

Health Insurance Coverage in the 2017 CPS ASEC Research File

Edward R. Berchick and Heide M. Jackson
Health and Disability Statistics Branch
Social, Economic, and Housing Statistics Branch
U.S. Census Bureau¹

SEHSD WORKING PAPER 2019-01

Highlights

- The 2017 CPS ASEC Research File was produced using the new system of editing, imputing, and weighting data, including health insurance data.
- The new processing system featured in the 2017 CPS ASEC Research File introduces new demographic changes to account for same sex couples, revised procedures for imputing income, and new procedures for estimating annual health insurance from subannual data.
- Health insurance estimates are different across the research and production files.
- During calendar year 2016, the percentage of people with no health insurance at any time was 7.9 percent using the 2017 CPS ASEC Research File, lower than the production file estimate of 8.8 percent.
- The research file contains information about whether a person had selected types of coverage for none, part, or all of the calendar year. Subannual data reveal that most people have health insurance coverage for all or none of a year in 2016.
- The research file can be linked to other CPS ASEC files and supplements, including the 2017 CPS ASEC Production File. Comparing records from the research file with the production file shows that about 4.8 percent of the sample have a change in their health insurance status.

Overview

After Census Bureau field representatives collect Current Population Survey Annual Social and Economic Supplement (CPS ASEC) survey responses,² the Census Bureau processes the

¹ This paper is released to inform interested parties of research and evaluation and to encourage discussion. The views expressed on statistical, measurement, or methodological issues are those of the authors and not necessarily those of the U.S. Census Bureau. The Census Bureau reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release. CBDRB-FY19-ROSS-B0048.

² Interviews are largely conducted through computer-assisted personal interview. However, some respondents are interviewed via computer-assisted telephone interview.

information into a data file that can provide a nationally-representative snapshot of the U.S. population.³ Specifically, the Census Bureau uses a series of statistical programs to resolve logically inconsistent answers (clean), fill in missing or incomplete information (impute), and account for survey sampling (weight). Collectively, these efforts are known as data processing.

Revisions to the data processing system were informed by over a decade of research and evaluation. Prior research identified limitations related to health insurance content in the CPS ASEC and suggested that improvements to the questionnaire and processing would produce higher quality estimates of health insurance coverage. After a decade of research showing that the CPS ASEC captured less health insurance coverage in comparison with other federal surveys (Hess et al., 2001; Kenney & Lynch, 2010; Pascale, 1999), the Census Bureau introduced a redesigned questionnaire in the 2014 survey. Data based on the new questionnaire became available during the traditional fall release (e.g., Smith & Medalia, 2014, 2015; Barnett & Vornovitsky, 2016; Barnett & Berchick, 2017; Berchick, Hood, & Barnett, 2018). However, to release these data in this timely manner, the Census Bureau used the existing processing system, which could not fully incorporate the richer information contained in the redesigned questionnaire.

The 2017 CPS ASEC Research File builds on the redesigned questionnaire and debuts the revised data processing system. That is, the Census Bureau has produced two CPS ASEC data files for this survey year. The first file is the 2017 CPS ASEC Production File which used existing programs of cleaning, imputing, and weighting the data.⁴ The second file, the 2017 CPS ASEC Research File, introduces a new method of estimating coverage which builds from subannual estimates to determine whether a person was covered at any point in the previous calendar year, and refines the methods by which respondents' data are cleaned, imputed and weighted to better reflect the rich economic and demographic data collected by the CPS ASEC. As most health insurance coverage questions concern the previous calendar year, coverage estimates in both 2017 CPS ASEC files corresponds to coverage in 2016.

In this working paper, we describe key changes implemented in the new processing system and present health insurance coverage estimates from the 2017 CPS ASEC Research File and 2017 CPS ASEC Production File.

³ For information on confidentiality protection, sampling error, nonsampling error, and definitions in the Current Population Survey, see <<https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar17.pdf>>.

⁴ This file was released in September 2017. We use the term "production file" to more clearly distinguish it from the research file, but the file has not been formally labeled as such.

Key Differences Between the Production and Research Files

The 2017 CPS ASEC Research File differs from the 2017 CPS ASEC Production File in three key ways. Specifically, the files differ with respect to (1) the population that the health insurance estimates describe; (2) the imputation process for households with incomplete and missing data; and (3) the availability of subannual measures to capture when in the calendar year a person had health insurance coverage. The research file and production file also include different types of coverage in their definitions of public, private, and military coverage. We discuss each of these high-level changes below.

The file also contains a number of new variables related to types of coverage that a person held at the time of interview and Marketplace coverage (including whether the premium was subsidized), as well as alternative measures of medical expenditures. The imputation procedure for subjective health status was also improved. A discussion of these additional changes lies outside the scope of the present working paper, but some information about these variables is available in the Data User Notes for Health Insurance, Health Status, and Medical Expenditures Content (<<https://www.census.gov/data/datasets/time-series/demo/income-poverty/data-extracts.html>>).

Universe for Health Insurance Content

In the production file, prior-year health insurance coverage information is available for all people. In the research file, however, prior-year information is not provided for a small number of infants who were born after the end of the prior calendar year. For example, a child born in January 2017 could neither be insured nor uninsured during 2016, but would be present in the household at the time of 2017 ASEC interview (and, therefore, could be currently insured or uninsured).⁵

When comparing estimates between the production and research files later in this working paper, we present two sets of comparisons. First, we compare the production estimates of health insurance coverage with the research file estimates. Second, we compare estimates after removing information from respondents born after the end of the prior calendar year.

Imputation Procedure

In both the production and research file, cases with missing health insurance information have values allocated to them through hotdeck imputation. Hotdeck imputation refers to a process by which we match a case with a missing value to an observation that is most similar to them

⁵ In the CPS ASEC, weighted estimates correspond to the population as of March of the survey year. Therefore, the weighted estimates of the total number of people with and without insurance will slightly differ from other population estimates.

on other observed characteristics. (See Andridge and Little, 2010 for more information about this imputation method.)

In the production file, hotdeck imputation was done at the individual level, separately for each insurance type, and without consideration of other coverage types reported or imputed for an individual or others within their household. This approach has the disadvantages of allowing improbable combinations of dual coverage and of allowing people who likely share coverage to have different coverage types.

In the research file, the imputation procedure starts by identifying people within a household who are eligible to share insurance coverage. These people are grouped together as a health insurance unit (HIU) (see State Health Access Data Assistance Center, 2012 for more information). If a person is missing information but another person within their HIU has data reported, the person with missing data is assumed to have the same coverage information as the person with reported data. If all people within a health insurance unit are missing data, hotdeck imputation matches a health insurance unit with missing data to a similar health insurance unit with reported data. In both procedures, all health insurance information is imputed simultaneously. This has the advantage of allowing imputed data to better match reported data for individuals and for health insurance units.

Moreover, the imputation procedure in the research file includes a wider range of characteristics than the procedures in the current production file. These characteristics address many limitations noted by prior research (see Davern, 2007) and make the imputation strategy more robust to a changing health insurance landscape. As a result of these two changes, the distribution of coverage and coverage types should be similar for comparable cases with missing information and cases with reported information.

Subannual Measures

Although CPS ASEC collects coverage information at the subannual level, the production file does not contain any information about the duration of coverage during the calendar year.

To offer a balance between providing subannual health insurance data while ensuring respondents' confidentiality, the public-use 2017 ASEC research file contains a limited number of variables that indicate whether a person was covered for none, some, or all of the previous calendar year (in this case, 2016). Such subannual information is provided for any coverage, as well as for private coverage, government coverage, and Medicaid. However, subannual data are

used internally when determining whether or not a person had any health insurance coverage during the calendar year.⁶

Interpretation of Survey Responses

The CPS ASEC health insurance questionnaire asks respondents whether they had any type of health insurance coverage, who else was covered by that insurance, and when (between January of the prior calendar year and the month of interview) this coverage was held. With the adaptive design of the questionnaire to help report their coverage, persons could report their coverage in a multitude of ways.

To release the data in a timely fashion, the production file had to be mapped to the pre-redesign data structure, which necessitated some simplifying assumptions regarding the coverage reported. The research file considers more nuanced information when considering reports.

Definitions of Key Health Insurance Types

In the production file, people are considered to have government coverage if they have coverage through Medicare, Medicaid or military health insurance coverage. In the production file, all types of military coverage were considered to be government insurance because using the processing system necessitated combining these types of coverage.

The research file allows these types of coverage to be distinguished from one another. People are considered to have government coverage if they have coverage through Medicare, Medicaid, or the VA / CHAMPVA. The research file also contains new variables allowing users to identify the types of military coverage. For more information on these changes, please see USER NOTE.

With these differences in mind, the remainder of this working paper provides high-level estimates to highlight estimates from 2017 CPS ASEC Research File and compare them with estimates from the 2017 CPS ASEC Production File.

⁶ The Census Bureau is currently considering ways to make additional subannual information available to researchers and policymakers in a way that sufficiently protects respondents' privacy and confidentiality. Evaluation is ongoing and these evaluations will be posted to the Census Bureau website when they are available. See in the Data User Notes for Health Insurance, Health Status, and Medical Expenditures Content (<<https://www.census.gov/data/datasets/time-series/demo/income-poverty/data-extracts.html>>) for additional details.

Key Estimates

Annual Estimates of Coverage

Table 1 compares the percentage of people with health insurance coverage in 2016 as estimated by the 2017 CPS ASEC Production File and by the 2017 CPS ASEC Research File.

The health insurance coverage rate in the 2017 research file is higher than in the production file. More specifically, the research file estimates 92.1 percent of people had health insurance coverage at some point in 2016, while the production file estimates 91.2 percent of people had coverage during this time period. The prevalence of coverage types also changed between files. The percentage of people who only had private coverage was higher in the research file than in the production file, while the percentage of people with government coverage during 2016 was lower in the research file.

Table 1. Health Insurance Coverage Numbers and Rates by Type of Coverage and Data File Source: 2016

Coverage type	2017 Research File (RF)				2017 Production File (P)				Difference (RF-P)	
	Number	MOE ¹ (±)	Rate	MOE ¹ (±)	Number	MOE ¹ (±)	Rate	MOE ¹ (±)	Number	Rate
Total.....	319,794	119	X	X	320,372	96	X	X	-578*	X
Any health plan.....	294,657	534	92.1	0.2	292,320	541	91.2	0.2	2,337*	0.9*
Any private plan ^{2,3}	217,638	1139	68.1	0.4	216,203	1,145	67.5	0.4	1,435*	0.6*
Employment-based ²	176,137	1241	55.1	0.4	178,455	1,130	55.7	0.4	-2,318*	-0.6*
Direct-purchase ²	37,819	698	11.8	0.2	51,961	874	16.2	0.3	-14,142*	-4.4*
..TRICARE.....	8,129	455	2.5	0.1	X	X	X	X	X	X
Any government plan ^{2,4}	110,714	978	34.6	0.3	119,361	1,018	37.3	0.3	-8,646*	-2.6*
Medicare ²	53,932	405	16.9	0.1	53,372	396	16.7	0.1	560*	0.2
Medicaid ²	60,461	963	18.9	0.3	62,303	931	19.4	0.3	-1,843*	-0.5*
Military excluding TRICARE ^{2,5}	3,054	174	1.0	0.1	X	X	X	X	X	X
..Military including TRICARE.....	10,951	499	3.4	0.2	14,638	575	4.6	0.2	-3,687	-1.1*
Uninsured⁶.....	25,137	498	7.9	0.2	28,052	519	8.8	0.2	-2,915*	-0.9*

Numbers in thousands, margins of error in thousands or percentage points as appropriate. Population as of March 2017. For information on confidentiality protection, sampling error, nonsampling error, and definitions, see www2.census.gov/programs-surveys/cps/techdocs/cpsmar17.pdf

X Not applicable.

* Changes between the estimates are statistically different from zero at the 90 percent confidence level.

¹A margin of error (MOE) is a measure of an estimate's variability. The larger the MOE in relation to the size of the estimate, the less reliable the estimate. This number, when added to and subtracted from the estimate, forms the 90 percent confidence interval. MOEs shown in this table are based on standard errors calculated using replicate weights. For more information, see "Standard Errors and Their Use" at www2.census.gov/library/publications/2018/demo/p60-264sa.pdf.

²The estimates by type of coverage are not mutually exclusive; people can be covered by more than one type of health insurance during the year.

³In the production file, private health insurance includes coverage provided through an employer or union, coverage purchased directly by an individual from an insurance company, or coverage through someone outside the household. In the research file, private health insurance also includes TRICARE.

⁴In the research file, government health insurance coverage includes Medicaid, Medicare, CHAMPVA (Civilian Health and Medical Program of the Department of Veterans Affairs), as well as care provided by the Department of Veterans Affairs and the military. In the production file, government health insurance also includes TRICARE.

⁵Military health care includes CHAMPVA, as well as care provided by the Department of Veterans Affairs and the military.

⁶Individuals are considered to be uninsured if they do not have health insurance coverage for the entire calendar year.

Source: U.S. Census Bureau, 2017 Current Population Survey Annual Social and Economic Supplement and 2017 Current Population Survey Annual Social and Economic Supplement Research File

Importantly, as noted earlier, these two estimates correspond to two slightly different populations. The research file excludes infants who were born before the time of interview but before the end of the 2016 calendar year. To highlight that these differences are not exclusively driven by differences in population of interest, Table A1 in the appendix replicates Table 1 but excludes this population of infants from the production file. With this comparison, the percentage of people with any coverage, private coverage, and public coverage differs between files once again.

As the data collected from respondents does not differ between the production and research files, these improvements to estimates of overall coverage rates are likely attributable to changes in the data processing system.⁷ Therefore, we now turn to estimates that highlight key aspects of changes to the processing system.

Changes in Cases Insurance Coverage

The production and research files can be linked to CPS Basic and other CPS supplements, as well as to each other. Data users who do this file-to-file linkage will see that some respondents' health insurance status and/or type of coverage differs between the two files. Table 2 summarizes the percent of cases that experienced a change in that health insurance type for cases with health insurance information in both files.⁸

Table 2. Person-level Differences in Health Insurance Coverage between the 2017 CPS ASEC Production File and 2017 CPS ASEC Research File

	% With Change	% Without Change
Insured	4.8	95.2
Private Insurance	9.2	90.8
Employer Sponsored Insurance	10.6	89.4
Direct Purchase	8.1	91.9
Government Coverage	8.4	91.6
Medicare	1.5	98.5
Medicaid	7.2	92.8
Military Coverage	2.4	97.6

Source: U.S. Census Bureau, 2017 Current Population Survey Annual Social and Economic Supplement and 2017 Current Population Survey Annual Social and Economic Supplement Research File

Overall, approximately 4.8% percent of the sample changed from having coverage during the calendar year to not having coverage during the calendar year or vice versa. Changes in an

⁷ Due to stochastic nature of matching in hotdeck imputation, there may be minor differences between two files using the same processing system due to chance.

⁸ Not all cases have health insurance information across the production and research files due to a change in universe across files and changes to in-sample criteria.

individual’s insurance coverage status may have occurred for a number of reasons. First, cleaning program changes allowed use of more information about respondents and their other question responses to help assign them a particular insurance type when information was incomplete or inconsistent. Second, changes to demographic and economic characteristics have affected the variables used to match cases as part of imputation procedures. Third, other refinements made to the imputation procedure allowed more respondent characteristics to be used when imputing health insurance coverage.

Differences in Outcomes for Imputed Cases

Population-level and person-level differences are expected in light of improvements to the way in which various questionnaire reports are prioritized and missing/incomplete data are imputed. People who had their health insurance information fully imputed were more likely to be classified as having health insurance coverage in the research file relative to the production file. In the research file, there is also a notable decline in the percentage of imputed cases that had both private and public coverage during 2016. In the research file, 9.5% of imputed cases had private and public coverage while in the production file, 20.3% of imputed cases had private and public coverage. Table 3 summarizes these findings.

Table 3. Insurance Coverage in 2016 by Imputation Status between the 2017 CPS ASEC Production File and 2017 CPS ASEC Research File

	Research File ¹				Production File ¹			
	Full Supplement Impute		Non-Full Supplement Impute		Full Supplement Impute		Non-Full Supplement Impute	
	%	SE ²	%	SE ²	%	SE ²	%	SE ²
Any Coverage	91.6	0.2	92.3	0.1	86.8	0.3	92.6	0.1
Private Coverage ³ Only	56.0	0.4	55.7	0.2	48.6	0.5	55.7	0.2
Public Coverage ³ Only	26.1	0.4	26.2	0.2	17.8	0.3	25.4	0.2
Private and Public Coverage	9.5	0.2	10.3	0.1	20.3	0.3	11.5	0.1

¹This table includes information for all respondents in the poverty and health insurance universes with data in both the research and production files.

²SE stands for standard error.

³Private and public coverage are defined to be consistent across files. Respondents are classified as having private coverage if they have direct purchase or employer-sponsored coverage. Respondents are classified as having public coverage if they have military, Medicaid, or Medicare coverage.

This preliminary evaluation suggests that the research file imputations improve data quality. In the research file, differences between reported and imputed cases are smaller than those observed in the production file.

Subannual Estimates of Coverage

To offer a balance between providing subannual health insurance data while ensuring respondents' confidentiality, the public-use 2017 ASEC research file contains a limited number of variables that indicate whether a person was covered for none, some, or all of the previous calendar year (in this case, 2016). Such subannual information is provided for any coverage, as well as for private coverage, government coverage, and Medicaid. Table 4 summarizes these frequencies and shows that, only a small percentage of people (less than 5 percent) had coverage for only part of the year.

Table 4. Duration of Selected Types of Health Insurance Coverage, 2016

	Duration of Health Insurance Coverage					
	No months		1-11 months		All 12 months	
	%	SE ³	%	SE ³	%	SE ³
Any coverage	7.9	0.1	4.2	0.1	88.0	0.1
Private Coverage ¹	31.9	0.2	3.7	0.1	64.4	0.2
Government Coverage ²	65.4	0.2	2.6	0.1	32.0	0.2
Medicaid	81.1	0.2	2.2	0.1	16.7	0.2

¹ Includes employer-based, direct-purchase, and TRICARE coverage

² Includes Medicare, Medicaid, and VA/CHAMPVA coverage

³ SE stands for standard error.

Source: U.S. Census Bureau, 2017 Current Population Survey Annual Social and Economic Supplement and 2017 Current Population Survey Annual Social and Economic Supplement Research File

Conclusion

The Census Bureau has made a number of refinements to data cleaning, imputation, and weighting procedures in the CPS ASEC. These changes, reflected in the 2017 CPS ASEC Research File have affected estimates of health insurance coverage. Early evaluation suggests that the data processing changes have improved data quality, but additional evaluation is ongoing.

References

- Andridge, R.R., and R.J.A. Little. 2010. "A Review of Hot Deck Imputation for Survey Non-response." *International Statistical Review* 78(1):40-64.
- Barnett, J.C., and E.R. Berchick. 2017. *Health Insurance Coverage in the United States: 2016*. Government Printing Office: Washington, DC.
- Barnett, J.C., and M.S. Vornovitsky. 2016. *Health Insurance Coverage in the United States: 2015*. Government Printing Office: Washington, DC.
- Berchick, E.R., E. Hood, and J. C. Barnett. 2018. *Health Insurance Coverage in the United States: 2017*. Government Printing Office: Washington, DC.
- Davern, M, H. Rodin, L.A. Blewett, and K.T. Call. 2007. "Are the Current Population Survey uninsurance estimates too high? An examination of the imputation process." *Health Services Research* 42(5): 2038-2055.
- Hess, J., J. Moore, J. Pascale, J. Rothgeb, and C. 2001. "The effects of person-level versus household-level questionnaire design on survey estimates and data quality." *Public Opinion Quarterly* 65(4): 574-584.
- Kenney, G., and V. Lynch. 2010. "Monitoring children's health insurance coverage under CHIPRA using federal surveys." In *Databases for Estimating Health Insurance Coverage for Children: A Workshop Summary*. Washington, DC: The National Academies Press.
- Pascale, J. 1999. "Methodological Issues in Measuring the Uninsured." *Proceedings of the Seventh Health Survey Research Methods Conference*, pp. 167–173.
- Smith, J.C., and C. Medalia. 2014. *Health Insurance Coverage in the United States: 2013*. Government Printing Office: Washington, DC.
- Smith, J.C., and c. Medalia. 2015. *Health Insurance Coverage in the United States: 2014*. Government Printing Office: Washington, DC.
- State Health Access Data Assistance Center. 2012. "Defining Family for Studies of Health Insurance Coverage." Issue Brief #27. Minneapolis, MN: University of Minnesota.

APPENDIX

Table A1. Health Insurance Coverage Numbers and Rates by Type of Coverage and Data File Source, Excluding Infants Born After the Calendar Year: 2016

Coverage type	2017 Research File (RF)				2017 Production File (P)				Difference (RF-P)	
	Number	MOE ¹ (±)	Rate	MOE ¹ (±)	Number	MOE ¹ (±)	Rate	MOE ¹ (±)	Number	Rate
Total	319,794	119	X	X	319,634	108	X	X	160	X
Any health plan	294,657	534	92.1	0.2	291,875	549	91.3	0.2	2,782*	0.8*
Any private plan excluding TRICARE ^{2,3}	210,679	1,203	65.9	0.4	215,997	1,145	67.6	0.4	-5,317*	-1.7*
Any private plan including TRICARE ^{2,3}	217,638	1,139	68.1	0.4	X	X	X	X	X	X
Employment-based ²	176,137	1,241	55.1	0.4	178,277	1,128	55.8	0.4	-2,140*	-0.7*
Direct-purchase ²	37,819	698	11.8	0.2	51,919	874	16.2	0.3	-14,101*	-4.4*
TRICARE.....	8,129	455	2.5	0.1	X	X	X	X	X	X
Any government plan including TRICARE ^{2,4}	116,421	1,060	36.4	0.3	119,065	1,014	37.3	0.3	-2,644*	-0.8*
Any government plan excluding TRICARE ^{2,4}	110,714	978	34.6	0.3	X	X	X	X	X	X
Medicare ²	53,932	405	16.9	0.1	53,367	396	16.7	0.1	565*	0.2
Medicaid ²	60,461	963	18.9	0.3	62,027	927	19.4	0.3	-1,567*	-0.5*
Military excluding TRICARE ^{2,5}	3,054	174	1.0	0.1	X	X	X	X	X	X
Military including TRICARE.....	10,951	499	3.4	0.2	14,610	573	4.6	0.2	-3,659*	-1.1*
Uninsured⁶	25,137	498	7.9	0.2	27,759	516	8.7	0.2	-2,662*	-0.8*

Numbers in thousands, margins of error in thousands or percentage points as appropriate. Population as of March of the following year. For information on confidentiality protection, sampling error, nonsampling error, and definitions, see www2.census.gov/programs-surveys/cps/techdocs/cpsmar17.pdf
 X Not applicable.

* Changes between the estimates are statistically different from zero at the 90 percent confidence level.

¹A margin of error (MOE) is a measure of an estimate's variability. The larger the MOE in relation to the size of the estimate, the less reliable the estimate. This number, when added to and subtracted from the estimate, forms the 90 percent confidence interval. MOEs shown in this table are based on standard errors calculated using replicate weights.

²The estimates by type of coverage are not mutually exclusive; people can be covered by more than one type of health insurance during the year.

³In the production file, private health insurance includes coverage provided through an employer or union, coverage purchased directly by an individual from an insurance company, or coverage through someone outside the household. In the research file, private health insurance also includes TRICARE.

⁴In the research file, government health insurance coverage includes Medicaid, Medicare, CHAMPVA (Civilian Health and Medical Program of the Department of Veterans Affairs), as well as care provided by the Department of Veterans Affairs and the military. In the production file, government health insurance also includes TRICARE.

⁵Military health care includes CHAMPVA, as well as care provided by the Department of Veterans Affairs and the military.

⁶Individuals are considered to be uninsured if they do not have health insurance coverage for the entire calendar year.

Source: U.S. Census Bureau, 2017 Current Population Survey Annual Social and Economic Supplement and 2017 Current Population Survey Annual Social and Economic Supplement Research File