quad (\text{B.22})$$

The Wave 4 adult Phase 1 base weights were further adjusted for nonresponse to the Phase 2 screener. Weighting classes were formed using the same information for forming the nonresponse adjustment cells at the household level, in addition to the person-level data collected in the Wave 4 household screener as described in Section B.1.1. The resulting adult weight (denoted as  $W4AP1NRWT_{ijkl}$ ), adjusted for nonresponse between Phases 1 and 2 of the adult sampling procedure, was calculated for respondents to the Phase 2 screener as shown in equation B.23:

$$W4AP1NRWT_{ijkl} = W4AP1BWT_{ijkl} \times W4AP1N_c \quad (\text{B.23})$$

where  $W4AP1N_c$  is the sum of  $W4AP1BWT_{ijkl}$  for adults sampled at Phase 1 in weighting class  $c$  of which adult  $l$  is a member, divided by the sum of  $W4AP1BWT_{ijkl}$  for all adults responding to the Phase 2 screener in that weighting class.

The Phase 2 weight, denoted as  $W4AP2WT_{ijkl}$ , was calculated as shown in equation B.24:

$$W4AP2WT_{ijkl} = W4AP1NRWT_{ijkl} \times \frac{1}{\text{Probability adult } l \text{ from household } (ijk) \text{ selected at Phase 2}} \times P2FACTOR_{d_2} \quad (\text{B.24})$$

where  $P2FACTOR_{d_2}$  is the Phase 2 probability adjustment factor for Phase 2 sampling domain  $d_2$  of which adult  $l$  is a member as shown in equation B.21.

### **Youth Weights**

The number of youth respondents from the Wave 4 replenishment sample was also monitored to ensure that the targeted number of respondents would be met. This monitoring took into account not only the original targeted number for the replenishment sample but also the yield from the Wave 1 Cohort. If the yield from the Wave 1 Cohort was higher than expected, the sampling rate for the replenishment sample was adjusted to compensate.

At one point in the field period, the Wave 4 projected yield for youth was high enough that youth were not needed from every household in the replenishment sample. Partway through data collection, a subsampling rate was applied at the household level only for the purpose of selecting youth. This subsampling decision was subsequently reversed, i.e., all replenishment sample households screened later in the field period were considered eligible for youth selection, when projected yield from the Wave 1 Cohort decreased based on updated information. An adjustment factor was incorporated into the replenishment sample youth weighting process to account for the subsampling that occurred.

Similar to the adult weighting, the raked household-level weight was used as the foundation for calculating the replenishment sample youth weight. The youth base weight, denoted as  $W4YBWT_{ijkl}$ , was computed as the product of the three quantities as shown in equation B.25: the final household weight  $W4HRKWT_{ijk}$ ; the inverse of the probability that youth were eligible for selection from Wave 4 household  $k$  of PSU  $l$  and segment  $j$ , denoted as  $W4PY_{ijk}$ ; and the inverse of the within-household probability of selection for youth  $l$  within household  $k$  of PSU  $i$  and segment  $j$ :

$$W4YBWT_{ijkl} = W4HRKWT_{ijk} \times \frac{1}{W4PY_{ijk}} \times \frac{1}{\text{Probability youth } l \text{ selected from household } (ijk)}. \quad (\text{B.25})$$

The youth base weights were adjusted to account for nonresponding youth. Weighting classes were formed using the same information for forming the nonresponse adjustment cells at the household level, in addition to person-level data collected during the household screener. The nonresponse-adjusted weight for responding youth is shown in equation B.26:

$$W4YNRWT_{ijkl} = W4YBWT_{ijkl} \times W4YN_c \quad (\text{B.26})$$

where  $W4YN_c$  is the nonresponse-adjustment factor calculated as the sum of  $W4YBWT_{ijkl}$  for all sampled youth in weighting class  $c$  of which youth  $l$  is a member, divided by the sum of  $W4YBWT_{ijkl}$  for all responding youth in that weighting class.

#### **B.4.3.2 Compositing Weights**

The weights of the adult and youth respondents to the Wave 4 replenishment sample were combined with the weights of the Wave 4 respondents from the Wave 1 Cohort who were members of the CNP at Wave 4. Most members of the Wave 1 Cohort are also members of the Wave 4 Cohort, and most youth and adults selected for the replenishment sample at Wave 4 were eligible for selection at Wave 1. To account for persons having multiple chances of selection for the PATH Study, the sample weights were appropriately composited to reflect the Wave 4 CNP.

Addresses sampled at the time of Wave 1 were not given the opportunity to be selected for the Wave 4 replenishment sample. In addition, for addresses identified as having a chance of selection at Wave 1, their Wave 4 probabilities of selection for the replenishment sample were the same as for Wave 4 addresses that did not have a chance of selection at the time of Wave 1.

Questions were asked of the Wave 4 replenishment sample interview respondents to help determine if they were members of the CNP at the time of Wave 1. If so, they had two chances of selection for the PATH Study, through either the Wave 1 sample or the Wave 4 replenishment sample. To account for this, the weights of the Wave 4 Cohort members with two chances of selection were composited as discussed below. All other members of the Wave 4 Cohort had only a single chance of selection for the study (through the replenishment sample). See Section B.4.3.3 for a description of the weights for the AYS replenishment sample interview respondents with a single chance of selection.

### ***Establishing Initial Weights for the Compositing Process***

For compositing, the weights used for the Wave 1 Cohort members who are also in the Wave 4 Cohort were the nonresponse-adjusted weights created as part of the single-wave weighting process described in Section B.4.2. The weights used for the replenishment sample respondents who were also in the Wave 1 CNP were the nonresponse-adjusted weights described in Section B.4.3.1. The two sets of weights were independently raked to the same population totals from the 2016 ACS PUMS. The raking was done using census region, age, race/ethnicity, sex, and educational attainment (for adults only).

The resulting preliminary raked weight, denoted as  $W4PCRWT_{ls}$ , was calculated for person  $l$  from source  $s$  (Wave 1 Cohort or Wave 4 replenishment sample) as shown in equation B.27:

$$W4PCRWT_{ls} = W4NRWT_{ls} \times W4PCR_{ls} \quad (\text{B.27})$$

where  $W4PCR_{ls}$  is the preliminary raking adjustment factor for person  $l$  from source  $s$ . For the Wave 1 Cohort,  $W4NRWT_{ls}$  is the nonresponse-adjusted weight created in the Wave 4 single-wave weighting process; for the Wave 4 Cohort,  $W4NRWT_{ls}$  is  $W4AP2WT_{ijkl}$  for adults and  $W4YNRWT_{ijkl}$  for youth.

### ***Creating and Applying Compositing Process***

The compositing factors were calculated in two steps. First, effective sample sizes were calculated using the preliminary raked weights for each source (Wave 1 Cohort, Wave 4 replenishment sample) by respondent characteristics or “compositing domains.” There were four compositing domains for adults, defined by the cross-classification of race (Black, non-Black) and age (ages 18-24, ages 25 and older), and six compositing domains for youth, corresponding to the single years of age from 12 to 17.<sup>44</sup>

Within a compositing domain  $m$ , the effective sample size for the Wave 1 Cohort, denoted as  $n\_eff_m^1$ , was calculated as the number of respondents (in the Wave 4 CNP) in domain  $m$  divided by

---

<sup>44</sup>The compositing domains for adults correspond to the dimensions of the adult sampling domains for which reliable population control totals were available for raking purposes.

the design effect associated with the variation of these respondents' preliminary raked weights, as shown in equation B.28:

$$n_{effm}^1 = \frac{n_m^1}{deff_m^1} \quad (B.28)$$

where  $n_m^1$  is the number of respondents (in the Wave 4 CNP) from the Wave 1 Cohort in domain  $m$ , and  $deff_m^1$  is the design effect associated with their preliminary raked weights, calculated as shown in equation B.29:

$$deff_m^1 = 1 + (cv_m^1)^2 \quad (B.29)$$

where  $cv_m^1$  is the coefficient of variation for the preliminary raked weights for the Wave 1 Cohort respondents (in the Wave 4 CNP) in domain  $m$ .

The effective sample size for a domain  $m$  of the Wave 4 replenishment sample, denoted as  $n_{effm}^4$ , was calculated using the same formulae described above, but based on the number of respondents from the replenishment sample (who were also in the Wave 1 CNP) in domain  $m$  and the design effect of their preliminary raked weights.

For a domain  $m$  of the Wave 1 Cohort, the compositing factor  $\alpha_m$  ( $0 < \alpha_m < 1$ ) was then calculated as shown in equation B.30:

$$\alpha_m = \frac{n_{effm}^1}{n_{effm}^1 + n_{effm}^4}. \quad (B.30)$$

For the Wave 4 replenishment sample, the compositing factor for domain  $m$  was calculated as  $(1 - \alpha_m)$ . The values of  $\alpha_m$  ranged from 0.720 to 0.895; the corresponding values of  $(1 - \alpha_m)$  ranged from 0.280 to 0.105. These compositing factors were applied to the preliminary raked weights for the Wave 1 Cohort and the Wave 4 replenishment sample, respectively.

The composited weight, denoted as  $W4CWT_{lsm}$ , was calculated for person  $l$  from source  $s$  (Wave 1 Cohort, Wave 4 replenishment sample) in domain  $m$  as the product of the preliminary raked weight and the compositing factor,  $\alpha_m$  or  $(1 - \alpha_m)$  for domain  $m$ , as shown in equation B.31:

$$W4CWT_{lsm} = \begin{cases} W4PCRKWT_{lsm} \times \alpha_m, & \text{if } s \text{ is Wave 1 cohort} \\ W4PCRKWT_{lsm} \times (1 - \alpha_m), & \text{if } s \text{ is replenishment sample} \end{cases} \quad (\text{B.31})$$

#### **B.4.3.3 Creating Final Wave 4 Weights for the Wave 4 Cohort**

The Wave 4 Cohort final weights were created by first making some weight adjustments for the Wave 4 respondents not included in the compositing process and then combining them with the Wave 4 respondents with composited weights before the final raking and trimming step.

##### ***Weights for Wave 4 Cohort Members Not Included in the Compositing Process***

The Wave 4 replenishment sample respondents who reported that they were not members of the Wave 1 CNP have a weight adjusted for nonresponse (as discussed in Section B.4.3.1) but were not part of the compositing effort (discussed in Section B.4.3.2) because they had only a single chance of being a member of the Wave 4 Cohort. With only a single chance of selection, the weights for this subgroup were generally larger than for other members of the Wave 4 Cohort whose weights were reduced by a compositing factor. There were two subgroups to which this set of respondents belonged: those in the military at Wave 1 and recent immigrants.

No issues with respect to the nonresponse-adjusted weights of those in the military at Wave 1 were identified. It was expected that recent immigrants would be concentrated in certain race/ethnicity groups. After examination of the distribution of the nonresponse-adjusted weights of recent immigrants by race/ethnicity, it appeared that they did not accurately represent recent immigrants at the time of Wave 4 in the population. So that these weights could more accurately represent this population, the weights were poststratified to corresponding 2016 ACS PUMS totals for persons who recently entered the United States. The poststratification was performed separately for youth and adults. The control totals for youth were based on race/ethnicity; those for adults were based on the cross-classification of age group and race/ethnicity. After this step, a few outlier weights remained.

In order to ensure that such outliers would not have a disproportionately large impact on PATH Study estimates, a particular concern at the subgroup level, these poststratified non-composited weights were then evaluated in comparison to the distribution of the composited weights for the remainder of the Wave 4 Cohort. This was done separately for youth and adult respondents. Any non-composited youth weights that exceeded the largest composited youth weight among the other members of the Wave 4 Cohort were trimmed to the size of this largest youth weight.

Corresponding trimming was performed for the poststratified non-composited adult weights. Fewer than 10 weights were trimmed in each group.

These cases (along with their poststratified and trimmed weights) were combined with the other Wave 4 Cohort adults and youth (along with their composited weights) before the last step of the weighting process, raking and trimming.

### ***Raking and Trimming***

The final raking and trimming step of the weighting process was conducted iteratively. First, the weights of all Wave 4 Cohort members were raked to independent population totals based on data from the 2016 ACS PUMS. For adults, the raking was done using cross-classifications of census region, age, race/ethnicity, sex, and educational attainment. For youth, the cross-classifications were based on census region, single year of age, race/ethnicity, and sex. These variables were imputed if missing, using methods similar to those used at Wave 1. (See Section 5.3 for more information about this imputation.)

After raking, a trimming step was performed to bring any extreme weights down to the median weight plus four times the interquartile range within groups, or trimming cells. Trimming cells were formed separately for shadow youth, youth, and adults. For each sample group, the cells were based on recruitment wave and for youth and adults selected at Wave 4, whether or not they were in the Wave 1 CNP. Table B-1 provides detailed definitions of the trimming cells.

**Table B-1.** Trimming cells used in creating the Wave 4 final weights for the Wave 4 Cohort

Recruitment wave	Sample group	Trimming cell definition
Wave 1	Shadow youth	Whether or not they were selected with certainty
	Youth	Whether or not they were selected with certainty
	Adults	Sampling domain defined by age group, tobacco usage, and race (see Section 2.1.2.4)
Wave 4	Shadow youth	Whether or not they were part of the sample in which only shadow youth were selected (i.e., SO or AYS sample)
	Youth not in Wave 1 CNP	Race/ethnicity and whether or not they were selected with certainty
	Adults not in Wave 1 CNP	Sampling domain defined by age group, tobacco usage, and race (see Section 2.1.2.4)
	Youth in the Wave 1 CNP	Whether or not they were selected with certainty
	Adults in the Wave 1 CNP	Age group, tobacco usage, and race (see Section 2.1.2.4)

After trimming, the sums of weights no longer matched the control totals, so raking and trimming were repeated until the resulting weights summed to the 2016 ACS PUMS totals for the raking dimensions and the weights were within the bounds defined by the interquartile range criterion.

After the iterative raking and trimming process, the final weight for person  $l$ , denoted as  $W4CRKWT_l$ , was calculated as shown in equation B.32:

$$W4CRKWT_l = W4CWT_l \times W4RT_l \quad (\text{B.32})$$

where  $W4RT_l$  is the combined raking and trimming adjustment for person  $l$ . The variable representing the final cross-sectional weight for a PATH Study participant in the Wave 4 Cohort responding to the Wave 4 adult interview (“A”) or youth interview (“Y”) is named R04\_x\_C04WGT. These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for variance estimation.

#### B.4.4 Creation of the Wave 4 Shadow Youth Weights

In preparation for the shadow youth completing a youth interview in later waves, Wave 4 weights were created for the shadow youth sample so that Wave 4 shadow youth would have a basis for the future-wave weights. These weights are not included in the data files.

The raked household weight was used as the foundation for calculating the shadow youth weight. As mentioned in Section 2.3.2, shadow youth were selected via the SO (shadow youth only) replenishment sample and the AYS (adults, youth, and shadow youth) replenishment sample. The

raked household weights were created for the SO replenishment sample similar to the AYS replenishment sample described in Section B.4.3.1. Each set of raked household weights (SO, AYS) summed to the estimated number of households from the 2016 ACS PUMS.

To create the Wave 4 shadow youth weights, the raked SO household weights and the raked AYS household weights were first combined. For the combined raked household weights to reflect the estimated number of households at the national level, adjustment factors were applied to the SO raked household weights and the AYS raked household weights. The final adjusted raked household weight was computed as the product of raked household weight and the inverse of the adjustment factor, where the adjustment factor was 1.512 for the SO replenishment sample and 2.952 for the AYS replenishment sample.

The shadow youth base weight was computed as the product of the final adjusted household weight and the inverse of the within-household probability of selection for shadow youth. Nonresponse adjustment was performed to account for nonresponse to the parent consent using a combination of census 2010 and ACS 5-year (2009-2013) data and person-level data collected during the household screener. Weighting classes were formed separately for shadow youth from the SO and AYS samples based on the following characteristics: census region; the percent of occupied housing units that were owner-occupied, the percent of the population who were Black or African American, the percent of the population who were Hispanic, the percent of the population who were Asian, and the percent of the population ages 25 and older in the census block containing the address; the median housing unit costs in the census tract containing the address; an urban segment indicator; the age and sex of the household screener respondent; and the number of enumerated members in the household.

Nonresponse-adjusted weights were raked to population totals from the 2016 ACS PUMS using census region, single year of age, race/ethnicity, and sex as raking variables. However, because no interview is conducted for shadow youth until they reach age 12, the only information available for this adjustment was from the household screener.

## B.5 Wave 4.5 Weights

At Wave 4.5, only youth ages 12 to 17 were interviewed. There are two longitudinal weights available for analysis of Wave 4.5 data for the Wave 1 Cohort: the all-waves weight and the single-wave weight. The Wave 4.5 all-waves weight was assigned to Wave 4.5 respondents who responded at Waves 1, 2, 3, and 4. The Wave 4.5 single-wave weight was assigned to all Wave 4.5 youth respondents in the Wave 1 Cohort who completed an interview at Wave 1 whether or not they responded at Wave 2, Wave 3, or Wave 4. In addition, there is a single-wave weight for all Wave 4.5 youth respondents in the Wave 4 Cohort.

The standard approach for computing single-wave and all-waves weights was employed at Wave 4.5. Beginning with an initial weight, the two steps involved:

- Adjusting the initial weight for nonresponse; and
- Raking the resulting nonresponse-adjusted weights to sample-based and population-based control totals with trimming incorporated as part of the raking process.

Sections B.5.1 and B.5.2 describe the creation of the Wave 4.5 all-waves and single-wave weights for the Wave 1 Cohort, respectively. Section B.5.3 describes the creation of the Wave 4.5 single-wave weights for the Wave 4 Cohort.

### B.5.1 Creation of the Wave 4.5 All-Waves Weights for the Wave 1 Cohort

The all-waves weights were created for Wave 4.5 youth respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, and Wave 4. This includes Wave 4.5 youth respondents who either:

- Completed a youth interview at Wave 1, Wave 2, Wave 3, and Wave 4; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1 and completed an interview for all waves in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed.

The Wave 4.5 all-waves weighting process began with the “all-wave respondents” at Wave 4 and the nonresponse-adjusted weights assigned to them during the Wave 4 all-waves weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 4.5 interview. Wave 4 interview variables were used in the nonresponse adjustment process, except for Wave 1

shadow youth who were responding shadow youth at Wave 4 where the household screener variables were used.

The Wave 3 all-waves weighting process described in Section B.3.1 applies to that for Wave 4.5, substituting “Wave 4” where “Wave 2” has been referenced and “Wave 4.5” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 4 age, forming nonresponse adjustment cells, computing the nonresponse-adjusted weight, and raking to control totals to form the final Wave 4.5 all-waves weight.

The variable representing the final all-waves weight for each Wave 1 Cohort member responding to the Wave 4.5 youth interview is named X04\_Y\_A01WGT. This variable is provided separately from the youth/parent data file on a weight file that also contains variables for use in variance estimation (see Section 5.2). Note that only those Wave 4.5 youth respondents who also responded at Wave 1, Wave 2, Wave 3, and Wave 4 will have a Wave 4.5 all-waves weight (as described at the beginning of this section) and corresponding records on this weight file.

## **B.5.2 Creation of the Wave 4.5 Single-Wave Weights for the Wave 1 Cohort**

Wave 4.5 single-wave weights were assigned to all Wave 4.5 youth respondents in the Wave 1 Cohort who completed an interview at Wave 1 regardless of their response status at Wave 2, Wave 3, or Wave 4. These weights may be used for longitudinal analyses that use Wave 1 information to estimate an outcome at Wave 4.5.

The final Wave 1 weight served as the initial weight for use in developing the Wave 4.5 single-wave weight. The initial weights were adjusted to account for nonresponse to the Wave 4.5 interview. PATH Study participants known to be permanently ineligible at Wave 4 (including those who became permanently ineligible at Wave 2 or Wave 3) were excluded from the Wave 4.5 nonresponse adjustment process.

The Wave 3 single-wave weighting process described in Section B.3.2 applies to that for Wave 4.5, substituting “Wave 4” where “Wave 2” has been referenced and “Wave 4.5” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning

the sample into groups defined by Wave 1 age, forming nonresponse adjustment cells, computing the nonresponse-adjusted weight, and raking to control totals to form the final Wave 4.5 single-wave weight.

The variable representing the final single-wave weight for each Wave 1 Cohort member completing an interview at Wave 1 and the Wave 4.5 youth interview is named X04\_Y\_S01WGT. This variable is provided separately from the youth/parent data file on a weight file that also contains variables for use in variance estimation (see Section 5.2).

### **B.5.3 Creation of the Wave 4.5 Single-Wave Weights for the Wave 4 Cohort**

Wave 4.5 single-wave weights were assigned to all Wave 4.5 youth respondents in the Wave 4 Cohort. The Wave 4 cross-sectional weight assigned to the Wave 4 Cohort (see Section B.4.3) served as the initial weight for use in developing the Wave 4.5 single-wave weight. The initial weights were adjusted to account for nonresponse to the Wave 4.5 interview. The Wave 2 weighting approach pertaining to youth respondents in Section B.2 is applicable to that for Wave 4.5, substituting “Wave 4” where “Wave 1” has been referenced and “Wave 4.5” where “Wave 2” has been referenced.

The nonresponse adjustment process involved the following steps: partitioning the sampled participants into groups defined by Wave 4 age (11, 12, 13, 14, 15, 16), forming nonresponse adjustment cells, adjusting the initial weights for nonresponse, and raking the nonresponse-adjusted weights to the same population-based control totals that were used at Wave 4 (census region, race/ethnicity, single year of age, and sex). The nonresponse-adjusted weights were also raked to sample-based control totals created from Wave 4 estimates of past 12-month tobacco use (for respondents who were at least age 13 at Wave 4) and Wave 4 estimates of past 12-month e-product use (for respondents who were at least age 14 at Wave 4).

After raking, a trimming step was performed to bring any extreme weights down to the median weight plus four times the interquartile range within groups, or trimming cells. For youth sampled as shadow youth at Wave 4, two trimming cells were formed based on whether or not they were part of the sample in which only shadow youth were selected. Youth selected at Wave 4 who were not in the Wave 1 CNP comprised a separate trimming cell. For all other youth, the trimming cells were

based on recruitment wave and the probability of selection. After trimming, the sums of weights no longer matched the control totals. Raking and trimming were repeated until the resulting weights summed to the control totals for the raking dimensions and the weights were within the bounds defined by the interquartile range criterion.

The variable representing the final single-wave weight for a PATH Study participant in the Wave 4 Cohort responding to the Wave 4.5 youth interview is named X04\_Y\_S04WGT. This variable is provided separately from the youth/parent data file on a weight file that also contains variables for variance estimation.

## B.6 Wave 5 Weights

There are three longitudinal weights available for analysis of Wave 5 data for the Wave 1 Cohort: two all-waves weights and a single-wave weight. The Wave 5 all-waves weight was assigned to Wave 5 respondents who also responded at Wave 1, Wave 2, Wave 3, and Wave 4. The Wave 1 Cohort special collection all-waves weight was assigned to all Wave 5 respondents who also responded at Waves 1, 2, 3, 4, and the special data collection wave, Wave 4.5. The Wave 5 single-wave weight was assigned to all Wave 5 respondents in the Wave 1 Cohort who completed an interview at Wave 1 whether or not they responded at Wave 2, Wave 3, or Wave 4.

For the Wave 4 Cohort, two longitudinal weights are available for analysis of Wave 5 data: a special collection all-waves weight and a single-wave weight. The special collection all-waves weight was assigned to all Wave 5 respondents in the Wave 4 Cohort who also responded at Wave 4 and the special data collection wave, Wave 4.5. (Note that youth not yet age 12 at the time of Wave 4.5 were not asked to verify their information with the study in the special collection wave, but were considered respondents at Wave 4.5 for the purpose of creating these Wave 5 special collection weights). In addition, there is a Wave 5 single-wave weight for all Wave 5 respondents in the Wave 4 Cohort.

As in previous waves, the standard approach for computing single-wave and all-waves weights was employed at Wave 5. Beginning with an initial weight, the two steps involved:

- Adjusting the initial weight for nonresponse; and
- Raking the resulting nonresponse-adjusted weights to sample-based and population-based control totals with trimming incorporated as part of the raking process.

Sections B.6.1, B.6.2, and B.6.3 describe the creation of the Wave 5 all-waves, special collection all-waves, and single-wave weights for the Wave 1 Cohort, respectively. Sections B.6.4 and B.6.5 describe the creation of the Wave 5 special collection all-waves and single-wave weights for the Wave 4 Cohort, respectively.

### B.6.1 Creation of the Wave 5 All-Waves Weights for the Wave 1 Cohort

The all-waves weights were created for Wave 5 respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, and Wave 4. This includes Wave 5 respondents who either:

- Completed an interview (adult or youth) at Wave 1, Wave 2, Wave 3, and Wave 4; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1, were at least age 14 at the time of the Wave 5 interview, and completed an interview for all waves in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed.

The Wave 5 all-waves weighting process began with the “all-waves respondents” at Wave 4 and the nonresponse-adjusted weights assigned to them during the Wave 4 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 5 interview. Wave 4 interview variables were used in the nonresponse adjustment process, except for shadow youth respondents at Wave 4 for whom the household screener variables were used.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 5, substituting “Wave 4” where “Wave 2” has been referenced and “Wave 5” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 4 age and forming nonresponse adjustment cells, and adjusting for nonresponse and raking to control totals to form the final Wave 5 all-waves weights.

The variable representing the final all-waves weight for each Wave 1 Cohort member responding to the Wave 5 adult interview (“A”) or youth interview (“Y”) is named R05\_x\_A01WGT (where  $x$  is

either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 5 respondents who also responded at Wave 1, Wave 2, Wave 3, and Wave 4 will have a Wave 5 all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

## B.6.2 Creation of the Wave 5 Special Collection All-Waves Weights for the Wave 1 Cohort

The special collection all-waves weights were created for Wave 5 respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 4.5. This includes Wave 5 respondents who either:

- Completed a youth interview at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 4.5; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1, were at least age 14 at the time of the Wave 5 interview, and completed an interview for all waves in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed.

The Wave 5 special collection all-waves weighting process began with the “all-waves respondents” at Wave 4.5 in the Wave 1 Cohort and the nonresponse-adjusted weights assigned to them during the Wave 4.5 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 5 interview. Wave 4.5 interview variables were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 5 special collection, substituting “Wave 4.5” where “Wave 2” has been referenced and “Wave 5” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 4.5 age and forming nonresponse adjustment cells, and adjusting for nonresponse and raking to control totals to form the final Wave 5 special collection all-waves weights.

The variable representing the final special collection all-waves weight for each Wave 1 Cohort member responding to the Wave 5 adult interview (“A”) or youth interview (“Y”) is named R05\_x\_AX01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided

separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 5 respondents who also responded at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 4.5 will have a Wave 5 special collection all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

### **B.6.3 Creation of the Wave 5 Single-Wave Weights for the Wave 1 Cohort**

Wave 5 single-wave weights were assigned to all Wave 5 respondents in the Wave 1 Cohort who completed an interview at Wave 1 regardless of their response status at Wave 2, Wave 3, Wave 4, or Wave 4.5. These weights may be used for longitudinal analyses that use Wave 1 information to estimate an outcome at Wave 5.

The final Wave 1 weights served as the initial (“base”) weights for use in developing the Wave 5 single-wave weights. These initial weights were then adjusted to account for nonresponse to the Wave 5 interview. PATH Study participants known to be permanently ineligible at Wave 4 (including those who became permanently ineligible at Wave 2 or Wave 3) were excluded from the Wave 5 nonresponse adjustment process for the single-wave weight.

The Wave 3 single-wave weighting process described in Section B.3.2 is applicable to that for Wave 5. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 1 age and forming nonresponse adjustment cells, and adjusting for nonresponse and raking to control totals to form the final Wave 5 single-wave weights.

The variable representing the final single-wave weight for each Wave 1 Cohort member completing an interview at Wave 1 and either the Wave 5 adult interview (“A”) or youth interview (“Y”) is named R05\_x\_S01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2).

## B.6.4 Creation of the Wave 5 Special Collection All-Waves Weights for the Wave 4 Cohort

The special collection all-waves weights were created for Wave 5 respondents in the Wave 4 Cohort who also responded at Wave 4 and Wave 4.5. This includes Wave 5 respondents who either:

- Completed a youth interview at Wave 4 and Wave 4.5; or
- Were in the shadow sample of persons ages 10 and 11 at Wave 4, were at least age 12 at the time of the Wave 5 interview, and completed a Wave 4.5 interview (if old enough to do so). By design, Wave 4 Cohort shadow youth who were age 10 at Wave 4 were not contacted for the Wave 4.5 special data collection. However, if they completed an interview at Wave 5, they were assigned this weight.

The Wave 5 special collection all-waves weighting process began with the “all-waves respondents” at Wave 4.5 in the Wave 4 Cohort and the nonresponse-adjusted weights assigned to them during the Wave 4.5 weighting process. For Wave 4 Cohort shadow youth who were age 10 at Wave 4, their raked weights created during the Wave 4 weighting process (see Section B.4.4) served as the base weights. These initial weights were then adjusted to account for nonresponse to the Wave 5 interview. For Wave 4.5 interview respondents, information used in the nonresponse adjustment process included variables from the Wave 4.5 interview; for youth too young to complete the Wave 4.5 interview, a smaller set of variables from the household screener, census 2010 and ACS 5-year (2009-2013) data was considered.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 5, substituting “Wave 4.5” where “Wave 2” has been referenced and “Wave 5” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 4.5 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking and trimming in an iterative process using the control totals and trimming cells described in Section B.5.3 to form the final Wave 5 special collection all-waves weights.

The variable representing the final special collection all-waves weight for each Wave 4 Cohort member responding to the Wave 5 adult interview (“A”) or youth interview (“Y”) is named R05\_x\_AX04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that Wave 5 respondents who responded at

Wave 4 and Wave 4.5 will have a Wave 5 special collection all-waves weight, additionally, those shadow youth who were age 10 at Wave 4 and responded at Wave 5 will also have this weight (as described at the beginning of this section) and corresponding records on these weight files.

### B.6.5 Creation of the Wave 5 Single-Wave Weights for the Wave 4 Cohort

Wave 5 single-wave weights were assigned to all Wave 5 respondents in the Wave 4 Cohort. The Wave 4 cross-sectional weight assigned to the Wave 4 Cohort (see Section B.4.3) served as the initial weight for use in developing the Wave 5 single-wave weight. The initial weights were adjusted to account for nonresponse to the Wave 5 interview. The nonresponse adjustment process involved partitioning the sample into groups defined by Wave 4 age (10, 11, 12, 13, 14, 15, 16, 17, 18-24, and 25+), forming nonresponse adjustment cells, and computing the nonresponse-adjusted weights.

The nonresponse-adjusted weights were raked to the same population-based control totals that were used for the Wave 4 Cohort at Wave 4 (census region, race/ethnicity, age group, education, and sex for adults at Wave 4; census region, race/ethnicity, single year of age, and sex for youth at Wave 4). The nonresponse-adjusted weights were also raked to sample-based control totals created from Wave 4 estimates of past 12-month tobacco use (for respondents who were at least age 13 at Wave 4) and Wave 4 estimates of past 12-month e-product use (for respondents who were at least age 14 at Wave 4).

After raking, a trimming step was performed to bring any extreme weights down to the median weight plus four times the interquartile range within groups, using the same trimming cells described in Table B-1. After trimming, the sums of weights no longer matched the control totals. Raking and trimming were repeated until the resulting weights summed to the control totals for the raking dimensions and the weights were within the bounds defined by the interquartile range criterion.

The variable representing the final single-wave weight for each Wave 4 Cohort member completing either the Wave 5 adult interview (“A”) or youth interview (“Y”) is named R05\_x\_S04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2).

## B.7 Wave 5.5 Weights

At Wave 5.5, only youth ages 13 to 17 and young adults ages 18 to 19 in the Wave 4 Cohort were interviewed. Even though data collection was focused on the Wave 4 Cohort, the Wave 1 Cohort longitudinal all-waves weight is available for analysis of Wave 5.5 data, given the amount of overlap between the two cohorts; this weight is assigned to Wave 5.5 respondents who also responded at Waves 1, 2, 3, 4, 4.5, and 5. There are two longitudinal weights available for analysis of Wave 5.5 data for the Wave 4 Cohort: an all-waves weight and a single-wave weight. The Wave 4 Cohort all-waves weight was assigned to Wave 5.5 respondents who also responded at Wave 4, Wave 4.5, and Wave 5. The Wave 4 Cohort single-wave weight was assigned to all Wave 5.5 respondents whether or not they responded at Wave 4.5 or Wave 5. These weights were assigned strictly to telephone interview respondents, representing 13- to 19-year-olds in the latter portion of 2020.

The standard approach for computing single-wave and all-waves weights was employed at Wave 5.5 for the “first-time” telephone respondents (i.e., those whose first Wave 5.5 interview was the telephone interview). Beginning with an initial weight, the two steps involved:

- Adjusting the initial weight for nonresponse; and
- Raking the resulting nonresponse-adjusted weights to sample-based and population-based control totals with trimming incorporated as part of the raking process.

For computing single-wave and all-waves weights for the “re-interviewed” respondents (i.e., Wave 5.5 in-person respondents re-interviewed by telephone), additional steps were implemented. Because only persons who responded to the in-person interview were approached for a re-interview, the weighting process for this group began with an adjustment for nonresponse to the in-person interview. After adjusting the initial weight for nonresponse to the Wave 5.5 in-person interview, these nonresponse-adjusted weights were adjusted for nonresponse to the telephone re-interview. Nonresponse-adjusted weights for the “re-interviewed” respondents and the Wave 5.5 “first-time” respondents were combined prior to raking.

Section B.7.1 describes the creation of the Wave 5.5 all-waves weights for the Wave 1 Cohort. Sections B.7.2 and B.7.3 describe the creation of the Wave 5.5 all-waves and single-wave weights for the Wave 4 Cohort, respectively.

## B.7.1 Creation of the Wave 5.5 All-Waves Weights for the Wave 1 Cohort

The all-waves weights were created for Wave 5.5 telephone interview respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5, and Wave 5. This includes Wave 5.5 respondents who either:

- Completed an interview (adult or youth) at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5, and Wave 5; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1, were at least age 15 at the time of the Wave 5.5 interview, and completed an interview for all waves in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed.

The Wave 5.5 all-waves weighting process began with the “all-waves respondents” at Wave 5 described in Section B.6.2 and the nonresponse-adjusted special collection weights assigned to them during the Wave 5 weighting process. The initial weights were adjusted to account for nonresponse to the telephone interview, separately for the “first-time” respondents and the “re-interviewed” respondents. For the “first-time” respondents, the initial weights were adjusted to account for nonresponse to the Wave 5.5 telephone interview.

As mentioned in the beginning of Section B.7, a two-step nonresponse adjustment process was implemented for the “re-interviewed” respondents. The initial weights were first adjusted to account for nonresponse to the Wave 5.5 in-person interview. These nonresponse-adjusted weights were further adjusted for nonresponse to the telephone re-interview.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 5.5, substituting “Wave 5” where “Wave 2” has been referenced and “Wave 5.5” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 5 age, forming nonresponse adjustment cells using prior wave information through Wave 5, and adjusting for nonresponse.

Nonresponse-adjusted weights for the “re-interviewed” respondents and the “first-time” respondents were combined with the nonresponse-adjusted PATH-ATS weights described in Section B.8.1 prior to the iterative raking and trimming procedure. The control totals used were the same as those used in creating the Wave 3. The trimming cells for the Wave 5.5 respondents were also the same as those used in the Wave 3 weighting process; the trimming cells used for the

PATH-ATS respondents are described in Section B.8.1. Raking the Wave 5.5 and PATH-ATS respondents together allows the two samples to be pooled for analyses representing the target population ages 15 and older, or some other combination of adults from the two samples, using the respective Wave 1 Cohort all-waves weights.

After raking and trimming, the Wave 5.5 respondent weight records were separated from the PATH-ATS records. The variable representing the final all-waves weight for each Wave 1 Cohort member responding to the Wave 5.5 adult interview (“A”) or youth interview (“Y”) is named X05\_x\_A01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 5.5 telephone interview respondents who also responded at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5, and Wave 5 will have this Wave 5.5 all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

## B.7.2 Creation of the Wave 5.5 All-Waves Weights for the Wave 4 Cohort

The all-waves weights were created for Wave 5.5 telephone interview respondents in the Wave 4 Cohort who also responded at Wave 4, Wave 4.5, and Wave 5. This includes Wave 5.5 respondents who either:

- Completed an interview (adult or youth) at Wave 4, Wave 4.5, and Wave 5; or
- Were in the shadow sample of persons ages 10 and 11 at Wave 4, were at least age 13 at the time of the Wave 5.5 interview, and completed an interview in Wave 4.5 (if old enough to do so) and in Wave 5.

The Wave 5.5 all-waves weighting process began with the “all-waves respondents” described in Section B.6.4 at Wave 5 in the Wave 4 Cohort and the nonresponse-adjusted special collection weights assigned to them during the Wave 5 weighting process. For the “first-time” respondents, these initial weights were adjusted to account for nonresponse to the Wave 5.5 telephone interview. For the “re-interviewed” respondents, the initial weights were first adjusted to account for nonresponse to the Wave 5.5 in-person interview; these nonresponse-adjusted weights were then adjusted to account for nonresponse to the Wave 5.5 telephone re-interview.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 5.5, substituting “Wave 5” where “Wave 2” has been referenced and “Wave 5.5” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 5 age, forming nonresponse adjustment cells using prior wave information through Wave 5, and adjusting for nonresponse.

Nonresponse-adjusted weights for the “re-interviewed” respondents and the “first-time” respondents were combined with the nonresponse-adjusted PATH-ATS weights described in Section B.8.2 prior to the iterative raking and trimming procedure. The control totals used were those described in Section B.6.5. The trimming cells for the Wave 5.5 respondents are described in Section B.5.3; the trimming cells used for the PATH-ATS respondents are described in Section B.8.2. Raking the Wave 5.5 and PATH-ATS respondents together allows the two samples to be pooled for analyses representing the target population ages 13 and older, or some other combination of adults from the two samples, using the respective Wave 4 Cohort all-waves weights.

After raking and trimming, the Wave 5.5 respondent weight records were separated from the PATH-ATS records. The variable representing the final all-waves weight for each Wave 4 Cohort member responding to the Wave 5.5 adult interview (“A”) or youth interview (“Y”) is named X05\_x\_AX04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 5.5 telephone interview respondents in the Wave 4 Cohort who responded at Wave 4, Wave 4.5, and Wave 5 will have this Wave 5.5 all-waves weight. In addition, those shadow youth who were age 10 at Wave 4 and responded at Wave 5 and Wave 5.5 will also have this weight (as described at the beginning of this section) and corresponding records on these weight files.

### B.7.3 Creation of the Wave 5.5 Single-Wave Weights for the Wave 4 Cohort

The Wave 4 Cohort single-wave weight was assigned to Wave 5.5 respondents whether or not they responded at Wave 4.5 or Wave 5. The Wave 4 cross-sectional weight (see Section B.4.3) served as the initial weight for developing the Wave 5.5 single-wave weight. The initial weights were adjusted to account for nonresponse to the telephone interview, separately for the “re-interviewed”

respondents and the “first-time” respondents. The weighting process for the “re-interviewed” respondents began with an adjustment for nonresponse to the in-person interview. These nonresponse-adjusted weights were then used as the initial weights for a second adjustment to account for nonresponse to the telephone re-interview. The nonresponse adjustment process involved partitioning the sample into groups defined by Wave 4 age, forming nonresponse adjustment cells, and computing the nonresponse-adjusted weights.

The nonresponse-adjusted weights for the “re-interviewed” and “first-time” respondents were combined and raked together using the iterative raking and trimming procedure, control totals and trimming cells described in Section B.5.3.

The variable representing the final single-wave weight for each Wave 4 Cohort member completing either the Wave 5.5 adult interview (“A”) or youth interview (“Y”) is named X05\_x\_S04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contains variables for use in variance estimation (see Section 5.2).

## B.8 PATH-ATS Weights

For the PATH-ATS, a subsample of adults ages 20 and older in the Wave 4 Cohort who responded to the Wave 5 adult interview were interviewed. All-waves weights are available for analysis of PATH-ATS data, representing adults ages 20 and older in the latter portion of 2020. The Wave 4 Cohort all-waves weight was assigned to all PATH-ATS respondents. Even though data collection was focused on the Wave 4 Cohort, the Wave 1 Cohort longitudinal all-waves weight is available for analysis of PATH-ATS data given the amount of overlap between the two cohorts; this weight is assigned to PATH-ATS respondents who also responded at Waves 1, 2, 3, 4, and 5.

Wave 5 nonresponse-adjusted weights were multiplied by the inverse of the PATH-ATS probability of selection to create the initial weights. Beginning with these initial weights, the standard all-waves weighting approach was applied:

- Adjusting the initial weight for nonresponse; and
- Raking the resulting nonresponse-adjusted weights to sample-based and population-based control totals with trimming incorporated as part of the raking process.

Nonresponse-adjusted weights for the PATH-ATS and the Wave 5.5 telephone interview respondents were combined prior to raking. Sections B.8.1 and B.8.2 describe the creation of the PATH-ATS all-waves weights for the Wave 1 Cohort and Wave 4 Cohort, respectively.

### B.8.1 Creation of the PATH-ATS All-Waves Weights for the Wave 1 Cohort

The all-waves weights were created for PATH-ATS respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 5.

The PATH-ATS all-waves weighting process began with the Wave 1 Cohort “all-waves respondents” at Wave 5 selected into the PATH-ATS sample and the nonresponse-adjusted weights assigned to them during the Wave 5 weighting process described in Section B.6.1. The initial weight was calculated as shown in equation B.33:

$$ATSIWT\_AC1_l = W5NRWT\_A1C_l \times \frac{1}{\text{Probability adult } l \text{ selected into the PATH-ATS sample}} \quad (\text{B.33})$$

where  $W5NRWT\_A1C_l$  is the Wave 5 nonresponse-adjusted all-waves weight assigned to adult  $l$  in the Wave 1 Cohort. The initial weights were then adjusted to account for nonresponse to the PATH-ATS interview.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for the PATH-ATS, substituting “Wave 5” where “Wave 2” has been referenced and “PATH-ATS” where “Wave 3” has been referenced. The nonresponse adjustment process involved partitioning the sample into groups defined by Wave 5 age, forming nonresponse adjustment cells using prior wave information through Wave 5, and adjusting for nonresponse.

Prior to the iterative raking and trimming process, the nonresponse-adjusted weights for PATH-ATS respondents were combined with the nonresponse-adjusted weights for the Wave 5.5 respondents as described in Section B.7.1. The control totals used for raking were the same as those used for Wave 3. The trimming cells for the PATH-ATS respondents were based on whether or not they were selected with certainty and the sampling domain defined by age (ages 20 to 24, ages 25 and older), tobacco product use (ENDS, cigarettes), and frequency of use (ever, past 12 months, past 30 days). Raking the PATH-ATS and Wave 5.5 respondents together allows the two samples to be

pooled for analyses representing the target population ages 15 and older, or some other combination of adults from the two samples, using the respective Wave 1 Cohort all-waves weights.

After raking and trimming, the PATH-ATS respondent weight records were separated from the Wave 5.5 records. The variable representing the final all-waves weight for each Wave 1 Cohort member responding to the PATH-ATS interview is named T05\_A\_A01WGT. This variable is provided on the weight file that also contains variables for use in variance estimation (see Section 5.2). Note that only those PATH-ATS respondents who also responded at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 5 will have a PATH-ATS all-waves weight and corresponding records on the weight file.

## B.8.2 Creation of the PATH-ATS All-Waves Weights for the Wave 4 Cohort

The all-waves weights were created for all PATH-ATS respondents. The PATH-ATS all-waves weighting process began with all participants selected into the PATH-ATS sample. The initial weight was calculated as shown in equation B.34:

$$ATSIWT_{AC4l} = W5NRWT_{S4C_l} \times \frac{1}{\text{Probability adult } l \text{ selected into the PATH-ATS sample}} \quad (\text{B.34})$$

where  $W5NRWT_{S4C_l}$  is the Wave 5 nonresponse-adjusted weight assigned to adult  $l$  in the Wave 4 Cohort as described in Section B.6.5.

The initial weights were adjusted to account for nonresponse to the PATH-ATS interview as described in Section B.8.1. Prior to the iterative raking and trimming process, the nonresponse-adjusted weights for PATH-ATS respondents were combined with the nonresponse-adjusted weights for the Wave 5.5 respondents as described in Section B.7.1. The control totals used were those described in Section B.6.5. The trimming cells for the PATH-ATS respondents were based on whether or not they were selected with certainty and the sampling domain defined by age (ages 20 to 24, ages 25 and older), tobacco product use (ENDS, cigarettes), and frequency of use (ever, past 12 months, past 30 days). Raking the PATH-ATS and Wave 5.5 respondents together allows the two samples to be pooled for analyses representing the target population ages 13 and older, or some other combination of adults from the two samples, using the respective Wave 4 Cohort all-waves weights.

After raking and trimming, the PATH-ATS respondent weight records were separated from the Wave 5.5 records. The variable representing the final all-waves weight for each Wave 4 Cohort member responding to the PATH-ATS is named T05\_A\_A04WGT. This variable is provided on the weight file that also contains variables for use in variance estimation (see Section 5.2).

## B.9 Wave 6 Weights

There are three longitudinal weights available for analysis of Wave 6 data for the Wave 1 Cohort: two all-waves weights and a single-wave weight. The Wave 6 all-waves weight was assigned to Wave 6 respondents who also responded at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 5. The Wave 1 Cohort special collection all-waves weight was assigned to all Wave 6 respondents who also responded at Waves 1, 2, 3, 4, 5 and the special data collection waves as appropriate: Wave 4.5 (for youth ages 12 to 17 at the Wave 4.5 interview) and either Wave 5.5 (for those at least 13 years old but under 20 on August 31, 2020) or PATH-ATS (for selected adults at least 20 years old on August 31, 2020). The Wave 6 single-wave weight was assigned to all Wave 6 respondents in the Wave 1 Cohort who completed an interview at Wave 1 whether or not they responded at Wave 2, Wave 3, Wave 4, Wave 5, or the special data collection waves.

Similarly for the Wave 4 Cohort, three longitudinal weights are available for analysis of Wave 6 data: two all-waves weights and a single-wave weight. The Wave 6 all-waves weight was assigned to Wave 6 respondents who also responded at Wave 4 and Wave 5. The special collection all-waves weight was assigned to all Wave 6 respondents in the Wave 4 Cohort who also responded at Wave 4 and Wave 5 and the special data collection waves Wave 4.5 as appropriate (see above) and either Wave 5.5 or PATH-ATS as appropriate (see above). In addition, a Wave 6 single-wave weight was assigned to all Wave 6 respondents in the Wave 4 Cohort whether or not they responded at Wave 4.5, Wave 5, or Wave 5.5/PATH-ATS.

As in previous waves, the standard approach for computing single-wave and all-waves weights was employed at Wave 6. Beginning with an initial weight, the two steps involved:

- Adjusting the initial weight for nonresponse; and
- Raking the resulting nonresponse-adjusted weights to sample-based and population-based control totals with trimming incorporated as part of the raking process.

Sections B.9.1, B.9.2, and B.9.3 describe the creation of the Wave 6 all-waves, special collection all-waves, and single-wave weights for the Wave 1 Cohort, respectively. Sections B.9.4, B.9.5, and B.9.6 describe the creation of the Wave 6 all-waves, special collection all-waves, and single-wave weights for the Wave 4 Cohort, respectively.

### B.9.1 Creation of the Wave 6 All-Waves Weights for the Wave 1 Cohort

The all-waves weights were created for Wave 6 respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 5. This includes Wave 6 respondents who either:

- Completed an interview (adult or youth) at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 5; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1 and completed an interview for all primary waves in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed.

The Wave 6 all-waves weighting process began with the “all-waves respondents” at Wave 5 and the nonresponse-adjusted weights assigned to them during the Wave 5 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 6 interview. Wave 5 interview variables were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 6, substituting “Wave 5” where “Wave 2” has been referenced and “Wave 6” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 5 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to control totals to form the final Wave 6 all-waves weights.

The variable representing the final all-waves weight for each Wave 1 Cohort member responding to the Wave 6 adult interview (“A”) or youth interview (“Y”) is named R06\_x\_A01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 6 respondents who also responded at Wave 1, Wave 2, Wave 3, Wave 4, and Wave 5 will have a Wave 6 all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

## B.9.2 Creation of the Wave 6 Special Collection All-Waves Weights for the Wave 1 Cohort

The special collection all-waves weights were created for Wave 6 respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5 as appropriate (see the beginning of Section B.9), Wave 5, and Wave 5.5/PATH-ATS as appropriate (see the beginning of Section B.9). This includes Wave 6 respondents who either:

- Completed an interview (adult or youth) at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5 as appropriate, Wave 5, and Wave 5.5/PATH-ATS as appropriate; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1 and completed an interview for all primary waves and special data collections in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed.

The Wave 6 special collection all-waves weighting process began with the “all-waves respondents” at Wave 5.5/PATH-ATS in the Wave 1 Cohort and the nonresponse-adjusted weights assigned to them during the Wave 5.5/PATH-ATS weighting process. The Wave 5.5 all-waves nonresponse-adjusted weights (see Section B.7.1) and the PATH-ATS all-waves nonresponse-adjusted weights (see Section B.8.1) were combined during the Wave 6 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 6 interview. Wave 5.5/PATH-ATS interview variables were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 6 special collection, substituting “Wave 5.5/PATH-ATS” where “Wave 2” has been referenced and “Wave 6” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 5.5/PATH-ATS indicator and Wave 5.5 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to control totals to form the final Wave 6 special collection all-waves weights.

The variable representing the final special collection all-waves weight for each Wave 1 Cohort member responding to the Wave 6 adult interview (“A”) or youth interview (“Y”) is named R06\_x\_AX01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 6 respondents who also responded at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5 as appropriate, Wave 5, and

Wave 5.5/PATH-ATS as appropriate will have a Wave 6 special collection all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

### **B.9.3 Creation of the Wave 6 Single-Wave Weights for the Wave 1 Cohort**

Wave 6 single-wave weights were assigned to all Wave 6 respondents in the Wave 1 Cohort who completed an interview at Wave 1 regardless of their response status at Wave 2, Wave 3, Wave 4, Wave 4.5, Wave 5, or Wave 5.5/PATH-ATS. These weights may be used for longitudinal analyses that use Wave 1 information to estimate an outcome at Wave 6.

The final Wave 1 weights served as the initial weights for use in developing the Wave 6 single-wave weights. These initial weights were then adjusted to account for nonresponse to the Wave 6 interview. PATH Study participants known to be permanently ineligible at Wave 5 (including those who became permanently ineligible at Wave 2, Wave 3, or Wave 4) were excluded from the Wave 6 nonresponse adjustment process for the single-wave weight.

The Wave 3 single-wave weighting process described in Section B.3.2 is applicable to that for Wave 6. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 1 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to control totals to form the final Wave 6 single-wave weights.

The variable representing the final single-wave weight for each Wave 1 Cohort member completing an interview at Wave 1 and either the Wave 6 adult interview (“A”) or youth interview (“Y”) is named R06\_x\_S01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2).

### **B.9.4 Creation of the Wave 6 All-Waves Weights for the Wave 4 Cohort**

The all-waves weights were created for Wave 6 respondents in the Wave 4 Cohort who also responded at Wave 4 and Wave 5. This includes Wave 6 respondents who either:

- Completed an interview (adult or youth) at Wave 4 and Wave 5; or
- Were in the shadow sample of persons ages 10 and 11 at Wave 4, were at least age 14 at the time of the Wave 6 interview, and completed an interview in Wave 5.

The Wave 6 all-waves weighting process began with the “all-waves respondents” at Wave 5 and the nonresponse-adjusted weights assigned to them during the Wave 5 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 6 interview. Wave 5 interview variables were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 6, substituting “Wave 5” where “Wave 2” has been referenced and “Wave 6” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 5 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to control totals to form the final Wave 6 all-waves weights.

The variable representing the final all-waves weight for each Wave 4 Cohort member responding to the Wave 6 adult interview (“A”) or youth interview (“Y”) is named R06\_x\_A04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 6 respondents who also responded at Wave 4 and Wave 5 will have a Wave 6 all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

## B.9.5 Creation of the Wave 6 Special Collection All-Waves Weights for the Wave 4 Cohort

The special collection all-waves weights were created for Wave 6 respondents in the Wave 4 Cohort who also responded at Wave 4, Wave 4.5 as appropriate (see the beginning of Section B.9), Wave 5, and Wave 5.5/PATH-ATS as appropriate (see the beginning of Section B.9). This includes Wave 6 respondents who either:

- Completed an interview (adult or youth) at Wave 4, Wave 4.5 as appropriate, Wave 5, and Wave 5.5/PATH-ATS as appropriate; or
- Were in the shadow sample of persons ages 10 and 11 at Wave 4, were at least age 14 at the time of the Wave 6 interview, and completed an interview for all waves in which they were old enough to do so or verified their information with the study for waves in which they were not old enough to be interviewed. By design, Wave 4 Cohort shadow youth who were age 10 at Wave 4 were not contacted for the Wave 4.5 special data collection. However, if they completed an interview at Wave 5 and Wave 5.5, they were assigned this weight.

The Wave 6 special collection all-waves weighting process began with the “all-waves respondents” at Wave 5.5/PATH-ATS in the Wave 4 Cohort and the nonresponse-adjusted weights assigned to them during the Wave 5.5/PATH-ATS weighting process. The Wave 5.5 all-waves nonresponse-adjusted weights (see Section B.7.2) and the PATH-ATS all-waves nonresponse-adjusted weights (see Section B.8.2) were combined during the Wave 6 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 6 interview. Wave 5.5/PATH-ATS interview variables were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 6 special collection, substituting “Wave 5.5/PATH-ATS” where “Wave 2” has been referenced and “Wave 6” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 5.5/PATH-ATS indicator and Wave 5.5 age, forming nonresponse adjustment cells, adjusting for nonresponse, and raking to control totals to form the final Wave 6 special collection all-waves weights.

The variable representing the final special collection all-waves weight for each Wave 4 Cohort member responding to the Wave 5 adult interview (“A”) or youth interview (“Y”) is named R06\_x\_AX04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that Wave 6 respondents who responded at Wave 4, Wave 4.5 as appropriate, Wave 5, and Wave 5.5/PATH-ATS as appropriate will have a Wave 6 special collection all-waves weight; additionally, those shadow youth who were age 10 at Wave 4 and responded at Wave 5 and Wave 5.5 will also have this weight (as described at the beginning of this section) and corresponding records on these weight files.

## B.9.6 Creation of the Wave 6 Single-Wave Weights for the Wave 4 Cohort

Wave 6 single-wave weights were assigned to all Wave 6 respondents in the Wave 4 Cohort regardless of their response status at Wave 4.5, Wave 5, or Wave 5.5/PATH-ATS. The Wave 4 cross-sectional weights assigned to the Wave 4 Cohort (see Section B.4.3) served as the initial weight for use in developing the Wave 6 single-wave weights. The initial weights were then adjusted to account for nonresponse to the Wave 6 interview. PATH Study participants known to be

permanently ineligible at Wave 5 were excluded from the Wave 6 nonresponse adjustment process for the single-wave weight.

The Wave 3 single-wave weighting process described in Section B.3.2 is applicable to that for Wave 6. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 4 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to control totals to form the final Wave 6 single-wave weights.

The variable representing the final single-wave weight for each Wave 4 Cohort member completing an interview at Wave 4 and either the Wave 6 adult interview (“A”) or youth interview (“Y”) is named R06\_x\_S04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2).

## B.10 Wave 7 Weights

There are three longitudinal weights available for analysis of Wave 7 data for the Wave 1 Cohort: two all-waves weights and a single-wave weight. The Wave 7 all-waves weight was assigned to Wave 7 respondents who also responded at all previous primary waves. The Wave 1 Cohort special collection all-waves weight was assigned to all Wave 7 respondents who also responded at all previous primary waves and the special data collection waves as appropriate: Wave 4.5 (for youth ages 12 to 17 at Wave 4.5) and either Wave 5.5 (for those at least 13 years old but under 20 on August 31, 2020) or PATH-ATS (for selected adults at least 20 years old on August 31, 2020). The Wave 7 single-wave weight was assigned to all Wave 7 respondents in the Wave 1 Cohort who completed an interview at Wave 1 whether they responded at interviewing waves or the special data collection waves.

Similarly for the Wave 4 Cohort, three longitudinal weights are available for analysis of Wave 7 data: two all-waves weights and a single-wave weight. The Wave 7 all-waves weight was assigned to Wave 7 respondents who also responded at Wave 4, Wave 5, and Wave 6. The special collection all-waves weight was assigned to all Wave 7 respondents in the Wave 4 Cohort who also responded at Wave 4, Wave 5, and Wave 6 and the special data collection waves Wave 4.5 as appropriate (see above) and either Wave 5.5 or PATH-ATS as appropriate (see above). In addition, a Wave 7 single-

wave weight was assigned to all Wave 7 respondents in the Wave 4 Cohort whether they responded at Wave 4.5, Wave 5, Wave 5.5/PATH-ATS, or Wave 6.

As in previous waves, the standard approach for computing single-wave and all-waves weights was employed at Wave 7. Beginning with an initial weight, the two steps involved:

- Adjusting the initial weight for nonresponse; and
- Raking the resulting nonresponse-adjusted weights to sample-based and population-based control totals with trimming incorporated as part of the raking process.

In addition, there is a cross-sectional weight for all Wave 7 respondents in the Wave 7 Cohort. The Wave 7 cross-sectional weight was assigned to adult and youth respondents in the Wave 7 replenishment sample and Wave 7 respondents ages 15 and older from the Wave 4 Cohort who were in the CNP at the time of Wave 7. Also created, but not included with the data files, are weights for the shadow youth respondents from the Wave 7 replenishment sample.

Sections B.10.1, B.10.2, and B.10.3 describe the creation of the Wave 7 all-waves, special collection all-waves, and single-wave weights for the Wave 1 Cohort, respectively. Sections B.10.4, B.10.5, and B.10.6 describe the creation of the Wave 7 all-waves, special collection all-waves, and single-wave weights for the Wave 4 Cohort, respectively. Section B.10.7 describes the creation of the Wave 7 cross-sectional weight for the Wave 7 Cohort. Section B.10.8 describes the weight created for shadow youth respondents in preparation of their completing a youth interview at later waves.

### **B.10.1 Creation of the Wave 7 All-Waves Weights for the Wave 1 Cohort**

The all-waves weights were created for Wave 7 respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, Wave 4, Wave 5, and Wave 6. This includes Wave 7 respondents who either:

- Completed an interview (adult or youth) at Wave 1, Wave 2, Wave 3, Wave 4, Wave 5, and Wave 6; or
- Were in the shadow sample of persons ages 9 to 11 at Wave 1 and completed an interview for all primary waves at which they were old enough to do so or verified their information with the study for waves at which they were not old enough to be interviewed.

The Wave 7 all-waves weighting process began with the “all-waves respondents” at Wave 6 and the nonresponse-adjusted weights assigned to them during the Wave 6 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 7 interview. Wave 6 interview variables were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 7, substituting “Wave 6” where “Wave 2” has been referenced and “Wave 7” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 6 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to control totals to form the final Wave 7 all-waves weights. The control totals described in Section B.2.3 were modified for youth participants at Wave 1. Raking dimensions involving tobacco use and e-cigarette use were expanded to include youth ages 12 to 15 and ages 12 to 14 at Wave 1, respectively. Two new raking dimensions defined by region, tobacco use and e-cigarette use, and age were created for youth ages 12 to 17.

The variable representing the final all-waves weight for each Wave 1 Cohort member responding to the Wave 7 adult interview (“A”) or youth interview (“Y”) is named R07\_x\_A01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 7 respondents who also responded at Wave 1, Wave 2, Wave 3, Wave 4, Wave 5, and Wave 6 will have a Wave 7 all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

## **B.10.2 Creation of the Wave 7 Special Collection All-Waves Weights for the Wave 1 Cohort**

The special collection all-waves weights were created for Wave 7 respondents in the Wave 1 Cohort who also responded at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5 as appropriate (see the beginning of Section B.10), Wave 5, Wave 5.5/PATH-ATS as appropriate (see the beginning of Section B.10), and Wave 6. This includes Wave 7 respondents who either:

- Completed an interview (adult or youth) at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5 as appropriate, Wave 5, Wave 5.5/PATH-ATS as appropriate, and Wave 6; or

- Were in the shadow sample of persons ages 9 to 11 at Wave 1 and completed an interview for all primary waves and special data collections at which they were old enough to do so or verified their information with the study for waves at which they were not old enough to be interviewed.

The Wave 7 special collection all-waves weighting process began with the “all-waves respondents” at Wave 6 in the Wave 1 Cohort and the nonresponse-adjusted weights assigned to them during the Wave 6 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 7 interview. Wave 6 interview variables were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 7 special collection all-waves weighting, substituting “Wave 6” where “Wave 2” has been referenced and “Wave 7” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 6 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to the revised control totals noted in Section B.10.1 to form the final Wave 7 special collection all-waves weights.

The variable representing the final special collection all-waves weight for each Wave 1 Cohort member responding to the Wave 7 adult interview (“A”) or youth interview (“Y”) is named R07\_x\_AX01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 7 respondents who also responded at Wave 1, Wave 2, Wave 3, Wave 4, Wave 4.5 as appropriate, Wave 5, Wave 5.5/PATH-ATS as appropriate, and Wave 6 will have a Wave 7 special collection all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

### B.10.3 Creation of the Wave 7 Single-Wave Weights for the Wave 1 Cohort

Wave 7 single-wave weights were assigned to all Wave 7 respondents in the Wave 1 Cohort who completed an interview at Wave 1 regardless of their response status at Wave 2, Wave 3, Wave 4, Wave 4.5, Wave 5, Wave 5.5/PATH-ATS, or Wave 6. These weights may be used for longitudinal analyses that use Wave 1 information to estimate an outcome at Wave 7.

The final Wave 1 weights served as the initial weights for use in developing the Wave 7 single-wave weights. These initial weights were then adjusted to account for nonresponse to the Wave 7

interview. PATH Study participants known to be permanently ineligible at Wave 6 (including those who became permanently ineligible at Wave 2, Wave 3, Wave 4, or Wave 5) were excluded from the Wave 7 nonresponse adjustment process for the single-wave weight.

The Wave 3 single-wave weighting process described in Section B.3.2 is applicable to that for Wave 7. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 1 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to the revised control totals noted in Section B.10.1 to form the final Wave 7 single-wave weights.

The variable representing the final single-wave weight for each Wave 1 Cohort member completing an interview at Wave 1 and either the Wave 7 adult interview (“A”) or youth interview (“Y”) is named R07\_x\_S01WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2).

#### B.10.4 Creation of the Wave 7 All-Waves Weights for the Wave 4 Cohort

The all-waves weights were created for Wave 7 respondents in the Wave 4 Cohort who also responded at Wave 4, Wave 5, and Wave 6. This includes Wave 7 respondents who either:

- Completed an interview (adult or youth) at Wave 4, Wave 5, and Wave 6; or
- Were in the shadow sample of persons ages 10 and 11 at Wave 4 and completed an interview at Wave 5 and Wave 6.

The Wave 7 all-waves weighting process began with the “all-waves respondents” at Wave 6 and the nonresponse-adjusted weights assigned to them during the Wave 6 weighting process. These initial weights were then adjusted to account for nonresponse to the Wave 7 interview. Wave 6 interview variables and recruitment wave were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 7, substituting “Wave 6” where “Wave 2” has been referenced and “Wave 7” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 6 age and forming nonresponse adjustment cells, adjusting for nonresponse, and raking to control totals to form the final Wave 7 all-waves weights. The

control totals described in Section B.5.3 were modified for youth participants at Wave 4. Sample-based control totals created from Wave 4 estimates of past 12-month tobacco use were expanded to include youth respondents who were age 12 at Wave 4. Similarly, youth respondents who were ages 12 and 13 at Wave 4 were included in the control totals created from Wave 4 estimates of past 12-month e-product use. Two new raking dimensions defined by region, past 12-month tobacco use and past 12-month e-product use, and age were created for youth ages 12 to 17.

The variable representing the final all-waves weight for each Wave 4 Cohort member responding to the Wave 7 adult interview (“A”) or youth interview (“Y”) is named R07\_x\_A04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that only those Wave 7 respondents who also responded at Wave 4, Wave 5, and Wave 6 will have a Wave 7 all-waves weight (as described at the beginning of this section) and corresponding records on these weight files.

### **B.10.5 Creation of the Wave 7 Special Collection All-Waves Weights for the Wave 4 Cohort**

The special collection all-waves weights were created for Wave 7 respondents in the Wave 4 Cohort who also responded at Wave 4, Wave 4.5 as appropriate (see the beginning of Section B.10), Wave 5, Wave 5.5/PATH-ATS as appropriate (see the beginning of Section B.10), and Wave 6. This includes Wave 7 respondents who either:

- Completed an interview (adult or youth) at Wave 4, Wave 4.5 as appropriate, Wave 5, Wave 5.5/PATH-ATS as appropriate, and Wave 6; or
- Were in the shadow sample of persons ages 10 and 11 at Wave 4, and completed an interview for all waves at which they were old enough to do so or verified their information with the study for waves at which they were not old enough to be interviewed. By design, Wave 4 Cohort shadow youth who were age 10 at Wave 4 were not contacted for the Wave 4.5 special data collection. However, if they completed an interview at Wave 5, Wave 5.5, and Wave 6, they were assigned this weight.

The Wave 7 special collection all-waves weighting process began with the “all-waves respondents” at Wave 6 in the Wave 4 Cohort and the nonresponse-adjusted weights assigned to them during the Wave 6 weighting process. These initial weights were then adjusted to account for nonresponse to

the Wave 7 interview. Wave 6 interview variables and recruitment wave were used in the nonresponse adjustment process.

The Wave 3 all-waves weighting process described in Section B.3.1 is applicable to that for Wave 7 special collection all-waves weighting, substituting “Wave 6” where “Wave 2” has been referenced and “Wave 7” where “Wave 3” has been referenced. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 6 age, forming nonresponse adjustment cells, adjusting for nonresponse, and raking to the revised control totals noted in Section B.10.4 to form the final Wave 7 special collection all-waves weights.

The variable representing the final special collection all-waves weight for each Wave 4 Cohort member responding to the Wave 7 adult interview (“A”) or youth interview (“Y”) is named R07\_x\_AX04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2). Note that Wave 7 respondents who responded at Wave 4, Wave 4.5 as appropriate, Wave 5, Wave 5.5/PATH-ATS as appropriate, and Wave 6 will have a Wave 7 special collection all-waves weight; additionally, those shadow youth who were age 10 at Wave 4 and responded at Wave 5, Wave 5.5, and Wave 6 will also have this weight (as described at the beginning of this section) and corresponding records on these weight files.

## B.10.6 Creation of the Wave 7 Single-Wave Weights for the Wave 4 Cohort

Wave 7 single-wave weights were assigned to all Wave 7 respondents in the Wave 4 Cohort regardless of their response status at Wave 4.5, Wave 5, Wave 5.5/PATH-ATS, or Wave 6. The Wave 4 cross-sectional weights assigned to the Wave 4 Cohort (see Section B.4.3) served as the initial weights for use in developing the Wave 7 single-wave weights. The initial weights were then adjusted to account for nonresponse to the Wave 7 interview. PATH Study participants known to be permanently ineligible at Wave 6 were excluded from the Wave 7 nonresponse adjustment process for the single-wave weight.

The Wave 3 single-wave weighting process described in Section B.3.2 is applicable to that for Wave 7. The nonresponse adjustment process consisted of the following steps: partitioning the sample into groups defined by Wave 4 age and forming nonresponse adjustment cells, adjusting for

nonresponse, and raking to the revised control totals noted in Section B.10.4 to form the final Wave 7 single-wave weights.

The variable representing the final single-wave weight for each Wave 4 Cohort member completing an interview at Wave 4 and either the Wave 7 adult interview (“A”) or youth interview (“Y”) is named R07\_x\_S04WGT (where  $x$  is either “A” or “Y” as appropriate). These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for use in variance estimation (see Section 5.2).

### **B.10.7 Creation of the Wave 7 Cross-Sectional Adult and Youth Weights for the Wave 7 Cohort**

The target population for the Wave 7 Cohort is the CNP who were ages 9 and older at the time of Wave 7. Thus, the weights for the Wave 7 Cohort are cross-sectional at Wave 7. As with Wave 4, the Wave 7 replenishment sample alone does not support the development of national estimates. Instead, the Wave 7 Cohort is formed by combining the Wave 7 replenishment sample respondents and the Wave 7 respondents ages 15 and older from the Wave 4 Cohort who were in the Wave 7 CNP. Wave 4 Cohort respondents at Wave 7 under the age of 15 were excluded from the Wave 7 Cohort because they were few in number.

The approach for creating the cross-sectional adult and youth weights is as follows:

1. Create nonresponse adjusted weights for the adult and youth replenishment sample respondents.
2. Rake the nonresponse-adjusted weights (from step 1) to population control totals. This creates the first set of Wave 7 respondents with preliminary raked weights.
3. Identify those Wave 4 Cohort members who were both Wave 7 interview respondents ages 15 and older and members of the Wave 7 CNP, and rake their Wave 7 single-wave nonresponse-adjusted weights to the same population control totals used in step 2. This creates the second set of Wave 7 respondents with preliminary raked weights.
4. Partition the respondents from step 2 into two groups: (a) those who were members of the Wave 4 CNP and (b) those who were not in the Wave 4 CNP.
5. Apply compositing factors to the weights established in step 3 and the weights for group (a) in step 4.

6. Combine all those with composited weights from step 5 with group (b) from step 4 (i.e., the set of Wave 7 replenishment sample respondents not included in the compositing process). Together this set of respondents forms the Wave 7 Cohort.
7. Rake and trim the Wave 7 Cohort weights from step 6.

Section B.10.7.1 describes step 1, Section B.10.7.2 describes steps 2 through 5, and Section B.10.7.3 describes steps 6 and 7.

### **B.10.7.1 Creating Nonresponse Adjusted Adult and Youth Replenishment Sample Weights**

The sections below describe the creation of the household weights as well as the adult and youth nonresponse adjusted weights for the Wave 7 replenishment sample. The weights were created using an approach similar to that used for the Wave 4 AYS sample except that some additional variables from the PATH Study's address vendor were used to form nonresponse adjustment cells with the goal to further reduce potential nonresponse bias.

#### ***Household Weights***

The process for creating household weights was similar to that described in Section B.1.1 for Wave 1. Base weights were calculated for all sampled addresses as the inverse of the probability of selection. Adjustments were then made to the base weights to account for (1) sampled addresses whose residential and occupancy status was unknown, and (2) nonresponse among addresses corresponding to residential households. These adjustments were performed separately within weighting classes based on information available for all sampled addresses, including both appended variables on the address sampling frame as well as census 2020 and ACS 5-year (2016-2021) data pertaining to the segments, tracts, and blocks in which they were located. In particular, census 2020 data were used to calculate the percent of occupied housing units that were owner-occupied, the percent of the population who were Black or African American, the percent of the population who were Asian, the percent of the population who were Hispanic, and the percent of the population ages 25 and older in the census block containing the address. The ACS 5-year data were used to estimate the median monthly housing unit costs in the census tract containing the address. Census

region, urbanicity of the block, an address type indicator,<sup>45</sup> and a set of appended variables<sup>46</sup> from the address vendor were also used when forming the weighting classes. The nonresponse-adjusted household weights were then raked to household counts from the 2021 ACS PUMS by census region and presence of children (i.e., persons younger than age 18).

### ***Adult Weights***

The number of adult respondents from the Wave 7 replenishment sample was monitored throughout the data collection period. Due to the post-pandemic environment, both household and adult response rates were lower than expected. Therefore, the within-household sampling rates for adults were adjusted during the data collection to bring the numbers of respondents, both overall and by domain (defined by crossing age group (18-24, 25 and older), race (Black or African American, all other races), and tobacco use (yes, no)), closer to the targets. This adjustment was conducted by “release group” (i.e., the random subsamples of Wave 7 sampled addresses released periodically during the data collection period) so that the within-household selection probabilities could be averaged across the full sample. There were three release groups for the Wave 7 replenishment sample. The within-household sampling rates for the third release group were set higher than those for the previous two groups. After data collection, the empirical within-household selection probabilities were computed separately for Phase 1 and Phase 2 and within cells formed to account for two factors: (1) the domain in which the sampled adult was a member; and (2) the sum of within-household selection probabilities of the enumerated adults in the household containing the sampled adult. Within each cell, the empirical within-household selection probability for a sampled person was the ratio of the number of sampled persons to the number of enumerated persons. This approach reduced the variation of the adult base weights.

The raked household-level weight was used as the foundation for calculating the replenishment sample adult weight. The adult Phase 1 base weight, denoted as  $W7AP1BWT_{ijkl}$ , was computed as

<sup>45</sup>The address type indicator distinguished between the different sources of mailing addresses for sampling.

<sup>46</sup>These variables reflect the characteristics of the sampled addresses, including age, sex, ethnicity, education, and marital status of the head of the household, presence of children, presence of a person age 18 to 24, presence of a person age 25 to 35, presence of a person age 35 to 64, presence of a person age 65 or older, whether the head of the household had an Asian surname, whether the head of the household had a Hispanic surname, household income, number of adults in the household, and tenure status. These appended variables came from various sources with varying accuracy and completeness.

the final household weight  $W7HRKWT_{ijk}$  divided by the empirical within-household Phase 1 probability of selection for adult  $l$  within Wave 7 household  $k$  of PSU  $i$  and segment  $j$ :

$$W7AP1BWT_{ijkl} = \frac{W7HRKWT_{ijk}}{\text{Empirical probability adult } l \text{ selected at Phase 1 from household } (ijk)}. \quad (\text{B.35})$$

The Wave 7 adult Phase 1 base weight was further adjusted for nonresponse to the Phase 2 screener. Weighting classes were formed using the same information for forming the nonresponse adjustment cells at the household level, in addition to the person-level data collected in the Wave 7 household screener as described in Section B.1.1. The resulting adult weight (denoted as  $W7AP1NRWT_{ijkl}$ ), adjusted for nonresponse between Phases 1 and 2 of the adult sampling procedure, was calculated for respondents to the Phase 2 screener as shown in equation B.36:

$$W7AP1NRWT_{ijkl} = W7AP1BWT_{ijkl} \times W7AP1N_c \quad (\text{B.36})$$

where  $W7AP1N_c$  is the sum of  $W7AP1BWT_{ijkl}$  for adults sampled at Phase 1 in weighting class  $c$  of which adult  $l$  is a member, divided by the sum of  $W7AP1BWT_{ijkl}$  for all adults responding to the Phase 2 screener in that weighting class.

The Phase 2 weight, denoted as  $W7AP2WT_{ijkl}$ , was calculated as shown in equation B.37:

$$W7AP2WT_{ijkl} = \frac{W7AP1NRWT_{ijkl}}{\text{Empirical probability adult } l \text{ from household } (ijk) \text{ selected at Phase 2}}. \quad (\text{B.37})$$

### **Youth Weights**

Similar to the adult weighting, the raked household-level weight was used as the foundation for calculating the replenishment sample youth weight. The youth base weight, denoted as  $W7YBWT_{ijkl}$ , was computed as the final household weight  $W7HRKWT_{ijk}$  divided by the within-household probability of selection for youth  $l$  within household  $k$  of PSU  $i$  and segment  $j$ , as shown in equation B.38:

$$W7YBWT_{ijkl} = \frac{W7HRKWT_{ijk}}{\text{Probability youth } l \text{ selected from household } (ijk)}. \quad (\text{B.38})$$

The youth base weights were adjusted to account for nonresponding youth. Weighting classes were formed using the same information for forming the nonresponse adjustment cells at the household

level, in addition to person-level data collected during the household screener. The nonresponse-adjusted weight for responding youth is shown in equation B.39:

$$W7YNRWT_{ijkl} = W7YBWT_{ijkl} \times W7YN_c \quad (\text{B.39})$$

where  $W7YN_c$  is the nonresponse-adjustment factor calculated as the sum of  $W7YBWT_{ijkl}$  for all sampled youth in weighting class  $c$  of which youth  $l$  is a member, divided by the sum of  $W7YBWT_{ijkl}$  for all responding youth in that weighting class.

### B.10.7.2 Compositing Weights

The weights of the adult and youth Wave 7 replenishment sample respondents were combined with the weights of the Wave 7 respondents from the Wave 4 Cohort who were ages 15 and older and members of the CNP at Wave 7. Most members of the Wave 4 Cohort are also members of the Wave 7 Cohort, and most youth and adults selected for the replenishment sample at Wave 7 were eligible for selection at Wave 4. To account for persons having multiple chances of selection, the sample weights were appropriately composited to reflect the Wave 7 CNP.

Addresses sampled at the time of Wave 4 were not given the opportunity to be selected for the Wave 7 replenishment sample. In addition, addresses identified as having a chance of selection at Wave 4 had the same Wave 7 probabilities of selection for the replenishment sample as Wave 7 addresses that did not have a chance of selection at the time of Wave 4.

Questions were asked of the Wave 7 replenishment sample interview respondents to help determine if they were members of the CNP at the time of Wave 4. If so, they had two chances of selection for the Wave 7 Cohort, through either the Wave 4 Cohort or the Wave 7 replenishment sample. To account for this, the weights of the Wave 7 Cohort members with two chances of selection were composited as discussed below. All other members of the Wave 7 Cohort had only a single chance of selection (through the Wave 7 replenishment sample). See Section B.10.7.3 for a description of the weights for the adult and youth interview respondents with a single chance of selection.

### ***Establishing Initial Weights for the Compositing Process***

Two sets of respondents form the Wave 7 Cohort: (1) Wave 4 Cohort members who were ages 15 and older and in the Wave 7 CNP; and (2) the Wave 7 replenishment sample respondents. Before compositing, the nonresponse-adjusted weights for these two sets of respondents were independently raked to the same population totals from the 2021 ACS PUMS. The raking was done using census region, age, race/ethnicity, sex, and (for adults only) educational attainment.

The resulting preliminary raked weight, denoted as  $W7PCRKWT_{ls}$ , was calculated for person  $l$  from source  $s$  (Wave 4 Cohort or Wave 7 replenishment sample) as shown in equation B.40:

$$W7PCRKWT_{ls} = W7NRWT_{ls} \times W7PCR_{ls} \quad (\text{B.40})$$

where  $W7PCR_{ls}$  is the preliminary raking adjustment factor for person  $l$  from source  $s$ . For the Wave 4 Cohort,  $W7NRWT_{ls}$  is the nonresponse-adjusted weight created in the Wave 7 single-wave weighting process; for the Wave 7 replenishment sample,  $W7NRWT_{ls}$  is  $W7AP2WT_{ijkl}$  for adults and  $W7YNRWT_{ijkl}$  for youth.

### ***Creating and Applying Compositing Process***

The compositing process was applied to two sets of respondents ages 15 or older: (1) Wave 4 Cohort members who were also in the Wave 7 CNP; and (2) Wave 7 replenishment sample members who were also in the Wave 4 CNP. The compositing factors were calculated in two steps. First, effective sample sizes were calculated using the preliminary raked weights for each source (Wave 4 Cohort, Wave 7 replenishment sample) by respondent characteristics or “compositing domains.” There were four compositing domains for adults, defined by the cross-classification of race (Black or African American adults, all other races) and age group (18-24, 25 and older), and three compositing domains for youth, corresponding to the single years of age from 15 to 17.<sup>47</sup>

Within a compositing domain  $m$ , the effective sample size for the Wave 4 Cohort, denoted as  $n\_eff_m^4$ , was calculated as the number of respondents (in the Wave 7 CNP) in domain  $m$  divided by

---

<sup>47</sup>The compositing domains for adults correspond to the dimensions of the adult sampling domains for which reliable population control totals were available for raking purposes.

the design effect associated with the variation of these respondents' preliminary raked weights, as shown in equation B.41:

$$n_{effm}^4 = \frac{n_m^4}{deff_m^4} \quad (\text{B.41})$$

where  $n_m^4$  is the number of respondents (in the Wave 7 CNP) from the Wave 4 Cohort in domain  $m$ , and  $deff_m^4$  is the design effect associated with their preliminary raked weights, calculated as shown in equation B.42:

$$deff_m^4 = 1 + (cv_m^4)^2 \quad (\text{B.42})$$

where  $cv_m^4$  is the coefficient of variation for the preliminary raked weights for the Wave 4 Cohort respondents (in the Wave 7 CNP) in domain  $m$ .

The effective sample size for a domain  $m$  of the Wave 7 replenishment sample, denoted as  $n_{effm}^7$ , was calculated using the same formulae described above, but based on the number of respondents from the replenishment sample (who were also in the Wave 4 CNP) in domain  $m$  and the design effect of their preliminary raked weights.

For a domain  $m$  of the Wave 4 Cohort, the compositing factor  $\alpha_m$  ( $0 < \alpha_m < 1$ ) was then calculated as shown in equation B.43:

$$\alpha_m = \frac{n_{effm}^4}{n_{effm}^4 + n_{effm}^7}. \quad (\text{B.43})$$

For the Wave 7 replenishment sample, the compositing factor for domain  $m$  was calculated as  $(1 - \alpha_m)$ . For adults, the values of  $\alpha_m$  ranged from 0.901 to 0.947; the corresponding values of  $(1 - \alpha_m)$  ranged from 0.053 to 0.099. For youth ages 15 to 17, the values of  $\alpha_m$  ranged from 0.527 to 0.630; the corresponding values of  $(1 - \alpha_m)$  ranged from 0.370 to 0.473. These compositing factors were applied to the preliminary raked weights for the Wave 4 Cohort and the Wave 7 replenishment sample, respectively.

The composited weight, denoted as  $W7CWT_{lsm}$ , was calculated for person  $l$  from source  $s$  (Wave 4 Cohort, Wave 7 replenishment sample) in domain  $m$  as the product of the preliminary raked weight and the compositing factor,  $\alpha_m$  or  $(1 - \alpha_m)$  for domain  $m$ , as shown in equation B.44:

$$W7CWT_{lsm} = \begin{cases} W7PCRKWT_{lsm} \times \alpha_m, & \text{if } s \text{ is Wave 4 Cohort} \\ W7PCRKWT_{lsm} \times (1 - \alpha_m), & \text{if } s \text{ is Wave 7 replenishment sample} \end{cases} \quad (\text{B.44})$$

The Wave 7 respondents from the Wave 4 Cohort who were under the age of 15 (at Wave 7) were excluded from the Wave 7 Cohort because they were few in number. If included in the compositing process described above and the compositing domains for youth expanded to those under 15, their compositing factors and resulting weights would be too small for their data to meaningfully contribute to study estimates.

### B.10.7.3 Creating Final Wave 7 Weights for the Wave 7 Cohort

The Wave 7 Cohort final weights were created by first combining the Wave 7 respondents that were not included in the compositing process and the Wave 7 respondents with composited weights, and then conducting final raking and trimming.

#### ***Weights for Wave 7 Cohort Members Not Included in the Compositing Process***

The Wave 7 replenishment sample respondents who either reported that they were not members of the Wave 4 CNP or were ages 12 to 14 have a preliminary raked weight but were not part of the compositing effort (discussed in Section B.10.5.2) because they had only a single chance of being a member of the Wave 7 Cohort. There were three subgroups to which this set of respondents belonged: (1) those in the military at Wave 4; (2) recent immigrants; and (3) those ages 12 to 14. With only a single change of selection, the weights for those in groups (1) and (2) were generally larger than for those 15 and older whose weights were reduced by a compositing factor. All these participants were combined with the other Wave 7 Cohort adults and youth (along with their composited weights) before the last step of the weighting process, raking and trimming.

### ***Raking and Trimming***

The final raking and trimming step of the weighting process was conducted iteratively. First, the weights of all Wave 7 Cohort members were raked to independent population totals based on data from the 2021 ACS PUMS. For adults, the raking was done using cross-classifications of census region, age, race/ethnicity, sex, and educational attainment. For youth, the cross-classifications were based on census region, single year of age, race/ethnicity, and sex. These variables were imputed if missing, using methods similar to those used at Wave 1. (See Section 5.3 for more information about this imputation.)

After raking, a trimming step was performed to bring any extreme weights down to the median weight plus six times (for adults) or five times (for youth) the interquartile range within groups, or trimming cells. Trimming cells were formed separately for youth and adults. These trimming cells were different from those used for the Wave 4 Cohort. For the Wave 7 Cohort cross-sectional weighting, the trimming cells were formed only based on the key domains of interest (defined by race/ethnicity, age, sex, and tobacco use), which was necessary to eliminate some outlier weights. Therefore, the new trimming approach successfully reduced the variation in weights without introducing any substantial potential bias in the overall estimates or the estimates for the key domains of interest.

After trimming, the sums of weights no longer matched the control totals, so raking and trimming were repeated until the resulting weights summed to the 2021 ACS PUMS totals for the raking dimensions and the weights were within the bounds defined by the interquartile range criterion.

After the iterative raking and trimming process, the final weight for person  $m$ , denoted as  $W7CRKWT_l$ , was calculated as shown in equation B.45:

$$W7CRKWT_l = W7CWT_l \times W7RT_l \quad (\text{B.45})$$

where  $W7RT_l$  is the combined raking and trimming adjustment for person  $l$ . The variable representing the final cross-sectional weight for a PATH Study participant in the Wave 7 Cohort responding to the Wave 7 adult interview (“A”) or youth interview (“Y”) is named R07\_x\_C07WGT. These variables are provided separately from the adult and youth/parent data files on weight files that also contain variables for variance estimation.

## B.10.8 Creation of the Wave 7 Shadow Youth Weights

In preparation for the shadow youth completing a youth interview at later waves, Wave 7 weights were created for the shadow youth sample so that Wave 7 shadow youth would have a basis for the future-wave weights. These Wave 7 weights are not included in the data files.

The Wave 7 shadow youth respondents were all from the Wave 7 replenishment sample. All the Wave 7 replenishment sample addresses were eligible for shadow youth screening and sampling. The raked household weight was used as the foundation for calculating the shadow youth weight. The shadow youth base weight was computed as the raked household weight divided by the within-household probability of selection for the shadow youth. Nonresponse adjustment was performed to account for nonresponse to the parent consent using a combination of census 2020 and ACS 5-year (2016-2021) data, appended variables from the address vendor, and person-level data collected during the household screener. Weighting classes were formed using the characteristics similar to those for adults and youth.

Nonresponse-adjusted weights were raked to population totals from the 2021 ACS PUMS using census region, single year of age, race/ethnicity, and sex as raking variables. However, because no interview is conducted for shadow youth until they reach age 12, the only information available for this adjustment was from the household screener.

## Appendix C

### Example Program Code for Analysis

---

This appendix contains example SAS, SUDAAN, Stata, R, and SPSS program code for use with the PATH Study data. Section C.1 provides example code for generating popular statistics and Section C.2 provides example code for calculating and specifying the number of degrees of freedom for testing the significance of estimates. These examples are not meant to be exhaustive or to provide instruction for users unfamiliar with a particular software package.

#### C.1 Example Code for Generating Popular Statistics

This section contains example SAS, SUDAAN, Stata, R, and SPSS program code for generating statistics using the PATH Study data. A few notes about the examples:

- Text in italics represents placeholders for actual dataset and variable names.
- These examples use the adult full-sample and replicate weights for Wave 1, but analyses of youth/parent data for Wave 1 or data from other waves would be the same with the respective full-sample and replicate weights.
- In the examples using subdomain analysis, the domain of interest (e.g., Asian youth or adults ages 18-24) is identified by setting *domainvar* = 1 for those cases, and setting *domainvar* = 0 for all other cases.

These examples are provided primarily to illustrate the correct specifications for creating appropriate variance estimates.

##### C.1.1 SAS

The following code creates tables including the unweighted frequencies of categorical variables *var1*, *var2*, *var3*, and *var4*, and weighted estimates of population totals and population proportions for each level of those variables (using the weight R01\_A\_PWT) along with the

standard errors of these estimates (using the replicate weights R01\_A\_PWGT1 – R01\_A\_PWGT100) and modified Wilson confidence intervals:

```
proc surveyfreq data=analysis_dataset
varmethod=BRR (fay=0.3);
  tables var1 var2 var3 var4/cl(type=Wilson truncate=no);
  weight R01_A_PWGT;
  repweights R01_A_PWGT1 - R01_A_PWGT100;
  title "PROC SURVEYFREQ using BRR-Fay Replication";
run;
```

The following code creates the weighted mean of continuous variable *var5* (using the weight R01\_A\_PWGT) along with the standard error of that estimate (using the replicate weights R01\_A\_PWGT1 – R01\_A\_PWGT100):

```
proc surveymeans data=analysis_dataset
varmethod=BRR (fay=0.3);
  var var5;
  weight R01_A_PWGT;
  repweights R01_A_PWGT1 - R01_A_PWGT100;
  title "PROC SURVEYMEANS using BRR-Fay Replication";
run;
```

The following code fits a linear regression model using continuous variable *respvar* as the outcome variable and continuous variables *cov1* and *cov2* as the predictors. All parameter estimates are weighted (using the weight R01\_A\_PWGT) and the standard errors are calculated using the replicate weights (R01\_A\_PWGT1 – R01\_A\_PWGT100).

```
proc surveyreg data=analysis_dataset
varmethod=BRR (fay=0.3);
  model respvar = cov1 cov2/solution;
  weight R01_A_PWGT;
  repweights R01_A_PWGT1 - R01_A_PWGT100;
  title "PROC SURVEYREG using BRR-Fay Replication";
run;
```

The following code fits a logistic regression model using dichotomous variable *respvar2* (with values of 0 and 1) as the outcome variable and continuous variables *cov1* and *cov2* as the

predictors. All parameter estimates are weighted (using the weight R01\_A\_PWGT) and the standard errors are calculated using the replicate weights (R01\_A\_PWGT1 – R01\_A\_PWGT100).

```
proc surveylogistic data=analysis_dataset
varmethod=BRR (fay=0.3);
  model respvar2 (event='1')= cov1 cov2;
  weight R01_A_PWGT;
  repweights R01_A_PWGT1 - R01_A_PWGT100;
  title "PROC SURVEYLOGISTIC using BRR-Fay Replication";
run;
```

Note that the current SAS/STAT software documentation states that for a complex survey design, analyses for a subpopulation should specify the subpopulation through a DOMAIN statement in proc surveymeans, proc surveyreg, or proc surveylogistic, or as the first dimension in the TABLES statement in proc surveyfreq (see SAS Institute Inc. 2018, pp.9819-9820). However, a simplification is possible for analysis of PATH Study data using the replicate weights (Replication is the recommended method of variance estimation for the PATH Study.) **Specifically, the DOMAIN statement is not required for analyses using the replication method for variance estimation with the replicate weights provided with the data files.** The appropriateness of this simpler approach has been confirmed with software developers at SAS Institute, Inc.

**In some situations, specifying the DOMAIN statement when using the replication method can result in inaccurate variance estimates.** Thus, when using the replicate weights provided with the PATH Study data, the data should be subset to the subpopulation of interest outside the SAS/STAT software procedures, for example in a separate data step, or within the procedures using a BY or WHERE statement to produce the correct standard errors for confidence intervals and hypothesis tests even though, in some older versions of SAS, a warning may appear in the program log stating that the method “does not provide a statistically valid subpopulation or domain analysis.” For this reason, the example SAS code provided in this user guide does not include a DOMAIN statement.

The following code provides examples of domain analyses in SAS. SAS will produce output for all levels of the domain variable. The Output Delivery System (ODS) statements may also be used to output the domain cases of interest for further processing.

```

proc surveyfreq data=analysis_dataset
varmethod=BRR (fay=0.3);
  table var1/cl(type=Wilson truncate=no);
  by domainvar;
  weight R01_A_PWGT;
  repweights R01_A_PWGT1 - R01_A_PWGT100;
  title "Domain Analysis for PROC SURVEYFREQ using BRR-Fay
Replication";
run;

proc surveymeans data=analysis_dataset
varmethod=BRR (fay=0.3);
  var var5;
  by domainvar;
  weight R01_A_PWGT;
  repweights R01_A_PWGT1 - R01_A_PWGT100;
  title "Domain Analysis for PROC SURVEYMEANS using BRR-Fay
Replication";
run;

proc surveyreg data=analysis_dataset
varmethod=BRR (fay=0.3);
  model respvar = cov1 cov2/solution;
  by domainvar;
  weight R01_A_PWGT;
  repweights R01_A_PWGT1 - R01_A_PWGT100;
  title "Domain Analysis for PROC SURVEYREG using BRR-Fay
Replication";
run;

proc surveylogistic data=analysis_dataset
varmethod=BRR (fay=0.3);
  model respvar2 (event='1')= cov1 cov2;
  by domainvar;
  weight R01_A_PWGT;
  repweights R01_A_PWGT1 - R01_A_PWGT100;
  title "Domain Analysis for PROC SURVEYLOGISTIC using BRR-Fay
Replication";
run;
```

## C.1.2 SUDAAN

When using Fay's method of BRR in SUDAAN, the BRR-Fay factor must be calculated outside of the software as shown in equation D.1:

$$ADJFAY = 1/[(1 - k)^2] \quad (C.1)$$

when  $k = 0.3$ , as with the PATH Study,  $ADJFAY = 2.040816$ .

The following code creates tables including the unweighted frequencies of categorical variables *var1*, *var2*, *var3*, and *var4*, and weighted estimates of population totals and population proportions for each level of those variables (using the weight R01\_A\_PWGT) along with the standard errors of these estimates (using the replicate weights R01\_A\_PWGT1 – R01\_A\_PWGT100):

```
proc crosstab data=analysis_dataset filetype=sas design=brr;
  weight R01_A_PWGT;
  repwgt R01_A_PWGT1 - R01_A_PWGT100/adjfay=2.040816;
  tables var1 var2 var3 var4;
  class var1 var2 var3 var4;
  print/style=nchs tablecell=all;
  title 'SUDAAN proc crosstab using BRR-Fay Replication';
run;
```

By default, SUDAAN generates the logit transformed confidence intervals with the above code. To obtain the Clopper-Pearson small percentage confidence intervals, the SMCONF option is specified in the PROC statement. For example, SMCONF = 25 produces the Clopper-Pearson confidence intervals for percentages less than or equal to 25% or percentages greater than or equal to 75%. For percentages between 25% and 75%, the logit transformed confidence intervals are calculated. To print the Clopper-Pearson confidence intervals for small percentages, the ROWSPCI option is

specified in the PRINT statement. The following code provides an example for creating Clopper-Pearson confidence intervals:

```
proc crosstab data=analysis_dataset filetype=sas design=brr
smconf=25;
  weight R01_A_PWGT;
  repwgt R01_A_PWGT1 - R01_A_PWGT100/adjfay=2.040816;
  tables var1 var2 var3 var4;
  class var1 var2 var3 var4;
  print rowspci/ style=nchs tablecell=all;
  title 'SUDAAN proc crosstab using BRR-Fay Replication and
producing Clopper-Pearson confidence intervals';
run;
```

The following code creates the weighted mean of continuous variable *var5* (using the weight R01\_A\_PWGT) along with the standard error of that estimate (using the replicate weights R01\_A\_PWGT1 – R01\_A\_PWGT100):

```
proc descript data=analysis_dataset filetype=sas design=brr;
  weight R01_A_PWGT;
  repwgt R01_A_PWGT1- R01_A_PWGT100/adjfay=2.040816;
  var var5;
  print/style=nchs;
  title 'SUDAAN proc descript using BRR-Fay Replication';
run;
```

The following code fits a linear regression model using continuous variable *respvar* as the outcome variable and continuous variables *cov1* and *cov2* as the predictors. All parameter estimates are weighted (using the weight R01\_A\_PWGT) and the standard errors are calculated using the replicate weights (R01\_A\_PWGT1 – R01\_A\_PWGT100).

```
proc regress data=analysis_dataset filetype=sas design=brr;
  weight R01_A_PWGT;
  repwgt R01_A_PWGT1 - R01_A_PWGT100/adjfay=2.040816;
  model respvar = cov1 cov2;
  title 'SUDAAN proc regress using BRR-Fay Replication';
run;
```

The following code fits a logistic regression model using dichotomous variable *respvar2* (with values of 0 and 1) as the outcome variable and continuous variables *cov1* and *cov2* as the

predictors. All parameter estimates are weighted (using the weight R01\_A\_PWGT) and the standard errors are calculated using the replicate weights (R01\_A\_PWGT1 – R01\_A\_PWGT100).

```
proc rlogist data=analysis_dataset filetype=sas design=brr;
  weight R01_A_PWGT;
  repwgt R01_A_PWGT1 - R01_A_PWGT100/adjfay=2.040816;
  model respvar2 = cov1 cov2;
  title 'SUDAAN proc rlogist using BRR-Fay Replication';
run;
```

Domain analyses can be performed using the same examples above with the inclusion of a subpopn statement. For example, for frequencies *var1*, *var2*, *var3*, and *var4* for just those respondents with *domainvar* = 1, the following code could be used:

```
proc crosstab data=analysis_dataset filetype=sas design=brr;
  subpopn domainvar = 1;
  weight R01_A_PWGT;
  repwgt R01_A_PWGT1 - R01_A_PWGT100/adjfay=2.040816;
  tables var1 var2 var3 var4;
  class var1 var2 var3 var4;
  print/style=nchs tablecell=all;
  title 'SUDAAN proc crosstab using BRR-Fay Replication';
run;
```

### C.1.3 Stata

The full-sample weight, variance estimation method, BRR-Fay replicate weights, and Fay's factor are relayed to Stata using the svyset statement. The following statement should be used with the PATH Study data:

```
svyset [pweight= r01_a_pwgt], brr(r01_a_pwgt1 - r01_a_pwgt100)
vce(brr) mse fay(.3)
```

Assuming this svyset is used, the following code creates tables including the unweighted frequencies of categorical variables *var1*, *var2*, *var3*, and *var4*, weighted estimates of population totals and population proportions for each level of those variables (using the weight R01\_A\_PWGT) along with the standard errors of these estimates (using the replicate weights R01\_A\_PWGT1 – R01\_A\_PWGT100), and modified Wilson confidence intervals of estimated

population proportions for a categorical variable *var1* (using the weight R01\_A\_PWGT and the replicate weights R01\_A\_PWGT1 – R01\_A\_PWGT100):

For weighted frequencies/estimates of population totals:

```
svy: tabulate var1, count se
svy: tabulate var2, count se
svy: tabulate var3, count se
svy: tabulate var4, count se
```

For weighted frequencies/estimates of population proportions:

```
svy: tabulate var1, se obs percent
svy: tabulate var2, se obs percent
svy: tabulate var3, se obs percent
svy: tabulate var4, se obs percent
```

Note that if all four variables are used in a single tabulate statement, the result is a multi-dimensional table containing all four variables rather than four one-dimensional tables.

To create these estimates for respondents with *domainvar* = 1, the following code may be used for each variable:

```
svy, subpop(domainvar): tabulate var1, count se
svy, subpop(domainvar): tabulate var1, se obs percent
```

Note that in the example above, Stata assumes that the subdomain of interest is *domainvar* = 1.

For modified Wilson confidence intervals of estimated population proportions:

```
svy: proportion var1, ciptype(wilson)
```

To create modified Wilson confidence intervals of estimated population proportions for each level of *domainvar*, the following code may be used:

```
svy: proportion var1, over(domainvar) ciptype(wilson)
```

Assuming the `svyset` statement above is used, the following code creates the weighted mean of continuous variable `var5` (using the weight `R01_A_PWGT`) along with the standard error of that estimate (using the replicate weights `R01_A_PWGT1 – R01_A_PWGT100`):

```
svy: mean var5
```

To create these estimates for each of the levels of `domainvar`, the following code may be used:

```
svy: mean var5, over(domainvar)
```

Assuming the `svyset` statement above is used, the following code fits a linear regression model using continuous variable `respvar` as the outcome variable and continuous variables `cov1` and `cov2` as the predictors. All parameter estimates are weighted (using the weight `R01_A_PWGT`) and the standard errors are calculated using the replicate weights (`R01_A_PWGT1 – R01_A_PWGT100`):

```
svy: regress respvar cov1 cov2
```

To perform this regression for respondents with `domainvar = 1`, the following code may be used:

```
svy, subpop(domainvar): regress respvar cov1 cov2
```

The following code fits a logistic regression model using dichotomous variable `respvar2` (with values of 0 and 1) as the outcome variable and continuous variables `cov1` and `cov2` as the predictors. All parameter estimates are weighted (using the weight `R01_A_PWGT`) and the standard errors are calculated using the replicate weights (`R01_A_PWGT1 – R01_A_PWGT100`).

For odds ratios:

```
svy: logistic respvar2 cov1 cov2
```

For coefficient estimates:

```
svy: logit respvar2 cov1 cov2
```

To perform this regression for respondents with *domainvar* = 1, the following code may be used:

```
svy, subpop(domainvar): logistic respvar2 cov1 cov2
svy, subpop(domainvar): logit respvar2 cov1 cov2
```

## C.1.4 R

The full-sample weight, variance estimation method, BRR-Fay replicate weights, and Fay's factor are relayed to the R survey package using the svrepdesign function. The following should be used with the PATH Study data:

```
options(survey.replicates.mse=TRUE)
y <-
  svrepdesign(
    id = ~PERSONID,
    weights = ~R01_A_PWT,
    repweights = "R01_A_PWT[1-9]+",
    type = "Fay",
    rho = 0.3,
    data = analysis_dataset
  )
```

Assuming the options and svrepdesign function above are executed, the following code creates the estimates of population totals and population proportions of the categorical variables *var1*, *var2*, *var3*, and *var4* (using the weight R01\_A\_PWT) along with the standard errors of these estimates (using the replicate weights R01\_A\_PWT1 – R01\_A\_PWT100):

For weighted frequencies/estimates of population totals:

```
svytotal(~factor(var1), design=y, na.rm=T)
svytotal(~factor(var2), design=y, na.rm=T)
svytotal(~factor(var3), design=y, na.rm=T)
svytotal(~factor(var4), design=y, na.rm=T)
```

For weighted frequencies/estimates of population proportions:

```
svymean(~factor(var1), design=y, na.rm=T)
svymean(~factor(var2), design=y, na.rm=T)
svymean(~factor(var3), design=y, na.rm=T)
svymean(~factor(var4), design=y, na.rm=T)
```

Note that this code should be executed for one variable at a time or the function will output estimates using only those records that have non-missing values across all the variables specified.

For confidence intervals:

```
svyciprop(~var1, y, method = "lo", level = 0.95)
svyciprop(~var2, y, method = "lo", level = 0.95)
svyciprop(~var3, y, method = "lo", level = 0.95)
svyciprop(~var4, y, method = "lo", level = 0.95)
```

To create these estimates for each of the levels of *domainvar*, the following code may be used:

```
svyby(~factor(var1), ~domainvar, svytot, design=y, na.rm=T)
svyby(~factor(var1), ~domainvar, svymean, design=y, na.rm=T)
```

Assuming the options and svrepdesign function above are executed, the following code creates the weighted mean of continuous variable *var5* (using the weight R01\_A\_PWGT) along with the appropriate standard errors of that estimate (using the replicate weights R01\_A\_PWGT1 – R01\_A\_PWGT100):

```
svymean(~var5, design = y, na.rm=T)
```

To create these estimates for each of the levels of *domainvar*, the following code may be used:

```
svyby(~var5, ~domainvar, svymean, design=y, na.rm=T)
```

Assuming the options and svrepdesign function above are executed, the following code fits a linear regression model using continuous variable *respvar* as the outcome variable and continuous variables *cov1* and *cov2* as the predictors. All parameter estimates are weighted (using the weight R01\_A\_PWGT) and the standard errors are calculated using the replicate weights (R01\_A\_PWGT1 – R01\_A\_PWGT100).

For coefficient estimates:

```
svyglm(respvar ~ cov1 + cov2, design=y)
```

For standard errors and significance test of model coefficients:

```
summary(svyglm(respvar ~ cov1 + cov2, design=y))
```

The following code fits a logistic regression model using dichotomous variable *respvar2* (with values of 0 and 1) as the outcome variable and continuous variables *cov1* and *cov2* as the predictors. All parameter estimates are weighted (using the weight R01\_A\_PWGT) and the standard errors are calculated using the replicate weights (R01\_A\_PWGT1 – R01\_A\_PWGT100).

For coefficient estimates:

```
svyglm(respvar2 ~ cov1 + cov2, design=y, family=binomial)
```

For standard errors and significance test of model coefficients:

```
summary(svyglm(respvar2 ~ cov1 + cov2, design=y,  
family=binomial))
```

To perform these regressions for respondents with *domainvar* = 1, the design statement should be changed as follows:

```
design=subset(y, domainvar=="1"))
```

This design statement applies to both linear and logistic regressions.

## C.1.5 SPSS

SPSS does not have the functionality to create variance estimates using the replication method, the preferred method for the PATH Study. As such, only the Taylor series linearization approach to variance estimation may be used with this software package. The appropriate design variables and full-sample weights are relayed to SPSS through a plan file created using the csplan function.

The following code should be used to create a plan file for use with the PATH Study data:

```
csplan analysis
/plan file="c:\myspace\myplan.csaplan"
/planvars analysisweight=R01_A_PWGT
/print plan
/design strata=VARSTRAT cluster=VARPSU
/estimator type=wr.
```

Assuming the plan file is created as specified above, the following code creates tables including the frequencies of categorical variables *var1*, *var2*, *var3*, and *var4*, and estimates of population totals and population proportions for each level of those variables (using the weight R01\_A\_PWGT) along with the standard errors of these estimates (created using linearization):

```
cstabulate
/plan file="c:\myspace\myplan.csaplan"
/tables variables=var1 var2 var3 var4
/cells popsize tablepct
/statistics count se cin(95).
```

Because SPSS only supports the Taylor series linearization approach to variance estimation, analyses for a subpopulation must be made using all cases rather than conducting the analysis only on a subset of cases of analytic interest. Estimates for each of the levels of *domainvar* may be created by including the following statement:

```
/subpop table=domainvar
```

Assuming the plan file is created as specified above, the following code creates the weighted mean of continuous variable *var5* (using the weight R01\_A\_PWGT) along with the appropriate standard error of that estimate (using linearization):

```
csdescriptives
/plan file="c:\myspace\myplan.csaplan"
/summary variables var5
/mean
/statistics se cin(95).
```

Estimates for each of the levels of *domainvar* may be created by including the following statement:

```
/subpop table=domainvar
```

Assuming the plan file is created as specified above, the following code fits a linear regression model using continuous variable *respvar* as the outcome variable and continuous variables *cov1* and *cov2* as the predictors. All parameter estimates are weighted (using the weight R01\_A\_PWGT) and the standard errors are calculated using linearization.

```
csglm respvar with cov1 cov2  
/plan file="c:\myspace\myplan.csaplan"  
/model cov1 cov2  
/statistics parameter se cinterval.
```

The following code fits a logistic regression model using dichotomous variable *respvar2* (with values of 0 and 1) as the outcome variable and continuous variables *cov1* and *cov2* as the predictors. All parameter estimates are weighted (using the weight R01\_A\_PWGT) and the standard errors are calculated using linearization.

```
cslogistic respvar2 with cov1 cov2  
/plan file="c:\myspace\myplan.csaplan"  
/model cov1 cov2  
/statistics parameter se cinterval.
```

To perform either of these regressions for respondents with *domainvar* = 1, the following statement may be inserted:

```
/domain variable domainvar (1)
```

## C.2 Example Code for Calculating and Specifying Degrees of Freedom for Analysis

This section contains example SAS, SUDAAN, Stata, R, and SPSS program code for calculating and specifying the number of degrees of freedom for testing the significance of estimates from the PATH Study data as discussed in Section 6.7. A few of notes about the examples:

- The variable *characteristic* is a variable with two categories corresponding to the values of 1 and 0.
- Text in italics represents placeholders for actual dataset names, variable names, and the calculated degrees of freedom.
- Most examples indicate the analysis of only one of the two categories; however, researchers are expected to repeat the analyses for all categories of interest, using the appropriate degrees of freedom for each category.
- These examples use the youth/parent full-sample and replicate weights from Wave 1, but analyses of youth/parent or adult data from other waves would be the same with the respective full-sample and replicate weights.
- *degrees\_of\_freedom1* corresponds to the number of degrees of freedom calculated for *characteristic* category 1.
- *degrees\_of\_freedom\_min* corresponds to the smallest number of degrees of freedom calculated across all categories of *characteristic*.

These examples are provided primarily to illustrate one method to calculate the degrees of freedom and the specifications for adjusting the degrees of freedom for significance testing in each software package. There may be other ways to accomplish the same task in each particular software package.

### C.2.1 SAS

The following code sorts and de-duplicates the data, first by *characteristic*, VARSTRAT, and VARPSU, then by *characteristic* and VARSTRAT. This creates a dataset with one

record for each pseudo-PSU (*uniquepsus*) and another with one record for each pseudo-strata (*uniquestrat*) for each category of *characteristic*.

```
proc sort nodupkey data=analysis_data out=uniquepsus;
  by characteristic VARSTRAT VARPSU;
run;
proc sort nodupkey data=analysis_data out=uniquestrat;
  by characteristic VARSTRAT;
run;
```

Using the input *uniquepsus* and *uniquestrat*, the following statements calculate the frequency of pseudo-PSUs and pseudo-strata by *characteristic*. This results in two datasets, *numpsusfreq* and *numstratfreq*, each having two records with counts of PSUs and strata by *characteristic*.

```
proc freq data=uniquepsus;
  table characteristic/out = numpsusfreq;
  title 'Number of pseudo-PSUs by group';
run;
proc freq data=uniquestrat;
  table characteristic/out = numstratfreq;
  title 'Number of pseudo-strata by group';
run;
```

The user could manually calculate the number of degrees of freedom based on the tables generated by the statements above. To do this programmatically, this final code below merges the counts of pseudo-PSUs and pseudo-strata from *numpsusfreq* and *numstratfreq*, by *characteristic* and then calculates *degrees\_of\_freedom* by subtracting the number of pseudo-strata from the number of pseudo-PSUs for each category of *characteristic*.

```
data df;
  merge numpsusfreq(rename=(count=numpsus))
    numstratfreq(rename=(count=numstrat));
  by characteristic;
  degrees_of_freedom=numpsus-numstrat;
run;
```

To view the calculated values, print the dataset *df*.

```
proc print data=df;
run;
```

The following code creates tables including the unweighted frequencies of categorical variables *var1*, *var2*, *var3*, and *var4*, and estimates of population proportions for each level of those variables (using the weight R01\_Y\_PWGT) along with the standard errors of these estimates (using the replicate weights R01\_Y\_PWGT1 – R01\_Y\_PWGT100) and modified Wilson confidence intervals for respondents with *characteristic* = 1:

```
proc surveyfreq data=analysis_data
  varmethod=BRR (fay=0.3);
  tables var1 var2 var3 var4/
    df = degrees_of_freedom1 cl(type=Wilson truncate=no);
  where characteristic = 1;
  weight R01_Y_PWGT;
  repweights R01_Y_PWGT1 - R01_Y_PWGT100;
  title "PROC SURVEYFREQ using BRR-Fay Replication with
    adjusted degrees of freedom";
run;
```

The following code creates the weighted mean of continuous variable *var5* (using the weight R01\_Y\_PWGT) along with the standard error of that estimate (using the replicate weights R01\_Y\_PWGT1 – R01\_Y\_PWGT100) and a confidence interval around the mean for respondents with *characteristic* = 1:

```
proc surveymeans data=analysis_data
  varmethod=BRR (fay=0.3)mean clm;
  var var5;
  where characteristic = 1;
  weight R01_Y_PWGT;
  repweights R01_Y_PWGT1 - R01_Y_PWGT100/
    df = degrees_of_freedom1;
  title "PROC SURVEYMEANS using BRR-Fay Replication with
    adjusted degrees of freedom";
run;
```

The following code fits a linear regression model using continuous variable *respvar* as the outcome variable and *characteristic* as the predictor. All parameter estimates are weighted (using the weight R01\_Y\_PWGT), the standard errors are calculated using the replicate weights

(R01\_Y\_PWGT1 – R01\_Y\_PWGT100), and the significance testing is conducted using the smallest degrees of freedom calculated across the categories of *characteristic*:

```
proc surveyreg data=analysis_data
varmethod=BRR (fay=0.3);
  model respvar = characteristic/
    df = degrees_of_freedom_min;
  weight R01_Y_PWGT;
  repweights R01_Y_PWGT1 - R01_Y_PWGT100;
  title "PROC SURVEYREG using BRR-Fay Replication with adjusted
    degrees of freedom";
run;
```

## C.2.2 SUDAAN

When using Fay's method of BRR in SUDAAN, the BRR-Fay factor must be calculated outside of the software as shown in equation D.2:

$$ADJFAY = 1/[(1 - k)^2] \quad (\text{D.2})$$

When  $k = 0.3$ , as with the PATH Study,  $ADJFAY = 2.040816$ . The degrees of freedom must also be calculated outside of the software.

The following code creates tables including the unweighted frequencies of categorical variables *var1*, *var2*, *var3*, and *var4*, estimates of population proportions for each level of those variables (using the weight R01\_Y\_PWGT) along with the standard errors of these estimates (using the replicate weights R01\_Y\_PWGT1 – R01\_Y\_PWGT100) and confidence intervals for respondents with *characteristic* = 1:

```
proc crosstab data= analysis_data filetype=sas design=brr
ddf = degrees_of_freedom1;
  weight R01_Y_PWGT;
  repwgt R01_Y_PWGT1 - R01_Y_PWGT100/adjfay=2.040816;
  tables var1 var2 var3 var4;
  class var1 var2 var3 var4;
  subpopn characteristic = 1;
  print/style=nchs tablecell=all;
  title 'SUDAAN proc crosstab using BRR-Fay Replication with
    adjusted degrees of freedom';
run;
```

The following code creates the weighted mean of continuous variable *var5* (using the weight R01\_Y\_PWGT) along with the standard error of that estimate (using the replicate weights R01\_Y\_PWGT1 – R01\_Y\_PWGT100) and a confidence interval around the mean for respondents with *characteristic* = 1:

```
proc descript data= analysis_data filetype=sas design=brr ddf =
degrees_of_freedom1;
  weight R01_Y_PWGT;
  repwgt R01_Y_PWGT1- R01_Y_PWGT100/adjfay=2.040816;
  var var5;
  subpopn characteristic = 1;
  print/style=nchs;
  title 'SUDAAN proc descript using BRR-Fay Replication with
adjusted degrees of freedom';
run;
```

The following code fits a linear regression model using continuous variable *respvar* as the outcome variable and *characteristic* as the predictor. All parameter estimates are weighted (using the weight R01\_Y\_PWGT), the standard errors are calculated using the replicate weights (R01\_Y\_PWGT1 – R01\_Y\_PWGT100), and the significance testing is conducted using the smallest degrees of freedom calculated across the categories of *characteristic*:

```
proc regress data=analysis_data filetype=sas design=brr
ddf = degrees_of_freedom_min;
  weight R01_Y_PWGT;
  repwgt R01_Y_PWGT1 - R01_Y_PWGT100/adjfay=2.040816;
  model respvar = characteristic;
  title 'SUDAAN proc regress using BRR-Fay Replication with
adjusted degrees of freedom';
run;
```

### C.2.3 Stata

Starting with the initial *analysis\_data* dataset, the following code identifies or tabulates unique combinations of *characteristic*, VARSTRAT, and VARPSU.

```
egen uniquepsus = tag(characteristic VARSTRAT VARPSU)
```

This creates the variable *uniquepsus* with only the unique combinations of the number of pseudo-PSUs. Using the variable *uniquepsus*, the following statement calculates the number of pseudo-PSUs by *characteristic* and saves them in the variable *numpsus*.

```
egen numpsus = total(uniquepsus), by(characteristic)
```

The code below displays the pseudo-PSUs by *characteristic*.

```
tabdisp characteristic, c(numpsus)
```

Starting with the initial *analysis\_data* dataset, the process then needs to be repeated to calculate the number of pseudo-strata:

```
egen uniquestrat = tag(characteristic VARSTRAT)
egen numstrat = total(uniquestrat), by(characteristic)
tabdisp characteristic, c(numstrat)
```

The user then must subtract the values displayed in *numstrat* from the values in *numpsus* based on the tables generated by the statements above to calculate the degrees of freedom for each level of *characteristic*.

The full-sample weight, variance estimation method, BRR-Fay replicate weights, and Fay's factor are relayed to Stata using the *svyset* statement. The following statement should be used with the PATH Study youth data:

```
svyset [pweight= R01_Y_PWT], brrweight(R01_Y_PWT1 -
R01_Y_PWT100) vce(brr) mse fay(.3)
```

Assuming this *svyset* is used, the following code creates tables including the weighted estimates of population proportions for each level of categorical variables *var1*, *var2*, *var3*, and *var4* (using the weight *R01\_Y\_PWT*) along with the standard errors of these estimates (using the

replicate weights R01\_Y\_PWGT1 – R01\_Y\_PWGT100) and logit confidence intervals of estimated population proportions for respondents with *characteristic* = 1:

```
svy, subpop(characteristic): tabulate var1, percent ci se obs
dof(degrees_of_freedom1)
svy, subpop(characteristic): tabulate var2, percent ci se obs
dof(degrees_of_freedom1)
svy, subpop(characteristic): tabulate var3, percent ci se obs
dof(degrees_of_freedom1)
svy, subpop(characteristic): tabulate var4, percent ci se obs
dof(degrees_of_freedom1)
```

To create modified Wilson confidence intervals of estimated population proportions for each level of *characteristic*, the following code may be used with the degrees of freedom specified in a **svyset** statement:

```
svyset [pweight= R01_Y_PWGT], brrweight(R01_Y_PWGT1 -
R01_Y_PWGT100) dof(degrees_of_freedom1) vce(brr) mse fay(.3)
svy: proportion var1, over(characteristic) citype(Wilson)
```

For calculating means or regression estimates, the degrees of freedom must be specified in the **svyset** procedure. The following code includes a **svyset** statement and creates the weighted mean of continuous variable *var5* (using the weight R01\_Y\_PWGT) along with the standard error of that estimate (using the replicate weights R01\_Y\_PWGT1 – R01\_Y\_PWGT100), and confidence interval around the mean for respondents with *characteristic* = 1:

```
svyset [pweight= R01_Y_PWGT], brrweight(R01_Y_PWGT1 -
R01_Y_PWGT100) dof(degrees_of_freedom1) vce(brr) mse fay(.3)
svy, subpop(characteristic): mean var5
```

The following code includes a new **svyset** statement and fits a linear regression model using continuous variable *respvar* as the outcome variable and *characteristic* as the predictor. All parameter estimates are weighted (using the weight R01\_Y\_PWGT), the standard errors are calculated using the replicate weights (R01\_Y\_PWGT1 – R01\_Y\_PWGT100), and the significance testing is conducted using the smallest degrees of freedom calculated across the categories of *characteristic*.

```
svyset [pweight= R01_Y_PWGT], brrweight(R01_Y_PWGT1 -
R01_Y_PWGT100) dof(degrees_of_freedom_min) vce(brr) mse fay(.3)
svy: regress respvar i.characteristic
```

## C.2.4 R

The following code sorts and de-duplicates the data, first by *characteristic*, VARSTRAT, and VARPSU, then by *characteristic* and VARSTRAT. This leaves only the unique combinations of the number of pseudo-PSUs (*uniquepsus*) and the number of pseudo-strata (*uniquestrat*) for each category of *characteristic*.

```
uniquepsus <- analysis_data
[!duplicated(analysis_data [c("characteristic",
"VARSTRAT", "VARPSU")]), ]
uniquestrat <- analysis_data
[!duplicated(analysis_data [c("characteristic",
"VARSTRAT")])], ]
```

Using the objects *uniquepsus* and *uniquestrat*, the following statements calculate the frequency of pseudo-PSUs and pseudo-strata by *characteristic*. This results in two objects, *numpsusfreq* and *numstratfreq*, with counts of PSUs and strata by *characteristic*.

```
numpsusfreq <- table(uniquepsus$characteristic)
numstratfreq <- table(uniquestrat$characteristic)
```

The following code calculates the values of *degrees\_of\_freedom* for each value of *characteristic* by subtracting the number of pseudo-strata from the number of pseudo-PSUs.

```
degrees_of_freedom <- numpsusfreq - numstratfreq
```

To view the calculated values, display *degrees\_of\_freedom*.

The full-sample weight, variance estimation method, BRR-Fay replicate weights, and Fay's factor are relayed to the R survey package using the svrepdesign function. The following should be used with the PATH Study data:

```
options(survey.replicates.mse=TRUE)
y <-
  svrepdesign(
    id = ~PERSONID,
    weights = ~R01_Y_PWGT,
    repweights = "R01_Y_PWGT[1-9]+",
    type = "Fay",
    rho = 0.3,
    data = analysis_data
  )
```

Assuming the options and svrepdesign function above are executed, the following code creates the estimates of population proportions for each level of the categorical variables *var1*, *var2*, *var3*, and *var4* (using the weight R01\_Y\_PWGT) along with the standard errors of these estimates (using the replicate weights R01\_Y\_PWGT1 – R01\_Y\_PWGT100), and confidence intervals for respondents with *characteristic* = 1:

```
y1=subset(y, characteristic ==1)
```

For estimates and standard errors:

```
svymean(~factor(var1), design=y1, na.rm=T)
svymean(~factor(var2), design=y1, na.rm=T)
svymean(~factor(var3), design=y1, na.rm=T)
svymean(~factor(var4), design=y1, na.rm=T)
```

For confidence intervals:

```
svyciprop(~var1, y1, method = "lo", level = 0.95,
df = degrees_of_freedom1, na.rm=T)
svyciprop(~var2, y1, method = "lo", level = 0.95,
df = degrees_of_freedom1, na.rm=T)
svyciprop(~var3, y1, method = "lo", level = 0.95,
df = degrees_of_freedom1, na.rm=T)
svyciprop(~var4, y1, method = "lo", level = 0.95,
df = degrees_of_freedom1, na.rm=T)
```

Assuming the options and `svrepdesign` function above are executed, the following code creates the weighted mean of continuous variable `var5` (using the weight `R01_Y_PWGT`) along with the appropriate standard error of that estimate (using the replicate weights `R01_Y_PWGT1 – R01_Y_PWGT100`) and a confidence interval for respondents with `characteristic = 1`:

For the mean and standard error of the mean:

```
svymean(~var5, design = y1, na.rm=T)
```

For the confidence interval around the mean:

```
confint(svymean(~var5, design = y1, na.rm=T), level = 0.95,
df = degrees_of_freedom1)
```

Assuming the options and `svrepdesign` function above are executed, the following code fits a linear regression model using continuous variable `respvar` as the outcome variable and `characteristic` as the predictor. All parameter estimates are weighted (using the weight `R01_Y_PWGT`), the standard errors are calculated using the replicate weights (`R01_Y_PWGT1 – R01_Y_PWGT100`), and the significance testing is conducted using the smallest degrees of freedom calculated across the categories of `characteristic`.

For coefficient estimates:

```
svyglm(respvar ~ characteristic, design=y)
```

For standard errors and significance test of model coefficients:

```
summary(svyglm(respvar ~ characteristic, design=y),
df.resid = degrees_of_freedom_min)
```

## C.2.5 SPSS

Starting with the `analysis_data` dataset, the following code sorts the data by `characteristic`, `VARSTRAT`, and `VARPSU`. This leaves only the unique combinations of the

number of pseudo-PSUs for each category of *characteristic* and saves the result as *sortedpsus*.

```
SORT CASES BY characteristic VARSTRAT VARPSU.  
AGGREGATE OUTFILE= 'c:\myspace\sortedpsus.sav' /BREAK=ALL/N=n.
```

The following code starts with *sortedpsus* and then de-duplicates the file by *characteristic*, VARSTRAT, and VARPSU.

```
MATCH FILES/FILE=*  
/TABLE='c:\myspace\sortedpsus.sav'  
/BY characteristic VARSTRAT VARPSU.  
EXECUTE.
```

```
MATCH FILES/FILE=*  
/FIRST = top/BY characteristic VARSTRAT VARPSU.  
SELECT IF top.  
EXECUTE.
```

Using the working dataset, the following statement calculates the frequency of pseudo-PSUs by *characteristic* and displays the result.

```
FREQUENCIES VARIABLES=characteristic/ORDER=ANALYSIS.
```

Starting with the *analysis\_data* dataset, the process then needs to be repeated to calculate the number of pseudo-strata:

```
SORT CASES BY characteristic VARSTRAT.  
AGGREGATE OUTFILE= 'c:\myspace\sortedstrat.sav' /BREAK=ALL/N=n.
```

```
MATCH FILES/FILE=*  
/TABLE='c:\myspace\sortedstrat.sav'  
/BY characteristic VARSTRAT.  
EXECUTE.
```

```
MATCH FILES/FILE=*  
/FIRST = top/BY characteristic VARSTRAT.  
SELECT IF top.  
EXECUTE.
```

```
FREQUENCIES VARIABLES=characteristic/ORDER=ANALYSIS.
```

The user then must subtract the values of number of pseudo-strata from the second table from the number of pseudo-PSUs created in the first table to calculate the degrees of freedom for each level of *characteristic*.

SPSS does not have the functionality to create variance estimates using the replication method, the preferred method for the PATH Study. As such, only the Taylor series linearization approach to variance estimation may be used with this software package. The appropriate design variables and full-sample weights are relayed to SPSS through a plan file created using the csplan function. The following code should be used to create a plan file for use with the PATH Study data:

```
csplan analysis  
/plan file="c:\myspace\myplan.csaplan"  
/planvars analysisweight=R01_Y_PWGT  
/print plan  
/design strata=VARSTRAT cluster=VARPSU  
/estimator type=wr.
```

Assuming the plan file is created as specified above, SPSS will automatically use the appropriate degrees of freedom (i.e., those calculated as indicated in Appendix A) in creating confidence intervals for proportions and means when the subpop statement is specified. The following code creates tables including the frequencies of categorical variables var1, var2, var3, and var4, and estimates of population proportions for each level of those variables (using the weight R01\_Y\_PWGT) along with the standard errors of these estimates (created using linearization) and confidence intervals for each level of *characteristic*:

```
cstabulate  
/plan file="c:\myspace\myplan.csaplan"  
/tables variables=var1 var2 var3 var4  
/subpop table= characteristic display=layered  
/cells tablepct  
/statistics count se.
```

Assuming the plan file is created as specified above, the following code creates the weighted mean of continuous variable *var5* (using the weight R01\_Y\_PWGT) along with the appropriate standard error (using linearization) and a confidence interval for each level of *characteristic*:

```
csdescriptives  
/plan file="c:\myspace\myplan.csaplan"  
/summary variables var5  
/subpop table=characteristic display=layered  
/statistics se cin(95)  
/mean.
```

Assuming the plan file is created as specified above, the following code fits a linear regression model using continuous variable *respvar* as the outcome variable and *characteristic* as the predictor. In this case, the minimum degrees of freedom across the two levels of *characteristic* must be specified. All parameter estimates are weighted (using the weight R01\_Y\_PWGT) and the standard errors are calculated using linearization.

```
csglm var1 with characteristic  
/plan file="c:\myspace\myplan.csaplan"  
/model characteristic  
/intercept include=yes show=yes  
/statistics parameter se cinterval  
/print covb summary variableinfo sampleinfo  
/test type=f padjust=lsd  
/missing classmissing=exclude  
/criteria cilevel=95 df= degrees_of_freedom_min.
```

## Appendix D Imputation

---

### D.1 Wave 1 Imputation

The sections below provide the methods used for imputing sex (Section D.1.1), age (Section D.1.2), education (Section D.1.3), race (Section D.1.4), and ethnicity (Section D.1.5). Section 7.4 provides a description of the variable naming convention used for the adult data files and youth/parent data files.

#### D.1.1 Sex

Sex was assigned for adults and youth from information provided in the extended interview (questionnaire item R01\_AM0004 for adults and item R01\_YM0004 for youth). If this was not available, either because the sampled person refused to provide it or because the sampled person did not respond to the interview, sex was assigned based on the information provided in the household screener.

After reviewing these sources, however, sex remained unavailable for a small number of sampled persons. Common demographic variables (for example, census region, age, education, race) are not indicators of sex. In most cases, this information was also unavailable if information on sex could not be obtained from the extended interview or household screener, so sex was randomly assigned so that approximately one half of the cases with missing sex was assigned as female and the other half was assigned as male.

The values of R01\_AM0004 and R01\_YM0004 are provided in the variables R01R\_A\_SEX for adults and R01R\_Y\_SEX for youth. All values, imputed and unimputed, are contained in the variable R01R\_x\_SEX\_IMP, where  $x$  is “A” (adult) or “Y” (youth); the variable R01R\_x\_SEX\_IMPFLAG indicates which values are from the interview, i.e., not imputed (R01R\_x\_SEX\_IMPFLAG = 0), from the household screener (R01R\_x\_SEX\_IMPFLAG = 1), or the result of the random imputation method described above (R01R\_x\_SEX\_IMPFLAG = 2).

## D.1.2 Age

A single-year-of-age variable was created for both adults and youth. For adults, age was first calculated based on the date of birth provided during the interview. If date of birth was not provided in the interview, the age in years was used if available.<sup>48</sup> For youth, age was first calculated based on the date of birth provided by the parent during the consent process. If date of birth was not provided, the age in years (also requested during the consent process) was used.

If age was still missing (because the responding adult refused to provide their date of birth or age, the responding youth was an emancipated minor and no parental consent was required, or the sampled person did not respond to the extended interview), age in years provided in the household screener was used.

For nonresponding youth, if age was missing after the above processing, a value was statistically imputed since age in years was used in the youth weighting process. Given that the only information available about these youth is from the household screener and that information (such as race, ethnicity, or sex) is not indicative of a child's age, the age in years was randomly assigned using the youth population distribution according to the 2013 ACS 1-year Public Use Microdata Sample (PUMS).

For adults, a four-level age category (18-24, 25-44, 45-64, and 65 and older) was used for the weighting process, but was not included on the data file. If the imputed single year of age was missing after the above processing, but the sampled adult indicated in the Phase 2 screener that they were in one of the age categories under age 30, that age category was used.<sup>49</sup> If that information was not available, the age category indicated in the household screener was used.

As with youth, if none of this information was available for an adult, the age category was imputed using the population distribution according to the 2013 ACS PUMS data. Because a higher proportion of adults in the older age categories are female, the random assignment was conducted separately by sex. Age range information, provided in the Phase 2 screener, was used for some adults who indicated they were age 30 or older. For these adults, the age category was randomly assigned

---

<sup>48</sup>Only adults who did not provide their date of birth were asked to provide their age in years.

<sup>49</sup>Questionnaire item R01\_AM0003 requesting an age category was asked in the interview only if the sampled adult refused to provide a date of birth or age in years. The highest age category was "30 or older."

using the population distribution by sex according to the 2013 ACS PUMS data for that age group. If it was unknown whether the sampled adult was age 30 or older, the age category was randomly assigned using the population distribution by sex for those 18 and older.

Neither the single year of age variable nor its imputed counterpart is provided on the PUFs. Instead, age range variables, R01R\_A\_AGECAT7 for adults and R01R\_Y\_AGECAT2 for youth, are provided along with their imputed counterparts R01R\_A\_AGECAT7\_IMP and R01R\_Y\_AGECAT2\_IMP, respectively. The age categories were created from the single-year of age variables, with the exception of adults with missing single-year of age who indicated that they were in one of the age categories under 30 years old in the Phase 2 screener; these cases were assigned to the appropriate age category based on that information as described above. All values for respondents, whether imputed or not, are contained in the imputed variables; the variable R01R\_x\_AGECAT#\_IMPFLAG indicates which values are from the interview (R01R\_x\_AGECAT#\_IMPFLAG = 0), from the household screener (R01R\_x\_AGECAT#\_IMPFLAG = 1), or are missing because no further information is available (adults only) (R01R\_x\_AGECAT#\_IMPFLAG = 3), where x is “A” (adult) or “Y” (youth) and # is “2” or “7” as appropriate.

### D.1.3 Education (Adults Only)

Adult respondents were asked questionnaire item R01\_AM0018 regarding the highest level of education they attained. A five-level education variable was used for the adult weighting and created by collapsing the 11 response categories as follows: less than high school or GED; high school graduate; some college but no degree, or associate degree; bachelor’s degree; advanced degree. Missing values were imputed using hot deck imputation because demographic characteristics that are indicative of educational attainment were available.<sup>50</sup> The imputation cells were formed by cross-classifying categories of census region, age, and sex, using imputed values of age and sex (calculated as described in the preceding sections) as appropriate. Neither the education variable described above nor its imputed counterpart is provided on the adult PUF. Instead, a six-level education

---

<sup>50</sup>Hot deck imputation is a method for handling missing data in which each missing value is replaced with an observed response from a “similar” unit in the imputation cell.

variable (R01R\_A\_AM0018) is provided with a separate category for respondents indicating “GED” as their highest level of education.

#### D.1.4 Race

A four-level race variable was used for both the adult and youth weighting. In the extended interview, respondents were asked to indicate which of 14 race categories applied to them (questionnaire items with the prefix R01\_AM0006 for adults and items with the prefix R01\_YM0006 for youth). The responses were combined into a single variable with four categories: White alone, Black or African American alone, Asian alone (including multiple Asian categories), and other (including multi-racial).

If the sampled adult or youth did not respond to the race question, or did not respond to the extended interview, the race information was assigned according to the information provided in the household screener.<sup>51</sup> The race information gathered for the sampled person in the household screener was considered first; if that was not available, the race information that the household screener respondent reported about themselves was used as a proxy for the sampled person.

If none of this information was available, race was imputed using hot deck imputation because demographic characteristics that are indicative of race were available. For adults, imputation cells were formed by cross-classifying categories of census region, age, and education, using imputed values of age and education (calculated as described in the preceding sections) as appropriate; for youth, the hot deck imputation cells were formed by the four categories of census region.

Variables indicating three levels of race (White alone, Black alone, and other) are provided on the PUFs. All values, imputed and unimputed, are contained in the variable R01R\_x\_RACE\_IMP, where  $x$  is “A” (adult) or “Y” (youth). The variables R01R\_A\_RACE\_IMPFLAG and R01R\_Y\_RACE\_CAT3\_IMPFLAG indicate which values are from the interview (R01R\_x\_RACE\_CAT3\_IMPFLAG = 0), from the household screener (R01R\_x\_RACE\_CAT3\_IMPFLAG = 1), or the result of the imputation method described above (R01R\_x\_RACE\_CAT3\_IMPFLAG = 2).

---

<sup>51</sup>There were only five race category responses in the household screener, but it was possible to uniquely code them into the four race categories used in weighting.

## D.1.5 Ethnicity

A two-level ethnicity variable indicating whether the sampled person is of Hispanic origin was used for both the adult and youth weighting. This variable was initially created from the Hispanic origin question asked in the extended interview (questionnaire item R01\_AM0005\_01 for adults and item R01\_YM0005\_01 for youth). The variables R01R\_A\_HISP and R01R\_Y\_HISP indicate whether the respondent is of Hispanic origin based on these variables for adults and youth, respectively.

If the sampled adult or youth did not respond to this question, or did not respond to the extended interview, the ethnicity information was assigned according to the information provided in the household screener: The ethnicity information gathered for the sampled person in the household screener was considered first; if that was not available, the ethnicity information that the household screener respondent reported about themselves was used as a proxy for the sampled person.

If none of this information was available, ethnicity was imputed using hot deck imputation because demographic characteristics that are indicative of ethnicity were available. For adults, imputation cells were formed by cross-classifying categories of census region, age, and education, using imputed values of age and education (calculated as described in the preceding sections) as appropriate. In Wave 1, while the ethnicity variable was imputed for some adults, imputation was not necessary for the youth.

All values for respondents, imputed or not, are contained in the variable R01R\_x\_HISP\_IMP, where  $x$  is “A” (adult) or “Y” (youth). The variables R01R\_A\_HISP\_IMPFLAG and R01R\_Y\_HISP\_IMPFLAG indicate which values are from the interview ( $R01R_x\_HISP\_IMPFLAG = 0$ ), from the household screener ( $R01R_x\_HISP\_IMPFLAG = 1$ ), or the result of the imputation method described above ( $R01R_x\_HISP\_IMPFLAG = 2$ ).

## D.2 Wave 4 Imputation

All Wave 4 respondents had non-missing values for age. The sections below provide the methods used for imputing sex (Section D.2.1), education (Section D.2.2), race (Section D.2.3), and ethnicity (Section D.2.4). The imputation of these Wave 4 characteristics was performed in the order they are presented below except that for the Wave 1 Cohort, race and ethnicity were imputed before

education. This sequential approach allowed the imputed values for earlier variables to be used in the imputation procedure for later variables, thus preserving correlations among the characteristics.

## D.2.1 Sex

The variables R04R\_A\_SEX and R04R\_Y\_SEX contain the sex responses for adults and youth, respectively. All values, imputed and unimputed, are contained in the variable R04R\_x\_SEX\_IMP, where  $x$  is “A” (adult) or “Y” (youth); the variable R04R\_x\_SEX\_IMPFLAG indicates which values are from the interview, i.e., not imputed ( $R04R_x\_SEX\_IMPFLAG = 0$ ), from the household screener ( $R04R_x\_SEX\_IMPFLAG = 1$ ); no statistical imputation was required for these items at Wave 4. The sections below describe the imputation of missing values of sex for the replenishment sample and the Wave 1 Cohort, respectively.

### D.2.1.1 Replenishment Sample

Sex was assigned for adults and youth by first looking at the information provided in the Wave 4 extended interview (questionnaire item R04\_AM0004\_RS for adults and item R04\_YM0004\_NB for youth). If this information was not provided, then sex was assigned according to the information provided in the Wave 4 household screener.

### D.2.1.2 Wave 1 Cohort

Members of the Wave 1 Cohort were asked to self-report their sex at Wave 1 only if they were an adult or youth at that time. This information was not requested from those who were shadow youth at Wave 1 until they completed their first youth interview at a subsequent wave. For Wave 4 adults, sex was assigned by first considering information provided in the Wave 1 extended interview (questionnaire item R01\_AM0004 for adults and item R01\_YM0004 for youth). If sex was still missing, then the Wave 1 imputed value of sex was assigned. For Wave 4 youth, sex was assigned by initially considering information provided in their first youth interview (R01\_YM0004 for Wave 1, R02\_YM0004\_NB for Wave 2, R03\_YM0004\_NB for Wave 3, and R04\_YM0004\_NB for Wave 4). If sex was still missing, then the Wave 1 imputed value of sex was assigned.

## D.2.2 Education (Adults Only)

A four-level education variable was used for the adult weighting and created with the following categories: less than high school or GED; high school graduate; some college but no degree, or associates degree; bachelor's degree and beyond. The sections below describe the imputation of missing values of education for the replenishment sample and the Wave 1 Cohort, respectively.

Neither the education variable described above nor its imputed counterpart is provided on the adult PUF. Instead, a five-level education variable (R04R\_A\_AM0018\_V2) is provided with a separate category for respondents indicating "GED" as their highest level of education.

### D.2.2.1 Replenishment Sample

All adult respondents in the replenishment sample were asked for their highest level of education, with 11 valid response categories presented.<sup>52</sup> Responses were combined into the four categories described above. Missing education values were imputed using hot deck imputation because demographic characteristics that are indicative of educational attainment were available. Imputation cells were formed by cross-classifying categories of census region, age, and sex, using imputed values of sex (calculated as described in the preceding section) as appropriate.

### D.2.2.2 Wave 1 Cohort

The four-level education variable described above was also created for adult respondents from the Wave 1 Cohort, based on the same questionnaire item responses. Missing education values were imputed using hot deck imputation. However, because the educational attainment item was also included in the Waves 1, 2, and 3 adult interviews, available historical information was used in the imputation process. The imputation cells were formed by cross-classifying categories of age, sex, race/ethnicity,<sup>53</sup> imputed highest education level at Wave 1, highest education level at Wave 2, and highest education level at Wave 3, using imputed values of sex and race/ethnicity (calculated as

---

<sup>52</sup>If the adult completed a parent interview prior to their adult interview, this information was captured in questionnaire item R04\_PM0001; otherwise, this information was captured in questionnaire item R04\_AM0018.

<sup>53</sup>For the Wave 1 Cohort, race and ethnicity were imputed prior to the imputation of education.

described in the preceding sections) as appropriate. As a final check, any imputed value less than the highest education level reported in the previous wave was replaced with the previous wave's value.

### D.2.3 Race

A four-level race variable was used for both the adult and youth weighting and created with the following categories: White alone; Black alone; Asian alone, including multiple Asian categories; other, including multi-racial. The sections below describe the imputation of missing values of race for the replenishment sample and the Wave 1 Cohort, respectively.

Variables indicating three levels of race (White alone, Black alone, and other) are provided on the PUFs. The variables R04R\_A\_RACECAT3 and R04R\_Y\_RACECAT3 contain the race responses for adults and youth, respectively. All values, imputed and unimputed, are contained in the variable R04R\_x\_RACECAT3\_IMP, where  $x$  is "A" (adult) or "Y" (youth). The variables R04R\_A\_RACECAT3\_IMPFLAG and R04R\_Y\_RACECAT3\_IMPFLAG indicate which values are from the interview ( $R04R_x_RACECAT3_IMPFLAG = 0$ ), from the household screener ( $R04R_x_RACECAT3_IMPFLAG = 1$ ), or the result of statistical imputation ( $R04R_x_RACECAT3_IMPFLAG = 2$ ).

#### D.2.3.1 Replenishment Sample

In the extended interview, respondents were asked to select as many of the 14 valid race categories as applied to them (questionnaire items with the prefix R04\_AM0006\_RS for adults and items with the prefix R04\_YM0006\_NB for youth). Responses were combined into a single variable with four categories described above.

If the responding adult or youth did not answer the race question, race was assigned according to the information provided in the Wave 4 household screener.<sup>54</sup> The race information gathered for the respondent in the household screener was considered first; if that was not available, the race information that the household screener respondent reported about themselves was used as a proxy for the respondent's race.

---

<sup>54</sup>There were five race categories presented in the household screener but they may be uniquely coded into the four race categories used in weighting.

If none of this information was available, race was imputed using hot deck imputation because demographic characteristics that are indicative of race were available. For adults, imputation cells were formed by cross-classifying categories of census region, age, and education, using imputed values of education (calculated as described in the preceding section) as appropriate; for youth, imputation cells were formed by the four categories of census region.

#### **D.2.3.2 Wave 1 Cohort**

Members of the Wave 1 Cohort were asked to self-report their race at Wave 1 only if they were an adult or youth at that time. This information was not requested from those who were shadow youth at Wave 1 until they completed their first youth interview at a subsequent wave. For Wave 4 adults, race was assigned by first considering information provided in the Wave 1 extended interview (questionnaire items with the prefix R01\_AM0006 for adults and items with the prefix R01\_YM0006 for youth). If race was still missing, then the Wave 1 imputed value of race was assigned. For Wave 4 youth, race was assigned by initially considering information provided in their first youth interview (items with prefix R01\_YM0006 for Wave 1, items with prefix R02\_YM0006\_NB for Wave 2, items with prefix R03\_YM0006\_NB for Wave 3, and items with prefix R04\_YM0006\_NB for Wave 4). If race was still missing, then the Wave 1 imputed value of race was assigned.

#### **D.2.4 Ethnicity**

The variables R04R\_A\_HISP and R04R\_Y\_HISP contain the ethnicity responses for adults and youth, respectively. All values for respondents, imputed or not, are contained in the variable R04R\_x\_HISP\_IMP, where  $x$  is “A” (adult) or “Y” (youth). The variables R04R\_A\_HISP\_IMPFLAG and R04R\_Y\_HISP\_IMPFLAG indicate which values are from the interview ( $R04R_x\_HISP\_IMPFLAG = 0$ ), from the household screener ( $R04R_x\_HISP\_IMPFLAG = 1$ ), or the result of statistical imputation ( $R04R_x\_HISP\_IMPFLAG = 2$ ). The sections below describe the imputation of missing values of ethnicity for the replenishment sample and the Wave 1 Cohort, respectively.

#### D.2.4.1 Replenishment Sample

In the extended interview, respondents were asked to select as many of the five valid ethnicity categories as applied to them (questionnaire items with the prefix R04\_AM0005\_RS for adults and items with the prefix R04\_YM0005\_NB for youth). Responses were combined into a single variable with two categories (Hispanic, not Hispanic) for weighting purposes.

If the responding adult or youth did not answer the ethnicity question, ethnicity was assigned according to the information provided in the Wave 4 household screener. The ethnicity information gathered for the respondent in the household screener was considered first; if that was not available, the ethnicity information that the household screener respondent reported about themselves was used as a proxy for the respondent's ethnicity.

#### D.2.4.2 Wave 1 Cohort

Members of the Wave 1 Cohort were asked to self-report their ethnicity at Wave 1 only if they were an adult or youth at that time. This information was not requested from those who were shadow youth at Wave 1 until they completed their first youth interview at a subsequent wave. For Wave 4 adults, ethnicity was assigned by first considering information provided in the Wave 1 extended interview (questionnaire items with the prefix R01\_AM0005 for adults and items with the prefix R01\_YM0005 for youth). If ethnicity was still missing, then the Wave 1 imputed value of ethnicity was assigned. For Wave 4 youth, ethnicity was assigned by initially considering information provided in their first youth interview (items with prefix R01\_YM0005 for Wave 1, items with prefix R02\_YM0005\_NB for Wave 2, items with prefix R03\_YM0005\_NB for Wave 3, and items with prefix R04\_YM0005\_NB for Wave 4). If ethnicity was still missing, then the Wave 1 imputed value of ethnicity was assigned.

### D.3 Wave 7 Imputation

All Wave 7 respondents had non-missing values for age. The sections below provide the methods used for imputing sex (Section D.3.1), education (Section D.3.2), race (Section D.3.3), and ethnicity (Section D.3.4). The imputation of these Wave 7 characteristics was performed in the order they are presented below except that for members of the Wave 1 and/or Wave 4 Cohorts, race and ethnicity values were imputed before education, using the results of the imputation procedures described in

Sections D.1 and D.2. This sequential approach allowed the imputed values for earlier variables to be used in the imputation procedure for later variables, thus preserving correlations among the characteristics.

### **D.3.1 Sex**

The variables R07R\_A\_SEX and R07R\_Y\_SEX contain the sex responses for adults and youth, respectively. All values, imputed and unimputed, are contained in the variable R07R\_x\_SEX\_IMP, where  $x$  is “A” (adult) or “Y” (youth); the variable R07R\_x\_SEX\_IMPFLAG indicates which values are from the interview, i.e., not imputed ( $R07R_x\_SEX\_IMPFLAG = 0$ ) or from the household screener ( $R07R_x\_SEX\_IMPFLAG = 1$ ); no statistical imputation was required for these items at Wave 7. The sections below describe the imputation of missing values of sex for the replenishment sample and members of the earlier cohorts, respectively.

#### **D.3.1.1 Wave 7 Replenishment Sample**

Sex was assigned for adults and youth by first looking to the information provided in the Wave 7 extended interview (questionnaire item R07\_AM0004\_RS for adults and item R07\_YM0004\_NB for youth). If this information was not provided, then sex was assigned according to the information provided in the Wave 7 household screener.

#### **D.3.1.2 Wave 1 and/or Wave 4 Cohorts**

Members of the Wave 1 and/or Wave 4 Cohorts were asked to self-report their sex at their first adult or youth interview. For Wave 7 respondents, their first interview may have happened at any wave, depending on when they were selected, their age at that time, and the number of waves missed due to nonresponse (if any). If the respondent did not report their sex at their first interview, then the Wave 4 imputed value of sex (described in Section D.2.1.2) was assigned. If sex was still missing, then the Wave 1 imputed value of sex created for the Wave 1 weighting process (described in Section D.1.1) was assigned.

## D.3.2 Education (Adults Only)

A four-level education variable was used for the adult weighting and created with the following categories: less than high school or GED; high school graduate; some college but no degree, or associates degree; bachelor's degree and beyond. The sections below describe the imputation of missing values of education for the replenishment sample and the Wave 1 and 4 Cohorts, respectively.

Neither the education variable described above nor its imputed counterpart is provided on the adult PUF. Instead, a five-level education variable (R07R\_A\_AM0018\_V2) is provided with a separate category for respondents indicating "GED" as their highest level of education.

### D.3.2.1 Wave 7 Replenishment Sample

All Wave 7 adult respondents were asked a question regarding their highest level of education, with 11 valid response categories presented.<sup>55</sup> Responses were combined into the four categories described above. Missing values were imputed using hot deck imputation because demographic characteristics that are indicative of educational attainment were available. The imputation cells were formed by cross-classifying categories of census region (Northeast, Midwest, South, West), age, and sex, using imputed values (calculated as described above) as appropriate.

### D.3.2.2 Wave 1 and/or Wave 4 Cohorts

The four-level education variable described above was also created for adult respondents from the Wave 1 and/or Wave 4 Cohorts, based on the same questionnaire item responses. Missing values were imputed using hot deck imputation. However, because the educational attainment item was also included in all previous waves' adult interviews, available historical information was used in the imputation process. The imputation cells were formed by cross-classifying categories of age, sex, race/ethnicity,<sup>56</sup> imputed highest education level from Wave 4, and highest education level at Wave 5 and/or Wave 6, using imputed values of age, sex, and race/ethnicity (calculated as described in the

---

<sup>55</sup>If the adult completed a parent interview prior to their adult interview, this information was captured in questionnaire item R07\_PM0001; otherwise, this information was captured in questionnaire item R07\_AM0018.

<sup>56</sup>For the Wave 1 and/or Wave 4 Cohorts, race and ethnicity were imputed prior to the imputation of education.

preceding sections) as appropriate. As a final check, any imputed value less than the highest education level reported in the previous wave was replaced with the previous wave's value.

### D.3.3 Race

A four-level race variable was used for both the adult and youth weighting and created with the following categories: White alone; Black alone; Asian alone, including multiple Asian categories; other, including multi-racial. The sections below describe the imputation of missing values of race for the replenishment sample and the Wave 1 and 4 Cohorts, respectively.

Variables indicating three levels of race (White alone, Black alone, and other) are provided on the PUFs. The variables R07R\_A\_RACECAT3 and R07R\_Y\_RACECAT3 contain the race responses for adults and youth, respectively. All values, imputed and unimputed, are contained in the variable R07R\_x\_RACECAT3\_IMP, where  $x$  is "A" (adult) or "Y" (youth). The variables R07R\_A\_RACECAT3\_IMPFLAG and R07R\_Y\_RACECAT3\_IMPFLAG indicate which values are from the interview (R07R\_x\_RACECAT3\_IMPFLAG = 0), from the household screener (R07R\_x\_RACECAT3\_IMPFLAG = 1), or the result of statistical imputation (R07R\_x\_RACECAT3\_IMPFLAG = 2).

#### D.3.3.1 Wave 7 Replenishment Sample

In the extended interview, respondents were asked to select as many of the 14 valid race categories as applied to them (questionnaire items with the prefix R07\_AM0006\_RS for adults and items with the prefix R07\_YM0006\_NB for youth). Responses were combined into a single variable with four categories described above.

If the responding adult or youth did not answer the race question, race was assigned according to the information provided in the Wave 7 household screener.<sup>57</sup> The race information gathered for the respondent in the household screener was considered first; if that was not available, the race information that the household screener respondent reported about themselves was used as a proxy for the respondent's race.

---

<sup>57</sup>There were five race categories presented in the household screener but they may be uniquely coded into the four race categories used in weighting.

If none of this information was available, race was imputed using hot deck imputation because demographic characteristics that are indicative of race were available. For adults, imputation cells were formed by cross-classifying categories of census region, age, and education, using imputed values (calculated as described in the preceding sections) as appropriate; for youth, the imputation cells were the four categories of census region.

### **D.3.3.2 Wave 1 and/or Wave 4 Cohorts**

Members of the Wave 1 and/or Wave 4 Cohorts were asked to self-report their race at their first adult or youth interview. For Wave 7 respondents, their first interview may have happened at any wave, depending on when they were selected, their age at that time, and the number of waves missed due to nonresponse (if any). If the respondent did not report their race at their first interview, then the Wave 4 imputed value of race (described in Section D.2.3.2) was assigned. If race was still missing, then the Wave 1 imputed value of race created for the Wave 1 weighting process (described in Section D.1.4) was assigned.

## **D.3.4 Ethnicity**

The variables R07R\_A\_HISP and R07R\_Y\_HISP contain the ethnicity responses for adults and youth, respectively. All values for respondents, imputed or not, are contained in the variable R07R\_x\_HISP\_IMP, where  $x$  is “A” (adult) or “Y” (youth). The variables R07R\_A\_HISP\_IMPFLAG and R07R\_Y\_HISP\_IMPFLAG indicate which values are from the interview ( $R07R_x\_HISP\_IMPFLAG = 0$ ), from the household screener ( $R07R_x\_HISP\_IMPFLAG = 1$ ), or the result of statistical imputation ( $R07R_x\_HISP\_IMPFLAG = 2$ ). The sections below describe the imputation of missing values of ethnicity for the replenishment sample and the earlier cohorts, respectively.

### **D.3.4.1 Wave 7 Replenishment Sample**

In the extended interview, respondents were asked to select as many of the five valid ethnicity categories as applied to them (questionnaire items with the prefix R07\_AM0005\_RS for adults and items with the prefix R07\_YM0005\_NB for youth). Responses were combined into a single variable with two categories (Hispanic, not Hispanic) for weighting purposes.

If the responding adult or youth did not answer the ethnicity question, ethnicity was assigned according to the information provided in the Wave 7 household screener. The ethnicity information gathered for the respondent in the household screener was considered first; if that was not available, the ethnicity information that the household screener respondent reported about themselves was used as a proxy for the respondent's ethnicity. If ethnicity was not available from these sources, ethnicity was imputed using hot deck imputation with imputation cells formed by cross-classifying categories of census region, age, and education, using imputed values (calculated as described in the preceding sections) as appropriate.

#### **D.3.4.2    Wave 1 and/or Wave 4 Cohorts**

Members of the Wave 1 and/or Wave 4 Cohorts were asked to self-report their ethnicity at their first adult or youth interview. For Wave 7 respondents, their first interview may have happened at any wave, depending on when they were selected, their age at that time, and the number of waves missed due to nonresponse (if any). If the respondent did not report their ethnicity at their first interview, then the Wave 4 imputed value of ethnicity (described in Section D.2.4.2) was assigned. If ethnicity was still missing, then the Wave 1 imputed value of ethnicity created for the Wave 1 weighting process (described in Section D.1.5) was assigned.

## Appendix E

### Excluded Questionnaire Variables

---

Table E-1. Wave 1 questionnaire variables entirely excluded from the PUFs

Questionnaire variable	Instrument	Questionnaire variable description
R01_AC1033MC_OS	Adult	Retail location where your cigarette is purchased most of the time – specify
R01_AC1033RY_OS	Adult	Retail location where your roll your own cigarette tobacco is purchased most of the time – specify
R01_AC1049MC	Adult	Brand of cigarettes usually/last smoked – specify
R01_AC1049RY	Adult	Brand of roll-your-own cigarette tobacco usually/last smoked – specify
R01_AC1071MC	Adult	Sub-brand of cigarette product usually/last smoked – specify
R01_AC1071RY	Adult	Name of roll-your-own cigarette tobacco product usually/last smoked – specify
R01_AD1033_OS	Adult	Retail location where your dissolvable tobacco is purchased most of the time – specify
R01_AD1049	Adult	Brand of dissolvable tobacco usually/last used – specify
R01_AD1071	Adult	Sub-brand of dissolvable tobacco usually/last smoked – specify
R01_AE1033_OS	Adult	Retail location where your e-cigarettes and e-liquid/e-cigarette cartridges is purchased most of the time – specify
R01_AE1049	Adult	Brand of e-cigarettes usually/last used – specify
R01_AE9003	Adult	Brand of e-cigarette owned – specify
R01_AG1033CG_OS	Adult	Retail location where your cigarillos/cigarillos for blunts is purchased most of the time – specify
R01_AG1033FC_OS	Adult	Retail location where your filtered cigars/filtered cigars for blunts is purchased most of the time – specify
R01_AG1033TC_OS	Adult	Retail location where your traditional cigars are purchased most of the time – specify
R01_AG1049CG	Adult	Brand of cigarillos usually/last smoked – specify
R01_AG1049FC	Adult	Brand of filtered cigars usually/last smoked – specify
R01_AG1049TC	Adult	Brand of traditional cigars usually/last smoked – specify
R01_AG1071CG	Adult	Sub-brand of cigarillo product usually/last smoked – specify
R01_AG1071FC	Adult	Sub-brand of filtered cigar product usually/last smoked – specify
R01_AG1071TC	Adult	Sub-brand of traditional cigar product usually/last smoked – specify
R01_AH1033_OS	Adult	Retail location where your hookah tobacco is purchased most of the time – specify
R01_AH1049	Adult	Brand of hookah tobacco usually/last smoked – specify
R01_AH1071	Adult	Sub-brand of hookah tobacco product usually/last smoked – specify
R01_AH9011_OS	Adult	Place where usually smoke/smoked a hookah: Somewhere else – specify
R01_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R01_AM0008	Adult	Currently on active duty in the U.S. Armed Forces

**Table E-1. Wave 1 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R01_AM0010_01	Adult	When served on active duty in the U.S. Armed Forces: September 2001 or later
R01_AM0010_02	Adult	When served on active duty in the U.S. Armed Forces: August 1990 to August 2001 (including Persian Gulf War)
R01_AM0010_03	Adult	When served on active duty in the U.S. Armed Forces: September 1980 to July 1990
R01_AM0010_04	Adult	When served on active duty in the U.S. Armed Forces: May 1975 to August 1980
R01_AM0010_05	Adult	When served on active duty in the U.S. Armed Forces: Vietnam era (August 1964 to April 1975)
R01_AM0010_06	Adult	When served on active duty in the U.S. Armed Forces: March 1961 to July 1964
R01_AM0010_07	Adult	When served on active duty in the U.S. Armed Forces: February 1955 to February 1961
R01_AM0010_08	Adult	When served on active duty in the U.S. Armed Forces: Korean War (July 1950 to January 1955)
R01_AM0010_09	Adult	When served on active duty in the U.S. Armed Forces: January 1947 to June 1950
R01_AM0010_10	Adult	When served on active duty in the U.S. Armed Forces: World War II (December 1941 to December 1946)
R01_AM0010_11	Adult	When served on active duty in the U.S. Armed Forces: November 1941 or earlier
R01_AM0011_01	Adult	Branch served when on active duty: Army
R01_AM0011_02	Adult	Branch served when on active duty: Navy
R01_AM0011_03	Adult	Branch served when on active duty: Air Force
R01_AM0011_04	Adult	Branch served when on active duty: Marine Corps
R01_AM0011_05	Adult	Branch served when on active duty: Coast Guard
R01_AM0012	Adult	Ever been enrolled in VA Health Care
R01_AM0013_01	Adult	Received any VA Health Care Benefits: I have received services at the VA
R01_AM0013_02	Adult	Received any VA Health Care Benefits: VA has paid for some or all of my health care
R01_AM0013_03	Adult	Received any VA Health Care Benefits: I have not received any VA health care benefits
R01_AM0016	Adult	Main job title or occupation
R01_AM0021	Adult	Sexual attraction to gender
R01_AM0025	Adult	Sexual orientation – specify
R01_AM0050	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard (Confirmation)
R01_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time – specify
R01_AP1049	Adult	Brand of pipe tobacco usually/last smoked – specify
R01_AP1071	Adult	Sub-brand of pipe tobacco product usually/last smoked – specify
R01_AS1033SM_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time – specify
R01_AS1033SU_OS	Adult	Retail location where your snus pouches is purchased most of the time – specify
R01_AS1049SM	Adult	Brand of smokeless tobacco usually/last used – specify

**Table E-1. Wave 1 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R01_AS1049SU	Adult	Brand of snus pouch usually/last used – specify
R01_AS1071SM	Adult	Sub-brand of smokeless tobacco product usually/last smoked – specify
R01_AS1071SU	Adult	Sub-brand of snus pouch product usually/last smoked – specify
R01_AX0190	Adult	Is deaf or has serious difficulty hearing
R01_AX0191	Adult	Is blind or has serious difficulty seeing, even when wearing glasses
R01_AX0197_OS	Adult	Brand of favorite tobacco advertisement – specify
R01_AX0217_10A	Adult	Tobacco Product 10, bar code scan
R01_AX0217_1A	Adult	Tobacco Product 1, bar code scan
R01_AX0217_2A	Adult	Tobacco Product 2, bar code scan
R01_AX0217_3A	Adult	Tobacco Product 3, bar code scan
R01_AX0217_4A	Adult	Tobacco Product 4, bar code scan
R01_AX0217_5A	Adult	Tobacco Product 5, bar code scan
R01_AX0217_6A	Adult	Tobacco Product 6, bar code scan
R01_AX0217_7A	Adult	Tobacco Product 7, bar code scan
R01_AX0217_8A	Adult	Tobacco Product 8, bar code scan
R01_AX0217_9A	Adult	Tobacco Product 9, bar code scan
R01_AX0309	Adult	Calendar year of most recent pregnancy
R01_LAD01	Adult	Respondent's preferred language to complete ACASI interview
R01_LCAD01	Adult	Language in which CAPI portions of adult interview were conducted
R01_CPT07	Parent	Youth date of birth
R01_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
R01_PM0016	Parent	Youth was identified as a twin or part of a multiple birth
R01_PM0017	Parent	Youth and sibling are identical twins
R01_PM0019	Parent	Any sibling in multiple birth identical to youth
R01_PM0021	Parent	First name of sibling that youth is a twin of
R01_PM0030	Parent	First names of siblings in multiple birth that are identical to youth
R01_PM0035	Parent	First names of siblings that are in the multiple birth with youth
R01_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative – specify
R01_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative – specify
R01_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative – specify
R01_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative – specify
R01_PX0186	Parent	Youth has serious difficulty walking or climbing stairs
R01_PX0190	Parent	Youth is deaf or has serious difficulty hearing
R01_PX0191	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
R01_EYOUTH	Youth	Respondent is an emancipated youth
R01LEY01	Youth	Emancipated Youth's preferred language to complete ACASI interview
R01_LYH01	Youth	Youth's preferred language to complete ACASI interview
R01_YB1118BD_OS	Youth	How you usually got your own bidis in the past 30 days – specify

**Table E-1. Wave 1 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R01_YB1033BD_OS	Youth	Retail location where your bidis are purchased most of the time – specify
R01_YB1118KK_OS	Youth	How you usually got your own kreteks in the past 30 days – specify
R01_YB1033KK_OS	Youth	Retail location where your kreteks are purchased most of the time – specify
R01_YB1049KK	Youth	Brand of kreteks usually/last smoked – specify
R01_YB1071KK	Youth	Sub-brand of kretek product usually/last smoked – specify
R01_YC1118_OS	Youth	How you usually got your own cigarettes in the past 30 days – specify
R01_YC1033_OS	Youth	Retail location where your cigarettes are purchased most of the time – specify
R01_YC1049	Youth	Brand of cigarettes usually/last smoked – specify
R01_YC1071	Youth	Sub-brand of cigarette product usually/last smoked – specify
R01_YB1049BD	Youth	Brand of bidis usually/last smoked – specify
R01_YD1118_OS	Youth	How you usually got your own dissolvable tobacco in the past 30 days – specify
R01_YD1033_OS	Youth	Retail location where you usually buy dissolvable tobacco most of the time – specify
R01_YD1049	Youth	Brand of dissolvable tobacco usually/last used – specify
R01_YD1071	Youth	Sub-brand of dissolvable tobacco product usually/last smoked – specify
R01_YE1118_OS	Youth	How you usually got your own e-cigarettes/cartridges and e-liquid in the past 30 days – specify
R01_YE1033_OS	Youth	Retail location where your e-cigarettes/cartridges and e-liquid are purchased most of the time – specify
R01_YE1049	Youth	Brand of e-cigarettes usually/last used – specify
R01_YG1118CL_OS	Youth	How you usually got your own cigarillos in the past 30 days – specify
R01_YG1033CL_OS	Youth	Retail location where your cigarillos are purchased most of the time – specify
R01_YG1118FC_OS	Youth	How you usually got your own filtered cigars in the past 30 days – specify
R01_YG1033FC_OS	Youth	Retail location where your filtered cigars are purchased most of the time – specify
R01_YG1118TC_OS	Youth	How you usually got your own traditional cigars in the past 30 days – specify
R01_YG1033TC_OS	Youth	Retail location where your traditional cigars are purchased most of the time – specify
R01_YG1049CL	Youth	Brand of cigarillos usually/last smoked – specify
R01_YG1049FC	Youth	Brand of filtered cigars usually/last smoked – specify
R01_YG1049TC	Youth	Brand of traditional cigars usually/last smoked – specify
R01_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked – specify
R01_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked – specify
R01_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked – specify
R01_YH9011_OS	Youth	Usually smoke a hookah: Somewhere else – specify
R01_YH1118_OS	Youth	How you usually got your own shisha or hookah tobacco in the past 30 days – specify

**Table E-1. Wave 1 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R01_YH1033_OS	Youth	Retail location where your hookah tobacco is purchased most of the time – specify
R01_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
R01_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
R01_YM0020	Youth	Last grade/year in school completed
R01_YM0021	Youth	Sexual attraction to gender
R01_YM0022	Youth	Sexual orientation (initial prompt)
R01_YM0023	Youth	Sexual orientation – other description
R01_YM0028	Youth	Sexual orientation (second prompt)
R01_YM0024	Youth	Reason for not providing sexual orientation
R01_YM0025	Youth	Sexual orientation – specify
R01_YP1118_OS	Youth	How you usually got your own pipe tobacco in the past 30 days – specify
R01_YP1033_OS	Youth	Retail location where your pipe tobacco are purchased most of the time – specify
R01_YP1049	Youth	Brand of pipe tobacco usually/last smoked – specify
R01_YP1071	Youth	Sub-brand of pipe tobacco product usually/last smoked – specify
R01_YS1118SM_OS	Youth	How you usually got your own smokeless tobacco in the past 30 days – specify
R01_YS1033SM_OS	Youth	Retail location where your smokeless tobacco is purchased most of the time – specify
R01_YS1118SU_OS	Youth	How you usually got your own snus pouches in the past 30 days – specify
R01_YS1033SU_OS	Youth	Retail location where your snus pouches are purchased most of the time – specify
R01_YS1049SM	Youth	Brand of smokeless tobacco usually/last used – specify
R01_YS1049SU	Youth	Brand of snus pouches usually/last used – specify
R01_YS1071SM	Youth	Sub-brand of smokeless tobacco product usually/last smoked – specify
R01_YS1071SU	Youth	Sub-brand of snus pouch product usually/last smoked – specify
R01_YX0136	Youth	Currently pregnant
R01_YX0137_NN	Youth	Number of weeks/months pregnant – Number
R01_YX0137_UN	Youth	Number of weeks/months pregnant – Unit
R01_YX0186	Youth	Has serious difficulty walking or climbing stairs
R01_YX0190	Youth	Is deaf or has serious difficulty hearing
R01_YX0191	Youth	Is blind or has serious difficulty seeing, even when wearing glasses
R01_YX0197_OS	Youth	Brand of favorite tobacco advertisement – specify
R01_YX0478_OS	Youth	Received discount coupons from: Some other way – specify
R01_YX0480_OS	Youth	How tobacco company sent you the information other than coupons: Some other way – specify
R01 YY0601_OS	Youth	First type of tobacco you tried – specify

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_AM0001 R01_AM0002 R01_AM0003	Adult	R01R_A_AGECAT7	Age range when interviewed (7 levels)
R01_AM0005_01 R01_AM0005_02 R01_AM0005_03 R01_AM0005_04 R01_AM0005_05	Adult	R01R_A_HISP	Hispanic origin from the interview (2 levels)
R01_AM0006_WH R01_AM0006_AA R01_AM0006_AI R01_AM0006_IN R01_AM0006_CH R01_AM0006_FI R01_AM0006_JA R01_AM0006_KO R01_AM0006_VI R01_AM0006_OA R01_AM0006_HI R01_AM0006_GU R01_AM0006_SA R01_AM0006_OP	Adult	R01R_A_RACECAT3	Race from the interview (3 levels)
R01_AM0018	Adult	R01R_A_AM0018	Highest grade or level of school completed (6 levels)
R01_AM0030	Adult	R01R_A_AM0030	Total household income in the past 12 months (5 levels)
R01_AX0135_01 R01_AX0135_02 R01_AX0135_03 R01_AX0135_04 R01_AX0135_05	Adult	R01R_A_AX0135	Ever had pregnancy outcome of miscarriages, induced abortions, ectopic or tubal pregnancies or stillbirths (2 levels)
R01_AX0300_01 R01_AX0300_02 R01_AX0300_03 R01_AX0300_04 R01_AX0300_05 R01_AX0300_06 R01_AX0300_07	Adult	R01R_A_AX0300	Indicator of Ever had outcome for live birth of Preterm birth, low birth weight, birth defects, Placenta Previa, Placenta Abruptio, or Pre-eclampsia (2 levels)
R01_AC1006 R01_AC1120	Adult	R01R_A_AC1006	Age range when first smoked part or all of a cigarette (6 levels)
R01_AC1007 R01_AC1121	Adult	R01R_A_AC1007	Age range when first started smoking cigarettes fairly regularly (6 levels)
R01_AC1020 R01_AC1122	Adult	R01R_A_AC1020	Age range when first started smoking cigarettes every day (6 levels)
R01_AD1006 R01_AD1120	Adult	R01R_A_AD1006	Age range when first used dissolvable tobacco, even one or two times (6 levels)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_AD1007 R01_AD1121	Adult	R01R_A_AD1007	Age range when first started using dissolvable tobacco fairly regularly (6 levels)
R01_AD1020 R01_AD1122	Adult	R01R_A_AD1020	Age range when first started using dissolvable tobacco every day (6 levels)
R01_AE1006 R01_AE1120	Adult	R01R_A_AE1006	Age range when first time used an e-cigarette, even one or two times (6 levels)
R01_AE1007 R01_AE1121	Adult	R01R_A_AE1007	Age range when first started using e-cigarettes fairly regularly (6 levels)
R01_AE1020 R01_AE1122	Adult	R01R_A_AE1020	Age range when first started using e-cigarettes every day (6 levels)
R01_AG1006CG R01_AG1120CG	Adult	R01R_A_AG1006CG	Age range when first time smoked part or all of a cigarillo, even one or two puffs (6 levels)
R01_AG1006FC R01_AG1120FC	Adult	R01R_A_AG1006FC	Age range when first smoked part or all of a filtered cigar, even one or two puffs (6 levels)
R01_AG1006TC R01_AG1120TC	Adult	R01R_A_AG1006TC	Age range when first smoked part or all of a traditional cigar, even one or two puffs (6 levels)
R01_AG1007CG R01_AG1121CG	Adult	R01R_A_AG1007CG	Age range when first started smoking cigarillos fairly regularly (6 levels)
R01_AG1007FC R01_AG1121FC	Adult	R01R_A_AG1007FC	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R01_AG1007TC R01_AG1121TC	Adult	R01R_A_AG1007TC	Age range when first started smoking traditional cigars fairly regularly (6 levels)
R01_AG1020CG R01_AG1122CG	Adult	R01R_A_AG1020CG	Age range when first started smoking cigarillos every day (6 levels)
R01_AG1020FC R01_AG1122FC	Adult	R01R_A_AG1020FC	Age range when first started smoking filtered cigars every day (6 levels)
R01_AG1020TC R01_AG1122TC	Adult	R01R_A_AG1020TC	Age range when first started smoking traditional cigars every day (6 levels)
R01_AH1006 R01_AH1120	Adult	R01R_A_AH1006	Age range when first smoked hookah, even one or two puffs (6 levels)
R01_AH1007 R01_AH1121	Adult	R01R_A_AH1007	Age range when first started smoking hookah fairly regularly (6 levels)
R01_AH1020 R01_AH1122	Adult	R01R_A_AH1020	Age range when first started smoking hookah every day (6 levels)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_AP1006 R01_AP1120	Adult	R01R_A_AP1006	Age range when first smoked part or all of a pipe filled with tobacco, even one or two puffs (6 levels)
R01_AP1007 R01_AP1121	Adult	R01R_A_AP1007	Age range when first started smoking a pipe filled with tobacco fairly regularly (6 levels)
R01_AP1020 R01_AP1122	Adult	R01R_A_AP1020	Age range when first started smoking a pipe filled with tobacco every day (6 levels)
R01_AS1006SM R01_AS1120SM	Adult	R01R_A_AS1006SM	Age range when first used smokeless tobacco, even one or two times (6 levels)
R01_AS1006SU R01_AS1120SU	Adult	R01R_A_AS1006SU	Age range when first used snus pouches, even one or two times (6 levels)
R01_AS1007SM R01_AS1121SM	Adult	R01R_A_AS1007SM	Age range when first started using smokeless tobacco fairly regularly (6 levels)
R01_AS1007SU R01_AS1121SU	Adult	R01R_A_AS1007SU	Age range when first started using snus pouches fairly regularly (6 levels)
R01_AS1020SM R01_AS1122SM	Adult	R01R_A_AS1020SM	Age range when first started using smokeless tobacco every day (6 levels)
R01_AS1020SU R01_AS1122SU	Adult	R01R_A_AS1020SU	Age range when first started using snus pouches every day (6 levels)
R01_AX0074 R01_AX0270	Adult	R01R_A_AX0074	Age range when first alcoholic drink was consumed (6 levels)
R01_AX0079 R01_AX0271	Adult	R01R_A_AX0079	Age range when first used marijuana, hash, THC or grass (6 levels)
R01_AX0082_01 R01_AX0272_01	Adult	R01R_A_AX0082_01	Age range when first started using: Ritalin or Adderall (6 levels)
R01_AX0082_02 R01_AX0272_02	Adult	R01R_A_AX0082_02	Age range when first started using: Painkillers, sedatives or tranquilizers (6 levels)
R01_AX0082_03 R01_AX0272_03	Adult	R01R_A_AX0082_03	Age range when first started using: Cocaine or crack (6 levels)
R01_AX0082_04 R01_AX0272_04	Adult	R01R_A_AX0082_04	Age range when first started using: Stimulants like methamphetamine or speed (6 levels)
R01_AX0082_05 R01_AX0272_05	Adult	R01R_A_AX0082_05	Age range when first started using: Any other drugs like heroin, inhalants, solvents, or hallucinogens (6 levels)
R01_AX0086 R01_AX0087	Adult	R01R_A_AX0086	Age range when first drank alcohol at all, counting small tastes or sips (6 levels)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_AX0112 R01_AX0252	Adult	R01R_A_AX0112	Age range when you were first told you had a heart attack (6 levels)
R01_AX0114 R01_AX0253	Adult	R01R_A_AX0114	Age range when you were first told you had high blood pressure (6 levels)
R01_AX0115 R01_AX0254	Adult	R01R_A_AX0115	Age range when you were first told you had high cholesterol (6 levels)
R01_AX0116 R01_AX0255	Adult	R01R_A_AX0116	Age range when you were first told you had congestive heart failure (6 levels)
R01_AX0117 R01_AX0256	Adult	R01R_A_AX0117	Age range when you were first told you had a stroke (6 levels)
R01_AX0120 R01_AX0257	Adult	R01R_A_AX0120	Age range when you were first told you had COPD (6 levels)
R01_AX0121 R01_AX0258	Adult	R01R_A_AX0121	Age range when you were first told you had chronic bronchitis (6 levels)
R01_AX0123 R01_AX0259	Adult	R01R_A_AX0123	Age range when you were first told you had emphysema (6 levels)
R01_AX0124 R01_AX0260	Adult	R01R_A_AX0124	Age range when you were first told you had asthma (6 levels)
R01_AX0131 R01_AX0261	Adult	R01R_A_AX0131	Age range when you were first told you had gum disease (6 levels)
R01_AX0133 R01_AX0262	Adult	R01R_A_AX0133	Age range when you were first told you had pre-cancerous oral lesions (6 levels)
R01_AX0143 R01_AX0264	Adult	R01R_A_AX0143	Age range when you were first told you had an ulcer (6 levels)
R01_AX0148 R01_AX0266	Adult	R01R_A_AX0148	Age range when you were first told you had stomach or gastro-intestinal bleeding (6 levels)
R01_AX0150 R01_AX0267	Adult	R01R_A_AX0150	Age range when you were first told you had osteoporosis (6 levels)
R01_AX0152 R01_AX0269	Adult	R01R_A_AX0152	Age range when you were first told you had a cataract or glaucoma (6 levels)
R01_AX0198 R01_AX0268	Adult	R01R_A_AX0198	Age range when you were first told you had a bone fracture because you have fragile bones (6 levels)
R01_AX0280 R01_AX0263	Adult	R01R_A_AX0280	Age range when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes (6 levels)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_AM0026_01 R01_AM0026_02 R01_AM0026_03 R01_AM0026_04 R01_AM0026_05 R01_AM0026_06 R01_AM0026_07 R01_AM0026_08	Adult	R01R_A_AM0026	Currently covered by health insurance or health coverage plan (5 levels)
R01_AX0308	Adult	R01R_A_AX0308	Number of pregnancies that have resulted in a live birth (2 levels)
R01_AX0310 R01_AX0108_FT R01_AX0108_IN R01_AX0311 R01_AX0109 R01_AX0312	Adult	R01R_A_BMI	Body mass index
R01_AX0145_01 R01_AX0145_06 R01_AX0145_07 R01_AX0145_08 R01_AX0145_10 R01_AX0145_11 R01_AX0145_13 R01_AX0145_14 R01_AX0145_17 R01_AX0145_20 R01_AX0145_22 R01_AX0145_26 R01_AX0145_28	Adult	R01R_A_AX0145_TOB	Type of cancer is a tobacco related cancer

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_AX0146_01	Adult	R01R_A_AX0146_TOB	Age when first tobacco-related cancer was diagnosed
R01_AX0146_06			
R01_AX0146_07			
R01_AX0146_08			
R01_AX0146_10			
R01_AX0146_11			
R01_AX0146_13			
R01_AX0146_14			
R01_AX0146_17			
R01_AX0146_20			
R01_AX0146_22			
R01_AX0146_26			
R01_AX0146_28			
R01_AX0265_01			
R01_AX0265_06			
R01_AX0265_07			
R01_AX0265_08			
R01_AX0265_10			
R01_AX0265_11			
R01_AX0265_13			
R01_AX0265_14			
R01_AX0265_17			
R01_AX0265_20			
R01_AX0265_22			
R01_AX0265_26			
R01_AX0265_28			
R01_AX0145_02	Adult	R01R_A_AX0145_NONTOB	Type of cancer is a non-tobacco related cancer
R01_AX0145_03			
R01_AX0145_04			
R01_AX0145_05			
R01_AX0145_09			
R01_AX0145_12			
R01_AX0145_15			
R01_AX0145_16			
R01_AX0145_18			
R01_AX0145_19			
R01_AX0145_21			
R01_AX0145_23			
R01_AX0145_24			
R01_AX0145_25			
R01_AX0145_27			
R01_AX0145_29			
R01_AX0145_30			
R01_AX0145_31			

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_AX0146_02	Adult	R01R_A_AX0146_NONTOB	Age when first non-tobacco-related cancer was diagnosed
R01_AX0146_03			
R01_AX0146_04			
R01_AX0146_05			
R01_AX0146_09			
R01_AX0146_12			
R01_AX0146_15			
R01_AX0146_16			
R01_AX0146_18			
R01_AX0146_19			
R01_AX0146_21			
R01_AX0146_23			
R01_AX0146_24			
R01_AX0146_25			
R01_AX0146_27			
R01_AX0146_29			
R01_AX0146_30			
R01_AX0146_31			
R01_AX0265_02			
R01_AX0265_03			
R01_AX0265_04			
R01_AX0265_05			
R01_AX0265_09			
R01_AX0265_12			
R01_AX0265_15			
R01_AX0265_16			
R01_AX0265_18			
R01_AX0265_19			
R01_AX0265_21			
R01_AX0265_23			
R01_AX0265_24			
R01_AX0265_25			
R01_AX0265_27			
R01_AX0265_29			
R01_AX0265_30			
R01_AX0265_31			
R01_AM0022	Adult	R01R_A_SEXORIENT2	Adult sexual orientation (2 levels)
R01_AM0023			
R01_AM0024			
R01_AM0028			
R01_PT0001	Parent	R01R_Y_PT0001	Parent or guardian relationship to youth (4 levels)
R01_PT0002	Parent	R01R_Y_PT0002	Parent's spouse or partner relationship to youth (4 levels)
R01_PM0001	Parent	R01R_Y_PM0001	Highest grade or year of school completed by parent (5 levels)
R01_PT0047	Parent	R01R_Y_PT0047	Parent or guardian: marital status (3 levels)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_CPT07	Parent	R01R_Y_AGECAT2	Age range when interviewed (2 levels)
R01_YM0005_01 R01_YM0005_02 R01_YM0005_03 R01_YM0005_04 R01_YM0005_05	Youth	R01R_Y_HISP	Hispanic origin from the interview (2 levels)
R01_YM0006_01 R01_YM0006_02 R01_YM0006_03 R01_YM0006_04 R01_YM0006_05 R01_YM0006_06 R01_YM0006_07 R01_YM0006_08 R01_YM0006_09 R01_YM0006_10 R01_YM0006_11 R01_YM0006_12 R01_YM0006_13 R01_YM0006_14	Youth	R01R_Y_RACECAT3	Race from the interview (3 levels)
R01_YX0671_01 R01_YX0671_02 R01_YX0671_03 R01_YX0671_04 R01_YX0671_05	Youth	R01R_Y_YX0671	Anyone who lives with you now use tobacco (3 levels)
R01_YX0086 R01_YX0087	Youth	R01R_Y_YX0086	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)
R01_YX0074 R01_YX0270	Youth	R01R_Y_YX0074	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
R01_YX0079 R01_YX0271	Youth	R01R_Y_YX0079	Age range when first used marijuana, hash, THC or grass (3 levels)
R01_YE1006 R01 YE1120	Youth	R01R_Y_YE1006	Age range when first tried an e-cigarette, even one or two times (3 levels)
R01_YP1006 R01_YP1120	Youth	R01R_Y_YP1006	Age range when first tried pipe tobacco, even one or two puffs (3 levels)
R01_YH1006 R01_YH1120	Youth	R01R_Y_YH1006	Age range when first tried smoking a hookah, even one or two puffs (3 levels)
R01_YB1006BD R01_YB1120BD	Youth	R01R_Y_YB1006BD	Age range when first tried a bidi, even one or two puffs (3 levels)
R01_YB1006KK R01_YB1120KK	Youth	R01R_Y_YB1006KK	Age range when first tried a kretek, even one or two puffs (3 levels)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_YC1006 R01_YC1120	Youth	R01R_Y_YC1006	Age range when first tried cigarette smoking, even one or two puffs (3 levels)
R01_YD1006 R01_YD1120	Youth	R01R_Y_YD1006	Age range when first tried a dissolvable tobacco product, even one or two times (3 levels)
R01_YG1006CL R01_YG1120CL	Youth	R01R_Y_YG1006CL	Age range when first tried a cigarillo, even one or two puffs (3 levels)
R01_YG1006FC R01_YG1120FC	Youth	R01R_Y_YG1006FC	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
R01_YG1006TC R01_YG1120TC	Youth	R01R_Y_YG1006TC	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
R01_YS1006SM R01_YS1120SM	Youth	R01R_Y_YS1006SM	Age range when first tried smokeless tobacco, even one or two times (3 levels)
R01_YS1006SU R01_YS1120SU	Youth	R01R_Y_YS1006SU	Age range when first tried snus pouches, even one or two times (3 levels)
R01_YX0082_01 R01_YX0272_01	Youth	R01R_Y_YX0082_01	Age range when first used: Ritalin or Adderall (3 levels)
R01_YX0082_02 R01_YX0272_02	Youth	R01R_Y_YX0082_02	Age range when first used: Painkillers, sedatives or tranquilizers (3 levels)
R01_YX0082_03 R01_YX0272_03	Youth	R01R_Y_YX0082_03	Age range when first used: Cocaine or crack (3 levels)
R01_YX0082_04 R01_YX0272_04	Youth	R01R_Y_YX0082_04	Age range when first used: Stimulants like methamphetamine or speed (3 levels)
R01_YX0082_05 R01_YX0272_05	Youth	R01R_Y_YX0082_05	Age range when first used: Any other drugs like heroin, inhalants, solvents or hallucinogens (3 levels)
R01_PT0019 R01_YT0019	Youth/ Parent	R01R_Y_PY0019	Youth's grade performance in school in past 12 months (as reported by parent or emancipated youth)
R01_PT0022 R01_YT0022	Youth/ Parent	R01R_Y_PY0022	Youth has taken medications regularly for asthma in past 12 months (as reported by parent or emancipated youth)
R01_PT0030 R01_YT0030	Youth/ Parent	R01R_Y_PY0030	How often youth missed school due to illness in past 12 months (as reported by parent or emancipated youth)
R01_PT0031 R01_YT0031	Youth/ Parent	R01R_Y_PY0031	Child has ever been told by a doctor or other health professional that he/she has asthma (as reported by parent or emancipated youth)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_PT0033 R01_YT0033	Youth/ Parent	R01R_Y_PY0033	Youth has been told by a doctor or other health professional in past 12 months that he/she has bronchitis, pneumonia or chronic cough (as reported by parent or emancipated youth)
R01_PT0034 R01_YT0034	Youth/ Parent	R01R_Y_PY0034	Number of visits youth made to ER or urgent care for a health problem in past 12 months (as reported by parent or emancipated youth)
R01_PT0035 R01_YT0035	Youth/ Parent	R01R_Y_PY0035	Youth's overall health status (as reported by parent or emancipated youth)
R01_PT0036_01 R01_YT0036_01	Youth/ Parent	R01R_Y_PY0036_01	Medications youth has taken for asthma in past 12 months: Quick-relief inhaler (as reported by parent or emancipated youth)
R01_PT0036_02 R01_YT0036_02	Youth/ Parent	R01R_Y_PY0036_02	Medications youth has taken for asthma in past 12 months: Controller/long-acting inhaler (as reported by parent or emancipated youth)
R01_PT0036_03 R01_YT0036_03	Youth/ Parent	R01R_Y_PY0036_03	Medications youth has taken for asthma in past 12 months: Other controlling medication (as reported by parent or emancipated youth)
R01_PT0036_04 R01_YT0036_04	Youth/ Parent	R01R_Y_PY0036_04	Medications youth has taken for asthma in past 12 months: Oral or injected steroid medication (as reported by parent or emancipated youth)
R01_PT0036_05 R01_YT0036_05	Youth/ Parent	R01R_Y_PY0036_05	Medications youth has taken for asthma in past 12 months: Oxygen therapy (as reported by parent or emancipated youth)
R01_PT0036_06 R01_YT0036_06	Youth/ Parent	R01R_Y_PY0036_06	Medications youth has taken for asthma in past 12 months: Other asthma medication (as reported by parent or emancipated youth)
R01_PT0039 R01_YT0039	Youth/ Parent	R01R_Y_PY0039	Number of asthma attacks youth has had in past 12 months that required use of an oral or injected steroid medication (as reported by parent or emancipated youth)
R01_PT0040 R01_YT0040	Youth/ Parent	R01R_Y_PY0040	Youth has ever been told by a doctor or other health professional that he/she has bronchitis, pneumonia or chronic cough (as reported by parent or emancipated youth)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_PT0044 R01_YT0044	Youth/ Parent	R01R_Y_PY0044	Youth has ever been told by a doctor, dentist or other health professional that he/she has dental health issues (as reported by parent or emancipated youth)
R01_PT0048 R01_YT0048	Youth/ Parent	R01R_Y_PY0048	Youth has taken medications regularly for ADHD or ADD in past 12 months (as reported by parent or emancipated youth)
R01_PT0049 R01_YT0049	Youth/ Parent	R01R_Y_PY0049	Youth has been told by a doctor, dentist or other health professional in past 12 months that he/she has dental health issues (as reported by parent or emancipated youth)
R01_PT0050 R01_YT0050	Youth/ Parent	R01R_Y_PY0050	Youth has ever been told by a doctor or other health professional that he/she has high blood pressure (as reported by parent or emancipated youth)
R01_PT0051 R01_YT0051	Youth/ Parent	R01R_Y_PY0051	Youth has ever been told by a doctor or other health professional that he/she has high cholesterol (as reported by parent or emancipated youth)
R01_PT0052 R01_YT0052	Youth/ Parent	R01R_Y_PY0052	Youth has ever been told by a doctor or other health professional that he/she has ADHD or ADD (as reported by parent or emancipated youth response)
R01_PT0125 R01_YT0125	Youth/ Parent	R01R_Y_PY0125	Youth had an asthma attack in past 12 months that required use of an oral or injected steroid medication (as reported by parent or emancipated youth)
R01_PT0127 R01_YT0127	Youth/ Parent	R01R_Y_PY0127	Youth has visited an emergency room or urgent care center in past 12 months because of asthma (as reported by parent or emancipated youth)
R01_PX0187 R01_YX0187	Youth/ Parent	R01R_Y_PY0187	Youth has difficulty doing errands alone because of a physical, mental or emotional condition (as reported by parent or emancipated youth0
R01_PX0188 R01_YX0188	Youth/ Parent	R01R_Y_PY0188	Youth has difficulty dressing or bathing (as reported by parent or emancipated youth)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R01_PX0189 R01_YX0189	Youth/ Parent	R01R_Y_PY0189	Youth has serious difficulty concentrating, remembering or making decisions because of a physical, mental or emotional condition (as reported by parent or emancipated youth)
R01_PT0281 R01_YT0281	Youth/ Parent	R01R_Y_PY0281	Youth has ever been told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (as reported by parent or emancipated youth)
R01_PT0282 R01_YT0282	Youth/ Parent	R01R_Y_PY0282	Number of times youth has visited an emergency room or urgent care center in past 12 months because of asthma (as reported by parent or emancipated youth)
R01_PX0302 R01_YX0302	Youth/ Parent	R01R_Y_PY0302	Youth visited ER or urgent care for a health problem in past 12 months (as reported by parent or emancipated youth)
R01_PT0038 R01_PT0260 R01_YT0038 R01_YT0260	Youth/ Parent	R01R_Y_PY0038	Age youth was first told he/she has asthma (as reported by parent or emancipated youth)
R01_PT0041 R01_PT0253 R01_YT0041 R01_YT0253	Youth/ Parent	R01R_Y_PY0041	Age youth was first told he/she had high blood pressure (as reported by parent or emancipated youth)
R01_PT0042 R01_PT0263 R01_YT0042 R01_YT0263	Youth/ Parent	R01R_Y_PY0042	Age youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (as reported by parent or emancipated youth)
R01_PT0043 R01_PT0254 R01_YT0043 R01_YT0254	Youth/ Parent	R01R_Y_PY0043	Age youth was first told he/she has high cholesterol (as reported by parent or emancipated youth)

**Table E-2. Wave 1 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

Questionnaire variable	Instrument	PUF derived variable	PUF derived variable description
R01_PT0007_FT	Youth/ Parent	R01R_Y_BMI	Body mass index
R01_PT0007_IN			
R01_PT0007_MT			
R01_PT0008_LB			
R01_PT0008_KG			
R01_YT0007_FT			
R01_YT0007_IN			
R01_YX0310			
R01_YX0311			
R01_YT0008			
R01_YX0312			
R01_YM0018	Youth	R01R_Y_YM0018	Grade level (If on holiday or break – grade level entering when returning to school) (7 levels)
R01_YM0019			

**Table E-3. Wave 2 questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R02_CAD10	Adult	Confirm respondent DOB
R02_LAD01	Adult	Respondent's preferred language to complete ACASI interview
R02_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R02_AM0011_01	Adult	Branch served when on active duty: Army
R02_AM0011_02	Adult	Branch served when on active duty: Navy
R02_AM0011_03	Adult	Branch served when on active duty: Air Force
R02_AM0011_04	Adult	Branch served when on active duty: Marine Corps
R02_AM0011_05	Adult	Branch served when on active duty: Coast Guard
R02_A09035_OS	Adult	Ever used the following electronic nicotine product: Something else – specify
R02_AC1033RY_OS	Adult	Retail location where your roll-your-own cigarette tobacco is purchased most of the time – specify
R02_AC1049RY	Adult	Brand of roll-your-own cigarette tobacco usually/last smoked – specify
R02_AC1071RY	Adult	Sub-brand of roll-your-own cigarette tobacco product usually/last smoked – specify
R02_AC1033MC_OS	Adult	Retail location where your cigarette is purchased most of the time – specify
R02_AC1049MC	Adult	Brand of cigarettes usually/last smoked – specify
R02_AC1071MC	Adult	Sub-brand of cigarette product usually/last smoked – specify
R02_AE9003	Adult	Brand of e-cigarette owned – specify
R02_AE1033_OS	Adult	Retail location where your e-cigarettes/e-cigarette cartridges/e-liquid is purchased most of the time – specify
R02_AE1049	Adult	Brand of e-cigarettes usually/last used – specify
R02_AG1033TC_OS	Adult	Retail location where your traditional cigars is purchased most of the time – specify
R02_AG1049TC	Adult	Brand of traditional cigars usually/last smoked – specify
R02_AG1071TC	Adult	Sub-brand of traditional cigar product usually/last smoked – specify
R02_AG1033CG_OS	Adult	Retail location where your cigarillos is purchased most of the time – specify
R02_AG1049CG	Adult	Brand of cigarillos usually/last smoked – specify
R02_AG1071CG	Adult	Sub-brand of cigarillo product usually/last smoked – specify
R02_AG1033FC_OS	Adult	Retail location where your filtered cigars is purchased most of the time – specify
R02_AG1049FC	Adult	Brand of filtered cigars usually/last smoked – specify
R02_AG1071FC	Adult	Sub-brand of filtered cigar product usually/last smoked – specify
R02_AG1033TJ_OS	Adult	Retail location where your traditional cigars for blunts is purchased most of the time – specify
R02_AG1049TJ	Adult	Brand of traditional cigars for blunts usually/last smoked – specify
R02_AG1071TJ	Adult	Sub-brand of traditional cigar as a blunt product usually/last smoked – specify

**Table E-3. Wave 2 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R02_AG1033GJ_OS	Adult	Retail location where your cigarillos for blunts is purchased most of the time – specify
R02_AG1049GJ	Adult	Brand of cigarillos for blunts usually/last smoked – specify
R02_AG1071GJ	Adult	Sub-brand of cigarillo as a blunt product usually/last smoked – specify
R02_AG1033FJ_OS	Adult	Retail location where your filtered cigars for blunts is purchased most of the time – specify
R02_AG1049FJ	Adult	Brand of filtered cigars for blunts usually/last smoked – specify
R02_AG1071FJ	Adult	Sub-brand of filtered cigar as a blunt product usually/last smoked – specify
R02_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time – specify
R02_AP1049	Adult	Brand of pipe tobacco usually/last smoked – specify
R02_AP1071	Adult	Sub-brand of pipe tobacco product usually/last smoked – specify
R02_AH9011_OS	Adult	Place where usually smoke/smoked hookah: Somewhere else – specify
R02_AH1033_OS	Adult	Retail location where your shisha or hookah tobacco is purchased most of the time – specify
R02_AH1049	Adult	Brand of shisha or hookah tobacco usually/last smoked – specify
R02_AH1071	Adult	Sub-brand of hookah tobacco product usually/last smoked – specify
R02_AS1033SU_OS	Adult	Retail location where your snus pouches is purchased most of the time – specify
R02_AS1049SU	Adult	Brand of snus pouch usually/last used – specify
R02_AS1071SU	Adult	Sub-brand of snus pouch product usually/last used – specify
R02_AS1033SM_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time – specify
R02_AS1049SM	Adult	Brand of smokeless tobacco usually/last used – specify
R02_AS1071SM	Adult	Sub-brand of smokeless tobacco product usually/last used – specify
R02_AD1033_OS	Adult	Retail location where your dissolvable tobacco is purchased most of the time – specify
R02_AD1049	Adult	Brand of dissolvable tobacco usually/last used – specify
R02_AD1071	Adult	Sub-brand of dissolvable tobacco usually/last smoked – specify
R02_AX0190_NB	Adult	Is deaf or has serious difficulty hearing
R02_AX0191_NB	Adult	Is blind or has serious difficulty seeing, even when wearing glasses
R02_AX0297_OS	Adult	Brand of your favorite e-cigarette advertisement – Something else – specify
R02_AX0197_OS	Adult	Brand of favorite tobacco advertisement – specify
R02_AM0012	Adult	Ever been enrolled in VA Health Care
R02_AM0021	Adult	Sexual attraction to gender
R02_AM0061	Adult	Consider yourself to be transgender

**Table E-3. Wave 2 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R02_AM0062	Adult	Consider self to be male-to-female, female-to-male, or non-conforming
R02_LCAD01	Adult	Language in which CAPI portions of adult interview were conducted
R02_AX0217_1A	Adult	Tobacco Product 1, bar code scan
R02_AX0217_2A	Adult	Tobacco Product 2, bar code scan
R02_AX0217_3A	Adult	Tobacco Product 3, bar code scan
R02_AX0217_4A	Adult	Tobacco Product 4, bar code scan
R02_AX0217_5A	Adult	Tobacco Product 5, bar code scan
R02_AX0217_6A	Adult	Tobacco Product 6, bar code scan
R02_AX0217_7A	Adult	Tobacco Product 7, bar code scan
R02_AX0217_8A	Adult	Tobacco Product 8, bar code scan
R02_AX0217_9A	Adult	Tobacco Product 9, bar code scan
R02_AX0217_10A	Adult	Tobacco Product 10, bar code scan
R02_CPT05	Parent	Confirm DOB for Wave 1 shadow youth who is expected to be a youth at Wave 2
R02_CPT05C	Parent	Confirm DOB for Wave 1 youth who is expected to be a youth at Wave 2
R02_CPT05D	Parent	Confirm DOB for Wave 1 shadow youth who is expected to be a shadow youth at Wave 2
R02_CPT07	Parent	Youth date of birth (Corrected)
R02_PT0046_NEW_FNAME	Parent	First name of spouse/partner that lives with parent in DU – specify
R02_PT0046_NEW_LNAME	Parent	Last name of spouse/partner that lives with parent in DU – specify
R02_PM0058_NEW_FNAME	Parent	First name of other parental figure or guardian – specify
R02_PM0058_NEW_LNAME	Parent	Last name of other parental figure or guardian – specify
R02_PM0021_NB	Parent	First name of sibling that youth is a twin of
R02_PM0030_NB	Parent	First names of siblings in multiple birth that are identical to youth
R02_PM0035_NB	Parent	First names of siblings that are in the multiple birth with youth
R02_PARENT_PERSONID	Parent	Parent/guardian Participant ID Number
R02_PM0053	Parent	Confirm parent's relationship to youth
R02_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative
R02_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative
R02_PM0052	Parent	Confirm parent's spouse is same as reported in Wave 1
R02_PT0046_PERSONID	Parent	Parent's spouse/partner's participant ID number (if different from Wave 1)
R02_PT0046_NEW_AGE	Parent	Age of new spouse/partner
R02_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative – specify
R02_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative – specify
R02_PM0058_PERSONID	Parent	Other parental figure/guardian participant ID number
R02_PM0058_NEW_AGE	Parent	Age of other parental figure/guardian
R02_PM0059_OS_1	Parent	Other parental figure/guardian's relationship to sampled youth: Other relative – specify

Table E-3. Wave 2 questionnaire variables entirely excluded from the PUFs (continued)

Questionnaire variable	Instrument	Questionnaire variable description
R02_PM0059_OS_2	Parent	Other parental figure/guardian's relationship to sampled youth: Other non-relative - specify
R02_PM0016_NB	Parent	Youth was identified as a twin or part of a multiple birth
R02_PM0017_NB	Parent	Youth and sibling are identical twins
R02_PM0019_NB	Parent	Any sibling in multiple birth identical to youth
R02_PX0186_NB	Parent	Youth has serious difficulty walking or climbing stairs
R02_PX0191_NB	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
R02_PX0190_NB	Parent	Youth is deaf or has serious difficulty hearing
R02_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
R02_E_YOUTH_LD	Youth	Emancipated youth
R02_CONTINUING_EYOUTH_LD	Youth	Emancipated Youth respondent at Wave 1 is an Emancipated Youth respondent at Wave 2
R02_NEW_BASELINE_EYOUTH_LD	Youth	Youth respondent at Wave 1 is an Emancipated Youth respondent at Wave 2
R02_LYH01	Youth	Youth's preferred language to complete ACASI interview
R02LEY01	Youth	Emancipated Youth's preferred language to complete ACASI interview
R02_YC1118_OS	Youth	In past 30 days, how you usually got your own cigarettes - specify
R02_YC1033_OS	Youth	Retail location where your cigarettes are bought most of the time - specify
R02_YC1049	Youth	Brand of cigarettes usually/last smoked - specify
R02_YC1071	Youth	Sub-brand of cigarette product usually/last smoked - specify
R02_Y09035_OS	Youth	Ever used some other electronic nicotine product - specify
R02_YE1118_OS	Youth	In past 30 days, how you usually got your own e-cigarettes/cartridges and e-liquid - specify
R02_YE1033_OS	Youth	Retail location where your e-cigarettes/cartridges and e-liquid are bought most of the time - specify
R02_YE1049	Youth	Brand of e-cigarettes/cartridges/e-liquid usually/last used - specify
R02_YG1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars - specify
R02_YG1033TC_OS	Youth	Retail location where your traditional cigars are bought most of the time - specify
R02_YG1049TC	Youth	Brand of traditional cigars usually/last smoked - specify
R02_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked - specify
R02_YG1118CL_OS	Youth	In past 30 days, how you usually got your own cigarillos - specify
R02_YG1033CL_OS	Youth	Retail location where your cigarillos are bought most of the time - specify
R02_YG1049CL	Youth	Brand of cigarillos usually/last smoked - specify
R02_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked - specify
R02_YG1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars - specify

**Table E-3. Wave 2 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R02_YG1033FC_OS	Youth	Retail location where your filtered cigars are bought most of the time – specify
R02_YG1049FC	Youth	Brand of filtered cigars usually/last smoked – specify
R02_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked – specify
R02_YG1118TJ_OS	Youth	In past 30 days, how you usually got your own traditional cigars as a blunt – specify
R02_YG1033TJ_OS	Youth	Retail location where your traditional cigars for blunts are bought most of the time – specify
R02_YG1049TJ	Youth	Brand of traditional cigars usually/last smoked as a blunt – specify
R02_YG1071TJ	Youth	Sub-brand of traditional cigar product usually/last smoked as a blunt – specify
R02_YG1118GJ_OS	Youth	In past 30 days, how you usually got your own cigarillo as a blunt – specify
R02_YG1033GJ_OS	Youth	Retail location where your cigarillos for blunts are bought most of the time – specify
R02_YG1049GJ	Youth	Brand of cigarillos usually/last smoked as a blunt – specify
R02_YG1071GJ	Youth	Sub-brand of cigarillo product usually/last smoked as a blunt – specify
R02_YG1118FJ_OS	Youth	In past 30 days, how you usually got your own filtered cigars as a blunt – specify
R02_YG1033FJ_OS	Youth	Retail location where your filtered cigars for blunts are bought most of the time – specify
R02_YG1049FJ	Youth	Brand of filtered cigars usually/last smoked as a blunt – specify
R02_YG1071FJ	Youth	Sub-brand of filtered cigar product usually/last smoked as a blunt – specify
R02_YP1118_OS	Youth	In past 30 days, how you usually got your own pipe tobacco – specify
R02_YP1033_OS	Youth	Retail location where your pipe tobacco are bought most of the time – specify
R02_YH9011_OS	Youth	Usually smoke hookah: Somewhere else – specify
R02_YH1118_OS	Youth	In past 30 days, how you usually got your own shisha or hookah tobacco – specify
R02_YH1033_OS	Youth	Retail location where your shisha or hookah tobacco is bought most of the time – specify
R02_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
R02_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
R02_YH1073	Youth	Smoke hookah because: It is part of my cultural tradition
R02_YS1118SU_OS	Youth	In past 30 days, how you usually got your own snus pouches – specify
R02_YS1033SU_OS	Youth	Retail location where your snus pouches are bought most of the time – specify
R02_YS1049SU	Youth	Brand of snus pouches usually/last used – specify
R02_YS1071SU	Youth	Sub-brand of snus pouch product usually/last used – specify

**Table E-3. Wave 2 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R02_YS1118SM_OS	Youth	In past 30 days, how you usually got your own smokeless tobacco – specify
R02_YT033SM_OS	Youth	Retail location where your smokeless tobacco is bought most of the time – specify
R02_YT049SM	Youth	Brand of smokeless tobacco usually/last used – specify
R02_YT071SM	Youth	Sub-brand of smokeless tobacco product usually/last used – specify
R02_YY0601_OS	Youth	First type of tobacco you tried – specify
R02_YX0302	Youth	In past 12 months, visited an emergency room or urgent care center for a health problem, accident or injury
R02_YT0034	Youth	In past 12 months, number of visits to emergency room or urgent care center
R02_YX0186_NB	Youth	Has serious difficulty walking or climbing stairs
R02_YX0188_NB	Youth	Has difficulty dressing or bathing
R02_YX0191_NB	Youth	Is blind or has serious difficulty seeing, even when wearing glasses
R02_YX0190_NB	Youth	Is deaf or has serious difficulty hearing
R02_YX0189_NB	Youth	Has serious difficulty concentrating, remembering or making decisions because of a physical, mental or emotional condition
R02_YX0187_NB	Youth	Has difficulty doing errands alone such as visiting a doctor's office or shopping because of a physical, mental or emotional condition
R02_YT0050_12M	Youth	In past 12 months, been told by a doctor, nurse or other health professional that you have high blood pressure
R02_YT0050_NB	Youth	Ever been told by a doctor, nurse or other health professional that you have high blood pressure
R02_YT0041_NB	Youth	Age when you were first told you had high blood pressure
R02_YT0253_NB	Youth	Nonresponse follow-up probe: Age group you were in when you were first told you had high blood pressure
R02_YT0051_12M	Youth	In past 12 months, been told by a doctor, nurse or other health professional that you have high cholesterol
R02_YT0051_NB	Youth	Ever been told by a doctor, nurse or other health professional that you have high cholesterol
R02_YT0043_NB	Youth	Age when you were first told you had high cholesterol
R02_YT0254_NB	Youth	Nonresponse follow-up probe: Age group you were in when you were first told you had high cholesterol
R02_YT0031_12M	Youth	In past 12 months, been told by a doctor, nurse or other health professional that you have asthma
R02_YT0033_12M	Youth	In past 12 months, been told by a doctor, nurse or other health professional that you have bronchitis, pneumonia or chronic cough
R02_YT0040_NB	Youth	Ever been told by a doctor, nurse or other health professional that you have bronchitis, pneumonia or chronic cough
R02_YT0052_NB	Youth	Ever been told by a doctor, nurse or other health professional that you have ADHD or ADD
R02_YT0052_12M	Youth	In past 12 months, been told by a doctor, nurse or other health professional that you have ADHD or ADD

**Table E-3. Wave 2 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R02_YT0048	Youth	In past 12 months, took medications regularly for ADHD or ADD
R02_YT0044_NB	Youth	Ever been told by a doctor, dentist or other health professional that you have dental health issues
R02_YT0049_12M	Youth	In past 12 months, been told by a doctor, dentist or other health professional that you have dental health issues
R02_YT0281_NB	Youth	Ever been told by a doctor, nurse or other health professional that you have diabetes, sugar diabetes, high blood sugar or borderline diabetes
R02_YT0042_NB	Youth	Age when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes
R02_YT0263_NB	Youth	Nonresponse follow-up probe: Age group you were in when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes
R02_YT0281_12M	Youth	In past 12 months, been told by a doctor, nurse or other health professional that you have diabetes, sugar diabetes, high blood sugar or borderline diabetes
R02_YX0136	Youth	Currently pregnant
R02_YX0137_NN	Youth	Number of weeks/months pregnant – Number
R02_YX0137_UN	Youth	Number of weeks/months pregnant – Unit
R02_YX0478_OS	Youth	Received discount coupons from: Some other way – specify
R02_YX0480_OS	Youth	How tobacco company sent you the information other than coupons: Some other way – specify
R02_YX0297_OS	Youth	Brand of your favorite e-cigarette advertisement – Something else – specify
R02_YX0197_OS	Youth	Brand of favorite tobacco advertisement – specify
R02_YM0020	Youth	Last grade/year in school completed
R02_YT0019	Youth	In past 12 months, academic performance at school
R02_YT0030	Youth	In past 12 months, how often missed school due to illness
R02_YM0030	Youth	In past 12 months, total household income category
R02_YM0031	Youth	In past 12 months, total household income above or below \$50,000
R02_YL0040	Youth	Home is owned or rented
R02_YM0021	Youth	Sexual attraction to gender
R02_YM0063	Youth	Sexual orientation
R02_YM0061	Youth	Transgender
R02_YM0062	Youth	Transgender category
R02_LCYS01	Youth	Language in which CAPI portions of youth interview were conducted

**Table E-4. Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R02_AM0001	Adult	R02R_A_AGECAT7	Age range when interviewed (7 levels)
R02_AM0002			
R02_AM0003			
R02_AC1007_NB	Adult	R02R_A_AC1007_NB	Age range when first started smoking cigarettes fairly regularly (6 levels)
R02_AC1121_NB			
R02_AC1020_NB	Adult	R02R_A_AC1020_NB	Age range when first started smoking cigarettes every day (6 levels)
R02_AC1122_NB			
R02_AE1007_NB	Adult	R02R_A_AE1007_NB	Age range when first started using e-cigarettes fairly regularly (6 levels)
R02_AE1121_NB			
R02_AE1020_NB	Adult	R02R_A_AE1020_NB	Age range when first started using e-cigarettes every day (6 levels)
R02_AE1122_NB			
R02_AG1007TC_NB	Adult	R02R_A_AG1007TC_NB	Age range when first started smoking traditional cigars fairly regularly (6 levels)
R02_AG1121TC_NB			
R02_AG1020TC_NB	Adult	R02R_A_AG1020TC_NB	Age range when first started smoking cigarillos every day (6 levels)
R02_AG1122TC_NB			
R02_AG1007CG_NB	Adult	R02R_A_AG1007CG_NB	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R02_AG1121CG_NB			
R02_AG1020CG_NB	Adult	R02R_A_AG1020CG_NB	Age range when first started smoking cigarillos every day (6 levels)
R02_AG1122CG_NB			
R02_AG1007FC_NB	Adult	R02R_A_AG1007FC_NB	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R02_AG1121FC_NB			
R02_AG1020FC_NB	Adult	R02R_A_AG1020FC_NB	Age range when first started smoking filtered cigars every day (6 levels)
R02_AG1122FC_NB			
R02_AP1007_NB	Adult	R02R_A_AP1007_NB	Age range when first started smoking a pipe filled with tobacco fairly regularly (6 levels)
R02_AP1121_NB			
R02_AP1020_NB	Adult	R02R_A_AP1020_NB	Age range when first started smoking a pipe filled with tobacco every day (6 levels)
R02_AP1122_NB			
R02_AH1007_NB	Adult	R02R_A_AH1007_NB	Age range when first started smoking hookah fairly regularly (6 levels)
R02_AH1121_NB			
R02_AH1020_NB	Adult	R02R_A_AH1020_NB	Age range when first started smoking hookah every day (6 levels)
R02_AS1007SU_NB	Adult	R02R_A_AS1007SU_NB	Age range when first started using snus pouches fairly regularly (6 levels)
R02_AS1121SU_NB			
R02_AS1020SU_NB	Adult	R02R_A_AS1020SU_NB	Age range when first started using snus pouches every day (6 levels)
R02_AS1122SU_NB			
R02_AS1007SM_NB	Adult	R02R_A_AS1007SM_NB	Age range when first started using smokeless tobacco fairly regularly (6 levels)
R02_AS1121SM_NB			
R02_AS1020SM_NB	Adult	R02R_A_AS1020SM_NB	Age range when first started using smokeless tobacco every day (6 levels)
R02_AS1122SM_NB			

**Table E-4. Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R02_AD1007_NB R02_AD1121_NB	Adult	R02R_A_AD1007_NB	Age range when first started using dissolvable tobacco fairly regularly (6 levels)
R02_AD1020_NB R02_AD1122_NB	Adult	R02R_A_AD1020_NB	Age range when first started using dissolvable tobacco every day (6 levels)
R02_AX0114_NB R02_AX0253_NB	Adult	R02R_A_AX0114_NB	Age range when you were first told you had high blood pressure (6 levels)
R02_AX0115_NB R02_AX0254_NB	Adult	R02R_A_AX0115_NB	Age range when you were first told you had high cholesterol (6 levels)
R02_AX0116_NB R02_AX0255_NB	Adult	R02R_A_AX0116_NB	Age range when you were first told you had congestive heart failure (6 levels)
R02_AX0117_NB R02_AX0256_NB	Adult	R02R_A_AX0117_NB	Age range when you were first told you had a stroke (6 levels)
R02_AX0112_NB R02_AX0252_NB	Adult	R02R_A_AX0112_NB	Age range when you were first told you had a heart attack (6 levels)
R02_AX0120_NB R02_AX0257_NB	Adult	R02R_A_AX0120_NB	Age range when you were first told you had COPD (6 levels)
R02_AX0121_NB R02_AX0258_NB	Adult	R02R_A_AX0121_NB	Age range when you were first told you had chronic bronchitis (6 levels)
R02_AX0123_NB R02_AX0259_NB	Adult	R02R_A_AX0123_NB	Age range when you were first told you had emphysema (6 levels)
R02_AX0124_NB R02_AX0260_NB	Adult	R02R_A_AX0124_NB	Age range when you were first told you had asthma (6 levels)
R02_AX0131_NB R02_AX0261_NB	Adult	R02R_A_AX0131_NB	Age range when you were first told you had gum disease (6 levels)
R02_AX0133_NB R02_AX0262_NB	Adult	R02R_A_AX0133_NB	Age range when you were first told you had pre-cancerous oral lesions (6 levels)
R02_AX0280_NB R02_AX0263_NB	Adult	R02R_A_AX0280_NB	Age range when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes (6 levels)
R02_AX0143_NB R02_AX0264_NB	Adult	R02R_A_AX0143_NB	Age range when you were first told you had an ulcer (6 levels)
R02_AX0148_NB R02_AX0266_NB	Adult	R02R_A_AX0148_NB	Age range when you were first told you had stomach or gastro-intestinal bleeding (6 levels)
R02_AX0150_NB R02_AX0267_NB	Adult	R02R_A_AX0150_NB	Age range when you were first told you had osteoporosis (6 levels)

**Table E-4. Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R02_AX0198_NB R02_AX0268_NB	Adult	R02R_A_AX0198_NB	Age range when you were first told you had a bone fracture because you have fragile bones (6 levels)
R02_AX0152_NB R02_AX0269_NB	Adult	R02R_A_AX0152_NB	Age range when you were first told you had a cataract or glaucoma (6 levels)
R02_AX0145_02 R02_AX0145_03 R02_AX0145_04 R02_AX0145_05 R02_AX0145_09 R02_AX0145_12 R02_AX0145_15 R02_AX0145_18 R02_AX0145_19 R02_AX0145_21 R02_AX0145_16 R02_AX0145_23 R02_AX0145_24 R02_AX0145_25 R02_AX0145_27 R02_AX0145_29 R02_AX0145_30 R02_AX0145_31	Adult	R02R_A_AX0145_NONTOB	Type of cancer is a non-tobacco related cancer (2 levels)
R02_AX0145_01 R02_AX0145_06 R02_AX0145_07 R02_AX0145_08 R02_AX0145_10 R02_AX0145_11 R02_AX0145_13 R02_AX0145_14 R02_AX0145_17 R02_AX0145_20 R02_AX0145_22 R02_AX0145_26 R02_AX0145_28	Adult	R02R_A_AX0145_TOB	Type of cancer is a tobacco related cancer (2 levels)

**Table E-4. Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R02_AX0146_NB_02			
R02_AX0265_NB_02			
R02_AX0146_NB_03			
R02_AX0265_NB_03			
R02_AX0146_NB_04			
R02_AX0265_NB_04			
R02_AX0146_NB_05			
R02_AX0265_NB_05			
R02_AX0146_NB_09			
R02_AX0265_NB_09			
R02_AX0146_NB_12			
R02_AX0265_NB_12			
R02_AX0146_NB_15			
R02_AX0265_NB_15			
R02_AX0146_NB_18			
R02_AX0265_NB_18			
R02_AX0146_NB_19			
R02_AX0265_NB_19			
R02_AX0146_NB_21			
R02_AX0265_NB_21			
R02_AX0146_NB_16			
R02_AX0265_NB_16			
R02_AX0146_NB_23			
R02_AX0265_NB_23			
R02_AX0146_NB_24			
R02_AX0265_NB_24			
R02_AX0146_NB_25			
R02_AX0265_NB_25			
R02_AX0146_NB_27			
R02_AX0265_NB_27			
R02_AX0146_NB_29			
R02_AX0265_NB_29			
R02_AX0146_NB_30			
R02_AX0265_NB_30			
R02_AX0146_NB_31			
R02_AX0265_NB_31			

**Table E-4. Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R02_AX0146_NB_01	Adult	R02R_A_AX0146_TOB	Age range when tobacco related cancer was diagnosed (6 levels)
R02_AX0265_NB_01			
R02_AX0146_NB_06			
R02_AX0265_NB_06			
R02_AX0146_NB_07			
R02_AX0265_NB_07			
R02_AX0146_NB_08			
R02_AX0265_NB_08			
R02_AX0146_NB_10			
R02_AX0265_NB_10			
R02_AX0146_NB_11			
R02_AX0265_NB_11			
R02_AX0146_NB_13			
R02_AX0265_NB_13			
R02_AX0146_NB_14			
R02_AX0265_NB_14			
R02_AX0146_NB_17			
R02_AX0265_NB_17			
R02_AX0146_NB_20			
R02_AX0265_NB_20			
R02_AX0146_NB_22			
R02_AX0265_NB_22			
R02_AX0146_NB_26			
R02_AX0265_NB_26			
R02_AX0146_NB_28			
R02_AX0265_NB_28			
R02_AX0135_12M	Adult	R02R_A_AX0135_12M	Outcome of last pregnancy was miscarriage, induced abortion, ectopic or tubal pregnancy or stillbirth (2 levels)
R02_AX0300_12M_01	Adult	R02R_A_AX0300_12M	Outcome for live birth was preterm birth, low birth weight, birth defects, placenta previa, placenta abruption, or preeclampsia (2 levels)
R02_AX0300_12M_02			
R02_AX0300_12M_03			
R02_AX0300_12M_04			
R02_AX0300_12M_05			
R02_AX0300_12M_06			
R02_AX0300_12M_07			
R02_AM0018	Adult	R02R_A_AM0018	Highest grade or level of school completed (6 levels)
R02_AM0030	Adult	R02R_A_AM0030	Total household income in the past 12 months (5 levels)
R02_AM0038	Adult	R02R_A_AM0038	Type of current residence (7 levels)
R02_AM0033	Adult	R02R_A_AM0033	Highest grade or year of school completed by mother, step-mother or mother-figure (6 levels)
R02_AM0034	Adult	R02R_A_AM0034	Highest grade or year of school completed by father, step-father or father-figure (6 levels)
R02_AM0036	Adult	R02R_A_AM0036	In past 12 months, parents' total household income (5 levels)

**Table E-4. Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R02_AM0026_01	Adult	R02R_A_AM0026	Currently covered by health insurance or health coverage plan (5 levels)
R02_AM0026_02			
R02_AM0026_03			
R02_AM0026_04			
R02_AM0026_05			
R02_AM0026_06			
R02_AM0026_07			
R02_AM0026_08			
R02_AM0063	Adult	R02R_A_SEXORIENT2	Adult sexual orientation (2 levels)
R02_AX0313	Adult	R02R_A_BMI	Body mass index
R02_AX0679_FT			
R02_AX0679_IN			
R02_AX0316			
R02_AX0314			
R02_AX0109			
R02_AX0312			
R02_PT0001	Parent	R02R_Y_PT0001	Parent or guardian relationship to youth (4 levels)
R02_PT0047	Parent	R02R_Y_PT0047	Parent or guardian marital status (3 levels)
R02_PT0045	Parent	R02R_P_OTHPAR_INHH	Youth has any parental figures/guardians other than the parent/guardian in the house
R02_PM0057			
R02_PT0002	Parent	R02R_Y_PT0002	Parent's spouse or partner relationship to youth (4 levels)
R02_PM0059	Parent	R02R_Y_PM0059	Other parental figure/guardian's relationship to youth (4 levels)
R02_PT0041_NB	Parent	R02R_Y_PT0041_NB	Age range youth was first told he/she had high blood pressure (3 levels)
R02_PT0253_NB			
R02_PT0043_NB	Parent	R02R_Y_PT0043_NB	Age range youth was first told he/she has high cholesterol (3 levels)
R02_PT0254_NB			
R02_PT0038_NB	Parent	R02R_Y_PT0038_NB	Age range youth was first told he/she has asthma (3 levels)
R02_PT0260_NB			
R02_PT0042_NB	Parent	R02R_Y_PT0042_NB	Age range youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (3 levels)
R02_PT0263_NB			
R02_PM0001	Parent	R02R_Y_PM0001	Highest grade or year of school completed by parent (6 levels)
R02_PM0118	Parent	R02R_Y_PM0118	Highest grade or year of school completed by spouse/guardian (6 levels)
R02_PM0130	Parent	R02R_Y_PM0130	Total household income in past 12 months (5 levels)
R02_YM0005_NB_01	Youth	R02R_Y_HISP	Hispanic origin from the interview (2 levels)
R02_YM0005_NB_02			
R02_YM0005_NB_03			
R02_YM0005_NB_04			
R02_YM0005_NB_05			

**Table E-4. Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

Questionnaire variable	Instrument	PUF derived variable	PUF derived variable description
R02_YM0006_NB_01 R02_YM0006_NB_02 R02_YM0006_NB_03 R02_YM0006_NB_04 R02_YM0006_NB_05 R02_YM0006_NB_06 R02_YM0006_NB_07 R02_YM0006_NB_08 R02_YM0006_NB_09 R02_YM0006_NB_10 R02_YM0006_NB_11 R02_YM0006_NB_12 R02_YM0006_NB_13 R02_YM0006_NB_14	Youth	R02R_Y_RACECAT3	Race from the interview (3 levels)
R02_YC1006_NB R02_YC1120_NB	Youth	R02R_Y_YC1006_NB	Age range when first tried cigarette smoking, even one or two puffs (3 levels)
R02_YC1007 R02_YC1121	Youth	R02R_Y_YC1007	Age range when first started smoking cigarettes fairly regularly (3 levels)
R02 YE1006_NB R02 YE1120_NB	Youth	R02R_Y_YE1006_NB	Age range when first tried an e-cigarette, even one or two times (3 levels)
R02 YE1007 R02 YE1121	Youth	R02R_Y_YE1007	Age range when first started using e-cigarettes fairly regularly (3 levels)
R02_YG1006TC_NB R02_YG1120TC_NB	Youth	R02R_Y_YG1006TC_NB	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
R02_YG1007TC R02_YG1121TC	Youth	R02R_Y_YG1007TC	Age range when first started smoking traditional cigars fairly regularly (3 levels)
R02_YG1006CL_NB R02_YG1120CL_NB	Youth	R02R_Y_YG1006CL_NB	Age range when first tried a cigarillo, even one or two puffs (3 levels)
R02_YG1007CL R02_YG1121CL	Youth	R02R_Y_YG1007CL	Age range when first started smoking cigarillos fairly regularly (3 levels)
R02_YG1006FC_NB R02_YG1120FC_NB	Youth	R02R_Y_YG1006FC_NB	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
R02_YG1007FC R02_YG1121FC	Youth	R02R_Y_YG1007FC	Age range when first started smoking filtered cigars fairly regularly (3 levels)
R02_YH1006_NB R02_YH1120_NB	Youth	R02R_Y_YH1006_NB	Age range when first tried smoking hookah, even one or two puffs (3 levels)
R02_YH1007 R02_YH1121	Youth	R02R_Y_YH1007	Age range when first started smoking a hookah fairly regularly (3 levels)
R02_YS1006SU_NB R02_YS1120SU_NB	Youth	R02R_Y_YS1006SU_NB	Age range when first tried snus pouches, even one or two times (3 levels)
R02_YS1007SU R02_YS1121SU	Youth	R02R_Y_YS1007SU	Age range when first started using snus pouches fairly regularly (3 levels)
R02_YS1006SM_NB R02_YS1120SM_NB	Youth	R02R_Y_YS1006SM_NB	Age range when first tried smokeless tobacco, even one or two times (3 levels)

**Table E-4. Wave 2 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R02_YS1007SM R02_YS1121SM	Youth	R02R_Y_YS1007SM	Age range when first started using smokeless tobacco fairly regularly (3 levels)
R02_YX0671_01 R02_YX0671_02 R02_YX0671_03 R02_YX0671_04 R02_YX0671_05	Youth	R02R_Y_YX0671	Anyone who lives with you now use tobacco (3 levels)
R02_YX0086_NB R02_YX0087_NB	Youth	R02R_Y_YX0086_NB	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)
R02_YT0038_NB R02_YT0260_NB	Youth	R02R_Y_YT0038_NB	Age range when you were first told you had asthma (3 levels)
R02_YX0074_NB R02_YX0270_NB	Youth	R02R_Y_YX0074_NB	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
R02_YX0079_NB R02_YX0271_NB	Youth	R02R_Y_YX0079_NB	Age range when first used marijuana, hash, THC or grass (3 levels)
R02_YX0082_NB_01 R02_YX0272_NB_01	Youth	R02R_Y_YX0082_NB_01	Age range when first used: Ritalin or Adderall (3 levels)
R02_YX0082_NB_02 R02_YX0272_NB_02	Youth	R02R_Y_YX0082_NB_02	Age range when first used: Painkillers, sedatives or tranquilizers (3 levels)
R02_YX0082_NB_03 R02_YX0272_NB_03	Youth	R02R_Y_YX0082_NB_03	Age range when first used: Cocaine or crack (3 levels)
R02_YX0082_NB_04 R02_YX0272_NB_04	Youth	R02R_Y_YX0082_NB_04	Age range when first used: Stimulants like methamphetamine or speed (3 levels)
R02_YX0272_NB_05 R02_YX0082_NB_05	Youth	R02R_Y_YX0082_NB_05	Age range when first used: Any other drugs like heroin, inhalants, solvents or hallucinogens (3 levels)
R02_YM0019 R02_YM0018	Youth	R02R_Y_YM0018	Grade level (If on holiday or break – grade level entering when returning to school) (7 levels)
R02_PT0007_FT R02_PT0007_IN R02_PT0007_MT R02_PT0008_LB R02_PT0008_KG R02_YX0310 R02_YT0007_FT R02_YT0007_IN R02_YX0311 R02_YT0008 R02_YX0312	Youth/ Parent	R02R_Y_BMI	Body mass index

**Table E-5. Wave 3 questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R03_CAD10	Adult	Confirm respondent DOB
R03_LAD01	Adult	Respondent's preferred language to complete ACASI interview
R03_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R03_AM0011_01	Adult	Branch served when on active duty: Army
R03_AM0011_02	Adult	Branch served when on active duty: Navy
R03_AM0011_03	Adult	Branch served when on active duty: Air Force
R03_AM0011_04	Adult	Branch served when on active duty: Marine Corps
R03_AM0011_05	Adult	Branch served when on active duty: Coast Guard
R03_AM0072_OS	Adult	Language other than English spoken at home - specify
R03_AZ1002_OS	Adult	Ever used any other tobacco products - specify
R03_AC1033MC_OS	Adult	Retail location where your cigarette is purchased most of the time - specify
R03_AC1049MC	Adult	Brand of cigarettes usually/last smoked - specify
R03_AC1071MC	Adult	Sub-brand of cigarette product usually/last smoked - specify
R03_AC1033RY_OS	Adult	Retail location where your roll-your-own cigarette tobacco is purchased most of the time - specify
R03_AC1049RY	Adult	Brand of roll-your-own cigarette tobacco usually/last smoked - specify
R03_AC1071RY	Adult	Sub-brand of roll-your-own cigarette tobacco product usually/last smoked - specify
R03_AV9003	Adult	Brand of [primary electronic nicotine product] owned - specify
R03_AV1011_OS	Adult	Flavor of [primary electronic nicotine product] when first started using - specify
R03_AV1131_OS	Adult	In past 30 days, [electronic nicotine products/cartridges/e-liquid] flavor used - specify
R03_AV1033_OS	Adult	Retail location where your [electronic nicotine products/cartridges/e-liquid] are bought most of the time - specify
R03_AV1012_OS	Adult	Flavor of regular brand/brand last used - specify
R03_AV1049	Adult	Brand of [electronic nicotine products/cartridges/e-liquid] usually/last used - specify
R03_AV2011_OS	Adult	When first started using [secondary electronic nicotine product], flavor used - specify
R03_AJ1033TC_OS	Adult	Retail location where your traditional cigars for blunts is purchased most of the time - specify
R03_AJ1049TC	Adult	Brand of traditional cigars for blunts usually/last smoked - specify
R03_AJ1071TC	Adult	Sub-brand of traditional cigar as a blunt product usually/last smoked - specify
R03_AJ1033CG_OS	Adult	Retail location where your cigarillos for blunts is purchased most of the time - specify
R03_AJ1049CG	Adult	Brand of cigarillos for blunts usually/last smoked - specify

**Table E-5. Wave 3 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R03_AJ1071CG	Adult	Sub-brand of cigarillo as a blunt product usually/last smoked – specify
R03_AJ1033FC_OS	Adult	Retail location where your filtered cigars for blunts is purchased most of the time – specify
R03_AJ1049FC	Adult	Brand of filtered cigars for blunts usually/last smoked – specify
R03_AJ1071FC	Adult	Sub-brand of filtered cigar as a blunt product usually/last smoked – specify
R03_AG1033TC_OS	Adult	Retail location where your traditional cigars is/were purchased most of the time – specify
R03_AG1049TC	Adult	Brand of traditional cigars usually/last smoked – specify
R03_AG1071TC	Adult	Sub-brand of traditional cigar product usually/last smoked – specify
R03_AG1033CG_OS	Adult	Retail location where your cigarillos is/were purchased most of the time – specify
R03_AG1049CG	Adult	Brand of cigarillos usually/last smoked – specify
R03_AG1071CG	Adult	Sub-brand of cigarillo product usually/last smoked – specify
R03_AG1033FC_OS	Adult	Retail location where your filtered cigars is/were purchased most of the time – specify
R03_AG1049FC	Adult	Brand of filtered cigars usually/last smoked – specify
R03_AG1071FC	Adult	Sub-brand of filtered cigar product usually/last smoked – specify
R03_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time – specify
R03_AP1049	Adult	Brand of pipe tobacco usually/last smoked – specify
R03_AP1071	Adult	Sub-brand of pipe tobacco product usually/last smoked – specify
R03_AH9011_OS	Adult	Place where usually smoke/smoked a hookah – specify
R03_AH1033_OS	Adult	Retail location where your hookah tobacco is purchased most of the time – specify
R03_AH1049	Adult	Brand of shisha or hookah tobacco usually/last smoked – specify
R03_AH1071	Adult	Sub-brand of hookah tobacco product usually/last smoked – specify
R03_AU1033_OS	Adult	Retail location where your snus pouches is purchased most of the time – specify
R03_AU1049	Adult	Brand of snus pouches usually/last used – specify
R03_AU1071	Adult	Sub-brand of snus pouch product usually/last used – specify
R03_AS1033_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time – specify
R03_AS1049	Adult	Brand of smokeless tobacco usually/last used – specify
R03_AS1071	Adult	Sub-brand of smokeless tobacco product usually/last used – specify

**Table E-5. Wave 3 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R03_AX0190_NB	Adult	Is deaf or has serious difficulty hearing
R03_AX0191_NB	Adult	Is blind or has serious difficulty seeing, even when wearing glasses
R03_AV9035_OS	Adult	Electronic nicotine product you have used – specify
R03_AX0203_OS	Adult	In past 30 days, place where noticed e-cigarettes or other electronic nicotine products being advertised – specify
R03_AX0677_OS	Adult	In past 30 days, noticed cigarettes or other tobacco products being advertised – specify
R03_AX0686_OS	Adult	In past 12 months, liked or followed brand on social media sites – specify
R03_AM0067	Adult	Enrolled in High School
R03_AM0068	Adult	Current grade in school
R03_AM0020_OS	Adult	Type of degree program currently enrolled in – specify
R03_AM0012	Adult	Ever been enrolled in VA Health Care
R03_AM0021	Adult	Sexual attraction to gender
R03_AM0061	Adult	Consider yourself to be transgender
R03_AM0062	Adult	Consider self to be male-to-female, female-to-male, or non-conforming
R03_LCAD01	Adult	Language in which CAPI portions of adult interview were conducted
R03_AX0217_1A	Adult	Tobacco Product 1, bar code scan
R03_AX0217_2A	Adult	Tobacco Product 2, bar code scan
R03_AX0217_3A	Adult	Tobacco Product 3, bar code scan
R03_AX0217_4A	Adult	Tobacco Product 4, bar code scan
R03_AX0217_5A	Adult	Tobacco Product 5, bar code scan
R03_AX0217_6A	Adult	Tobacco Product 6, bar code scan
R03_AX0217_7A	Adult	Tobacco Product 7, bar code scan
R03_AX0217_8A	Adult	Tobacco Product 8, bar code scan
R03_AX0217_9A	Adult	Tobacco Product 9, bar code scan
R03_AX0217_10A	Adult	Tobacco Product 10, bar code scan
R03_AM0001	Adult	Date of birth (Corrected)
R03_AM0002	Adult	Respondent age
R03_CPT05	Parent	Confirm DOB for Wave 1 shadow youth who is expected to be a youth at Wave 3
R03_CPT05C	Parent	Confirm DOB for Wave 1 youth who is expected to be a youth at Wave 3
R03_CPT05D	Parent	Confirm DOB for Wave 1 shadow youth who is expected to be a shadow youth at Wave 3
R03_CPT07	Parent	Corrected Youth DOB
R03_PM0021_NB	Parent	First names of siblings in multiple birth that are identical to youth
R03_PM0030_NB	Parent	First name of sibling that youth is a twin of
R03_PM0035_NB	Parent	First names of siblings that are in the multiple birth with youth
R03_PM0058_NEW_FNAME	Parent	First name of other parental figure or guardian
R03_PM0058_NEW_LNAME	Parent	Last name of other parental figure or guardian
R03_PT0046_NEW_FNAME	Parent	First name of spouse/partner

**Table E-5. Wave 3 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R03_PT0046_NEW_LNAME	Parent	Last name of spouse/partner
R03_PARENT_PERSONID	Parent	Wave 3 Parent/guardian Participant ID Number
R03_PM0053	Parent	Confirm parent's relationship to youth
R03_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative
R03_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative
R03_PM0052	Parent	Confirm parent's spouse is same as reported in previous wave
R03_PT0046_NEW AGE	Parent	Age of new spouse/partner
R03_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative – specify
R03_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative – specify
R03_PM0058_PERSONID	Parent	Other parental figure/guardian participant ID number
R03_PM0058_NEW AGE	Parent	Age of other parental figure/guardian
R03_PM0059_OS_1	Parent	Other parental figure/guardian's relationship to sampled youth: Other relative – specify
R03_PM0059_OS_2	Parent	Other parental figure/guardian's relationship to sampled youth: Other non-relative – specify
R03_PM0016_NB	Parent	Youth was identified as a twin or part of a multiple birth
R03_PM0017_NB	Parent	Youth and sibling are identical twins
R03_PM0019_NB	Parent	Any sibling in multiple birth identical to youth
R03_PX0186_NB	Parent	Youth has serious difficulty walking or climbing stairs
R03_PX0191_NB	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
R03_PX0190_NB	Parent	Youth is deaf or has serious difficulty hearing
R03_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
R03_E_YOUTH_LD	Youth	Respondent is an emancipated youth
R03_LYH01	Youth	Youth's preferred language to complete ACASI interview
R03LEY01	Youth	Emancipated Youth's preferred language to complete ACASI interview
R03_YM0066	Youth	How well speak English
R03_YM0070	Youth	How well read English
R03_YM0073	Youth	How well write in English
R03_YM0072_OS	Youth	Language other than English spoken at home: Other language – specify
R03_YC1118_OS	Youth	In past 30 days, how you usually got your own cigarettes – specify
R03_YC1033_OS	Youth	Retail location where your cigarettes are bought most of the time – specify
R03_YC1049	Youth	Brand of cigarettes usually/last smoked – specify
R03_YC1071	Youth	Sub-brand of cigarette product usually/last smoked – specify

**Table E-5. Wave 3 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R03_YV9035_OS	Youth	Ever used some other electronic nicotine product – specify
R03_YV1011_OS	Youth	Flavor of first [primary electronic nicotine product] used: Specify.
R03_YV1131_OS	Youth	In past 30 days, used [electronic nicotine products/cartridges/e-liquid] flavored to taste like some other flavor – specify
R03_YV1118_OS	Youth	In past 30 days, how you usually got your own [electronic nicotine products/cartridges/e-liquid] – specify
R03_YV1033_OS	Youth	Where[electronic nicotine products/cartridges/e-liquid] is purchased – specify
R03_YV1049	Youth	Brand of e-liquid usually/last used – specify
R03_YV2011_OS	Youth	Flavor of first [secondary electronic nicotine product] used – specify
R03_YJ1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars as a blunt – specify
R03_YJ1033TC_OS	Youth	Retail location where your traditional cigars for blunts are bought most of the time – specify
R03_YJ1049TC	Youth	Brand of traditional cigars usually/last smoked as a blunt – specify
R03_YJ1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked as a blunt – specify
R03_YJ1118CG_OS	Youth	In past 30 days, how you usually got your own cigarillo as a blunt – specify
R03_YJ1033CG_OS	Youth	Retail location where your cigarillos for blunts are bought most of the time – specify
R03_YJ1049CG	Youth	Brand of cigarillos usually/last smoked as a blunt – specify
R03_YJ1071CG	Youth	Sub-brand of cigarillo product usually/last smoked as a blunt – specify
R03_YJ1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars as a blunt – specify
R03_YJ1033FC_OS	Youth	Retail location where your filtered cigars for blunts are bought most of the time – specify
R03_YJ1049FC	Youth	Brand of filtered cigars usually/last smoked as a blunt – specify
R03_YJ1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked as a blunt – specify
R03_YG1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars – specify
R03_YG1033TC_OS	Youth	Retail location where your traditional cigars are purchased most of the time – specify
R03_YG1049TC	Youth	Brand of traditional cigars usually/last smoked – specify
R03_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked – specify
R03_YG1118CL_OS	Youth	In past 30 days, how you usually got your own cigarillos – specify

**Table E-5. Wave 3 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R03_YG1033CL_OS	Youth	Retail location where your cigarillos are purchased most of the time – specify
R03_YG1049CL	Youth	Brand of cigarillos usually/last smoked – specify
R03_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked – specify
R03_YG1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars – specify
R03_YG1033FC_OS	Youth	Retail location where your filtered cigars are purchased most of the time – specify
R03_YG1049FC	Youth	Brand of filtered cigars usually/last smoked – specify
R03_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked – specify
R03_YP1118_OS	Youth	In past 30 days, how you usually got your own pipe tobacco – specify
R03_YP1033_OS	Youth	Retail location where your pipe tobacco are bought most of the time – specify
R03_YH9011_OS	Youth	Usually smoke a hookah: Somewhere else – specify
R03_YH1118_OS	Youth	In past 30 days, how you usually got your own shisha or hookah tobacco – specify
R03_YH1033_OS	Youth	Retail location where your hookah tobacco is purchased most of the time – specify
R03_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
R03_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
R03_YH1073	Youth	Smoke hookah because: It is part of my cultural tradition
R03_YU1118_OS	Youth	In past 30 days, how you usually got your own snus pouches – specify
R03_YU1033_OS	Youth	Retail location where your snus pouches are purchased most of the time – specify
R03_YU1049	Youth	Brand of snus pouches usually/last used – specify
R03_YU1071	Youth	Sub-brand of snus pouches usually/last used – specify
R03_YS1118_OS	Youth	In past 30 days, how you usually got your own smokeless tobacco – specify
R03_YS1033_OS	Youth	Retail location where your smokeless tobacco is bought most of the time – specify
R03_YS1049	Youth	Brand of smokeless tobacco usually/last used – specify
R03_YS1071	Youth	Sub-brand of smokeless tobacco product usually/last used – specify
R03_YZ1002_OS	Youth	Ever used any other tobacco products – specify
R03_YY0601_OS	Youth	First type of tobacco you tried – specify
R03_YX0686_OS	Youth	In past 12 months, liked or followed brand on social media sites: Other – specify
R03_YX0136	Youth	Currently pregnant
R03_YX0137_NN	Youth	Number of weeks/months pregnant – Number
R03_YX0137_UN	Youth	Number of weeks/months pregnant – Unit

**Table E-5. Wave 3 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R03_YX0203_OS	Youth	In past 30 days, has noticed e-cigarettes or other electronic nicotine products being advertised – specify
R03_YX0677_OS	Youth	In past 30 days, noticed cigarettes or other tobacco products being advertised – specify
R03_YM0020	Youth	Last grade/year in school completed
R03_YM0021	Youth	Sexual attraction to gender
R03_YM0063	Youth	Sexual orientation
R03_YM0061	Youth	Transgender
R03_YM0062	Youth	Transgender category
R03_LCYS01	Youth	Language in which CAPI portions of youth interview were conducted

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R03_AM0001	Adult	R03R_A_AGECAT7	Age range when interviewed (7 levels)
R03_AM0002			
R03_AM0003			
R03_AM0069	Adult	R03R_A_AM0069	Respondent is a citizen of the United States (2 levels)
R03_AM0065	Adult	R03R_A_AM0065	Number of years lived in the United States (3 levels)
R03_AC1007_NB	Adult	R03R_A_AC1007_NB	Age range when first started smoking cigarettes fairly regularly (6 levels)
R03_AC1121_NB			
R03_AC1020_NB	Adult	R03R_A_AC1020_NB	Age range when first started smoking cigarettes every day (6 levels)
R03_AC1122_NB			
R03_AV1007_NB	Adult	R03R_A_AV1007_NB	Age range when first started using [EPRODTYPE1]s fairly regularly (6 levels)
R03_AV1121_NB			
R03_AV1020_NB	Adult	R03R_A_AV1020_NB	Age range when first started using [EPRODTYPE1]s every day (6 levels)
R03_AV1122_NB			
R03_AG1007TC_NB	Adult	R03R_A_AG1007TC_NB	Age range when first started smoking traditional cigars fairly regularly (6 levels)
R03_AG1121TC_NB			
R03_AG1020TC_NB	Adult	R03R_A_AG1020TC_NB	Age range when first started smoking cigarillos every day (6 levels)
R03_AG1122TC_NB			
R03_AG1007CG_NB	Adult	R03R_A_AG1007CG_NB	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R03_AG1121CG_NB			
R03_AG1020CG_NB	Adult	R03R_A_AG1020CG_NB	Age range when first started smoking cigarillos every day (6 levels)
R03_AG1122CG_NB			
R03_AG1007FC_NB	Adult	R03R_A_AG1007FC_NB	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R03_AG1121FC_NB			
R03_AG1020FC_NB	Adult	R03R_A_AG1020FC_NB	Age range when first started smoking filtered cigars every day (6 levels)
R03_AG1122FC_NB			
R03_AP1007_NB	Adult	R03R_A_AP1007_NB	Age range when first started smoking a pipe filled with tobacco fairly regularly (6 levels)
R03_AP1121_NB			
R03_AP1020_NB	Adult	R03R_A_AP1020_NB	Age range when first started smoking a pipe filled with tobacco every day (6 levels)
R03_AP1122_NB			
R03_AH1007_NB	Adult	R03R_A_AH1007_NB	Age range when first started smoking hookah fairly regularly (6 levels)
R03_AH1121_NB			
R03_AH1020_NB	Adult	R03R_A_AH1020_NB	Age range when first started smoking hookah every day (6 levels)
R03_AH1122_NB			

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R03_AU1007_NB R03_AU1121_NB	Adult	R03R_A_AU1007_NB	Age range when first started smoking snus pouches fairly regularly (6 levels)
R03_AU1020_NB R03_AU1122_NB	Adult	R03R_A_AU1020_NB	Age range when first started smoking snus pouches every day (6 levels)
R03_AS1007_NB R03_AS1121_NB	Adult	R03R_A_AS1007_NB	Age range when first started using smokeless tobacco fairly regularly (6 levels)
R03_AS1020_NB R03_AS1122_NB	Adult	R03R_A_AS1020_NB	Age range when first started using smokeless tobacco every day (6 levels)
R03_AX0114_NB R03_AX0253_NB	Adult	R03R_A_AX0114_NB	Age range when you were first told you had high blood pressure (6 levels)
R03_AX0115_NB R03_AX0254_NB	Adult	R03R_A_AX0115_NB	Age range when you were first told you had high cholesterol (6 levels)
R03_AX0116_NB R03_AX0255_NB	Adult	R03R_A_AX0116_NB	Age range when you were first told you had congestive heart failure (6 levels)
R03_AX0117_NB R03_AX0256_NB	Adult	R03R_A_AX0117_NB	Age range when you were first told you had a stroke (6 levels)
R03_AX0112_NB R03_AX0252_NB	Adult	R03R_A_AX0112_NB	Age range when you were first told you had a heart attack (6 levels)
R03_AX0120_NB R03_AX0257_NB	Adult	R03R_A_AX0120_NB	Age range when you were first told you had COPD (6 levels)
R03_AX0121_NB R03_AX0258_NB	Adult	R03R_A_AX0121_NB	Age range when you were first told you had chronic bronchitis (6 levels)
R03_AX0123_NB R03_AX0259_NB	Adult	R03R_A_AX0123_NB	Age range when you were first told you had emphysema (6 levels)
R03_AX0124_NB R03_AX0260_NB	Adult	R03R_A_AX0124_NB	Age range when you were first told you had asthma (6 levels)
R03_AX0131_NB R03_AX0261_NB	Adult	R03R_A_AX0131_NB	Age range when you were first told you had gum disease (6 levels)
R03_AX0133_NB R03_AX0262_NB	Adult	R03R_A_AX0133_NB	Age range when you were first told you had pre-cancerous oral lesions (6 levels)
R03_AX0280_NB R03_AX0263_NB	Adult	R03R_A_AX0280_NB	Age range when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes (6 levels)
R03_AX0143_NB R03_AX0264_NB	Adult	R03R_A_AX0143_NB	Age range when you were first told you had an ulcer (6 levels)
R03_AX0148_NB R03_AX0266_NB	Adult	R03R_A_AX0148_NB	Age range when you were first told you had stomach or gastro-intestinal bleeding (6 levels)
R03_AX0150_NB R03_AX0267_NB	Adult	R03R_A_AX0150_NB	Age range when you were first told you had osteoporosis (6 levels)

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R03_AX0198_NB R03_AX0268_NB	Adult	R03R_A_AX0198_NB	Age range when you were first told you had a bone fracture because you have fragile bones (6 levels)
R03_AX0152_NB R03_AX0269_NB	Adult	R03R_A_AX0152_NB	Age range when you were first told you had a cataract or glaucoma (6 levels)
R03_AX0703 R03_AX0704	Adult	R03R_A_AX0703	Age range when you were first told you had macular degeneration (6 levels)
R03_AX0145_02 R03_AX0145_03 R03_AX0145_04 R03_AX0145_05 R03_AX0145_09 R03_AX0145_12 R03_AX0145_15 R03_AX0145_18 R03_AX0145_19 R03_AX0145_21 R03_AX0145_16 R03_AX0145_23 R03_AX0145_24 R03_AX0145_25 R03_AX0145_27 R03_AX0145_29 R03_AX0145_30 R03_AX0145_31	Adult	R03R_A_AX0145_NONTOB	Type of cancer is a non-tobacco related cancer (2 levels)
R03_AX0145_01 R03_AX0145_06 R03_AX0145_07 R03_AX0145_08 R03_AX0145_10 R03_AX0145_11 R03_AX0145_13 R03_AX0145_14 R03_AX0145_17 R03_AX0145_20 R03_AX0145_22 R03_AX0145_26 R03_AX0145_28	Adult	R03R_A_AX0145_TOB	Type of cancer is a tobacco related cancer (2 levels)

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

Questionnaire variable	Instrument	PUF derived variable	PUF derived variable description
R03_AX0146_NB_02	Adult	R03R_A_AX0146_NONTOB	Age range when nontobacco related cancer was diagnosed (6 levels)
R03_AX0265_NB_02			
R03_AX0146_NB_03			
R03_AX0265_NB_03			
R03_AX0146_NB_04			
R03_AX0265_NB_04			
R03_AX0146_NB_05			
R03_AX0265_NB_05			
R03_AX0146_NB_09			
R03_AX0265_NB_09			
R03_AX0146_NB_12			
R03_AX0265_NB_12			
R03_AX0146_NB_15			
R03_AX0265_NB_15			
R03_AX0146_NB_18			
R03_AX0265_NB_18			
R03_AX0146_NB_19			
R03_AX0265_NB_19			
R03_AX0146_NB_21			
R03_AX0265_NB_21			
R03_AX0146_NB_16			
R03_AX0265_NB_16			
R03_AX0146_NB_23			
R03_AX0265_NB_23			
R03_AX0146_NB_24			
R03_AX0265_NB_24			
R03_AX0146_NB_25			
R03_AX0265_NB_25			
R03_AX0146_NB_27			
R03_AX0265_NB_27			
R03_AX0146_NB_29			
R03_AX0265_NB_29			
R03_AX0146_NB_30			
R03_AX0265_NB_30			
R03_AX0146_NB_31			
R03_AX0265_NB_31			

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R03_AX0146_NB_01	Adult	R03R_A_AX0146_TOB	Age range when tobacco related cancer was diagnosed (6 levels)
R03_AX0265_NB_01			
R03_AX0146_NB_06			
R03_AX0265_NB_06			
R03_AX0146_NB_07			
R03_AX0265_NB_07			
R03_AX0146_NB_08			
R03_AX0265_NB_08			
R03_AX0146_NB_10			
R03_AX0265_NB_10			
R03_AX0146_NB_11			
R03_AX0265_NB_11			
R03_AX0146_NB_13			
R03_AX0265_NB_13			
R03_AX0146_NB_14			
R03_AX0265_NB_14			
R03_AX0146_NB_17			
R03_AX0265_NB_17			
R03_AX0146_NB_20			
R03_AX0265_NB_20			
R03_AX0146_NB_22			
R03_AX0265_NB_22			
R03_AX0146_NB_26			
R03_AX0265_NB_26			
R03_AX0146_NB_28			
R03_AX0265_NB_28			
R03_AX0135_12M	Adult	R03R_A_AX0135_12M	Outcome of last pregnancy was miscarriage, induced abortion, ectopic or tubal pregnancy or stillbirth (2 levels)
R03_AX0300_12M_01	Adult	R03R_A_AX0300_12M	Outcome for live birth was preterm birth, low birth weight, birth defects, placenta previa, placenta abruption, preeclampsia, or cleft lip or palate (2 levels)
R03_AX0300_12M_02			
R03_AX0300_12M_03			
R03_AX0300_12M_04			
R03_AX0300_12M_05			
R03_AX0300_12M_06			
R03_AX0300_12M_07			
R03_AX0706			
R03_AM0018	Adult	R03R_A_AM0018	Highest grade or level of school completed (6 levels)
R03_AM0030	Adult	R03R_A_AM0030	Total household income in the past 12 months (5 levels)
R03_AM0038	Adult	R03R_A_AM0038	Type of current residence (7 levels)
R03_AM0033	Adult	R03R_A_AM0033	Highest grade or year of school completed by mother, step-mother or mother-figure (6 levels)
R03_AM0034	Adult	R03R_A_AM0034	Highest grade or year of school completed by father, step-father or father-figure (6 levels)

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R03_AM0036	Adult	R03R_A_AM0036	In past 12 months, parents' total household income (5 levels)
R03_AM0026_01 R03_AM0026_02 R03_AM0026_03 R03_AM0026_04 R03_AM0026_05 R03_AM0026_06 R03_AM0026_07 R03_AM0026_08	Adult	R03R_A_AM0026	Currently covered by health insurance or health coverage plan (5 levels)
R03_AM0063	Adult	R03R_A_SEXORIENT2	Adult sexual orientation (2 levels)
R03_AX0313 R03_AX0679_FT R03_AX0679_IN R03_AX0316 R03_AX0109 R03_AX0312	Adult	R03R_A_BMI	Body mass index
R03_AT0047	Adult	R03R_A_AT0047	Recoded marital status (3 levels)
R03_PT0001	Parent	R03R_Y_PT0001	Recoded parent or guardian relationship to youth (4 levels)
R03_PT0047	Parent	R03R_Y_PT0047	Recoded parent or guardian marital status (3 levels)
R03_PT0045 R03_PM0057	Parent	R03R_P_OTHPAR_INHH	Recoded youth has any parental figures/guardians other than the parent/guardian in the house (2 levels)
R03_PT0002	Parent	R03R_Y_PT0002	Recoded parent's spouse or partner relationship to youth (4 levels)
R03_PM0059	Parent	R03R_Y_PM0059	Recoded other parental figure/guardian's relationship to youth (4 levels)
R03_PT0041_NB R03_PT0253_NB	Parent	R03R_Y_PT0041_NB	Age range youth was first told he/she had high blood pressure (3 levels)
R03_PT0043_NB R03_PT0254_NB	Parent	R03R_Y_PT0043_NB	Age range youth was first told he/she has high cholesterol (3 levels)
R03_PT0038_NB R03_PT0260_NB	Parent	R03R_Y_PT0038_NB	Age range youth was first told he/she has asthma (3 levels)
R03_PT0042_NB R03_PT0263_NB	Parent	R03R_Y_PT0042_NB	Age range youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (3 levels)
R03_PM0069	Parent	R03R_Y_PM0069	Recoded parent respondent is a citizen of the United States (2 levels)

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R03_PM0065	Parent	R03R_Y_PM0065	Recoded number of years parents lived in the United States (3 levels)
R03_PM0001	Parent	R03R_Y_PM0001	Recoded highest grade or year of school completed by parent (6 levels)
R03_PM0118	Parent	R03R_Y_PM0118	Recoded highest grade or year of school completed by spouse/guardian (6 levels)
R03_PM0130	Parent	R03R_Y_PM0130	Recoded total household income in past 12 months (5 levels)
R03_PL0040	Parent	R03R_Y_PL0040	Recoded home is owned or rented (2 levels)
R03_YM0069	Youth	R03R_Y_YM0069	Recoded citizen of the United States (2 levels)
R03_YM0065	Youth	R03R_Y_YM0065	Recoded number of years youth lived in the United States (3 levels)
R03_YC1006_NB R03_YC1120_NB	Youth	R03R_Y_YC1006_NB	Age range when first tried cigarette smoking, even one or two puffs (3 levels)
R03_YC1007 R03_YC1121	Youth	R03R_Y_YC1007	Age range when first started smoking cigarettes fairly regularly (3 levels)
R03_YV1006_NB R03_YV1120_NB	Youth	R03R_Y_YV1006_NB	Age range when first tried [primary electronic nicotine product], even one or two times, even one or two puffs (3 levels)
R03_YV1007 R03_YV1121	Youth	R03R_Y_YV1007	Age range when first started using [primary electronic nicotine product]s fairly regularly (3 levels)
R03_YG1006TC_NB R03_YG1120TC_NB	Youth	R03R_Y_YG1006TC_NB	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
R03_YG1007TC R03_YG1121TC	Youth	R03R_Y_YG1007TC	Age range when first started smoking traditional cigars fairly regularly (3 levels)
R03_YG1006CL_NB R03_YG1120CL_NB	Youth	R03R_Y_YG1006CL_NB	Age range when first tried a cigarillo, even one or two puffs (3 levels)
R03_YG1007CL R03_YG1121CL	Youth	R03R_Y_YG1007CL	Age range when first started smoking cigarillos fairly regularly (3 levels)
R03_YG1006FC_NB R03_YG1120FC_NB	Youth	R03R_Y_YG1006FC_NB	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
R03_YG1007FC R03_YG1121FC	Youth	R03R_Y_YG1007FC	Age range when first started smoking filtered cigars fairly regularly (3 levels)

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R03_YH1006_NB R03_YH1120_NB	Youth	R03R_Y_YH1006_NB	Age range when first tried smoking hookah, even one or two puffs (3 levels)
R03_YH1007 R03_YH1121	Youth	R03R_Y_YH1007	Age range when first started smoking a hookah fairly regularly (3 levels)
R03_YU1006_NB R03_YU1120_NB	Youth	R03R_Y_YU1006_NB	Age range when first tried snus pouches, even one or two times (3 levels)
R03_YU1007 R03_YU1121	Youth	R03R_Y_YU1007	Age range when first started using snus pouches fairly regularly (3 levels)
R03_YS1006_NB R03_YS1120_NB	Youth	R03R_Y_YS1006_NB	Age range when first tried smokeless tobacco, even one or two times (3 levels)
R03_YS1007 R03_YS1121	Youth	R03R_Y_YS1007	Age range when first started using smokeless tobacco fairly regularly (3 levels)
R03_YX0671_01 R03_YX0671_02 R03_YX0671_03 R03_YX0671_04 R03_YX0671_05 R03_YX0671_06 R03_YX0671_07 R03_YX0671_08	Youth	R03R_Y_YX0671	Recoded anyone who lives with you now use tobacco (4 levels)
R03_YT0038_NB R03_YT0260_NB	Youth	R03R_Y_YT0038_NB	Age range when you were first told you had asthma (3 levels)
R03_YX0086_NB R03_YX0087_NB	Youth	R03R_Y_YX0086_NB	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)
R03_YX0074_NB R03_YX0270_NB	Youth	R03R_Y_YX0074_NB	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
R03_YX0079_NB R03_YX0271_NB	Youth	R03R_Y_YX0079_NB	Age range when first used marijuana, hash, THC or grass (3 levels)
R03_YX0082_NB_01 R03_YX0272_NB_01	Youth	R03R_Y_YX0082_NB_01	Age range when first used: Ritalin or Adderall (3 levels)
R03_YX0082_NB_02 R03_YX0272_NB_02	Youth	R03R_Y_YX0082_NB_02	Age range when first used: Painkillers, sedatives or tranquilizers (3 levels)
R03_YX0082_NB_03 R03_YX0272_NB_03	Youth	R03R_Y_YX0082_NB_03	Age range when first used: Cocaine or crack (3 levels)
R03_YX0082_NB_04 R03_YX0272_NB_04	Youth	R03R_Y_YX0082_NB_04	Age range when first used: Stimulants like methamphetamine or speed (3 levels)

**Table E-6. Wave 3 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R03_YX0082_NB_05 R03_YX0272_NB_05	Youth	R03R_Y_YX0082_NB_05	Age range when first used: Any other drugs like heroin, inhalants, solvents or hallucinogens (3 levels)
R03_YM0019 R03_YM0018	Youth	R03R_Y_YM0018	Recoded grade level (If on holiday or break – grade level entering when returning to school) (7 levels)
R03_YM0005_NB_01 R03_YM0005_NB_02 R03_YM0005_NB_03 R03_YM0005_NB_04 R03_YM0005_NB_05	Youth	R03R_Y_HISP	DERIVED – Hispanic origin from the interview (2 levels)
R03_YM0006_NB_01 R03_YM0006_NB_02 R03_YM0006_NB_03 R03_YM0006_NB_04 R03_YM0006_NB_05 R03_YM0006_NB_06 R03_YM0006_NB_07 R03_YM0006_NB_08 R03_YM0006_NB_09 R03_YM0006_NB_10 R03_YM0006_NB_11 R03_YM0006_NB_12 R03_YM0006_NB_13 R03_YM0006_NB_14	Youth	R03R_Y_RACECAT3	Recoded Race from the interview (3 levels)
R03_PT0007_FT R03_PT0007_IN R03_PT0007_MT R03_PT0008_LB R03_PT0008_KG R03_YX0310 R03_YT0007_FT R03_YT0007_IN R03_YX0311 R03_YT0008 R03_YX0312	Youth/ Parent	R03R_Y_BMI	Body mass index

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_CAD10	Adult	Confirm respondent DOB
R04_LAD01	Adult	Respondent's preferred language to complete ACASI interview
R04_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R04_AM0011_NB_01	Adult	Branch served when on active duty: Army
R04_AM0011_NB_02	Adult	Branch served when on active duty: Navy
R04_AM0011_NB_03	Adult	Branch served when on active duty: Air Force
R04_AM0011_NB_04	Adult	Branch served when on active duty: Marine Corps
R04_AM0011_NB_05	Adult	Branch served when on active duty: Coast Guard
R04_AM0072_OS	Adult	Language other than English spoken at home – specify
R04_AZ1002_OS	Adult	Ever used any other tobacco products – specify
R04_AC1033MC_OS	Adult	Retail location where your cigarette is purchased most of the time – specify
R04_AC1049MC	Adult	Brand of cigarettes usually/last smoked – specify
R04_AC1071MC	Adult	Sub-brand of cigarette product usually/last smoked – specify
R04_AC1033RY_OS	Adult	Retail location where your roll-your-own cigarette tobacco is purchased most of the time – specify
R04_AC1049RY	Adult	Brand of roll-your-own cigarette tobacco usually/last smoked – specify
R04_AC1071RY	Adult	Sub-brand of roll-your-own cigarette tobacco product usually/last smoked – specify
R04_AV9003	Adult	Brand of [primary electronic nicotine product] owned – specify
R04_AV1011_OS	Adult	Flavor of [primary electronic nicotine product] when first started using – specify
R04_AV1131_OS	Adult	In past 30 days, [electronic nicotine products/cartridges/e-liquid] flavor used – specify
R04_AV1033_OS	Adult	Retail location where your [electronic nicotine products/cartridges/e-liquid] are bought most of the time – specify
R04_AV1012_OS	Adult	Flavor of regular brand/brand last used – specify
R04_AV1049	Adult	Brand of [electronic nicotine products/cartridges/e-liquid] usually/last used – specify
R04_AJ1033TC_OS	Adult	Retail location where your traditional cigars for blunts is purchased most of the time – specify
R04_AJ1049TC	Adult	Brand of traditional cigars for blunts usually/last smoked – specify
R04_AJ1071TC	Adult	Sub-brand of traditional cigar as a blunt product usually/last smoked – specify
R04_AJ1033CG_OS	Adult	Retail location where your cigarillos for blunts is purchased most of the time – specify
R04_AJ1049CG	Adult	Brand of cigarillos for blunts usually/last smoked – specify
R04_AJ1071CG	Adult	Sub-brand of cigarillo as a blunt product usually/last smoked – specify

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_AJ1033FC_OS	Adult	Retail location where your filtered cigars for blunts is purchased most of the time – specify
R04_AJ1049FC	Adult	Brand of filtered cigars for blunts usually/last smoked – specify
R04_AJ1071FC	Adult	Sub-brand of filtered cigar as a blunt product usually/last smoked – specify
R04_AG1033TC_OS	Adult	Retail location where your traditional cigars is/were purchased most of the time – specify
R04_AG1049TC	Adult	Brand of traditional cigars usually/last smoked – specify
R04_AG1071TC	Adult	Sub-brand of traditional cigar product usually/last smoked – specify
R04_AG1033CG_OS	Adult	Retail location where your cigarillos is/were purchased most of the time – specify
R04_AG1049CG	Adult	Brand of cigarillos usually/last smoked – specify
R04_AG1071CG	Adult	Sub-brand of cigarillo product usually/last smoked – specify
R04_AG1033FC_OS	Adult	Retail location where your filtered cigars is/were purchased most of the time – specify
R04_AG1049FC	Adult	Brand of filtered cigars usually/last smoked – specify
R04_AG1071FC	Adult	Sub-brand of filtered cigar product usually/last smoked – specify
R04_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time – specify
R04_AP1049	Adult	Brand of pipe tobacco usually/last smoked – specify
R04_AP1071	Adult	Sub-brand of pipe tobacco product usually/last smoked – specify
R04_AH9011_OS	Adult	Place where usually smoke/smoked a hookah – specify
R04_AH1033_OS	Adult	Retail location where your hookah tobacco is purchased most of the time – specify
R04_AH1049	Adult	Brand of shisha or hookah tobacco usually/last smoked – specify
R04_AH1071	Adult	Sub-brand of hookah tobacco product usually/last smoked – specify
R04_AU1033_OS	Adult	Retail location where your snus pouches is purchased most of the time – specify
R04_AU1049	Adult	Brand of snus pouches usually/last used – specify
R04_AU1071	Adult	Sub-brand of snus pouch product usually/last used – specify
R04_AS1033_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time – specify
R04_AS1049	Adult	Brand of smokeless tobacco usually/last used – specify
R04_AS1071	Adult	Sub-brand of smokeless tobacco product usually/last used – specify
R04_AX0686_OS	Adult	In past 12 months, liked or followed brand on social media sites – specify
R04_AX0190_NB	Adult	Is deaf or has serious difficulty hearing
R04_AX0191_NB	Adult	Is blind or has serious difficulty seeing, even when wearing glasses
R04_AX0203_OS	Adult	In past 30 days, place where noticed e-cigarettes or other electronic nicotine products being advertised – specify
R04_AX0677_OS	Adult	In past 30 days, noticed cigarettes or other tobacco products being advertised – specify

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_AM0067	Adult	Enrolled in High School
R04_AM0068	Adult	Current grade in school
R04_AM0020_OS	Adult	Type of degree program currently enrolled in – specify
R04_AM0012_12M	Adult	Ever been enrolled in VA Health Care
R04_AM0012_NB	Adult	Ever been enrolled in VA Health Care
R04_AM0021	Adult	Sexual attraction to gender
R04_AM0061	Adult	Consider yourself to be transgender
R04_AM0062	Adult	Consider self to be male-to-female, female-to-male, or non-conforming
R04_LCAD01	Adult	Language in which CAPI portions of adult interview were conducted
R04_AX0217_1A	Adult	Tobacco Product 1, bar code scan
R04_AX0217_2A	Adult	Tobacco Product 2, bar code scan
R04_AX0217_3A	Adult	Tobacco Product 3, bar code scan
R04_AX0217_4A	Adult	Tobacco Product 4, bar code scan
R04_AX0217_5A	Adult	Tobacco Product 5, bar code scan
R04_AX0217_6A	Adult	Tobacco Product 6, bar code scan
R04_AX0217_7A	Adult	Tobacco Product 7, bar code scan
R04_AX0217_8A	Adult	Tobacco Product 8, bar code scan
R04_AX0217_9A	Adult	Tobacco Product 9, bar code scan
R04_AX0217_10A	Adult	Tobacco Product 10, bar code scan
R04_AM0001	Adult	Date of birth (Corrected)
R04_AM0002	Adult	Respondent age
R04_AG1011CG_OS	Adult	Flavor of cigarillo when first started smoking: Some other flavor – specify
R04_AG1011FC_OS	Adult	Flavor of filtered cigar when first started smoking: Some other flavor – specify
R04_AG1011TC_OS	Adult	Flavor of traditional cigar when first started smoking: Some other flavor – specify
R04_AG1012CG_OS	Adult	Cigarillo flavor usually/last smoked: Some other flavor – specify
R04_AG1012FC_OS	Adult	Filtered cigar flavor usually/last smoked: Some other flavor – specify
R04_AG1012TC_OS	Adult	Traditional cigar flavor usually/last smoked: Some other flavor – specify
R04_AG1131CG_OS	Adult	In past 30 days, cigarillo flavor smoked: Some other flavor – specify
R04_AG1131FC_OS	Adult	In past 30 days, filtered cigar flavor smoked: Some other flavor – specify
R04_AG1131TC_OS	Adult	In past 30 days, traditional cigar flavor smoked: Some other flavor – specify
R04_AH1011_OS	Adult	Flavor of hookah tobacco when first started smoking: Some other flavor – specify
R04_AH1012_OS	Adult	Hookah tobacco flavor usually/last smoked: Some other flavor – specify
R04_AH1131_OS	Adult	In past 30 days, hookah tobacco flavor smoked: Some other flavor – specify
R04_AJ1011CG_OS	Adult	Flavor of cigarillo used as blunts when first started smoking: Some other flavor – specify
R04_AJ1011FC_OS	Adult	Flavor of filtered cigar used as blunts when first started smoking: Some other flavor – specify

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_AJ1011TC_OS	Adult	Flavor of traditional cigar used as blunts when first started smoking: Some other flavor – specify
R04_AJ1012CG_OS	Adult	Cigarillo as blunts flavor usually/last smoked: Some other flavor – specify
R04_AJ1012FC_OS	Adult	Filtered cigar as blunts flavor usually/last smoked: Some other flavor – specify
R04_AJ1012TC_OS	Adult	Traditional cigar as blunts flavor usually/last smoked: Some other flavor – specify
R04_AJ1131CG_OS	Adult	In past 30 days, cigarillo as blunts flavor smoked: Some other flavor – specify
R04_AJ1131FC_OS	Adult	In past 30 days, filtered cigar as blunts flavor smoked: Some other flavor – specify
R04_AJ1131TC_OS	Adult	In past 30 days, traditional cigar as blunts flavor smoked: Some other flavor – specify
R04_AN0336_OS	Adult	Main reason you used a nicotine patch, gum, inhaler, nasal spray, lozenge or pill: Some other reason – specify (current established, current experimental or recent former established non-electronic tobacco users)
R04_AN0336E_OS	Adult	Main reason you used a nicotine patch, gum, inhaler, nasal spray, lozenge or pill: Some other reason – specify (current established, recent former established or current experimental electronic nicotine product users)
R04_AP1011_OS	Adult	Flavor of pipe tobacco when first started smoking: Some other flavor – specify
R04_AP1012_OS	Adult	Pipe tobacco flavor usually/last smoked: Some other flavor – specify
R04_AP1131_OS	Adult	In past 30 days, pipe tobacco flavor smoked: Some other flavor – specify
R04_AS1011_OS	Adult	Flavor of smokeless tobacco when first started using: Some other flavor – specify
R04_AS1012_OS	Adult	Smokeless tobacco flavor usually/last used: Some other flavor – specify
R04_AS1131_OS	Adult	In past 30 days, smokeless tobacco flavor used: Some other flavor – specify
R04_AU1011_OS	Adult	Flavor of snus when first started using: Some other flavor – specify
R04_AU1012_OS	Adult	Snus flavor usually/last used: Some other flavor – specify
R04_AU1131_OS	Adult	In past 30 days, snus flavor used: Some other flavor – specify
R04_AM0008_RS	Adult	Currently on active duty in the U.S. Armed Forces
R04_AM0050_RS	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard (Confirmation)
R04_AM0074_RS	Adult	Was on full-time active duty in the U.S. Armed Forces, Military Reserves or National Guard for the entire year in 2014
R04_AM0075_RS	Adult	Lived in the U.S. in any of the 50 states or the District of Columbia during 2014

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_CPT05	Parent	Confirm DOB for Wave 1 shadow youth who is expected to be a youth at Wave 4
R04_CPT05C	Parent	Confirm DOB for Wave 1 youth who is expected to be a youth at Wave 4
R04_CPT05D	Parent	Confirm DOB for Wave 1 shadow youth who is expected to be a shadow youth at Wave 4
R04_CPT07	Parent	Corrected Youth DOB
R04_PM0021_NB	Parent	First names of siblings in multiple birth that are identical to youth
R04_PM0030_NB	Parent	First name of sibling that youth is a twin of
R04_PM0035_NB	Parent	First names of siblings that are in the multiple birth with youth
R04_PM0058_NEW_FNAME	Parent	First name of other parental figure or guardian
R04_PM0058_NEW_LNAME	Parent	Last name of other parental figure or guardian
R04_PM0061_NEW_FNAME	Parent	First name of the second other parental figure or guardian
R04_PM0061_NEW_LNAME	Parent	Last name of the second other parental figure or guardian
R04_PT0046_NEW_FNAME	Parent	First name of spouse/partner
R04_PT0046_NEW_LNAME	Parent	Last name of spouse/partner
R04_PARENT_PERSONID	Parent	Wave 4 Parent/guardian Participant ID Number
R04_PM0053	Parent	Confirm parent's relationship to youth
R04_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative
R04_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative
R04_PM0052	Parent	Confirm parent's spouse is same as reported in previous wave
R04_PT0046_NEW_AGE	Parent	Age of new spouse/partner
R04_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative – specify
R04_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative – specify
R04_PM0058_PERSONID	Parent	Other parental figure/guardian participant ID number
R04_PM0058_NEW_AGE	Parent	Age of other parental figure/guardian
R04_PM0061_PERSONID	Parent	Second other parental figure/guardian PID
R04_PM0061_NEW_AGE	Parent	Age of second other parental figure/guardian
R04_PM0059_OS_1	Parent	Other parental figure/guardian's relationship to sampled youth: Other relative – specify
R04_PM0059_OS_2	Parent	Other parental figure/guardian's relationship to sampled youth: Other non-relative – specify
R04_PM0062_OS_1	Parent	Second other parental figure/guardian's relationship to youth: Other relative – specify
R04_PM0062_OS_2	Parent	Second other parental figure/guardian's relationship to youth: Other non-relative – specify
R04_PM0072_OS	Parent	Other language spoken at home: Some other language – specify (parent respondent)
R04_PM0016_NB	Parent	Youth was identified as a twin or part of a multiple birth
R04_PM0017_NB	Parent	Youth and sibling are identical twins
R04_PM0019_NB	Parent	Any sibling in multiple birth identical to youth
R04_PX0186_NB	Parent	Youth has serious difficulty walking or climbing stairs
R04_PX0191_NB	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
R04_PX0190_NB	Parent	Youth is deaf or has serious difficulty hearing

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
R04_E_YOUTH	Youth	Respondent is an emancipated youth
R04_LYH01	Youth	Youth's preferred language to complete ACASI interview
R04LEY01	Youth	Emancipated Youth's preferred language to complete ACASI interview
R04_YM0007	Youth	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R04_YM0011_01	Youth	Branch served when on active duty: Army
R04_YM0011_02	Youth	Branch served when on active duty: Navy
R04_YM0011_03	Youth	Branch served when on active duty: Air Force
R04_YM0011_04	Youth	Branch served when on active duty: Marine Corps
R04_YM0011_05	Youth	Branch served when on active duty: Coast Guard
R04_YM0066	Youth	How well speak English
R04_YM0070	Youth	How well read English
R04_YM0073	Youth	How well write in English
R04_YM0072_OS	Youth	Language other than English spoken at home: Other language – specify
R04_YC1118_OS	Youth	In past 30 days, how you usually got your own cigarettes – specify
R04_YC1033_OS	Youth	Retail location where your cigarettes are bought most of the time – specify
R04_YC1049	Youth	Brand of cigarettes usually/last smoked – specify
R04_YC1071	Youth	Sub-brand of cigarette product usually/last smoked – specify
R04_YV9035_OS	Youth	Ever used some other electronic nicotine product – specify
R04_YV1011_OS	Youth	Flavor of first [primary electronic nicotine product] used: Specify.
R04_YV1131_OS	Youth	In past 30 days, used [electronic nicotine products/cartridges/e-liquid] flavored to taste like some other flavor – specify
R04_YV1118_OS	Youth	In past 30 days, how you usually got your own [electronic nicotine products/cartridges/e-liquid] – specify
R04_YV1033_OS	Youth	Where[electronic nicotine products/cartridges/e-liquid] is purchased – specify
R04_YV1049	Youth	Brand of e-liquid usually/last used – specify
R04_YJ1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars as a blunt – specify
R04_YJ1033TC_OS	Youth	Retail location where your traditional cigars for blunts are bought most of the time – specify
R04_YJ1049TC	Youth	Brand of traditional cigars usually/last smoked as a blunt – specify
R04_YJ1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked as a blunt – specify
R04_YJ1118CG_OS	Youth	In past 30 days, how you usually got your own cigarillo as a blunt – specify
R04_YJ1033CG_OS	Youth	Retail location where your cigarillos for blunts are bought most of the time – specify

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_YJ1049CG	Youth	Brand of cigarillos usually/last smoked as a blunt – specify
R04_YJ1071CG	Youth	Sub-brand of cigarillo product usually/last smoked as a blunt – specify
R04_YJ1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars as a blunt – specify
R04_YJ1033FC_OS	Youth	Retail location where your filtered cigars for blunts are bought most of the time – specify
R04_YJ1049FC	Youth	Brand of filtered cigars usually/last smoked as a blunt – specify
R04_YJ1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked as a blunt – specify
R04_YG1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars – specify
R04_YG1033TC_OS	Youth	Retail location where your traditional cigars are purchased most of the time – specify
R04_YG1049TC	Youth	Brand of traditional cigars usually/last smoked – specify
R04_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked – specify
R04_YG1118CL_OS	Youth	In past 30 days, how you usually got your own cigarillos – specify
R04_YG1033CL_OS	Youth	Retail location where your cigarillos are purchased most of the time – specify
R04_YG1049CL	Youth	Brand of cigarillos usually/last smoked – specify
R04_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked – specify
R04_YG1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars – specify
R04_YG1033FC_OS	Youth	Retail location where your filtered cigars are purchased most of the time – specify
R04_YG1049FC	Youth	Brand of filtered cigars usually/last smoked – specify
R04_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked – specify
R04_YP1118_OS	Youth	In past 30 days, how you usually got your own pipe tobacco – specify
R04_YP1033_OS	Youth	Retail location where your pipe tobacco are bought most of the time – specify
R04_YH9011_OS	Youth	Usually smoke a hookah: Somewhere else – specify
R04_YH1118_OS	Youth	In past 30 days, how you usually got your own shisha or hookah tobacco – specify
R04_YH1033_OS	Youth	Retail location where your hookah tobacco is purchased most of the time – specify
R04_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
R04_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
R04_YH1073	Youth	Smoke hookah because: It is part of my cultural tradition
R04_YU1118_OS	Youth	In past 30 days, how you usually got your own snus pouches – specify
R04_YU1033_OS	Youth	Retail location where your snus pouches are purchased most of the time – specify
R04_YU1049	Youth	Brand of snus pouches usually/last used – specify
R04_YU1071	Youth	Sub-brand of snus pouches usually/last used – specify

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_YS1118_OS	Youth	In past 30 days, how you usually got your own smokeless tobacco – specify
R04_YS1033_OS	Youth	Retail location where your smokeless tobacco is bought most of the time – specify
R04_YS1049	Youth	Brand of smokeless tobacco usually/last used – specify
R04_YS1071	Youth	Sub-brand of smokeless tobacco product usually/last used – specify
R04_YZ1002_OS	Youth	Ever used any other tobacco products – specify
R04_YY0601_OS	Youth	First type of tobacco you tried – specify
R04_YX0686_OS	Youth	In past 12 months, liked or followed brand on social media sites: Other – specify
R04_YX0136	Youth	Currently pregnant
R04_YX0137_NN	Youth	Number of weeks/months pregnant – Number
R04_YX0137_UN	Youth	Number of weeks/months pregnant – Unit
R04_YX0203_OS	Youth	In past 30 days, has noticed e-cigarettes or other electronic nicotine products being advertised – specify
R04_YX0677_OS	Youth	In past 30 days, noticed cigarettes or other tobacco products being advertised – specify
R04_YM0020	Youth	Last grade/year in school completed
R04_YM0021	Youth	Sexual attraction to gender
R04_YM0063	Youth	Sexual orientation
R04_YM0061	Youth	Transgender
R04_YM0062	Youth	Transgender category
R04_LCYS01	Youth	Language in which CAPI portions of youth interview were conducted
R04_YG1011CL_OS	Youth	Flavor of first cigarillo smoked: Some other flavor – specify
R04_YG1011FC_OS	Youth	Flavor of first filtered cigar smoked: Some other flavor – specify
R04_YG1011TC_OS	Youth	Flavor of first traditional cigar smoked: Some other flavor – specify
R04_YG1131CL_OS	Youth	In past 30 days, smoked cigarillos flavored to taste like: Some other flavor – specify
R04_YG1131FC_OS	Youth	In past 30 days, smoked filtered cigars flavored to taste like: Some other flavor – specify
R04_YG1131TC_OS	Youth	In past 30 days, smoked traditional cigars flavored to taste like: Some other flavor – specify
R04_YH1011_OS	Youth	Flavor of first shisha or hookah tobacco smoked: Some other flavor – specify
R04_YH1131_OS	Youth	In past 30 days, smoked hookah tobacco flavored to taste like: Some other flavor – specify
R04_YJ1011CG_OS	Youth	Flavor of first cigarillo smoked as a blunt: Some other flavor – specify
R04_YJ1011FC_OS	Youth	Flavor of first filtered cigar smoked as a blunt: Some other flavor – specify
R04_YJ1011TC_OS	Youth	Flavor of first traditional cigar smoked as a blunt: Some other flavor – specify
R04_YJ1131CG_OS	Youth	In past 30 days, smoked cigarillos as blunts flavored to taste like: Some other flavor – specify
R04_YJ1131FC_OS	Youth	In past 30 days, smoked filtered cigars as blunts flavored to taste like: Some other flavor – specify

**Table E-7. Wave 4 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R04_YJ1131TC_OS	Youth	In past 30 days, smoked traditional cigars as blunts flavored to taste like: Some other flavor – specify
R04_YS1011_OS	Youth	Flavor of first smokeless tobacco product used: Some other flavor – specify
R04_YS1131_OS	Youth	In past 30 days, used smokeless tobacco flavored to taste like: Some other flavor – specify
R04_YU1011_OS	Youth	Flavor of first snus used: Some other flavor – specify
R04_YU1131_OS	Youth	In past 30 days, used snus flavored to taste like: Some other flavor – specify

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AM0001	Adult	R04R_A_AGECAT6	Age range when interviewed (6 levels)
R04_AM0002			
R04_AM0003			
R04_AM0069	Adult	R04R_A_AM0069	Respondent is a citizen of the United States (2 levels)
R04_AM0065	Adult	R04R_A_AM0065_V2	Number of years lived in the United States (2 levels)
R04_AC1007_NB	Adult	R04R_A_AC1007_NB	Age range when first started smoking cigarettes fairly regularly (6 levels)
R04_AC1121_NB			
R04_AC1020_NB	Adult	R04R_A_AC1020_NB	Age range when first started smoking cigarettes every day (6 levels)
R04_AC1122_NB			
R04_AV1007_NB	Adult	R04R_A_AV1007_NB	Age range when first started using [EPRODTYPE1]s fairly regularly (6 levels)
R04_AV1121_NB			
R04_AV1020_NB	Adult	R04R_A_AV1020_NB	Age range when first started using [EPRODTYPE1]s every day (6 levels)
R04_AG1007TC_NB	Adult	R04R_A_AG1007TC_NB	Age range when first started smoking traditional cigars fairly regularly (6 levels)
R04_AG1121TC_NB			
R04_AG1020TC_NB	Adult	R04R_A_AG1020TC_NB	Age range when first started smoking cigarillos every day (6 levels)
R04_AG1122TC_NB			
R04_AG1007CG_NB	Adult	R04R_A_AG1007CG_NB	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R04_AG1121CG_NB			
R04_AG1020CG_NB	Adult	R04R_A_AG1020CG_NB	Age range when first started smoking cigarillos every day (6 levels)
R04_AG1122CG_NB			
R04_AG1007FC_NB	Adult	R04R_A_AG1007FC_NB	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R04_AG1121FC_NB			
R04_AG1020FC_NB	Adult	R04R_A_AG1020FC_NB	Age range when first started smoking filtered cigars every day (6 levels)
R04_AG1122FC_NB			
R04_AP1007_NB	Adult	R04R_A_AP1007_NB	Age range when first started smoking a pipe filled with tobacco fairly regularly (6 levels)
R04_AP1121_NB			
R04_AP1020_NB	Adult	R04R_A_AP1020_NB	Age range when first started smoking a pipe filled with tobacco every day (6 levels)
R04_AP1122_NB			
R04_AH1007_NB	Adult	R04R_A_AH1007_NB	Age range when first started smoking hookah fairly regularly (6 levels)
R04_AH1121_NB			
R04_AH1020_NB	Adult	R04R_A_AH1020_NB	Age range when first started smoking hookah every day (6 levels)
R04_AH1122_NB			

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AU1007_NB R04_AU1121_NB	Adult	R04R_A_AU1007_NB	Age range when first started smoking snus pouches fairly regularly (6 levels)
R04_AU1020_NB R04_AU1122_NB	Adult	R04R_A_AU1020_NB	Age range when first started smoking snus pouches every day (6 levels)
R04_AS1007_NB R04_AS1121_NB	Adult	R04R_A_AS1007_NB	Age range when first started using smokeless tobacco fairly regularly (6 levels)
R04_AS1020_NB R04_AS1122_NB	Adult	R04R_A_AS1020_NB	Age range when first started using smokeless tobacco every day (6 levels)
R04_AX0114_NB R04_AX0253_NB	Adult	R04R_A_AX0114_NB	Age range when you were first told you had high blood pressure (6 levels)
R04_AX0115_NB R04_AX0254_NB	Adult	R04R_A_AX0115_NB	Age range when you were first told you had high cholesterol (6 levels)
R04_AX0116_NB R04_AX0255_NB	Adult	R04R_A_AX0116_NB	Age range when you were first told you had congestive heart failure (6 levels)
R04_AX0117_NB R04_AX0256_NB	Adult	R04R_A_AX0117_NB	Age range when you were first told you had a stroke (6 levels)
R04_AX0112_NB R04_AX0252_NB	Adult	R04R_A_AX0112_NB	Age range when you were first told you had a heart attack (6 levels)
R04_AX0120_NB R04_AX0257_NB	Adult	R04R_A_AX0120_NB	Age range when you were first told you had COPD (6 levels)
R04_AX0121_NB R04_AX0258_NB	Adult	R04R_A_AX0121_NB	Age range when you were first told you had chronic bronchitis (6 levels)
R04_AX0123_NB R04_AX0259_NB	Adult	R04R_A_AX0123_NB	Age range when you were first told you had emphysema (6 levels)
R04_AX0124_NB R04_AX0260_NB	Adult	R04R_A_AX0124_NB	Age range when you were first told you had asthma (6 levels)
R04_AX0131_NB R04_AX0261_NB	Adult	R04R_A_AX0131_NB	Age range when you were first told you had gum disease (6 levels)
R04_AX0133_NB R04_AX0262_NB	Adult	R04R_A_AX0133_NB	Age range when you were first told you had pre-cancerous oral lesions (6 levels)
R04_AX0280_NB R04_AX0263_NB	Adult	R04R_A_AX0280_NB	Age range when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes (6 levels)
R04_AX0143_NB R04_AX0264_NB	Adult	R04R_A_AX0143_NB	Age range when you were first told you had an ulcer (6 levels)
R04_AX0148_NB R04_AX0266_NB	Adult	R04R_A_AX0148_NB	Age range when you were first told you had stomach or gastro-intestinal bleeding (6 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AX0150_NB R04_AX0267_NB	Adult	R04R_A_AX0150_NB	Age range when you were first told you had osteoporosis (6 levels)
R04_AX0198_NB R04_AX0268_NB	Adult	R04R_A_AX0198_NB	Age range when you were first told you had a bone fracture because you have fragile bones (6 levels)
R04_AX0152_NB R04_AX0269_NB	Adult	R04R_A_AX0152_NB	Age range when you were first told you had a cataract or glaucoma (6 levels)
R04_AX0703 R04_AX0704	Adult	R04R_A_AX0703	Age range when you were first told you had macular degeneration (6 levels)
R04_AX0145_02 R04_AX0145_03 R04_AX0145_04 R04_AX0145_05 R04_AX0145_09 R04_AX0145_12 R04_AX0145_15 R04_AX0145_18 R04_AX0145_19 R04_AX0145_21 R04_AX0145_16 R04_AX0145_23 R04_AX0145_24 R04_AX0145_25 R04_AX0145_27 R04_AX0145_29 R04_AX0145_30 R04_AX0145_31	Adult	R04R_A_AX0145_NONTOB	Type of cancer is a non-tobacco related cancer (2 levels)
R04_AX0145_01 R04_AX0145_06 R04_AX0145_07 R04_AX0145_08 R04_AX0145_10 R04_AX0145_11 R04_AX0145_13 R04_AX0145_14 R04_AX0145_17 R04_AX0145_20 R04_AX0145_22 R04_AX0145_26 R04_AX0145_28	Adult	R04R_A_AX0145_TOB	Type of cancer is a tobacco related cancer (2 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AX0146_NB_02			
R04_AX0265_NB_02			
R04_AX0146_NB_03			
R04_AX0265_NB_03			
R04_AX0146_NB_04			
R04_AX0265_NB_04			
R04_AX0146_NB_05			
R04_AX0265_NB_05			
R04_AX0146_NB_09			
R04_AX0265_NB_09			
R04_AX0146_NB_12			
R04_AX0265_NB_12			
R04_AX0146_NB_15			
R04_AX0265_NB_15			
R04_AX0146_NB_18			
R04_AX0265_NB_18			
R04_AX0146_NB_19			
R04_AX0265_NB_19			
R04_AX0146_NB_21			
R04_AX0265_NB_21			
R04_AX0146_NB_16			
R04_AX0265_NB_16			
R04_AX0146_NB_23			
R04_AX0265_NB_23			
R04_AX0146_NB_24			
R04_AX0265_NB_24			
R04_AX0146_NB_25			
R04_AX0265_NB_25			
R04_AX0146_NB_27			
R04_AX0265_NB_27			
R04_AX0146_NB_29			
R04_AX0265_NB_29			
R04_AX0146_NB_30			
R04_AX0265_NB_30			
R04_AX0146_NB_31			
R04_AX0265_NB_31			

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AX0146_NB_01	Adult	R04R_A_AX0146_TOB	Age range when tobacco related cancer was diagnosed (6 levels)
R04_AX0265_NB_01			
R04_AX0146_NB_06			
R04_AX0265_NB_06			
R04_AX0146_NB_07			
R04_AX0265_NB_07			
R04_AX0146_NB_08			
R04_AX0265_NB_08			
R04_AX0146_NB_10			
R04_AX0265_NB_10			
R04_AX0146_NB_11			
R04_AX0265_NB_11			
R04_AX0146_NB_13			
R04_AX0265_NB_13			
R04_AX0146_NB_14			
R04_AX0265_NB_14			
R04_AX0146_NB_17			
R04_AX0265_NB_17			
R04_AX0146_NB_20			
R04_AX0265_NB_20			
R04_AX0146_NB_22			
R04_AX0265_NB_22			
R04_AX0146_NB_26			
R04_AX0265_NB_26			
R04_AX0146_NB_28			
R04_AX0265_NB_28			
R04_AX0135_12M	Adult	R04R_A_AX0135_12M	Outcome of last pregnancy was miscarriage, induced abortion, ectopic or tubal pregnancy or stillbirth (2 levels)
R04_AX0300_12M_01	Adult	R04R_A_AX0300_12M	Outcome for live birth was preterm birth, low birth weight, birth defects, placenta previa, placenta abruption, preeclampsia, or cleft lip or palate (2 levels)
R04_AX0300_12M_02			
R04_AX0300_12M_03			
R04_AX0300_12M_04			
R04_AX0300_12M_05			
R04_AX0300_12M_06			
R04_AX0300_12M_07			
R04_AX0706			
R04_AM0018	Adult	R04R_A_AM0018_V2	Highest grade or level of school completed (5 levels)
R04_AM0030	Adult	R04R_A_AM0030	Total household income in the past 12 months (5 levels)
R04_AM0033	Adult	R04R_A_AM0033_V2	Highest grade or year of school completed by mother, step-mother or mother-figure (5 levels)
R04_AM0034	Adult	R04R_A_AM0034_V2	Highest grade or year of school completed by father, step-father or father-figure (5 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AM0036	Adult	R04R_A_AM0036	In past 12 months, parents' total household income (5 levels)
R04_AM0026_01 R04_AM0026_02 R04_AM0026_03 R04_AM0026_04 R04_AM0026_05 R04_AM0026_06 R04_AM0026_07 R04_AM0026_08	Adult	R04R_A_AM0026_V2	Currently covered by health insurance or health coverage plan (2 levels)
R04_AM0063	Adult	R04R_A_SEXORIENT2	Adult sexual orientation (2 levels)
R04_AX0313 R04_AX0679_FT R04_AX0679_IN R04_AX0316 R04_AX0109 R04_AX0312	Adult	R04R_A_BMI	Body mass index
R04_AT0047	Adult	R04R_A_AT0047	Recoded marital status (3 levels)
R04_AL0040	Adult	R04R_A_AL0040	Indicator of home ownership (2 levels)
R04_AM0017	Adult	R04R_A_AM0017	Recoded reason for not working for pay (5 levels)
R04_AM0042	Adult	R04R_A_AM0042	Recoded where you currently live (3 levels)
R04_AM0072_02 R04_AM0072_03 R04_AM0072_04 R04_AM0072_05 R04_AM0072_06 R04_AM0072_07 R04_AM0072_08 R04_AM0072_09 R04_AM0072_10	Adult	R04R_A_AM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (adult)
R04_AX0066_01 R04_AX0066_02 R04_AX0066_03 R04_AX0066_04 R04_AX0066_05 R04_AX0066_06 R04_AX0066_07 R04_AX0066_08 R04_AX0066_09	Adult	R04R_A_AX0066	Recoded anyone who lives with you now who uses tobacco (4 levels)
R04_AX0093	Adult	R04R_A_AX0093	Extent to which you are able to carry out your everyday physical activities (2 levels)
R04_AX0691	Adult	R04R_A_AX0691	Recoded type of living space currently living in (4 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AX0757 R04_AX0758	Adult	R04R_A_AX0757	Age range when you were first told by a doctor, therapist or other mental health professional that you had schizophrenia, schizoaffective disorder or psychosis (6 levels)
R04_AX0762 R04_AX0772	Adult	R04R_A_AX0762	Age range when you were first told by a doctor, therapist or other mental health professional that you had a psychotic illness or episode (6 levels)
R04_AC1006_RS R04_AC1120_RS	Adult	R04R_A_AC1006_RS	Age group when first smoked part or all of a cigarette (6 levels)
R04_AG1006TC_RS R04_AG1120TC_RS	Adult	R04R_A_AG1006TC_RS	Age group when first smoked part or all of a traditional cigar, even one or two puffs (6 levels)
R04_AG1006CG_RS R04_AG1120CG_RS	Adult	R04R_A_AG1006CG_RS	Age group when first smoked part or all of a cigarillo, even one or two puffs (6 levels)
R04_AG1006FC_RS R04_AG1120FC_RS	Adult	R04R_A_AG1006FC_RS	Age group when first smoked part or all of a filtered cigar, even one or two puffs (6 levels)
R04_AV1006_RS R04_AV1120_RS	Adult	R04R_A_AV1006_RS	Age group when first used an electronic nicotine product, even one or two times (6 levels)
R04_AU1006_RS R04_AU1120_RS	Adult	R04R_A_AU1006_RS	Age group when first used snus, even one or two times (6 levels)
R04_AP1006_RS R04_AP1120_RS	Adult	R04R_A_AP1006_RS	Age group when you first smoked a pipe filled with tobacco, even one or two puffs (6 levels)
R04_AH1006_RS R04_AH1120_RS	Adult	R04R_A_AH1006_RS	Age group when first smoked tobacco in a hookah, even one or two puffs (6 levels)
R04_AS1006_RS R04_AS1120_RS	Adult	R04R_A_AS1006_RS	Age group when you first used smokeless tobacco, even one or two times (6 levels)
R04_AX0135_RS_01 R04_AX0135_RS_02 R04_AX0135_RS_03 R04_AX0135_RS_04 R04_AX0135_RS_05	Adult	R04R_A_AX0135_RS	Indicator for adverse pregnancy outcomes resulting in no birth (2 levels)
R04_AX0308_RS	Adult	R04R_A_AX0308_RS	Indicator for any pregnancies resulting in a live birth (2 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AX0300_RS_01 R04_AX0300_RS_02 R04_AX0300_RS_03 R04_AX0300_RS_04 R04_AX0300_RS_05 R04_AX0300_RS_06 R04_AX0300_RS_07	Adult	R04R_A_AX0300_RS	Indicator for adverse pregnancy outcomes resulting in risky birth (2 levels)
R04_AX0086_RS R04_AX0087_RS	Adult	R04R_A_AX0086_RS	Age group when you first drank alcohol at all, counting small tastes or sips (6 levels)
R04_AX0074_RS R04_AX0270_RS	Adult	R04R_A_AX0074_RS	Age group when first alcoholic drink was consumed, other than small tastes or sips (6 levels)
R04_AX0079_RS R04_AX0271_RS	Adult	R04R_A_AX0079_RS	Age group when first used marijuana, hash, THC, grass, pot or weed (6 levels)
R04_AX0082_RS_01 R04_AX0272_RS_01	Adult	R04R_A_AX0082_RS_01	Age range when first used: Ritalin or Adderall (6 levels)
R04_AX0082_RS_02 R04_AX0272_RS_02	Adult	R04R_A_AX0082_RS_02	Age range when first used: Painkillers, sedatives, or tranquilizers (6 levels)
R04_AX0082_RS_03 R04_AX0272_RS_03	Adult	R04R_A_AX0082_RS_03	Age range when first used: Cocaine or crack (6 levels)
R04_AX0082_RS_04 R04_AX0272_RS_04	Adult	R04R_A_AX0082_RS_04	Age range when first used: Stimulants like methamphetamine or speed (6 levels)
R04_AX0082_RS_05 R04_AX0272_RS_05	Adult	R04R_A_AX0082_RS_05	Age range when first used: Other drugs like heroin, inhalants, solvents, or hallucinogens (6 levels)
R04_AM0005_RS_01 R04_AM0005_RS_02 R04_AM0005_RS_03 R04_AM0005_RS_04 R04_AM0005_RS_05	Adult	R04R_A_HISP	Hispanic origin from the interview (2 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_AM0006_RS_01 R04_AM0006_RS_02 R04_AM0006_RS_03 R04_AM0006_RS_04 R04_AM0006_RS_05 R04_AM0006_RS_06 R04_AM0006_RS_07 R04_AM0006_RS_08 R04_AM0006_RS_09 R04_AM0006_RS_10 R04_AM0006_RS_11 R04_AM0006_RS_12 R04_AM0006_RS_13 R04_AM0006_RS_14	Adult	R04R_A_RACECAT3	Recoded Race from the interview (3 levels)
R04_PT0001	Parent	R04R_Y_PT0001	Recoded parent or guardian relationship to youth (4 levels)
R04_PT0047	Parent	R04R_Y_PT0047	Recoded parent or guardian marital status (3 levels)
R04_PT0045 R04_PM0057 R04_PM0060	Parent	R04R_P_OTHPAR_INHH	Recoded youth has any parental figures/guardians other than the parent/guardian in the house (2 levels)
R04_PT0002	Parent	R04R_Y_PT0002	Recoded parent's spouse or partner relationship to youth (4 levels)
R04_PM0059	Parent	R04R_Y_PM0059	Recoded other parental figure/guardian's relationship to youth (4 levels)
R04_PM0062	Parent	R04R_Y_PM0062	Second other parental figure/guardian's relationship to youth (4 levels)
R04_PT0041_NB R04_PT0253_NB	Parent	R04R_Y_PT0041_NB	Age range youth was first told he/she had high blood pressure (3 levels)
R04_PT0043_NB R04_PT0254_NB	Parent	R04R_Y_PT0043_NB	Age range youth was first told he/she has high cholesterol (3 levels)
R04_PT0038_NB R04_PT0260_NB	Parent	R04R_Y_PT0038_NB	Age range youth was first told he/she has asthma (3 levels)
R04_PT0042_NB R04_PT0263_NB	Parent	R04R_Y_PT0042_NB	Age range youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (3 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_PM0069	Parent	R04R_Y_PM0069	Recoded parent respondent is a citizen of the United States (2 levels)
R04_PM0065 R04_PM0065_NN	Parent	R04R_Y_PM0065_V2	Recoded number of years parents lived in the United States (2 levels)
R04_PM0001	Parent	R04R_Y_PM0001	Recoded highest grade or year of school completed by parent (6 levels)
R04_PM0118	Parent	R04R_Y_PM0118	Recoded highest grade or year of school completed by spouse/guardian (6 levels)
R04_PM0130	Parent	R04R_Y_PM0130	Recoded total household income in past 12 months (5 levels)
R04_PL0040	Parent	R04R_Y_PL0040	Recoded home is owned or rented (2 levels)
R04_PM0072_02 R04_PM0072_03 R04_PM0072_04 R04_PM0072_05 R04_PM0072_06 R04_PM0072_07 R04_PM0072_08 R04_PM0072_09 R04_PM0072_10	Parent	R04R_Y_PM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (parent respondent)
R04_PX0757 R04_PX0758	Parent	R04R_Y_PX0757	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had schizophrenia, schizoaffective disorder or psychosis (3 levels)
R04_PX0762 R04_PX0772	Parent	R04R_Y_PX0762	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had a psychotic illness or episode (3 levels)
R04_YM0069	Youth	R04R_Y_YM0069	Recoded citizen of the United States (2 levels)
R04_YM0065	Youth	R04R_Y_YM0065_V2	Recoded number of years youth lived in the United States (2 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_YC1006_NB R04_YC1120_NB	Youth	R04R_Y_YC1006_NB	Age range when first tried cigarette smoking, even one or two puffs (3 levels)
R04_YC1007 R04_YC1121	Youth	R04R_Y_YC1007	Age range when first started smoking cigarettes fairly regularly (3 levels)
R04_YV1006_NB R04_YV1120_NB	Youth	R04R_Y_YV1006_NB	Age range when first tried [primary electronic nicotine product], even one or two times, even one or two puffs (3 levels)
R04_YV1007 R04_YV1121	Youth	R04R_Y_YV1007	Age range when first started using [primary electronic nicotine products]s fairly regularly (3 levels)
R04_YG1006TC_NB R04_YG1120TC_NB	Youth	R04R_Y_YG1006TC_NB	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
R04_YG1007TC R04_YG1121TC	Youth	R04R_Y_YG1007TC	Age range when first started smoking traditional cigars fairly regularly (3 levels)
R04_YG1006CL_NB R04_YG1120CL_NB	Youth	R04R_Y_YG1006CL_NB	Age range when first tried a cigarillo, even one or two puffs (3 levels)
R04_YG1007CL R04_YG1121CL	Youth	R04R_Y_YG1007CL	Age range when first started smoking cigarillos fairly regularly (3 levels)
R04_YG1006FC_NB R04_YG1120FC_NB	Youth	R04R_Y_YG1006FC_NB	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
R04_YG1007FC R04_YG1121FC	Youth	R04R_Y_YG1007FC	Age range when first started smoking filtered cigars fairly regularly (3 levels)
R04_YH1006_NB R04_YH1120_NB	Youth	R04R_Y_YH1006_NB	Age range when first tried smoking hookah, even one or two puffs (3 levels)
R04_YH1007 R04_YH1121	Youth	R04R_Y_YH1007	Age range when first started smoking a hookah fairly regularly (3 levels)
R04_YU1006_NB R04_YU1120_NB	Youth	R04R_Y_YU1006_NB	Age range when first tried snus pouches, even one or two times (3 levels)
R04_YU1007 R04_YU1121	Youth	R04R_Y_YU1007	Age range when first started using snus pouches fairly regularly (3 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_YS1006_NB R04_YS1120_NB	Youth	R04R_Y_YS1006_NB	Age range when first tried smokeless tobacco, even one or two times (3 levels)
R04_YS1007 R04_YS1121	Youth	R04R_Y_YS1007	Age range when first started using smokeless tobacco fairly regularly (3 levels)
R04_YX0671_01 R04_YX0671_02 R04_YX0671_03 R04_YX0671_04 R04_YX0671_05 R04_YX0671_06 R04_YX0671_07 R04_YX0671_08	Youth	R04R_Y_YX0671	Recoded anyone who lives with you now use tobacco (4 levels)
R04_YT0038_NB R04_YT0260_NB	Youth	R04R_Y_YT0038_NB	Age range when you were first told you had asthma (3 levels)
R04_YX0086_NB R04_YX0087_NB	Youth	R04R_Y_YX0086_NB	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)
R04_YX0074_NB R04_YX0270_NB	Youth	R04R_Y_YX0074_NB	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
R04_YX0079_NB R04_YX0271_NB	Youth	R04R_Y_YX0079_NB	Age range when first used marijuana, hash, THC or grass (3 levels)
R04_YX0082_NB_01 R04_YX0272_NB_01	Youth	R04R_Y_YX0082_NB_01	Age range when first used: Ritalin or Adderall (3 levels)
R04_YX0082_NB_02 R04_YX0272_NB_02	Youth	R04R_Y_YX0082_NB_02	Age range when first used: Painkillers, sedatives or tranquilizers (3 levels)
R04_YX0082_NB_03 R04_YX0272_NB_03	Youth	R04R_Y_YX0082_NB_03	Age range when first used: Cocaine or crack (3 levels)
R04_YX0082_NB_04 R04_YX0272_NB_04	Youth	R04R_Y_YX0082_NB_04	Age range when first used: Stimulants like methamphetamine or speed (3 levels)
R04_YX0082_NB_05 R04_YX0272_NB_05	Youth	R04R_Y_YX0082_NB_05	Age range when first used: Any other drugs like heroin, inhalants, solvents or hallucinogens (3 levels)
R04_YM0019 R04_YM0018	Youth	R04R_Y_YM0018	Recoded grade level (If on holiday or break – grade level entering when returning to school) (7 levels)

**Table E-8. Wave 4 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R04_YM0005_NB_01 R04_YM0005_NB_02 R04_YM0005_NB_03 R04_YM0005_NB_04 R04_YM0005_NB_05	Youth	R04R_Y_HISP	DERIVED – Hispanic origin from the interview (2 levels)
R04_YM0006_NB_01 R04_YM0006_NB_02 R04_YM0006_NB_03 R04_YM0006_NB_04 R04_YM0006_NB_05 R04_YM0006_NB_06 R04_YM0006_NB_07 R04_YM0006_NB_08 R04_YM0006_NB_09 R04_YM0006_NB_10 R04_YM0006_NB_11 R04_YM0006_NB_12 R04_YM0006_NB_13 R04_YM0006_NB_14	Youth	R04R_Y_RACECAT3	Recoded Race from the interview (3 levels)
R04_PT0007_FT R04_PT0007_IN R04_PT0007_MT R04_PT0008_LB R04_PT0008_KG R04_YX0310 R04_YT0007_FT R04_YT0007_IN R04_YX0311 R04_YT0008 R04_YX0312	Youth/ Parent	R04R_Y_BMI	Body mass index
R04_YM0072_02 R04_YM0072_03 R04_YM0072_04 R04_YM0072_05 R04_YM0072_06 R04_YM0072_07 R04_YM0072_08 R04_YM0072_09 R04_YM0072_10	Youth	R04R_Y_YM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (youth)

**Table E-9. Wave 4.5 questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X04_CPT05	Parent	Confirm DOB for Wave 1 shadow youth who is expected to be a youth at Wave 4.5
X04_CPT05C	Parent	Confirm DOB for Wave 4 youth who is expected to be a youth at Wave 4.5
X04_CPT07	Parent	Corrected Youth DOB
X04_PM0021_NB	Parent	First names of siblings in multiple birth that are identical to youth
X04_PM0030_NB	Parent	First name of sibling that youth is a twin of
X04_PM0035_NB	Parent	First names of siblings that are in the multiple birth with youth
X04_PM0058_NEW_FNAME	Parent	First name of other parental figure or guardian
X04_PM0058_NEW_LNAME	Parent	Last name of other parental figure or guardian
X04_PM0061_NEW_FNAME	Parent	First name of the second other parental figure or guardian
X04_PM0061_NEW_LNAME	Parent	Last name of the second other parental figure or guardian
X04_PT0046_NEW_FNAME	Parent	First name of spouse/partner
X04_PT0046_NEW_LNAME	Parent	Last name of spouse/partner
X04_PARENT_PERSONID	Parent	Wave 4.5 Parent/guardian Participant ID Number
X04_PM0053	Parent	Confirm parent's relationship to youth
X04_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative
X04_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative
X04_PM0052	Parent	Confirm parent's spouse is same as reported in previous wave
X04_PT0046_NEW_AGE	Parent	Age of new spouse/partner
X04_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative – specify
X04_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative – specify
X04_PM0058_PERSONID	Parent	Other parental figure/guardian participant ID number
X04_PM0058_NEW_AGE	Parent	Age of other parental figure/guardian
X04_PM0061_PERSONID	Parent	Second other parental figure/guardian PID
X04_PM0061_NEW_AGE	Parent	Age of second other parental figure/guardian
X04_PM0059_OS_1	Parent	Other parental figure/guardian's relationship to sampled youth: Other relative – specify
X04_PM0059_OS_2	Parent	Other parental figure/guardian's relationship to sampled youth: Other non-relative – specify
X04_PM0062_OS_1	Parent	Second other parental figure/guardian's relationship to youth: Other relative – specify
X04_PM0062_OS_2	Parent	Second other parental figure/guardian's relationship to youth: Other non-relative – specify
X04_PM0072_OS	Parent	Other language spoken at home: Some other language – specify (parent respondent)
X04_PM0016_NB	Parent	Youth was identified as a twin or part of a multiple birth
X04_PM0017_NB	Parent	Youth and sibling are identical twins
X04_PM0019_NB	Parent	Any sibling in multiple birth identical to youth
X04_PX0186_NB	Parent	Youth has serious difficulty walking or climbing stairs
X04_PX0191_NB	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
X04_PX0190_NB	Parent	Youth is deaf or has serious difficulty hearing

**Table E-9. Wave 4.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X04_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
X04_E_YOUTH	Youth	Respondent is an emancipated youth
X04_LYH01	Youth	Youth's preferred language to complete ACASI interview
X04LEY01	Youth	Emancipated Youth's preferred language to complete ACASI interview
X04_YM0007	Youth	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
X04_YM0069	Youth	Youth respondent is a citizen of the United States
X04_YM0011_01	Youth	Branch served when on active duty: Army
X04_YM0011_02	Youth	Branch served when on active duty: Navy
X04_YM0011_03	Youth	Branch served when on active duty: Air Force
X04_YM0011_04	Youth	Branch served when on active duty: Marine Corps
X04_YM0011_05	Youth	Branch served when on active duty: Coast Guard
X04_YM0066	Youth	How well speak English
X04_YM0070	Youth	How well read English
X04_YM0073	Youth	How well write in English
X04_YM0072_OS	Youth	Language other than English spoken at home: Other language – specify
X04_YC1118_OS	Youth	In past 30 days, how you usually got your own cigarettes – specify
X04_YC1033_OS	Youth	Retail location where your cigarettes are bought most of the time – specify
X04_YC1049	Youth	Brand of cigarettes usually/last smoked – specify
X04_YC1071	Youth	Sub-brand of cigarette product usually/last smoked – specify
X04_YV9035_OS	Youth	Ever used some other electronic nicotine product – specify
X04_YV1011_OS	Youth	Flavor of first [primary electronic nicotine product] used: Specify.
X04_YV1131_OS	Youth	In past 30 days, used [electronic nicotine products/cartridges/e-liquid] flavored to taste like some other flavor – specify
X04_YV1118_OS	Youth	In past 30 days, how you usually got your own [electronic nicotine products/cartridges/e-liquid] – specify
X04_YV1033_OS	Youth	Where[electronic nicotine products/cartridges/e-liquid] is purchased – specify
X04_YV1049	Youth	Brand of e-liquid usually/last used – specify
X04_YJ1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars as a blunt – specify
X04_YJ1033TC_OS	Youth	Retail location where your traditional cigars for blunts are bought most of the time – specify
X04_YJ1049TC	Youth	Brand of traditional cigars usually/last smoked as a blunt – specify
X04_YJ1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked as a blunt – specify
X04_YJ1118CG_OS	Youth	In past 30 days, how you usually got your own cigarillo as a blunt – specify
X04_YJ1033CG_OS	Youth	Retail location where your cigarillos for blunts are bought most of the time – specify

**Table E-9. Wave 4.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X04_YJ1049CG	Youth	Brand of cigarillos usually/last smoked as a blunt – specify
X04_YJ1071CG	Youth	Sub-brand of cigarillo product usually/last smoked as a blunt – specify
X04_YJ1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars as a blunt – specify
X04_YJ1033FC_OS	Youth	Retail location where your filtered cigars for blunts are bought most of the time – specify
X04_YJ1049FC	Youth	Brand of filtered cigars usually/last smoked as a blunt – specify
X04_YJ1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked as a blunt – specify
X04_YG1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars – specify
X04_YG1033TC_OS	Youth	Retail location where your traditional cigars are purchased most of the time – specify
X04_YG1049TC	Youth	Brand of traditional cigars usually/last smoked – specify
X04_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked – specify
X04_YG1118CL_OS	Youth	In past 30 days, how you usually got your own cigarillos – specify
X04_YG1033CL_OS	Youth	Retail location where your cigarillos are purchased most of the time – specify
X04_YG1049CL	Youth	Brand of cigarillos usually/last smoked – specify
X04_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked – specify
X04_YG1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars – specify
X04_YG1033FC_OS	Youth	Retail location where your filtered cigars are purchased most of the time – specify
X04_YG1049FC	Youth	Brand of filtered cigars usually/last smoked – specify
X04_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked – specify
X04_YP1118_OS	Youth	In past 30 days, how you usually got your own pipe tobacco – specify
X04_YP1033_OS	Youth	Retail location where your pipe tobacco are bought most of the time – specify
X04_YH9011_OS	Youth	Usually smoke a hookah: Somewhere else – specify
X04_YH1118_OS	Youth	In past 30 days, how you usually got your own shisha or hookah tobacco – specify
X04_YH1033_OS	Youth	Retail location where your hookah tobacco is purchased most of the time – specify
X04_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
X04_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
X04_YH1073	Youth	Smoke hookah because: It is part of my cultural tradition
X04_YU1118_OS	Youth	In past 30 days, how you usually got your own snus pouches – specify
X04_YU1033_OS	Youth	Retail location where your snus pouches are purchased most of the time – specify
X04_YU1049	Youth	Brand of snus pouches usually/last used – specify
X04_YU1071	Youth	Sub-brand of snus pouches usually/last used – specify

**Table E-9. Wave 4.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X04_YS1118_OS	Youth	In past 30 days, how you usually got your own smokeless tobacco – specify
X04_YS1033_OS	Youth	Retail location where your smokeless tobacco is bought most of the time – specify
X04_YS1049	Youth	Brand of smokeless tobacco usually/last used – specify
X04_YS1071	Youth	Sub-brand of smokeless tobacco product usually/last used – specify
X04_YZ1002_OS	Youth	Ever used any other tobacco products – specify
X04_YY0601_OS	Youth	First type of tobacco you tried – specify
X04_YX0686_OS	Youth	In past 12 months, liked or followed brand on social media sites: Other – specify
X04_YX0136	Youth	Currently pregnant
X04_YX0137_NN	Youth	Number of weeks/months pregnant – Number
X04_YX0137_UN	Youth	Number of weeks/months pregnant – Unit
X04_YX0203_OS	Youth	In past 30 days, has noticed e-cigarettes or other electronic nicotine products being advertised – specify
X04_YX0677_OS	Youth	In past 30 days, noticed cigarettes or other tobacco products being advertised – specify
X04_YM0020	Youth	Last grade/year in school completed
X04_YM0021	Youth	Sexual attraction to gender
X04_YM0063	Youth	Sexual orientation
X04_YM0061	Youth	Transgender
X04_YM0062	Youth	Transgender category
X04_LCYS01	Youth	Language in which CAPI portions of youth interview were conducted
X04_YG1011CL_OS	Youth	Flavor of first cigarillo smoked: Some other flavor – specify
X04_YG1011FC_OS	Youth	Flavor of first filtered cigar smoked: Some other flavor – specify
X04_YG1011TC_OS	Youth	Flavor of first traditional cigar smoked: Some other flavor – specify
X04_YG1131CL_OS	Youth	In past 30 days, smoked cigarillos flavored to taste like: Some other flavor – specify
X04_YG1131FC_OS	Youth	In past 30 days, smoked filtered cigars flavored to taste like: Some other flavor – specify
X04_YG1131TC_OS	Youth	In past 30 days, smoked traditional cigars flavored to taste like: Some other flavor – specify
X04_YH1011_OS	Youth	Flavor of first shisha or hookah tobacco smoked: Some other flavor – specify
X04_YH1131_OS	Youth	In past 30 days, smoked hookah tobacco flavored to taste like: Some other flavor – specify
X04_YJ1011CG_OS	Youth	Flavor of first cigarillo smoked as a blunt: Some other flavor – specify
X04_YJ1011FC_OS	Youth	Flavor of first filtered cigar smoked as a blunt: Some other flavor – specify
X04_YJ1011TC_OS	Youth	Flavor of first traditional cigar smoked as a blunt: Some other flavor – specify
X04_YJ1131CG_OS	Youth	In past 30 days, smoked cigarillos as blunts flavored to taste like: Some other flavor – specify
X04_YJ1131FC_OS	Youth	In past 30 days, smoked filtered cigars as blunts flavored to taste like: Some other flavor – specify

**Table E-9. Wave 4.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X04_YJ1131TC_OS	Youth	In past 30 days, smoked traditional cigars as blunts flavored to taste like: Some other flavor – specify
X04_YS1011_OS	Youth	Flavor of first smokeless tobacco product used: Some other flavor – specify
X04_YS1131_OS	Youth	In past 30 days, used smokeless tobacco flavored to taste like: Some other flavor – specify
X04_YU1011_OS	Youth	Flavor of first snus used: Some other flavor – specify
X04_YU1131_OS	Youth	In past 30 days, used snus flavored to taste like: Some other flavor – specify

**Table E-10. Wave 4.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X04_PT0001	Parent	X04R_Y_PT0001_V2	Recoded parent or guardian relationship to youth (3 levels)
X04_PT0047	Parent	X04R_Y_PT0047	Recoded parent or guardian marital status (3 levels)
X04_PT0045 X04_PM0057 X04_PM0060	Parent	X04R_P_OTHPAR_INHH	Recoded youth has any parental figures/guardians other than the parent/guardian in the house (2 levels)
X04_PT0002	Parent	X04R_Y_PT0002_V2	Recoded parent's spouse or partner relationship to youth (3 levels)
X04_PM0059	Parent	X04R_Y_PM0059_V2	Recoded other parental figure/guardian's relationship to youth (3 levels)
X04_PM0062	Parent	X04R_Y_PM0062_V2	Second other parental figure/guardian's relationship to youth (3 levels)
X04_PT0041_NB X04_PT0253_NB	Parent	X04R_Y_PT0041_NB	Age range youth was first told he/she had high blood pressure (3 levels)
X04_PT0043_NB X04_PT0254_NB	Parent	X04R_Y_PT0043_NB	Age range youth was first told he/she has high cholesterol (3 levels)
X04_PT0038_NB X04_PT0260_NB	Parent	X04R_Y_PT0038_NB	Age range youth was first told he/she has asthma (3 levels)
X04_PT0042_NB X04_PT0263_NB	Parent	X04R_Y_PT0042_NB	Age range youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (3 levels)
X04_PM0069	Parent	X04R_Y_PM0069	Recoded parent respondent is a citizen of the United States (2 levels)
X04_PM0065 X04_PM0065_NN	Parent	X04R_Y_PM0065_V2	Recoded number of years parents lived in the United States (2 levels)
X04_PM0001 X04_PM0118	Parent	X04R_P_PARSP_EDUC	Recoded highest grade or year of school completed by parent/spouse/guardian (6 levels)
X04_PM0130	Parent	X04R_Y_PM0130	Recoded total household income in past 12 months (5 levels)
X04_PL0040	Parent	X04R_Y_PL0040	Recoded home is owned or rented (2 levels)
X04_RX0762_NB X04_RX0772_NB	Parent	X04R_Y_RX0762_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had a psychotic illness or episode (3 levels)

**Table E-10. Wave 4.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X04_PX0757_NB X04_PX0758_NB	Parent	X04R_Y_PX0757_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had schizophrenia, schizoaffective disorder or psychosis (3 levels)
X04_PM0072_02 X04_PM0072_03 X04_PM0072_04 X04_PM0072_05 X04_PM0072_06 X04_PM0072_07 X04_PM0072_08 X04_PM0072_09 X04_PM0072_10	Parent	X04R_Y_PM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (parent respondent)
X04_YM0065	Youth	X04R_Y_YM0065_V2	Recoded number of years youth lived in the United States (2 levels)
X04_YC1006_NB X04_YC1120_NB	Youth	X04R_Y_YC1006_NB	Age range when first tried cigarette smoking, even one or two puffs (3 levels)
X04_YC1007 X04_YC1121	Youth	X04R_Y_YC1007	Age range when first started smoking cigarettes fairly regularly (3 levels)
X04_YV1006_NB X04_YV1120_NB	Youth	X04R_Y_YV1006_NB	Age range when first tried [primary electronic nicotine product], even one or two times, even one or two puffs (3 levels)
X04_YV1007 X04_YV1121	Youth	X04R_Y_YV1007	Age range when first started using [primary electronic nicotine product]s fairly regularly (3 levels)
X04_YG1006TC_NB X04_YG1120TC_NB	Youth	X04R_Y_YG1006TC_NB	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
X04_YG1007TC X04_YG1121TC	Youth	X04R_Y_YG1007TC	Age range when first started smoking traditional cigars fairly regularly (3 levels)
X04_YG1006CL_NB X04_YG1120CL_NB	Youth	X04R_Y_YG1006CL_NB	Age range when first tried a cigarillo, even one or two puffs (3 levels)
X04_YG1007CL X04_YG1121CL	Youth	X04R_Y_YG1007CL	Age range when first started smoking cigarillos fairly regularly (3 levels)
X04_YG1006FC_NB X04_YG1120FC_NB	Youth	X04R_Y_YG1006FC_NB	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
X04_YG1007FC X04_YG1121FC	Youth	X04R_Y_YG1007FC	Age range when first started smoking filtered cigars fairly regularly (3 levels)

**Table E-10. Wave 4.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X04_YH1006_NB X04_YH1120_NB	Youth	X04R_Y_YH1006_NB	Age range when first tried smoking hookah, even one or two puffs (3 levels)
X04_YH1007 X04_YH1121	Youth	X04R_Y_YH1007	Age range when first started smoking a hookah fairly regularly (3 levels)
X04_YU1006_NB X04_YU1120_NB	Youth	X04R_Y_YU1006_NB	Age range when first tried snus pouches, even one or two times (3 levels)
X04_YU1007 X04_YU1121	Youth	X04R_Y_YU1007	Age range when first started using snus pouches fairly regularly (3 levels)
X04_YS1006_NB X04_YS1120_NB	Youth	X04R_Y_YS1006_NB	Age range when first tried smokeless tobacco, even one or two times (3 levels)
X04_YS1007 X04_YS1121	Youth	X04R_Y_YS1007	Age range when first started using smokeless tobacco fairly regularly (3 levels)
X04_YX0671_01 X04_YX0671_02 X04_YX0671_03 X04_YX0671_04 X04_YX0671_05 X04_YX0671_06 X04_YX0671_07 X04_YX0671_08	Youth	X04R_Y_YX0671	Recoded anyone who lives with you now use tobacco (4 levels)
X04_YT0038_NB X04_YT0260_NB	Youth	X04R_Y_YT0038_NB	Age range when you were first told you had asthma (3 levels)
X04_YX0086_NB X04_YX0087_NB	Youth	X04R_Y_YX0086_NB	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)
X04_YX0074_NB X04_YX0270_NB	Youth	X04R_Y_YX0074_NB	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
X04_YX0079_NB X04_YX0271_NB	Youth	X04R_Y_YX0079_NB	Age range when first used marijuana, hash, THC or grass (3 levels)
X04_YX0082_NB_01 X04_YX0272_NB_01	Youth	X04R_Y_YX0082_NB_01	Age range when first used: Ritalin or Adderall (3 levels)
X04_YX0082_NB_02 X04_YX0272_NB_02	Youth	X04R_Y_YX0082_NB_02	Age range when first used: Painkillers, sedatives or tranquilizers (3 levels)
X04_YX0082_NB_03 X04_YX0272_NB_03	Youth	X04R_Y_YX0082_NB_03	Age range when first used: Cocaine or crack (3 levels)
X04_YX0082_NB_04 X04_YX0272_NB_04	Youth	X04R_Y_YX0082_NB_04	Age range when first used: Stimulants like methamphetamine or speed (3 levels)

**Table E-10. Wave 4.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X04_YX0082_NB_05 X04_YX0272_NB_05	Youth	X04R_Y_XY0082_NB_05	Age range when first used: Any other drugs like heroin, inhalants, solvents or hallucinogens (3 levels)
X04_YM0019 X04_YM0018	Youth	X04R_Y_YM0018_V2	Recoded grade level (If on holiday or break – grade level entering when returning to school) (6 levels)
X04_YM0005_NB_01 X04_YM0005_NB_02 X04_YM0005_NB_03 X04_YM0005_NB_04 X04_YM0005_NB_05	Youth	X04R_Y_HISP	DERIVED – Hispanic origin from the interview (2 levels)
X04_YM0006_NB_01 X04_YM0006_NB_02 X04_YM0006_NB_03 X04_YM0006_NB_04 X04_YM0006_NB_05 X04_YM0006_NB_06 X04_YM0006_NB_07 X04_YM0006_NB_08 X04_YM0006_NB_09 X04_YM0006_NB_10 X04_YM0006_NB_11 X04_YM0006_NB_12 X04_YM0006_NB_13 X04_YM0006_NB_14	Youth	X04R_Y_RACECAT3	Recoded Race from the interview (3 levels)
X04_PT0007_FT X04_PT0007_IN X04_PT0007_MT X04_PT0008_LB X04_PT0008_KG X04_YX0310 X04_YT0007_FT X04_YT0007_IN X04_YX0311 X04_YT0008 X04_YX0312	Youth/ Parent	X04R_Y_BMI	Body mass index
X04_YM0072_02 X04_YM0072_03 X04_YM0072_04 X04_YM0072_05 X04_YM0072_06 X04_YM0072_07 X04_YM0072_08 X04_YM0072_09 X04_YM0072_10	Youth	X04R_Y_YM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (youth)

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_CAD10	Adult	Confirm respondent DOB
R05_LAD01	Adult	Respondent's preferred language to complete ACASI interview
R05_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R05_AM0011_NB_01	Adult	Branch served when on active duty: Army
R05_AM0011_NB_02	Adult	Branch served when on active duty: Navy
R05_AM0011_NB_03	Adult	Branch served when on active duty: Air Force
R05_AM0011_NB_04	Adult	Branch served when on active duty: Marine Corps
R05_AM0011_NB_05	Adult	Branch served when on active duty: Coast Guard
R05_AM0072_OS	Adult	Other language spoken at home: Some other language – specify
R05_AZ1002_OS	Adult	Ever used any other tobacco products – specify
R05_AC1033MC_OS	Adult	Retail location where your cigarettes are purchased most of the time: Somewhere else – specify
R05_AC1049MC	Adult	Brand of cigarettes usually/last smoked – specify
R05_AC1071MC	Adult	Sub-brand of cigarette product usually/last smoked – specify
R05_AC1033RY_OS	Adult	Retail location where your roll-your-own cigarette tobacco is purchased most of the time: Somewhere else – specify
R05_AC1049RY	Adult	Brand of roll-your-own cigarette tobacco usually/last smoked – specify
R05_AC1071RY	Adult	Sub-brand of roll-your-own cigarette tobacco product usually/last smoked – specify
R05_AV9003	Adult	Brand of electronic nicotine product owned – specify
R05_AV1011_OS	Adult	Flavor of electronic nicotine product when first started using: Some other flavor – specify
R05_AV9035_OS	Adult	Electronic nicotine product you have used: Something else – specify
R05_AV1131_OS	Adult	In past 30 days, [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavor used: Some other flavor – specify
R05_AV1033_OS	Adult	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
R05_AV1012_OS	Adult	Flavor of electronic nicotine product brand regularly/last used: Some other flavor – specify
R05_AV1049	Adult	Brand of electronic nicotine products/electronic nicotine cartridges/e-liquid usually/last used – specify
R05_AJ1033TC_OS	Adult	Retail location where your traditional cigars for blunts are purchased most of the time: Somewhere else – specify
R05_AJ1049TC	Adult	Brand of traditional cigars for blunts usually/last smoked – specify
R05_AJ1071TC	Adult	Sub-brand of traditional cigar as a blunt product usually/last smoked – specify
R05_AJ1033CG_OS	Adult	Retail location where your cigarillos for blunts are purchased most of the time: Somewhere else – specify
R05_AJ1049CG	Adult	Brand of cigarillos for blunts usually/last smoked – specify
R05_AJ1071CG	Adult	Sub-brand of cigarillo as a blunt product usually/last smoked – specify

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_AJ1033FC_OS	Adult	Retail location where your filtered cigars for blunts are purchased most of the time: Somewhere else – specify
R05_AJ1049FC	Adult	Brand of filtered cigars for blunts usually/last smoked – specify
R05_AJ1071FC	Adult	Sub-brand of filtered cigar as a blunt product usually/last smoked – specify
R05_AG1033TC_OS	Adult	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
R05_AG1049TC	Adult	Brand of traditional cigars usually/last smoked – specify
R05_AG1071TC	Adult	Sub-brand of traditional cigar product usually/last smoked – specify
R05_AG1033CG_OS	Adult	Retail location where your cigarillos are purchased most of the time: Somewhere else – specify
R05_AG1049CG	Adult	Brand of cigarillos usually/last smoked – specify
R05_AG1071CG	Adult	Sub-brand of cigarillo product usually/last smoked – specify
R05_AG1033FC_OS	Adult	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
R05_AG1049FC	Adult	Brand of filtered cigars usually/last smoked – specify
R05_AG1071FC	Adult	Sub-brand of filtered cigar product usually/last smoked – specify
R05_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time: Somewhere else – specify
R05_AP1049	Adult	Brand of pipe tobacco usually/last smoked – specify
R05_AP1071	Adult	Sub-brand of pipe tobacco product usually/last smoked – specify
R05_AH9011_OS	Adult	Place where usually smoke/smoked a hookah: Somewhere else -Specify
R05_AH1033_OS	Adult	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify
R05_AH1049	Adult	Brand of shisha or hookah tobacco usually/last smoked – specify
R05_AH1071	Adult	Sub-brand of hookah tobacco product usually/last smoked – specify
R05_AU1033_OS	Adult	Retail location where your snus is purchased most of the time: Somewhere else – specify
R05_AU1049	Adult	Brand of snus usually/last used – specify
R05_AU1071	Adult	Sub-brand of snus product usually/last used – specify
R05_AS1033_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else – specify
R05_AS1049	Adult	Brand of smokeless tobacco usually/last used – specify
R05_AS1071	Adult	Sub-brand of smokeless tobacco product usually/last used – specify
R05_AX0686_OS	Adult	In past 12 months, liked or followed brand on social media sites: Some other brand – specify
R05_AX0190_NB	Adult	Is deaf or has serious difficulty hearing
R05_AX0191_NB	Adult	Is blind or has serious difficulty seeing, even when wearing glasses
R05_AX0203_OS	Adult	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_AX0677_OS	Adult	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
R05_AM0067	Adult	Enrolled in High School
R05_AM0068	Adult	Current grade in school
R05_AM0020_OS	Adult	Type of degree program currently enrolled in: Other type of degree program – specify
R05_AM0012_12M	Adult	In past 12 months, been enrolled in VA Health Care
R05_AM0012_NB	Adult	Ever been enrolled in VA Health Care
R05_AM0021	Adult	Sexual attraction to gender
R05_AM0061	Adult	Consider yourself to be transgender
R05_AM0062	Adult	Consider self to be male-to-female, female-to-male, or non-conforming
R05_LCAD01	Adult	Language in which CAPI portions of adult interview were conducted
R05_AX0217_1A	Adult	Tobacco Product 1, bar code scan
R05_AX0217_2A	Adult	Tobacco Product 2, bar code scan
R05_AX0217_3A	Adult	Tobacco Product 3, bar code scan
R05_AX0217_4A	Adult	Tobacco Product 4, bar code scan
R05_AX0217_5A	Adult	Tobacco Product 5, bar code scan
R05_AX0217_6A	Adult	Tobacco Product 6, bar code scan
R05_AX0217_7A	Adult	Tobacco Product 7, bar code scan
R05_AX0217_8A	Adult	Tobacco Product 8, bar code scan
R05_AX0217_9A	Adult	Tobacco Product 9, bar code scan
R05_AX0217_10A	Adult	Tobacco Product 10, bar code scan
R05_AG1011CG_OS	Adult	Flavor of cigarillo when first started smoking: Some other flavor – specify
R05_AG1011FC_OS	Adult	Flavor of filtered cigar when first started smoking: Some other flavor – specify
R05_AG1011TC_OS	Adult	Flavor of traditional cigar when first started smoking: Some other flavor – specify
R05_AG1012CG_OS	Adult	Cigarillo flavor usually/last smoked: Some other flavor – specify
R05_AG1012FC_OS	Adult	Filtered cigar flavor usually/last smoked: Some other flavor – specify
R05_AG1012TC_OS	Adult	Traditional cigar flavor usually/last smoked: Some other flavor – specify
R05_AG1131CG_OS	Adult	In past 30 days, cigarillo flavor smoked: Some other flavor – specify
R05_AG1131FC_OS	Adult	In past 30 days, filtered cigar flavor smoked: Some other flavor – specify
R05_AG1131TC_OS	Adult	In past 30 days, traditional cigar flavor smoked: Some other flavor – specify
R05_AH1011_OS	Adult	Flavor of hookah tobacco when first started smoking: Some other flavor – specify
R05_AH1012_OS	Adult	Hookah tobacco flavor usually/last smoked: Some other flavor – specify

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_AH1131_OS	Adult	In past 30 days, hookah tobacco flavor smoked: Some other flavor – specify
R05_AJ1011CG_OS	Adult	Flavor of cigarillo used as blunts when first started smoking: Some other flavor – specify
R05_AJ1011FC_OS	Adult	Flavor of filtered cigar used as blunts when first started smoking: Some other flavor – specify
R05_AJ1011TC_OS	Adult	Flavor of traditional cigar used as blunts when first started smoking: Some other flavor – specify
R05_AJ1012CG_OS	Adult	Cigarillo as blunts flavor usually/last smoked: Some other flavor – specify
R05_AJ1012FC_OS	Adult	Filtered cigar as blunts flavor usually/last smoked: Some other flavor – specify
R05_AJ1012TC_OS	Adult	Traditional cigar as blunts flavor usually/last smoked: Some other flavor – specify
R05_AJ1131CG_OS	Adult	In past 30 days, cigarillo as blunts flavor smoked: Some other flavor – specify
R05_AJ1131FC_OS	Adult	In past 30 days, filtered cigar as blunts flavor smoked: Some other flavor – specify
R05_AJ1131TC_OS	Adult	In past 30 days, traditional cigar as blunts flavor smoked: Some other flavor – specify
R05_AN0336_OS	Adult	Main reason you used a nicotine patch, gum, inhaler, nasal spray, or lozenge: Some other reason – specify (current established, current experimental or recent former established non-electronic tobacco users)
R05_AP1011_OS	Adult	Flavor of pipe tobacco when first started smoking: Some other flavor – specify
R05_AP1012_OS	Adult	Pipe tobacco flavor usually/last smoked: Some other flavor – specify
R05_AP1131_OS	Adult	In past 30 days, pipe tobacco flavor smoked: Some other flavor – specify
R05_AS1011_OS	Adult	Flavor of smokeless tobacco when first started using: Some other flavor – specify
R05_AS1012_OS	Adult	Smokeless tobacco flavor usually/last used: Some other flavor – specify
R05_AS1131_OS	Adult	In past 30 days, smokeless tobacco flavor used: Some other flavor – specify
R05_AU1011_OS	Adult	Flavor of snus when first started using: Some other flavor – specify
R05_AU1012_OS	Adult	Snus flavor usually/last used: Some other flavor – specify
R05_AU1131_OS	Adult	In past 30 days, snus flavor used: Some other flavor – specify
R05_AV1133_OS	Adult	In past 30 days, flavor of electronic nicotine product/electronic nicotine cartridge/e-liquid used most often: Some other flavor – specify
R05_AV9001_OS	Adult	Type of electronic nicotine product used most often: Something else – specify
R05_AU1133_OS	Adult	In past 30 days, flavor of snus used most often: Some other flavor – specify
R05_AV0335_OS	Adult	Main reason you stopped using electronic nicotine products: Some other reason – specify
R05_AM0069	Adult	Adult respondent is a citizen of the United States

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_AG1133CG_OS	Adult	In past 30 days, flavor of cigarillo smoked most often: Some other flavor – specify
R05_AG1133FC_OS	Adult	In past 30 days, flavor of filtered cigar smoked most often: Some other flavor – specify
R05_AG1133TC_OS	Adult	In past 30 days, flavor of traditional cigar smoked most often: Some other flavor – specify
R05_AH1133_OS	Adult	In past 30 days, flavor of hookah tobacco smoked most often: Some other flavor – specify
R05_AH9045_OS	Adult	Location when first tried a hookah: Somewhere else – specify
R05_AJ1133CG_OS	Adult	In past 30 days, flavor of cigarillos smoked as blunts most often: Some other flavor – specify
R05_AJ1133FC_OS	Adult	In past 30 days, flavor of filtered cigars smoked as blunts most often: Some other flavor – specify
R05_AJ1133TC_OS	Adult	In past 30 days, flavor of traditional cigars smoked as blunts most often: Some other flavor – specify
R05_AN0349_OS	Adult	Strength of nicotine patch last used (current or recent former established non-electronic tobacco users): Some other strength – specify
R05_AN0349E_OS	Adult	Strength of nicotine patch last used (current or recent former established non-marijuana electronic nicotine product users): Some other strength – specify
R05_AN0349H_OS	Adult	Strength of nicotine patch last used (former experimental or long-term former established tobacco users): Some other strength – specify
R05_AP1133_OS	Adult	In past 30 days, flavor of pipe tobacco smoked most often: Some other flavor – specify
R05_AS1133_OS	Adult	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor – specify
R05_CPT05	Parent	Confirm DOB for Wave 4.5 shadow youth who is expected to be a youth at Wave 5
R05_CPT05C	Parent	Confirm DOB for Wave 4.5 youth who is expected to be a youth at Wave 5
R05_CPT05D	Parent	Confirm DOB for shadow youth who is expected to be a shadow youth at Wave 5
R05_CPT07	Parent	Corrected Youth DOB
R05_E_YOUTH	Youth	Emancipated youth
R05_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
R05_LCY501	Youth	Language in which CAPI portions of youth interview were conducted
R05LEY01	Youth	Emancipated Youth's preferred language to complete ACASI interview
R05_LYH01	Youth	Youth's preferred language to complete ACASI interview
R05_PARENT_PERSONID	Parent	Wave 5 Parent/guardian Participant ID Number
R05_PM0016_NB	Parent	Youth was identified as a twin or part of a multiple birth
R05_PM0017_NB	Parent	Youth and sibling are identical twins
R05_PM0019_NB	Parent	Any sibling in multiple birth identical to youth
R05_PM0021_NB	Parent	First names of siblings in multiple birth that are identical to youth
R05_PM0030_NB	Parent	First name of sibling that youth is a twin of

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_PM0035_NB	Parent	First names of siblings that are in the multiple birth with youth
R05_PM0052	Parent	Confirm parent's spouse is same as reported in previous wave(s)
R05_PM0053	Parent	Confirm parent's relationship to youth
R05_PM0058_NEW_AGE	Parent	Age of other parental figure/guardian
R05_PM0058_NEW_FNAME	Parent	First name of other parental figure or guardian
R05_PM0058_NEW_LNAME	Parent	Last name of other parental figure or guardian
R05_PM0058_PERSONID	Parent	Other parental figure/guardian PERSONID
R05_PM0059_OS_1	Parent	Other parental figure/guardian's relationship to sampled youth: Other relative – specify
R05_PM0059_OS_2	Parent	Other parental figure/guardian's relationship to sampled youth: Other non-relative – specify
R05_PM0061_NEW_AGE	Parent	Age of second other parental figure/guardian
R05_PM0061_NEW_FNAME	Parent	First name of the second other parental figure or guardian
R05_PM0061_NEW_LNAME	Parent	Last name of the second other parental figure or guardian
R05_PM0061_PERSONID	Parent	Second other parental figure/guardian PERSONID
R05_PM0062_OS_1	Parent	Second other parental figure/guardian's relationship to youth: Other relative – specify
R05_PM0062_OS_2	Parent	Second other parental figure/guardian's relationship to youth: Other non-relative – specify
R05_PM0067	Parent	Parent respondent is currently enrolled in high school
R05_PM0068	Parent	Parent respondent grade (if currently enrolled in high school)
R05_PM0072_OS	Parent	Other language spoken at home: Some other language – specify (parent respondent)
R05_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative – specify
R05_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative – specify
R05_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative – specify
R05_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative – specify
R05_PT0046_NEW_AGE	Parent	Age of new spouse/partner
R05_PT0046_NEW_FNAME	Parent	First name of spouse/partner
R05_PT0046_NEW_LNAME	Parent	Last name of spouse/partner
R05_PX0186_NB	Parent	Youth has serious difficulty walking or climbing stairs
R05_PX0190_NB	Parent	Youth is deaf or has serious difficulty hearing
R05_PX0191_NB	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
R05_YC1033_OS	Youth	Retail location where your cigarettes are purchased most of the time: Somewhere else – specify
R05_YC1049	Youth	Brand of cigarettes usually/last smoked – specify
R05_YC1071	Youth	Sub-brand of cigarette product usually/last smoked – specify
R05_YC1118_OS	Youth	In past 30 days, how you usually got your own cigarettes – specify
R05_YG1011CL_OS	Youth	Flavor of first cigarillo smoked: Some other flavor – specify
R05_YG1011FC_OS	Youth	Flavor of first filtered cigar smoked: Some other flavor – specify

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_YG1011TC_OS	Youth	Flavor of first traditional cigar smoked: Some other flavor – specify
R05_YG1033CL_OS	Youth	Retail location where your cigarillos are purchased most of the time: Somewhere else – specify
R05_YG1033FC_OS	Youth	Retail location where your filtered cigars are purchased most of the time: Somewhere else – specify
R05_YG1033TC_OS	Youth	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
R05_YG1049CL	Youth	Brand of cigarillos usually/last smoked – specify
R05_YG1049FC	Youth	Brand of filtered cigars usually/last smoked – specify
R05_YG1049TC	Youth	Brand of traditional cigars usually/last smoked – specify
R05_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked – specify
R05_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked – specify
R05_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked – specify
R05_YG1118CL_OS	Youth	In past 30 days, how you usually got your own cigarillos – specify
R05_YG1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars – specify
R05_YG1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars – specify
R05_YG1131CL_OS	Youth	In past 30 days, smoked cigarillos flavored to taste like: Some other flavor – specify
R05_YG1131FC_OS	Youth	In past 30 days, smoked filtered cigars flavored to taste like: Some other flavor – specify
R05_YG1131TC_OS	Youth	In past 30 days, smoked traditional cigars flavored to taste like: Some other flavor – specify
R05_YG1133CL_OS	Youth	In past 30 days, flavor of cigarillos smoked most often: Some other flavor – specify
R05_YG1133FC_OS	Youth	In past 30 days, flavor of filtered cigars smoked most often: Some other flavor – specify
R05_YG1133TC_OS	Youth	In past 30 days, flavor of traditional cigars smoked most often: Some other flavor – specify
R05_YH1011_OS	Youth	Flavor of first shisha or hookah tobacco smoked: Some other flavor – specify
R05_YH1033_OS	Youth	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify
R05_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
R05_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
R05_YH1073	Youth	Smoke hookah because: It is part of my cultural tradition
R05_YH1118_OS	Youth	In past 30 days, how you usually got your own shisha or hookah tobacco – specify
R05_YH1131_OS	Youth	In past 30 days, smoked hookah tobacco flavored to taste like: Some other flavor – specify
R05_YH1133_OS	Youth	In past 30 days, flavor of hookah smoked most often: Some other flavor – specify
R05_YH9011_OS	Youth	Usually smoke a hookah: Somewhere else – specify
R05_YH9045_OS	Youth	Where you were when you first tried a hookah, even one or two puffs – specify

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_YJ1011CG_OS	Youth	Flavor of first cigarillo smoked as a blunt: Some other flavor – specify
R05_YJ1011FC_OS	Youth	Flavor of first filtered cigar smoked as a blunt: Some other flavor – specify
R05_YJ1011TC_OS	Youth	Flavor of first traditional cigar smoked as a blunt: Some other flavor – specify
R05_YJ1033CG_OS	Youth	Retail location where your cigarillos for blunts are purchased most of the time: Somewhere else – specify
R05_YJ1033FC_OS	Youth	Retail location where your filtered cigars for blunts are purchased most of the time: Somewhere else – specify
R05_YJ1033TC_OS	Youth	Retail location where your traditional cigars for blunts are purchased most of the time: Somewhere else – specify
R05_YJ1049CG	Youth	Brand of cigarillos usually/last smoked as a blunt – specify
R05_YJ1049FC	Youth	Brand of filtered cigars usually/last smoked as a blunt – specify
R05_YJ1049TC	Youth	Brand of traditional cigars usually/last smoked as a blunt – specify
R05_YJ1071CG	Youth	Sub-brand of cigarillo product usually/last smoked as a blunt – specify
R05_YJ1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked as a blunt – specify
R05_YJ1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked as a blunt – specify
R05_YJ1118CG_OS	Youth	In past 30 days, how you usually got your own cigarillo as a blunt – specify
R05_YJ1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars as a blunt – specify
R05_YJ1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars as a blunt – specify
R05_YJ1131CG_OS	Youth	In past 30 days, smoked cigarillos as blunts flavored to taste like: Some other flavor – specify
R05_YJ1131FC_OS	Youth	In past 30 days, smoked filtered cigars as blunts flavored to taste like: Some other flavor – specify
R05_YJ1131TC_OS	Youth	In past 30 days, smoked traditional cigars as blunts flavored to taste like: Some other flavor – specify
R05_YJ1133CG_OS	Youth	In past 30 days, flavor of cigarillos smoked as blunts most often: Some other flavor – specify
R05_YJ1133FC_OS	Youth	In past 30 days, flavor of filtered cigars smoked as blunts most often: Some other flavor – specify
R05_YJ1133TC_OS	Youth	In past 30 days, flavor of traditional cigars smoked as blunts most often: Some other flavor – specify
R05_YM0007	Youth	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R05_YM0011_01	Youth	Branch served when on active duty: Army
R05_YM0011_02	Youth	Branch served when on active duty: Navy
R05_YM0011_03	Youth	Branch served when on active duty: Air Force
R05_YM0011_04	Youth	Branch served when on active duty: Marine Corps
R05_YM0011_05	Youth	Branch served when on active duty: Coast Guard
R05_YM0020	Youth	Last grade/year in school completed

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_YM0021	Youth	Sexual attraction to gender
R05_YM0061	Youth	Consider yourself to be transgender
R05_YM0062	Youth	Transgender category
R05_YM0063	Youth	Sexual orientation
R05_YM0066	Youth	How well you speak English (youth)
R05_YM0069	Youth	Youth respondent is a citizen of the United States
R05_YM0070	Youth	How well you read English (youth)
R05_YM0072_OS	Youth	Other language spoken at home: Some other language - specify (youth respondent)
R05_YM0073	Youth	How well you write in English (youth)
R05_YN0129_OS	Youth	In past 12 months, tried to completely stop using: Other - specify
R05_YN0336_OS	Youth	Main reason you used a nicotine patch, nicotine gum, nicotine inhaler, nicotine nasal spray or lozenge (past year tobacco users) - specify
R05_YN0349_OS	Youth	Strength of nicotine patch last used (past year tobacco users): Some other strength - specify
R05_YN0349H_OS	Youth	Strength of nicotine patch last used (past year tobacco non-users): Some other strength - specify
R05_YP1033_OS	Youth	Retail location where your pipe tobacco is purchased most of the time: Somewhere else - specify
R05_YP1118_OS	Youth	In past 30 days, how you usually got your own pipe tobacco - specify
R05_YS1011_OS	Youth	Flavor of first smokeless tobacco product used: Some other flavor - specify
R05_YS1033_OS	Youth	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else - specify
R05_YS1049	Youth	Brand of smokeless tobacco usually/last used - specify
R05_YS1071	Youth	Sub-brand of smokeless tobacco product usually/last used - specify
R05_YS1118_OS	Youth	In past 30 days, how you usually got your own smokeless tobacco - specify
R05_YS1131_OS	Youth	In past 30 days, used smokeless tobacco flavored to taste like: Some other flavor - specify
R05_YS1133_OS	Youth	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor - specify
R05_YU1011_OS	Youth	Flavor of first snus used: Some other flavor - specify
R05_YU1033_OS	Youth	Retail location where your snus is purchased most of the time: Somewhere else - specify
R05_YU1049	Youth	Brand of snus usually/last used - specify
R05_YU1071	Youth	Sub-brand of snus usually/last used - specify
R05_YU1118_OS	Youth	In past 30 days, how you usually got your own snus - specify
R05_YU1131_OS	Youth	In past 30 days, used snus flavored to taste like: Some other flavor - specify

**Table E-11. Wave 5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R05_YU1133_OS	Youth	In past 30 days, flavor of snus used most often: Some other flavor – specify
R05_YV1011_OS	Youth	Flavor of first electronic nicotine product used: Some other flavor – specify
R05_YV1033_OS	Youth	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
R05_YV1049	Youth	Brand of electronic nicotine products/electronic nicotine cartridges/e-liquid usually/last used – specify
R05_YV1118_OS	Youth	In past 30 days, how you usually got your own [electronic nicotine products/electronic nicotine cartridges/e-liquid] – specify
R05_YV1131_OS	Youth	In past 30 days, used [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavored to taste like: Some other flavor – specify
R05_YV1133_OS	Youth	In past 30 days, flavor of electronic nicotine product/cartridges/e-liquid used most often: Some other flavor – specify
R05_YV9001_OS	Youth	Type of electronic nicotine product used most often – specify
R05_YV9035_OS	Youth	Ever used some other electronic nicotine product – specify
R05_YX0136	Youth	Currently pregnant
R05_YX0137_NN	Youth	Number of weeks/months pregnant – Number
R05_YX0137_UN	Youth	Number of weeks/months pregnant – Unit
R05_YX0203_OS	Youth	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify
R05_YX0677_OS	Youth	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
R05_YX0686_OS	Youth	In past 12 months, liked or followed tobacco or electronic nicotine product brand on social media sites: Some other brand – specify
R05_YY0601_OS	Youth	First type of tobacco you tried – specify
R05_YZ1002_OS	Youth	Ever used any other tobacco products – specify

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_AM0001	Adult	R05R_A_AGECAT6	Age range when interviewed (6 levels)
R05_AM0002			
R05_AM0003			
R05_AM0065	Adult	R05R_A_AM0065_V2	Number of years lived in the United States (2 levels)
R05_AC1007_NB	Adult	R05R_A_AC1007_NB	Age range when first started smoking cigarettes fairly regularly (6 levels)
R05_AC1121_NB			
R05_AC1020_NB	Adult	R05R_A_AC1020_NB	Age range when first started smoking cigarettes every day (6 levels)
R05_AC1122_NB			
R05_AV1007_NB	Adult	R05R_A_AV1007_NB	Age range when first started using [EPRODTYPE1]s fairly regularly (6 levels)
R05_AV1121_NB			
R05_AV1020_NB	Adult	R05R_A_AV1020_NB	Age range when first started using [EPRODTYPE1]s every day (6 levels)
R05_AV1122_NB			
R05_AG1007TC_NB	Adult	R05R_A_AG1007TC_NB	Age range when first started smoking traditional cigars fairly regularly (6 levels)
R05_AG1121TC_NB			
R05_AG1020TC_NB	Adult	R05R_A_AG1020TC_NB	Age range when first started smoking cigarillos every day (6 levels)
R05_AG1122TC_NB			
R05_AG1007CG_NB	Adult	R05R_A_AG1007CG_NB	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R05_AG1121CG_NB			
R05_AG1020CG_NB	Adult	R05R_A_AG1020CG_NB	Age range when first started smoking cigarillos every day (6 levels)
R05_AG1122CG_NB			
R05_AG1007FC_NB	Adult	R05R_A_AG1007FC_NB	Age range when first started smoking filtered cigars fairly regularly (6 levels)
R05_AG1121FC_NB			
R05_AG1020FC_NB	Adult	R05R_A_AG1020FC_NB	Age range when first started smoking filtered cigars every day (6 levels)
R05_AG1122FC_NB			
R05_AP1007_NB	Adult	R05R_A_AP1007_NB	Age range when first started smoking a pipe filled with tobacco fairly regularly (6 levels)
R05_AP1121_NB			
R05_AP1020_NB	Adult	R05R_A_AP1020_NB	Age range when first started smoking a pipe filled with tobacco every day (6 levels)
R05_AP1122_NB			
R05_AH1007_NB	Adult	R05R_A_AH1007_NB	Age range when first started smoking hookah fairly regularly (6 levels)
R05_AH1121_NB			
R05_AH1020_NB	Adult	R05R_A_AH1020_NB	Age range when first started smoking hookah every day (6 levels)
R05_AH1122_NB			
R05_AU1007_NB	Adult	R05R_A_AU1007_NB	Age range when first started smoking snus pouches fairly regularly (6 levels)
R05_AU1121_NB			
R05_AU1020_NB	Adult	R05R_A_AU1020_NB	Age range when first started smoking snus pouches every day (6 levels)
R05_AU1122_NB			
R05_AS1007_NB	Adult	R05R_A_AS1007_NB	Age range when first started using smokeless tobacco fairly regularly (6 levels)
R05_AS1121_NB			
R05_AS1020_NB	Adult	R05R_A_AS1020_NB	Age range when first started using smokeless tobacco every day (6 levels)
R05_AS1122_NB			
R05_AX0114_NB	Adult	R05R_A_AX0114_NB	Age range when you were first told you had high blood pressure (6 levels)
R05_AX0253_NB			
R05_AX0115_NB	Adult	R05R_A_AX0115_NB	Age range when you were first told you had high cholesterol (6 levels)
R05_AX0254_NB			

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_AX0116_NB R05_AX0255_NB	Adult	R05R_A_AX0116_NB	Age range when you were first told you had congestive heart failure (6 levels)
R05_AX0117_NB R05_AX0256_NB	Adult	R05R_A_AX0117_NB	Age range when you were first told you had a stroke (6 levels)
R05_AX0112_NB R05_AX0252_NB	Adult	R05R_A_AX0112_NB	Age range when you were first told you had a heart attack (6 levels)
R05_AX0120_NB R05_AX0257_NB	Adult	R05R_A_AX0120_NB	Age range when you were first told you had COPD (6 levels)
R05_AX0121_NB R05_AX0258_NB	Adult	R05R_A_AX0121_NB	Age range when you were first told you had chronic bronchitis (6 levels)
R05_AX0123_NB R05_AX0259_NB	Adult	R05R_A_AX0123_NB	Age range when you were first told you had emphysema (6 levels)
R05_AX0124_NB R05_AX0260_NB	Adult	R05R_A_AX0124_NB	Age range when you were first told you had asthma (6 levels)
R05_AX0131_NB R05_AX0261_NB	Adult	R05R_A_AX0131_NB	Age range when you were first told you had gum disease (6 levels)
R05_AX0133_NB R05_AX0262_NB	Adult	R05R_A_AX0133_NB	Age range when you were first told you had pre-cancerous oral lesions (6 levels)
R05_AX0280_NB R05_AX0263_NB	Adult	R05R_A_AX0280_NB	Age range when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes (6 levels)
R05_AX0143_NB R05_AX0264_NB	Adult	R05R_A_AX0143_NB	Age range when you were first told you had an ulcer (6 levels)
R05_AX0148_NB R05_AX0266_NB	Adult	R05R_A_AX0148_NB	Age range when you were first told you had stomach or gastro-intestinal bleeding (6 levels)
R05_AX0150_NB R05_AX0267_NB	Adult	R05R_A_AX0150_NB	Age range when you were first told you had osteoporosis (6 levels)
R05_AX0198_NB R05_AX0268_NB	Adult	R05R_A_AX0198_NB	Age range when you were first told you had a bone fracture because you have fragile bones (6 levels)
R05_AX0152_NB R05_AX0269_NB	Adult	R05R_A_AX0152_NB	Age range when you were first told you had a cataract or glaucoma (6 levels)
R05_AX0703 R05_AX0704	Adult	R05R_A_AX0703	Age range when you were first told you had macular degeneration (6 levels)

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_AX0145_02	Adult	R05R_A_AX0145_NONTOB	Type of cancer is a non-tobacco related cancer (2 levels)
R05_AX0145_03			
R05_AX0145_04			
R05_AX0145_05			
R05_AX0145_09			
R05_AX0145_12			
R05_AX0145_15			
R05_AX0145_18			
R05_AX0145_19			
R05_AX0145_21			
R05_AX0145_16			
R05_AX0145_23			
R05_AX0145_24			
R05_AX0145_25			
R05_AX0145_27			
R05_AX0145_29			
R05_AX0145_30			
R05_AX0145_31			
R05_AX0145_01	Adult	R05R_A_AX0145_TOB	Type of cancer is a tobacco related cancer (2 levels)
R05_AX0145_06			
R05_AX0145_07			
R05_AX0145_08			
R05_AX0145_10			
R05_AX0145_11			
R05_AX0145_13			
R05_AX0145_14			
R05_AX0145_17			
R05_AX0145_20			
R05_AX0145_22			
R05_AX0145_26			
R05_AX0145_28			

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_AX0146_NB_02	Adult	R05R_A_AX0146_NONTOB	Age range when nontobacco related cancer was diagnosed (6 levels)
R05_AX0265_NB_02			
R05_AX0146_NB_03			
R05_AX0265_NB_03			
R05_AX0146_NB_04			
R05_AX0265_NB_04			
R05_AX0146_NB_05			
R05_AX0265_NB_05			
R05_AX0146_NB_09			
R05_AX0265_NB_09			
R05_AX0146_NB_12			
R05_AX0265_NB_12			
R05_AX0146_NB_15			
R05_AX0265_NB_15			
R05_AX0146_NB_18			
R05_AX0265_NB_18			
R05_AX0146_NB_19			
R05_AX0265_NB_19			
R05_AX0146_NB_21			
R05_AX0265_NB_21			
R05_AX0146_NB_16			
R05_AX0265_NB_16			
R05_AX0146_NB_23			
R05_AX0265_NB_23			
R05_AX0146_NB_24			
R05_AX0265_NB_24			
R05_AX0146_NB_25			
R05_AX0265_NB_25			
R05_AX0146_NB_27			
R05_AX0265_NB_27			
R05_AX0146_NB_29			
R05_AX0265_NB_29			
R05_AX0146_NB_30			
R05_AX0265_NB_30			
R05_AX0146_NB_31			
R05_AX0265_NB_31			

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_AX0146_NB_01	Adult	R05R_A_AX0146_TOB	Age range when tobacco related cancer was diagnosed (6 levels)
R05_AX0265_NB_01			
R05_AX0146_NB_06			
R05_AX0265_NB_06			
R05_AX0146_NB_07			
R05_AX0265_NB_07			
R05_AX0146_NB_08			
R05_AX0265_NB_08			
R05_AX0146_NB_10			
R05_AX0265_NB_10			
R05_AX0146_NB_11			
R05_AX0265_NB_11			
R05_AX0146_NB_13			
R05_AX0265_NB_13			
R05_AX0146_NB_14			
R05_AX0265_NB_14			
R05_AX0146_NB_17			
R05_AX0265_NB_17			
R05_AX0146_NB_20			
R05_AX0265_NB_20			
R05_AX0146_NB_22			
R05_AX0265_NB_22			
R05_AX0146_NB_26			
R05_AX0265_NB_26			
R05_AX0146_NB_28			
R05_AX0265_NB_28			
R05_AX0135_12M	Adult	R05R_A_AX0135_12M	Outcome of last pregnancy was miscarriage, induced abortion, ectopic or tubal pregnancy or stillbirth (2 levels)
R05_AX0300_12M_01	Adult	R05R_A_AX0300_12M	Outcome for live birth was preterm birth, low birth weight, birth defects, placenta previa, placenta abruption, preeclampsia, or cleft lip or palate (2 levels)
R05_AX0300_12M_02			
R05_AX0300_12M_03			
R05_AX0300_12M_04			
R05_AX0300_12M_05			
R05_AX0300_12M_06			
R05_AX0300_12M_07			
R05_AX0706			
R05_AM0018	Adult	R05R_A_AM0018_V2	Highest grade or level of school completed (5 levels)
R05_AM0030	Adult	R05R_A_AM0030	Total household income in the past 12 months (5 levels)
R05_AM0033	Adult	R05R_A_AM0033_V2	Highest grade or year of school completed by mother, step-mother or mother-figure (5 levels)
R05_AM0034	Adult	R05R_A_AM0034_V2	Highest grade or year of school completed by father, step-father or father-figure (5 levels)
R05_AM0036	Adult	R05R_A_AM0036	In past 12 months, parents' total household income (5 levels)

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_AM0026_01	Adult	R05R_A_AM0026_V2	Currently covered by health insurance or health coverage plan (2 levels)
R05_AM0026_02			
R05_AM0026_03			
R05_AM0026_04			
R05_AM0026_05			
R05_AM0026_06			
R05_AM0026_07			
R05_AM0026_08			
R05_AM0063	Adult	R05R_A_SEXORIENT2	Adult sexual orientation (2 levels)
R05_AX0313	Adult	R05R_A_BMI	Body mass index
R05_AX0679_FT			
R05_AX0679_IN			
R05_AX0316			
R05_AX0109			
R05_AX0312			
R05_AT0047	Adult	R05R_A_AT0047	Recoded marital status (3 levels)
R05_AL0040	Adult	R05R_A_AL0040	Indicator of home ownership (2 levels)
R05_AM0017	Adult	R05R_A_AM0017	Recoded reason for not working for pay (5 levels)
R05_AM0042	Adult	R05R_A_AM0042	Recoded where you currently live (3 levels)
R05_AM0072_02	Adult	R05R_A_AM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (adult)
R05_AM0072_03			
R05_AM0072_04			
R05_AM0072_05			
R05_AM0072_06			
R05_AM0072_07			
R05_AM0072_08			
R05_AM0072_09			
R05_AM0072_10			
R05_AX0066_01	Adult	R05R_A_AX0066	Recoded anyone who lives with you now: Cigarettes, cigars, cigarillos or filtered cigars, pipe tobacco, e-products exclusively, other tobacco products, including smokeless, snus and hookah, or no one living in the home uses tobacco (4 levels)
R05_AX0066_02			
R05_AX0066_03			
R05_AX0066_04			
R05_AX0066_05			
R05_AX0066_06			
R05_AX0066_07			
R05_AX0066_08			
R05_AX0066_09			
R05_AX0093	Adult	R05R_A_AX0093	Extent to which you are able to carry out your everyday physical activities (2 levels)
R05_AX0691	Adult	R05R_A_AX0691	Recoded type of living space currently living in (4 levels)
R05_AX0757_NB	Adult	R05R_A_AX0757_NB	Age range when you were first told by a doctor, therapist or other mental health professional that you had schizophrenia, schizoaffective disorder or psychosis (6 levels)
R05_AX0758_NB			

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_AX0762_NB R05_AX0772_NB	Adult	R05R_A_AX0762_NB	Age range when you were first told by a doctor, therapist or other mental health professional that you had a psychotic illness or episode (6 levels)
R05_AX0411 R05_AX0412	Adult	R05R_A_AX0411	Age range when you were first told you have poor circulation (PAD or PVD) (6 levels)
R05_PT0001	Parent	R05R_Y_PT0001_V2	Recoded parent or guardian relationship to youth (3 levels)
R05_PT0047	Parent	R05R_Y_PT0047	Recoded parent or guardian marital status (3 levels)
R05_PT0045 R05_PM0057 R05_PM0060	Parent	R05R_P_OTHPAR_INHH	Recoded youth has any parental figures/guardians other than the parent/guardian in the house (2 levels)
R05_PT0002	Parent	R05R_Y_PT0002_V2	Recoded parent's spouse or partner relationship to youth (3 levels)
R05_PM0059	Parent	R05R_Y_PM0059_V2	Recoded other parental figure/guardian's relationship to youth (3 levels)
R05_PM0062	Parent	R05R_Y_PM0062_V2	Second other parental figure/guardian's relationship to youth (3 levels)
R05_PT0041_NB R05_PT0253_NB	Parent	R05R_Y_PT0041_NB	Age range youth was first told he/she had high blood pressure (3 levels)
R05_PT0043_NB R05_PT0254_NB	Parent	R05R_Y_PT0043_NB	Age range youth was first told he/she has high cholesterol (3 levels)
R05_PT0038_NB R05_PT0260_NB	Parent	R05R_Y_PT0038_NB	Age range youth was first told he/she has asthma (3 levels)
R05_PT0042_NB R05_PT0263_NB	Parent	R05R_Y_PT0042_NB	Age range youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (3 levels)
R05_PM0069	Parent	R05R_Y_PM0069	Recoded parent respondent is a citizen of the United States (2 levels)
R05_PM0065 R05_PM0065_NN	Parent	R05R_Y_PM0065_V2	Recoded number of years parents lived in the United States (2 levels)
R05_PM0001 R05_PM0118	Parent	R05R_P_PARSP_EDUC	Recoded highest grade or year of school completed by parent/spouse/guardian (6 levels)
R05_PM0130	Parent	R05R_Y_PM0130	Recoded total household income in past 12 months (5 levels)
R05_PL0040	Parent	R05R_Y_PL0040	Recoded home is owned or rented (2 levels)

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_PM0072_02 R05_PM0072_03 R05_PM0072_04 R05_PM0072_05 R05_PM0072_06 R05_PM0072_07 R05_PM0072_08 R05_PM0072_09 R05_PM0072_10	Parent	R05R_Y_PM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (parent respondent)
R05_PX0757_NB R05_PX0758_NB	Parent	R05R_Y_PX0757_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had schizophrenia, schizoaffective disorder or psychosis (3 levels)
R05_PX0762_NB R05_PX0772_NB	Parent	R05R_Y_PX0762_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had a psychotic illness or episode (3 levels)
R05_YM0065	Youth	R05R_Y_YM0065_V2	Recoded number of years youth lived in the United States (2 levels)
R05_YC1006_NB R05_YC1120_NB	Youth	R05R_Y_YC1006_NB	Age range when first tried cigarette smoking, even one or two puffs (3 levels)
R05_YC1007 R05_YC1121	Youth	R05R_Y_YC1007	Age range when first started smoking cigarettes fairly regularly (3 levels)
R05_YV1006_NB R05_YV1120_NB	Youth	R05R_Y_YV1006_NB	Age range when first tried [primary electronic nicotine product], even one or two times, even one or two puffs (3 levels)
R05_YV1007 R05_YV1121	Youth	R05R_Y_YV1007	Age range when first started using [primary electronic nicotine product]s fairly regularly (3 levels)
R05_YG1006TC_NB R05_YG1120TC_NB	Youth	R05R_Y_YG1006TC_NB	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
R05_YG1007TC R05_YG1121TC	Youth	R05R_Y_YG1007TC	Age range when first started smoking traditional cigars fairly regularly (3 levels)
R05_YG1006CL_NB R05_YG1120CL_NB	Youth	R05R_Y_YG1006CL_NB	Age range when first tried a cigarillo, even one or two puffs (3 levels)
R05_YG1007CL R05_YG1121CL	Youth	R05R_Y_YG1007CL	Age range when first started smoking cigarillos fairly regularly (3 levels)
R05_YG1006FC_NB R05_YG1120FC_NB	Youth	R05R_Y_YG1006FC_NB	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
R05_YG1007FC R05_YG1121FC	Youth	R05R_Y_YG1007FC	Age range when first started smoking filtered cigars fairly regularly (3 levels)
R05_YH1006_NB R05_YH1120_NB	Youth	R05R_Y_YH1006_NB	Age range when first tried smoking hookah, even one or two puffs (3 levels)

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_YH1007 R05_YH1121	Youth	R05R_Y_YH1007	Age range when first started smoking a hookah fairly regularly (3 levels)
R05_YU1006_NB R05_YU1120_NB	Youth	R05R_Y_YU1006_NB	Age range when first tried snus pouches, even one or two times (3 levels)
R05_YU1007 R05_YU1121	Youth	R05R_Y_YU1007	Age range when first started using snus pouches fairly regularly (3 levels)
R05_YS1006_NB R05_YS1120_NB	Youth	R05R_Y_YS1006_NB	Age range when first tried smokeless tobacco, even one or two times (3 levels)
R05_YS1007 R05_YS1121	Youth	R05R_Y_YS1007	Age range when first started using smokeless tobacco fairly regularly (3 levels)
R05_YX0671_01 R05_YX0671_02 R05_YX0671_03 R05_YX0671_04 R05_YX0671_05 R05_YX0671_06 R05_YX0671_07 R05_YX0671_08	Youth	R05R_Y_YX0671	Recoded anyone who lives with you now use tobacco (4 levels)
R05_YT0038_NB R05_YT0260_NB	Youth	R05R_Y_YT0038_NB	Age range when you were first told you had asthma (3 levels)
R05_YX0086_NB R05_YX0087_NB	Youth	R05R_Y_YX0086_NB	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)
R05_YX0074_NB R05_YX0270_NB	Youth	R05R_Y_YX0074_NB	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
R05_YX0079_NB R05_YX0271_NB	Youth	R05R_Y_YX0079_NB	Age range when first used marijuana, hash, THC or grass (3 levels)
R05_YX0082_NB_01 R05_YX0272_NB_01	Youth	R05R_Y_YX0082_NB_01	Age range when first used: Ritalin or Adderall (3 levels)
R05_YX0082_NB_02 R05_YX0272_NB_02	Youth	R05R_Y_YX0082_NB_02	Age range when first used: Painkillers, sedatives or tranquilizers (3 levels)
R05_YX0082_NB_03 R05_YX0272_NB_03	Youth	R05R_Y_YX0082_NB_03	Age range when first used: Cocaine or crack (3 levels)
R05_YX0082_NB_04 R05_YX0272_NB_04	Youth	R05R_Y_YX0082_NB_04	Age range when first used: Stimulants like methamphetamine or speed (3 levels)
R05_YX0082_NB_05 R05_YX0272_NB_05	Youth	R05R_Y_YX0082_NB_05	Age range when first used: Any other drugs like heroin, inhalants, solvents or hallucinogens (3 levels)
R05_YM0019 R05_YM0018	Youth	R05R_Y_YM0018_V2	Recoded grade level (If on holiday or break – grade level entering when returning to school) (6 levels)
R05_YM0004_NB	Youth	R05R_Y_SEX	Gender

**Table E-12. Wave 5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R05_YM0005_NB_01 R05_YM0005_NB_02 R05_YM0005_NB_03 R05_YM0005_NB_04 R05_YM0005_NB_05	Youth	R05R_Y_HISP	DERIVED – Hispanic origin from the interview (2 levels)
R05_YM0006_NB_01 R05_YM0006_NB_02 R05_YM0006_NB_03 R05_YM0006_NB_04 R05_YM0006_NB_05 R05_YM0006_NB_06 R05_YM0006_NB_07 R05_YM0006_NB_08 R05_YM0006_NB_09 R05_YM0006_NB_10 R05_YM0006_NB_11 R05_YM0006_NB_12 R05_YM0006_NB_13 R05_YM0006_NB_14	Youth	R05R_Y_RACECAT3	Recoded Race from the interview (3 levels)
R05_PT0007_FT R05_PT0007_IN R05_PT0007_MT R05_PT0008_LB R05_PT0008_KG R05_YX0310 R05_YT0007_FT R05_YT0007_IN R05_YX0311 R05_YT0008 R05_YX0312	Youth/ Parent	R05R_Y_BMI	Body mass index
R05_YM0072_02 R05_YM0072_03 R05_YM0072_04 R05_YM0072_05 R05_YM0072_06 R05_YM0072_07 R05_YM0072_08 R05_YM0072_09 R05_YM0072_10	Youth	R05R_Y_YM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (youth)

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_CAD10	Adult	Confirm DOB (for continuing respondents)
X05_LAD01	Adult	Respondent's preferred language to complete ACASI interview
X05_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
X05_AM0011_NB_01	Adult	Branch served when on active duty: Army
X05_AM0011_NB_02	Adult	Branch served when on active duty: Navy
X05_AM0011_NB_03	Adult	Branch served when on active duty: Air Force
X05_AM0011_NB_04	Adult	Branch served when on active duty: Marine Corps
X05_AM0011_NB_05	Adult	Branch served when on active duty: Coast Guard
X05_AM0017	Adult	Reason for not working for pay
X05_AM0072_OS	Adult	Other language spoken at home: Some other language – specify
X05_AT0047	Adult	Marital Status
X05_AZ1002_OS	Adult	Ever used any other tobacco products – specify
X05_AC1033MC_OS	Adult	Retail location where your cigarettes are purchased most of the time: Somewhere else – specify
X05_AC1049MC	Adult	Brand of cigarettes usually/last smoked – specify
X05_AC1071MC	Adult	Sub-brand of cigarette product usually/last smoked – specify
X05_AC1033RY_OS	Adult	Retail location where your roll-your-own cigarette tobacco is purchased most of the time: Somewhere else – specify
X05_AC1049RY	Adult	Brand of roll-your-own cigarette tobacco usually/last smoked – specify
X05_AC1071RY	Adult	Sub-brand of roll-your-own cigarette tobacco product usually/last smoked – specify
X05_AV9003	Adult	Brand of electronic nicotine product owned – specify
X05_AV1011_OS	Adult	Flavor of electronic nicotine product when first started using: Some other flavor – specify
X05_AV9035_OS	Adult	Electronic nicotine product you have used: Something else – specify
X05_AV1131_OS	Adult	In past 30 days, [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavor used: Some other flavor – specify
X05_AV1033_OS	Adult	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
X05_AV1012_OS	Adult	Flavor of electronic nicotine product brand regularly/last used: Some other flavor – specify
X05_AV1049	Adult	Brand of electronic nicotine products/electronic nicotine cartridges/e-liquid usually/last used – specify
X05_AJ1033TC_OS	Adult	Retail location where your traditional cigars for blunts are purchased most of the time: Somewhere else – specify
X05_AJ1049TC	Adult	Brand of traditional cigars for blunts usually/last smoked – specify
X05_AJ1071TC	Adult	Sub-brand of traditional cigar as a blunt product usually/last smoked – specify
X05_AJ1033CG_OS	Adult	Retail location where your cigarillos for blunts are purchased most of the time: Somewhere else – specify
X05_AJ1049CG	Adult	Brand of cigarillos for blunts usually/last smoked – specify
X05_AJ1071CG	Adult	Sub-brand of cigarillo as a blunt product usually/last smoked – specify
X05_AJ1033FC_OS	Adult	Retail location where your filtered cigars for blunts are purchased most of the time: Somewhere else – specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_AJ1049FC	Adult	Brand of filtered cigars for blunts usually/last smoked – specify
X05_AJ1071FC	Adult	Sub-brand of filtered cigar as a blunt product usually/last smoked – specify
X05_AG1033TC_OS	Adult	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
X05_AG1049TC	Adult	Brand of traditional cigars usually/last smoked – specify
X05_AG1071TC	Adult	Sub-brand of traditional cigar product usually/last smoked – specify
X05_AG1033CG_OS	Adult	Retail location where your cigarillos are purchased most of the time: Somewhere else – specify
X05_AG1049CG	Adult	Brand of cigarillos usually/last smoked – specify
X05_AG1071CG	Adult	Sub-brand of cigarillo product usually/last smoked – specify
X05_AG1033FC_OS	Adult	Retail location where your filtered cigars are purchased most of the time: Somewhere else – specify
X05_AG1049FC	Adult	Brand of filtered cigars usually/last smoked – specify
X05_AG1071FC	Adult	Sub-brand of filtered cigar product usually/last smoked – specify
X05_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time: Somewhere else – specify
X05_AP1049	Adult	Brand of pipe tobacco usually/last smoked – specify
X05_AP1071	Adult	Sub-brand of pipe tobacco product usually/last smoked – specify
X05_AH9011_OS	Adult	Place where usually smoke/smoked hookah: Somewhere else – specify
X05_AH1033_OS	Adult	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify
X05_AH1049	Adult	Brand of shisha or hookah tobacco usually/last smoked – specify
X05_AH1071	Adult	Sub-brand of hookah tobacco product usually/last smoked – specify
X05_AU1033_OS	Adult	Retail location where your snus is purchased most of the time: Somewhere else – specify
X05_AU1049	Adult	Brand of snus usually/last used – specify
X05_AU1071	Adult	Sub-brand of snus product usually/last used – specify
X05_AS1033_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else – specify
X05_AS1049	Adult	Brand of smokeless tobacco usually/last used – specify
X05_AS1071	Adult	Sub-brand of smokeless tobacco product usually/last used – specify
X05_AX0190_NB	Adult	Is deaf or has serious difficulty hearing
X05_AX0191_NB	Adult	Is blind or has serious difficulty seeing, even when wearing glasses
X05_AX0203_OS	Adult	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify
X05_AX0677_OS	Adult	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
X05_AM0067	Adult	Enrolled in High School
X05_AM0068	Adult	Current grade in school
X05_AM0020_OS	Adult	Type of degree program currently enrolled in: Other type of degree program – specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_AM0012_12M	Adult	In past 12 months, been enrolled in VA Health Care
X05_AM0012_NB	Adult	Ever been enrolled in VA Health Care
X05_AM0021	Adult	Sexual attraction to gender
X05_AM0061	Adult	Consider yourself to be transgender
X05_AM0062	Adult	Consider self to be male-to-female, female-to-male, or non-conforming
X05_AX0064_01	Adult	Work setting: Indoors
X05_AX0064_02	Adult	Work setting: Outdoors
X05_AX0064_03	Adult	Work setting: In a vehicle
X05_AX0135_12M	Adult	What was the outcome of your last pregnancy?
X05_AX0136	Adult	Are you pregnant now?
X05_LCAD01	Adult	Language in which CAPI portions of adult interview were conducted
X05_AX0214	Adult	OK to scan tobacco products
X05_AX0218	Adult	Number of tobacco products given by respondent to scan
X05_AX0217_1A	Adult	Wave 5.5 Adult tobacco Product 1, barcode scan 1
X05_AX0233_01	Adult	First bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_2A	Adult	Wave 5.5 Adult tobacco Product 2, barcode scan 1
X05_AX0233_02	Adult	Second bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_3A	Adult	Wave 5.5 Adult tobacco Product 3, barcode scan 1
X05_AX0233_03	Adult	Third bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_4A	Adult	Wave 5.5 Adult tobacco Product 4, barcode scan 1
X05_AX0233_04	Adult	Fourth bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_5A	Adult	Wave 5.5 Adult tobacco Product 5, barcode scan 1
X05_AX0233_05	Adult	Fifth bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_6A	Adult	Wave 5.5 Adult tobacco Product 6, barcode scan 1
X05_AX0233_06	Adult	Sixth bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_7A	Adult	Wave 5.5 Adult tobacco Product 7, barcode scan 1
X05_AX0233_07	Adult	Seventh bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_8A	Adult	Wave 5.5 Adult tobacco Product 8, barcode scan 1
X05_AX0233_08	Adult	Eighth bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_9A	Adult	Wave 5.5 Adult tobacco Product 9, barcode scan 1
X05_AX0233_09	Adult	Ninth bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AX0217_10A	Adult	Wave 5.5 Adult tobacco Product 10, barcode scan 1
X05_AX0233_10	Adult	Tenth bar code scan: Last time used this type of tobacco it was from this specific package or container
X05_AG1011CG_OS	Adult	Flavor of cigarillo when first started smoking: Some other flavor – specify
X05_AG1011FC_OS	Adult	Flavor of filtered cigar when first started smoking: Some other flavor – specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_AG1011TC_OS	Adult	Flavor of traditional cigar when first started smoking: Some other flavor – specify
X05_AG1012CG_OS	Adult	Cigarillo flavor usually/last smoked: Some other flavor – specify
X05_AG1012FC_OS	Adult	Filtered cigar flavor usually/last smoked: Some other flavor – specify
X05_AG1012TC_OS	Adult	Traditional cigar flavor usually/last smoked: Some other flavor – specify
X05_AG1131CG_OS	Adult	In past 30 days, cigarillo flavor smoked: Some other flavor – specify
X05_AG1131FC_OS	Adult	In past 30 days, filtered cigar flavor smoked: Some other flavor – specify
X05_AG1131TC_OS	Adult	In past 30 days, traditional cigar flavor smoked: Some other flavor – specify
X05_AH1011_OS	Adult	Flavor of hookah tobacco when first started smoking: Some other flavor – specify
X05_AH1012_OS	Adult	Hookah tobacco flavor usually/last smoked: Some other flavor – specify
X05_AH1131_OS	Adult	In past 30 days, hookah tobacco flavor smoked: Some other flavor – specify
X05_AJ1011CG_OS	Adult	Flavor of cigarillo used as blunts when first started smoking: Some other flavor – specify
X05_AJ1011FC_OS	Adult	Flavor of filtered cigar used as blunts when first started smoking: Some other flavor – specify
X05_AJ1011TC_OS	Adult	Flavor of traditional cigar used as blunts when first started smoking: Some other flavor – specify
X05_AJ1012CG_OS	Adult	Cigarillo as blunts flavor usually/last smoked: Some other flavor – specify
X05_AJ1012FC_OS	Adult	Filtered cigar as blunts flavor usually/last smoked: Some other flavor – specify
X05_AJ1012TC_OS	Adult	Traditional cigar as blunts flavor usually/last smoked: Some other flavor – specify
X05_AJ1131CG_OS	Adult	In past 30 days, cigarillo as blunts flavor smoked: Some other flavor – specify
X05_AJ1131FC_OS	Adult	In past 30 days, filtered cigar as blunts flavor smoked: Some other flavor – specify
X05_AJ1131TC_OS	Adult	In past 30 days, traditional cigar as blunts flavor smoked: Some other flavor – specify
X05_AN0336_OS	Adult	Main reason you used a nicotine patch, gum, inhaler, nasal spray, lozenge or pill: Some other reason – specify (current established, current experimental or recent former established non-electronic tobacco users)
X05_AP1011_OS	Adult	Flavor of pipe tobacco when first started smoking: Some other flavor – specify
X05_AP1012_OS	Adult	Pipe tobacco flavor usually/last smoked: Some other flavor – specify
X05_AP1131_OS	Adult	In past 30 days, pipe tobacco flavor smoked: Some other flavor – specify
X05_AS1011_OS	Adult	Flavor of smokeless tobacco when first started using: Some other flavor – specify
X05_AS1012_OS	Adult	Smokeless tobacco flavor usually/last used: Some other flavor – specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_AS1131_OS	Adult	In past 30 days, smokeless tobacco flavor used: Some other flavor – specify
X05_AU1011_OS	Adult	Flavor of snus when first started using: Some other flavor – specify
X05_AU1012_OS	Adult	Snus flavor usually/last used: Some other flavor – specify
X05_AU1131_OS	Adult	In past 30 days, snus flavor used: Some other flavor – specify
X05_AV1133_OS	Adult	In past 30 days, flavor of electronic nicotine product/electronic nicotine cartridge/e-liquid used most often: Some other flavor – specify
X05_AV9001_OS	Adult	Type of electronic nicotine product used most often: Something else – specify
X05_AU1133_OS	Adult	In past 30 days, flavor of snus used most often: Some other flavor – specify
X05_AV0335_OS	Adult	Main reason you stopped using electronic nicotine products: Some other reason – specify
X05_AM0069	Adult	Adult respondent is a citizen of the United States
X05_AG1133CG_OS	Adult	In past 30 days, flavor of cigarillo smoked most often: Some other flavor – specify
X05_AG1133FC_OS	Adult	In past 30 days, flavor of filtered cigar smoked most often: Some other flavor – specify
X05_AG1133TC_OS	Adult	In past 30 days, flavor of traditional cigar smoked most often: Some other flavor – specify
X05_AH1133_OS	Adult	In past 30 days, flavor of hookah tobacco smoked most often: Some other flavor – specify
X05_AH9045_OS	Adult	Location when first tried a hookah: Somewhere else – specify
X05_AJ1133CG_OS	Adult	In past 30 days, flavor of cigarillos smoked as blunts most often: Some other flavor – specify
X05_AJ1133FC_OS	Adult	In past 30 days, flavor of filtered cigars smoked as blunts most often: Some other flavor – specify
X05_AJ1133TC_OS	Adult	In past 30 days, flavor of traditional cigars smoked as blunts most often: Some other flavor – specify
X05_AN0349_OS	Adult	Strength of nicotine patch last used (current or recent former established non-electronic tobacco users): Some other strength – specify
X05_AN0349E_OS	Adult	Strength of nicotine patch last used (current or recent former established non-marijuana electronic nicotine product users): Some other strength – specify
X05_AN0349H_OS	Adult	Strength of nicotine patch last used (former experimental or long-term former established tobacco users): Some other strength – specify
X05_AP1133_OS	Adult	In past 30 days, flavor of pipe tobacco smoked most often: Some other flavor – specify
X05_AS1133_OS	Adult	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor – specify
X05_AX0601_OS	Adult	In past 30 days, saw anyone using e-cigarettes or electronic nicotine products in or around your school: Yes, somewhere else at school not listed here – specify
X05_AX0602_OS	Adult	In past 30 days, saw anyone smoking a cigarette in or around your school: Yes, somewhere else at school not listed here – specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_CPT05	Parent	Confirm DOB for shadow youth who is expected to be a youth at Wave 5.5
X05_CPT05C	Parent	Confirm DOB for youth from a prior wave who is expected to be a youth at Wave 5.5
X05_CPT07	Parent	Date of birth (corrected values for continuing youth/new values for new baseline youth)
X05_E_YOUTH	Youth	Emancipated youth
X05_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
X05_LCYS01	Youth	Language in which CAPI portions of youth interview were conducted
X05LEY01	Youth	Emancipated Youth's preferred language to complete ACASI interview
X05_LYH01	Youth	Youth's preferred language to complete ACASI interview
X05_PARENT_PERSONID	Parent	Wave 5.5 Parent/guardian Participant ID Number
X05_PM0016_NB	Parent	Youth was identified as a twin or part of a multiple birth
X05_PM0017_NB	Parent	Youth and sibling are identical twins
X05_PM0019_NB	Parent	Any sibling in multiple birth identical to youth
X05_PM0021_NB	Parent	First name of sibling that youth is a twin of
X05_PM0030_NB	Parent	First names of siblings in multiple birth that are identical to youth
X05_PM0035_NB	Parent	First names of siblings that are in the multiple birth with youth
X05_PM0052	Parent	Confirm parent's spouse is same as reported in previous wave(s)
X05_PM0053	Parent	Confirm parent's relationship to youth
X05_PM0058_NEW_AGE	Parent	Age of other parental figure/guardian
X05_PM0058_NEW_FNAME	Parent	First name of other parental figure or guardian - specify
X05_PM0058_NEW_LNAME	Parent	Last name of other parental figure or guardian - specify
X05_PM0058_PERSONID	Parent	Other parental figure/guardian PERSONID
X05_PM0059_OS_1	Parent	Other parental figure/guardian's relationship to sampled youth: Other relative - specify
X05_PM0059_OS_2	Parent	Other parental figure/guardian's relationship to sampled youth: Other non-relative - specify
X05_PM0061_NEW_AGE	Parent	Age of second other parental figure/guardian
X05_PM0061_NEW_FNAME	Parent	First name of second other parental figure or guardian - specify
X05_PM0061_NEW_LNAME	Parent	Last name of second other parental figure or guardian - specify
X05_PM0061_PERSONID	Parent	Second other parental figure/guardian PERSONID
X05_PM0062_OS_1	Parent	Second other parental figure/guardian's relationship to youth: Other relative - specify
X05_PM0062_OS_2	Parent	Second other parental figure/guardian's relationship to youth: Other non-relative - specify
X05_PM0067	Parent	Parent respondent is currently enrolled in high school
X05_PM0068	Parent	Parent respondent grade (if currently enrolled in high school)
X05_PM0072_OS	Parent	Other language spoken at home: Some other language - specify (parent respondent)
X05_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative - specify
X05_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative - specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative - specify
X05_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative - specify
X05_PT0046_NEW_AGE	Parent	Age of new spouse/partner
X05_PT0046_PERSONID	Parent	Parent's spouse/partner's PERSONID (if different from previous waves)
X05_PT0046_NEW_FNAME	Parent	First name of spouse/partner that lives with parent in DU - specify
X05_PT0046_NEW_LNAME	Parent	Last name of spouse/partner that lives with parent in DU - specify
X05_PX0186_NB	Parent	Youth has serious difficulty walking or climbing stairs
X05_PX0190_NB	Parent	Youth is deaf or has serious difficulty hearing
X05_PX0191_NB	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
X05_YC1033_OS	Youth	Retail location where your cigarettes are purchased most of the time: Somewhere else - specify
X05_YC1049	Youth	Brand of cigarettes usually/last smoked - specify
X05_YC1071	Youth	Sub-brand of cigarette product usually/last smoked - specify
X05_YC1118_OS	Youth	In past 30 days, how you usually got your own cigarettes - specify
X05_YG1011CL_OS	Youth	Flavor of first cigarillo smoked: Some other flavor - specify
X05_YG1011FC_OS	Youth	Flavor of first filtered cigar smoked: Some other flavor - specify
X05_YG1011TC_OS	Youth	Flavor of first traditional cigar smoked: Some other flavor - specify
X05_YG1033CL_OS	Youth	Retail location where your cigarillos are purchased most of the time: Somewhere else - specify
X05_YG1033FC_OS	Youth	Retail location where your filtered cigars are purchased most of the time: Somewhere else - specify
X05_YG1033TC_OS	Youth	Retail location where your traditional cigars are purchased most of the time: Somewhere else - specify
X05_YG1049CL	Youth	Brand of cigarillos usually/last smoked - specify
X05_YG1049FC	Youth	Brand of filtered cigars usually/last smoked - specify
X05_YG1049TC	Youth	Brand of traditional cigars usually/last smoked - specify
X05_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked - specify
X05_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked - specify
X05_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked - specify
X05_YG1118CL_OS	Youth	In past 30 days, how you usually got your own cigarillos - specify
X05_YG1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars - specify
X05_YG1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars - specify
X05_YG1131CL_OS	Youth	In past 30 days, smoked cigarillos flavored to taste like: Some other flavor - specify
X05_YG1131FC_OS	Youth	In past 30 days, smoked filtered cigars flavored to taste like: Some other flavor - specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_YG1131TC_OS	Youth	In past 30 days, smoked traditional cigars flavored to taste like: Some other flavor – specify
X05_YG1133CL_OS	Youth	In past 30 days, flavor of cigarillos smoked most often: Some other flavor – specify
X05_YG1133FC_OS	Youth	In past 30 days, flavor of filtered cigars smoked most often: Some other flavor – specify
X05_YG1133TC_OS	Youth	In past 30 days, flavor of traditional cigars smoked most often: Some other flavor – specify
X05_YH1011_OS	Youth	Flavor of first shisha or hookah tobacco smoked: Some other flavor – specify
X05_YH1033_OS	Youth	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify
X05_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
X05_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
X05_YH1073	Youth	Smoke hookah because: It is part of my cultural tradition
X05_YH1118_OS	Youth	In past 30 days, how you usually got your own shisha or hookah tobacco – specify
X05_YH1131_OS	Youth	In past 30 days, smoked hookah tobacco flavored to taste like: Some other flavor – specify
X05_YH1133_OS	Youth	In past 30 days, flavor of hookah smoked most often: Some other flavor – specify
X05_YH9011_OS	Youth	Usually smoke a hookah: Somewhere else – specify
X05_YH9045_OS	Youth	Where you were when you first tried a hookah, even one or two puffs – specify
X05_YJ1011CG_OS	Youth	Flavor of first cigarillo smoked as a blunt: Some other flavor – specify
X05_YJ1011FC_OS	Youth	Flavor of first filtered cigar smoked as a blunt: Some other flavor – specify
X05_YJ1011TC_OS	Youth	Flavor of first traditional cigar smoked as a blunt: Some other flavor – specify
X05_YJ1033CG_OS	Youth	Retail location where your cigarillos for blunts are purchased most of the time: Somewhere else – specify
X05_YJ1033FC_OS	Youth	Retail location where your filtered cigars for blunts are purchased most of the time: Somewhere else – specify
X05_YJ1033TC_OS	Youth	Retail location where your traditional cigars for blunts are purchased most of the time: Somewhere else – specify
X05_YJ1049CG	Youth	Brand of cigarillos usually/last smoked as a blunt – specify
X05_YJ1049FC	Youth	Brand of filtered cigars usually/last smoked as a blunt – specify
X05_YJ1049TC	Youth	Brand of traditional cigars usually/last smoked as a blunt – specify
X05_YJ1071CG	Youth	Sub-brand of cigarillo product usually/last smoked as a blunt – specify
X05_YJ1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked as a blunt – specify
X05_YJ1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked as a blunt – specify
X05_YJ1118CG_OS	Youth	In past 30 days, how you usually got your own cigarillo as a blunt – specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_YJ1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars as a blunt – specify
X05_YJ1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars as a blunt – specify
X05_YJ1131CG_OS	Youth	In past 30 days, smoked cigarillos as blunts flavored to taste like: Some other flavor – specify
X05_YJ1131FC_OS	Youth	In past 30 days, smoked filtered cigars as blunts flavored to taste like: Some other flavor – specify
X05_YJ1131TC_OS	Youth	In past 30 days, smoked traditional cigars as blunts flavored to taste like: Some other flavor – specify
X05_YJ1133CG_OS	Youth	In past 30 days, flavor of cigarillos smoked as blunts most often: Some other flavor – specify
X05_YJ1133FC_OS	Youth	In past 30 days, flavor of filtered cigars smoked as blunts most often: Some other flavor – specify
X05_YJ1133TC_OS	Youth	In past 30 days, flavor of traditional cigars smoked as blunts most often: Some other flavor – specify
X05_YM0007	Youth	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
X05_YM0011_01	Youth	Branch served when on active duty: Army
X05_YM0011_02	Youth	Branch served when on active duty: Navy
X05_YM0011_03	Youth	Branch served when on active duty: Air Force
X05_YM0011_04	Youth	Branch served when on active duty: Marine Corps
X05_YM0011_05	Youth	Branch served when on active duty: Coast Guard
X05_YM0020	Youth	Last grade/year in school completed
X05_YM0021	Youth	Sexual attraction to gender
X05_YM0061	Youth	Consider yourself to be transgender
X05_YM0062	Youth	Transgender category
X05_YM0063	Youth	Sexual orientation
X05_YM0066	Youth	How well you speak English (youth)
X05_YM0069	Youth	Youth respondent is a citizen of the United States
X05_YM0070	Youth	How well you read English (youth)
X05_YM0072_OS	Youth	Other language spoken at home: Some other language – specify (youth respondent)
X05_YM0073	Youth	How well you write in English (youth)
X05_YN0129_OS	Youth	In past 12 months, tried to completely stop using: Other – specify
X05_YN0336_OS	Youth	Main reason you used a nicotine patch, nicotine gum, nicotine inhaler, nicotine nasal spray or lozenge (past year tobacco users): Some other reason – specify
X05_YN0349_OS	Youth	Strength of nicotine patch last used (past year tobacco users): Some other strength – specify
X05_YN0349H_OS	Youth	Strength of nicotine patch last used (past year tobacco non-users): Some other strength – specify
X05_YP1033_OS	Youth	Retail location where your pipe tobacco is purchased most of the time: Somewhere else – specify
X05_YP1118_OS	Youth	In past 30 days, how you usually got your own pipe tobacco – specify
X05_YS1011_OS	Youth	Flavor of first smokeless tobacco product used: Some other flavor – specify
X05_YS1033_OS	Youth	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else – specify
X05_YS1049	Youth	Brand of smokeless tobacco usually/last used – specify

**Table E-13. Wave 5.5 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
X05_YS1071	Youth	Sub-brand of smokeless tobacco product usually/last used – specify
X05_YS1118_OS	Youth	In past 30 days, how you usually got your own smokeless tobacco – specify
X05_YS1131_OS	Youth	In past 30 days, used smokeless tobacco flavored to taste like: Some other flavor – specify
X05_YS1133_OS	Youth	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor – specify
X05_YU1011_OS	Youth	Flavor of first snus used: Some other flavor – specify
X05_YU1033_OS	Youth	Retail location where your snus is purchased most of the time: Somewhere else – specify
X05_YU1049	Youth	Brand of snus usually/last used – specify
X05_YU1071	Youth	Sub-brand of snus usually/last used – specify
X05_YU1118_OS	Youth	In past 30 days, how you usually got your own snus – specify
X05_YU1131_OS	Youth	In past 30 days, used snus flavored to taste like: Some other flavor – specify
X05_YU1133_OS	Youth	In past 30 days, flavor of snus used most often: Some other flavor – specify
X05_YV1011_OS	Youth	Flavor of first electronic nicotine product used: Some other flavor – specify
X05_YV1033_OS	Youth	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
X05_YV1049	Youth	Brand of electronic nicotine products/electronic nicotine cartridges/e-liquid usually/last used – specify
X05_YV1118_OS	Youth	In past 30 days, how you usually got your own [electronic nicotine products/electronic nicotine cartridges/e-liquid] – specify
X05_YV1131_OS	Youth	In past 30 days, used [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavored to taste like: Some other flavor – specify
X05_YV1133_OS	Youth	In past 30 days, flavor of electronic nicotine product/cartridges/e-liquid used most often: Some other flavor – specify
X05_YV9001_OS	Youth	Type of electronic nicotine product used most often – specify
X05_YV9035_OS	Youth	Ever used some other electronic nicotine product – specify
X05_YX0136	Youth	Currently pregnant
X05_YX0137_NN	Youth	Number of weeks/months pregnant – Number
X05_YX0137_UN	Youth	Number of weeks/months pregnant – Unit
X05_YX0203_OS	Youth	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify
X05_YX0677_OS	Youth	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
X05_YY0601_OS	Youth	First type of tobacco you tried – specify
X05_YZ1002_OS	Youth	Ever used any other tobacco products – specify
X05_YX0601_OS	Youth	In past 30 days, saw anyone using e-cigarettes or electronic nicotine products in or around your school: Yes, somewhere else at school not listed here – specify
X05_YX0602_OS	Youth	In past 30 days, saw anyone smoking a cigarette in or around your school: Yes, somewhere else at school not listed here – specify

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_AM0001	Adult	X05R_A_AGECAT6	Wave 5.5 Adult Age (6 levels)
X05_AM0002			
X05_AM0003			
X05_AM0065	Adult	X05R_A_AM0065_V2	Recoded number of years in US (2 levels)
X05_AC1007_NB	Adult	X05R_A_AC1007_NB	Age group when you first started smoking cigarettes fairly regularly (6 levels)
X05_AC1121_NB			
X05_AC1020_NB	Adult	X05R_A_AC1020_NB	Age group when you first started smoking cigarettes every day (6 levels)
X05_AC1122_NB			
X05_AV1007_NB	Adult	X05R_A_AV1007_NB	Age group when you first started using electronic nicotine products fairly regularly (6 levels)
X05_AV1121_NB			
X05_AV1020_NB	Adult	X05R_A_AV1020_NB	Age group when you first started using electronic nicotine products every day (6 levels)
X05_AV1122_NB			
X05_AG1007TC_NB	Adult	X05R_A_AG1007TC_NB	Age group when you first started smoking traditional cigars fairly regularly (6 levels)
X05_AG1121TC_NB			
X05_AG1020TC_NB	Adult	X05R_A_AG1020TC_NB	Age group when you first started smoking traditional cigars every day (6 levels)
X05_AG1122TC_NB			
X05_AG1007CG_NB	Adult	X05R_A_AG1007CG_NB	Age group when you first started smoking cigarillos fairly regularly (6 levels)
X05_AG1121CG_NB			
X05_AG1020CG_NB	Adult	X05R_A_AG1020CG_NB	Age group when you first started smoking cigarillos every day (6 levels)
X05_AG1122CG_NB			
X05_AG1007FC_NB	Adult	X05R_A_AG1007FC_NB	Age group when you first started smoking filtered cigars fairly regularly (6 levels)
X05_AG1121FC_NB			
X05_AG1020FC_NB	Adult	X05R_A_AG1020FC_NB	Age group when you first started smoking filtered cigars every day (6 levels)
X05_AG1122FC_NB			
X05_AP1007_NB	Adult	X05R_A_AP1007_NB	Age group when you first started smoking a pipe filled with tobacco fairly regularly (6 levels)
X05_AP1121_NB			
X05_AP1020_NB	Adult	X05R_A_AP1020_NB	Age group when you first started smoking a pipe filled with tobacco every day (6 levels)
X05_AP1122_NB			
X05_AH1007_NB	Adult	X05R_A_AH1007_NB	Age group when you first started smoking hookah fairly regularly (6 levels)
X05_AH1121_NB			
X05_AH1020_NB	Adult	X05R_A_AH1020_NB	Age group when you first started smoking hookah every day (6 levels)
X05_AH1122_NB			
X05_AU1007_NB	Adult	X05R_A_AU1007_NB	Age group when you first started using snus pouches fairly regularly (6 levels)
X05_AU1121_NB			

**Table E-14.** Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_AU1020_NB X05_AU1122_NB	Adult	X05R_A_AU1020_NB	Age group when you first started using snus pouches every day (6 levels)
X05_AS1007_NB X05_AS1121_NB	Adult	X05R_A_AS1007_NB	Age group when you first started using smokeless tobacco fairly regularly (6 levels)
X05_AS1020_NB X05_AS1122_NB	Adult	X05R_A_AS1020_NB	Age group when you first started using smokeless tobacco every day (6 levels)
X05_AX0114_NB X05_AX0253_NB	Adult	X05R_A_AX0114_NB	Age range you were in when you were first told you had high blood pressure (6 levels)
X05_AX0115_NB X05_AX0254_NB	Adult	X05R_A_AX0115_NB	Age range you were in when you were first told you had high cholesterol (6 levels)
X05_AX0116_NB X05_AX0255_NB	Adult	X05R_A_AX0116_NB	Age range you were in when you were first told you had congestive heart failure (6 levels)
X05_AX0117_NB X05_AX0256_NB	Adult	X05R_A_AX0117_NB	Age range you were in when you were first told you had a stroke (6 levels)
X05_AX0112_NB X05_AX0252_NB	Adult	X05R_A_AX0112_NB	Age range you were in when you were first told you had a heart attack (6 levels)
X05_AX0120_NB X05_AX0257_NB	Adult	X05R_A_AX0120_NB	Age range you were in when you were first told you had COPD (6 levels)
X05_AX0121_NB X05_AX0258_NB	Adult	X05R_A_AX0121_NB	Age range you were in when you were first told you had chronic bronchitis (6 levels)
X05_AX0123_NB X05_AX0259_NB	Adult	X05R_A_AX0123_NB	Age range you were in when you were first told you had emphysema (6 levels)
X05_AX0124_NB X05_AX0260_NB	Adult	X05R_A_AX0124_NB	Age range you were in when you were first told you had asthma (6 levels)
X05_AX0131_NB X05_AX0261_NB	Adult	X05R_A_AX0131_NB	Age range you were in when you were first told you had gum disease (6 levels)
X05_AX0133_NB X05_AX0262_NB	Adult	X05R_A_AX0133_NB	Age range you were in when you were first told you had pre-cancerous oral lesions (6 levels)
X05_AX0280_NB X05_AX0263_NB	Adult	X05R_A_AX0280_NB	Age range you were in when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes (6 levels)
X05_AX0143_NB X05_AX0264_NB	Adult	X05R_A_AX0143_NB	Age range you were in when you were first told you had an ulcer (6 levels)
X05_AX0148_NB X05_AX0266_NB	Adult	X05R_A_AX0148_NB	Age range you were in when you were first told you had stomach or gastro-intestinal bleeding (6 levels)

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_AX0150_NB X05_AX0267_NB	Adult	X05R_A_AX0150_NB	Age range you were in when you were first told you had osteoporosis (6 levels)
X05_AX0198_NB X05_AX0268_NB	Adult	X05R_A_AX0198_NB	Age range you were in when you were first told you had a bone fracture because you have fragile bones (6 levels)
X05_AX0152_NB X05_AX0269_NB	Adult	X05R_A_AX0152_NB	Age range you were in when you were first told you had cataract or glaucoma (6 levels)
X05_AX0145_02 X05_AX0145_03 X05_AX0145_04 X05_AX0145_05 X05_AX0145_09 X05_AX0145_12 X05_AX0145_15 X05_AX0145_18 X05_AX0145_19 X05_AX0145_21 X05_AX0145_16 X05_AX0145_23 X05_AX0145_24 X05_AX0145_25 X05_AX0145_27 X05_AX0145_29 X05_AX0145_30 X05_AX0145_31	Adult	X05R_A_AX0145_NONTOB	Type of cancer is a non-tobacco related cancer (2 levels)
X05_AX0145_01 X05_AX0145_06 X05_AX0145_07 X05_AX0145_08 X05_AX0145_10 X05_AX0145_11 X05_AX0145_13 X05_AX0145_14 X05_AX0145_17 X05_AX0145_20 X05_AX0145_22 X05_AX0145_26 X05_AX0145_28	Adult	X05R_A_AX0145_TOB	Type of cancer is a tobacco related cancer (2 levels)

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_AX0146_NB_02	Adult	X05R_A_AX0146_NONTOB	Age range when first non tobacco-related cancer was diagnosed (6 levels)
X05_AX0265_NB_02			
X05_AX0146_NB_03			
X05_AX0265_NB_03			
X05_AX0146_NB_04			
X05_AX0265_NB_04			
X05_AX0146_NB_05			
X05_AX0265_NB_05			
X05_AX0146_NB_09			
X05_AX0265_NB_09			
X05_AX0146_NB_12			
X05_AX0265_NB_12			
X05_AX0146_NB_15			
X05_AX0265_NB_15			
X05_AX0146_NB_18			
X05_AX0265_NB_18			
X05_AX0146_NB_19			
X05_AX0265_NB_19			
X05_AX0146_NB_21			
X05_AX0265_NB_21			
X05_AX0146_NB_16			
X05_AX0265_NB_16			
X05_AX0146_NB_23			
X05_AX0265_NB_23			
X05_AX0146_NB_24			
X05_AX0265_NB_24			
X05_AX0146_NB_25			
X05_AX0265_NB_25			
X05_AX0146_NB_27			
X05_AX0265_NB_27			
X05_AX0146_NB_29			
X05_AX0265_NB_29			
X05_AX0146_NB_30			
X05_AX0265_NB_30			
X05_AX0146_NB_31			
X05_AX0265_NB_31			

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_AX0146_NB_01 X05_AX0265_NB_01 X05_AX0146_NB_06 X05_AX0265_NB_06 X05_AX0146_NB_07 X05_AX0265_NB_07 X05_AX0146_NB_08 X05_AX0265_NB_08 X05_AX0146_NB_10 X05_AX0265_NB_10 X05_AX0146_NB_11 X05_AX0265_NB_11 X05_AX0146_NB_13 X05_AX0265_NB_13 X05_AX0146_NB_14 X05_AX0265_NB_14 X05_AX0146_NB_17 X05_AX0265_NB_17 X05_AX0146_NB_20 X05_AX0265_NB_20 X05_AX0146_NB_22 X05_AX0265_NB_22 X05_AX0146_NB_26 X05_AX0265_NB_26 X05_AX0146_NB_28 X05_AX0265_NB_28	Adult	X05R_A_AX0146_TOB	Age range when first tobacco-related cancer was diagnosed (6 levels)
X05_AX0300_12M_01 X05_AX0300_12M_02 X05_AX0300_12M_03 X05_AX0300_12M_04 X05_AX0300_12M_05 X05_AX0300_12M_06 X05_AX0300_12M_07 X05_AX0706	Adult	X05R_A_AX0300_12M	Indicator for adverse pregnancy outcomes resulting in risky birth (2 levels)
X05_AM0018	Adult	X05R_A_AM0018_X1	Recoded education level (2 levels)
X05_AM0030	Adult	X05R_A_AM0030	Recoded total household income in the past 12 months (5 levels)
X05_AM0033	Adult	X05R_A_AM0033_V2	Recoded education level of mother (5 levels)
X05_AM0034	Adult	X05R_A_AM0034_V2	Recoded education level of father (5 levels)
X05_AM0036	Adult	X05R_A_AM0036	Recoded parents' total household income in the past 12 months (5 levels)
X05_AM0026_01 X05_AM0026_02 X05_AM0026_03 X05_AM0026_04 X05_AM0026_05 X05_AM0026_06 X05_AM0026_07 X05_AM0026_08	Adult	X05R_A_AM0026_V2	Indicator of health insurance (2 levels)
X05_AM0063	Adult	X05R_A_SEXORIENT2	Recoded sexual orientation (2 levels)

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_AX0313 X05_AX0679_FT X05_AX0679_IN X05_AX0316 X05_AX0109 X05_AX0312	Adult	X05R_A_BMI	Body mass index
X05_AL0040	Adult	X05R_A_AL0040	Indicator of home ownership (2 levels)
X05_AM0041_02 X05_AM0041_03 X05_AM0041_04 X05_AM0041_05	Adult	X05R_A_AM0041_X1	Who you currently live with (2 levels)
X05_AM0042	Adult	X05R_A_AM0042_X1	Recoded where you currently live (2 levels)
X05_AM0072_02 X05_AM0072_03 X05_AM0072_04 X05_AM0072_05 X05_AM0072_06 X05_AM0072_07 X05_AM0072_08 X05_AM0072_09 X05_AM0072_10	Adult	X05R_A_AM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (adult)
X05_AX0066_01 X05_AX0066_02 X05_AX0066_03 X05_AX0066_04 X05_AX0066_05 X05_AX0066_06 X05_AX0066_07 X05_AX0066_08 X05_AX0066_09	Adult	X05R_A_AX0066	Recoded anyone who lives with you now who uses tobacco (4 levels)
X05_AX0093	Adult	X05R_A_AX0093	Extent to which you are able to carry out your everyday physical activities (2 levels)
X05_AX0757_NB X05_AX0758_NB	Adult	X05R_A_AX0757_NB	Age range when you were first told you had schizophrenia, schizoaffective disorder or psychosis (6 levels)
X05_AX0762_NB X05_AX0772_NB	Adult	X05R_A_AX0762_NB	Age range when you were first told you had a psychotic illness or episode (6 levels)
X05_PT0001	Parent	X05R_Y_PT0001_V2	Recoded parent or guardian relationship to youth (3 levels)
X05_PT0047	Parent	X05R_Y_PT0047	Recoded parent or guardian marital status (3 levels)
X05_PT0045 X05_PM0057 X05_PM0060	Parent	X05R_P_OTHPAR_INHH	Wave 5.5 Recoded youth has any parental figures/guardians other than the parent/guardian in the house (2 levels)
X05_PT0002	Parent	X05R_Y_PT0002_V2	Recoded parent's spouse or partner relationship to youth (3 levels)

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_PM0059	Parent	X05R_Y_PM0059_V2	Recoded other parental figure/guardian's relationship to youth (3 levels)
X05_PM0062	Parent	X05R_Y_PM0062_V2	Recoded second other parental figure/guardian's relationship to youth (3 levels)
X05_PT0041_NB X05_PT0253_NB	Parent	X05R_Y_PT0041_NB	Age range youth was first told he/she had high blood pressure (3 levels)
X05_PT0043_NB X05_PT0254_NB	Parent	X05R_Y_PT0043_NB	Age range youth was first told he/she has high cholesterol (3 levels)
X05_PT0038_NB X05_PT0260_NB	Parent	X05R_Y_PT0038_NB	Age range youth was first told he/she has asthma (3 levels)
X05_PT0042_NB X05_PT0263_NB	Parent	X05R_Y_PT0042_NB	Age range youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (3 levels)
X05_PM0069	Parent	X05R_Y_PM0069	Recoded parent respondent is a citizen of the United States (2 levels)
X05_PM0065 X05_PM0065_NN	Parent	X05R_Y_PM0065_V2	Recoded number of years parents lived in the United States (2 levels)
X05_PM0001 X05_PM0118	Parent	X05R_P_PARSP_EDUC	Recoded highest grade or year of school completed by parent/spouse/guardian (6 levels)
X05_PM0130	Parent	X05R_Y_PM0130	Recoded total household income in past 12 months (5 levels)
X05_PL0040	Parent	X05R_Y_PL0040	Recoded home is owned or rented (2 levels)
X05_PM0072_02 X05_PM0072_03 X05_PM0072_04 X05_PM0072_05 X05_PM0072_06 X05_PM0072_07 X05_PM0072_08 X05_PM0072_09 X05_PM0072_10	Parent	X05R_Y_PM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (parent respondent)
X05_PX0757_NB X05_PX0758_NB	Parent	X05R_Y_PX0757_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had schizophrenia, schizoaffective disorder or psychosis (3 levels)
X05_PX0762_NB X05_PX0772_NB	Parent	X05R_Y_PX0762_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had a psychotic illness or episode (3 levels)
X05_YM0065	Youth	X05R_Y_YM0065_V2	Recoded number of years youth lived in the United States (2 levels)
X05_YC1006_NB X05_YC1120_NB	Youth	X05R_Y_YC1006_NB	Age range when first tried cigarette smoking, even one or two puffs (3 levels)

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_YC1007 X05_YC1121	Youth	X05R_Y_YC1007	Age range when first started smoking cigarettes fairly regularly (3 levels)
X05_YV1006_NB X05_YV1120_NB	Youth	X05R_Y_YV1006_NB	Age range when first tried [primary electronic nicotine product], even one or two times, even one or two puffs (3 levels)
X05_YV1007 X05_YV1121	Youth	X05R_Y_YV1007	Age range when first started using [primary electronic nicotine product]s fairly regularly (3 levels)
X05_YG1006TC_NB X05_YG1120TC_NB	Youth	X05R_Y_YG1006TC_NB	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
X05_YG1007TC X05_YG1121TC	Youth	X05R_Y_YG1007TC	Age range when first started smoking traditional cigars fairly regularly (3 levels)
X05_YG1006CL_NB X05_YG1120CL_NB	Youth	X05R_Y_YG1006CL_NB	Age range when first tried a cigarillo, even one or two puffs (3 levels)
X05_YG1007CL X05_YG1121CL	Youth	X05R_Y_YG1007CL	Age range when first started smoking cigarillos fairly regularly (3 levels)
X05_YG1006FC_NB X05_YG1120FC_NB	Youth	X05R_Y_YG1006FC_NB	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
X05_YG1007FC X05_YG1121FC	Youth	X05R_Y_YG1007FC	Age range when first started smoking filtered cigars fairly regularly (3 levels)
X05_YH1006_NB X05_YH1120_NB	Youth	X05R_Y_YH1006_NB	Age range when first tried smoking hookah, even one or two puffs (3 levels)
X05_YH1007 X05_YH1121	Youth	X05R_Y_YH1007	Age range when first started smoking a hookah fairly regularly (3 levels)
X05_YU1006_NB X05_YU1120_NB	Youth	X05R_Y_YU1006_NB	Age range when first tried snus pouches, even one or two times (3 levels)
X05_YU1007 X05_YU1121	Youth	X05R_Y_YU1007	Age range when first started using snus pouches fairly regularly (3 levels)
X05_YS1006_NB X05_YS1120_NB	Youth	X05R_Y_YS1006_NB	Age range when first tried smokeless tobacco, even one or two times (3 levels)
X05_YS1007 X05_YS1121	Youth	X05R_Y_YS1007	Age range when first started using smokeless tobacco fairly regularly (3 levels)
X05_YX0671_01 X05_YX0671_02 X05_YX0671_03 X05_YX0671_04 X05_YX0671_05 X05_YX0671_06 X05_YX0671_07 X05_YX0671_08	Youth	X05R_Y_YX0671	Recoded anyone who lives with you now uses tobacco (4 levels)
X05_YT0038_NB X05_YT0260_NB	Youth	X05R_Y_YT0038_NB	Age range when you were first told you had asthma (3 levels)
X05_YX0086_NB X05_YX0087_NB	Youth	X05R_Y_YX0086_NB	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_YX0074_NB X05_YX0270_NB	Youth	X05R_Y_YX0074_NB	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
X05_YX0079_NB X05_YX0271_NB	Youth	X05R_Y_YX0079_NB	Age range when first used marijuana, hash, THC or grass (3 levels)
X05_YX0082_NB_01 X05_YX0272_NB_01	Youth	X05R_Y_YX0082_NB_01	Age range when first used: Ritalin or Adderall (3 levels)
X05_YX0082_NB_02 X05_YX0272_NB_02	Youth	X05R_Y_YX0082_NB_02	Age range when first used: Painkillers (3 levels)
X05_YX0082_NB_03 X05_YX0272_NB_03	Youth	X05R_Y_YX0082_NB_03	Age range when first used: Sedatives or tranquilizers (3 levels)
X05_YX0082_NB_04 X05_YX0272_NB_04	Youth	X05R_Y_YX0082_NB_04	Age range when first used: Cocaine or crack (3 levels)
X05_YX0082_NB_05 X05_YX0272_NB_05	Youth	X05R_Y_YX0082_NB_05	Age range when first used: Methamphetamine or speed (3 levels)
X05_YX0082_NB_06 X05_YX0272_NB_06	Youth	X05R_Y_YX0082_NB_06	Age range when first used: Heroin (3 levels)
X05_YX0082_NB_07 X05_YX0272_NB_07	Youth	X05R_Y_YX0082_NB_07	Age range when first used: Inhalants or solvents (3 levels)
X05_YX0082_NB_08 X05_YX0272_NB_08	Youth	X05R_Y_YX0082_NB_08	Age range when first used: Hallucinogens (3 levels)
X05_YM0019 X05_YM0018	Youth	X05R_Y_YM0018_V2	Recoded grade level (If on holiday or break – grade level entering when returning to school) (6 levels)
X05_YM0004_NB	Youth	X05R_Y_SEX	Gender
X05_YM0005_NB_01 X05_YM0005_NB_02 X05_YM0005_NB_03 X05_YM0005_NB_04 X05_YM0005_NB_05	Youth	X05R_Y_HISP	DERIVED – Hispanic origin from the interview (2 levels)
X05_YM0006_NB_01 X05_YM0006_NB_02 X05_YM0006_NB_03 X05_YM0006_NB_04 X05_YM0006_NB_05 X05_YM0006_NB_06 X05_YM0006_NB_07 X05_YM0006_NB_08 X05_YM0006_NB_09 X05_YM0006_NB_10 X05_YM0006_NB_11 X05_YM0006_NB_12 X05_YM0006_NB_13 X05_YM0006_NB_14	Youth	X05R_Y_RACECAT3	Recoded Race from the interview (3 levels)

**Table E-14. Wave 5.5 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
X05_PT0007_FT X05_PT0007_IN X05_PT0007_MT X05_PT0008_LB X05_PT0008_KG X05_YX0310 X05_YT0007_FT X05_YT0007_IN X05_YX0311 X05_YT0008 X05_YX0312	Youth/Parent	X05R_Y_BMI	Body mass index
X05_YM0072_02 X05_YM0072_03 X05_YM0072_04 X05_YM0072_05 X05_YM0072_06 X05_YM0072_07 X05_YM0072_08 X05_YM0072_09 X05_YM0072_10	Youth	X05R_Y_YM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (youth)

**Table E-15. PATH-ATS questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
T05_AM0027	Adult	Confirm DOB (for continuing respondents)
T05_LCAD01	Adult	PATH-ATS interview language
T05_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
T05_AM0067	Adult	Enrolled in High School
T05_AM0068	Adult	Current grade in school
T05_AC1033_OS	Adult	Retail location where your cigarettes are purchased most of the time: Somewhere else – specify
T05_AC1048_OS	Adult	Brand of cigarettes usually smoked – specify
T05_AC1070_OS	Adult	Sub-brand of cigarette product usually smoked – specify
T05_AV1011_OS	Adult	Flavor of electronic nicotine product when first started using: Some other flavor – specify
T05_AV9001_OS	Adult	Type of electronic nicotine product used most often: Something else – specify
T05_AV1131_OS	Adult	In past 30 days, [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavor used: Some other flavor – specify
T05_AV1133_OS	Adult	In past 30 days, flavor of electronic nicotine product/electronic nicotine cartridge/e-liquid used most often: Some other flavor – specify
T05_AV1033_OS	Adult	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
T05_AV1048_OS	Adult	Brand of [electronic nicotine products/electronic nicotine cartridges/e-liquid] usually used – specify
T05_AG1011TC_OS	Adult	Flavor of traditional cigar when first started smoking: Some other flavor – specify
T05_AG1131TC_OS	Adult	In past 30 days, traditional cigar flavor smoked: Some other flavor – specify
T05_AG1133TC_OS	Adult	In past 30 days, flavor of traditional cigar smoked most often: Some other flavor – specify
T05_AG1033TC_OS	Adult	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
T05_AG1048TC_OS	Adult	Brand of traditional cigars usually smoked – specify
T05_AG1070TC_OS	Adult	Sub-brand of traditional cigar product usually smoked – specify
T05_AG1011CG_OS	Adult	Flavor of cigarillo when first started smoking: Some other flavor – specify
T05_AG1131CG_OS	Adult	In past 30 days, cigarillo flavor smoked: Some other flavor – specify
T05_AG1133CG_OS	Adult	In past 30 days, flavor of cigarillo smoked most often: Some other flavor – specify
T05_AG1033CG_OS	Adult	Retail location where your cigarillos are purchased most of the time: Somewhere else – specify
T05_AG1048CG_OS	Adult	Brand of cigarillos usually smoked – specify
T05_AG1070CG_OS	Adult	Sub-brand of cigarillo product usually smoked – specify
T05_AG1011FC_OS	Adult	Flavor of filtered cigar when first started smoking: Some other flavor – specify
T05_AG1131FC_OS	Adult	In past 30 days, filtered cigar flavor smoked: Some other flavor – specify
T05_AG1133FC_OS	Adult	In past 30 days, flavor of filtered cigar smoked most often: Some other flavor – specify

**Table E-15. PATH-ATS questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
T05_AG1033FC_OS	Adult	Retail location where your filtered cigars are purchased most of the time: Somewhere else – specify
T05_AG1048FC_OS	Adult	Brand of filtered cigars usually smoked – specify
T05_AG1070FC_OS	Adult	Sub-brand of filtered cigar product usually smoked – specify
T05_AP1011_OS	Adult	Flavor of pipe tobacco when first started smoking: Some other flavor – specify
T05_AP1131_OS	Adult	In past 30 days, pipe tobacco flavor smoked: Some other flavor – specify
T05_AP1133_OS	Adult	In past 30 days, flavor of pipe tobacco smoked most often: Some other flavor – specify
T05_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time: Somewhere else – specify
T05_AP1048_OS	Adult	Brand of pipe tobacco usually smoked – specify
T05_AP1070_OS	Adult	Sub-brand of pipe tobacco usually smoked – specify
T05_AH1011_OS	Adult	Flavor of hookah tobacco when first started smoking: Some other flavor – specify
T05_AH1131_OS	Adult	In past 30 days, hookah tobacco flavor smoked: Some other flavor – specify
T05_AH1133_OS	Adult	In past 30 days, flavor of hookah tobacco smoked most often: Some other flavor – specify
T05_AH1033_OS	Adult	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify
T05_AH1048_OS	Adult	Brand of shisha or hookah tobacco usually smoked – specify
T05_AH1070_OS	Adult	Sub-brand of hookah tobacco usually smoked – specify
T05_AU1011_OS	Adult	Flavor of snus when first started using: Some other flavor – specify
T05_AU1131_OS	Adult	In past 30 days, snus flavor used: Some other flavor – specify
T05_AU1133_OS	Adult	In past 30 days, flavor of snus used most often: Some other flavor – specify
T05_AU1033_OS	Adult	Retail location where your snus is purchased most of the time: Somewhere else – specify
T05_AU1048_OS	Adult	Brand of snus usually used – specify
T05_AU1070_OS	Adult	Sub-brand of snus product usually used – specify
T05_AS1011_OS	Adult	Flavor of smokeless tobacco when first started using: Some other flavor – specify
T05_AS1131_OS	Adult	In past 30 days, smokeless tobacco flavor used: Some other flavor – specify
T05_AS1133_OS	Adult	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor – specify
T05_AS1033_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else – specify
T05_AS1048_OS	Adult	Brand of smokeless tobacco usually used – specify
T05_AS1070_OS	Adult	Sub-brand of smokeless tobacco product usually used – specify
T05_AN0263_OS	Adult	Flavor of e-liquid used [when stopped/last time tried to quit/when quit] smoking cigarettes: Some other flavor – specify
T05_AN0264_OS	Adult	Flavor of e-liquid used most often [when stopped/last time tried to quit/when quit] smoking cigarettes: Some other flavor – specify
T05_AX0203_OS	Adult	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify
T05_AX0677_OS	Adult	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
T05_AM0014_OS	Adult	Current employment status – specify

**Table E-15. PATH-ATS questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
T05_AM0020_OS	Adult	Type of degree program currently enrolled in: Other type of degree program – specify
T05_AM1935_OS	Adult	Main reason not currently working for pay – specify
T05_AX0109	Adult	Current weight: Pounds

**Table E-16. PATH-ATS questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
T05_AM0001	Adult	T05R_A_AGECAT6	Age range when interviewed (6 levels)
T05_AM0002			
T05_AM0018	Adult	T05R_A_AM0018_V2	Recoded education level (5 levels)
T05_AM0030	Adult	T05R_A_AM0030	Recoded total household income in the past 12 months (5 levels)
T05_AT0047	Adult	T05R_A_AT0047	Recoded marital status (3 levels)
T05_AL0040	Adult	T05R_A_AL0040	Indicator of home ownership (2 levels)
T05_AM0017	Adult	T05R_A_AM0017	Recoded reason for not working for pay (5 levels)
T05_AX0093	Adult	T05R_A_AX0093	Extent to which you are able to carry out your everyday physical activities (2 levels)
T05_AM1935	Adult	T05R_A_AM1935	Main reason not currently working for pay (10 levels)

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_CAD10	Adult	Confirm DOB (for continuing respondents)
R06_LAD01	Adult	Respondent's preferred language to complete ACASI interview
R06_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R06_AM0011_NB_01	Adult	Branch served when on active duty: Army
R06_AM0011_NB_02	Adult	Branch served when on active duty: Navy
R06_AM0011_NB_03	Adult	Branch served when on active duty: Air Force
R06_AM0011_NB_04	Adult	Branch served when on active duty: Marine Corps
R06_AM0011_NB_05	Adult	Branch served when on active duty: Coast Guard
R06_AM0069	Adult	Adult respondent is a citizen of the United States
R06_AM0072_OS	Adult	Other language spoken at home: Some other language - specify
R06_AQ9203_OS	Adult	In past 30 days, noticed IQOS or HeatSticks advertised: Somewhere else - specify
R06_AQ9033_OS	Adult	Where IQOS was bought: Somewhere else - specify
R06_AQ1033_OS	Adult	Retail location where your HeatSticks are purchased most of the time: Somewhere else - specify
R06_AZ1002_OS	Adult	Ever used any other tobacco products - specify
R06_AC1118MC_OS	Adult	How you usually get your own cigarettes: Some other way - specify
R06_AC1033MC_OS	Adult	Retail location where your cigarettes are purchased most of the time: Somewhere else - specify
R06_AC8800_OS	Adult	Where you usually throw away cigarette butts after you finish smoking a cigarette: Somewhere else - specify
R06_AC8801_OS	Adult	What you usually do with a cigarette pack or carton when it's empty: Something else - specify
R06_AC1049MC	Adult	Brand of cigarettes usually/last smoked - specify
R06_AC1071MC	Adult	Sub-brand of cigarette product usually/last smoked - specify
R06_AC1118RY_OS	Adult	How you usually get your own roll-your-own cigarette tobacco: Some other way - specify
R06_AC1033RY_OS	Adult	Retail location where your roll-your-own cigarette tobacco is purchased most of the time: Somewhere else - specify
R06_AC8802_OS	Adult	What you usually do with a pouch of roll-your-own cigarette tobacco when it's empty: Something else - specify
R06_AC1049RY	Adult	Brand of roll-your-own cigarette tobacco usually/last smoked - specify
R06_AC1071RY	Adult	Sub-brand of roll-your-own cigarette tobacco product usually/last smoked - specify
R06_AV1011_OS	Adult	Flavor of electronic nicotine product when first started using: Some other flavor - specify
R06_AV9035_OS	Adult	Electronic nicotine product you have used: Something else - specify
R06_AV9044_OS	Adult	Electronic nicotine product used most often: Something else - specify
R06_AV9039_OS	Adult	Type of electronic nicotine product used: Something else - specify
R06_AV9001_OS	Adult	Type of electronic nicotine product used most often: Something else - specify
R06_AV1131_OS	Adult	In past 30 days, [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavor used: Some other flavor - specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_AV1133_OS	Adult	In past 30 days, flavor of electronic nicotine product/electronic nicotine cartridge/e-liquid used most often: Some other flavor – specify
R06_AV1118_OS	Adult	How you usually get your own electronic nicotine products/pods or cartridges/e-liquid: Some other way – specify
R06_AV1033_OS	Adult	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
R06_AV8810_OS	Adult	What you usually do with your disposable electronic nicotine product when it is empty: Something else – specify
R06_AV8811_OS	Adult	What you usually do with a pod or cartridge when it is empty: Something else – specify
R06_AV8812_OS	Adult	What you usually do with the battery for your electronic nicotine product after it no longer works or is needed: Something else – specify
R06_AV8814_OS	Adult	What you usually do with the coils or atomizers for your electronic nicotine product after they no longer work: Something else – specify
R06_AV8816_OS	Adult	What you usually do with a bottle or container of e-liquid when it is empty: Something else – specify
R06_AV8815_OS	Adult	What you usually do with leftover or unused e-liquid: Something else – specify
R06_AV1012_OS	Adult	Flavor of electronic nicotine product brand regularly/last used: Some other flavor – specify
R06_AV1049	Adult	Brand of electronic nicotine products/pods or cartridges/e-liquid usually/last used – specify
R06_AV0335_OS	Adult	Main reason you stopped using electronic nicotine products: Some other reason – specify
R06_AJ1011TC_OS	Adult	Flavor of traditional cigar used as blunts when first started smoking: Some other flavor – specify
R06_AJ1131TC_OS	Adult	In past 30 days, traditional cigar as blunts flavor smoked: Some other flavor – specify
R06_AJ1133TC_OS	Adult	In past 30 days, flavor of traditional cigars smoked as blunts most often: Some other flavor – specify
R06_AJ1118TC_OS	Adult	How you usually get your own traditional cigars for when you smoke blunts: Some other way – specify
R06_AJ1033TC_OS	Adult	Retail location where your traditional cigars for blunts are purchased most of the time: Somewhere else – specify
R06_AJ1012TC_OS	Adult	Traditional cigar as blunts flavor usually/last smoked: Some other flavor – specify
R06_AJ1049TC	Adult	Brand of traditional cigars for blunts usually/last smoked – specify
R06_AJ1071TC	Adult	Sub-brand of traditional cigar as a blunt product usually/last smoked – specify
R06_AJ1011CG_OS	Adult	Flavor of cigarillo used as blunts when first started smoking: Some other flavor – specify
R06_AJ1131CG_OS	Adult	In past 30 days, cigarillo as blunts flavor smoked: Some other flavor – specify
R06_AJ1133CG_OS	Adult	In past 30 days, flavor of cigarillos smoked as blunts most often: Some other flavor – specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_AJ1118CG_OS	Adult	How you usually get your own cigarillos for when you smoke blunts: Some other way – specify
R06_AJ1033CG_OS	Adult	Retail location where your cigarillos for blunts are purchased most of the time: Somewhere else – specify
R06_AJ1012CG_OS	Adult	Cigarillo as blunts flavor usually/last smoked: Some other flavor – specify
R06_AJ1049CG	Adult	Brand of cigarillos for blunts usually/last smoked – specify
R06_AJ1071CG	Adult	Sub-brand of cigarillo as a blunt product usually/last smoked – specify
R06_AJ1011FC_OS	Adult	Flavor of filtered cigar used as blunts when first started smoking: Some other flavor – specify
R06_AJ1131FC_OS	Adult	In past 30 days, filtered cigar as blunts flavor smoked: Some other flavor – specify
R06_AJ1133FC_OS	Adult	In past 30 days, flavor of filtered cigars smoked as blunts most often: Some other flavor – specify
R06_AJ1118FC_OS	Adult	How you usually get your own filtered cigars for when you smoke blunts: Some other way – specify
R06_AJ1033FC_OS	Adult	Retail location where your filtered cigars for blunts are purchased most of the time: Somewhere else – specify
R06_AJ1012FC_OS	Adult	Filtered cigar as blunts flavor usually/last smoked: Some other flavor – specify
R06_AJ1049FC	Adult	Brand of filtered cigars for blunts usually/last smoked – specify
R06_AJ1071FC	Adult	Sub-brand of filtered cigar as a blunt product usually/last smoked – specify
R06_AG1011TC_OS	Adult	Flavor of traditional cigar when first started smoking: Some other flavor – specify
R06_AG1131TC_OS	Adult	In past 30 days, traditional cigar flavor smoked: Some other flavor – specify
R06_AG1133TC_OS	Adult	In past 30 days, flavor of traditional cigar smoked most often: Some other flavor – specify
R06_AG1118TC_OS	Adult	How you usually get your own traditional cigars: Some other way – specify
R06_AG1033TC_OS	Adult	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
R06_AG1012TC_OS	Adult	Traditional cigar flavor usually/last smoked: Some other flavor – specify
R06_AG1049TC	Adult	Brand of traditional cigars usually/last smoked – specify
R06_AG1071TC	Adult	Sub-brand of traditional cigar product usually/last smoked – specify
R06_AG1011CG_OS	Adult	Flavor of cigarillo when first started smoking: Some other flavor – specify
R06_AG1131CG_OS	Adult	In past 30 days, cigarillo flavor smoked: Some other flavor – specify
R06_AG1133CG_OS	Adult	In past 30 days, flavor of cigarillo smoked most often: Some other flavor – specify
R06_AG1118CG_OS	Adult	How you usually get your own cigarillos: Some other way – specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_AG1033CG_OS	Adult	Retail location where your cigarillos are purchased most of the time: Somewhere else – specify
R06_AG1012CG_OS	Adult	Cigarillo flavor usually/last smoked: Some other flavor – specify
R06_AG1049CG	Adult	Brand of cigarillos usually/last smoked – specify
R06_AG1071CG	Adult	Sub-brand of cigarillo product usually/last smoked – specify
R06_AG1011FC_OS	Adult	Flavor of filtered cigar when first started smoking: Some other flavor – specify
R06_AG1131FC_OS	Adult	In past 30 days, filtered cigar flavor smoked: Some other flavor – specify
R06_AG1133FC_OS	Adult	In past 30 days, flavor of filtered cigar smoked most often: Some other flavor – specify
R06_AG1118FC_OS	Adult	How you usually get your own filtered cigars: Some other way – specify
R06_AG1033FC_OS	Adult	Retail location where your filtered cigars are purchased most of the time: Somewhere else – specify
R06_AG1012FC_OS	Adult	Filtered cigar flavor usually/last smoked: Some other flavor – specify
R06_AG1049FC	Adult	Brand of filtered cigars usually/last smoked – specify
R06_AG1071FC	Adult	Sub-brand of filtered cigar product usually/last smoked – specify
R06_AP1011_OS	Adult	Flavor of pipe tobacco when first started smoking: Some other flavor – specify
R06_AP1131_OS	Adult	In past 30 days, pipe tobacco flavor smoked: Some other flavor – specify
R06_AP1133_OS	Adult	In past 30 days, flavor of pipe tobacco smoked most often: Some other flavor – specify
R06_AP1118_OS	Adult	How you usually get your own pipe tobacco: Some other way – specify
R06_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time: Somewhere else – specify
R06_AP1012_OS	Adult	Pipe tobacco flavor usually/last smoked: Some other flavor – specify
R06_AP1049	Adult	Brand of pipe tobacco usually/last smoked – specify
R06_AP1071	Adult	Sub-brand of pipe tobacco product usually/last smoked – specify
R06_AH1011_OS	Adult	Flavor of hookah tobacco when first started smoking: Some other flavor – specify
R06_AH9045_OS	Adult	Location when first tried a hookah: Somewhere else – specify
R06_AH9011_OS	Adult	Place where usually smoke/smoked hookah: Somewhere else – specify
R06_AH1131_OS	Adult	In past 30 days, hookah tobacco flavor smoked: Some other flavor – specify
R06_AH1133_OS	Adult	In past 30 days, flavor of hookah tobacco smoked most often: Some other flavor – specify
R06_AH1118_OS	Adult	How you usually get your own hookah tobacco: Some other way – specify
R06_AH1033_OS	Adult	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify
R06_AH1012_OS	Adult	Hookah tobacco flavor usually/last smoked: Some other flavor – specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_AH1049	Adult	Brand of shisha or hookah tobacco usually/last smoked – specify
R06_AH1071	Adult	Sub-brand of hookah tobacco product usually/last smoked – specify
R06_AU1011_OS	Adult	Flavor of snus when first started using: Some other flavor – specify
R06_AU1131_OS	Adult	In past 30 days, snus flavor used: Some other flavor – specify
R06_AU1133_OS	Adult	In past 30 days, flavor of snus used most often: Some other flavor – specify
R06_AU1118_OS	Adult	How you usually get your own snus: Some other way – specify
R06_AU1033_OS	Adult	Retail location where your snus is purchased most of the time: Somewhere else – specify
R06_AU1012_OS	Adult	Snus flavor usually/last used: Some other flavor – specify
R06_AU1049	Adult	Brand of snus usually/last used – specify
R06_AU1071	Adult	Sub-brand of snus product usually/last used – specify
R06_AS1011_OS	Adult	Flavor of smokeless tobacco when first started using: Some other flavor – specify
R06_AS1131_OS	Adult	In past 30 days, smokeless tobacco flavor used: Some other flavor – specify
R06_AS1133_OS	Adult	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor – specify
R06_AS1118_OS	Adult	How you usually get your own smokeless tobacco: Some other way – specify
R06_AS1033_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else – specify
R06_AS1012_OS	Adult	Smokeless tobacco flavor usually/last used: Some other flavor – specify
R06_AS1049	Adult	Brand of smokeless tobacco usually/last used – specify
R06_AS1071	Adult	Sub-brand of smokeless tobacco product usually/last used – specify
R06_AN0349E_OS	Adult	Strength of nicotine patch last used (current or recent former established non-marijuana electronic nicotine product users): Some other strength – specify
R06_AN0263_OS	Adult	Flavor of e-liquid used during last quit attempt: Some other flavor – specify
R06_AN0264_OS	Adult	Flavor of e-liquid used most often during last quit attempt: Some other flavor – specify
R06_AN0265_OS	Adult	Type of e-cigarette or electronic nicotine product used most often during last quit attempt: Something else – specify
R06_AN0336_OS	Adult	Main reason you used a nicotine patch, gum, inhaler, nasal spray, or lozenge: Some other reason – specify (Current or recent former non-electronic tobacco users)
R06_AN0349_OS	Adult	Strength of nicotine patch last used (current or recent former established non-electronic tobacco users): Some other strength – specify
R06_AX0601_OS	Adult	In past 30 days, saw anyone using e-cigarettes or electronic nicotine products in or around your school: Yes, somewhere else at school not listed here – specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_AX0602_OS	Adult	In past 30 days, saw anyone smoking a cigarette in or around your school: Yes, somewhere else at school not listed here – specify
R06_AX0190_NB	Adult	Is deaf or has serious difficulty hearing
R06_AX0191_NB	Adult	Is blind or has serious difficulty seeing, even when wearing glasses
R06_AN0349H_OS	Adult	Strength of nicotine patch last used: Some other strength – specify (Long-term former tobacco users)
R06_AX0290_OS	Adult	In past 30 days, way in which you used marijuana: Use marijuana some other way – specify
R06_AX0373_OS	Adult	In past 30 days, saw information on social media related to tobacco products, e-cigarettes, or other electronic nicotine products: Something else – specify
R06_AX0374_OS	Adult	In past 30 days, when you saw posts of people using or talking about tobacco products, e-cigarettes, or other electronic nicotine products, it was posted by: Other – specify
R06_AX0375_OS	Adult	In past 30 days, when you saw ads promoting or encouraging the use of tobacco products, e-cigarettes, or other electronic nicotine products on social media, it was posted by: Other – specify
R06_AX0378_OS	Adult	In past 30 days, product seen on social media: Some other type of tobacco product – specify
R06_AX0379_OS	Adult	In past 30 days, product seen most often on social media: Some other type of tobacco product – specify
R06_AX0203_OS	Adult	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify
R06_AX0677_OS	Adult	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
R06_AM0067	Adult	Enrolled in High School
R06_AM0068	Adult	Current grade in school
R06_AM0020_OS	Adult	Type of degree program currently enrolled in: Other type of degree program – specify
R06_AM0021	Adult	Sexual attraction to gender
R06_AM0061	Adult	Consider yourself to be transgender
R06_AM0062	Adult	Consider self to be male-to-female, female-to-male, or non-conforming
R06_AM1935_OS	Adult	Main reason you are not currently working for pay: Some other reason – specify
R06_LCAD01	Adult	Language in which CAPI portions of adult interview were conducted
R06_AX0214	Adult	OK to scan tobacco products
R06_AX0218	Adult	Number of tobacco products given by respondent to scan
R06_AX0217_1A	Adult	Wave 6 Adult tobacco Product 1, barcode scan/entry 1
R06_AX0233_01	Adult	First bar code scan: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_01	Adult	First bar code scan/entry: Same as brand you usually use
R06_AX0217_2A	Adult	Wave 6 Adult tobacco Product 2, barcode scan/entry 1
R06_AX0233_02	Adult	Second bar code scan/entry: Last time used this type of tobacco it was from this specific package or container

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_AX0237_02	Adult	Second bar code scan/entry: Same as brand you usually use
R06_AX0217_3A	Adult	Wave 6 Adult tobacco Product 3, barcode scan/entry 1
R06_AX0233_03	Adult	Third bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_03	Adult	Third bar code scan/entry: Same as brand you usually use
R06_AX0217_4A	Adult	Wave 6 Adult tobacco Product 4, barcode scan/entry 1
R06_AX0233_04	Adult	Fourth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_04	Adult	Fourth bar code scan/entry: Same as brand you usually use
R06_AX0217_5A	Adult	Wave 6 Adult tobacco Product 5, barcode scan/entry 1
R06_AX0233_05	Adult	Fifth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_05	Adult	Fifth bar code scan/entry: Same as brand you usually use
R06_AX0217_6A	Adult	Wave 6 Adult tobacco Product 6, barcode scan/entry 1
R06_AX0233_06	Adult	Sixth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_06	Adult	Sixth bar code scan/entry: Same as brand you usually use
R06_AX0217_7A	Adult	Wave 6 Adult tobacco Product 7, barcode scan/entry 1
R06_AX0233_07	Adult	Seventh bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_07	Adult	Seventh bar code scan/entry: Same as brand you usually use
R06_AX0217_8A	Adult	Wave 6 Adult tobacco Product 8, barcode scan/entry 1
R06_AX0233_08	Adult	Eighth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_08	Adult	Eighth bar code scan/entry: Same as brand you usually use
R06_AX0217_9A	Adult	Wave 6 Adult tobacco Product 9, barcode scan/entry 1
R06_AX0233_09	Adult	Ninth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_09	Adult	Ninth bar code scan/entry: Same as brand you usually use
R06_AX0217_10A	Adult	Wave 6 Adult tobacco Product 10, barcode scan/entry 1
R06_AX0233_10	Adult	Tenth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R06_AX0237_10	Adult	Tenth bar code scan/entry: Same as brand you usually use
R06_AX0232	Adult	Interviewer manually entered one or more barcode values
R06_CPT05	Parent	Confirm DOB for shadow youth who is expected to be a youth at Wave 6
R06_CPT05C	Parent	Confirm DOB for youth from a prior wave who is expected to be a youth at Wave 6
R06_CPT07	Parent	Date of birth (corrected values for continuing youth/new values for new baseline youth)
R06_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
R06_E_YOUTH	Youth	Emancipated youth
R06_LCY01	Youth	Language in which CAPI portions of youth interview were conducted
R06LEY01	Youth	Emancipated Youth's preferred language to complete ACASI interview
R06LYH01	Youth	Youth's preferred language to complete ACASI interview
R06_PARENT_PERSONID	Parent	Wave 6 Parent/guardian Participant ID Number

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_PM0016_NB	Parent	Youth was identified as a twin or part of a multiple birth
R06_PM0017_NB	Parent	Youth and sibling are identical twins
R06_PM0019_NB	Parent	Any sibling in multiple birth identical to youth
R06_PM0021_NB	Parent	First name of sibling that youth is a twin of
R06_PM0030_NB	Parent	First names of siblings in multiple birth that are identical to youth
R06_PM0035_NB	Parent	First names of siblings that are in the multiple birth with youth
R06_PM0052	Parent	Confirm parent's spouse is same as reported in previous wave(s)
R06_PM0053	Parent	Confirm parent's relationship to youth
R06_PM0058_NEW_AGE	Parent	Age of other parental figure/guardian
R06_PM0058_NEW_FNAME	Parent	First name of other parental figure or guardian - specify
R06_PM0058_NEW_LNAME	Parent	Last name of other parental figure or guardian - specify
R06_PM0058_PERSONID	Parent	Other parental figure/guardian PERSONID
R06_PM0059_OS_1	Parent	Other parental figure/guardian's relationship to sampled youth: Other relative - specify
R06_PM0059_OS_2	Parent	Other parental figure/guardian's relationship to sampled youth: Other non-relative - specify
R06_PM0061_NEW_AGE	Parent	Age of second other parental figure/guardian
R06_PM0061_NEW_FNAME	Parent	First name of second other parental figure or guardian - specify
R06_PM0061_NEW_LNAME	Parent	Last name of second other parental figure or guardian - specify
R06_PM0061_PERSONID	Parent	Second other parental figure/guardian PERSONID
R06_PM0062_OS_1	Parent	Second other parental figure/guardian's relationship to youth: Other relative - specify
R06_PM0062_OS_2	Parent	Second other parental figure/guardian's relationship to youth: Other non-relative - specify
R06_PM0067	Parent	Parent respondent is currently enrolled in high school
R06_PM0068	Parent	Parent respondent grade (if currently enrolled in high school)
R06_PM0072_OS	Parent	Other language spoken at home: Some other language - specify (parent respondent)
R06_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative - specify
R06_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative - specify
R06_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative - specify
R06_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative - specify
R06_PT0046_NEW_AGE	Parent	Age of new spouse/partner
R06_PT0046_PERSONID	Parent	Parent's spouse/partner's PERSONID (if different from previous waves)
R06_PT0046_NEW_FNAME	Parent	First name of spouse/partner that lives with parent in DU - specify
R06_PT0046_NEW_LNAME	Parent	Last name of spouse/partner that lives with parent in DU - specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_PX0186_NB	Parent	Youth has serious difficulty walking or climbing stairs
R06_PX0190_NB	Parent	Youth is deaf or has serious difficulty hearing
R06_PX0191_NB	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
R06_YC1033_OS	Youth	Retail location where your cigarettes are purchased most of the time: Somewhere else – specify
R06_YC1049	Youth	Brand of cigarettes usually/last smoked – specify
R06_YC1071	Youth	Sub-brand of cigarette product usually/last smoked – specify
R06_YC1118_OS	Youth	In past 30 days, how you usually got your own cigarettes: Some other way – specify
R06_YG1011CL_OS	Youth	Flavor of first cigarillo smoked: Some other flavor – specify
R06_YG1011FC_OS	Youth	Flavor of first filtered cigar smoked: Some other flavor – specify
R06_YG1011TC_OS	Youth	Flavor of first traditional cigar smoked: Some other flavor – specify
R06_YG1033CL_OS	Youth	Retail location where your cigarillos are purchased most of the time: Somewhere else – specify
R06_YG1033FC_OS	Youth	Retail location where your filtered cigars are purchased most of the time: Somewhere else – specify
R06_YG1033TC_OS	Youth	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
R06_YG1049CL	Youth	Brand of cigarillos usually/last smoked – specify
R06_YG1049FC	Youth	Brand of filtered cigars usually/last smoked – specify
R06_YG1049TC	Youth	Brand of traditional cigars usually/last smoked – specify
R06_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked – specify
R06_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked – specify
R06_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked – specify
R06_YG1118CL_OS	Youth	In past 30 days, how you usually got your own cigarillos: Some other way – specify
R06_YG1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars: Some other way – specify
R06_YG1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars: Some other way – specify
R06_YG1131CL_OS	Youth	In past 30 days, smoked cigarillos flavored to taste like: Some other flavor – specify
R06_YG1131FC_OS	Youth	In past 30 days, smoked filtered cigars flavored to taste like: Some other flavor – specify
R06_YG1131TC_OS	Youth	In past 30 days, smoked traditional cigars flavored to taste like: Some other flavor – specify
R06_YG1133CL_OS	Youth	In past 30 days, flavor of cigarillos smoked most often: Some other flavor – specify
R06_YG1133FC_OS	Youth	In past 30 days, flavor of filtered cigars smoked most often: Some other flavor – specify
R06_YG1133TC_OS	Youth	In past 30 days, flavor of traditional cigars smoked most often: Some other flavor – specify
R06_YH1011_OS	Youth	Flavor of first shisha or hookah tobacco smoked: Some other flavor – specify
R06_YH1033_OS	Youth	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
R06_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
R06_YH1073	Youth	Smoke hookah because: It is part of my cultural tradition
R06_YH1118_OS	Youth	In past 30 days, how you usually got your own shisha or hookah tobacco: Some other way – specify
R06_YH1131_OS	Youth	In past 30 days, smoked hookah tobacco flavored to taste like: Some other flavor – specify
R06_YH1133_OS	Youth	In past 30 days, flavor of hookah smoked most often: Some other flavor – specify
R06_YH9011_OS	Youth	Usually smoke a hookah: Somewhere else – specify
R06_YH9045_OS	Youth	Where you were when you first tried a hookah, even one or two puffs – specify
R06_YJ1011CG_OS	Youth	Flavor of first cigarillo smoked as a blunt: Some other flavor – specify
R06_YJ1011FC_OS	Youth	Flavor of first filtered cigar smoked as a blunt: Some other flavor – specify
R06_YJ1011TC_OS	Youth	Flavor of first traditional cigar smoked as a blunt: Some other flavor – specify
R06_YJ1033CG_OS	Youth	Retail location where your cigarillos for blunts are purchased most of the time: Somewhere else – specify
R06_YJ1033FC_OS	Youth	Retail location where your filtered cigars for blunts are purchased most of the time: Somewhere else – specify
R06_YJ1033TC_OS	Youth	Retail location where your traditional cigars for blunts are purchased most of the time: Somewhere else – specify
R06_YJ1049CG	Youth	Brand of cigarillos usually/last smoked as a blunt – specify
R06_YJ1049FC	Youth	Brand of filtered cigars usually/last smoked as a blunt – specify
R06_YJ1049TC	Youth	Brand of traditional cigars usually/last smoked as a blunt – specify
R06_YJ1071CG	Youth	Sub-brand of cigarillo product usually/last smoked as a blunt – specify
R06_YJ1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked as a blunt – specify
R06_YJ1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked as a blunt – specify
R06_YJ1118CG_OS	Youth	In past 30 days, how you usually got your own cigarillo as a blunt: Some other way – specify
R06_YJ1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars as a blunt: Some other way – specify
R06_YJ1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars as a blunt: Some other way – specify
R06_YJ1131CG_OS	Youth	In past 30 days, smoked cigarillos as blunts flavored to taste like: Some other flavor – specify
R06_YJ1131FC_OS	Youth	In past 30 days, smoked filtered cigars as blunts flavored to taste like: Some other flavor – specify
R06_YJ1131TC_OS	Youth	In past 30 days, smoked traditional cigars as blunts flavored to taste like: Some other flavor – specify
R06_YJ1133CG_OS	Youth	In past 30 days, flavor of cigarillos smoked as blunts most often: Some other flavor – specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_YJ1133FC_OS	Youth	In past 30 days, flavor of filtered cigars smoked as blunts most often: Some other flavor – specify
R06_YJ1133TC_OS	Youth	In past 30 days, flavor of traditional cigars smoked as blunts most often: Some other flavor – specify
R06_YM0007	Youth	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R06_YM0011_01	Youth	Branch served when on active duty: Army
R06_YM0011_02	Youth	Branch served when on active duty: Navy
R06_YM0011_03	Youth	Branch served when on active duty: Air Force
R06_YM0011_04	Youth	Branch served when on active duty: Marine Corps
R06_YM0011_05	Youth	Branch served when on active duty: Coast Guard
R06_YM0020	Youth	Last grade/year in school completed
R06_YM0021	Youth	Sexual attraction to gender
R06_YM0061	Youth	Consider yourself to be transgender
R06_YM0062	Youth	Transgender category
R06_YM0063	Youth	Sexual orientation
R06_YM0066	Youth	How well you speak English (youth)
R06_YM0069	Youth	Youth respondent is a citizen of the United States
R06_YM0070	Youth	How well you read English (youth)
R06_YM0072_OS	Youth	Other language spoken at home: Some other language – specify (youth respondent)
R06_YM0073	Youth	How well you write in English (youth)
R06_YN0129_OS	Youth	In past 12 months, tried to completely stop using: Other – specify
R06_YN0336_OS	Youth	Main reason you used a nicotine patch, nicotine gum, nicotine inhaler, nicotine nasal spray or lozenge (past year tobacco users): Some other reason – specify
R06_YN0349_OS	Youth	Strength of nicotine patch last used (past year tobacco users): Some other strength – specify
R06_YN0349H_OS	Youth	Strength of nicotine patch last used (past year tobacco non-users): Some other strength – specify
R06_YP1033_OS	Youth	Retail location where your pipe tobacco is purchased most of the time: Somewhere else – specify
R06_YP1118_OS	Youth	In past 30 days, how you usually got your own pipe tobacco: Some other way – specify
R06_YS1011_OS	Youth	Flavor of first smokeless tobacco product used: Some other flavor – specify
R06_YS1033_OS	Youth	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else – specify
R06_YS1049	Youth	Brand of smokeless tobacco usually/last used – specify
R06_YS1071	Youth	Sub-brand of smokeless tobacco product usually/last used – specify
R06_YS1118_OS	Youth	In past 30 days, how you usually got your own smokeless tobacco: Some other way – specify
R06_YS1131_OS	Youth	In past 30 days, used smokeless tobacco flavored to taste like: Some other flavor – specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_YS1133_OS	Youth	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor – specify
R06_YQ9203_OS	Youth	In past 30 days, noticed IQOS or HeatSticks advertised: Somewhere else – specify
R06_YQ9033_OS	Youth	Where IQOS was bought: Somewhere else – specify
R06_YQ1033_OS	Youth	Retail location where your HeatSticks are purchased most of the time: Somewhere else – specify
R06_YU1011_OS	Youth	Flavor of first snus used: Some other flavor – specify
R06_YU1033_OS	Youth	Retail location where your snus is purchased most of the time: Somewhere else – specify
R06_YU1049	Youth	Brand of snus usually/last used – specify
R06_YU1071	Youth	Sub-brand of snus usually/last used – specify
R06_YU1118_OS	Youth	In past 30 days, how you usually got your own snus: Some other way – specify
R06_YU1131_OS	Youth	In past 30 days, used snus flavored to taste like: Some other flavor – specify
R06_YU1133_OS	Youth	In past 30 days, flavor of snus used most often: Some other flavor – specify
R06_YV1011_OS	Youth	Flavor of first electronic nicotine product used: Some other flavor – specify
R06_YV1033_OS	Youth	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
R06_YV1049	Youth	Brand of electronic nicotine products/pods or cartridges/e-liquid usually/last – specify
R06_YV1118_OS	Youth	In past 30 days, how you usually got your own [electronic nicotine products/pods or cartridges/e-liquid]: Some other way – specify
R06_YV1131_OS	Youth	In past 30 days, used [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavored to taste like: Some other flavor – specify
R06_YV1133_OS	Youth	In past 30 days, flavor of electronic nicotine product/cartridges/e-liquid used most often: Some other flavor – specify
R06_YV9001_OS	Youth	Type of electronic nicotine product used most often – specify
R06_YV9035_OS	Youth	Ever used some other electronic nicotine product – specify
R06_YV9044_OS	Youth	Type of electronic nicotine product used most often: Something else – specify
R06_YV9039_OS	Youth	Type of electronic nicotine product used: Something else – specify
R06_YX0136	Youth	Currently pregnant
R06_YX0137_NN	Youth	Number of weeks/months pregnant – Number
R06_YX0137_UN	Youth	Number of weeks/months pregnant – Unit
R06_YX0373_OS	Youth	In past 30 days, saw information on social media related to tobacco products, e-cigarettes, or other electronic nicotine products: Something else – specify
R06_YX0374_OS	Youth	In past 30 days, when you saw posts of people using or talking about tobacco products, e-cigarettes, or other electronic nicotine products, it was posted by: Other – specify

**Table E-17. Wave 6 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R06_YX0375_OS	Youth	In past 30 days, when you saw ads promoting or encouraging the use of tobacco products, e-cigarettes, or other electronic nicotine products on social media, it was posted by: Other – specify
R06_YX0378_OS	Youth	In past 30 days, product seen on social media: Some other type of tobacco product – specify
R06_YX0379_OS	Youth	In past 30 days, product seen most often on social media: Some other type of tobacco product – specify
R06_YX0203_OS	Youth	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify
R06_YX0290_OS	Youth	In past 30 days, way in which you used marijuana: Use marijuana some other way – specify
R06_YX0677_OS	Youth	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
R06_YY0601_OS	Youth	First type of tobacco you tried: Other – specify
R06_YZ1002_OS	Youth	Ever used any other tobacco products – specify
R06_YX0601_OS	Youth	In past 30 days, saw anyone using e-cigarettes or electronic nicotine products in or around your school: Yes, somewhere else at school not listed here – specify
R06_YX0602_OS	Youth	In past 30 days, saw anyone smoking a cigarette in or around your school: Yes, somewhere else at school not listed here – specify

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_AM0001	Adult	R06R_A_AGECAT6	Wave 6 Adult Age (6 levels)
R06_AM0002			
R06_AM0003			
R06_AT0047	Adult	R06R_A_AT0047	Recoded marital status (3 levels)
R06_AM0065	Adult	R06R_A_AM0065_V2	Recoded number of years in US (2 levels)
R06_AC1006	Adult	R06R_A_AC1006	Age group when first smoked part or all of a cigarette (6 levels)
R06_AC1120			
R06_AC1007	Adult	R06R_A_AC1007	Age group when you first started smoking cigarettes fairly regularly (6 levels)
R06_AC1121			
R06_AC1020	Adult	R06R_A_AC1020	Age group when you first started smoking cigarettes every day (6 levels)
R06_AC1122			
R06_AV1006	Adult	R06R_A_AV1006	Age group when first used an electronic nicotine product, even one or two times (6 levels)
R06_AV1120			
R06_AV1007	Adult	R06R_A_AV1007	Age group when you first started using electronic nicotine products fairly regularly (6 levels)
R06_AV1121			
R06_AV1020	Adult	R06R_A_AV1020	Age group when you first started using electronic nicotine products every day (6 levels)
R06_AV1122			
R06_AV7701	Adult	R06R_A_AV7701	Age group when you were last evaluated or treated by a doctor for a respiratory condition related to your use of electronic nicotine product (6 levels)
R06_AV7702			
R06_AG1006TC	Adult	R06R_A_AG1006TC	Age group when first smoked part or all of a traditional cigar, even one or two puffs (6 levels)
R06_AG1120TC			
R06_AG1007TC	Adult	R06R_A_AG1007TC	Age group when you first started smoking traditional cigars fairly regularly (6 levels)
R06_AG1121TC			
R06_AG1020TC	Adult	R06R_A_AG1020TC	Age group when you first started smoking traditional cigars every day (6 levels)
R06_AG1122TC			
R06_AG1006CG	Adult	R06R_A_AG1006CG	Age group when first smoked part or all of a cigarillo, even one or two puffs (6 levels)
R06_AG1120CG			
R06_AG1007CG	Adult	R06R_A_AG1007CG	Age group when you first started smoking cigarillos fairly regularly (6 levels)
R06_AG1121CG			
R06_AG1020CG	Adult	R06R_A_AG1020CG	Age group when you first started smoking cigarillos every day (6 levels)
R06_AG1122CG			
R06_AG1006FC	Adult	R06R_A_AG1006FC	Age group when first smoked part or all of a filtered cigar, even one or two puffs (6 levels)
R06_AG1120FC			
R06_AG1007FC	Adult	R06R_A_AG1007FC	Age group when you first started smoking filtered cigars fairly regularly (6 levels)
R06_AG1121FC			

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_AG1020FC R06_AG1122FC	Adult	R06R_A_AG1020FC	Age group when you first started smoking filtered cigars every day (6 levels)
R06_AP1006 R06_AP1120	Adult	R06R_A_AP1006	Age group when you first smoked a pipe filled with tobacco, even one or two puffs (6 levels)
R06_AP1007 R06_AP1121	Adult	R06R_A_AP1007	Age group when you first started smoking a pipe filled with tobacco fairly regularly (6 levels)
R06_AP1020 R06_AP1122	Adult	R06R_A_AP1020	Age group when you first started smoking a pipe filled with tobacco every day (6 levels)
R06_AH1006 R06_AH1120	Adult	R06R_A_AH1006	Age group when first smoked tobacco in a hookah, even one or two puffs (6 levels)
R06_AH1007 R06_AH1121	Adult	R06R_A_AH1007	Age group when you first started smoking hookah fairly regularly (6 levels)
R06_AH1020 R06_AH1122	Adult	R06R_A_AH1020	Age group when you first started smoking hookah every day (6 levels)
R06_AU1006 R06_AU1120	Adult	R06R_A_AU1006	Age group when first used snus, even one or two times (6 levels)
R06_AU1007 R06_AU1121	Adult	R06R_A_AU1007	Age group when you first started using snus pouches fairly regularly (6 levels)
R06_AU1020 R06_AU1122	Adult	R06R_A_AU1020	Age group when you first started using snus pouches every day (6 levels)
R06_AS1006 R06_AS1120	Adult	R06R_A_AS1006	Age group when you first used smokeless tobacco, even one or two times (6 levels)
R06_AS1007 R06_AS1121	Adult	R06R_A_AS1007	Age group when you first started using smokeless tobacco fairly regularly (6 levels)
R06_AS1020 R06_AS1122	Adult	R06R_A_AS1020	Age group when you first started using smokeless tobacco every day (6 levels)
R06_AX0114 R06_AX0253	Adult	R06R_A_AX0114	Age range you were in when you were first told you had high blood pressure (6 levels)
R06_AX0115 R06_AX0254	Adult	R06R_A_AX0115	Age range you were in when you were first told you had high cholesterol (6 levels)
R06_AX0116 R06_AX0255	Adult	R06R_A_AX0116	Age range you were in when you were first told you had congestive heart failure (6 levels)
R06_AX0117 R06_AX0256	Adult	R06R_A_AX0117	Age range you were in when you were first told you had a stroke (6 levels)

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_AX0112 R06_AX0252	Adult	R06R_A_AX0112	Age range you were in when you were first told you had a heart attack (6 levels)
R06_AX0120 R06_AX0257	Adult	R06R_A_AX0120	Age range you were in when you were first told you had COPD (6 levels)
R06_AX0121 R06_AX0258	Adult	R06R_A_AX0121	Age range you were in when you were first told you had chronic bronchitis (6 levels)
R06_AX0123 R06_AX0259	Adult	R06R_A_AX0123	Age range you were in when you were first told you had emphysema (6 levels)
R06_AX0124 R06_AX0260	Adult	R06R_A_AX0124	Age range you were in when you were first told you had asthma (6 levels)
R06_AX0131 R06_AX0261	Adult	R06R_A_AX0131	Age range you were in when you were first told you had gum disease (6 levels)
R06_AX0133 R06_AX0262	Adult	R06R_A_AX0133	Age range you were in when you were first told you had pre-cancerous oral lesions (6 levels)
R06_AX0280 R06_AX0263	Adult	R06R_A_AX0280	Age range you were in when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes (6 levels)
R06_AX0143 R06_AX0264	Adult	R06R_A_AX0143	Age range you were in when you were first told you had an ulcer (6 levels)
R06_AX0148 R06_AX0266	Adult	R06R_A_AX0148	Age range you were in when you were first told you had stomach or gastro-intestinal bleeding (6 levels)
R06_AX0150 R06_AX0267	Adult	R06R_A_AX0150	Age range you were in when you were first told you had osteoporosis (6 levels)
R06_AX0198 R06_AX0268	Adult	R06R_A_AX0198	Age range you were in when you were first told you had a bone fracture because you have fragile bones (6 levels)
R06_AX0152 R06_AX0269	Adult	R06R_A_AX0152	Age range you were in when you were first told you had cataract or glaucoma (6 levels)
R06_AX0411 R06_AX0412	Adult	R06R_A_AX0411	Age group when you were first told you have poor circulation (PAD or PVD) (6 levels)
R06_AX0703 R06_AX0704	Adult	R06R_A_AX0703	Age group when first told you had macular degeneration (6 levels)
R06_AX0086 R06_AX0087	Adult	R06R_A_AX0086	Age group when you first drank alcohol at all, counting small tastes or sips (6 levels)

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_AX0074 R06_AX0270	Adult	R06R_A_AX0074	Age group when you had your first alcoholic drink, other than small tastes or sips (6 levels)
R06_AX0079 R06_AX0271	Adult	R06R_A_AX0079	Age group when you first used marijuana (6 levels)
R06_AW7701 R06_AW7702	Adult	R06R_A_AW7701	Age group when you were last evaluated or treated by a doctor for a respiratory condition related to your use of marijuana in electronic nicotine products (6 levels)
R06_AX0082_01 R06_AX0272_01	Adult	R06R_A_AX0082_01	Age group when first used Ritalin® or Adderall® (6 levels)
R06_AX0082_02 R06_AX0272_02	Adult	R06R_A_AX0082_02	Age group when first used painkillers (6 levels)
R06_AX0082_03 R06_AX0272_03	Adult	R06R_A_AX0082_03	Age group when first used sedatives or tranquilizers (6 levels)
R06_AX0082_04 R06_AX0272_04	Adult	R06R_A_AX0082_04	Age group when first used cocaine or crack (6 levels)
R06_AX0082_05 R06_AX0272_05	Adult	R06R_A_AX0082_05	Age group when first used stimulants like methamphetamine or speed (6 levels)
R06_AX0082_06 R06_AX0272_06	Adult	R06R_A_AX0082_06	Age group when first used heroin (6 levels)
R06_AX0082_07 R06_AX0272_07	Adult	R06R_A_AX0082_07	Age group when first used inhalants or solvents (6 levels)
R06_AX0082_08 R06_AX0272_08	Adult	R06R_A_AX0082_08	Age group when first used hallucinogens (6 levels)
R06_AX0300_12M_01 R06_AX0300_12M_02 R06_AX0300_12M_03 R06_AX0300_12M_04 R06_AX0300_12M_05 R06_AX0300_12M_06 R06_AX0300_12M_07 R06_AX0706	Adult	R06R_A_AX0300_12M	Indicator for adverse pregnancy outcomes resulting in risky birth (2 levels)
R06_AX0135_12M	Adult	R06R_A_AX0135_12M	Outcome of last pregnancy was miscarriage, induced abortion, ectopic or tubal pregnancy or stillbirth (2 levels)
R06_AX0691	Adult	R06R_A_AX0691	Recoded type of living space currently living in (4 levels)
R06_AM0017	Adult	R06R_A_AM0017	Recoded reason for not working for pay (5 levels)
R06_AM0018	Adult	R06R_A_AM0018_V2	Recoded education level (5 levels)
R06_AM0030	Adult	R06R_A_AM0030	Recoded total household income in the past 12 months (5 levels)
R06_AM0033	Adult	R06R_A_AM0033_V2	Recoded education level of mother (5 levels)
R06_AM0034	Adult	R06R_A_AM0034_V2	Recoded education level of father (5 levels)

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_AM0036	Adult	R06R_A_AM0036	Recoded parents' total household income in the past 12 months (5 levels)
R06_AM0026_01 R06_AM0026_02 R06_AM0026_03 R06_AM0026_04 R06_AM0026_05 R06_AM0026_06 R06_AM0026_07 R06_AM0026_08	Adult	R06R_A_AM0026_V2	Indicator of health insurance (2 levels)
R06_AX0145_02 R06_AX0145_03 R06_AX0145_04 R06_AX0145_05 R06_AX0145_09 R06_AX0145_12 R06_AX0145_15 R06_AX0145_18 R06_AX0145_19 R06_AX0145_21 R06_AX0145_16 R06_AX0145_23 R06_AX0145_24 R06_AX0145_25 R06_AX0145_27 R06_AX0145_29 R06_AX0145_30 R06_AX0145_31	Adult	R06R_A_AX0145_NONTOB	Type of cancer is a non-tobacco related cancer (2 levels)
R06_AX0145_01 R06_AX0145_06 R06_AX0145_07 R06_AX0145_08 R06_AX0145_10 R06_AX0145_11 R06_AX0145_13 R06_AX0145_14 R06_AX0145_17 R06_AX0145_20 R06_AX0145_22 R06_AX0145_26 R06_AX0145_28	Adult	R06R_A_AX0145_TOB	Type of cancer is a tobacco related cancer (2 levels)
R06_AX0313 R06_AX0679_FT R06_AX0679_IN R06_AX0316 R06_AX0109 R06_AX0312	Adult	R06R_A_BMI	Body mass index
R06_AX0093	Adult	R06R_A_AX0093	Extent to which you are able to carry out your everyday physical activities (2 levels)

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_AL0040	Adult	R06R_A_AL0040	Indicator of home ownership (2 levels)
R06_AM0063	Adult	R06R_A_SEXORIENT2	Recoded sexual orientation (2 levels)
R06_AM0042	Adult	R06R_A_AM0042	Recoded where you currently live (3 levels)
R06_AX0146_02 R06_AX0265_02 R06_AX0146_03 R06_AX0265_03 R06_AX0146_04 R06_AX0265_04 R06_AX0146_05 R06_AX0265_05 R06_AX0146_09 R06_AX0265_09 R06_AX0146_12 R06_AX0265_12 R06_AX0146_15 R06_AX0265_15 R06_AX0146_18 R06_AX0265_18 R06_AX0146_19 R06_AX0265_19 R06_AX0146_21 R06_AX0265_21 R06_AX0146_16 R06_AX0265_16 R06_AX0146_23 R06_AX0265_23 R06_AX0146_24 R06_AX0265_24 R06_AX0146_25 R06_AX0265_25 R06_AX0146_27 R06_AX0265_27 R06_AX0146_29 R06_AX0265_29 R06_AX0146_30 R06_AX0265_30 R06_AX0146_31 R06_AX0265_31	Adult	R06R_A_AX0146_NONTOB	Age range when first non tobacco-related cancer was diagnosed (6 levels)

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_AX0146_01	Adult	R06R_A_AX0146_TOB	Age range when first tobacco-related cancer was diagnosed (6 levels)
R06_AX0265_01			
R06_AX0146_06			
R06_AX0265_06			
R06_AX0146_07			
R06_AX0265_07			
R06_AX0146_08			
R06_AX0265_08			
R06_AX0146_10			
R06_AX0265_10			
R06_AX0146_11			
R06_AX0265_11			
R06_AX0146_13			
R06_AX0265_13			
R06_AX0146_14			
R06_AX0265_14			
R06_AX0146_17			
R06_AX0265_17			
R06_AX0146_20			
R06_AX0265_20			
R06_AX0146_22			
R06_AX0265_22			
R06_AX0146_26			
R06_AX0265_26			
R06_AX0146_28			
R06_AX0265_28			
R06_AX0066_01	Adult	R06R_A_AX0066_V2	Recoded anyone who lives with you now who uses tobacco (4 levels)
R06_AX0066_02			
R06_AX0066_03			
R06_AX0066_04			
R06_AX0066_05			
R06_AX0066_06			
R06_AX0066_07			
R06_AX0066_08			
R06_AX0066_09			
R06_AX0066_10			
R06_AM0072_02	Adult	R06R_A_AM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (adult)
R06_AM0072_03			
R06_AM0072_04			
R06_AM0072_05			
R06_AM0072_06			
R06_AM0072_07			
R06_AM0072_08			
R06_AM0072_09			
R06_AM0072_10			
R06_AX0757	Adult	R06R_A_AX0757	Age range when you were first told you had schizophrenia, schizoaffective disorder or psychosis (6 levels)
R06_AX0758			

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_AX0762 R06_AX0772	Adult	R06R_A_AX0762	Age range when you were first told you had a psychotic illness or episode (6 levels)
R06_AM1935	Adult	R06R_A_AM1935	Main reason you are not currently working for pay (10 levels)
R06_PT0001	Parent	R06R_Y_PT0001_V2	Recoded parent or guardian relationship to youth (3 levels)
R06_PT0047	Parent	R06R_Y_PT0047	Recoded parent or guardian marital status (3 levels)
R06_PT0045 R06_PM0057 R06_PM0060	Parent	R06R_P_OTHPAR_INHH	Wave 6 Recoded youth has any parental figures/guardians other than the parent/guardian in the house (2 levels)
R06_PT0002	Parent	R06R_Y_PT0002_V2	Recoded parent's spouse or partner relationship to youth (3 levels)
R06_PM0059	Parent	R06R_Y_PM0059_V2	Recoded other parental figure/guardian's relationship to youth (3 levels)
R06_PM0062	Parent	R06R_Y_PM0062_V2	Recoded second other parental figure/guardian's relationship to youth (3 levels)
R06_PT0041_NB R06_PT0253_NB	Parent	R06R_Y_PT0041_NB	Age range youth was first told he/she had high blood pressure (3 levels)
R06_PT0043_NB R06_PT0254_NB	Parent	R06R_Y_PT0043_NB	Age range youth was first told he/she has high cholesterol (3 levels)
R06_PT0038_NB R06_PT0260_NB	Parent	R06R_Y_PT0038_NB	Age range youth was first told he/she has asthma (3 levels)
R06_PT0042_NB R06_PT0263_NB	Parent	R06R_Y_PT0042_NB	Age range youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (3 levels)
R06_PM0069	Parent	R06R_Y_PM0069	Recoded parent respondent is a citizen of the United States (2 levels)
R06_PM0065 R06_PM0065_NN	Parent	R06R_Y_PM0065_V2	Recoded number of years parents lived in the United States (2 levels)
R06_PM0001 R06_PM0118	Parent	R06R_P_PARSP_EDUC	Recoded highest grade or year of school completed by parent/spouse/guardian (6 levels)
R06_PM0130	Parent	R06R_Y_PM0130	Recoded total household income in past 12 months (5 levels)
R06_PL0040	Parent	R06R_Y_PL0040	Recoded home is owned or rented (2 levels)

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_PM0072_02 R06_PM0072_03 R06_PM0072_04 R06_PM0072_05 R06_PM0072_06 R06_PM0072_07 R06_PM0072_08 R06_PM0072_09 R06_PM0072_10	Parent	R06R_Y_PM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (parent respondent)
R06_PX0757_NB R06_PX0758_NB	Parent	R06R_Y_PX0757_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had schizophrenia, schizoaffective disorder or psychosis (3 levels)
R06_PX0762_NB R06_PX0772_NB	Parent	R06R_Y_PX0762_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had a psychotic illness or episode (3 levels)
R06_YM0065	Youth	R06R_Y_YM0065_V2	Recoded number of years youth lived in the United States (2 levels)
R06_YC1006 R06_YC1120	Youth	R06R_Y_YC1006	Age range when first tried cigarette smoking, even one or two puffs (3 levels)
R06_YC1007 R06_YC1121	Youth	R06R_Y_YC1007	Age range when first started smoking cigarettes fairly regularly (3 levels)
R06_YV1006 R06_YV1120	Youth	R06R_Y_YV1006	Age range when first tried an electronic nicotine product, even one or two times (3 levels)
R06_YV1007 R06_YV1121	Youth	R06R_Y_YV1007	Age range when first started using an electronic nicotine product fairly regularly (3 levels)
R06_YG1006TC R06_YG1120TC	Youth	R06R_Y_YG1006TC	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
R06_YG1007TC R06_YG1121TC	Youth	R06R_Y_YG1007TC	Age range when first started smoking traditional cigars fairly regularly (3 levels)
R06_YG1006CL R06_YG1120CL	Youth	R06R_Y_YG1006CL	Age range when first tried a cigarillo, even one or two puffs (3 levels)
R06_YG1007CL R06_YG1121CL	Youth	R06R_Y_YG1007CL	Age range when first started smoking cigarillos fairly regularly (3 levels)
R06_YG1006FC R06_YG1120FC	Youth	R06R_Y_YG1006FC	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
R06_YG1007FC R06_YG1121FC	Youth	R06R_Y_YG1007FC	Age range when first started smoking filtered cigars fairly regularly (3 levels)
R06_YH1006 R06_YH1120	Youth	R06R_Y_YH1006	Age range when first tried smoking hookah, even one or two puffs (3 levels)

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_YH1007 R06_YH1121	Youth	R06R_Y_YH1007	Age range when first started smoking a hookah fairly regularly (3 levels)
R06_YU1006 R06_YU1120	Youth	R06R_Y_YU1006	Age range when first tried snus pouches, even one or two times (3 levels)
R06_YU1007 R06_YU1121	Youth	R06R_Y_YU1007	Age range when first started using snus pouches fairly regularly (3 levels)
R06_YS1006 R06_YS1120	Youth	R06R_Y_YS1006	Age range when first tried smokeless tobacco, even one or two times (3 levels)
R06_YS1007 R06_YS1121	Youth	R06R_Y_YS1007	Age range when first started using smokeless tobacco fairly regularly (3 levels)
R06_YT0038 R06_YT0260	Youth	R06R_Y_YT0038	Age range when you were first told you had asthma (3 levels)
R06_YX0671_01 R06_YX0671_02 R06_YX0671_03 R06_YX0671_04 R06_YX0671_05 R06_YX0671_06 R06_YX0671_07 R06_YX0671_08 R06_YX0671_09	Youth	R06R_Y_YX0671_V2	Recoded anyone who lives with you now uses tobacco (4 levels)
R06_YX0086 R06_YX0087	Youth	R06R_Y_YX0086	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)
R06_YX0074 R06_YX0270	Youth	R06R_Y_YX0074	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
R06_YX0079 R06_YX0271	Youth	R06R_Y_YX0079	Age range when first used marijuana, hash, THC or grass (3 levels)
R06_YX0082_01 R06_YX0272_01	Youth	R06R_Y_YX0082_01	Age range when first used: Ritalin or Adderall (3 levels)
R06_YX0082_02 R06_YX0272_02	Youth	R06R_Y_YX0082_02	Age range when first used: Painkillers (3 levels)
R06_YX0082_03 R06_YX0272_03	Youth	R06R_Y_YX0082_03	Age range when first used: Sedatives or tranquilizers (3 levels)
R06_YX0082_04 R06_YX0272_04	Youth	R06R_Y_YX0082_04	Age range when first used: Cocaine or crack (3 levels)
R06_YX0082_05 R06_YX0272_05	Youth	R06R_Y_YX0082_05	Age range when first used: Methamphetamine or speed (3 levels)
R06_YX0082_06 R06_YX0272_06	Youth	R06R_Y_YX0082_06	Age range when first used: Heroin (3 levels)
R06_YX0082_07 R06_YX0272_07	Youth	R06R_Y_YX0082_07	Age range when first used: Inhalants or solvents (3 levels)
R06_YX0082_08 R06_YX0272_08	Youth	R06R_Y_YX0082_08	Age range when first used: Hallucinogens (3 levels)

**Table E-18. Wave 6 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R06_YM0019 R06_YM0018	Youth	R06R_Y_YM0018_V2	Recoded grade level (If on holiday or break – grade level entering when returning to school) (5 levels)
R06_YM0004_NB	Youth	R06R_Y_SEX	Gender
R06_YM0005_NB_01 R06_YM0005_NB_02 R06_YM0005_NB_03 R06_YM0005_NB_04 R06_YM0005_NB_05	Youth	R06R_Y_HISP	DERIVED – Hispanic origin from the interview (2 levels)
R06_YM0072_02 R06_YM0072_03 R06_YM0072_04 R06_YM0072_05 R06_YM0072_06 R06_YM0072_07 R06_YM0072_08 R06_YM0072_09 R06_YM0072_10	Youth	R06R_Y_YM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (youth)
R06_YM0006_NB_01 R06_YM0006_NB_02 R06_YM0006_NB_03 R06_YM0006_NB_04 R06_YM0006_NB_05 R06_YM0006_NB_06 R06_YM0006_NB_07 R06_YM0006_NB_08 R06_YM0006_NB_09 R06_YM0006_NB_10 R06_YM0006_NB_11 R06_YM0006_NB_12 R06_YM0006_NB_13 R06_YM0006_NB_14	Youth	R06R_Y_RACECAT3	Recoded Race from the interview (3 levels)
R06_PT0007_FT R06_PT0007_IN R06_PT0007_MT R06_PT0008_LB R06_PT0008_KG R06_YX0310 R06_YT0007_FT R06_YT0007_IN R06_YX0311 R06_YT0008 R06_YX0312	Youth/Parent	R06R_Y_BMI	Body mass index

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_CAD10	Adult	Confirm DOB (for continuing respondents)
R07_LAD01	Adult	Respondent's preferred language to complete ACASI interview
R07_AM0007	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R07_AM0011_NB_01	Adult	Branch served when on active duty: Army
R07_AM0011_NB_02	Adult	Branch served when on active duty: Navy
R07_AM0011_NB_03	Adult	Branch served when on active duty: Air Force
R07_AM0011_NB_04	Adult	Branch served when on active duty: Marine Corps
R07_AM0011_NB_05	Adult	Branch served when on active duty: Coast Guard
R07_AM0011_NB_06	Adult	Branch served when on active duty: Space Force
R07_AM0069	Adult	Adult respondent is a citizen of the United States
R07_AM0072_OS	Adult	Other language spoken at home: Some other language - specify
R07_AQ9203_OS	Adult	In past 30 days, noticed IQOS or HeatSticks advertised: Somewhere else - specify
R07_AQ9033_OS	Adult	Where IQOS was bought: Somewhere else - specify
R07_AQ1033_OS	Adult	Retail location where your HeatSticks are purchased most of the time: Somewhere else - specify
R07_AA9448NP_OS	Adult	Nicotine pouch brand used in past 30 days: Something else - specify
R07_AA0106NP_OS	Adult	Nicotine pouch brand used today, yesterday, or the day before yesterday: Something else - specify
R07_AA9449OT_OS	Adult	Type of oral nicotine product used in past 30 days: Something else - specify
R07_AA0106OT_OS	Adult	Type of oral nicotine product used today, yesterday, or the day before yesterday: Something else - specify
R07_AZ1002_OS	Adult	Ever used any other tobacco products - specify
R07_AC1118MC_OS	Adult	How you usually get your own cigarettes: Some other way - specify
R07_AC1033MC_OS	Adult	Retail location where your cigarettes are purchased most of the time: Somewhere else - specify
R07_AC8800_OS	Adult	Where you usually throw away cigarette butts after you finish smoking a cigarette: Somewhere else - specify
R07_AC8801_OS	Adult	What you usually do with a cigarette pack or carton when it's empty: Something else - specify
R07_AC1049MC	Adult	Brand of cigarettes usually/last smoked - specify
R07_AC1071MC	Adult	Sub-brand of cigarette product usually/last smoked - specify
R07_AC1118RY_OS	Adult	How you usually get your own roll-your-own cigarette tobacco: Some other way - specify
R07_AC1033RY_OS	Adult	Retail location where your roll-your-own cigarette tobacco is purchased most of the time: Somewhere else - specify
R07_AC8802_OS	Adult	What you usually do with a pouch of roll-your-own cigarette tobacco when it's empty: Something else - specify
R07_AC1049RY	Adult	Brand of roll-your-own cigarette tobacco usually/last smoked - specify
R07_AC1071RY	Adult	Sub-brand of roll-your-own cigarette tobacco product usually/last smoked - specify
R07_AV1011_OS	Adult	Flavor of electronic nicotine product when first started using: Some other flavor - specify
R07_AV9035_OS	Adult	Electronic nicotine product you have used: Something else - specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_AV9044_OS	Adult	Electronic nicotine product used most often: Something else – specify
R07_AV9039_OS	Adult	Type of electronic nicotine product used: Something else – specify
R07_AV9001_OS	Adult	Type of electronic nicotine product used most often: Something else – specify
R07_AV1131_OS	Adult	In past 30 days, [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavor used: Some other flavor – specify
R07_AV1133_OS	Adult	In past 30 days, flavor of electronic nicotine product/electronic nicotine cartridge/e-liquid used most often: Some other flavor – specify
R07_AV1118_OS	Adult	How you usually get your own electronic nicotine products/pods or cartridges/e-liquid: Some other way – specify
R07_AV1033_OS	Adult	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
R07_AV8810_OS	Adult	What you usually do with your disposable electronic nicotine product when it is empty: Something else – specify
R07_AV8811_OS	Adult	What you usually do with a pod or cartridge when it is empty: Something else – specify
R07_AV8812_OS	Adult	What you usually do with the battery for your electronic nicotine product after it no longer works or is needed: Something else – specify
R07_AV8814_OS	Adult	What you usually do with the coils or atomizers for your electronic nicotine product after they no longer work: Something else – specify
R07_AV8816_OS	Adult	What you usually do with a bottle or container of e-liquid when it is empty: Something else – specify
R07_AV8815_OS	Adult	What you usually do with leftover or unused e-liquid: Something else – specify
R07_AV1012_OS	Adult	Flavor of electronic nicotine product brand regularly/last used: Some other flavor – specify
R07_AV1049	Adult	Brand of electronic nicotine products/pods or cartridges/e-liquid usually/last used – specify
R07_AV0335_OS	Adult	Main reason you stopped using electronic nicotine products: Some other reason – specify
R07_AJ1011TC_OS	Adult	Flavor of traditional cigar used as blunts when first started smoking: Some other flavor – specify
R07_AJ1131TC_OS	Adult	In past 30 days, traditional cigar as blunts flavor smoked: Some other flavor – specify
R07_AJ1133TC_OS	Adult	In past 30 days, flavor of traditional cigars smoked as blunts most often: Some other flavor – specify
R07_AJ1118TC_OS	Adult	How you usually get your own traditional cigars for when you smoke blunts: Some other way – specify
R07_AJ1033TC_OS	Adult	Retail location where your traditional cigars for blunts are purchased most of the time: Somewhere else – specify
R07_AJ1012TC_OS	Adult	Traditional cigar as blunts flavor usually/last smoked: Some other flavor – specify
R07_AJ1049TC	Adult	Brand of traditional cigars for blunts usually/last smoked – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_AJ1071TC	Adult	Sub-brand of traditional cigar as a blunt product usually/last smoked – specify
R07_AJ1011CG_OS	Adult	Flavor of cigarillo used as blunts when first started smoking: Some other flavor – specify
R07_AJ1131CG_OS	Adult	In past 30 days, cigarillo as blunts flavor smoked: Some other flavor – specify
R07_AJ1133CG_OS	Adult	In past 30 days, flavor of cigarillos smoked as blunts most often: Some other flavor – specify
R07_AJ1118CG_OS	Adult	How you usually get your own cigarillos for when you smoke blunts: Some other way – specify
R07_AJ1033CG_OS	Adult	Retail location where your cigarillos for blunts are purchased most of the time: Somewhere else – specify
R07_AJ1012CG_OS	Adult	Cigarillo as blunts flavor usually/last smoked: Some other flavor – specify
R07_AJ1049CG	Adult	Brand of cigarillos for blunts usually/last smoked – specify
R07_AJ1071CG	Adult	Sub-brand of cigarillo as a blunt product usually/last smoked – specify
R07_AJ1011FC_OS	Adult	Flavor of filtered cigar used as blunts when first started smoking: Some other flavor – specify
R07_AJ1131FC_OS	Adult	In past 30 days, filtered cigar as blunts flavor smoked: Some other flavor – specify
R07_AJ1133FC_OS	Adult	In past 30 days, flavor of filtered cigars smoked as blunts most often: Some other flavor – specify
R07_AJ1118FC_OS	Adult	How you usually get your own filtered cigars for when you smoke blunts: Some other way – specify
R07_AJ1033FC_OS	Adult	Retail location where your filtered cigars for blunts are purchased most of the time: Somewhere else – specify
R07_AJ1012FC_OS	Adult	Filtered cigar as blunts flavor usually/last smoked: Some other flavor – specify
R07_AJ1049FC	Adult	Brand of filtered cigars for blunts usually/last smoked – specify
R07_AJ1071FC	Adult	Sub-brand of filtered cigar as a blunt product usually/last smoked – specify
R07_AG1011TC_OS	Adult	Flavor of traditional cigar when first started smoking: Some other flavor – specify
R07_AG1131TC_OS	Adult	In past 30 days, traditional cigar flavor smoked: Some other flavor – specify
R07_AG1133TC_OS	Adult	In past 30 days, flavor of traditional cigar smoked most often: Some other flavor – specify
R07_AG1118TC_OS	Adult	How you usually get your own traditional cigars: Some other way – specify
R07_AG1033TC_OS	Adult	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
R07_AG1012TC_OS	Adult	Traditional cigar flavor usually/last smoked: Some other flavor – specify
R07_AG1049TC	Adult	Brand of traditional cigars usually/last smoked – specify
R07_AG1071TC	Adult	Sub-brand of traditional cigar product usually/last smoked – specify
R07_AG1011CG_OS	Adult	Flavor of cigarillo when first started smoking: Some other flavor – specify
R07_AG1131CG_OS	Adult	In past 30 days, cigarillo flavor smoked: Some other flavor – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_AG1133CG_OS	Adult	In past 30 days, flavor of cigarillo smoked most often: Some other flavor – specify
R07_AG1118CG_OS	Adult	How you usually get your own cigarillos: Some other way – specify
R07_AG1033CG_OS	Adult	Retail location where your cigarillos are purchased most of the time: Somewhere else – specify
R07_AG1012CG_OS	Adult	Cigarillo flavor usually/last smoked: Some other flavor – specify
R07_AG1049CG	Adult	Brand of cigarillos usually/last smoked – specify
R07_AG1071CG	Adult	Sub-brand of cigarillo product usually/last smoked – specify
R07_AG1011FC_OS	Adult	Flavor of filtered cigar when first started smoking: Some other flavor – specify
R07_AG1131FC_OS	Adult	In past 30 days, filtered cigar flavor smoked: Some other flavor – specify
R07_AG1133FC_OS	Adult	In past 30 days, flavor of filtered cigar smoked most often: Some other flavor – specify
R07_AG1118FC_OS	Adult	How you usually get your own filtered cigars: Some other way – specify
R07_AG1033FC_OS	Adult	Retail location where your filtered cigars are purchased most of the time: Somewhere else – specify
R07_AG1012FC_OS	Adult	Filtered cigar flavor usually/last smoked: Some other flavor – specify
R07_AG1049FC	Adult	Brand of filtered cigars usually/last smoked – specify
R07_AG1071FC	Adult	Sub-brand of filtered cigar product usually/last smoked – specify
R07_AP1011_OS	Adult	Flavor of pipe tobacco when first started smoking: Some other flavor – specify
R07_AP1131_OS	Adult	In past 30 days, pipe tobacco flavor smoked: Some other flavor – specify
R07_AP1133_OS	Adult	In past 30 days, flavor of pipe tobacco smoked most often: Some other flavor – specify
R07_AP1118_OS	Adult	How you usually get your own pipe tobacco: Some other way – specify
R07_AP1033_OS	Adult	Retail location where your pipe tobacco is purchased most of the time: Somewhere else – specify
R07_AP1012_OS	Adult	Pipe tobacco flavor usually/last smoked: Some other flavor – specify
R07_AP1049	Adult	Brand of pipe tobacco usually/last smoked – specify
R07_AP1071	Adult	Sub-brand of pipe tobacco product usually/last smoked – specify
R07_AH1011_OS	Adult	Flavor of hookah tobacco when first started smoking: Some other flavor – specify
R07_AH9045_OS	Adult	Location when first tried a hookah: Somewhere else – specify
R07_AH9011_OS	Adult	Place where usually smoke/smoked hookah: Somewhere else – specify
R07_AH1131_OS	Adult	In past 30 days, hookah tobacco flavor smoked: Some other flavor – specify
R07_AH1133_OS	Adult	In past 30 days, flavor of hookah tobacco smoked most often: Some other flavor – specify
R07_AH1118_OS	Adult	How you usually get your own hookah tobacco: Some other way – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_AH1033_OS	Adult	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify
R07_AH1012_OS	Adult	Hookah tobacco flavor usually/last smoked: Some other flavor – specify
R07_AH1049	Adult	Brand of shisha or hookah tobacco usually/last smoked – specify
R07_AH1071	Adult	Sub-brand of hookah tobacco product usually/last smoked – specify
R07_AU1011_OS	Adult	Flavor of snus when first started using: Some other flavor – specify
R07_AU1131_OS	Adult	In past 30 days, snus flavor used: Some other flavor – specify
R07_AU1133_OS	Adult	In past 30 days, flavor of snus used most often: Some other flavor – specify
R07_AU1118_OS	Adult	How you usually get your own snus: Some other way – specify
R07_AU1033_OS	Adult	Retail location where your snus is purchased most of the time: Somewhere else – specify
R07_AU1012_OS	Adult	Snus flavor usually/last used: Some other flavor – specify
R07_AU1049	Adult	Brand of snus usually/last used – specify
R07_AU1071	Adult	Sub-brand of snus product usually/last used – specify
R07_AS1011_OS	Adult	Flavor of smokeless tobacco when first started using: Some other flavor – specify
R07_AS1131_OS	Adult	In past 30 days, smokeless tobacco flavor used: Some other flavor – specify
R07_AS1133_OS	Adult	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor – specify
R07_AS1118_OS	Adult	How you usually get your own smokeless tobacco: Some other way – specify
R07_AS1033_OS	Adult	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else – specify
R07_AS1012_OS	Adult	Smokeless tobacco flavor usually/last used: Some other flavor – specify
R07_AS1049	Adult	Brand of smokeless tobacco usually/last used – specify
R07_AS1071	Adult	Sub-brand of smokeless tobacco product usually/last used – specify
R07_AQ1118_OS	Adult	How you usually get your own HeatSticks: Some other way – specify
R07_AQ8800_OS	Adult	Where you usually throw away HeatStick after it is finished: Somewhere else – specify
R07_AQ8801_OS	Adult	What you usually do with a HeatStick pack or carton when it's empty: Something else – specify
R07_AQ8812_OS	Adult	What you usually do with the IQOS holder after it no longer works or is not needed: Something else – specify
R07_AQ8813_OS	Adult	What you usually do with the IQOS pocket charger after it longer works or is not needed: Something else – specify
R07_AQ0335_OS	Adult	Main reason you stopped using IQOS: Some other reason – specify
R07_AN0349E_OS	Adult	Strength of nicotine patch last used (current or recent former established non-marijuana electronic nicotine product users): Some other strength – specify
R07_AN0263_OS	Adult	Flavor of e-liquid used during last quit attempt: Some other flavor – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_AN0264_OS	Adult	Flavor of e-liquid used most often during last quit attempt: Some other flavor – specify
R07_AN0265_OS	Adult	Type of e-cigarette or electronic nicotine product used most often during last quit attempt: Something else – specify
R07_AN0268	Adult	Brand of e-cigarette or electronic nicotine product used most often to help during last quit attempt – specify
R07_AN0336_OS	Adult	Main reason you used a nicotine patch, gum, inhaler, nasal spray, or lozenge: Some other reason – specify (Current or recent former non-electronic tobacco users)
R07_AN0349_OS	Adult	Strength of nicotine patch last used (current or recent former established non-electronic tobacco users): Some other strength – specify
R07_AV9070	Adult	Description of situation when you most recently swallowed e-liquid, or got it in your mouth, on your skin, or in your eyes
R07_AV9080	Adult	Description of situation when any child age 5 to 12 most recently swallowed e-liquid, or got it in their mouth, on their skin, or in their eyes
R07_AV9085	Adult	Description of situation when any child age 13 to 17 most recently swallowed e-liquid, or got it in their mouth, on their skin, or in their eyes
R07_AX0601_OS	Adult	In past 30 days, saw anyone using e-cigarettes or electronic nicotine products in or around your school: Yes, somewhere else at school not listed here – specify
R07_AX0602_OS	Adult	In past 30 days, saw anyone smoking a cigarette in or around your school: Yes, somewhere else at school not listed here – specify
R07_AX0313	Adult	Height and weight responses – unit preference
R07_AX0679_FT	Adult	Height without shoes: Feet
R07_AX0679_IN	Adult	Height without shoes: Inches
R07_AX0316	Adult	Height without shoes: Meters
R07_AX0109	Adult	Current weight: Pounds
R07_AX0312	Adult	Current weight: Kilograms
R07_AX0190_NB	Adult	Is deaf or has serious difficulty hearing
R07_AX0191_NB	Adult	Is blind or has serious difficulty seeing, even when wearing glasses
R07_AN0349H_OS	Adult	Strength of nicotine patch last used: Some other strength – specify (Long-term former tobacco users)
R07_AX0119_NB_OS	Adult	Doctor, nurse or other health professional said you had a lung or respiratory condition: Some other lung or respiratory condition – specify
R07_AX0119_12M_OS	Adult	In past 12 months, doctor, nurse or other health professional said you had a lung or respiratory condition: Some other lung or respiratory condition – specify
R07_AX0290_OS	Adult	In past 30 days, way in which you used marijuana: Use marijuana some other way – specify
R07_AX0416_OS	Adult	Social media used in the past 7 days: Something else – specify
R07_AX0417_OS	Adult	Social media used most often: Something else – specify
R07_AX0420_OS	Adult	In past 7 days, type of information about e-cigarettes or other electronic nicotine products seen on social media: Something else – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_AX0421_OS	Adult	In past 7 days, who posted content on social media that you've seen related to e-cigarettes or other electronic nicotine products: Other – specify
R07_AX0424_OS	Adult	Social media platform you most recently posted or shared content about e-cigarettes or other electronic nicotine products: Something else – specify
R07_AX0426_OS	Adult	Tobacco brand you currently follow or regularly search on social media: Some other brand – specify
R07_AX0427_OS	Adult	In past 7 days, saw any content posted on any social media related to: Some other type of product – specify
R07_AX0203_OS	Adult	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify
R07_AX0677_OS	Adult	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
R07_AM0067	Adult	Enrolled in High School
R07_AM0068	Adult	Current grade in school
R07_AM0020_OS	Adult	Type of degree program currently enrolled in: Other type of degree program – specify
R07_AM0021	Adult	Sexual attraction to gender
R07_AM0021_OS	Adult	Sexual attraction to gender: Something else – specify
R07_AM0063_OS	Adult	Sexual orientation: Something else – specify
R07_AM0062	Adult	Consider self to be male-to-female, female-to-male, or non-conforming
R07_LCAD01	Adult	Language in which CAPI portions of adult interview were conducted
R07_AX0214	Adult	OK to scan tobacco products
R07_AX0218	Adult	Number of tobacco products given by respondent to scan
R07_AX0217_1A	Adult	Wave 7 Adult tobacco Product 1, barcode scan/entry 1
R07_AX0233_01	Adult	First bar code scan: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_01	Adult	First bar code scan/entry: Same as brand you usually use
R07_AX0217_2A	Adult	Wave 7 Adult tobacco Product 2, barcode scan/entry 1
R07_AX0233_02	Adult	Second bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_02	Adult	Second bar code scan/entry: Same as brand you usually use
R07_AX0217_3A	Adult	Wave 7 Adult tobacco Product 3, barcode scan/entry 1
R07_AX0233_03	Adult	Third bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_03	Adult	Third bar code scan/entry: Same as brand you usually use
R07_AX0217_4A	Adult	Wave 7 Adult tobacco Product 4, barcode scan/entry 1
R07_AX0233_04	Adult	Fourth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_04	Adult	Fourth bar code scan/entry: Same as brand you usually use
R07_AX0217_5A	Adult	Wave 7 Adult tobacco Product 5, barcode scan/entry 1
R07_AX0233_05	Adult	Fifth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_05	Adult	Fifth bar code scan/entry: Same as brand you usually use
R07_AX0217_6A	Adult	Wave 7 Adult tobacco Product 6, barcode scan/entry 1
R07_AX0233_06	Adult	Sixth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_06	Adult	Sixth bar code scan/entry: Same as brand you usually use
R07_AX0217_7A	Adult	Wave 7 Adult tobacco Product 7, barcode scan/entry 1

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_AX0233_07	Adult	Seventh bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_07	Adult	Seventh bar code scan/entry: Same as brand you usually use
R07_AX0217_8A	Adult	Wave 7 Adult tobacco Product 8, barcode scan/entry 1
R07_AX0233_08	Adult	Eighth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_08	Adult	Eighth bar code scan/entry: Same as brand you usually use
R07_AX0217_9A	Adult	Wave 7 Adult tobacco Product 9, barcode scan/entry 1
R07_AX0233_09	Adult	Ninth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_09	Adult	Ninth bar code scan/entry: Same as brand you usually use
R07_AX0217_10A	Adult	Wave 7 Adult tobacco Product 10, barcode scan/entry 1
R07_AX0233_10	Adult	Tenth bar code scan/entry: Last time used this type of tobacco it was from this specific package or container
R07_AX0237_10	Adult	Tenth bar code scan/entry: Same as brand you usually use
R07_AX0232	Adult	Interviewer manually entered one or more barcode values
R07_AM0005_RS_01	Adult	Ethnic origin: Not of Hispanic, Latino or Spanish origin
R07_AM0005_RS_02	Adult	Ethnic origin: Mexican, Mexican American, Chicano or Chicana
R07_AM0005_RS_03	Adult	Ethnic origin: Puerto Rican
R07_AM0005_RS_04	Adult	Ethnic origin: Cuban
R07_AM0005_RS_05	Adult	Ethnic origin: Another Hispanic, Latino or Latina, or Spanish
R07_AM0006_RS_01	Adult	Race: White
R07_AM0006_RS_02	Adult	Race: Black or African American
R07_AM0006_RS_03	Adult	Race: American Indian or Alaska Native
R07_AM0006_RS_04	Adult	Race: Asian Indian
R07_AM0006_RS_05	Adult	Race: Chinese
R07_AM0006_RS_06	Adult	Race: Filipino
R07_AM0006_RS_07	Adult	Race: Japanese
R07_AM0006_RS_08	Adult	Race: Korean
R07_AM0006_RS_09	Adult	Race: Vietnamese
R07_AM0006_RS_10	Adult	Race: Other Asian
R07_AM0006_RS_11	Adult	Race: Native Hawaiian
R07_AM0006_RS_12	Adult	Race: Guamanian or Chamorro
R07_AM0006_RS_13	Adult	Race: Samoan
R07_AM0006_RS_14	Adult	Race: Other Pacific Islander
R07_AM0008_RS	Adult	Currently on active duty in the U.S. Armed Forces
R07_AM0050_RS	Adult	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard (Confirmation)
R07_AM0074_RS	Adult	Was on full-time active duty in the U.S. Armed Forces, Military Reserves or National Guard for the entire year in 2017
R07_AM0075_RS	Adult	Lived in the U.S. in any of the 50 states or the District of Columbia during 2017
R07_CPT05	Parent	Confirm DOB for shadow youth who is expected to be a youth at Wave 7
R07_CPT05C	Parent	Confirm DOB for youth from a prior wave who is expected to be a youth at Wave 7
R07_CPT07	Parent	Date of birth (corrected values for continuing youth/new values for new baseline youth)
R07_E_YOUTH	Youth	Emancipated youth

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_LCPT01	Parent	Language in which the CAPI portions of the interview was completed (interviewer report)
R07_LCYSO1	Youth	Language in which CAPI portions of youth interview were conducted
R07_LYH01	Youth	Youth's preferred language to complete ACASI interview
R07_PARENT_PERSONID	Parent	Wave 7 Parent/guardian Participant ID Number
R07_PM0016_NB	Parent	Youth was identified as a twin or part of a multiple birth
R07_PM0017_NB	Parent	Youth and sibling are identical twins
R07_PM0019_NB	Parent	Any sibling in multiple birth identical to youth
R07_PM0021_NB	Parent	First name of sibling that youth is a twin of
R07_PM0030_NB	Parent	First names of siblings in multiple birth that are identical to youth
R07_PM0035_NB	Parent	First names of siblings that are in the multiple birth with youth
R07_PM0052	Parent	Confirm parent's spouse is same as reported in previous wave(s)
R07_PM0053	Parent	Confirm parent's relationship to youth
R07_PM0058	Parent	Other parental figure/guardian list selection
R07_PM0058_NEW_AGE	Parent	Age of other parental figure/guardian
R07_PM0058_NEW_FNAME	Parent	First name of other parental figure or guardian – specify
R07_PM0058_NEW_LNAME	Parent	Last name of other parental figure or guardian – specify
R07_PM0058_PERSONID	Parent	Other parental figure/guardian PERSONID
R07_PM0059_OS_1	Parent	Other parental figure/guardian's relationship to sampled youth: Other relative – specify
R07_PM0059_OS_2	Parent	Other parental figure/guardian's relationship to sampled youth: Other non-relative – specify
R07_PM0061	Parent	Second other parental figure/guardian list selection
R07_PM0061_NEW_AGE	Parent	Age of second other parental figure/guardian
R07_PM0061_NEW_FNAME	Parent	First name of second other parental figure or guardian – specify
R07_PM0061_NEW_LNAME	Parent	Last name of second other parental figure or guardian – specify
R07_PM0061_PERSONID	Parent	Second other parental figure/guardian PERSONID
R07_PM0062_OS_1	Parent	Second other parental figure/guardian's relationship to youth: Other relative – specify
R07_PM0062_OS_2	Parent	Second other parental figure/guardian's relationship to youth: Other non-relative – specify
R07_PM0067	Parent	Parent respondent is currently enrolled in high school
R07_PM0068	Parent	Parent respondent grade (if currently enrolled in high school)
R07_PM0072_OS	Parent	Other language spoken at home: Some other language – specify (parent respondent)
R07_PT0001_OS_1	Parent	Parent or guardian relationship to youth: Other relative – specify
R07_PT0001_OS_2	Parent	Parent or guardian relationship to youth: Other non-relative – specify
R07_PT0002_OS_1	Parent	Parent's spouse or partner relationship to youth: Other relative – specify
R07_PT0002_OS_2	Parent	Parent's spouse or partner relationship to youth: Non-relative – specify
R07_PT0007_FT	Parent	Youth's current height (feet)
R07_PT0007_IN	Parent	Youth's current height (inches)
R07_PT0007_MT	Parent	Youth's current height (meters)

Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)

Questionnaire variable	Instrument	Questionnaire variable description
R07_PT0008_KG	Parent	Youth's current weight (kilograms)
R07_PT0008_LB	Parent	Youth's current weight (pounds)
R07_PT0046_NEW_AGE	Parent	Age of new spouse/partner
R07_PT0046_PERSONID	Parent	Parent's spouse/partner's PERSONID (if different from previous waves)
R07_PT0046_NEW_FNAME	Parent	First name of spouse/partner that lives with parent in DU – specify
R07_PT0046_NEW_LNAME	Parent	Last name of spouse/partner that lives with parent in DU – specify
R07_PX0186_NB	Parent	Youth has serious difficulty walking or climbing stairs
R07_PX0190_NB	Parent	Youth is deaf or has serious difficulty hearing
R07_PX0191_NB	Parent	Youth is blind or has serious difficulty seeing even when wearing glasses
R07_YA0106NP_OS	Youth	Nicotine pouch brand used today, yesterday, or the day before yesterday: Something else – specify
R07_YA0106OT_OS	Youth	Type of oral nicotine product used today, yesterday, or the day before yesterday: Something else – specify
R07_YA9448NP_OS	Youth	Nicotine pouch brand used in past 30 days: Something else – specify
R07_YA9449OT_OS	Youth	Type of oral nicotine product used in past 30 days: Something else – specify
R07_YC1033_OS	Youth	Retail location where your cigarettes are purchased most of the time: Somewhere else – specify
R07_YC1049	Youth	Brand of cigarettes usually/last smoked – specify
R07_YC1071	Youth	Sub-brand of cigarette product usually/last smoked – specify
R07_YC1118_OS	Youth	In past 30 days, how you usually got your own cigarettes: Some other way – specify
R07_YG1011CL_OS	Youth	Flavor of first cigarillo smoked: Some other flavor – specify
R07_YG1011FC_OS	Youth	Flavor of first filtered cigar smoked: Some other flavor – specify
R07_YG1011TC_OS	Youth	Flavor of first traditional cigar smoked: Some other flavor – specify
R07_YG1033CL_OS	Youth	Retail location where your cigarillos are purchased most of the time: Somewhere else – specify
R07_YG1033FC_OS	Youth	Retail location where your filtered cigars are purchased most of the time: Somewhere else – specify
R07_YG1033TC_OS	Youth	Retail location where your traditional cigars are purchased most of the time: Somewhere else – specify
R07_YG1049CL	Youth	Brand of cigarillos usually/last smoked – specify
R07_YG1049FC	Youth	Brand of filtered cigars usually/last smoked – specify
R07_YG1049TC	Youth	Brand of traditional cigars usually/last smoked – specify
R07_YG1071CL	Youth	Sub-brand of cigarillo product usually/last smoked – specify
R07_YG1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked – specify
R07_YG1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked – specify
R07_YG1118CL_OS	Youth	In past 30 days, how you usually got your own cigarillos: Some other way – specify
R07_YG1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars: Some other way – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_YG1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars: Some other way – specify
R07_YG1131CL_OS	Youth	In past 30 days, smoked cigarillos flavored to taste like: Some other flavor – specify
R07_YG1131FC_OS	Youth	In past 30 days, smoked filtered cigars flavored to taste like: Some other flavor – specify
R07_YG1131TC_OS	Youth	In past 30 days, smoked traditional cigars flavored to taste like: Some other flavor – specify
R07_YG1133CL_OS	Youth	In past 30 days, flavor of cigarillos smoked most often: Some other flavor – specify
R07_YG1133FC_OS	Youth	In past 30 days, flavor of filtered cigars smoked most often: Some other flavor – specify
R07_YG1133TC_OS	Youth	In past 30 days, flavor of traditional cigars smoked most often: Some other flavor – specify
R07_YH1011_OS	Youth	Flavor of first shisha or hookah tobacco smoked: Some other flavor – specify
R07_YH1033_OS	Youth	Retail location where your shisha or hookah tobacco is purchased most of the time: Somewhere else – specify
R07_YH1049	Youth	Brand of hookah tobacco usually/last smoked – specify
R07_YH1071	Youth	Sub-brand of hookah tobacco product usually/last smoked – specify
R07_YH1073	Youth	Smoke hookah because: It is part of my cultural tradition
R07_YH1118_OS	Youth	In past 30 days, how you usually got your own shisha or hookah tobacco: Some other way – specify
R07_YH1131_OS	Youth	In past 30 days, smoked hookah tobacco flavored to taste like: Some other flavor – specify
R07_YH1133_OS	Youth	In past 30 days, flavor of hookah smoked most often: Some other flavor – specify
R07_YH9011_OS	Youth	Usually smoke a hookah: Somewhere else – specify
R07_YH9045_OS	Youth	Where you were when you first tried a hookah, even one or two puffs – specify
R07_YJ1011CG_OS	Youth	Flavor of first cigarillo smoked as a blunt: Some other flavor – specify
R07_YJ1011FC_OS	Youth	Flavor of first filtered cigar smoked as a blunt: Some other flavor – specify
R07_YJ1011TC_OS	Youth	Flavor of first traditional cigar smoked as a blunt: Some other flavor – specify
R07_YJ1033CG_OS	Youth	Retail location where your cigarillos for blunts are purchased most of the time: Somewhere else – specify
R07_YJ1033FC_OS	Youth	Retail location where your filtered cigars for blunts are purchased most of the time: Somewhere else – specify
R07_YJ1033TC_OS	Youth	Retail location where your traditional cigars for blunts are purchased most of the time: Somewhere else – specify
R07_YJ1049CG	Youth	Brand of cigarillos usually/last smoked as a blunt – specify
R07_YJ1049FC	Youth	Brand of filtered cigars usually/last smoked as a blunt – specify
R07_YJ1049TC	Youth	Brand of traditional cigars usually/last smoked as a blunt – specify
R07_YJ1071CG	Youth	Sub-brand of cigarillo product usually/last smoked as a blunt – specify
R07_YJ1071FC	Youth	Sub-brand of filtered cigar product usually/last smoked as a blunt – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_YJ1071TC	Youth	Sub-brand of traditional cigar product usually/last smoked as a blunt – specify
R07_YJ1118CG_OS	Youth	In past 30 days, how you usually got your own cigarillo as a blunt: Some other way – specify
R07_YJ1118FC_OS	Youth	In past 30 days, how you usually got your own filtered cigars as a blunt: Some other way – specify
R07_YJ1118TC_OS	Youth	In past 30 days, how you usually got your own traditional cigars as a blunt: Some other way – specify
R07_YJ1131CG_OS	Youth	In past 30 days, smoked cigarillos as blunts flavored to taste like: Some other flavor – specify
R07_YJ1131FC_OS	Youth	In past 30 days, smoked filtered cigars as blunts flavored to taste like: Some other flavor – specify
R07_YJ1131TC_OS	Youth	In past 30 days, smoked traditional cigars as blunts flavored to taste like: Some other flavor – specify
R07_YJ1133CG_OS	Youth	In past 30 days, flavor of cigarillos smoked as blunts most often: Some other flavor – specify
R07_YJ1133FC_OS	Youth	In past 30 days, flavor of filtered cigars smoked as blunts most often: Some other flavor – specify
R07_YJ1133TC_OS	Youth	In past 30 days, flavor of traditional cigars smoked as blunts most often: Some other flavor – specify
R07_YM0005_NB_01	Youth	Ethnic origin: Not of Hispanic, Latino/Latina or Spanish origin
R07_YM0005_NB_02	Youth	Ethnic origin: Mexican, Mexican American, Chicano or Chicana
R07_YM0005_NB_03	Youth	Ethnic origin: Puerto Rican
R07_YM0005_NB_04	Youth	Ethnic origin: Cuban
R07_YM0005_NB_05	Youth	Ethnic origin: Another Hispanic, Latino/Latina, or Spanish origin
R07_YM0006_NB_01	Youth	Race: White
R07_YM0006_NB_02	Youth	Race: Black or African American
R07_YM0006_NB_03	Youth	Race: American Indian or Alaska Native
R07_YM0006_NB_04	Youth	Race: Asian Indian
R07_YM0006_NB_05	Youth	Race: Chinese
R07_YM0006_NB_06	Youth	Race: Filipino
R07_YM0006_NB_07	Youth	Race: Japanese
R07_YM0006_NB_08	Youth	Race: Korean
R07_YM0006_NB_09	Youth	Race: Vietnamese
R07_YM0006_NB_10	Youth	Race: Other Asian
R07_YM0006_NB_11	Youth	Race: Native Hawaiian
R07_YM0006_NB_12	Youth	Race: Guamanian or Chamorro
R07_YM0006_NB_13	Youth	Race: Samoan
R07_YM0006_NB_14	Youth	Race: Other Pacific Islander
R07_YM0007	Youth	Ever served on active duty in the U.S. Armed Forces, Military Reserves or National Guard
R07_YM0011_01	Youth	Branch served when on active duty: Army
R07_YM0011_02	Youth	Branch served when on active duty: Navy
R07_YM0011_03	Youth	Branch served when on active duty: Air Force
R07_YM0011_04	Youth	Branch served when on active duty: Marine Corps
R07_YM0011_05	Youth	Branch served when on active duty: Coast Guard
R07_YM0011_06	Youth	Branch served when on active duty: Space Force
R07_YM0020	Youth	Last grade/year in school completed
R07_YM0021	Youth	Sexual attraction to gender

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_YM0021_OS	Youth	Sexual attraction to gender: Something else – specify
R07_YM0062	Youth	Transgender category
R07_YM0063	Youth	Sexual orientation
R07_YM0063_OS	Youth	Sexual orientation: Something else – specify
R07_YM0066	Youth	How well you speak English (youth)
R07_YM0069	Youth	Youth respondent is a citizen of the United States
R07_YM0070	Youth	How well you read English (youth)
R07_YM0072_OS	Youth	Other language spoken at home: Some other language – specify (youth respondent)
R07_YM0073	Youth	How well you write in English (youth)
R07_YN0129_OS	Youth	In past 12 months, tried to completely stop using: Other – specify
R07_YN0336_OS	Youth	Main reason you used a nicotine patch, nicotine gum, nicotine inhaler, nicotine nasal spray or lozenge (past year tobacco users): Some other reason – specify
R07_YN0349_OS	Youth	Strength of nicotine patch last used (past year tobacco users): Some other strength – specify
R07_YN0349H_OS	Youth	Strength of nicotine patch last used (past year tobacco non-users): Some other strength – specify
R07_YP1033_OS	Youth	Retail location where your pipe tobacco is purchased most of the time: Somewhere else – specify
R07_YP1118_OS	Youth	In past 30 days, how you usually got your own pipe tobacco: Some other way – specify
R07_YS1011_OS	Youth	Flavor of first smokeless tobacco product used: Some other flavor – specify
R07_YS1033_OS	Youth	Retail location where your smokeless tobacco is purchased most of the time: Somewhere else – specify
R07_YS1049	Youth	Brand of smokeless tobacco usually/last used – specify
R07_YS1071	Youth	Sub-brand of smokeless tobacco product usually/last used – specify
R07_YS1118_OS	Youth	In past 30 days, how you usually got your own smokeless tobacco: Some other way – specify
R07_YS1131_OS	Youth	In past 30 days, used smokeless tobacco flavored to taste like: Some other flavor – specify
R07_YS1133_OS	Youth	In past 30 days, flavor of smokeless tobacco used most often: Some other flavor – specify
R07_YT0007_FT	Youth	Current height: Feet
R07_YT0007_IN	Youth	Current height: Inches
R07_YT0008	Youth	Current weight: Pounds
R07_YQ9203_OS	Youth	In past 30 days, noticed IQOS or HeatSticks advertised: Somewhere else – specify
R07_YQ9033_OS	Youth	Where IQOS was bought: Somewhere else – specify
R07_YQ1033_OS	Youth	Retail location where your HeatSticks are purchased most of the time: Somewhere else – specify
R07_YQ1118_OS	Youth	In past 30 days, how you usually got your own HeatSticks: I got it some other way – specify
R07_YU1011_OS	Youth	Flavor of first snus used: Some other flavor – specify
R07_YU1033_OS	Youth	Retail location where your snus is purchased most of the time: Somewhere else – specify
R07_YU1049	Youth	Brand of snus usually/last used – specify
R07_YU1071	Youth	Sub-brand of snus usually/last used – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_YU1118_OS	Youth	In past 30 days, how you usually got your own snus: Some other way – specify
R07_YU1131_OS	Youth	In past 30 days, used snus flavored to taste like: Some other flavor – specify
R07_YU1133_OS	Youth	In past 30 days, flavor of snus used most often: Some other flavor – specify
R07_YV1011_OS	Youth	Flavor of first electronic nicotine product used: Some other flavor – specify
R07_YV1033_OS	Youth	Retail location where your [electronic nicotine products/electronic nicotine cartridges/e-liquid] are purchased most of the time: Somewhere else – specify
R07_YV1049	Youth	Brand of electronic nicotine products/pods or cartridges/e-liquid usually/last – specify
R07_YV1118_OS	Youth	In past 30 days, how you usually got your own [electronic nicotine products/pods or cartridges/e-liquid]: Some other way – specify
R07_YV1131_OS	Youth	In past 30 days, used [electronic nicotine products/electronic nicotine cartridges/e-liquid] flavored to taste like: Some other flavor – specify
R07_YV1133_OS	Youth	In past 30 days, flavor of electronic nicotine product/cartridges/e-liquid used most often: Some other flavor – specify
R07_YV9035_OS	Youth	Ever used some other electronic nicotine product – specify
R07_YV9044_OS	Youth	Type of electronic nicotine product used most often: Something else – specify
R07_YV9039_OS	Youth	Type of electronic nicotine product used: Something else – specify
R07_YX0136	Youth	Currently pregnant
R07_YX0137_NN	Youth	Number of weeks/months pregnant – Number
R07_YX0137_UN	Youth	Number of weeks/months pregnant – Unit
R07_YX0203_OS	Youth	In past 30 days, noticed e-cigarettes or other electronic nicotine products being advertised: Somewhere else – specify
R07_YX0290_OS	Youth	In past 30 days, way in which you used marijuana: Use marijuana some other way – specify
R07_YX0310	Youth	Height and weight responses – unit preference
R07_YX0311	Youth	Current height: Meters
R07_YX0312	Youth	Current weight: Kilograms
R07_YX0416_OS	Youth	Social media used in past 7 days: Something else – specify
R07_YX0417_OS	Youth	Social media used most often: Something else – specify
R07_YX0420_OS	Youth	In past 7 days, type of information about e-cigarettes or other electronic nicotine products seen on social media: Something else – specify
R07_YX0421_OS	Youth	In past 7 days, who posted content on social media that you've seen related to e-cigarettes or other electronic nicotine products: Other – specify
R07_YX0424_OS	Youth	Social media platform you most recently posted or shared content about e-cigarettes or other electronic nicotine products: Something else – specify
R07_YX0426_OS	Youth	Tobacco brand you currently follow or regularly search on social media: Some other brand – specify

**Table E-19. Wave 7 questionnaire variables entirely excluded from the PUFs (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>Questionnaire variable description</b>
R07_YX0427_OS	Youth	In past 7 days, saw any content posted on any social media related to: Some other type of product – specify
R07_YX0677_OS	Youth	In past 30 days, noticed cigarettes or other tobacco products being advertised: Somewhere else – specify
R07 YY0601_OS	Youth	First type of tobacco you tried: Other – specify
R07_YZ1002_OS	Youth	Ever used any other tobacco products – specify
R07_YX0601_OS	Youth	In past 30 days, saw anyone using e-cigarettes or electronic nicotine products in or around your school: Yes, somewhere else at school not listed here – specify
R07_YX0602_OS	Youth	In past 30 days, saw anyone smoking a cigarette in or around your school: Yes, somewhere else at school not listed here – specify

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_AM0001	Adult	R07R_A_AGECAT6	Wave 7 Adult Age (6 levels)
R07_AM0002			
R07_AM0003			
R07_AM0004_RS	Adult	R07R_A_SEX	Wave 7 Adult Gender (2 levels)
R07_AT0047	Adult	R07R_A_AT0047	Recoded marital status (3 levels)
R07_AM0065	Adult	R07R_A_AM0065_V2	Recoded number of years in US (2 levels)
R07_AC1006	Adult	R07R_A_AC1006	Age group when first smoked part or all of a cigarette (6 levels)
R07_AC1120			
R07_AC1007	Adult	R07R_A_AC1007	Age group when you first started smoking cigarettes fairly regularly (6 levels)
R07_AC1121			
R07_AC1020	Adult	R07R_A_AC1020	Age group when you first started smoking cigarettes every day (6 levels)
R07_AC1122			
R07_AV1006	Adult	R07R_A_AV1006	Age group when first used an electronic nicotine product, even one or two times (6 levels)
R07_AV1120			
R07_AV1007	Adult	R07R_A_AV1007	Age group when you first started using electronic nicotine products fairly regularly (6 levels)
R07_AV1121			
R07_AV1020	Adult	R07R_A_AV1020	Age group when you first started using electronic nicotine products every day (6 levels)
R07_AV1122			
R07_AG1006TC	Adult	R07R_A_AG1006TC	Age group when first smoked part or all of a traditional cigar, even one or two puffs (6 levels)
R07_AG1120TC			
R07_AG1007TC	Adult	R07R_A_AG1007TC	Age group when you first started smoking traditional cigars fairly regularly (6 levels)
R07_AG1121TC			
R07_AG1020TC	Adult	R07R_A_AG1020TC	Age group when you first started smoking traditional cigars every day (6 levels)
R07_AG1122TC			
R07_AG1006CG	Adult	R07R_A_AG1006CG	Age group when first smoked part or all of a cigarillo, even one or two puffs (6 levels)
R07_AG1120CG			
R07_AG1007CG	Adult	R07R_A_AG1007CG	Age group when you first started smoking cigarillos fairly regularly (6 levels)
R07_AG1121CG			
R07_AG1020CG	Adult	R07R_A_AG1020CG	Age group when you first started smoking cigarillos every day (6 levels)
R07_AG1122CG			
R07_AG1006FC	Adult	R07R_A_AG1006FC	Age group when first smoked part or all of a filtered cigar, even one or two puffs (6 levels)
R07_AG1120FC			
R07_AG1007FC	Adult	R07R_A_AG1007FC	Age group when you first started smoking filtered cigars fairly regularly (6 levels)
R07_AG1121FC			

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_AG1020FC R07_AG1122FC	Adult	R07R_A_AG1020FC	Age group when you first started smoking filtered cigars every day (6 levels)
R07_AP1006 R07_AP1120	Adult	R07R_A_AP1006	Age group when you first smoked a pipe filled with tobacco, even one or two puffs (6 levels)
R07_AP1007 R07_AP1121	Adult	R07R_A_AP1007	Age group when you first started smoking a pipe filled with tobacco fairly regularly (6 levels)
R07_AP1020 R07_AP1122	Adult	R07R_A_AP1020	Age group when you first started smoking a pipe filled with tobacco every day (6 levels)
R07_AH1006 R07_AH1120	Adult	R07R_A_AH1006	Age group when first smoked tobacco in a hookah, even one or two puffs (6 levels)
R07_AH1007 R07_AH1121	Adult	R07R_A_AH1007	Age group when you first started smoking hookah fairly regularly (6 levels)
R07_AH1020 R07_AH1122	Adult	R07R_A_AH1020	Age group when you first started smoking hookah every day (6 levels)
R07_AU1006 R07_AU1120	Adult	R07R_A_AU1006	Age group when first used snus, even one or two times (6 levels)
R07_AU1007 R07_AU1121	Adult	R07R_A_AU1007	Age group when you first started using snus pouches fairly regularly (6 levels)
R07_AU1020 R07_AU1122	Adult	R07R_A_AU1020	Age group when you first started using snus pouches every day (6 levels)
R07_AS1006 R07_AS1120	Adult	R07R_A_AS1006	Age group when you first used smokeless tobacco, even one or two times (6 levels)
R07_AS1007 R07_AS1121	Adult	R07R_A_AS1007	Age group when you first started using smokeless tobacco fairly regularly (6 levels)
R07_AS1020 R07_AS1122	Adult	R07R_A_AS1020	Age group when you first started using smokeless tobacco every day (6 levels)
R07_AQ1006 R07_AQ1120	Adult	R07R_A_AQ1006	Age group when you first used IQOS, even one or two times (6 levels)
R07_AQ1007 R07_AQ1121	Adult	R07R_A_AQ1007	Age group when you first started using IQOS fairly regularly (6 levels)
R07_AQ1020 R07_AQ1122	Adult	R07R_A_AQ1020	Age group when you first started using IQOS every day (6 levels)
R07_AX0114 R07_AX0253	Adult	R07R_A_AX0114	Age range you were in when you were first told you had high blood pressure (6 levels)
R07_AX0115 R07_AX0254	Adult	R07R_A_AX0115	Age range you were in when you were first told you had high cholesterol (6 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_AX0116 R07_AX0255	Adult	R07R_A_AX0116	Age range you were in when you were first told you had congestive heart failure (6 levels)
R07_AX0117 R07_AX0256	Adult	R07R_A_AX0117	Age range you were in when you were first told you had a stroke (6 levels)
R07_AX0112 R07_AX0252	Adult	R07R_A_AX0112	Age range you were in when you were first told you had a heart attack (6 levels)
R07_AX0113 R07_AX0249	Adult	R07R_A_AX0113	Age range you were in when you were first told you had atrial fibrillation or Afib (6 levels)
R07_AX0120 R07_AX0257	Adult	R07R_A_AX0120	Age range you were in when you were first told you had COPD (6 levels)
R07_AX0121 R07_AX0258	Adult	R07R_A_AX0121	Age range you were in when you were first told you had chronic bronchitis (6 levels)
R07_AX0123 R07_AX0259	Adult	R07R_A_AX0123	Age range you were in when you were first told you had emphysema (6 levels)
R07_AX0124 R07_AX0260	Adult	R07R_A_AX0124	Age range you were in when you were first told you had asthma (6 levels)
R07_AX0131 R07_AX0261	Adult	R07R_A_AX0131	Age range you were in when you were first told you had gum disease (6 levels)
R07_AX0133 R07_AX0262	Adult	R07R_A_AX0133	Age range you were in when you were first told you had pre-cancerous oral lesions (6 levels)
R07_AX0280 R07_AX0263	Adult	R07R_A_AX0280	Age range you were in when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes (6 levels)
R07_AX0143 R07_AX0264	Adult	R07R_A_AX0143	Age range you were in when you were first told you had an ulcer (6 levels)
R07_AX0148 R07_AX0266	Adult	R07R_A_AX0148	Age range you were in when you were first told you had stomach or gastro-intestinal bleeding (6 levels)
R07_AX0150 R07_AX0267	Adult	R07R_A_AX0150	Age range you were in when you were first told you had osteoporosis (6 levels)
R07_AX0198 R07_AX0268	Adult	R07R_A_AX0198	Age range you were in when you were first told you had a bone fracture because you have fragile bones (6 levels)
R07_AX0411 R07_AX0412	Adult	R07R_A_AX0411	Age group when you were first told you have poor circulation (PAD or PVD) (6 levels)
R07_AX0703 R07_AX0704	Adult	R07R_A_AX0703	Age group when first told you had macular degeneration (6 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_AX0289 R07_AX0291	Adult	R07R_A_AX0289	Age group when you had your first seizure (6 levels)
R07_AX0086 R07_AX0087	Adult	R07R_A_AX0086	Age group when you first drank alcohol at all, counting small tastes or sips (6 levels)
R07_AX0074 R07_AX0270	Adult	R07R_A_AX0074	Age group when you had your first alcoholic drink, other than small tastes or sips (6 levels)
R07_AX0079 R07_AX0271	Adult	R07R_A_AX0079	Age group when you first used marijuana (6 levels)
R07_AX0082_01 R07_AX0272_01	Adult	R07R_A_AX0082_01	Age group when first used Ritalin® or Adderall® (6 levels)
R07_AX0082_02 R07_AX0272_02	Adult	R07R_A_AX0082_02	Age group when first used painkillers (6 levels)
R07_AX0082_03 R07_AX0272_03	Adult	R07R_A_AX0082_03	Age group when first used sedatives or tranquilizers (6 levels)
R07_AX0082_04 R07_AX0272_04	Adult	R07R_A_AX0082_04	Age group when first used cocaine or crack (6 levels)
R07_AX0082_05 R07_AX0272_05	Adult	R07R_A_AX0082_05	Age group when first used stimulants like methamphetamine or speed (6 levels)
R07_AX0082_06 R07_AX0272_06	Adult	R07R_A_AX0082_06	Age group when first used heroin (6 levels)
R07_AX0082_07 R07_AX0272_07	Adult	R07R_A_AX0082_07	Age group when first used inhalants or solvents (6 levels)
R07_AX0082_08 R07_AX0272_08	Adult	R07R_A_AX0082_08	Age group when first used hallucinogens (6 levels)
R07_AX0145_02 R07_AX0145_03 R07_AX0145_04 R07_AX0145_05 R07_AX0145_09 R07_AX0145_12 R07_AX0145_15 R07_AX0145_18 R07_AX0145_19 R07_AX0145_21 R07_AX0145_16 R07_AX0145_23 R07_AX0145_24 R07_AX0145_25 R07_AX0145_27 R07_AX0145_29 R07_AX0145_30 R07_AX0145_31	Adult	R07R_A_AX0145_NONTOB	Type of cancer is a non-tobacco related cancer (2 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_AX0145_01 R07_AX0145_06 R07_AX0145_07 R07_AX0145_08 R07_AX0145_10 R07_AX0145_11 R07_AX0145_13 R07_AX0145_14 R07_AX0145_17 R07_AX0145_20 R07_AX0145_22 R07_AX0145_26 R07_AX0145_28	Adult	R07R_A_AX0145_TOB	Type of cancer is a tobacco related cancer (2 levels)
R07_AX0146_02 R07_AX0265_02 R07_AX0146_03 R07_AX0265_03 R07_AX0146_04 R07_AX0265_04 R07_AX0146_05 R07_AX0265_05 R07_AX0146_09 R07_AX0265_09 R07_AX0146_12 R07_AX0265_12 R07_AX0146_15 R07_AX0265_15 R07_AX0146_18 R07_AX0265_18 R07_AX0146_19 R07_AX0265_19 R07_AX0146_21 R07_AX0265_21 R07_AX0146_16 R07_AX0265_16 R07_AX0146_23 R07_AX0265_23 R07_AX0146_24 R07_AX0265_24 R07_AX0146_25 R07_AX0265_25 R07_AX0146_27 R07_AX0265_27 R07_AX0146_29 R07_AX0265_29 R07_AX0146_30 R07_AX0265_30 R07_AX0146_31 R07_AX0265_31	Adult	R07R_A_AX0146_NONTOB	Age range when first non tobacco-related cancer was diagnosed (6 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_AX0146_01	Adult	R07R_A_AX0146_TOB	Age range when first tobacco-related cancer was diagnosed (6 levels)
R07_AX0265_01			
R07_AX0146_06			
R07_AX0265_06			
R07_AX0146_07			
R07_AX0265_07			
R07_AX0146_08			
R07_AX0265_08			
R07_AX0146_10			
R07_AX0265_10			
R07_AX0146_11			
R07_AX0265_11			
R07_AX0146_13			
R07_AX0265_13			
R07_AX0146_14			
R07_AX0265_14			
R07_AX0146_17			
R07_AX0265_17			
R07_AX0146_20			
R07_AX0265_20			
R07_AX0146_22			
R07_AX0265_22			
R07_AX0146_26			
R07_AX0265_26			
R07_AX0146_28			
R07_AX0265_28			
R07_AX0300_12M_01	Adult	R07R_A_AX0300_12M	Indicator for adverse pregnancy outcomes resulting in risky birth (2 levels)
R07_AX0300_12M_02			
R07_AX0300_12M_03			
R07_AX0300_12M_04			
R07_AX0300_12M_05			
R07_AX0300_12M_06			
R07_AX0300_12M_07			
R07_AX0706			
R07_AX0135_12M	Adult	R07R_A_AX0135_12M	Outcome of last pregnancy was miscarriage, induced abortion, ectopic or tubal pregnancy or stillbirth (2 levels)
R07_AX0691	Adult	R07R_A_AX0691	Recoded type of living space currently living in (4 levels)
R07_AM0017	Adult	R07R_A_AM0017	Recoded reason for not working for pay (5 levels)
R07_AM0018	Adult	R07R_A_AM0018_V2	Recoded education level (5 levels)
R07_AM0030	Adult	R07R_A_AM0030	Recoded total household income in the past 12 months (5 levels)
R07_AM0033	Adult	R07R_A_AM0033_V3	Recoded education level of mother (4 levels)
R07_AM0034	Adult	R07R_A_AM0034_V3	Recoded education level of father (4 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_AM0036	Adult	R07R_A_AM0036	Recoded parents' total household income in the past 12 months (5 levels)
R07_AM0026_01 R07_AM0026_02 R07_AM0026_03 R07_AM0026_04 R07_AM0026_05 R07_AM0026_06 R07_AM0026_07 R07_AM0026_08	Adult	R07R_A_AM0026_V2	Indicator of health insurance (2 levels)
R07_AM0063	Adult	R07R_A_SEXORIENT2	Recoded sexual orientation (2 levels)
R07_AX0313 R07_AX0679_FT R07_AX0679_IN R07_AX0316 R07_AX0109 R07_AX0312	Adult	R07R_A_BMI	Body mass index
R07_AL0040	Adult	R07R_A_AL0040	Indicator of home ownership (2 levels)
R07_AM0042	Adult	R07R_A_AM0042_V2	Recoded where you currently live (2 levels)
R07_AM0072_02 R07_AM0072_03 R07_AM0072_04 R07_AM0072_05 R07_AM0072_06 R07_AM0072_07 R07_AM0072_08 R07_AM0072_09 R07_AM0072_10	Adult	R07R_A_AM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (adult)
R07_AX0066_01 R07_AX0066_02 R07_AX0066_03 R07_AX0066_04 R07_AX0066_05 R07_AX0066_06 R07_AX0066_07 R07_AX0066_08 R07_AX0066_09 R07_AX0066_10 R07_AX0066_11 R07_AX0066_12	Adult	R07R_A_AX0066_V3	Recoded anyone who lives with you now who uses tobacco (4 levels)
R07_AX0093	Adult	R07R_A_AX0093	Extent to which you are able to carry out your everyday physical activities (2 levels)
R07_AX0757 R07_AX0758	Adult	R07R_A_AX0757	Age range when you were first told you had schizophrenia, schizoaffective disorder or psychosis (6 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_AX0762 R07_AX0772	Adult	R07R_A_AX0762	Age range when you were first told you had a psychotic illness or episode (6 levels)
R07_AX0135_RS_01 R07_AX0135_RS_02 R07_AX0135_RS_03 R07_AX0135_RS_04 R07_AX0135_RS_05	Adult	R07R_A_AX0135_RS	Indicator for adverse pregnancy outcomes resulting in no birth (2 levels)
R07_AX0300_RS_01 R07_AX0300_RS_02 R07_AX0300_RS_03 R07_AX0300_RS_04 R07_AX0300_RS_05 R07_AX0300_RS_06 R07_AX0300_RS_07	Adult	R07R_A_AX0300_RS	Indicator for adverse pregnancy outcomes resulting in risky birth (2 levels)
R07_AX0308_RS	Adult	R07R_A_AX0308_RS	Indicator for any pregnancies resulting in a live birth (2 levels)
R07_PT0001	Parent	R07R_Y_PT0001_V2	Recoded parent or guardian relationship to youth (3 levels)
R07_PT0047	Parent	R07R_Y_PT0047	Recoded parent or guardian marital status (3 levels)
R07_PT0045 R07_PM0057 R07_PM0060	Parent	R07R_P_OTHPAR_INHH	Wave 7 Recoded youth has any parental figures/guardians other than the parent/guardian in the house (2 levels)
R07_PT0002	Parent	R07R_Y_PT0002_V2	Recoded parent's spouse or partner relationship to youth (3 levels)
R07_PM0059	Parent	R07R_Y_PM0059_V2	Recoded other parental figure/guardian's relationship to youth (3 levels)
R07_PM0062	Parent	R07R_Y_PM0062_V2	Recoded second other parental figure/guardian's relationship to youth (3 levels)
R07_PT0041_NB R07_PT0253_NB	Parent	R07R_Y_PT0041_NB	Age range youth was first told he/she had high blood pressure (3 levels)
R07_PT0043_NB R07_PT0254_NB	Parent	R07R_Y_PT0043_NB	Age range youth was first told he/she has high cholesterol (3 levels)
R07_PT0038_NB R07_PT0260_NB	Parent	R07R_Y_PT0038_NB	Age range youth was first told he/she has asthma (3 levels)
R07_PT0042_NB R07_PT0263_NB	Parent	R07R_Y_PT0042_NB	Age range youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes (3 levels)
R07_PM0069	Parent	R07R_Y_PM0069	Recoded parent respondent is a citizen of the United States (2 levels)
R07_PM0065 R07_PM0065_NN	Parent	R07R_Y_PM0065_V2	Recoded number of years parents lived in the United States (2 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_PM0001 R07_PM0118	Parent	R07R_P_PARSP_EDUC_V2	Recoded highest grade or year of school completed by parent/spouse/guardian (5 levels)
R07_PM0130	Parent	R07R_Y_PM0130	Recoded total household income in past 12 months (5 levels)
R07_PL0040	Parent	R07R_Y_PL0040	Recoded home is owned or rented (2 levels)
R07_PM1937_12M	Parent	R07R_Y_PM1937_12M	Recoded total number of days youth was hospitalized because of COVID-19 in past 12 months (7 levels)
R07_PM0072_02 R07_PM0072_03 R07_PM0072_04 R07_PM0072_05 R07_PM0072_06 R07_PM0072_07 R07_PM0072_08 R07_PM0072_09 R07_PM0072_10	Parent	R07R_Y_PM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (parent respondent)
R07_PX0289 R07_PX0291	Parent	R07R_Y_PX0289	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had schizophrenia, schizoaffective disorder or psychosis (3 levels)
R07_PX0757_NB R07_PX0758_NB	Parent	R07R_Y_PX0757_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had schizophrenia, schizoaffective disorder or psychosis (3 levels)
R07_PX0762_NB R07_PX0772_NB	Parent	R07R_Y_PX0762_NB	Age range when youth was first told by a doctor, therapist or other mental health professional that he/she had a psychotic illness or episode (3 levels)
R07_YM0065	Youth	R07R_Y_YM0065_V2	Recoded number of years youth lived in the United States (2 levels)
R07_YC1006 R07_YC1120	Youth	R07R_Y_YC1006	Age range when first tried cigarette smoking, even one or two puffs (3 levels)
R07_YC1007 R07_YC1121	Youth	R07R_Y_YC1007	Age range when first started smoking cigarettes fairly regularly (3 levels)
R07_YV1006 R07_YV1120	Youth	R07R_Y_YV1006	Age range when first tried an electronic nicotine product, even one or two times (3 levels)
R07_YV1007 R07_YV1121	Youth	R07R_Y_YV1007	Age range when first started using an electronic nicotine product fairly regularly (3 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_YG1006TC R07_YG1120TC	Youth	R07R_Y_YG1006TC	Age range when first tried a traditional cigar, even one or two puffs (3 levels)
R07_YG1007TC R07_YG1121TC	Youth	R07R_Y_YG1007TC	Age range when first started smoking traditional cigars fairly regularly (3 levels)
R07_YG1006CL R07_YG1120CL	Youth	R07R_Y_YG1006CL	Age range when first tried a cigarillo, even one or two puffs (3 levels)
R07_YG1007CL R07_YG1121CL	Youth	R07R_Y_YG1007CL	Age range when first started smoking cigarillos fairly regularly (3 levels)
R07_YG1006FC R07_YG1120FC	Youth	R07R_Y_YG1006FC	Age range when first tried a filtered cigar, even one or two puffs (3 levels)
R07_YG1007FC R07_YG1121FC	Youth	R07R_Y_YG1007FC	Age range when first started smoking filtered cigars fairly regularly (3 levels)
R07_YH1006 R07_YH1120	Youth	R07R_Y_YH1006	Age range when first tried smoking hookah, even one or two puffs (3 levels)
R07_YH1007 R07_YH1121	Youth	R07R_Y_YH1007	Age range when first started smoking a hookah fairly regularly (3 levels)
R07_YU1006 R07_YU1120	Youth	R07R_Y_YU1006	Age range when first tried snus pouches, even one or two times (3 levels)
R07_YU1007 R07_YU1121	Youth	R07R_Y_YU1007	Age range when first started using snus pouches fairly regularly (3 levels)
R07_YS1006 R07_YS1120	Youth	R07R_Y_YS1006	Age range when first tried smokeless tobacco, even one or two times (3 levels)
R07_YS1007 R07_YS1121	Youth	R07R_Y_YS1007	Age range when first started using smokeless tobacco fairly regularly (3 levels)
R07_YQ1006 R07_YQ1120	Youth	R07R_Y_YQ1006	Age range when first tried IQOS, even one or two times (3 levels)
R07_YQ1007 R07_YQ1121	Youth	R07R_Y_YQ1007	Age range when first started using IQOS fairly regularly (3 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_YX0671_01 R07_YX0671_02 R07_YX0671_03 R07_YX0671_04 R07_YX0671_05 R07_YX0671_06 R07_YX0671_07 R07_YX0671_08 R07_YX0671_09 R07_YX0671_10 R07_YX0671_11	Youth	R07R_Y_YX0671_V3	Recoded anyone who lives with you now uses tobacco (4 levels)
R07_YT0038 R07_YT0260	Youth	R07R_Y_YT0038	Age range when you were first told you had asthma (3 levels)
R07_YX0289 R07_YX0291	Youth	R07R_Y_YX0289	Age range when you had your first seizure (3 levels)
R07_YX0086 R07_YX0087	Youth	R07R_Y_YX0086	Age range when first drank alcohol at all, counting small tastes or sips (3 levels)
R07_YX0074 R07_YX0270	Youth	R07R_Y_YX0074	Age range when had first alcoholic drink, other than small tastes or sips (3 levels)
R07_YX0079 R07_YX0271	Youth	R07R_Y_YX0079	Age range when first used marijuana, hash, THC or grass (3 levels)
R07_YX0082_01 R07_YX0272_01	Youth	R07R_Y_YX0082_01	Age range when first used: Ritalin or Adderall (3 levels)
R07_YX0082_02 R07_YX0272_02	Youth	R07R_Y_YX0082_02	Age range when first used: Painkillers (3 levels)
R07_YX0082_03 R07_YX0272_03	Youth	R07R_Y_YX0082_03	Age range when first used: Sedatives or tranquilizers (3 levels)
R07_YX0082_04 R07_YX0272_04	Youth	R07R_Y_YX0082_04	Age range when first used: Cocaine or crack (3 levels)
R07_YX0082_05 R07_YX0272_05	Youth	R07R_Y_YX0082_05	Age range when first used: Methamphetamine or speed (3 levels)
R07_YX0082_06 R07_YX0272_06	Youth	R07R_Y_YX0082_06	Age range when first used: Heroin (3 levels)
R07_YX0082_07 R07_YX0272_07	Youth	R07R_Y_YX0082_07	Age range when first used: Inhalants or solvents (3 levels)
R07_YX0082_08 R07_YX0272_08	Youth	R07R_Y_YX0082_08	Age range when first used: Hallucinogens (3 levels)
R07_YM0019 R07_YM0018	Youth	R07R_Y_YM0018_V2	Recoded grade level (If on holiday or break – grade level entering when returning to school) (5 levels)
R07_YM1937_12M	Parent	R07R_Y_YM1937_12M	Recoded total number of days hospitalized because of COVID-19 in past 12 months (7 levels)
R07_YM1937	Parent	R07R_Y_YM1937	Recoded total number of days hospitalized because of COVID-19 (7 levels)

**Table E-20. Wave 7 questionnaire variables excluded from the PUFs that are replaced with a derived variable (continued)**

<b>Questionnaire variable</b>	<b>Instrument</b>	<b>PUF derived variable</b>	<b>PUF derived variable description</b>
R07_YM0004_NB	Youth	R07R_Y_SEX	New baseline youth respondent sex (2 levels)
R07_YM0005_NB_01 R07_YM0005_NB_02 R07_YM0005_NB_03 R07_YM0005_NB_04 R07_YM0005_NB_05	Youth	R07R_Y_HISP	DERIVED – Hispanic origin from the interview (2 levels)
R07_YM0006_NB_01 R07_YM0006_NB_02 R07_YM0006_NB_03 R07_YM0006_NB_04 R07_YM0006_NB_05 R07_YM0006_NB_06 R07_YM0006_NB_07 R07_YM0006_NB_08 R07_YM0006_NB_09 R07_YM0006_NB_10 R07_YM0006_NB_11 R07_YM0006_NB_12 R07_YM0006_NB_13 R07_YM0006_NB_14	Youth	R07R_Y_RACECAT3	Recoded Race from the interview (3 levels)
R07_pt0007_ft R07_PT0007_IN R07_PT0007_MT R07_PT0008_LB R07_PT0008_KG R07_YX0310 R07_YT0007_FT R07_YT0007_IN R07_YX0311 R07_YT0008 R07_YX0312	Youth/ Parent	R07r_y_bmi	Body mass index
R07_YM0072_02 R07_YM0072_03 R07_YM0072_04 R07_YM0072_05 R07_YM0072_06 R07_YM0072_07 R07_YM0072_08 R07_YM0072_09 R07_YM0072_10	Youth	R07R_Y_YM0072_02	Recoded other language(s) spoken at home: Chinese, Tagalog, Vietnamese, French, Korean, German, Arabic, Russian or some other language (youth)

## Appendix F

### Variables with Coded Outlier Values

---

Table F-1. Wave 1 questionnaire variables with coded outlier values

Variable name	Data file	Variable description
R01_AC1006	Adult	Age when first smoked part or all of a cigarette
R01_AC1007	Adult	Age when first started smoking cigarettes fairly regularly
R01_AC1020	Adult	Age when first started smoking cigarettes every day
R01_AC1009_NN	Adult	How long since you completely quit smoking cigarettes – Number
R01_AC9002_YR	Adult	How long smoking/smoked cigarettes fairly regularly – Years
R01_AC9002_MO	Adult	How long smoking/smoked cigarettes fairly regularly – Months
R01_AC9003_NN	Adult	How many cigarettes smoked per day when you smoked fairly regularly – Number
R01_AC1021_NN	Adult	Average number of cigarettes now smoked each day – Number
R01_AC1022	Adult	Number of days smoked cigarettes in past 30 days
R01_AC1023_NN	Adult	Average number of cigarettes smoked per day on days smoked in past 30 days – Number
R01_AC9005_NN	Adult	Average number of cigarettes smoked per day when smoked fairly regularly in the past – Number
R01_AC1024_NN	Adult	Time to first cigarette after waking – Number
R01_AC9006_NN	Adult	Average number of cigarettes smoked per day 12 months ago – Number
R01_AC1051RY_NN	Adult	How long been smoking regular brand of roll-your-own cigarette tobacco – Number
R01_AC1041_D	Adult	Amount usually paid for a carton of cigarettes – Dollars
R01_AC1041_C	Adult	Amount usually paid for a carton of cigarettes – Cents
R01_AC1042_D	Adult	Amount usually paid for a pack of cigarettes – Dollars
R01_AC1042_C	Adult	Amount usually paid for a pack of cigarettes – Cents
R01_AC1043_D	Adult	Amount usually paid for a single cigarette – Dollars
R01_AC1043_C	Adult	Amount usually paid for a single cigarette – Cents
R01_AC1051MC_NN	Adult	How long smoked regular brand of cigarettes – Number
R01_AE1006	Adult	Age when first time used an e-cigarette, even one or two times
R01_AE1009_NN	Adult	How long since you took last drag from an e-cigarette – Number
R01_AE9029_NN	Adult	Time since you took last puff from an e-cigarette – Number
R01_AE1024_NN	Adult	Time to first e-cigarette puff after waking – Number
R01_AE1051_NN	Adult	How long used regular brand of e-cigarettes – Number
R01_AG1006TC	Adult	Age when first smoked part or all of a traditional cigar, even one or two puffs
R01_AG1007TC	Adult	Age when first started smoking traditional cigars fairly regularly
R01_AG1009TC_NN	Adult	How long since you last smoked a traditional cigar – Number
R01_AG1022TC	Adult	Number of days smoked traditional cigars in past 30 days
R01_AG9029TC_NN	Adult	Time since you last smoked traditional cigars – Number
R01_AG1024TC_NN	Adult	Time to first traditional cigar after waking on days smoked – Number

**Table F-1. Wave 1 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R01_AG1098TC	Adult	Number of times used a coupon to buy traditional cigars in past 30 days
R01_AG1051TC_NN	Adult	How long smoked regular brand of traditional cigar – Number
R01_AG1006CG	Adult	Age when first time smoked part or all of a cigarillo, even one or two puffs
R01_AG1007CG	Adult	Age when first started smoking cigarillos fairly regularly
R01_AG1009CG_NN	Adult	How long since you last smoked a cigarillo – Number
R01_AG1022CG	Adult	Number of days smoked cigarillos in past 30 days
R01_AG9029CG_NN	Adult	Time since you last smoked cigarillos – Number
R01_AG1024CG_NN	Adult	Time to first cigarillo after waking on days smoked – Number
R01_AG1051CG_NN	Adult	How long smoked regular brand of cigarillo – Number
R01_AG1006FC	Adult	Age when first smoked part or all of a filtered cigar, even one or two puffs
R01_AG1007FC	Adult	Age when first started smoking filtered cigars fairly regularly
R01_AG1020FC	Adult	Age when first started smoking filtered cigars every day
R01_AG1009FC_NN	Adult	How long since you last smoked a filtered cigar – Number
R01_AG1022FC	Adult	Number of days smoked filtered cigars in past 30 days
R01_AG9029FC_NN	Adult	Time since you last smoked filtered cigars – Number
R01_AG1024FC_NN	Adult	Time to first filtered cigar after waking on days smoked – Number
R01_AG1051FC_NN	Adult	How long smoked regular brand of filtered cigar – Number
R01_AP1006	Adult	Age when first smoked part or all of a pipe filled with tobacco, even one or two puffs
R01_AP1007	Adult	Age when first started smoking a pipe filled with tobacco fairly regularly
R01_AP1009_NN	Adult	How long since you last smoked a pipe filled with tobacco – Number
R01_AP1022	Adult	Number of days smoked a pipe filled with tobacco in past 30 days
R01_AP9029_NN	Adult	Time since you last smoked a pipe filled with tobacco – Number
R01_AP1024_NN	Adult	Time to first pipe bowl after waking on days smoked – Number
R01_AH1006	Adult	Age when first smoked hookah, even one or two puffs
R01_AH1007	Adult	Age when first started smoking hookah fairly regularly
R01_AH1009_NN	Adult	How long since you last smoked a hookah – Number
R01_AH1024_NN	Adult	Time to first puff from hookah after waking on days smoked – Number
R01_AH1051_NN	Adult	How long used regular brand of hookah tobacco – Number
R01_AS1006SU	Adult	Age when first used snus pouches, even one or two times
R01_AS1009SU_NN	Adult	How long since you last used snus pouches – Number
R01_AS9029SU_NN	Adult	Time since you last used snus pouches – Number
R01_AS1024SU_NN	Adult	Time to first snus pouch after waking on days used – Number
R01_AS1051SU_NN	Adult	How long used regular brand of snus pouches – Number
R01_AS1006SM	Adult	Age when first used smokeless tobacco, even one or two times
R01_AS1007SM	Adult	Age when first started using smokeless tobacco fairly regularly
R01_AS1009SM_NN	Adult	How long since you last used smokeless tobacco – Number
R01_AS9029SM_NN	Adult	Time since you last used smokeless tobacco – Number
R01_AS1024SM_NN	Adult	Time to first use of smokeless tobacco on days used – Number
R01_AS1051SM_NN	Adult	How long used regular brand of smokeless tobacco – Number
R01_AD1006	Adult	Age when first used dissolvable tobacco, even one or two times
R01_AD1009_NN	Adult	How long since you last used dissolvable tobacco – Number

**Table F-1. Wave 1 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R01_AD9029_NN	Adult	Time since you last used dissolvable tobacco – Number
R01_AD1051_NN	Adult	How long used regular brand of dissolvable tobacco – Number
R01_AY0010_NN	Adult	Time to first use of any tobacco product after waking – Number
R01_AN0130_NN	Adult	Length of time you stopped smoking/using tobacco product(s) because you were trying to quit, in the past 12 months – Number
R01_AN0251_NN	Adult	How long been using nicotine medication to help quit – Number
R01_AN0309	Adult	Number of nicotine patches, gum, inhaler, nasal spray, lozenges or pills used today/yesterday/the day before yesterday
R01_AN0175_NN	Adult	How long ago did you stop using nicotine medication – Number
R01_AN0252_NN	Adult	How long have you been using a prescription drug to help quit – Number
R01_AN0205_NN	Adult	How long ago you stopped using prescription medication – Number
R01_AN0130E_NN	Adult	Length of time you stopped smoking/using tobacco product(s) because you were trying to quit, in the past 12 months – Number
R01_AX0068	Adult	Number of hours in past 7 days that you were in close contact with others when they were smoking
R01_AX0155	Adult	Number of times you visited an emergency room or urgent care center for a health problem of your own in past 12 months
R01_AX0114	Adult	Age when you were first told you had high blood pressure
R01_AX0115	Adult	Age when you were first told you had high cholesterol
R01_AX0116	Adult	Age when you were first told you had congestive heart failure
R01_AX0117	Adult	Age when you were first told you had a stroke
R01_AX0112	Adult	Age when you were first told you had a heart attack
R01_AX0120	Adult	Age when you were first told you had COPD
R01_AX0121	Adult	Age when you were first told you had chronic bronchitis
R01_AX0123	Adult	Age when you were first told you had emphysema
R01_AX0124	Adult	Age when you were first told you had asthma
R01_AX0131	Adult	Age when you were first told you had gum disease
R01_AX0133	Adult	Age when you were first told you had pre-cancerous oral lesions
R01_AX0280	Adult	Age when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes
R01_AX0143	Adult	Age when you were first told you had an ulcer
R01_AX0148	Adult	Age when you were first told you had stomach or gastro-intestinal bleeding
R01_AX0150	Adult	Age when you were first told you had osteoporosis
R01_AX0198	Adult	Age when you were first told you had a bone fracture because you have fragile bones
R01_AX0152	Adult	Age when you were first told you had a cataract or glaucoma
R01_AX0146_01	Adult	Age when bladder cancer was first diagnosed
R01_AX0146_04	Adult	Age when brain cancer was first diagnosed
R01_AX0146_06	Adult	Age when cervix (cervical) cancer was first diagnosed
R01_AX0146_07	Adult	Age when colon cancer was first diagnosed

**Table F-1. Wave 1 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R01_AX0146_09	Adult	Age when gallbladder cancer was first diagnosed
R01_AX0146_10	Adult	Age when kidney cancer was first diagnosed
R01_AX0146_16	Adult	Age when melanoma was first diagnosed
R01_AX0146_17	Adult	Age when mouth/tongue/lip cancer was first diagnosed
R01_AX0146_21	Adult	Age when prostate cancer was first diagnosed
R01_AX0146_24	Adult	Age when skin (unknown kind) cancer was first diagnosed
R01_AX0146_30	Adult	Age when uterus (uterine) cancer was first diagnosed
R01_AX0086	Adult	Age when first drank alcohol at all, counting small tastes or sips
R01_AX0074	Adult	Age when first alcoholic drink was consumed
R01_AX0079	Adult	Age when first used marijuana, hash, THC or grass
R01_AX0082_01	Adult	Age when first started using: Ritalin or Adderall
R01_AX0082_02	Adult	Age when first started using: Painkillers, sedatives or tranquilizers
R01_AX0082_03	Adult	Age when first started using: Cocaine or crack
R01_AX0082_04	Adult	Age when first started using: Stimulants like methamphetamine or speed
R01_AX0082_05	Adult	Age when first started using: Any other drugs like heroin, inhalants, solvents, or hallucinogens
R01_AX0137_NN	Adult	Weeks/Months pregnant - Number
R01_AX0309	Adult	Calendar year of most recent pregnancy
R01_AN0145_NN	Adult	Longest time period for which you stopped smoking/using tobacco product(s) because you were trying to quit, in the past 12 months - Number
R01_AN0145E_NN	Adult	Longest time period for which you stopped using e-cigarettes because you were trying to quit, in the past 12 months - Number
R01_AN0165_NN	Adult	Length of time used nicotine medication during last tobacco product(s) quit attempt - Number
R01_AN0165E_NN	Adult	Length of time used nicotine medication during last e-cigarette quit attempt - Number
R01_AN0195_NN	Adult	Length of time used prescription medication during last tobacco product(s) quit attempt - Number
R01_AN0195E_NN	Adult	Length of time used prescription medication during last e-cigarette quit attempt - Number
R01_AN0135_VALUE_INFO	Adult	Type of response provided for R01_AN0135
R01_AN0135E_VALUE_INFO	Adult	Type of response provided for R01_AN0135E
R01_AN0115	Adult	Number of times tried to quit smoking/using tobacco product(s) in past 12 months
R01_PT0007_FT	Youth	Youth's current height (feet)
R01_PT0007_IN	Youth	Youth's current height (inches)
R01_PT0008_LB	Youth	Youth's current weight (pounds)
R01_PT0043	Youth	Age youth was first told he/she has high cholesterol
R01_PT0038	Youth	Age youth was first told he/she has asthma
R01_PT0042	Youth	Age youth was first told by a doctor or other health professional that he/she has diabetes, sugar diabetes, high blood sugar or borderline diabetes
R01_YC1006	Youth	Age when first tried cigarette smoking, even one or two puffs
R01_YC1022	Youth	Number of days smoked cigarettes in past 30 days
R01_YC1051_NN	Youth	How long smoked regular brand of cigarettes - Number
R01 YE1006	Youth	Age when first tried an e-cigarette, even one or two times
R01 YE1022	Youth	Number of days used an e-cigarette in past 30 days

**Table F-1. Wave 1 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R01_YG1006TC	Youth	Age when first tried a traditional cigar, even one or two puffs
R01_YG1022TC	Youth	Number of days smoked a traditional cigar in past 30 days
R01_YG1006CL	Youth	Age when first tried a cigarillo, even one or two puffs
R01_YG1022CL	Youth	Number of days smoked a cigarillo in past 30 days
R01_YG1006FC	Youth	Age when first tried a filtered cigar, even one or two puffs
R01_YP1006	Youth	Age when first tried pipe tobacco, even one or two puffs
R01_YH1006	Youth	Age when first tried smoking a hookah, even one or two puffs
R01_YS1006SU	Youth	Age when first tried snus pouches, even one or two times
R01_YD1006	Youth	Age when first tried a dissolvable tobacco product, even one or two times
R01_YB1006BD	Youth	Age when first tried a bidi, even one or two puffs
R01_YB1006KK	Youth	Age when first tried a kretek, even one or two puffs
R01_YX0086	Youth	Age when first drank alcohol at all
R01_YX0074	Youth	Age when consumed first alcoholic drink
R01_YX0079	Youth	Age when first used marijuana, hash, THC or grass
R01_YX0082_01	Youth	Age when first used: Ritalin or Adderall
R01_YX0082_02	Youth	Age when first used: Painkillers, sedatives or tranquilizers
R01_YX0082_03	Youth	Age when first used: Cocaine or crack
R01_YX0082_04	Youth	Age when first used: Stimulants like methamphetamine or speed
R01_YX0082_05	Youth	Age when first used: Any other drugs like heroin, inhalants, solvents or hallucinogens

**Table F-2. Wave 2 questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R02_AC1009_NN	Adult	How long since last smoked a cigarette – Number
R02_A00103C	Adult	Number of puffs from e-cigarette today/yesterday/day before yesterday
R02_A00103H	Adult	Number of puffs from e-hookah today/yesterday/day before yesterday
R02_AE1009_NN	Adult	How long since you took last puff from an e-cigarette – Number
R02_AG1009TC_NN	Adult	How long since last smoked a traditional cigar – Number
R02_AG1009CG_NN	Adult	How long since last smoked a cigarillo – Number
R02_AG1009FC_NN	Adult	How long since last smoked a filtered cigar – Number
R02_AG1009TJ_NN	Adult	How long since last smoked a traditional cigar as a blunt – Number
R02_AG1009GJ_NN	Adult	How long since last smoked a cigarillo as a blunt – Number
R02_AG1009FJ_NN	Adult	How long since last smoked a filtered cigar as a blunt – Number
R02_AP1009_NN	Adult	How long since you last smoked a pipe filled with tobacco – Number
R02_AH1009_NN	Adult	How long since you last smoked a hookah – Number
R02_AS1009SU_NN	Adult	How long since you last used snus pouches – Number
R02_AS1009SM_NN	Adult	How long since you last used smokeless tobacco – Number
R02_AD1009_NN	Adult	How long since you last used dissolvable tobacco – Number
R02_AC1023_NN	Adult	In past 30 days, average number of cigarettes smoked per day on days smoked – Number
R02_AC1024_NN	Adult	Time to first cigarette after waking on days smoked – Number
R02_AC1051MC_NN	Adult	How long smoked regular brand of manufactured cigarettes – Number
R02_AE9029_NN	Adult	Time since last puff from an e-cigarette – Number
R02_AE1024_NN	Adult	Time to first e-cigarette puff after waking on days used – Number
R02_AE1051_NN	Adult	How long used regular brand of e-cigarette/e-cigarette cartridge/e-liquid – Number
R02_A09029_NN	Adult	Time since last puff from an electronic nicotine product (other than e-cigarettes) – Number
R02_AG9029TC_NN	Adult	Time since last smoked traditional cigars – Number
R02_AG1024TC_NN	Adult	Time to first traditional cigar after waking on days smoked – Number
R02_AG1051TC_NN	Adult	How long smoked regular brand of traditional cigar – Number
R02_AG9029CG_NN	Adult	Time since last smoked cigarillos – Number
R02_AG9029FC_NN	Adult	Time since last smoked filtered cigars – Number
R02_AG1024FC_NN	Adult	Time to first filtered cigar after waking on days smoked – Number
R02_AP9029_NN	Adult	Time since last smoked a pipe filled with tobacco – Number
R02_AH1024_NN	Adult	Time to first puff from hookah after waking on days smoked – Number
R02_AS9029SU_NN	Adult	Time since last used snus pouches – Number
R02_AS1024SU_NN	Adult	Time to first snus pouch after waking on days used – Number
R02_AS9029SM_NN	Adult	Time since last used smokeless tobacco – Number
R02_AS1024SM_NN	Adult	Time to first use of smokeless tobacco after waking on days used – Number
R02_AD9029_NN	Adult	Time since last used dissolvable tobacco – Number

**Table F-2. Wave 2 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R02_AY0010_NN	Adult	Time to first use of any tobacco product after waking – Number
R02_AN0130_NN	Adult	In past 12 months, length of time you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R02_AN0135_VALUE_INFO	Adult	Type of response provided for R02_AN0135
R02_AN0145_NN	Adult	In past 12 months, longest time period for which you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R02_AN0251_NN	Adult	How long been using nicotine medication to help quit – Number
R02_AN0175_NN	Adult	How long ago did you stop using nicotine medication – Number
R02_AN0195_NN	Adult	Length of time used prescription medication during last tobacco product(s) quit attempt – Number
R02_AN0252_NN	Adult	How long have you been using a prescription drug to help quit – Number
R02_AN0205_NN	Adult	How long ago you stopped using prescription medication – Number
R02_AN0130E_NN	Adult	In past 12 months, length of time you stopped using e-cigarettes because you were trying to quit – Number
R02_AN0135E_VALUE_INFO	Adult	Type of response provided for R02_AN0135E
R02_AN0145E_NN	Adult	In past 12 months, longest time period for which you stopped using e-cigarettes because you were trying to quit – Number
R02_AN0165E_NN	Adult	Length of time used nicotine medication during last e-cigarette quit attempt – Number
R02_AN0195E_NN	Adult	Length of time used prescription medication during last e-cigarette quit attempt – Number
R02_AX0068	Adult	In past 7 days, Number of hours that you were in close contact with others when they were smoking
R02_AX0679_FT	Adult	Height without shoes: Feet
R02_AX0679_IN	Adult	Height without shoes: Inches
R02_AX0316	Adult	Height without shoes: Meters
R02_AX0109	Adult	Current weight: Pounds
R02_AX0312	Adult	Current weight: Kilograms
R02_AX0242_NN	Adult	On days when you do physical activity or exercise of at least moderate intensity, how long do you do these activities – Number
R02_AX0137_NN	Adult	Weeks/Months pregnant – Number
R02_PT0007_FT	Youth	Youth's current height (feet)
R02_PT0007_IN	Youth	Youth's current height (inches)
R02_PT0007_MT	Youth	Youth's current height (meters)
R02_PT0008_LB	Youth	Youth's current weight (pounds)
R02_YE1007	Youth	Age when first started using e-cigarettes fairly regularly
R02_YE1041_D	Youth	Amount paid for last bottle or container of e-liquid – Dollars
R02_YE1041_C	Youth	Amount paid for last bottle or container of e-liquid – Cents
R02_YG1007TC	Youth	Age when first started smoking traditional cigars fairly regularly
R02_YG1034GJ_D	Youth	Amount paid when last time bought a box of cigarillos for blunts/pack of cigarillos for blunts/single cigarillo for blunts – Dollars

**Table F-2. Wave 2 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R02_YS1007SM	Youth	Age when first started using smokeless tobacco fairly regularly
R02_YT0007_FT	Youth	Current height: Feet
R02_YT0007_IN	Youth	Current height: Inches
R02_YX0311	Youth	Current height: Meters
R02_YT0008	Youth	Current weight: Pounds
R02_YX0312	Youth	Current weight: Kilograms
R02_YT0038_NB	Youth	Age when you were first told you had asthma
R02_YX0137_NN	Youth	Number of weeks/months pregnant – Number
R02_YX0086_NB	Youth	Age when first drank alcohol at all
R02_YX0074_NB	Youth	Age when consumed first alcoholic drink

**Table F-3. Wave 3 questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R03_AM0065	Adult	Number of years lived in the United States
R03_AC1009_NN	Adult	How long since last smoked a cigarette - Number
R03_AV1009EC_NN	Adult	How long since last took a puff from an e-cigarette - Number
R03_AV1009EG_NN	Adult	How long since last took a puff from an e-cigar - Number
R03_AV1009EP_NN	Adult	How long since last took a puff from an e-pipe - Number
R03_AV1009EH_NN	Adult	How long since last took a puff from an e-hookah - Number
R03_AV1009EN_NN	Adult	How long since last took a puff from an electronic nicotine product - Number
R03_AG1009TC_NN	Adult	How long since last smoked a traditional cigar - Number
R03_AG1009CG_NN	Adult	How long since last smoked a cigarillo - Number
R03_AG1009FC_NN	Adult	How long since last smoked a filtered cigar - Number
R03_AG1009TJ_NN	Adult	How long since last smoked a traditional cigar as a blunt - Number
R03_AG1009GJ_NN	Adult	How long since last smoked a cigarillo as a blunt - Number
R03_AG1009FJ_NN	Adult	How long since last smoked a filtered cigar as a blunt - Number
R03_AP1009_NN	Adult	How long since you last smoked a pipe filled with tobacco - Number
R03_AH1009_NN	Adult	How long since you last smoked a hookah - Number
R03_AU1009_NN	Adult	How long since you last used snus pouches - Number
R03_AS1009_NN	Adult	How long since you last used smokeless tobacco - Number
R03_AC1020_NB	Adult	Age when first started smoking cigarettes every day
R03_AC1024_NN	Adult	Time to first cigarette after waking on days smoked - Number
R03_AC1024_30D_NN	Adult	Time to first cigarette after waking on days smoked - Number
R03_AC0103MC	Adult	Number of cigarettes smoked today/yesterday/day before yesterday
R03_AC1040	Adult	Number of cigarette packs in the carton you usually buy
R03_AC1051MC_NN	Adult	How long smoked regular brand of manufactured cigarettes - Number
R03_AV1007_NB	Adult	Age when first started using [primary electronic nicotine product]s fairly regularly
R03_AV9045	Adult	Number of milliliters of e-liquid your tank system holds/held
R03_AM0103	Adult	How many puffs of marijuana from a [primary electronic nicotine product] have/did you taken/take
R03_AV9029_NN	Adult	Time since you took last puff from an [primary electronic nicotine product] - Number
R03_AV1024_NN	Adult	Time to first [primary electronic nicotine product] after waking - Number
R03_AV1024_30D_NN	Adult	Time to first [primary electronic nicotine product]s after waking - Number
R03_AV0103	Adult	Number of puffs taken from [primary electronic nicotine product] today/yesterday/day before yesterday
R03_AV1051_NN	Adult	How long used regular brand of [electronic nicotine products/cartridges/e-liquid] - Number
R03_AV2029_NN	Adult	How long since last took a puff from [secondary electronic nicotine product] - Number
R03_AJ9029TC_NN	Adult	Time since last smoked traditional cigar as blunt - Number
R03_AJ0103TC	Adult	Number of traditional cigars as a blunt smoked today/yesterday/day before yesterday
R03_AG9029TC_NN	Adult	Time since last smoked traditional cigars - Number

**Table F-3. Wave 3 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R03_AG1024TC_NN	Adult	Time to first traditional cigar after waking on days smoked – Number
R03_AG1024TC_30D_NN	Adult	Time to first traditional cigar after waking on days smoked – Number
R03_AG1040TC	Adult	Number of traditional cigars that come/came in box or pack usually buy/bought
R03_AG1007CG_NB	Adult	Age when first started smoking cigarillos fairly regularly
R03_AG9029CG_NN	Adult	Time since last smoked cigarillos – Number
R03_AG1024CG_NN	Adult	Time to first cigarillo after waking on days smoked – Number
R03_AG1024CG_30D_NN	Adult	Time to first cigarillo after waking on days smoked – Number
R03_AG1007FC_NB	Adult	Age when first started smoking filtered cigars fairly regularly
R03_AG9029FC_NN	Adult	Time since last smoked filtered cigars – Number
R03_AG1024FC_NN	Adult	Time to first filtered cigar after waking on days smoked – Number
R03_AG1024FC_30D_NN	Adult	Time to first filtered cigar after waking on days smoked – Number
R03_AP1024_NN	Adult	Time to first pipe bowl after waking on days smoked – Number
R03_AP1024_30D_NN	Adult	Time to first pipe bowl after waking on days smoked – Number
R03_AU1024_30D_NN	Adult	Time to first snus pouch of the day after waking on the days used – Number
R03_AU9002	Adult	Number of days to use up the container of snus
R03_AS1007_NB	Adult	Age when first started using smokeless tobacco fairly regularly
R03_AS9029_NN	Adult	Time since last used smokeless tobacco – Number
R03_AS1024_NN	Adult	Time to first use of smokeless tobacco after waking on days used – Number
R03_AY0010_NN	Adult	Time to first use of any tobacco product after waking – Number
R03_AN0130E_NN	Adult	In past 12 months, length of time you stopped using [e-cigarettes/electronic nicotine products] because you were trying to quit – Number
R03_AN0135E_VALUE_INFO	Adult	Type of response provided for R03_AN0135E
R03_AN0145E_NN	Adult	In past 12 months, longest time period for which you stopped using [e-cigarettes/electronic nicotine products] because you were trying to quit – Number
R03_AN0130_NN	Adult	In past 12 months, length of time you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R03_AN0135_VALUE_INFO	Adult	Type of response provided for R03_AN0135
R03_AN0145_NN	Adult	In past 12 months, longest time period for which you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R03_AN0251_NN	Adult	How long been using NRT – Number
R03_AN0205_NN	Adult	How long ago you stopped using prescription medication – Number
R03_AX0068	Adult	In past 7 days, number of hours that you were in close contact with others when they were smoking
R03_AX0679_FT	Adult	Height without shoes: Feet
R03_AX0679_IN	Adult	Height without shoes: Inches
R03_AX0316	Adult	Height without shoes: Meters
R03_AX0109	Adult	Current weight: Pounds
R03_AX0312	Adult	Current weight: Kilograms

**Table F-3. Wave 3 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R03_AX0242_NN	Adult	On days when you do physical activity or exercise of at least moderate intensity, how long do you do these activities – Number
R03_AX0155	Adult	In past 12 months, number of times you visited an emergency room or urgent care center for a health problem of your own
R03_AX0117_NB	Adult	Age when you were first told you had a stroke
R03_AX0112_NB	Adult	Age when you were first told you had a heart attack or needed bypass surgery
R03_AX0695	Adult	Number of permanent teeth removed because of tooth decay or gum disease
R03_AX0729_NN	Adult	Time since you had your teeth cleaned by a dentist, hygienist, or other health professional – Number
R03_AX0143_NB	Adult	Age when you were first told you had an ulcer
R03_AX0148_NB	Adult	Age when you were first told you had stomach or gastro-intestinal bleeding
R03_AX0150_NB	Adult	Age when you were first told you had osteoporosis
R03_AX0198_NB	Adult	Age when you were first told you had a bone fracture because you have fragile bones
R03_AX0703	Adult	Age when first told you had macular degeneration
R03_AX0137_NN	Adult	Weeks/Months pregnant – Number
R03_PT0007_MT	Youth	Youth's current height (meters)
R03_PT0008_LB	Youth	Youth's current weight (pounds)
R03_PM0065	Youth	Number of years lived in the United States
R03_YM0065	Youth	Number of years lived in the United States
R03_YV9045	Youth	Milliliters of e-liquid tank holds
R03_YV9040	Youth	In past 30 days, number of [electronic nicotine products/cartridges/milliliters of e-liquid] used per day on days used [primary electronic nicotine product]
R03_YV2029_NN	Youth	How long since you took a puff from [secondary electronic nicotine product] – Number
R03_YC1220	Youth	In past 12 months, number of times you stopped smoking for one day or longer because you were trying to quit smoking cigarettes for good
R03_YX0068	Youth	In past seven days, number of hours that you were in close contact with others when they were smoking
R03_YT0007_FT	Youth	Current height: Feet
R03_YT0007_IN	Youth	Current height: Inches
R03_YX0311	Youth	Current height: Meters
R03_YT0008	Youth	Current weight: Pounds
R03_YX0312	Youth	Current weight: Kilograms

**Table F-4. Wave 4 questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R04_AM0065	Adult	Number of years lived in the United States
R04_AC1009_NN	Adult	How long since last smoked a cigarette – Number
R04_AV1009_NN	Adult	How long since you last took a puff from an electronic nicotine product – Number
R04_AG1009TC_NN	Adult	How long since last smoked a traditional cigar – Number
R04_AG1009CG_NN	Adult	How long since last smoked a cigarillo – Number
R04_AG1009FC_NN	Adult	How long since last smoked a filtered cigar – Number
R04_AG1009TJ_NN	Adult	How long since last smoked a traditional cigar as a blunt – Number
R04_AG1009GJ_NN	Adult	How long since last smoked a cigarillo as a blunt – Number
R04_AG1009FJ_NN	Adult	How long since last smoked a filtered cigar as a blunt – Number
R04_AP1009_NN	Adult	How long since you last smoked a pipe filled with tobacco – Number
R04_AH1009_NN	Adult	How long since you last smoked tobacco in a hookah – Number
R04_AU1009_NN	Adult	How long since you last used snus – Number
R04_AS1009_NN	Adult	How long since you last used smokeless tobacco – Number
R04_AC1006_RS	Adult	Age when first smoked part or all of a cigarette
R04_AC1007_NB	Adult	Age when first started smoking cigarettes fairly regularly
R04_AC9002_NN	Adult	How long smoking/smoked cigarettes fairly regularly – Number
R04_AC9029_NN	Adult	How long since you last smoked a cigarette – Number (Some day or current experimental smokers)
R04_AC1024_NN	Adult	Time to first cigarette after waking on days smoked – Number (current smokers)
R04_AC1024_30D_NN	Adult	Time to first cigarette after waking on days smoked – Number (non-current smokers)
R04_AC1051MC_NN	Adult	How long smoked regular brand of manufactured cigarettes – Number
R04_AV1006_RS	Adult	Age when first used an electronic nicotine product, even one or two times
R04_AV1007_NB	Adult	Age when first started using electronic nicotine products fairly regularly
R04_AV1020_NB	Adult	Age when first started using electronic nicotine products every day
R04_AV9045	Adult	Number of milliliters of e-liquid your tank system holds/held
R04_AV9029_NN	Adult	Time since you took last puff from an electronic nicotine product – Number (Some day or current experimental users)
R04_AV1026	Adult	Average number of times you pick up your electronic nicotine product to take one or more puffs each day on days used
R04_AV1027	Adult	Number of puffs you take each time you pick up your electronic nicotine product to use it
R04_AV1024_NN	Adult	Time to first electronic nicotine product puff after waking – Number (current users)
R04_AV1024_30D_NN	Adult	Time to first electronic nicotine product puff after waking – Number (non-current users)
R04_AV1138_D	Adult	How much usually pay/paid for a box or pack of [electronic nicotine products/electronic nicotine cartridges/milliliters of e-liquid] – Dollars
R04_AV1138_C	Adult	How much usually pay/paid for a box or pack of [electronic nicotine products/electronic nicotine cartridges/milliliters of e-liquid] – Cents

**Table F-4. Wave 4 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R04_AV1051_NN	Adult	How long used regular brand of [electronic nicotine products/electronic nicotine cartridges/e-liquid] – Number
R04_AM0103	Adult	Number of puffs of marijuana taken from an electronic nicotine product today/yesterday/day before yesterday
R04_AJ9029TC_NN	Adult	Time since last smoked traditional cigar as blunt – Number (Some day or current experimental smokers)
R04_AJ1023TC	Adult	In past 30 days, average number of traditional cigars as blunts smoked per day on days smoked
R04_AJ9029CG_NN	Adult	Time since last smoked cigarillo as blunt – Number (Some day or current experimental smokers)
R04_AJ1023CG	Adult	In past 30 days, average number of cigarillos as blunts smoked per day on days smoked
R04_AJ9029FC_NN	Adult	Time since last smoked filtered cigar as blunt – Number (Some day or current experimental smokers)
R04_AG1006TC_RS	Adult	Age when first smoked part or all of a traditional cigar, even one or two puffs
R04_AG1007TC_NB	Adult	Age when first started smoking traditional cigars fairly regularly
R04_AG9029TC_NN	Adult	Time since last smoked traditional cigars – Number (Some day or current experimental smokers)
R04_AG1024TC_NN	Adult	Time to first traditional cigar after waking on days smoked – Number (current smokers)
R04_AG1024TC_30D_NN	Adult	Time to first traditional cigar after waking on days smoked – Number (non-current smokers)
R04_AG1006CG_RS	Adult	Age when first smoked part or all of a cigarillo, even one or two puffs
R04_AG1021CG	Adult	Average number of cigarillos now smoked each day
R04_AG9029CG_NN	Adult	Time since last smoked cigarillos – Number (Some day or current experimental smokers)
R04_AG1024CG_NN	Adult	Time to first cigarillo after waking on days smoked – Number (current smokers)
R04_AG1024CG_30D_NN	Adult	Time to first cigarillo after waking on days smoked – Number (non-current smokers)
R04_AG1040CG	Adult	Number of cigarillos that come/came in box or pack usually buy/bought
R04_AG1006FC_RS	Adult	Age when first smoked part or all of a filtered cigar, even one or two puffs
R04_AG1007FC_NB	Adult	Age when first started smoking filtered cigars fairly regularly
R04_AG9029FC_NN	Adult	Time since last smoked filtered cigars – Number (Some day or current experimental smokers)
R04_AG1024FC_NN	Adult	Time to first filtered cigar after waking on days smoked – Number (current smokers)
R04_AG1024FC_30D_NN	Adult	Time to first filtered cigar after waking on days smoked – Number (non-current smokers)
R04_AP1006_RS	Adult	Age when you first smoked a pipe filled with tobacco, even one or two puffs
R04_AP1007_NB	Adult	Age when first started smoking a pipe filled with tobacco fairly regularly
R04_AP9029_NN	Adult	Time since last smoked a pipe filled with tobacco – Number (Some day or current experimental smokers)

**Table F-4. Wave 4 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R04_AP1024_NN	Adult	Time to first pipe bowl after waking on days smoked – Number (current smokers)
R04_AP1024_30D_NN	Adult	Time to first pipe bowl after waking on days smoked – Number (non-current smokers)
R04_AP1051_NN	Adult	How long smoked regular brand of pipe tobacco – Number
R04_AH1006_RS	Adult	Age when first smoked tobacco in a hookah, even one or two puffs
R04_AU1006_RS	Adult	Age when first used snus, even one or two times
R04_AU9029_NN	Adult	How long since you last used snus – Number (Some day or current experimental users)
R04_AU1024_NN	Adult	Time to first snus product after waking on days used – Number (current users)
R04_AU1024_30D_NN	Adult	Time to first snus product of the day after waking on the days used – Number (non-current users)
R04_AS1006_RS	Adult	Age when you first used smokeless tobacco, even one or two times
R04_AS1021	Adult	Average number of times smokeless tobacco now used each day
R04_AS1022	Adult	In past 30 days, number of days used smokeless tobacco
R04_AS9029_NN	Adult	Time since last used smokeless tobacco – Number (Some day or current experimental users)
R04_AS1024_NN	Adult	Time to first use of smokeless tobacco after waking on days used – Number (Current users)
R04_AS1051_NN	Adult	How long used regular brand of smokeless tobacco – Number
R04_AY0010_NN	Adult	Time to first use of any tobacco product after waking – Number
R04_AN0115E	Adult	In past 12 months, number of times tried to quit using electronic nicotine products
R04_AN0130E_NN	Adult	In past 12 months, length of time you stopped using electronic nicotine products because you were trying to quit – Number
R04_AN0135E_VALUE_INFO	Adult	Type of response provided for R04_AN0135E
R04_AN0145E_NN	Adult	In past 12 months, longest time period for which you stopped using electronic nicotine products because you were trying to quit – Number
R04_AN0175E_NN	Adult	How long ago did you stop using NRT – Number (current established, recent former established or current experimental electronic nicotine product users)
R04_AN0115	Adult	In past 12 months, number of times tried to quit smoking/using tobacco product(s)
R04_AN0130_NN	Adult	In past 12 months, length of time you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R04_AN0135_VALUE_INFO	Adult	Type of response provided for R04_AN0135
R04_AN0145_NN	Adult	In past 12 months, longest time period for which you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R04_AN0251_NN	Adult	How long been using NRT – Number (past 12 month non-electronic tobacco quitters or quit attempters)
R04_AN0175_NN	Adult	How long ago did you stop using NRT – Number (current established or recent former established non-electronic tobacco users)
R04_AN0195_NN	Adult	Length of time used prescription medication during last tobacco product(s) quit attempt – Number

**Table F-4. Wave 4 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R04_AX0068	Adult	In past 7 days, number of hours that you were in close contact with others when they were smoking
R04_AX0679_FT	Adult	Height without shoes: Feet
R04_AX0679_IN	Adult	Height without shoes: Inches
R04_AX0316	Adult	Height without shoes: Meters
R04_AX0109	Adult	Current weight: Pounds
R04_AX0312	Adult	Current weight: Kilograms
R04_AX0242_NN	Adult	On days when you do physical activity or exercise of at least moderate intensity, how long do you do these activities – Number
R04_AX0114_NB	Adult	Age when you were first told you had high blood pressure
R04_AX0116_NB	Adult	Age when you were first told you had congestive heart failure
R04_AX0117_NB	Adult	Age when you were first told you had a stroke
R04_AX0112_NB	Adult	Age when you were first told you had a heart attack or needed bypass surgery
R04_AX0120_NB	Adult	Age when you were first told you had COPD
R04_AX0121_NB	Adult	Age when you were first told you had chronic bronchitis
R04_AX0123_NB	Adult	Age when you were first told you had emphysema
R04_AX0124_NB	Adult	Age when you were first told you had asthma
R04_AX0695_NB	Adult	Number of permanent teeth removed because of tooth decay or gum disease
R04_AX0695_12M	Adult	In past 12 months, number of permanent teeth removed because of tooth decay or gum disease
R04_AX0725	Adult	Number of times you brush your teeth in one day
R04_AX0696	Adult	In the last seven days, number of times used dental floss or any other device to clean between teeth
R04_AX0729_NN	Adult	Time since you had your teeth cleaned by a dentist, hygienist, or other health professional – Number
R04_AX0131_NB	Adult	Age when you were first told you had gum disease
R04_AX0133_NB	Adult	Age when you were first told you had pre-cancerous oral lesions
R04_AX0280_NB	Adult	Age when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes
R04_AX0143_NB	Adult	Age when you were first told you had an ulcer
R04_AX0150_NB	Adult	Age when you were first told you had osteoporosis
R04_AX0703	Adult	Age when first told you had macular degeneration
R04_AX0757	Adult	Age when you were first told by a doctor, therapist or other mental health professional that you had schizophrenia, schizoaffective disorder or psychosis
R04_AX0762	Adult	Age when you were first told by a doctor, therapist or other mental health professional that you had a psychotic illness or episode
R04_AX0137_NN	Adult	Weeks/Months pregnant – Number
R04_AX0086_RS	Adult	Age when first drank alcohol at all, counting small tastes or sips
R04_AX0074_RS	Adult	Age when first alcoholic drink was consumed, other than small tastes or sips
R04_AX0079_RS	Adult	Age when first used marijuana, hash, THC, grass, pot or weed
R04_PM0058_NEW_AGE	Youth	Age of other parental figure/guardian
R04_PM0061_NEW_AGE	Youth	Age of second other parental figure/guardian
R04_PM0065_NN	Youth	Number of years lived in the United States (parent respondent)
R04_YM0065	Youth	Number of years lived in the United States (youth)

**Table F-4. Wave 4 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R04_YC1006_NB	Youth	Age when first tried cigarette smoking, even one or two puffs
R04_YV1006_NB	Youth	Age when you first tried an electronic nicotine product, even one or two times
R04_YV1007	Youth	Age when first started using electronic nicotine products fairly regularly
R04_YV0103	Youth	Number of puffs taken from an electronic nicotine product today/yesterday/day before yesterday
R04_YV9040	Youth	In past 30 days, number of [electronic nicotine products/nicotine cartridges/milliliters of e-liquid] used per day on days used an electronic nicotine product
R04_YV1026	Youth	Average number of times you pick up your electronic nicotine product to use it for one or more puffs on days that you use
R04_YG1006TC_NB	Youth	Age when first tried a traditional cigar, even one or two puffs
R04_YG1007TC	Youth	Age when first started smoking traditional cigars fairly regularly
R04_YG1006FC_NB	Youth	Age when first tried a filtered cigar, even one or two puffs
R04_YH9006	Youth	Average number of times you smoke tobacco in a hookah each day
R04_YU1006_NB	Youth	Age when you first tried snus, even one or two times
R04_YS0103	Youth	Number of times used smokeless tobacco today/yesterday/day before yesterday
R04_YN0329	Youth	Number of Chantix, varenicline, Wellbutrin, Zyban or Bupropion taken today/yesterday/the day before yesterday (past 12 month tobacco users)
R04_YX0068	Youth	In past seven days, number of hours that you were in close contact with others when they were smoking
R04_YT0007_FT	Youth	Current height: Feet
R04_YT0007_IN	Youth	Current height: Inches
R04_YX0311	Youth	Current height: Meters
R04_YT0008	Youth	Current weight: Pounds
R04_YX0312	Youth	Current weight: Kilograms
R04_YN0309H	Youth	Number of nicotine patches, gum, inhaler, nasal spray, lozenges or pills used today/yesterday/the day before yesterday (past 12 month tobacco non-users)
R04_YN0329H	Youth	Number of Chantix, varenicline, Wellbutrin, Zyban or bupropion taken today/yesterday/the day before yesterday (past 12 month tobacco non-users)
R04_YX0086_NB	Youth	Age when first drank alcohol at all
R04_YX0074_NB	Youth	Age when consumed first alcoholic drink
R04_YX0079_NB	Youth	Age when first used marijuana, hash, THC or grass

**Table F-5. Wave 4.5 questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
X04_PT0007_MT	Youth	Youth's current height (meters)
X04_PT0008_KG	Youth	Youth's current weight (kilograms)
X04_YM0065	Youth	Number of years lived in the United States (youth)
X04_YV1007	Youth	Age when first started using electronic nicotine products fairly regularly
X04_YV1027	Youth	Number of puffs you take each time you pick up your electronic nicotine product to use it
X04_YJ0103TC	Youth	Number of traditional cigars as a blunt smoked today/yesterday/day before yesterday
X04_YJ0103CG	Youth	Number of cigarillos as blunts smoked today/yesterday/day before yesterday
X04_YG0103TC	Youth	Number of traditional cigars smoked today/yesterday/day before yesterday
X04_YH1007	Youth	Age when first started smoking tobacco in a hookah fairly regularly
X04_YH9004	Youth	Average number of times you smoke tobacco in a hookah in a month
X04_YU1007	Youth	Age when you first started using snus fairly regularly
X04_YS1007	Youth	Age when first started using smokeless tobacco fairly regularly
X04_YX0068	Youth	In past seven days, number of hours that you were in close contact with others when they were smoking
X04_YT0007_FT	Youth	Current height: Feet
X04_YT0007_IN	Youth	Current height: Inches
X04_YX0311	Youth	Current height: Meters
X04_YT0008	Youth	Current weight: Pounds
X04_YX0312	Youth	Current weight: Kilograms
X04_YX0074_NB	Youth	Age when consumed first alcoholic drink
X04_YX0082_NB_02	Youth	Age when first used: Painkillers, sedatives, or tranquilizers
X04_YX0082_NB_03	Youth	Age when first used: Cocaine or crack
X04_YX0082_NB_04	Youth	Age when first used: Stimulants like methamphetamine or speed

**Table F-6. Wave 5 questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R05_AM0065	Adult	Number of years lived in the United States
R05_AC1009_NN	Adult	How long since last smoked a cigarette – Number
R05_AV1009_NN	Adult	How long since you last took a puff from an electronic nicotine product – Number
R05_AG1009TC_NN	Adult	How long since last smoked a traditional cigar – Number
R05_AG1009CG_NN	Adult	How long since last smoked a cigarillo – Number
R05_AG1009FC_NN	Adult	How long since last smoked a filtered cigar – Number
R05_AG1009TJ_NN	Adult	How long since last smoked a traditional cigar as a blunt – Number
R05_AG1009GJ_NN	Adult	How long since last smoked a cigarillo as a blunt – Number
R05_AG1009FJ_NN	Adult	How long since last smoked a filtered cigar as a blunt – Number
R05_AP1009_NN	Adult	How long since you last smoked a pipe filled with tobacco – Number
R05_AH1009_NN	Adult	How long since you last smoked tobacco in a hookah – Number
R05_AU1009_NN	Adult	How long since you last used snus – Number
R05_AS1009_NN	Adult	How long since you last used smokeless tobacco – Number
R05_AC1020_NB	Adult	Age when first started smoking cigarettes every day
R05_AC9002_NN	Adult	How long smoking/smoked cigarettes fairly regularly – Number
R05_AC9029_NN	Adult	How long since you last smoked a cigarette – Number (Some day or current experimental smokers)
R05_AC1024_NN	Adult	Time to first cigarette after waking on days smoked – Number (current smokers)
R05_AC1024_30D_NN	Adult	Time to first cigarette after waking on days smoked – Number (non-current smokers)
R05_AC1051MC_NN	Adult	How long smoked regular brand of manufactured cigarettes – Number
R05_AV9029_NN	Adult	Time since you took last puff from an electronic nicotine product – Number (Some day or current experimental users)
R05_AV1026	Adult	Average number of times you pick up your electronic nicotine product to take one or more puffs each day on days used
R05_AV1027	Adult	Number of puffs you take each time you pick up your electronic nicotine product to use it
R05_AV1024_NN	Adult	Time to first electronic nicotine product puff after waking – Number (current users)
R05_AV1024_30D_NN	Adult	Time to first electronic nicotine product puff after waking – Number (non-current users)
R05_AV1040	Adult	Number of [electronic nicotine products/electronic nicotine cartridges/milliliters of e-liquid] that come in a box or pack you buy/bought
R05_AM0103	Adult	Number of puffs of marijuana taken from an electronic nicotine product today/yesterday/day before yesterday
R05_AJ9029TC_NN	Adult	Time since last smoked traditional cigar as blunt – Number (Some day or current experimental smokers)
R05_AJ9029CG_NN	Adult	Time since last smoked cigarillo as blunt – Number (Some day or current experimental smokers)
R05_AJ1023CG	Adult	In past 30 days, average number of cigarillos as blunts smoked per day on days smoked
R05_AJ1040CG	Adult	Number of cigarillos for blunts come in box or pack usually buy

**Table F-6. Wave 5 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R05_AJ9029FC_NN	Adult	Time since last smoked filtered cigar as blunt – Number (Some day or current experimental smokers)
R05_AG1007TC_NB	Adult	Age when first started smoking traditional cigars fairly regularly
R05_AG9029TC_NN	Adult	Time since last smoked traditional cigars – Number (Some day or current experimental smokers)
R05_AG1024TC_NN	Adult	Time to first traditional cigar after waking on days smoked – Number (current smokers)
R05_AG1051TC_NN	Adult	How long smoked regular brand of traditional cigar – Number
R05_AG9029CG_NN	Adult	Time since last smoked cigarillos – Number (Some day or current experimental smokers)
R05_AG1023CG	Adult	In past 30 days, average number of cigarillos smoked per day on days smoked
R05_AG1024CG_NN	Adult	Time to first cigarillo after waking on days smoked – Number (current smokers)
R05_AG1024CG_30D_NN	Adult	Time to first cigarillo after waking on days smoked – Number (non-current smokers)
R05_AG1051CG_NN	Adult	How long smoked regular brand of cigarillo – Number
R05_AG1020FC_NB	Adult	Age when first started smoking filtered cigars every day
R05_AG9029FC_NN	Adult	Time since last smoked filtered cigars – Number (Some day or current experimental smokers)
R05_AG1024FC_NN	Adult	Time to first filtered cigar after waking on days smoked – Number (current smokers)
R05_AG1051FC_NN	Adult	How long smoked regular brand of filtered cigar – Number
R05_AP1007_NB	Adult	Age when first started smoking a pipe filled with tobacco fairly regularly
R05_AP9029_NN	Adult	Time since last smoked a pipe filled with tobacco – Number (Some day or current experimental smokers)
R05_AP1024_NN	Adult	Time to first pipe bowl after waking on days smoked – Number (current smokers)
R05_AP1051_NN	Adult	How long smoked regular brand of pipe tobacco – Number
R05_AH1007_NB	Adult	Age when first started smoking hookah fairly regularly
R05_AH1020_NB	Adult	Age when first started smoking hookah every day
R05_AH9005	Adult	Average number of times you smoke a hookah in a week
R05_AH9029_NN	Adult	Time since last smoked hookah – Number
R05_AH1051_NN	Adult	How long smoked regular brand of shisha or hookah tobacco – Number
R05_AU1007_NB	Adult	Age when first started using snus fairly regularly
R05_AU1020_NB	Adult	Age when first started using snus every day
R05_AU9029_NN	Adult	How long since you last used snus – Number (Some day or current experimental users)
R05_AU1051_NN	Adult	How long used regular brand of snus – Number
R05_AS1020_NB	Adult	Age when first started using smokeless tobacco every day
R05_AS1022	Adult	In past 30 days, number of days used smokeless tobacco
R05_AS9029_NN	Adult	Time since last used smokeless tobacco – Number (Some day or current experimental users)
R05_AS9002	Adult	Number of days to use up the container of smokeless tobacco
R05_AN0135E_VALUE_INFO	Adult	Type of response provided for R05_AN0135E

**Table F-6. Wave 5 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R05_AN0130E_NN	Adult	In past 12 months, length of time you stopped using electronic nicotine products because you were trying to quit - Number
R05_AN0145E_NN	Adult	In past 12 months, longest time period for which you stopped using electronic nicotine products because you were trying to quit - Number
R05_AN0115	Adult	In past 12 months, number of times tried to quit smoking/using tobacco product(s)
R05_AN0135_VALUE_INFO	Adult	Type of response provided for R05_AN0135
R05_AN0130_NN	Adult	In past 12 months, length of time you stopped smoking/using tobacco product(s) because you were trying to quit - Number
R05_AN0145_NN	Adult	In past 12 months, longest time period for which you stopped smoking/using tobacco product(s) because you were trying to quit - Number
R05_AN0205_NN	Adult	How long ago you stopped using prescription medication - Number (current established, recent former established or current experimental non-electronic tobacco users)
R05_AN0251_NN	Adult	How long been using NRT - Number (past 12 month non-electronic tobacco quitters or quit attempters)
R05_AN0252_NN	Adult	How long have you been using a prescription drug - Number (past 12 month non-electronic tobacco quitters or quit attempters)
R05_AX0068	Adult	In past 7 days, number of hours that you were in close contact with others when they were smoking
R05_AX0679_FT	Adult	Height without shoes: Feet
R05_AX0679_IN	Adult	Height without shoes: Inches
R05_AX0316	Adult	Height without shoes: Meters
R05_AX0109	Adult	Current weight: Pounds
R05_AX0312	Adult	Current weight: Kilograms
R05_AX0242_NN	Adult	On days when you do physical activity or exercise of at least moderate intensity, how long do you do these activities - Number
R05_AX0116_NB	Adult	Age when you were first told you had congestive heart failure
R05_AX0112_NB	Adult	Age when you were first told you had a heart attack or needed bypass surgery
R05_AX0411	Adult	Age when you were first told you have poor circulation (PAD or PVD)
R05_AX0123_NB	Adult	Age when you were first told you had emphysema
R05_AX0124_NB	Adult	Age when you were first told you had asthma
R05_AX0729_NN	Adult	Time since you had your teeth cleaned by a dentist, hygienist, or other health professional - Number
R05_AX0725	Adult	Number of times you brush your teeth in one day
R05_AX0696	Adult	In the last seven days, number of times used dental floss or any other device to clean between teeth
R05_AX0131_NB	Adult	Age when you were first told you had gum disease
R05_AX0133_NB	Adult	Age when you were first told you had pre-cancerous oral lesions
R05_AX0280_NB	Adult	Age when you were first told you had diabetes, sugar diabetes, high blood sugar or borderline diabetes
R05_AX0143_NB	Adult	Age when you were first told you had an ulcer

**Table F-6. Wave 5 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R05_AX0148_NB	Adult	Age when you were first told you had stomach or gastro-intestinal bleeding
R05_AX0152_NB	Adult	Age when you were first told you had a cataract or glaucoma
R05_AX0150_NB	Adult	Age when you were first told you had osteoporosis
R05_AX0198_NB	Adult	Age when you were first told you had a bone fracture because you have fragile bones
R05_AX0703	Adult	Age when first told you had macular degeneration
R05_AN0359H	Adult	Number of nicotine patches, gum, inhaler, nasal spray, lozenges or pills used today/yesterday/the day before yesterday (former experimental or long-term former established tobacco users)
R05_PT0008_LB	Youth	Youth's current weight (pounds)
R05_PM0065_NN	Youth	Number of years lived in the United States (parent respondent)
R05_YM0065	Youth	Number of years lived in the United States (youth)
R05_YC1007	Youth	Age when first started smoking cigarettes fairly regularly
R05_YV1007	Youth	Age when first started using electronic nicotine products fairly regularly
R05_YJ0103CG	Youth	Number of cigarillos as blunts smoked today/yesterday/day before yesterday
R05_YG1007TC	Youth	Age when first started smoking traditional cigars fairly regularly
R05_YG0103CL	Youth	Number of cigarillos smoked today/yesterday/day before yesterday
R05_YG0103FC	Youth	Number of filtered cigars smoked today/yesterday/day before yesterday
R05_YH9005	Youth	Average number of times you smoke tobacco in a hookah in a week
R05_YU1007	Youth	Age when you first started using snus fairly regularly
R05_YX0068	Youth	In past seven days, number of hours that you were in close contact with others when they were smoking
R05_YT0007_FT	Youth	Current height: Feet
R05_YT0007_IN	Youth	Current height: Inches
R05_YX0311	Youth	Current height: Meters
R05_YT0008	Youth	Current weight: Pounds
R05_YX0312	Youth	Current weight: Kilograms
R05_YX0086_NB	Youth	Age when first drank alcohol at all
R05_YX0074_NB	Youth	Age when consumed first alcoholic drink

**Table F-7. Wave 5.5 questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
X05_AM0065	Adult	Number of years lived in the United States
X05_AC1009_NN	Adult	How long since last smoked a cigarette – Number
X05_AV1009_NN	Adult	How long since you last took a puff from an electronic nicotine product – Number
X05_AG1009TC_NN	Adult	How long since last smoked a traditional cigar – Number
X05_AG1009CG_NN	Adult	How long since last smoked a cigarillo – Number
X05_AG1009FC_NN	Adult	How long since last smoked a filtered cigar – Number
X05_AG1009GJ_NN	Adult	How long since last smoked a cigarillo as a blunt – Number
X05_AG1009FJ_NN	Adult	How long since last smoked a filtered cigar as a blunt – Number
X05_AP1009_NN	Adult	How long since you last smoked a pipe filled with tobacco – Number
X05_AH1009_NN	Adult	How long since you last smoked tobacco in a hookah – Number
X05_AU1009_NN	Adult	How long since you last used snus – Number
X05_AS1009_NN	Adult	How long since you last used smokeless tobacco – Number
X05_AC9029_NN	Adult	How long since you last smoked a cigarette – Number (Some day or current experimental smokers)
X05_AC1042_D	Adult	Amount usually paid for a pack of cigarettes – Dollars
X05_AC1042_C	Adult	Amount usually paid for a pack of cigarettes – Cents
X05_AV9029_NN	Adult	Time since you took last puff from an electronic nicotine product – Number (Some day or current experimental users)
X05_AV1026	Adult	Average number of times you pick up your electronic nicotine product to take one or more puffs each day on days used
X05_AJ9029TC_NN	Adult	Time since last smoked traditional cigar as blunt – Number (Some day or current experimental smokers)
X05_AJ9029CG_NN	Adult	Time since last smoked cigarillo as blunt – Number (Some day or current experimental smokers)
X05_AJ0103CG	Adult	Number of cigarillos as blunts smoked today/yesterday/day before yesterday
X05_AJ9029FC_NN	Adult	Time since last smoked filtered cigar as blunt – Number (Some day or current experimental smokers)
X05_AG9029TC_NN	Adult	Time since last smoked traditional cigars – Number (Some day or current experimental smokers)
X05_AG1024TC_30D_NN	Adult	Time to first traditional cigar after waking on days smoked – Number (non-current smokers)
X05_AG9029CG_NN	Adult	Time since last smoked cigarillos – Number (Some day or current experimental smokers)
X05_AG9029FC_NN	Adult	Time since last smoked filtered cigars – Number (Some day or current experimental smokers)
X05_AP9029_NN	Adult	Time since last smoked a pipe filled with tobacco – Number (Some day or current experimental smokers)
X05_AP0103	Adult	Number of bowls filled with pipe tobacco smoked today/yesterday/day before yesterday
X05_AH9029_NN	Adult	Time since last smoked hookah – Number
X05_AS9029_NN	Adult	Time since last used smokeless tobacco – Number (Some day or current experimental users)
X05_AX0090	Adult	Self perception of physical health
X05_AX0091	Adult	Self perception of mental health
X05_AX0088	Adult	Self perception of overall health
X05_AX0090_12M	Adult	Self perception of physical health now compared with 12 months ago
X05_AX0091_12M	Adult	Self perception of mental health now compared with 12 months ago

**Table F-7. Wave 5.5 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
X05_AX0088_12M	Adult	Self perception of overall health now compared with 12 months ago
X05_AX0158_01	Adult	Takes anti-inflammatory or pain medication: Aspirin
X05_AX0158_02	Adult	Takes anti-inflammatory or pain medication: Tylenol (Acetaminophen)
X05_AX0158_03	Adult	Takes anti-inflammatory or pain medication: Cox-2 inhibitors
X05_AX0158_04	Adult	Takes anti-inflammatory or pain medication: Other anti-inflammatory medications, for example: ibuprofen, Motrin, Advil, Naprosyn, or Aleve
X05_AX0134_12M	Adult	In past 12 months, been pregnant
X05_AX0161	Adult	Last time you had significant problems with: Feeling very trapped, lonely, sad, blue, depressed or hopeless about the future
X05_AX0162	Adult	Last time you had significant problems with: Sleep trouble – such as bad dreams, sleeping restlessly or falling asleep during the day
X05_AX0163	Adult	Last time you had significant problems with: Feeling very anxious, nervous, tense, scared, panicked or like something bad was going to happen
X05_AX0164	Adult	Last time you had significant problems with: Becoming very distressed and upset when something reminded you of the past
X05_AX0165	Adult	Last time you did the following two or more times: Lied or conned to get things you wanted or to avoid having to do something
X05_AX0166	Adult	Last time you did the following two or more times: Had a hard time paying attention at school, work or home
X05_AX0167	Adult	Last time you did the following two or more times: Had a hard time listening to instructions at school, work or home
X05_AX0168	Adult	Last time you did the following two or more times: Were a bully or threatened other people
X05_AX0169	Adult	Last time you did the following two or more times: Started physical fights with other people
X05_AX0250	Adult	Last time you did the following two or more times: Felt restless or the need to run around or climb on things
X05_AX0251	Adult	Last time you did the following two or more times: Gave answers before the other person finished asking the question
X05_PT0008_LB	Youth	Youth's current weight (pounds)
X05_YM0065	Youth	Number of years lived in the United States (youth)
X05_YV1007	Youth	Age when first started using electronic nicotine products fairly regularly
X05_YV1027	Youth	Number of puffs you take each time you pick up your electronic nicotine product to use it
X05_YG1007TC	Youth	Age when first started smoking traditional cigars fairly regularly
X05_YH9004	Youth	Average number of times you smoke tobacco in a hookah in a month
X05_YU1007	Youth	Age when you first started using snus fairly regularly
X05_YX0068	Youth	In past seven days, number of hours that you were in close contact with others when they were smoking
X05_YT0007_FT	Youth	Current height: Feet
X05_YT0007_IN	Youth	Current height: Inches
X05_YT0008	Youth	Current weight: Pounds
X05_YT0282	Youth	In past 12 months, number of times you visited an emergency room or urgent care center because of asthma

**Table F-8. PATH-ATS questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Variable description</b>
T05_AC1009_NN	How long since last smoked a cigarette – Number
T05_AV1009_NN	How long since you last took a puff from an electronic nicotine product – Number
T05_AG1009TC_NN	How long since last smoked a traditional cigar – Number
T05_AG1009CG_NN	How long since last smoked a cigarillo – Number
T05_AG1009FC_NN	How long since last smoked a filtered cigar – Number
T05_AG1009GJ_NN	How long since last smoked a cigarillo as a blunt – Number
T05_AP1009_NN	How long since you last smoked a pipe filled with tobacco – Number
T05_AH1009_NN	How long since you last smoked tobacco in a hookah – Number
T05_AU1009_NN	How long since you last used snus – Number
T05_AS1009_NN	How long since you last used smokeless tobacco – Number
T05_AC1021_NN	Average number of cigarettes now smoked each day – Number
T05_AC9029_NN	How long since you last smoked a cigarette – Number (Some day or current experimental smokers)
T05_AC1023_NN	In past 30 days, average number of cigarettes smoked per day on days smoked – Number
T05_AC1024_NN	Time to first cigarette after waking on days smoked – Number (current smokers)
T05_AC1024_30D_NN	Time to first cigarette after waking on days smoked – Number (non-current smokers)
T05_AV9029_NN	Time since you took last puff from an electronic nicotine product – Number (Some day or current experimental users)
T05_AV1026	Average number of times you pick up your electronic nicotine product to take one or more puffs each day on days used
T05_AV1027	Number of puffs you take each time you pick up your electronic nicotine product to use it
T05_AV1024_NN	Time to first electronic nicotine product puff after waking – Number (current users)
T05_AV1024_30D_NN	Time to first electronic nicotine product puff after waking – Number (non-current users)
T05_AG9029TC_NN	Time since last smoked traditional cigars – Number (Some day or current experimental smokers)
T05_AG1024TC_NN	Time to first traditional cigar after waking on days smoked – Number (current smokers)
T05_AG1024CG_NN	Time to first cigarillo after waking on days smoked – Number (current smokers)
T05_AG9029FC_NN	Time since last smoked filtered cigars – Number (Some day or current experimental smokers)
T05_AG1024FC_NN	Time to first filtered cigar after waking on days smoked – Number (current smokers)
T05_AP9029_NN	Time since last smoked a pipe filled with tobacco – Number (Some day or current experimental smokers)
T05_AP1024_30D_NN	Time to first pipe bowl after waking on days smoked – Number (non-current smokers)
T05_AH9029_NN	Time since last smoked hookah – Number
T05_AU9029_NN	How long since you last used snus – Number (Some day or current experimental users)
T05_AS9029_NN	Time since last used smokeless tobacco – Number (Some day or current experimental users)
T05_AS1024_NN	Time to first use of smokeless tobacco after waking on days used – Number (Current users)

**Table F-8. PATH-ATS questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Variable description</b>
T05_AN0130E_NN	In past 12 months, length of time you stopped using electronic nicotine products because you were trying to quit – Number
T05_AN0145E_NN	In past 12 months, longest time period for which you stopped using electronic nicotine products because you were trying to quit – Number
T05_AN0130_NN	In past 12 months, length of time you stopped smoking cigarettes because you were trying to quit – Number
T05_AN0135_VALUE_INFO	Type of response provided for T05_AN0135
T05_AN0145_NN	In past 12 months, longest time period for which you stopped smoking cigarettes because you were trying to quit – Number
T05_AX0126	In past 12 months, number of asthma attacks you have had that required the use of an oral or injected steroid

**Table F-9. Wave 6 questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R06_AM0065	Adult	Number of years lived in the United States
R06_AV1009_NN	Adult	How long since you last took a puff from an electronic nicotine product – Number
R06_AG1009TC_NN	Adult	How long since last smoked a traditional cigar – Number
R06_AG1009FC_NN	Adult	How long since last smoked a filtered cigar – Number
R06_AP1009_NN	Adult	How long since you last smoked a pipe filled with tobacco – Number
R06_AH1009_NN	Adult	How long since you last smoked tobacco in a hookah – Number
R06_AC1006	Adult	Age when first smoked part or all of a cigarette
R06_AC1020	Adult	Age when first started smoking cigarettes every day
R06_AC9029_NN	Adult	How long since you last smoked a cigarette – Number (Current smokers)
R06_AC1024_NN	Adult	Time to first cigarette after waking on days smoked – Number (current smokers)
R06_AC1051MC_NN	Adult	How long smoked regular brand of manufactured cigarettes – Number
R06_AV1006	Adult	Age when first used an electronic nicotine product, even one or two times
R06_AV1007	Adult	Age when first started using electronic nicotine products fairly regularly
R06_AV1020	Adult	Age when first started using electronic nicotine products every day
R06_AV9029_NN	Adult	Time since you took last puff from an electronic nicotine product – Number (Current users)
R06_AV1026	Adult	Average number of times you pick up your electronic nicotine product to take one or more puffs each day on days used
R06_AV1027	Adult	Number of puffs you take each time you pick up your electronic nicotine product to use it
R06_AV1024_NN	Adult	Time to first electronic nicotine product puff after waking – Number (current users)
R06_AV7701	Adult	Age when you were last evaluated or treated by a doctor for a respiratory condition related to your use of electronic nicotine product
R06_AV1051_NN	Adult	How long used regular brand of [electronic nicotine products/pods or cartridges/e-liquid] – Number
R06_AJ9029TC_NN	Adult	Time since last smoked traditional cigar as blunt – Number (Past 12 month smokers)
R06_AJ9029CG_NN	Adult	Time since last smoked cigarillo as blunt – Number (Past 12 month smokers)
R06_AJ9029FC_NN	Adult	Time since last smoked filtered cigar as blunt – Number (Past 12 month smokers)
R06_AG1006TC	Adult	Age when first smoked part or all of a traditional cigar, even one or two puffs
R06_AG1020TC	Adult	Age when first started smoking traditional cigars every day
R06_AG9029TC_NN	Adult	Time since last smoked traditional cigars – Number (Current smokers)
R06_AG1024TC_NN	Adult	Time to first traditional cigar after waking on days smoked – Number (current smokers)
R06_AG1051TC_NN	Adult	How long smoked regular brand of traditional cigar – Number
R06_AG1006CG	Adult	Age when first smoked part or all of a cigarillo, even one or two puffs

**Table F-9. Wave 6 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R06_AG1007CG	Adult	Age when first started smoking cigarillos fairly regularly
R06_AG1020CG	Adult	Age when first started smoking cigarillos every day
R06_AG9029CG_NN	Adult	Time since last smoked cigarillos – Number (Current smokers)
R06_AG1024CG_NN	Adult	Time to first cigarillo after waking on days smoked – Number (current smokers)
R06_AG1051CG_NN	Adult	How long smoked regular brand of cigarillo – Number
R06_AG1006FC	Adult	Age when first smoked part or all of a filtered cigar, even one or two puffs
R06_AG1007FC	Adult	Age when first started smoking filtered cigars fairly regularly
R06_AG1020FC	Adult	Age when first started smoking filtered cigars every day
R06_AG9029FC_NN	Adult	Time since last smoked filtered cigars – Number (Current smokers)
R06_AG1024FC_NN	Adult	Time to first filtered cigar after waking on days smoked – Number (current smokers)
R06_AP1006	Adult	Age when you first smoked a pipe filled with tobacco, even one or two puffs
R06_AP1007	Adult	Age when first started smoking a pipe filled with tobacco fairly regularly
R06_AP9029_NN	Adult	Time since last smoked a pipe filled with tobacco – Number (Current smokers)
R06_AP1024_NN	Adult	Time to first pipe bowl after waking on days smoked – Number (current smokers)
R06_AP1051_NN	Adult	How long smoked regular brand of pipe tobacco – Number
R06_AH1006	Adult	Age when first smoked tobacco in a hookah, even one or two puffs
R06_AH1020	Adult	Age when first started smoking tobacco in a hookah every day
R06_AH9029_NN	Adult	Time since last smoked hookah – Number
R06_AU1006	Adult	Age when first used snus, even one or two times
R06_AU1007	Adult	Age when first started using snus fairly regularly
R06_AU1020	Adult	Age when first started using snus every day
R06_AU9029_NN	Adult	How long since you last used snus – Number (Current users)
R06_AU1024_NN	Adult	Time to first snus product after waking on days used – Number (current users)
R06_AS1006	Adult	Age when you first used smokeless tobacco, even one or two times
R06_AS1007	Adult	Age when first started using smokeless tobacco fairly regularly
R06_AS1020	Adult	Age when first started using smokeless tobacco every day
R06_AS1021	Adult	Average number of times smokeless tobacco now used each day
R06_AS9029_NN	Adult	Time since last used smokeless tobacco – Number (Current users)
R06_AN0115E	Adult	In past 12 months, number of times tried to quit using electronic nicotine products
R06_AN0135E_VALUE_INFO	Adult	Type of response provided for R06_AN0135E
R06_AN0130E_NN	Adult	In past 12 months, length of time you stopped using electronic nicotine products because you were trying to quit – Number
R06_AN0145E_NN	Adult	In past 12 months, longest time period for which you stopped using electronic nicotine products because you were trying to quit – Number

**Table F-9. Wave 6 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R06_AN0205E_NN	Adult	How long ago you stopped using prescription medication – Number (Current or recent former ENP users)
R06_AN0115	Adult	In past 12 months, number of times tried to quit smoking/using tobacco product(s)
R06_AN0135_VALUE_INFO	Adult	Type of response provided for R06_AN0135
R06_AN0270_NN	Adult	How long used tobacco product along with e-cigarettes or other electronic nicotine products during last quit attempt – Number
R06_AN0271_NN	Adult	How long used e-cigarette or other electronic nicotine products to help after stopped using tobacco product – Number
R06_AN0130_NN	Adult	In past 12 months, length of time you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R06_AN0145_NN	Adult	In past 12 months, longest time period for which you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R06_AN0260_NN	Adult	How long used e-cigarettes or other electronic nicotine products to help you during last quit attempt – Number
R06_AN0375_NN	Adult	How long ago did you stop using the nicotine patch – Number
R06_AX0068	Adult	In past 7 days, number of hours that you were in close contact with others when they were smoking
R06_AX0679_FT	Adult	Height without shoes: Feet
R06_AX0679_IN	Adult	Height without shoes: Inches
R06_AX0109	Adult	Current weight: Pounds
R06_AX0242_NN	Adult	On days when you do physical activity or exercise of at least moderate intensity, how long do you do these activities – Number
R06_AX0114	Adult	Age when you were first told you had high blood pressure
R06_AX0115	Adult	Age when you were first told you had high cholesterol
R06_AX0116	Adult	Age when you were first told you had congestive heart failure
R06_AX0117	Adult	Age when you were first told you had a stroke
R06_AX0123	Adult	Age when you were first told you had emphysema
R06_AX0126	Adult	In past 12 months, number of asthma attacks you have had that required the use of an oral or injected steroid
R06_AX0729_NN	Adult	Time since you had your teeth cleaned by a dentist, hygienist, or other health professional – Number
R06_AX0696	Adult	In the last seven days, number of times used dental floss or any other device to clean between teeth
R06_AX0131	Adult	Age when you were first told you had gum disease
R06_AX0133	Adult	Age when you were first told you had pre-cancerous oral lesions
R06_AX0280	Adult	Age when you were first told you had diabetes, sugar diabetes, high blood sugar, or borderline diabetes
R06_AX0143	Adult	Age when you were first told you had an ulcer
R06_AX0148	Adult	Age when you were first told you had stomach or other gastro-intestinal bleeding
R06_AX0152	Adult	Age when you were first told you had a cataract or glaucoma
R06_AX0146_07	Adult	Age when colon cancer was first diagnosed
R06_AX0146_13	Adult	Age when liver cancer was first diagnosed
R06_AX0150	Adult	Age when you were first told you had osteoporosis, sometimes called thin or brittle bones
R06_AX0198	Adult	Age when you were first told you had a bone fracture because you have fragile bones

**Table F-9. Wave 6 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R06_AX0757	Adult	Age when you were first told that you had schizophrenia, schizoaffective disorder, or psychosis
R06_AX0762	Adult	Age when you were first told you that you had a psychotic illness or episode
R06_AX0137_NN	Adult	Weeks/Months pregnant – Number
R06_AX0086	Adult	Age when you first drank alcohol at all, counting small tastes or sips
R06_AX0074	Adult	Age when you had your first alcoholic drink, other than small tastes or sips
R06_AX0079	Adult	Age when you first used marijuana
R06_AX0288	Adult	In past 30 days, number of days you used marijuana
R06_AW1022	Adult	In past 30 days, number of days you used marijuana in an electronic nicotine product
R06_AW0103	Adult	Number of puffs of marijuana taken from electronic nicotine product today/yesterday/day before yesterday
R06_AW7701	Adult	Age when you were last evaluated or treated by a doctor for a respiratory condition related to your use of marijuana in electronic nicotine products
R06_AX0082_01	Adult	Age when first used Ritalin® or Adderall®
R06_AX0082_02	Adult	Age when first used painkillers
R06_AX0082_03	Adult	Age when first used sedatives or tranquilizers
R06_AX0082_04	Adult	Age when first used cocaine or crack
R06_AX0082_05	Adult	Age when first used stimulants like methamphetamine or speed
R06_AX0082_06	Adult	Age when first used heroin
R06_AX0082_07	Adult	Age when first used inhalants or solvents
R06_AX0082_08	Adult	Age when first used hallucinogens
R06_AM1937_12M	Adult	In past 12 months, total number of days hospitalized because of coronavirus symptoms
R06_AM1937	Adult	Total number of days hospitalized because of coronavirus symptoms
R06_PT0008_LB	Parent	Youth's current weight (pounds)
R06_PM0065_NN	Parent	Number of years lived in the United States (parent respondent)
R06_YM0065	Youth	Number of years lived in the United States (youth)
R06_YC1006	Youth	Age when first tried cigarette smoking, even one or two puffs
R06_YS9002	Youth	Number of days to use up the container of smokeless tobacco (that you usually buy)
R06_YT0007_FT	Youth	Current height: Feet
R06_YT0007_IN	Youth	Current height: Inches
R06_YT0008	Youth	Current weight: Pounds
R06_YX0086	Youth	Age when you first drank alcohol at all, counting small tastes or sips
R06_YX0074	Youth	Age when you had your first alcoholic drink, other than small tastes or sips
R06_YX0082_03	Youth	Age when first used sedatives or tranquilizers
R06_YX0082_06	Youth	Age when first used heroin
R06_YX0082_08	Youth	Age when first used hallucinogens
R06_YM1937_12M	Youth	In past 12 months, total number of days hospitalized because of coronavirus symptoms

**Table F-10. Wave 7 questionnaire variables with coded outlier values**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R07_AM0065	Adult	Number of years lived in the United States
R07_AC1009_NN	Adult	How long since last smoked a cigarette - Number
R07_AV1009_NN	Adult	How long since you last took a puff from an electronic nicotine product - Number
R07_AG1009TJ_NN	Adult	How long since last smoked a traditional cigar as a blunt - Number
R07_AG1009GJ_NN	Adult	How long since last smoked a cigarillo as a blunt - Number
R07_AU1009_NN	Adult	How long since you last used snus - Number
R07_AS1009_NN	Adult	How long since you last used smokeless tobacco - Number
R07_AA0103NP	Adult	Number of nicotine pouches used today/yesterday/day before yesterday
R07_AC1006	Adult	Age when first smoked part or all of a cigarette
R07_AC1020	Adult	Age when first started smoking cigarettes every day
R07_AC9006_RS_NN	Adult	Average number of cigarettes smoked per day 12 months ago - Number
R07_AC9029_NN	Adult	Time since last smoked a cigarette - Number (Current smokers, no past 30 day use)
R07_AC1051MC_NN	Adult	How long smoked regular brand of manufactured cigarettes - Number
R07_AC1051RY_NN	Adult	How long smoked regular brand of roll-your-own cigarette tobacco - Number
R07_AV1006	Adult	Age when first used an electronic nicotine product, even one or two times
R07_AV1007	Adult	Age when first started using electronic nicotine products fairly regularly
R07_AV1020	Adult	Age when first started using electronic nicotine products every day
R07_AV9029_NN	Adult	Time since last puff from an electronic nicotine product - Number (Current users, no past 30 day use)
R07_AV1026	Adult	Average number of times you pick up your electronic nicotine product to take one or more puffs each day on days used
R07_AV1027	Adult	Number of puffs you take each time you pick up your electronic nicotine product to use it
R07_AV1024_NN	Adult	Time to first electronic nicotine product puff after waking - Number (current users)
R07_AV1040	Adult	Number of [electronic nicotine products/electronic nicotine cartridges/milliliters of e-liquid] that come in a box or pack you buy/bought
R07_AJ9029TC_NN	Adult	Time since last smoked traditional cigar as blunt - Number (Past 12 month smokers, no past 30 day use)
R07_AJ9029CG_NN	Adult	Time since last smoked cigarillo as blunt - Number (Past 12 month smokers, no past 30 day use)
R07_AJ9029FC_NN	Adult	Time since last smoked filtered cigar as blunt - Number (Past 12 month smokers, no past 30 day use)
R07_AG1007TC	Adult	Age when first started smoking traditional cigars fairly regularly
R07_AG9029TC_NN	Adult	Time since last smoked traditional cigars - Number (Current smokers, no past 30 day use)
R07_AG1024TC_NN	Adult	Time to first traditional cigar after waking on days smoked - Number (current smokers)
R07_AG1007CG	Adult	Age when first started smoking cigarillos fairly regularly
R07_AG1020CG	Adult	Age when first started smoking cigarillos every day

**Table F-10. Wave 7 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R07_AG9029CG_NN	Adult	Time since last smoked cigarillos – Number (Current smokers, no past 30 day use)
R07_AG1024CG_NN	Adult	Time to first cigarillo after waking on days smoked – Number (current smokers)
R07_AG1006FC	Adult	Age when first smoked part or all of a filtered cigar, even one or two puffs
R07_AG1007FC	Adult	Age when first started smoking filtered cigars fairly regularly
R07_AG1020FC	Adult	Age when first started smoking filtered cigars every day
R07_AG9029FC_NN	Adult	Time since last smoked filtered cigars – Number (Current smokers, no past 30 day use)
R07_AP1006	Adult	Age when you first smoked a pipe filled with tobacco, even one or two puffs
R07_AP1007	Adult	Age when first started smoking a pipe filled with tobacco fairly regularly
R07_AP1020	Adult	Age when first started smoking a pipe filled with tobacco every day
R07_AP9029_NN	Adult	Time since last smoked a pipe filled with tobacco – Number (Current smokers, no past 30 day use)
R07_AH1006	Adult	Age when first smoked tobacco in a hookah, even one or two puffs
R07_AH1007	Adult	Age when first started smoking hookah fairly regularly
R07_AH1020	Adult	Age when first started smoking tobacco in a hookah every day
R07_AH9029_NN	Adult	Time since last smoked hookah – Number (Current smokers, no past 30 day use)
R07_AU1006	Adult	Age when first used snus, even one or two times
R07_AU1007	Adult	Age when first started using snus fairly regularly
R07_AU1020	Adult	Age when first started using snus every day
R07_AU9029_NN	Adult	How long since you last used snus – Number (Current users, no past 30 day use)
R07_AU1051_NN	Adult	How long used regular brand of snus – Number
R07_AS1006	Adult	Age when you first used smokeless tobacco, even one or two times
R07_AS1007	Adult	Age when first started using smokeless tobacco fairly regularly
R07_AS1020	Adult	Age when first started using smokeless tobacco every day
R07_AS9029_NN	Adult	Time since last used smokeless tobacco – Number (Current users, no past 30 day use)
R07_AS1024_NN	Adult	Time to first use of smokeless tobacco after waking on days used – Number (Current users)
R07_AS1051_NN	Adult	How long used regular brand of smokeless tobacco – Number
R07_AQ1006	Adult	Age when first used IQOS, even one or two times
R07_AQ1007	Adult	Age when first started using IQOS fairly regularly
R07_AQ1020	Adult	Age when first started using IQOS every day
R07_AQ9002_NN	Adult	How long you have used/used IQOS fairly regularly – Number
R07_AQ1021_NN	Adult	Average number of HeatSticks now used each day – Number
R07_AQ9029_NN	Adult	Time since last used IQOS – Number (Current users, no past 30 day use)
R07_AQ1023_NN	Adult	In past 30 days, average number of HeatSticks used per day on days used – Number
R07_AQ1024_NN	Adult	Time to first HeatStick use after waking on days used – Number (current users)
R07_AQ1040	Adult	Number of packs of HeatSticks in the carton you usually buy
R07_AQ1041_D	Adult	Amount usually paid for a carton of HeatSticks – Dollars

**Table F-10. Wave 7 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R07_AQ1042_D	Adult	Amount usually paid for a pack of HeatSticks – Dollars
R07_AN0329E	Adult	Number of Chantix, varenicline, Wellbutrin, Zyban or bupropion taken today/yesterday/the day before yesterday (current or recent former electronic nicotine product users)
R07_AN0130E_NN	Adult	In past 12 months, length of time you stopped using electronic nicotine products because you were trying to quit – Number
R07_AN0135E_VALUE_INFO	Adult	Type of response provided for R07_AN0135E
R07_AN0145E_NN	Adult	In past 12 months, longest time period for which you stopped using electronic nicotine products because you were trying to quit – Number
R07_AN0205E_NN	Adult	How long ago you stopped using prescription medication – Number (current established, recent former established or current experimental electronic nicotine product users)
R07_AN0251E_NN	Adult	How long been using NRT – Number (past 12 month electronic nicotine product quitters or quit attempters)
R07_AN0252E_NN	Adult	How long have you been using a prescription drug – Number (past 12 month electronic nicotine product quitters or quit attempters)
R07_AN0270_NN	Adult	How long used tobacco product along with e-cigarettes or other electronic nicotine products during last quit attempt – Number
R07_AN0271_NN	Adult	How long used e-cigarette or other electronic nicotine products to help after stopped using tobacco product – Number
R07_AN0359	Adult	Number of nicotine gum, inhaler, nasal spray, or lozenges used today/yesterday/the day before yesterday (current or recent former established non-electronic tobacco users)
R07_AN0130_NN	Adult	In past 12 months, length of time you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R07_AN0135_VALUE_INFO	Adult	Type of response provided for R07_AN0135
R07_AN0145_NN	Adult	In past 12 months, longest time period for which you stopped smoking/using tobacco product(s) because you were trying to quit – Number
R07_AN0260_NN	Adult	How long used e-cigarettes or other electronic nicotine products to help you during last quit attempt – Number
R07_AV9069_YR	Adult	Approximate year when you most recently swallowed e-liquid, or got it in your mouth, on your skin, or in your eyes
R07_AV9074_YR	Adult	Approximate year when any child under the age of 5 most recently swallowed e-liquid, or got it in their mouth, on their skin, or in their eyes
R07_AV9078_YR	Adult	Approximate year when any child age 5 to 12 most recently swallowed e-liquid, or got it in their mouth, on their skin, or in their eyes
R07_AV9084_YR	Adult	Approximate year when any child age 13 to 17 most recently swallowed e-liquid, or got it in their mouth, on their skin, or in their eyes
R07_AX0679_FT	Adult	Height without shoes: Feet
R07_AX0679_IN	Adult	Height without shoes: Inches
R07_AX0109	Adult	Current weight: Pounds
R07_AX0312	Adult	Current weight: Kilograms
R07_AX0316	Adult	Height without shoes: Meters

Table F-10. Wave 7 questionnaire variables with coded outlier values (continued)

Variable name	Data file	Variable description
R07_AX0242_NN	Adult	On days when you do physical activity or exercise of at least moderate intensity, how long do you do these activities – Number
R07_AX0114	Adult	Age when you were first told you had high blood pressure
R07_AX0115	Adult	Age when you were first told you had high cholesterol
R07_AX0116	Adult	Age when you were first told you had congestive heart failure
R07_AX0117	Adult	Age when you were first told you had a stroke
R07_AX0112	Adult	Age when you were first told you had a heart attack or needed bypass surgery
R07_AX0120	Adult	Age when you were first told you had COPD
R07_AX0123	Adult	Age when you were first told you had emphysema
R07_AX0124	Adult	Age when you were first told you had asthma
R07_AX0126	Adult	In past 12 months, number of asthma attacks you have had that required the use of an oral or injected steroid
R07_AX0729_NN	Adult	Time since you had your teeth cleaned by a dentist, hygienist, or other dental professional – Number
R07_AX0725	Adult	Number of times you brush your teeth in one day
R07_AX0696	Adult	In the last seven days, number of times used dental floss or any other device to clean between teeth
R07_AX0131	Adult	Age when you were first told you had gum disease
R07_AX0133	Adult	Age when you were first told you had pre-cancerous oral lesions
R07_AX0280	Adult	Age when you were first told you had diabetes, sugar diabetes, high blood sugar, or borderline diabetes
R07_AX0143	Adult	Age when you were first told you had an ulcer
R07_AX0148	Adult	Age when you were first told you had stomach or other gastrointestinal bleeding
R07_AX0146_02	Adult	Age when blood cancer was first diagnosed
R07_AX0146_16	Adult	Age when skin cancer (melanoma) was first diagnosed
R07_AX0146_19	Adult	Age when ovarian cancer was first diagnosed
R07_AX0146_22	Adult	Age when rectal cancer was first diagnosed
R07_AX0146_23	Adult	Age when skin (non-melanoma) cancer was first diagnosed
R07_AX0146_24	Adult	Age when skin (don't know what kind) cancer was first diagnosed
R07_AX0146_25	Adult	Age when soft tissue (muscle or fat) cancer was first diagnosed
R07_AX0146_26	Adult	Age when stomach cancer was first diagnosed
R07_AX0146_31	Adult	Age when other cancer was first diagnosed
R07_AX0150	Adult	Age when you were first told you had osteoporosis, sometimes called thin or brittle bones
R07_AX0198	Adult	Age when you were first told you had a bone fracture because you have fragile bones
R07_AX0703	Adult	Age when first told you had macular degeneration
R07_AX0289	Adult	Age when you had your first seizure
R07_AX0757	Adult	Age when you were first told that you had schizophrenia, schizoaffective disorder, or psychosis
R07_AX0762	Adult	Age when you were first told you that you had a psychotic illness or episode
R07_AX0137_NN	Adult	Weeks/Months pregnant – Number
R07_AX0086	Adult	Age when you first drank alcohol at all, counting small tastes or sips

**Table F-10. Wave 7 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R07_AX0074	Adult	Age when you had your first alcoholic drink, other than small tastes or sips
R07_AX0076	Adult	In past 30 days, number of alcoholic drinks usually consumed each day on days drank
R07_AX0079	Adult	Age when you first used marijuana
R07_AX0288	Adult	In past 30 days, number of days you used marijuana
R07_AW1022	Adult	In past 30 days, number of days you used marijuana in an electronic nicotine product
R07_AX0082_01	Adult	Age when first used Ritalin or Adderall
R07_AX0082_02	Adult	Age when first used painkillers
R07_AX0082_03	Adult	Age when first used sedatives or tranquilizers
R07_AX0082_04	Adult	Age when first used cocaine or crack
R07_AX0082_05	Adult	Age when first used stimulants like methamphetamine or speed
R07_AX0082_06	Adult	Age when first used heroin
R07_AX0082_07	Adult	Age when first used inhalants or solvents
R07_AX0082_08	Adult	Age when first used hallucinogens
R07_AM1937_12M	Adult	In past 12 months, total number of days hospitalized because of COVID-19
R07_AM1937	Adult	Total number of days hospitalized because of COVID-19
R07_YM0065	Youth	Number of years lived in the United States (youth)
R07_YC1006	Youth	Age when first tried cigarette smoking, even one or two puffs
R07_YC1007	Youth	Age when first started smoking cigarettes fairly regularly
R07_YG1007TC	Youth	Age when first started smoking traditional cigars fairly regularly
R07_YG1007CL	Youth	Age when first started smoking cigarillos fairly regularly
R07_YH1006	Youth	Age when first tried smoking tobacco in a hookah, even one or two puffs
R07_YU1006	Youth	Age when first tried snus, even one or two times
R07_YU1007	Youth	Age when you first started using snus fairly regularly
R07_YS1006	Youth	Age when first tried smokeless tobacco, even one or two times
R07_YQ1006	Youth	Age when you first tried IQOS, even one or two times
R07_YQ1007	Youth	Age when first started using IQOS fairly regularly
R07_YV9069_YR	Youth	Approximate year when you most recently swallowed e-liquid, or got it in your mouth, on your skin, or in your eyes
R07_YX0725	Youth	Number of times you brush your teeth in one day
R07_YT0007_FT	Youth	Current height: Feet
R07_YT0007_IN	Youth	Current height: Inches
R07_YT0008	Youth	Current weight: Pounds
R07_YX0311	Youth	Current height: Meters
R07_YX0312	Youth	Current weight: Kilograms
R07_YT0038	Youth	Age when you were first told you had asthma
R07_YX0289	Youth	Age when you had your first seizure
R07_YX0086	Youth	Age when you first drank alcohol at all, counting small tastes or sips
R07_YX0074	Youth	Age when you had your first alcoholic drink, other than small tastes or sips
R07_YX0076	Youth	In past 30 days, number of drinks usually consumed on days drank
R07_YX0079	Youth	Age when you first used marijuana
R07_YX0288	Youth	In past 30 days, number of days used marijuana
R07_YX0082_01	Youth	Age when first used Ritalin or Adderall
R07_YX0082_02	Youth	Age when first used painkillers

**Table F-10. Wave 7 questionnaire variables with coded outlier values (continued)**

<b>Variable name</b>	<b>Data file</b>	<b>Variable description</b>
R07_YX0082_04	Youth	Age when first used cocaine or crack
R07_YX0082_06	Youth	Age when first used heroin
R07_YX0082_07	Youth	Age when first used inhalants or solvents
R07_YX0082_08	Youth	Age when first used hallucinogens