MEPS HC-189
2016 Food Security

July 2018

Table of Contents

Table of Contents (continued)

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, which includes 5 Rounds of interviews covering 2 full calendar years, provides 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 and reflects an oversample of Blacks and Hispanics. 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. The linkage of the MEPS to the previous year's NHIS provides additional data for longitudinal analytic purposes.

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 can not 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 date filled and sources and 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 summary reports, micro data files, and tables via the MEPS website. Selected data can be analyzed through MEPSnet, an on-line interactive tool designed to give data users the capability to statistically analyze MEPS data in a menu-driven environment.

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 2016 food security data file from the Medical Expenditure Panel Survey Household Component (MEPS HC). Released as an ASCII file (with related SAS, SPSS, and Stata programming statements and data user information) and a SAS transport dataset, this public use file provides information collected on a nationally representative sample of the civilian noninstitutionalized population of the United States for calendar year 2016. The file contains 16 variables and has a logical record length of 51 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 4 of Panel 20 and Round 2 of Panel 21, in calendar year 2016, 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 2016 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 and a variable locator indicating the major MEPS data items on public use files that have been released to date can be found at the following link on the MEPS website.

2.0 Data File Information

This public use dataset contains variables and frequency distributions associated with 13,500 households who participated in the MEPS Household Component of the Medical Expenditure Panel Survey in 2016. These households received a positive family-level weight and were part of one of the two MEPS panels for whom food security data were collected in Round 4 of Panel 20 or Round 2 of Panel 21.

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
-9	NOT ASCERTAINED	Interviewer did not record the data

2.3　　　　Codebook Format

This codebook describes an ASCII data set and provides the following programming identifiers for each variable:

Identifier	Description
Name	Variable name (maximum of 8 characters)
Description	Variable descriptor (maximum 40 characters)
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. The last two characters denote the rounds of data collection, Round 4 of Panel 20 and Round 2 of Panel 21.

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 – PANEL)

HOMEIDX uniquely identifies each household on the file and consists of the Dwelling Unit ID (DUID) followed by the 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 five-digit random ID 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 either Panel 20 or Panel 21 for each person on the file. Panel 20 is the panel that started in 2015, and Panel 21 is the panel that started in 2016.

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 – FSNEDY42)

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 the food purchased didn't last and the person/household didn't have money to get more

FSAFRD42 - how often in the last 30 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 -9)

FSSKDY42 - how many meals were skipped in the last 30 days (coded as "-1 Inapplicable" when FSSKIP42 = 2, -7, -8, or -9 OR when FSOUT42, FSLAST42, and FSAFRD42 = 3, -7, -8, or -9)

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 -9)

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 -9)

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 -9)

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 -9; or when FSLESS42, FSHGRY42, and FSWTLS42 = 2, -7, -8, or -9)

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 -9; or when FSLESS42, FSHGRY42, and FSWTLS42 = 2, -7, -8, or -9; or when FSNEAT42 = 2, -7, -8, or -9)

2.6   Linking to Other Files

2.6.1   MEPS Public Use Files

The family (RU) level records in this public use file can be linked to the 2016 full year Consolidated file to obtain additional data for the families included in this file. To link the records with the Consolidated data file records, create an RU ID using the first six characters of HOMEIDX (columns 1-5 are DUID and column 6 is RU letter). On the full year Consolidated file concatenate DUID with RULETR42 (Round 4/2 RU letter) to obtain an equivalent ID. The records in both files can then be linked with this ID. The reference person of the RU can be identified in the Consolidated data file by the variable REFPRS42.

Food Security PUF:
Includes HOMEIDX(12345A4) as DUID(12345) + RULETER42(A) + Round(4)
        Construct RUID(12345A) as substr(HOMEIDX, 1,6)

Consolidated PUF: Includes DUID(12345) and RULETER42(A)
        Construct RUID(12345A) as DUID(12345) + RULETER42(A)

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 each panel, the longitudinal file comprises MEPS survey data obtained in Rounds 1 through 5 of the panel and can be used to analyze changes over a two-year period. 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 two years covered by that 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   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 MEPS data in this public use file pertain to calendar year 2016. The data were collected in Rounds 1, 2, and 3 for MEPS Panel 21 and Rounds 3, 4, and 5 for MEPS Panel 20. (Note that Round 3 for a MEPS panel is designed to overlap two calendar years, as illustrated below.) The 2016 food security data were collected only in Round 4 of Panel 20 and Round 2 of Panel 21.

	2015		2016		2017
	Jan		Jan	Dec	Jan
Panel 20 2015-2016	Round 1	Round 2	Round 3	Round 4	Round 5
		Panel 21 2016-2017	Round 1	Round 2	Round 3
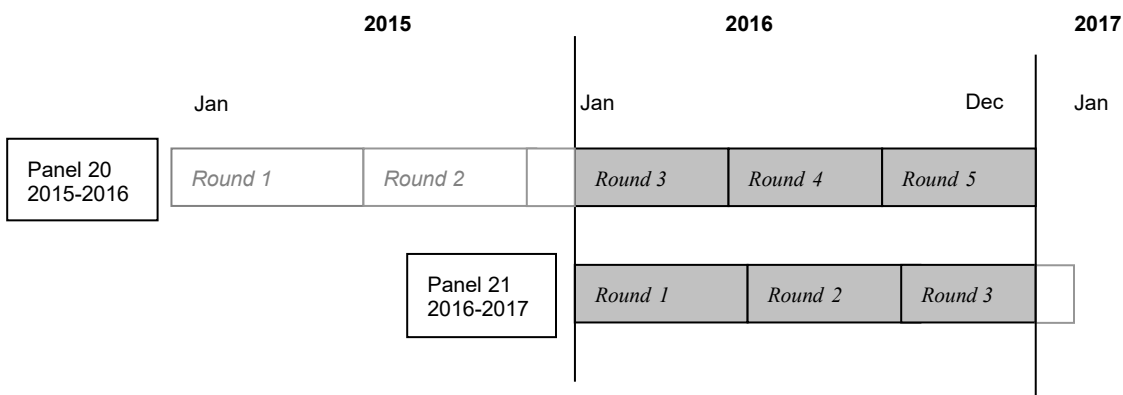
A sample design feature shared by both Panel 20 and Panel 21 involved the partitioning of the sample domain "Other" (serving as the catchall stratum, and consisting mainly of households with "White" members) into two sample domains. This was done for the first time in Panel 16. The two domains were defined as: 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 both MEPS panels the "partial" domain was sampled at a lower rate than the "complete" domain. This approach served to reduce survey costs, since the "partials" tend to have higher costs in gaining survey participation, but increased sample variability due to the resulting increased variance in sampling rates.

3.1.1   References

There have been some published reports on the MEPS sample design. For detailed information on the MEPS sample design, see Cohen, S., Sample Design of the 1997 Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Healthcare Research and Quality; 2000. MEPS Methodology Report, No. 11. AHRQ Pub. No. 01-0001 and Ezzati-Rice, T.M., Rohde, F., Greenblatt, J., (2008). Sample Design of the Medical Expenditure Panel Survey Household Component, 1998-2007, Methodology Report, No. 22. March 2008. Agency for Healthcare Research and Quality, Rockville, MD.

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

Each responding household found in this 2016 MEPS dataset is associated with one of two separate and overlapping MEPS panels, MEPS Panel 20 and MEPS Panel 21. These panels consist of subsamples of households participating in the 2014 and 2015 NHIS, respectively, and reflecting the NHIS sample design first implemented in 2006.

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. It may be noted that 2016 is the last year for which both MEPS panels reflect the 2006 NHIS sample design. MEPS Panel 22 (fielded in 2017) will reflect the new NHIS design, first implemented in 2016. To the extent that users compare MEPS estimates to 2016 NHIS estimates, they should be cognizant of this design change as it may affect the extent to which MEPS and NHIS data are comparable. An overview of the new CDC NHIS sample design can be found at the CDC website.

As background, the NHIS is a complex multi-stage sample design. A brief and simplified description of the NHIS design follows. The first stage of sample selection is an area sample of PSUs, where PSUs generally consist of one or more counties. Within PSUs, density strata are formed, generally reflecting the density of minority populations for single or groups of blocks or block equivalents that are assigned to the strata. Within each such density stratum "supersegments" are formed, consisting of clusters of housing units. Samples of supersegments are selected for use over a 10-year data collection period for the NHIS. Households within supersegments are selected for each calendar year the NHIS is carried out. In the NHIS sample design used since 2006, Asians are oversampled in addition to Hispanics and Blacks. These features of the NHIS complex survey design carry over to the MEPS. The only major difference in eligibility status for housing units between NHIS and MEPS is that college dorms represent ineligible housing units for MEPS. College aged students living away from home during the school year were interviewed at their place of residence for the NHIS but were identified by and linked to their parents' household for MEPS. (There is also a person-level stage of sampling for the NHIS, but that does not affect the MEPS sample design.)

The households (occupied DUs) selected for MEPS Panel 20 were a subsample of the 2014 NHIS responding households, while those in MEPS Panel 21 were a subsample of 2015 NHIS responding households. A MEPS household may contain one or more family units, each consisting of one or more individuals. Analysis using MEPS data can be undertaken using either the individual or the family as the unit of analysis.

There were 10,610 households (occupied DUs) selected for MEPS Panel 20, of which 10,571 were eligible for fielding (college dormitories were eliminated). They were randomly selected from among the households responding to the 2014 NHIS. A subsample of 9,700 households was randomly selected for MEPS Panel 21 from the households responding to the 2015 NHIS, of which 9,658 were fielded for MEPS after the elimination of college dorms.

3.1.3   Sample Weights and Variance Estimation

In the dataset MEPS HC-189: 2016 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.

3.2   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 both MEPS Panel 20 and Panel 21.

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 representing students living in student housing or consisting entirely of military personnel are deleted from the sample. For the NHIS, college students living in student housing are sampled independently from their families. For MEPS, such students are identified through the sample selection of their parents' RU. Removing from MEPS those college students found in college housing sampled for the NHIS eliminates the opportunity of multiple chances of selection for MEPS for these students. Military personnel not living in the same RU as civilians are ineligible for MEPS. After such 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 students and military members discussed above). Note that all response rates discussed here are unweighted.

Table 3.1.   Sample Size and Unweighted Response Rates for 2016 Full Year File (Panel 21 Rounds 1-3/Panel 20, Rounds 3-5)

	Panel 20	Panel 21	2016 Combined
A. Percentage of NHIS households designated for use in MEPS (those initially characterized as responding) *	75.1%	71.2%	—
B. Number of households sampled from the NHIS	10,610	9,700	—
C. Number of Households sampled from the NHIS and fielded for MEPS	10,571	9,658	—
D. Round 1 – Number of RUs eligible for interviewing	11,283	10,280	—
E. Round 1 – Number of RUs with completed interviews	8,287	7,643	—
F. Round 2 – Number of RUs eligible for interviewing	8,554	7,870	—
G. Round 2 – Number of RUs with completed interviews	7,991	7,319	—
H. Round 3 – Number of RUs eligible for interviewing	8,136	7,478	—
I. Round 3 – Number of RUs with completed interviews	7,743	7,035	—
J. Round 4 – Number of RUs eligible for interviewing	7,877	—	—
K. Round 4 – Number of RUs with completed interviews	7,621	—	—
L. Round 5 – Number of RUs eligible for interviewing	7,698	—	—
M. Round 5 – Number of RUs with completed interviews	7,421	—	—
Overall annual unweighted response rates P21: A x (E/D) x (G/F) x (I/H) P20: A x (E/D) x (G/F) x (I/H) x (K/J) x (M/L) Combined: 0.510 x P20 + 0.490 x P21	45.7% (Panel 20 through Round 5)	46.3% (Panel 21 through Round 3)	46.0%

*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.

3.2.1 Response Rates

In order to produce annual health care estimates for calendar year 2016 based on the full MEPS sample data from the MEPS Panel 20 and Panel 21, the two panels are combined. More specifically, full calendar year 2016 data collected in Rounds 3 through 5 for the MEPS Panel 20 sample are pooled with data from the first three rounds of data collection for the MEPS Panel 21 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 21 Round 2 the ratio of 7,319 (Row G) to 7,870 (Row F) multiplied by 100 represents the response rate for the round (93.0 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 the combined sample of Panel 20 and Panel 21 for 2016 was obtained by computing the products of the relative sample sizes and the corresponding overall panel response rates and then summing the two products. Panel 21 represents about 49.0 percent of the combined sample size while Panel 20 represents the remaining 51.0 percent. Thus, the combined response rate of 46.0 percent was computed as 0.510 times 45.7, the overall Panel 20 response rate through Round 5 plus 0.490 times 46.3, the overall Panel 21 response rate through Round 3.

3.2.2 Panel 21 Response Rates

For MEPS Panel 21 Round 1, 9,658 households were fielded in 2016 (Row C of Table 3.1), a randomly selected subsample of the households responding to the 2015 National Health Interview Survey (NHIS).

Table 3.1 shows the number of RUs eligible for interviewing in each Round of Panel 21 as well as the number of RUs completing the MEPS interview. Computing the individual round "conditional" response rates as described in section 3.2.1 and then taking the product of these three response rates and the factor 71.2 (the percentage of the NHIS sampled households designated for use in selecting a sample of households for MEPS) yields an overall response rate of 46.3 percent for Panel 21 through Round 3.

3.2.3 Panel 20 Response Rates

For MEPS Panel 20, 10,571 households were fielded in 2015 (as indicated in Row C of Table 3.1), a randomly selected subsample of the households responding to the 2014 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 20. The overall response rate for Panel 20 was computed in a similar fashion to that of Panel 21 but covering all five rounds of MEPS interviewing as well the factor representing the percentage of NHIS sampled households eligible for MEPS. The overall response rate for Panel 20 through Round 5 is 45.7 percent.

3.2.4 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 20 response rate was weighted by a factor of 0.510 and Panel 21 was weighted by a factor of 0.490, reflecting approximately the distribution of the overall sample between the two panels. The resulting combined response rate for the combined panels was computed as (0.510 x 45.7) plus (0.490 x 46.3) or 46.0 percent (as shown in Table 3.1).

3.2.5 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 oversampling of Hispanic, Black, and Asian households for the NHIS carries over to MEPS through the set of NHIS responding households eligible for sample selection for MEPS. In the NHIS under the old sample design utilized through 2005, Hispanic households were oversampled at a rate of roughly 2 to 1. That is, the probability of selecting a Hispanic household for participation in the NHIS was roughly twice that for households in the general population that were not oversampled. The oversampling rate for Black households under the old design was roughly 1.5 to 1. Under the NHIS sample design employed through 2015 (which is the sample design applicable for MEPS Panels 20 and 21), Asians, as well as Hispanics and Blacks, are oversampled. The average oversampling rates for the three minority groups have not yet been reported.

For both Panel 20 and Panel 21, all households in the Asian, Hispanic, and Black domains were sampled with certainty (i.e., all households assigned to those domains were included in the MEPS). For Panel 20, the "Other, complete" domain was sampled at a rate of about 84 percent while the "Other, partial complete" domain was sampled at a rate of about 53 percent. For Panel 21, the corresponding sampling rates for the "Other, complete" domain and the "Other, partial complete" domain were about 81 percent and 49 percent, respectively.

Within strata (domains) for both panels, responding NHIS households were selected for MEPS using a systematic sample selection procedure from among those eligible. For the "non-Other" strata households were all selected with certainty. Within strata involving "Others" (two strata for both panels) the selection 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. With the subsampling, households that were oversampled for MEPS in calendar year 2016 were those responding households in the NHIS identified as having members whose race/ethnicity was Hispanic, Black, or Asian for both panels.

Typically, sample allocations across sample domains change from one MEPS panel to another. The sample domains used may also vary by panel although this was not the case for Panel 20 and Panel 21. When one compares unweighted measures (e.g., response rates) between panels and years, one should take into account such differences. If, 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.

Within each domain (sample stratum) systematic samples of the MEPS-eligible households were selected from among the NHIS household respondents made available for MEPS sample selection purposes.

3.3   Food Security Weight (FSWT42)

The Food Security questionnaire was administered to RU level respondents in Round 4 of Panel 20 and Round 2 of Panel 21. A family weight specific to Rounds 4/2, FSWT42, was created for the purposes of analyzing the Food Security questionnaire data. To create FSWT42, first those RUs that were MEPS respondents in Round 4/2, were identified. The MEPS family weight (FAMWT16F) from the 2016 full year Consolidated data file of the RU reference person was

then assigned as the family weight of the RU. A small number of RUs (less than 2 percent) had a reference person with no MEPS family weight in the Consolidated data file because at least one member of the family became a nonrespondent for MEPS in the subsequent round. For these RUs, the reference person's full-year person weight was assigned as the family weight. The reference person of the RU in MEPS is defined as the household member 16 years of age or older who own or rents the home. Note that there are two different family weights provided in the 2016 Full Year Consolidated PUF: the MEPS family weight, FAMWT16F; and the family weight based on the Current Population Survey (CPS) definition of a family, FAMWT16C. The MEPS family weight (FAMWT16F) was used for the food security data because the MEPS family weight definition of family was closest to the MEPS Round 4/2 RU definition. For more information on the derivation of FAMWT16F and FAMWT16C, see the MEPS HC-192, 2016 Full Year Consolidated Data File 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 20	Panel 21	Combined	Population estimate (weighted total of combined sample)
Number	6,946	6,554	13,500	137,402,160

3.4   Variance Estimation

The MEPS is based on a complex sample design. 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. However, an option is also provided to apply the BRR approach when needed to develop variances for more complex estimators.

3.4.1   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, Stata, SAS (version 8.2 and higher), and SPSS (version 12.0 and higher). For complete information on the capabilities of each 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. However, beginning with the 2002 Point-in-Time PUF, the variance strata and PSUs were 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 were 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 2008 file and beyond are based on the new 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 in the event that data are pooled for several years.

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 from 2002 or later, one can use the variance strata numbering as is.

2. When pooling any year from 1996 to 2001 with any year from 2002 or later, use the pooled linkage public use file HC-036 that contains the proper variance structure to use when making estimates from MEPS data that have been pooled over multiple years and where one or more years are from 1996-2001.

3. The HC-036 file is updated every year to allow pooling of any year from 1996 to 2001 with any year from 2002 up to the latest year. Further details on the HC-036 file can be found in the public use documentation of the HC-036 file.

3.4.2   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. So 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 full developed BRR replicate weights. The Taylor Series approach does not fully capture the effects of the different weighting adjustments either.

The dataset HC-036BRR 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 file. For more information about creating BRR replicates, users can refer to the documentation for the HC-036BRR pooled linkage file.

D. Variable-Source Crosswalk

VARIABLE-SOURCE CROSSWALK
FOR MEPS HC-189: 2016 FOOD SECURITY DATA FILE

SURVEY ADMINISTRATION VARIABLES

VARIABLE	DESCRIPTION	SOURCE
HOMEIDX	HOME ID NUMBER (DUID + RU + ROUND)	Constructed
DUID	DWELLING UNIT ID	Assigned in Sampling
PANEL	PANEL	Constructed

FOOD SECURITY VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
FSOUT42	HOW OFTEN HAVE YOU RUN OUT OF FOOD	FS02
FSLAST42	HOW OFTEN DID FOOD NOT LAST	FS03
FSAFRD42	HOW OFTEN NOT AFFORD BALANCED MEALS	FS04
FSSKIP42	DID YOU EVER SKIP MEALS	FS05
FSSKDY42	HOW MANY DAYS WERE MEALS SKIPPED	FS06
FSLESS42	DID YOU EVER EAT LESS	FS07
FSHGRY42	DID YOU EVER GO HUNGRY	FS08
FSWTLS42	LOW FOOD MONEY CAUSE WEIGHT LOSS	FS09
FSNEAT42	DID YOU EVER NOT EAT	FS10
FSNEDY42	HOW MANY DAYS DID YOU NOT EAT	FS11

WEIGHTS VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
FSWT42	FOOD SECURITY WEIGHT	Constructed
VARSTR	VARIANCE ESTIMATION STRATUM - 2016	Constructed
VARPSU	VARIANCE ESTIMATION PSU - 2016	Constructed