

**MEPS HC-230
2021 Food Security**

July 2023

**Agency for Healthcare Research and Quality
Center for Financing, Access, and Cost Trends
5600 Fishers Lane
Rockville, MD 20857
(301) 427-1406**

Table of Contents

<u>Section</u>	<u>Page</u>
A. Data Use Agreement	A-1
B. Background	B-1
1.0 Household Component.....	B-1
2.0 Medical Provider Component.....	B-1
3.0 Survey Management and Data Collection	B-2
C. Technical and Programming Information.....	C-1
1.0 General Information.....	C-1
2.0 Data File Information.....	C-1
2.1 Codebook Structure	C-2
2.2 Reserved Codes	C-2
2.3 Codebook Format	C-2
2.4 Variable Naming	C-3
2.5 File Contents.....	C-3
2.5.1 Survey Administration Variables (HOMEIDX - RULETR42)	C-3
2.5.2 Food Security Variables (FSOUT42 - FSNEY42)	C-4
2.6 Linking to Other Files	C-5
2.6.1 MEPS Public Use Files	C-5
2.6.2 National Health Interview Survey.....	C-5
2.6.3 Longitudinal Analysis	C-5
3.0 Survey Sample Information	C-6
3.1 Discussion of Pandemic Effects on Quality of 2021 MEPS Data	C-6
3.1.1 Summary	C-6
3.1.2 Overview	C-6

<u>Section</u>	<u>Page</u>	
3.1.3	Data Quality Issues for MEPS in 2021 Directly Associated with Data Quality Concerns for the NHIS and CPS.....	C-6
3.1.4	Modifications to the MEPS HC 2021 Sample Design	C-7
3.1.5	Data Quality Issues for MEPS for FY 2021.....	C-8
3.2	Background on Sample Design and Response Rates	C-9
3.2.1	MEPS-Linked to the National Health Interview Survey (NHIS).....	C-10
3.2.2	Sample Weights and Variance Estimation	C-12
3.3	The MEPS Sampling Process and Response Rates: An Overview	C-12
3.3.1	Response Rates	C-15
3.3.2	Panel 26 Response Rates.....	C-16
3.3.3	Panel 25 Response Rates.....	C-16
3.3.4	Panel 24 Response Rates.....	C-17
3.3.5	Panel 23 Response Rates.....	C-17
3.3.6	Annual (Combined Panel) Response Rate	C-17
3.3.7	Oversampling	C-17
3.4	Food Security Weight (FSWT42)	C-19
3.4.1	Background and Target Population.....	C-19
3.4.2	Development of the Food Security Weights	C-19
3.5	Variance Estimation	C-20
3.5.1	Taylor-series Linearization Method.....	C-20
3.5.2	Balanced Repeated Replication (BRR) Method.....	C-22

<u>Section</u>		<u>Page</u>
	3.6 Using MEPS Data for Trend Analysis	C-23
	References.....	C-25
D.	Variable-Source Crosswalk	D-1

A. Data Use Agreement

Individual identifiers have been removed from the micro-data contained in these files. Nevertheless, under sections 308 (d) and 903 (c) of the Public Health Service Act (42 U.S.C. 242m and 42 U.S.C. 299 a-1), data collected by the Agency for Healthcare Research and Quality (AHRQ) and/or the National Center for Health Statistics (NCHS) may not be used for any purpose other than for the purpose for which they were supplied; any effort to determine the identity of any reported cases is prohibited by law.

Therefore in accordance with the above referenced Federal Statute, it is understood that:

1. No one is to use the data in this data set in any way except for statistical reporting and analysis; and
2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) the Director Office of Management AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity; and
3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel Survey or the National Health Interview Survey. Furthermore, linkage of the Medical Expenditure Panel Survey and the National Health Interview Survey may not occur outside the AHRQ Data Center, NCHS Research Data Center (RDC) or the U.S. Census RDC network.

By using these data you signify your agreement to comply with the above stated statutorily based requirements with the knowledge that deliberately making a false statement in any matter within the jurisdiction of any department or agency of the Federal Government violates Title 18 part 1 Chapter 47 Section 1001 and is punishable by a fine of up to \$10,000 or up to 5 years in prison.

The Agency for Healthcare Research and Quality requests that users cite AHRQ and the Medical Expenditure Panel Survey as the data source in any publications or research based upon these data.

B. Background

1.0 Household Component

The Medical Expenditure Panel Survey (MEPS) provides nationally representative estimates of health care use, expenditures, sources of payment, and health insurance coverage for the U.S. civilian noninstitutionalized population. The MEPS Household Component (HC) also provides estimates of respondents' health status, demographic and socio-economic characteristics, employment, access to care, and satisfaction with health care. Estimates can be produced for individuals, families, and selected population subgroups. The panel design of the survey includes five rounds of interviews covering two full calendar years. Additional rounds were added in 2020 and 2021, covering a third and fourth year respectively, to compensate for the smaller number of completed interviews in later panels. These extra rounds provide data for examining person level changes in selected variables such as expenditures, health insurance coverage, and health status. Using computer assisted personal interviewing (CAPI) technology, information about each household member is collected, and the survey builds on this information from interview to interview. All data for a sampled household are reported by a single household respondent.

The MEPS HC was initiated in 1996. Each year a new panel of sample households is selected. Because the data collected are comparable to those from earlier medical expenditure surveys conducted in 1977 and 1987, it is possible to analyze long-term trends. Each annual MEPS HC sample size is about 15,000 households. Data can be analyzed at either the person or event level. Data must be weighted to produce national estimates.

The set of households selected for each panel of the MEPS HC is a subsample of households participating in the previous year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The NHIS sampling frame provides a nationally representative sample of the U.S. civilian noninstitutionalized population. In 2006, the NHIS implemented a new sample design, which included Asian persons in addition to households with Black and Hispanic persons in the oversampling of minority populations. NHIS introduced a new sample design in 2016 that discontinued oversampling of these minority groups.

2.0 Medical Provider Component

Upon completion of the household CAPI interview and obtaining permission from the household survey respondents, a sample of medical providers are contacted by telephone to obtain information that household respondents cannot accurately provide. This part of the MEPS is called the Medical Provider Component (MPC) and information is collected on dates of visits, diagnosis and procedure codes, charges and payments. The Pharmacy Component (PC), a subcomponent of the MPC, does not collect charges or diagnosis and procedure codes but does collect drug detail information, including National Drug Code (NDC) and medicine name, as well as amounts of payment. The MPC is not designed to yield national estimates. It is primarily used as an imputation source to supplement/replace household reported expenditure information.

3.0 Survey Management and Data Collection

MEPS HC and MPC data are collected under the authority of the Public Health Service Act. Data are collected under contract with Westat, Inc. (MEPS HC) and Research Triangle Institute (MEPS MPC). Data sets and summary statistics are edited and published in accordance with the confidentiality provisions of the Public Health Service Act and the Privacy Act. The National Center for Health Statistics (NCHS) provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of micro data files and tables via the [MEPS website](#) and datatools.ahrq.gov.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857 (301-427-1406).

C. Technical and Programming Information

1.0 General Information

This documentation describes the 2021 food security data file from the Medical Expenditure Panel Survey Household Component (MEPS HC). Released as an ASCII file (with related SAS, SPSS, Stata, and R programming statements and data user information), a SAS dataset, a SAS transport dataset, a Stata dataset, and an Excel file, this public use file provides information collected on a nationally representative sample of the civilian noninstitutionalized population of the United States for calendar year 2021. The file contains 17 variables and has a logical record length of 57 with an additional 2-byte carriage return/line feed at the end of each record.

This file consists of MEPS survey data obtained in Round 8 of Panel 23, Round 6 of Panel 24, Round 4 of Panel 25, and Round 2 of Panel 26, and contains variables pertaining to food security.

The following documentation offers a brief overview of the types and levels of data provided, content and structure of the files, and programming information. It contains the following sections:

- Data File Information
- Survey Sample Information
- Variable-Source Crosswalk

Both weighted and unweighted frequencies of most variables included in the 2021 food security data file are provided in the accompanying codebook file. The exceptions to this are weight variables and variance estimation variables. Only unweighted frequencies of these variables are included in the accompanying codebook file. See the Weights Variables list in Section D, Variable-Source Crosswalk.

A database of all MEPS products released to date can be found on the [MEPS website](#).

2.0 Data File Information

This public use dataset contains variables and frequency distributions associated with 12,714 households who participated in the MEPS Household Component of the Medical Expenditure Panel Survey in 2021. These households received a positive family-level weight and were part of one of the four MEPS panels for whom food security data were collected in Round 8 of Panel 23, Round 6 of Panel 24, Round 4 of Panel 25, or Round 2 of Panel 26.

2.1 Codebook Structure

The codebook and data file sequence lists variables in the following order:

- Unique household identifiers and survey administration variables
- Food security variables
- Weight and variance estimation variables

2.2 Reserved Codes

The following reserved code values are used:

Value	Definition
-1 INAPPLICABLE	Question was not asked due to skip pattern
-7 REFUSED	Question was asked and respondent refused to answer question
-8 DK	Question was asked and respondent did not know answer or the information could not be ascertained
-15 CANNOT BE COMPUTED	Value cannot be derived from data

As part of the MEPS instrument design change in Spring of 2018, -9 (NOT ASCERTAINED) was removed from the MEPS instrument. This affected responses starting in Panel 23 Round 1, Panel 22 Round 3, and Panel 21 Round 5 and will continue in subsequent panels and rounds. Cases that used to contain -9 (NOT ASCERTAINED) in MEPS variables are now distributed between -8 (DK) and -15 (CANNOT BE COMPUTED). Most of the cases that were previously -9 (NOT ASCERTAINED) will now be assigned -8 (DK). However, -15 (CANNOT BE COMPUTED) will be assigned for MEPS variables that are constructed from MEPS instrument variables in cases where there is not enough information from the MEPS instrument to calculate the constructed MEPS variables. “Lack of information” is often the result of skip patterns in the data or from missing information resulting from -7 (REFUSED) or -8 (DK). Also note that reserved code -8 previously identified cases where respondent chose “don’t know” to a question. It now represents a broader category that includes cases where either the information from the question was “not ascertained” or the respondent chose “don’t know”.

2.3 Codebook Format

This codebook describes an ASCII data set (although the data are also being provided in a SAS data set, SAS transport file, Stata data set, and Excel file) and provides the following programming identifiers for each variable:

Variable Programming Identifiers

Identifier	Description
Name	Variable name
Description	Variable descriptor
Format	Number of bytes
Type	Type of data: numeric (indicated by NUM) or character (indicated by CHAR)
Start	Beginning column position of variable in record
End	Ending column position of variable in record

2.4 Variable Naming

Variable names reflect the content of the variable, with an eight-character limitation. Historically round dates have been indicated by two numbers following the variable name; the first number representing the round for second panel persons (Panel 25), the second number representing the round for first panel persons (Panel 26). The variable names in this 2021 file have not been renamed from prior years, despite the addition of Round 8 of Panel 23 and Round 6 of Panel 24, and the round numbers (8) and (6) are not included in the variable names.

Variables contained in this delivery were derived either from the questionnaire itself or from the CAPI. The source of each variable is identified in Appendix 1 “Variable-Source Crosswalk.” Sources for each variable are indicated in one of three ways: (1) variables derived from CAPI or assigned in sampling are so indicated; (2) variables collected at one or more specific questions have those numbers and questionnaire sections indicated in the “SOURCE” column; and (3) variables constructed from multiple questions using complex algorithms are labeled “Constructed” in the “SOURCE” column.

2.5 File Contents

2.5.1 Survey Administration Variables (HOMEIDX - RULETR42)

HOMEIDX uniquely identifies each household on the file and consists of the Dwelling Unit ID (DUID) followed by the Reporting Unit (RU) letter and round number.

The definitions of Dwelling Units (DUs) in the MEPS Household Survey are generally consistent with the definitions employed for the National Health Interview Survey (NHIS). The DUID is a seven-digit number consisting of a 2-digit panel number followed by a five-digit random number assigned after the case was sampled for MEPS.

PANEL is a constructed variable used to specify the panel number for the person. PANEL will indicate Panel 23, Panel 24, Panel 25, or Panel 26 for each person on the file. Panel 23 is the Panel that started in 2018, Panel 24 is the panel that started in 2019, Panel 25 is the panel that started in 2020, and Panel 26 is the panel that started in 2021.

An RU is a person or group of persons in the sampled DU who are related by blood, marriage, adoption, or other family association. Each RU was interviewed as a single entity for MEPS. Thus, the RU serves chiefly as a family-based “survey” operations unit rather than an analytic unit. Members of each RU within the DU in Round 8, Round 6, Round 4, or Round 2 are identified in the variable RULETR42.

Households are eligible for the Food Security PUF if the MEPS interview was completed by an RU member and if the household is not a student RU.

2.5.2 Food Security Variables (FSOUT42 - FSNEY42)

Respondents were asked:

- FSOUT42 - how often in the last 30 days anyone in the household worried whether food would run out before getting money to buy more
- FSLAST42 - how often in the last 30 days the food purchased didn't last and the person/household didn't have money to get more
- FSAFRD42 - how often in the last 30 days the person/household could not afford to eat balanced meals
- FSSKIP42 - in the last 30 days did the person/household reduce or skip meals because there wasn't enough money for food (coded as “-1 Inapplicable” when FSOUT42, FSLAST42, and FSAFRD42 = 3, -7, -8, or -15)
- FSSKDY42 - how many meals were skipped in the last 30 days (coded as “-1 Inapplicable” when FSSKIP42 = 2, -7, -8, or -15 OR when FSOUT42, FSLAST42, and FSAFRD42 = 3, -7, -8, or -15)
- FSLESS42 - in the last 30 days did the person/household ever eat less because there wasn't enough money for food (coded as “-1 Inapplicable” when FSOUT42, FSLAST42, and FSAFRD42 = 3, -7, -8, or -15)
- FSHGRY42 - in the last 30 days was the person/household ever hungry but didn't eat because there wasn't enough money for food (coded as “-1 Inapplicable” when FSOUT42, FSLAST42, and FSAFRD42 = 3, -7, -8, or -15)
- FSWTLS42 - in the last 30 days did anyone in the household lose weight because there wasn't enough money for food (coded as “-1 Inapplicable” when FSOUT42, FSLAST42, and FSAFRD42 = 3, -7, -8, or -15)
- FSNEAT42 - in the last 30 days did anyone in the household not eat for a whole day because there wasn't enough money for food (coded as “-1 Inapplicable” when FSOUT42, FSLAST42, and FSAFRD42 = 3, -7, -8, or -15; or when FSLESS42, FSHGRY42, and FSWTLS42 = 2, -7, -8, or -15)

FSNEDY42 - how many days in the last 30 days anyone in the household had not eaten for a whole day because there wasn't enough money for food (coded as "-1 Inapplicable" when FSOUT42, FSLAST42, and FSAFRD42 = 3, -7, -8, or -15; or when FSLESS42, FSHGRY42, and FSWTLS42 = 2, -7, -8, or -15; or when FSNEAT42 = 2, -7, -8, or -15)

2.6 Linking to Other Files

2.6.1 MEPS Public Use Files

This Food Security file can be linked to the 2021 full year Consolidated file by DUID and RULETR42 to obtain additional data for the families included in this file. The reference person of the RU can be identified in the Consolidated data file by the variable REFPRS42.

2.6.2 National Health Interview Survey

The set of households selected for MEPS is a subsample of those participating in the National Health Interview Survey (NHIS), thus, each MEPS panel can also be linked back to the previous year's NHIS public use data files. For information on obtaining MEPS/NHIS link files please see the [AHRQ website](#).

2.6.3 Longitudinal Analysis

Panel-specific longitudinal files are available for downloading in the data section of the MEPS website. For all four panels (Panel 23, Panel 24, Panel 25, and Panel 26), the longitudinal file comprises MEPS survey data obtained in all rounds of the panel and can be used to analyze changes over the entire length of the panel. For Panel 24, a file representing a three-year period will also be established and updated to cover four years with the release of 2022 data. For Panel 23, a file representing a four-year period will be established. Variables in the file pertaining to survey administration, demographics, employment, health status, disability days, quality of care, patient satisfaction, health insurance, and medical care use and expenditures were obtained from the MEPS full-year Consolidated files from the years covered by each panel.

For more details or to download the data files, please see Longitudinal Weight Files at the [AHRQ website](#).

3.0 Survey Sample Information

3.1 Discussion of Pandemic Effects on Quality of 2021 MEPS Data

3.1.1 Summary

The challenges associated with MEPS data collection in 2020 after the onset of the COVID-19 pandemic continued into 2021. The major modifications to the standard MEPS study design remained in effect, permitting data to be collected safely but with accompanying concerns related to the quality of the data obtained. These data quality issues are discussed below. The suggestion made in the documentation for the FY 2020 MEPS Consolidated PUF data (as well as for most federal major in-person surveys conducted in 2021 and 2020) still holds. Researchers are counseled to take care in the interpretation of estimates based on data collected from these two calendar years. This includes the comparison of such estimates to those of other years and corresponding trend analyses.

3.1.2 Overview

Section 3.1 of the documentation for the [2020 Full Year Consolidated Data File](#) provides a general discussion of the impact of the COVID-19 pandemic on several other major in-person federal surveys as well as on MEPS. In addition, it offers a detailed look at how MEPS was modified to permit safe data collection and the development of useful estimates at a time when the way the U.S. health care system functioned underwent many transformations in order to meet population needs.

In this corresponding 2021 document, focus is placed mostly on MEPS data quality in 2021. However, it also includes how data quality issues related to the two federal surveys most closely connected to it, the National Health Interview Survey (NHIS) carried out by the National Center for Health Statistics (NCHS) and the Current Population Survey (CPS) carried out by the Census Bureau, have an impact on the data quality issues of MEPS.

Specifically, the following discussion describes: 1) data quality issues experienced by the NHIS and CPS that affect MEPS; 2) modifications to the MEPS sample design in 2021 due to the continuing pandemic; and 3) potential data quality issues in the FY 2021 MEPS data related to the COVID-19 pandemic.

3.1.3 Data Quality Issues for MEPS in 2021 Directly Associated with Data Quality Concerns for the NHIS and CPS

Households fielded for Round 1 of MEPS in each year have been selected as a subsample from among the NHIS responding households from the prior year. The MEPS first year panel in 2021 was Panel 26. The households fielded for MEPS in Round 1 of Panel 26 were thus selected from NHIS responding households in 2020. It is important to note here that the NHIS households eligible for use in MEPS are restricted to the first three quarters of the NHIS as the fourth quarter

households cannot be made available in time for MEPS data collection early in the next calendar year.

The onset of the pandemic in 2020 at a national level took place in mid-March of that year, when the NHIS data collection for the first quarter of 2020 was virtually completed and that of the second quarter was about to begin. The NHIS had to make a rapid transition from in-person to telephone interviewing in order to attempt to gather NHIS data for the second quarter of 2020. While NCHS was able to make the transition, assessments made by NCHS at the time indicated a much lower response rate than is typically experienced during Quarter 2 and the quality of Quarter 2 data was of particular concern. NCHS thus modified the 2020 NHIS sample design for Quarters 3 and 4. A randomly selected subsample of the sampled housing units originally selected for fielding in Quarters 3 and 4 of 2020 was removed from the sample to be fielded. This reduced sample for Quarters 3 and 4 was then enhanced by randomly selecting responding households from the 2019 NHIS for interviewing in 2020 as well. In consideration of the data quality issues and sample design modifications associated with the 2020 NHIS, the MEPS sample design for FY 2021 was modified, as will be discussed shortly.

With respect to the CPS, the quality of CPS data is always of particular importance to MEPS as March CPS-ASEC estimates serve as the basis of control totals for the raking component of the MEPS weighting process. These control totals incorporate the following demographic variables: age, sex, race/ethnicity, region, MSA status, educational attainment, and poverty status. The CPS estimates of educational attainment and poverty status used in the development of the FY 2021 MEPS PUFs were of particular concern. Evaluations of these estimates undertaken by the Census Bureau have shown that they suffered from bias due to survey nonresponse with CPS income estimates being on the high side and the estimate of those under poverty being on the low side. The impact of these CPS estimates on the quality of MEPS estimates has been carefully considered. The approach used for the MEPS Full Year 2021 Consolidated PUF sample weights is discussed in Section 3.3.

A set of references (Bramlett et al., 2021; Dahlhamer et al., 2021; Lau et al, 2021; Rothbaum & Bee, 2021, 2022; Zuvekas & Kashihara, 2021) discussing the fielding of these surveys during the pandemic and possible bias concerns, can be found in the References section of this document.

3.1.4 Modifications to the MEPS HC 2021 Sample Design

Two key factors were thus expected to raise issues with MEPS plans for fielding a 2021 sample. First, 2020 NHIS data quality and sample size issues were of particular concern for Quarter 2 of that year. Second, roughly half of the NHIS sampled households for Quarter 3 would also have been respondents in the 2019 NHIS so that many of the Quarter 3 NHIS respondents were expected to have already been sampled and fielded for Panel 25 of MEPS. It thus became clear that it would be prudent to modify the 2021 MEPS sample design for MEPS Panel 26. Action had to be taken immediately because the MEPS sample selection from NHIS responding households begins in the late summer/early fall of each year.

AHRQ contacted NCHS, reviewing the various issues and asking if it would be possible that responding households in NHIS Panels 2 and 4 from Quarter 1 of 2020 be made available for

MEPS sample selection. Virtually all of these households were interviewed in-person prior to the major onset of the pandemic, so the Quarter 1 response rates for all four NHIS panels were consistent with prior years and the data quality issues associated with the pandemic could be avoided. NCHS was fully supportive of this approach and made NHIS Panels 2 and 4 for Quarter 1 available for use by MEPS. Thus, for MEPS Panel 26, the NHIS responding households subsampled from MEPS were selected from among all NHIS responding households in Quarter 1 as well as those responding in Quarter 3 that were not originally sampled for the 2019 NHIS.

As an adjunct to this modification, it was decided to take advantage of the additional PSUs (sampled localities) available from NHIS Panels 2 and 4 and appearing in the MEPS sample for the first time. State level estimation is of interest to MEPS, and the added PSUs would serve to increase the precision for state level estimates. State estimates that would be expected to benefit the most from these added PSUs were the “middle-sized” states. The largest states already had large sample sizes while precision for the smallest states would remain low. As a result, the MEPS sample focused on oversampling the “middle-sized” states rather than Hispanics, Blacks, and Asians, as has usually been the practice.

Finally, it was decided to collect data for Panels 23 and 24 for nine rounds, so that these two panels will ultimately contribute to MEPS estimates for four calendar years. In so doing, the number of respondents to MEPS will be kept at a relatively high level despite the decline in response rates due to the pandemic. The MEPS FY 2021 PUF records thus consist of data obtained from the following MEPS Panels and corresponding rounds: Panel 23, Rounds 7-9; Panel 24, Rounds 5-7; Panel 25, Rounds 3-5; and Panel 26, Rounds 1-3.

3.1.5 Data Quality Issues for MEPS for FY 2021

Three sources of potential bias were identified for MEPS for FY 2020: long recall period for Round 6 of Panel 23; switching from in-person to telephone interviewing which likely had a larger impact on Panel 25; and the impact of CPS bias on the MEPS weights. A number of statistically significant differences were found between panels for FY 2020. Those findings are discussed in MEPS HC-224.

With this in mind, there were a number of uncertainties for FY 2021 warranting examination. Would Panel 23 data quality increase substantially once the issue of an extensive recall period was eliminated? Would event reporting continue to be generally higher in Panel 25 compared to other panels? Since Panel 26 was the first year MEPS panel in 2021, would Panel 26 estimates tend to be different than those of the other three panels?

Preliminary analyses undertaken to examine the quality of MEPS FY 2021 data appearing on the Full Year 2021 Consolidated PUF have been focused on the comparison of health insurance status distribution (some private insurance, some public insurance, no health insurance) for the MEPS target population between the panels fielded. These comparisons were undertaken for the full sample and the three age groups of 0-17, 18-64, and 65+.

The analyses undertaken thus far suggest no major differences between the four panels for the distribution of health insurance status. Even though slight differences were observed with Panel 25 (e.g., the distribution associated with the age range 18-64 showed a higher percentage of all

public insurance compared to the other three panels while those at least 65 years of age showed a lower percentage of some private insurance compared to the other three panels), no statistically significant differences were detected.

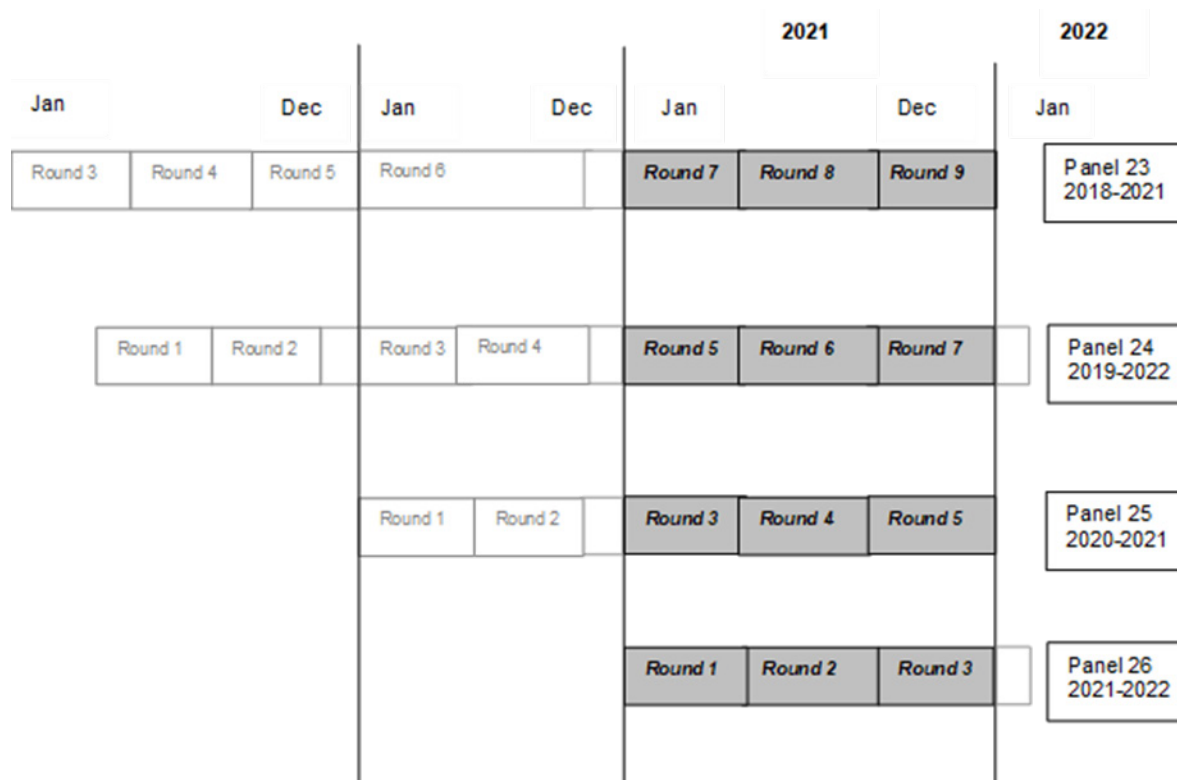
Further analyses of MEPS estimates will be conducted as part of the production of the FY 2021 Consolidated PUF to be released later in 2023..

3.2 Background on Sample Design and Response Rates

The MEPS is designed to produce estimates at the national and regional level over time for the civilian, noninstitutionalized population of the United States and some subpopulations of interest. The data in this public use file pertain to calendar year 2021. The data were collected in Rounds 1, 2, and 3 for MEPS Panel 26, Rounds 3, 4, and 5 for MEPS Panel 25, Rounds 5, 6 and 7 for MEPS Panel 24, and Rounds 7, 8 and 9 for MEPS Panel 23.

As usual, Round 3 for a MEPS panel (this time for Panel 26) has been designed to overlap two calendar years, as illustrated below. However it may be noted that, with the fielding of a third and fourth panel in 2021 (as indicated in the data quality discussion in Section 3.1), the structure of other rounds has changed. Round 7 of Panel 23 and Round 5 of Panel 24 serve the same purpose. Thus, Round 7 of Panel 23 was fielded in 2020 and designed to collect data for the remainder of 2020 as well as the period of time from January 1, 2021 up through the date of the Round 7 interview. Round 5 of Panel 24 was designed for the same purpose, collecting data associated with both 2020 and 2021. This was done to permit all three of these panels to provide data for the FY 2021 MEPS data sets as well as those for FY 2020.

For 2021, Panel 24 Round 6 represents the reference period from the date of the Round 6 interview back to January 1, 2021 (as discussed in the data quality subsection) and Panel 23 Round 8 represents the reference period from the date of the Round 8 interview back to January 1, 2021. The 2021 food security data were collected only in Round 8 of Panel 23, Round 6 of Panel 24, Round 4 of Panel 25, and Round 2 of Panel 26.



A sample design feature shared by Panel 23, Panel 24, and Panel 25 involved the partitioning of the sample domain “Other” (serving as the catchall stratum and consisting mainly of households with “White” members) into two sample domains. This was done for the first time in Panel 16. The two domains distinguished between those households characterized as “complete” respondents to the NHIS; and those characterized as “partial completes.” NHIS “partial completes” typically have a lower response rate to MEPS and for all three MEPS panels the “partial” domain was sampled at a lower rate than the “complete” domain. This approach has served to reduce survey costs, since the “partials” tend to have higher costs in gaining survey participation, but has also increased sample variance due to the resulting increased variance in sampling rates. Starting with Panel 25, the “Other, Partial” domain includes the NHIS households that have provided only a roster of household members. For detailed information on the MEPS sample design, see Chowdhury et al. (2019). This feature was not of particular emphasis in Panel 26, for reasons discussed in Section 3.1.4.

3.2.1 MEPS-Linked to the National Health Interview Survey (NHIS)

Each responding household found in this 2021 MEPS dataset is associated with one of the four separate and overlapping MEPS panels, MEPS Panel 23, MEPS Panel 24 MEPS Panel 25, and MEPS Panel 26. These panels consist of subsamples of households participating in the 2017, 2018, 2019, and 2020 NHIS, respectively. The Full Year 2018 PUF was the first one where both MEPS panels reflect the new NHIS sample design first implemented in 2016.

Whenever there is a change in sample or study design, it is good survey practice to assess whether such a change could affect the sample estimates. For example, increased coverage of the

target populations with an updated sample design based on data from the latest Census can improve the accuracy of the sample estimates. MEPS estimates have been and will continue to be evaluated to determine if an important change in the survey estimates might be associated with a change in design. Discussion on the potential effects of such MEPS design changes in 2021 appears in the data quality section, 3.1. Background on the two NHIS sample designs of interest here is provided next.

Background on the NHIS Sample Redesign Implemented in 2016

Beginning in 2016, NCHS implemented another new sample design for the NHIS, which differed substantially from the prior design. Each of the 50 states as well as the District of Columbia served as explicit strata for sample selection purposes with the intent of providing the capability of state-level NHIS estimates obtained through pooling across years if the sample size for a single year would result in unreliable estimates. In contrast to the previous design, households in areas with relatively high concentrations of minorities are not oversampled. PSUs are still formed at the county level. However, within sampled PSUs, the clusters of addresses that have been sampled for each year of the NHIS are not in the form of segments (consisting of one or more census blocks) as was done for the previous NHIS designs. For the 2016 NHIS, each such cluster consisted of roughly 25 subclusters selected using random systematic sampling across the full geography of the PSU. Each subcluster is made up of, generally, four nearby addresses or roughly 100 addresses in all. The number of subclusters per cluster can vary from year to year.

Another major change is that the list of DUs (addresses) was obtained from the Computerized Delivery Sequence File (CDSF) of the U.S. Postal Service, which is a different approach than the standard listing process for area probability samples used in the pre-2016 designs. While addresses in the CDSF provide very high coverage of most areas of the country, coverage in rural areas can be somewhat lower. For rural areas where this was a concern, address lists were created through the conventional listing process. A [description of the NHIS sample design](#) is found on the NHIS website.

Panel 23 Household Sample Size

A subsample of 9,700 households (occupied DUs) was randomly selected for MEPS Panel 23 from NHIS responding households in 2017, of which 9,694 were fielded for MEPS after the elimination of any units characterized as ineligible for fielding.

Panel 24 Household Sample Size

A subsample of 9,700 households was randomly selected for MEPS Panel 24 from the households responding to the 2018 NHIS, of which 9,684 were fielded for MEPS after the elimination of any units characterized as ineligible for fielding.

Panel 25 Household Sample Size

A subsample of 9,900 households was randomly selected for MEPS Panel 25 from the households responding to the 2019 NHIS, of which 9,888 were fielded for MEPS after the elimination of any units characterized as ineligible for fielding.

Panel 26 Household Sample Size

A subsample of 9,510 households was randomly selected for MEPS Panel 26 from the households responding to the 2020 NHIS, of which all 9,510 were fielded for MEPS since no units were characterized as ineligible for fielding.

Implications of the New Design on MEPS Estimates

Under the new design, MEPS sampled households reflect the clustering of the NHIS, as described above but to a somewhat lesser degree due to the sampling from NHIS respondents. Due to the spreading of the NHIS sample in small subclusters across the PSU and the sampling limited to only NHIS respondents, the impact of clustering on the variance of MEPS estimates may be more limited. Also, in contrast to the previous design, the NHIS sampling rates at the address level currently do not vary due to oversampling of minorities (although this could change in subsequent years). On balance, the overall variation in sampling rates/weights at the national level for the NHIS is expected to be lower with a corresponding positive impact on the precision of MEPS estimates. However, with a reduction in the sample sizes of minority households, precision levels of MEPS estimates for Hispanics, Blacks, and Asians may be reduced to some extent.

3.2.2 Sample Weights and Variance Estimation

In the dataset “MEPS HC-230: 2021 Food Security Data File,” a weight variable is provided for generating MEPS estimates of totals, means, percentages, and rates for families in the civilian noninstitutionalized population. Procedures and considerations associated with the construction and interpretation of family estimates using these and other variables are discussed below. It should be noted that the increase in eligible NHIS households for MEPS, discussed in section 3.1.4, has affected the MEPS variance structure. This is discussed in more detail in Section 3.5.1.

3.3 The MEPS Sampling Process and Response Rates: An Overview

For most MEPS panels, a sample representing about three-eighths of the NHIS responding households is made available for use in MEPS. This was the case for MEPS Panel 23, Panel 24, and Panel 25. For Panel 26, the sample made available for use in MEPS represented about five-eighths of the NHIS responding households for the reasons discussed in Section 3.1.4.

Because the MEPS subsampling has to be done soon after NHIS responding households are identified, a small percentage of the NHIS households initially characterized as NHIS

respondents are later classified as nonrespondents for the purposes of NHIS data analysis. This actually serves to increase the overall MEPS response rate slightly since the percentage of NHIS households designated for use in MEPS (all those characterized initially as respondents from the NHIS panels and quarters used by MEPS for a given year) is slightly larger than the final NHIS household-level response rate and some NHIS nonresponding households do participate in MEPS. However, as a result, these NHIS nonrespondents who are MEPS participants have no NHIS data available to link with MEPS data. Once the MEPS sample is selected from among the NHIS households characterized as NHIS respondents, RUs consisting entirely of military personnel are deleted from the sample. Military personnel not living in the same RU as civilians are ineligible for MEPS. After these exclusions, all RUs associated with households selected from among those identified as NHIS responding households are then fielded in the first round of MEPS.

Table 3.1 shows in Rows A, B, and C the three informational components just discussed. Row A indicates the percentage of NHIS households eligible for MEPS. Row B indicates the number of NHIS households sampled for MEPS. Row C indicates the number of sampled households actually fielded for MEPS (after dropping the military members discussed above and a small number of NHIS households sampled in error). Note that all response rates discussed here are unweighted.

Table 3.1. Sample Size and Unweighted Response Rates for 2021 Full Year File (Panel 26 Rounds 1-3/Panel 25 Rounds 3-5/Panel 24 Rounds 5-7/Panel 23 Rounds 7-9)

Components	Panel 23	Panel 24	Panel 25	Panel 26	2021 Combined
A. Percentage of NHIS households designated for use in MEPS (those initially characterized as responding) *	67.1%	64.3%	65.7%	60.6%	-
B. Number of households sampled from the NHIS	9,700	9,700	9,900	9,510	-
C. Number of Households sampled from the NHIS and fielded for MEPS	9,694	9,684	9,888	9,510	-
D. Round 1 - Number of RUs eligible for interviewing	10,089	10,090	10,152	9,795	-
E. Round 1 - Number of RUs with completed interviews	7,351	7,186	6,265	5,882	-
F. Round 2 - Number of RUs eligible for interviewing	7,492	7,323	5,958	6,045	-
G. Round 2 - Number of RUs with completed interviews	6,960	6,777	4,677	4,799	-
H. Round 3 - Number of RUs eligible for interviewing	7,082	6,890	5,189	4,876	-
I. Round 3 - Number of RUs with completed interviews	6,703	6,289	4,230	4,103	-
J. Round 4 - Number of RUs eligible for interviewing	6,776	6,371	4,307	-	-
K. Round 4 - Number of RUs with completed interviews	6,522	5,446	3,685	-	-
L. Round 5 - Number of RUs eligible for interviewing	6,536	5,495	3,706	-	-
M. Round 5 - Number of RUs with completed interviews	6,383	4,770	3,278	-	-
N. Round 6 - Number of RUs eligible for interviewing	6,480	4,808	-	-	-
O. Round 6 - Number of RUs with completed interviews	5,120	3,959	-	-	-
P. Round 7 - Number of RUs eligible for interviewing	5,170	4,002	-	-	-
Q. Round 7 - Number of RUs with completed interviews	4,513	3,500	-	-	-

Components	Panel 23	Panel 24	Panel 25	Panel 26	2021 Combined
R. Round 8 - Number of RUs eligible for interviewing	4,548	-	-	-	-
S. Round 8 - Number of RUs with completed interviews	3,984	-	-	-	-
T. Round 9 - Number of RUs eligible for interviewing	3,996	-	-	-	-
U. Round 9 - Number of RUs with completed interviews	3,603	-	-	-	-
Overall annual unweighted response rates P26: A x (E/D) x (G/F) x (I/H) P25: A x (E/D) x (G/F) x (I/H) x (K/J) x (M/L) P24: A x (E/D) x (G/F) x (I/H) x (K/J) x (M/L) x (O/N) x (Q/P) P23: A x (E/D) x (G/F) x (I/H) x (K/J) x (M/L) x (O/N) x (Q/P) x (S/R) x (U/T) Combined: $0.22 \times P23 + 0.22 \times P24 + 0.25 \times P25 + 0.31 \times P26$	22.0% (Panel 23 through Round 9)	20.7% (Panel 24 through Round 7)	19.6% (Panel 25 through Round 5)	24.3% (Panel 26 through Round 3)	21.8%

*Among the panels and quarters of the NHIS allocated to MEPS, the percentage of households that were considered to be NHIS respondents at the time the MEPS sample was selected.

3.3.1 Response Rates

To produce annual health care estimates for calendar year 2021 based on the full MEPS sample data from the MEPS Panel 23, Panel 24, Panel 25, and Panel 26, the four panels are combined. More specifically, full calendar year 2021 data collected in Rounds 7 through 9 for the MEPS Panel 23, Rounds 5, 6 and 7 for the MEPS Panel 24 and Rounds 3 through 5 for the MEPS Panel 25 sample are pooled with data from the first three rounds of data collection for the MEPS Panel 26 sample (the general approach is described below).

As mentioned above, all response rates discussed here are unweighted. To understand the calculation of MEPS response rates, some features related to MEPS data collection should be noted. When an RU is visited for a round of data collection, changes in RU membership are identified. Such changes include the formation of student RUs as well as other new RUs created when RU members from a previous round have moved to another location in the U.S. Thus, the number of RUs eligible for MEPS interviewing in a given round is determined after data collection is fully completed. The ratio of the number of RUs completing the MEPS interview in a given round to the number of RUs characterized as eligible to complete the interview for that round represents the “conditional” response rate for that round expressed as a proportion. It is “conditional” in that it pertains to the set of RUs characterized as eligible for MEPS for that

round and thus is “conditioned” on prior participation rather than representing the overall response rate through that round. For example, in Table 3.1, for Panel 26 Round 2 the ratio of 4,799 (Row G) to 6,045 (Row F) multiplied by 100 represents the response rate for the round (79.4 percent when computed), conditioned on the set of RUs characterized as eligible for MEPS for that round. Taking the product of the percentage of the NHIS sample eligible for MEPS (Row A) with the product of the ratios for a consecutive set of MEPS rounds beginning with Round 1 produces the overall response rate through the last MEPS round specified.

The overall unweighted response rate for 2021 for the combined sample after pooling the respondents across the four panels was obtained by computing the product of the compositing factor associated with each panel and the corresponding overall panel response rate and then summing the four products. Panel 26 represents about 31.5 percent of the combined sample size, Panel 25 represents about 24.7 percent of the combined sample size, Panel 24 represents about 22.1 percent, and Panel 23 represents the remaining 21.7 percent.

Thus, the combined response rate of 21.8 percent was computed as 0.22 times 22.0 (22.0 is the overall Panel 23 response rate through Round 9) plus 0.22 times 20.7 (20.7 is the overall Panel 24 response rate through Round 7) plus 0.25 times 19.6 (19.6 is the overall Panel 25 response rate through Round 5) plus 0.31 times 24.3 (24.3 is the overall Panel 26 response rate through Round 3.)

The overall response rate of 21.8 percent for 2021 is lower than that for 2020 (27.6 percent), reflecting the continued impact of the pandemic on data collection efforts.

3.3.2 Panel 26 Response Rates

For MEPS Panel 26 Round 1, 9,510 households were fielded in 2021 (Row C of Table 3.1), a randomly selected subsample of the households responding to the 2020 NHIS.

Table 3.1 shows the number of RUs eligible for interviewing in each Round of Panel 26 as well as the number of RUs completing the MEPS interview. Computing the individual round “conditional” response rates as described in Section 3.3.1 and then taking the product of these three response rates and the factor 60.6 (the percentage of the NHIS sampled households characterized as responding at the time of sample selection of households for MEPS) yields an overall response rate of 24.3 percent for Panel 26 through Round 3.

3.3.3 Panel 25 Response Rates

A total of 9,888 households were fielded in 2020 for MEPS Panel 25 (Row C of Table 3.1), a randomly selected subsample of the households responding to the 2019 National Health Interview Survey (NHIS).

Table 3.1 shows the number of RUs eligible for interviewing and the number completing the interview for all five rounds of Panel 25. The overall response rate for Panel 25 was computed in a similar fashion to that of Panel 26 but covering all five rounds of MEPS interviewing as well as

the factor representing the percentage of NHIS sampled households eligible for MEPS. The overall response rate for Panel 25 through Round 5 is 19.6 percent..

3.3.4 Panel 24 Response Rates

A total of 9,684 households were fielded in 2019 for MEPS Panel 24 (Row C of Table 3.1), a randomly selected subsample of the households responding to the 2018 NHIS.

Table 3.1 shows the number of RUs eligible for interviewing and the number completing the interview for all seven rounds of Panel 24. The overall response rate for Panel 24 was computed in a similar fashion to that of Panel 25 but covering all seven rounds of MEPS interviewing as well as the factor representing the percentage of NHIS sampled households eligible for MEPS. The overall response rate for Panel 24 through Round 5 is 20.7 percent.

3.3.5 Panel 23 Response Rates

A total of 9,694 households were fielded in 2019 for MEPS Panel 23 (Row C of Table 3.1), a randomly selected subsample of the households responding to the 2017 NHIS.

Table 3.1 shows the number of RUs eligible for interviewing and the number completing the interview for all nine rounds of Panel 23. The overall response rate for Panel 23 was computed in a similar fashion to that of Panel 24 but covering all nine rounds of MEPS interviewing as well as the factor representing the percentage of NHIS sampled households eligible for MEPS. The overall response rate for Panel 23 through Round 5 is 22.0 percent.

3.3.6 Annual (Combined Panel) Response Rate

A combined panel response rate for the survey respondents in this data set is obtained by taking a weighted average of the panel specific response rates. The Panel 23 response rate was weighted by a factor of 0.22, the Panel 24 response rate was weighted by a factor of 0.22, the Panel 25 response rate was weighted by a factor of 0.25, and the Panel 26 response rate was weighted by a factor of 0.31, reflecting approximately the distribution of the overall sample between the four panels. The resulting combined response rate for the combined panels was computed as $(0.22 \times 22.0) + (0.22 \times 20.7) + (0.25 \times 19.6) + (0.31 \times 24.3)$ or 21.8 percent (as shown in Table 3.1).

3.3.7 Oversampling

Oversampling is a feature of the MEPS sample design, helping to increase the precision of estimates for some subgroups of interest. Before going into details related to MEPS, the concept of oversampling will be discussed.

In a sample where all persons in a population are selected with the same probability and survey coverage of the population is high, the sample distribution is expected to be proportionate to the

population distribution. For example, if Hispanics represent 15 percent of the general population, one would expect roughly 15 percent of the persons sampled to be Hispanic. However, in order to improve the precision of estimates for specific subgroups of a population, one might decide to select samples from those subgroups at higher rates than the remainder of the population. Thus, one might select Hispanics at twice the rate (i.e., at double the probability) of persons not oversampled. As a result, an oversampled subgroup comprises a higher proportion of the sample than it represents in the general population. Sample weights ensure that population estimates are not distorted by a disproportionate contribution from oversampled subgroups. Base sample weights for oversampled groups will be smaller than for the portion of the population not oversampled. For example, if a subgroup is sampled at roughly twice the rate of sample selection for the remainder of the population not oversampled, members of the oversampled subgroup will receive base or initial sample weights (prior to nonresponse or poststratification adjustments) that are roughly half the size of the group not oversampled.

As mentioned above, oversampling is implemented to increase the sample sizes and thus improve the precision of survey estimates for particular subgroups of the population. The “cost” of oversampling is that the precision of estimates for the general population and subgroups not oversampled will be reduced to some extent compared to the precision one could have achieved if the same overall sample size were selected without any oversampling.

The NHIS no longer oversamples households with members who are Hispanic, Black, or Asian. Nevertheless, these minority groups are still of analytic interest for MEPS. As a result for Panels 23, 24, and 25, all households in the Hispanic, Black, and Asian domains were sampled with certainty (i.e., all households assigned to those domains were included in the MEPS). In addition, all households in Panel 23 who had a member who was a veteran were selected with certainty. Among all remaining households for Panel 23, the “Other, complete” domain was sampled at a rate of about 69 percent while the “Other, partial complete” domain was sampled at a rate of about 43 percent. For Panel 24, the corresponding sampling rates for the “Other, complete” domain and the “Other, partial complete” domain were about 79 percent and 50 percent, respectively. For Panel 25, the corresponding sampling rates for the “Other, complete” domain and the “Other, partial complete” domain were about 77 percent and 50 percent, respectively. The somewhat lower sampling rates for Panel 23 in the two “Other” domains arose due to the oversampling of veterans in that panel. With a specified overall sample size of 9,700, fewer were needed from those assigned to the “Other” domains in that panel.

Within the “noncertainty” strata (the “Other” domains) for the three panels, responding NHIS households were selected for MEPS using a systematic sample selection procedure from among those eligible. The selection of the households was with probability proportionate to size (pps) where the size measure was the inverse of the NHIS initial probability of selection. The pps sampling was undertaken to help reduce the variability in the MEPS weights incurred due to the variability of the NHIS sampling rates.

As discussed in Section 3.1.4, the Panel 26 sample design oversampled NHIS households from the “middle-sized” states rather than households with members who are Hispanic, Black, or Asian.

A note with respect to the interpretation of MEPS response rates, which are unweighted. Typically, sample allocations across sample domains change from one MEPS panel to another. The sample domains used may also vary by panel as is the case for Panel 23 versus Panel 24, Panel 25, and Panel 26. When one compares unweighted measures (e.g., response rates) between panels and years, one should take into account such differences. Suppose, for example, members of one domain have a lower propensity to respond than those of another domain. Then, if that domain has been allocated a higher proportion of the sample, the corresponding panel may have a lower unweighted response rate simply because of the differences in sample allocation.

3.4 Food Security Weight (FSWT42)

3.4.1 Background and Target Population

The Food Security questions are designed to learn more about existing food concerns among families in the U.S. These questions are incorporated into the MEPS survey instrument in the second round of data collection for each panel in a calendar year. Thus, for calendar year 2021, this took place in Round 8 of Panel 23, Round 6 of Panel 24, Round 4 of Panel 25, and Round 2 of Panel 26. To ensure that data reflected family circumstances as accurately as possible, data were collected from a member of the RU. For virtually all MEPS interviews, this was the RU respondent. However, a relative handful of MEPS interviews are conducted with proxy respondents. As a result, food security data were not collected for such RUs, and they are not part of the target population for the Food Security Weights. It may be noted that such families can be expected to be somewhat different from families generally. For example, a proxy respondent may be called for if his or her elderly parent was too sick to respond or had entered a nursing home.

Some RUs for which Food Security data were obtained may have gone out of scope prior to the end of 2021 while others may have become MEPS nonrespondents. As a result, the MEPS family weights established to reflect MEPS families in 2021 and appearing on the FY 2021 Consolidated PUF do not pertain to the target population associated with the Food Security weights. The Food Security weights were thus established with this in mind. The target population for the Food Security questions in 2021 can be described as MEPS families in the fall of 2021 that did not require a proxy respondent. As a reminder, single person RUs are considered a MEPS family as are partners who, though unrelated by marriage, consider themselves as a single family unit.

3.4.2 Development of the Food Security Weights

The weights for the 2021 Food Security data were established utilizing the MEPS family weights created for the FY 2021 Consolidated PUF which already compensated for MEPS nonresponse at the family level across MEPS rounds. To reflect such nonresponse for the Food Security weights, the initial Food Security weight assigned to an RU (MEPS family) was that of the weight of the corresponding MEPS family at the end of 2021, as established through the weighting of families for the FY 2021 Consolidated PUF. Specifically, these weights were assigned to each responding RU at Rounds 8/6/4/2 of Panels 23/24/25/26 where an RU member

completed the MEPS interview. Those MEPS families that responded in Rounds 8/6/4/2 but were a nonresponding RU at Round 9/7/5/3 would thus not receive a Food Security weight as they would not have received a MEPS family weight.

Proxy respondents at Rounds 8/6/4/2 were then removed from further consideration in the weighting process. Slightly under 1 percent of the MEPS family population at Rounds 8/6/4/2 had their MEPS data reported by proxy respondents.

Fewer than 20 RU respondents of the families otherwise eligible for the Food Security weights at Rounds 8/6/4/2 did not answer at least three of the 10 Food Security questions. These were treated as Food Security nonrespondents and a global adjustment factor was assigned to the weights of the respondents of the Food Security questions to determine the final value of weight variable FSWT42, the 2021 Food Security weight appearing on the 2021 Food Security PUF.

For information on the derivation of FAMWT21F, the weight variable representing the MEPS family population appearing on the 2021 Full Year Consolidated PUF, see MEPS HC-233, the corresponding PUF Documentation. Table 3.2 shows the number of families in the Food Security data file by panel and the weighted total number of the families.

Table 3.2. Numbers of families by Panel and the weighted total number of families

Panel 23	Panel 24	Panel 25	Panel 26	Combined	Population estimate (weighted total of combined sample)
3,162	3,094	2,810	3,648	12,714	139,142,489

3.5 Variance Estimation

To obtain estimates of variability (such as the standard error of sample estimates or corresponding confidence intervals) for MEPS estimates, analysts need to take into account the complex sample design of MEPS for both person-level and family-level analyses. Several methodologies have been developed for estimating standard errors for surveys with a complex sample design, including the Taylor-series linearization method, balanced repeated replication, and jackknife replication. Various software packages provide analysts with the capability of implementing these methodologies. MEPS analysts most commonly use the Taylor Series approach. Although this data file does not contain replicate weights, the capability of employing replicate weights constructed using the Balanced Repeated Replication (BRR) methodology is also provided if needed to develop variances for more complex estimators (see Section 3.5.2).

3.5.1 Taylor-series Linearization Method

The variables needed to calculate appropriate standard errors based on the Taylor-series linearization method are included on this and all other MEPS public use files. Software packages that permit the use of the Taylor-series linearization method include SUDAAN, R, Stata, SAS (version 8.2 and higher), and SPSS (version 12.0 and higher). For complete information on the

capabilities of a package, analysts should refer to the corresponding software user documentation.

Using the Taylor-series linearization method, variance estimation strata and the variance estimation PSUs within these strata must be specified. The variables VARSTR and VARPSU on this MEPS data file serve to identify the sampling strata and primary sampling units required by the variance estimation programs. Specifying a “with replacement” design in one of the previously mentioned computer software packages will provide estimated standard errors appropriate for assessing the variability of MEPS survey estimates. It should be noted that the number of degrees of freedom associated with estimates of variability indicated by such a package may not appropriately reflect the number available. For variables of interest distributed throughout the country (and thus the MEPS sample PSUs), one can generally expect to have at least 100 degrees of freedom associated with the estimated standard errors for national estimates based on this MEPS database.

Prior to 2002, MEPS variance strata and PSUs were developed independently from year to year, and the last two characters of the strata and PSU variable names denoted the year. Beginning with the 2002 Point-in-Time PUF, the approach changed with the intention that variance strata and PSUs would be developed to be compatible with all future PUFs until the NHIS design changed. Thus, when pooling data across years 2002 through the Panel 11 component of the 2007 files, the variance strata and PSU variables provided can be used without modification for variance estimation purposes for estimates covering multiple years of data. There are 203 variance estimation strata, each stratum with either two or three variance estimation PSUs.

From Panel 12 of the 2007 files, a new set of variance strata and PSUs were developed because of the introduction of a new NHIS design. There are 165 variance strata with either two or three variance estimation PSUs per stratum, starting from Panel 12. Therefore, there are a total of 368 (203+165) variance strata in the 2007 Full Year file as it consists of two panels that were selected under two independent NHIS sample designs. Since both MEPS panels in the Full Year files from 2008 through 2016 were based on the next NHIS design, there are only 165 variance strata. These variance strata (VARSTR values) have been numbered from 1001 to 1165 so that they can be readily distinguished from those developed under the former NHIS sample design if data are pooled for several years.

The NHIS sample design was changed again in 2016, effectively changing the MEPS design beginning with calendar year 2017. From Panel 22 of the 2017 files, a new set of variance strata and PSUs were developed. There are 117 variance strata with either two or three variance estimation PSUs per stratum. Therefore, there are a total of 282 (165+117) variance strata in the 2017 Full Year file as it consists of two panels that were selected under two independent NHIS sample designs. To make the pooling of data across multiple years of MEPS more straightforward, the numbering system for the variance strata has changed. Those strata associated with the new design were numbered from 2001 to 2117.

However, the NHIS sample design was further modified in 2018. With the modification in the 2018 NHIS sample design, the MEPS variance structure for the 2019 Full Year file was also modified, reducing the number of variance strata to 105. Consistency was maintained with the prior structure in that the 2019 Full Year file variance strata were also numbered within the range

of values from 2001-2117, although there are now gaps in the values assigned within this range. Due to the modification, each stratum could contain up to five variance estimation PSUs.

For Panel 26 in the 2021 Full Year file, additional NHIS sample was used for MEPS to account for increasing nonresponse during the pandemic (as discussed in Section 3.1.4). The additional sample was assigned to the existing variance strata, so the 2021 Full Year file continues to have 105 variance strata, numbered 2001-2117, with a few gaps in the values in that range. The additional sample was assigned to new variance estimation PSUs, so in the 2021 Full Year file, each stratum could contain up to eight variance estimation PSUs.

Some analysts may be interested in pooling data across multiple years of MEPS data. If pooling across years is to be undertaken, it should be noted that, to obtain appropriate standard errors when doing so, it is necessary to specify a common variance structure. Prior to 2002, each annual MEPS public use file was released with a variance structure unique to the particular MEPS sample in that year. Starting in 2002, the annual MEPS public use files were released with a common variance structure that allowed users to pool data from 2002 through 2018. However, with the need to modify the variance structure beginning with 2019, this can no longer be routinely done.

To ensure that variance strata are identified appropriately for variance estimation purposes when pooling MEPS data across several years, one can proceed as follows:

1. When pooling any year between 2002 through 2018, use the variance strata numbering as is.
2. When pooling (a) any year from 1996 to 2001 with any year from 2002 or later, or (b) the year 2019 and beyond with any earlier year, use the pooled linkage public use file HC-036 that contains the proper variance structure. The HC-036 file is updated every year so that appropriate variance structures are available with pooled data. Further details on the HC-036 file can be found in the public use documentation of the HC-036 file.

3.5.2 Balanced Repeated Replication (BRR) Method

BRR replicate weights are not provided on this MEPS PUF for the purposes of variance estimation. However, a file containing a BRR replication structure is made available so that the users can form replicate weights, if desired, from the final MEPS weight to compute variances of MEPS estimates using either BRR or Fay's modified BRR (Fay, 1989) methods. The replicate weights are useful to compute variances of complex non-linear estimators for which a Taylor linear form is not easy to derive and not available in commonly used software. For instance, it is not possible to calculate the variances of a median or the ratio of two medians using the Taylor linearization method. For these types of estimators, users may calculate a variance using BRR or Fay's modified BRR methods. However, it should be noted that the replicate weights have been derived from the final weight through a shortcut approach. Specifically, the replicate weights are not computed starting with the base weight and all adjustments made in different stages of weighting are not applied independently in each replicate. Thus, the variances computed using this one-step BRR do not capture the effects of all weighting adjustments that would be captured

in a set of fully developed BRR replicate weights. The Taylor Series approach does not fully capture the effects of the different weighting adjustments either.

The dataset HC-036BRR, MEPS 1996-2021 Replicates for Variance Estimation File, contains the information necessary to construct the BRR replicates. It contains a set of 128 flags (BRR1-BRR128) in the form of half sample indicators, each of which is coded 0 or 1 to indicate whether the person should or should not be included in that particular replicate. These flags can be used in conjunction with the full-year weight to construct the BRR replicate weights. For analysis of MEPS data pooled across years, the BRR replicates can be formed in the same way using the HC-036, MEPS 1996-2021 Pooled Linkage Variance Estimation File. For more information about creating BRR replicates, users can refer to the documentation for the [HC-036BRR pooled linkage file](#) on the AHRQ website.

3.6 Using MEPS Data for Trend Analysis

Note: While the Food Security PUF is only available starting in 2016, this section reviews major changes in MEPS that can affect trend analyses.

First, of course, we note that there are uncertainties associated with 2020 and 2021 data quality as discussed earlier in the data quality section (Section 3.1). Preliminary evaluations of a set of MEPS estimates of particular importance suggest that they are of reasonable quality. Nevertheless, analysts are advised to exercise caution in interpreting these estimates, particularly in terms of trend analyses since access to health care was substantially affected by the COVID-19 pandemic as were related factors such as health insurance and employment status for many people.

MEPS began in 1996, and the utility of the survey for analyzing health care trends expands with each additional year of data; however, when examining trends over time using MEPS, the length of time being analyzed should be considered. In particular, large shifts in survey estimates over short periods of time (e.g. from one year to the next) that are statistically significant should be interpreted with caution unless they are attributable to known factors such as changes in public policy, economic conditions, or MEPS survey methodology.

With respect to methodological considerations, in 2013 MEPS introduced an effort focused on field procedure changes such as interviewer training to obtain more complete information about health care utilization from MEPS respondents with full implementation in 2014. This effort likely resulted in improved data quality and a reduction in underreporting starting in the second half of 2013 and throughout 2014 full year files and have had some impact on analyses involving trends in utilization across years. The changes in the NHIS sample design in 2016 and 2018 could also potentially affect trend analyses. The new NHIS sample design is based on more up-to-date information related to the distribution of housing units across the U.S. As a result, it can be expected to better cover the full U.S. civilian, noninstitutionalized population, the target population for MEPS, as well as many of its subpopulations. Better coverage of the target population helps to reduce the potential for bias in both NHIS and MEPS estimates.

Another change with the potential to affect trend analyses involved major modifications to the MEPS instrument design and data collection process, particularly in the events sections of the instrument. These were introduced in the Spring of 2018 and thus affected data beginning with Round 1 of Panel 23, Round 3 of Panel 22, and Round 5 of Panel 21. Since the full-year 2017 PUFs were established from data collected in Rounds 1-3 of Panel 22 and Rounds 3-5 of Panel 21, they reflected two different instrument designs. In order to mitigate the effect of such differences within the same full year file, the Panel 22 Round 3 data and the Panel 21 Round 5 data were transformed to make them as consistent as possible with data collected under the previous design. The changes in the instrument were designed to make the data collection effort more efficient and easy to administer. In addition, expectations were that data on some items, such as those related to health care events, would be more complete with the potential of identifying more events. Increases in service use reported since the implementation of these changes are consistent with these expectations. ***Data users should be aware of possible impacts on the data and especially trend analyses for these data years due to the design transition.***

Process changes, such as data editing and imputation, may also affect trend analyses. For example, users should refer to Section 2.5.11 in the 2021 Consolidated file (HC-233) and, for more detail, the documentation for the prescription drug file (HC-229A) when analyzing prescription drug spending over time.

As always, it is recommended that data users review relevant sections of the documentation for descriptions of these types of changes that might affect the interpretation of changes over time before undertaking trend analyses.

Analysts may also wish to consider using statistical techniques to smooth or stabilize analyses of trends using MEPS data such as comparing pooled time periods (e.g. 1996-1997 versus 2011-2012), working with moving averages, or using modeling techniques with several consecutive years of MEPS data to test the fit of specified patterns over time.

Finally, statistical significance tests should be conducted to assess the likelihood that observed trends are not attributable to sampling variation. In addition, researchers should be aware of the impact of multiple comparisons on Type I error. Without making appropriate allowance for multiple comparisons, undertaking numerous statistical significance tests of trends increases the likelihood of concluding that a change has taken place when one has not.

References

- Bramlett, M.D., Dahlhamer, J.M., & Bose, J. (2021, September). [*Weighting Procedures and Bias Assessment for the 2020 National Health Interview Survey*](#). Centers for Disease Control and Prevention.
- Chowdhury, S.R., Machlin, S.R., Gwet, K.L. [*Sample Designs of the Medical Expenditure Panel Survey Household Component, 1996-2006 and 2007-2016*](#). Methodology Report #33. January 2019. Agency for Healthcare Research and Quality, Rockville, MD.
- [*Current Population Survey: 2021 Annual Social and Economic \(ASEC\) Supplement*](#). (2021). U.S. Census Bureau.
- Dahlhamer, J.M., Bramlett, M.D., Maitland, A., & Blumberg, S.J. (2021). [*Preliminary evaluation of nonresponse bias due to the COVID-19 pandemic on National Health Interview Survey estimates, April-June 2020*](#). National Center for Health Statistics.
- Fay, R.E. (1989). Theory and Application of Replicate Weighting for Variance Calculations. *Proceedings of the Survey Research Methods Sections, ASA*, 212-217.
- Lau, D.T., Sosa, P., Dasgupta, N., & He, H. (2021). [*Impact of the COVID-19 Pandemic on Public Health Surveillance and Survey Data Collections in the United States*](#). *American Journal of Public Health*, 111 (12), pp. 2118-2121.
- Rothbaum, J. & Bee, A. (2021, May 3). [*Coronavirus Infects Surveys, Too: Survey Nonresponse Bias and the Coronavirus Pandemic*](#). U.S. Census Bureau.
- Rothbaum, J. & Bee, A. (2022, September 13). [*How Has the Pandemic Continued to Affect Survey Response? Using Administrative Data to Evaluate Nonresponse in the 2022 Current Population Survey Annual Social and Economic Supplement*](#). U.S. Census Bureau.
- Zuvekas, S.H. & Kashihara, D. (2021). [*The Impacts of the COVID-19 Pandemic on the Medical Expenditure Panel Survey*](#). *American Journal of Public Health*, 111 (12), pp. 2157-2166.

D. Variable-Source Crosswalk

FOR MEPS HC-230: 2021 FOOD SECURITY DATA FILE

SURVEY ADMINISTRATION VARIABLES

VARIABLE	DESCRIPTION	SOURCE
HOMEIDX	HOME ID NUMBER (DUID + RU + ROUND)	Constructed
DUID	PANEL # + ENCRYPTED DU IDENTIFIER	Assigned in Sampling
PANEL	PANEL NUMBER	Constructed
RULETR42	RU LETTER - R4/2	CAPI Derived

FOOD SECURITY VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
FSOUT42	HOW OFTEN HAVE YOU RUN OUT OF FOOD	FS20
FSLAST42	HOW OFTEN DID FOOD NOT LAST	FS30
FSAFRD42	HOW OFTEN NOT AFFORD BALANCED MEALS	FS40
FSSKIP42	DID YOU EVER SKIP MEALS	FS50
FSSKDY42	HOW MANY DAYS WERE MEALS SKIPPED	FS60
FSLESS42	DID YOU EVER EAT LESS	FS70
FSHGRY42	DID YOU EVER GO HUNGRY	FS80
FSWTLS42	LOW FOOD MONEY CAUSE WEIGHT LOSS	FS90
FSNEAT42	DID YOU EVER NOT EAT	FS100
FSNEDY42	HOW MANY DAYS DID YOU NOT EAT	FS110

WEIGHTS VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
FSWT42	FOOD SECURITY WEIGHT	Constructed
VARSTR	VARIANCE ESTIMATION STRATUM - 2021	Constructed
VARPSU	VARIANCE ESTIMATION PSU - 2021	Constructed