

**MEPS HC-117:
2009 P13R3/P14R1
Population Characteristics
June 2010**

**Agency for Healthcare Research and Quality
Center for Financing, Access, and Cost Trends
540 Gaither Road
Rockville, MD 20850
(301) 427-1406**

Table of Contents

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-1
	2.2 Reserved Codes	C-2
	2.3 Codebook Format	C-2
	2.4 Variable Naming	C-2
	2.5 File Contents.....	C-3
	2.5.1 Survey Administration Variables.....	C-3
	2.5.2 Demographic Variables.....	C-10
	2.5.3 Health Status and Priority Condition Variables	C-15
	2.5.4 Employment Variables.....	C-18
	2.5.5 Health Insurance Variables	C-21
	2.6 Linking to Other Files	C-24
	2.6.1 National Health Interview Survey.....	C-24
	2.6.2 Longitudinal Analysis	C-24
	3.0 Survey Sample Information	C-24
	3.1 Sample Design and Response Rates.....	C-24
	3.1.1 The MEPS Sampling Process and Response Rates: An Overview	C-25
	3.1.2 Panel 13 Response.....	C-27
	3.1.3 Panel 14 Response.....	C-28
	3.1.4 Combined Panel Response.....	C-28
	3.1.5 Oversampling.....	C-28

Table of Contents (continued)

3.2	Sample Weights.....	C-29
3.2.1	Person-level Weight.....	C-29
3.2.2	Family-level Weight	C-30
3.2.2.1	Definition of MEPS Families	C-30
3.2.2.2	Assignment of Weights.....	C-31
3.2.2.3	Instructions to Create Family Estimates.....	C-31
3.2.3	Relationship between Person- and Family-Level Weights.....	C-32
3.3	Variance Estimation	C-32
3.4	Using MEPS Data for Trend Analysis	C-34
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.

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 non-institutionalized 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 non-institutionalized 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. MEPS further oversamples additional policy relevant subgroups such as low income households. 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 visit, 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 (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 web site: www.meps.ahrq.gov. 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, 540 Gaither Road, Rockville, MD 20850 (301-427-1406).

C. Technical and Programming Information

1.0 General Information

This documentation describes the 2009 point in time data file to be released from the Medical Expenditure Panel Survey Household Component (MEPS HC). Released as an ASCII file (with related SAS and SPSS programming statements and data user information), and a SAS transport data set, this public use file provides information on data collected on a nationally representative sample of the civilian, non-institutionalized population of the United States during the early part of 2009. The file contains 104 variables and has a logical record length of 217 with an additional 2-byte carriage return/line feed at the end of each record. The data consist of 2009 data obtained in Round 3 of Panel 13 and Round 1 of Panel 14 of the MEPS Household Component and contain variables pertaining to Survey Administration, Demographics, Health Status and Priority Conditions, Employment, and Health Insurance.

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

- Data File Information
- Survey Sample Information
- Variable/Questionnaire Crosswalk

Both weighted and unweighted frequencies of all the variables included on the 2009 point-in-time data file are provided in the accompanying codebook file.

MEPS survey questionnaires, the Household Survey Variable Locator indicating the major MEPS data items on public use files that have been released to date and a catalog of MEPS products are available on the MEPS web page (www.meps.ahrq.gov).

2.0 Data File Information

This public use data file contains variable and frequency distributions for a total of 39,654 persons (19,881 from Panel 13 Round 3 and 19,773 from Panel 14 Round 1). This count includes all household survey respondents who resided in eligible responding households. Of these persons, 37,970 were assigned a positive person-level weight (18,779 from Panel 13 Round 3 and 19,191 from Panel 14 Round 1). For each variable, both weighted and unweighted frequencies are provided. In conjunction with the weight variable (WGTSP13) provided on this file, data for these persons can be used to make estimates for the civilian, non-institutionalized U.S. population as of the first half of 2009.

2.1 Codebook Structure

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

- Unique person identifiers
- Demographic variables
- Health Status variables

- Employment variables
- Health Insurance 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
-10 HOURLY WAGE >= \$67.31	Variable was top-coded for confidentiality

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

In general, variable names reflect the content of the variable, with an 8 character limitation. All of the variables on this file (except some demographic variables and DUID, DUPERSID, PID, and KEYNESS) end in "13" to denote they are combination Panel 14 Round 1/Panel 13 Round 3 variables. For edited variables, the "13" is followed by an "X", and they are so noted in the variable label. Variables contained in this delivery were derived either from the questionnaire itself or from the CAPI. The source of each variable is identified in the section of the documentation entitled "D. Variable-Source Crosswalk". Sources for each variable are indicated in one of four ways: (1) variables derived from the CAPI or assigned in sampling are so indicated; (2) variables derived from complex algorithms associated with reenumeration are labeled "RE Section"; (3) variables that come from one or more specific questions have those

numbers listed in the “Source” column; and (4) 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

The Survey Administration variables contain information related to conducting the interview, household and family composition, and person-level and RU-level status codes. Data for the Survey Administration variables were derived from the sampling process, the CAPI programs, or were computed based on information provided by the respondent in the reenumeration section of the questionnaire. Most Survey Administration variables on this file are asked during every round of the MEPS interview. Variables in this delivery describe data for Panel 13 Round 3 and Panel 14 Round 1 in 2009.

The variable PANEL indicates the panel from which the data are derived. A value of 13 indicates Panel 13 Round 3 data and a value of 14 indicates Panel 14 Round 1 data.

Note that Round 3 of Panel 13 covers both the end of 2008 and the beginning of 2009. (When possible, the variables were constructed to represent data from the 2009 portion of Round 3.)

Dwelling Units, Reporting Units, and Families

The definition of Dwelling Units (DUs) in the MEPS Household Survey is generally consistent with the definition employed for the National Health Interview Survey (NHIS). The dwelling unit ID (DUID) is a five-digit random ID number assigned after the case was sampled for MEPS. The person number (PID) uniquely identifies all persons within the dwelling unit. The variable DUPERSID is a combination of the variables DUID and PID; thus it uniquely identifies each sampled person in MEPS.

A Reporting Unit (RU) is a person or group of persons in the sampled dwelling unit who are related by blood, marriage, adoption, foster care or other family association. Each RU is to be 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 are identified by the variable RULETR13. Regardless of the legal status of their association, two persons living together as a “family” unit were treated as a single reporting unit if they chose to be so identified. Examples of different types of reporting units are:

1. A married daughter and her husband living with her parents in the same dwelling unit constitute a single reporting unit.
2. A husband and wife and their unmarried daughter, age 18, who is living away from home while at college, constitute two reporting units.
3. Three unrelated persons living in the same dwelling unit would each constitute a distinct reporting unit, three reporting units in all.

Unmarried college students less than 24 years of age who usually live in the sampled household, but were living away from home and going to school at the time of the MEPS interview, were treated as a reporting unit separate from that of their parents for the purpose of data collection. The variable RUSIZE13 indicates the number of persons in each RU, treating each student as a single RU separate from their parents. Thus, students are not included in the RUSIZE13 count of their parents' RU. However, for many analytic objectives, the student reporting units would be combined with their parents' reporting unit, treating the combined entity as a single family. Family identifier and size variables are described below and include students with their parents' reporting unit.

The variable FAMID13 identifies a family (i.e., persons living together related to one another by blood, marriage, adoption, foster care, or self-identified as a single unit plus related students who are living away at post-secondary school) for each round. These family identifier variables use a letter and a DU identifier to indicate a person's family affiliation. In order to identify a person's family affiliation, users must create a unique family identification variable by concatenating the DU identifier (DUID) and the FAMID13 variable, as described in Section 3.2.2.3 Instructions to Create Family Estimates.

The variable FAMSIZ13 indicates the number of persons associated with a single family unit after students are linked to their associated parent RUs for analytical purposes. Family-level analyses should use the FAMSIZ13 variable. In a few cases, students were deleted from the file because attempts to contact them were unsuccessful, and no data were collected for them. However, these persons are accounted for in the FAMSIZ13 variable.

The family size (FAMSIZ13) and the reporting unit (RU) size (RUSIZE13) counts may not be consistent with the count of records on the file. There are some reporting units where the RU size variable (RUSIZE13) is not equal to the number of people in that RU actually included on the file. This occurs because people who did not respond for their entire period of eligibility were not included on the file. In addition, for a number of these reporting units, the reference person is not included on the file for this same reason.

The variable RURSLT13 indicates the RU response status for Round 3 for the Panel 13 sample and Round 1 for the Panel 14 sample. The values include the following:

- 60 Complete with RU member
- 61 Complete with proxy--all RU members deceased on or after 1/1/2009
- 62 Complete with proxy--all RU members institutionalized or deceased on or after 1/1/2009
- 63 Complete with proxy, other

There are several other variables that characterize the reporting unit. The variable RUCLAS13 indicates the RU classification. RUs are classified for fielding purposes as 1 "Standard", 2 "New RU", or 3 "Student RU". Standard RUs are the original RUs from NHIS. A new RU is one which has been created when members of the household leave the standard RU and are followed according to the rules of the survey. A student RU is one in which an unmarried college student under 24 years of age is considered a usual member of the household but was living away from

home while going to school and was treated as a Reporting Unit (RU) separate from that of their parents for the purpose of data collection.

Reference Period Dates

The reference period is the period of time for which data were collected in each round for each person. The reference period dates were determined during the interview for each person by the CAPI program.

The round-specific beginning reference period dates are included for each person. These variables include BEGRFD13, BEGRFM13, and BEGRFY13. The reference period for Panel 14 Round 1 for most persons identified at NHIS began on January 1, 2009 and ended on the date of the Round 1 interview. Persons who joined the RU after January 1, 2009 have their beginning reference date for the round as the day they joined the RU.

For Panel 13 Round 3, the reference period for most persons began on the date of the previous round's interview and ended on the date of the current round's interview. Persons who joined after the previous round's interview had their beginning reference date for the round set as the day they joined the RU.

The dates of the interview and the ending reference period dates are included for each person. These variables include ENDRFD13, ENDRFM13, ENDRFY13, RUENDDD13, RUENDM13, and RUENDY13. In general, the date of the interview is the reference period end date for most persons. Note that the end date of the reference period is prior to the date of the interview if the person was deceased during the round, left the country, was institutionalized prior to that round's interview, or joined the military during the round and was not living with someone else who was eligible. Because of this, it is possible for a person whose reference period for Round 3 ended in 2008 to be included in this delivery. While these few persons do not have a positive person-level weight for 2009, they are included in this file because they do have a positive family-level weight for 2009. If a person left the RU and that person was key and in-scope, the person was followed in the new RU to which he or she moved and his or her reference period dates pertain to the new RU.

Reference Person Identifiers

The variable RNDREF13 identifies the reference person for the RU. In general, the reference person is defined as the household member 16 years of age or older who owns or rents the home. If the person identified as the reference person in a previous round (at NHIS if Round 1) still lives in the RU at the date of the current interview, then this person automatically continues to be the reference person for the current round. Only when the previously identified reference person is no longer living in the RU, the household respondent is asked to identify another person from the RU fitting this definition. If the respondent is unable to identify a new reference person then the questionnaire asks for the head of household among the DU members fitting this definition and the person selected or added is then considered the reference person for that RU. This information was collected in the reenumeration section of the CAPI questionnaire.

Respondent Identifiers

The respondent is the person who answered the interview questions for the reporting unit (RU). The round-specific variable RDRESP13 identifies the respondent. Only one respondent is identified for each RU. In instances where the interview was completed in more than one session, only the first respondent is indicated.

There are two types of respondents. The respondent can be either an RU member or a non-RU member proxy. The variable PROXY13 identifies the type of respondent.

Person Status

A number of variables describe the various components reflecting each person's status for each round of data collection. These variables provide information about a person's in-scope status, keyness status, eligibility status, and disposition status. These variables include: KEYNESS, INSCOP13, and PSTAT13. These variables are set based on sampling information and responses provided in the reenumeration section of the CAPI questionnaire.

Through the reenumeration section of the CAPI questionnaire, each member of a reporting unit was classified as "key" or "non-key", "in-scope" or "out-of-scope", and "eligible" or "ineligible" for MEPS data collection. To be included in the set of persons used in the derivation of MEPS person-level estimates, a person had to be a member of the U.S. civilian, non-institutionalized population for at least one day during 2009. Because a person's eligibility for the survey might have changed since the NHIS interview, a reenumeration of household membership was conducted at the start of each round's interview. Only persons who were "in-scope" sometime during 2009, "key", and responded for the full period in which they were in-scope were assigned person-level weights and thus are to be used in the derivation of person-level national estimates from the MEPS.

In-Scope

A person is considered as in-scope during a round if he or she is a member of the U.S. civilian, non-institutionalized population at some time during that round. The variable INSCOP13 indicates a person's in-scope status, specifically indicating whether a person was ever in-scope during the 2009 portion of the round.

Keyness

The term "keyness" is related to an individual's chance of being included in MEPS for purposes of making estimates about the U. S. civilian, non-institutionalized population. A person is key if that person is linked for sampling purposes to the set of NHIS sampled households designated for inclusion in MEPS. Specifically, a key person either was a member of an NHIS household at the time of the NHIS interview, or was a family member who began living with a member of such a household after being out-of-scope prior to joining that household. (Examples of the latter situation include newborns and persons returning from military service, an institution, or living outside the United States.)

A non-key person is one whose chance of selection for the NHIS (and MEPS) was associated with a household eligible but not sampled for the NHIS, and who later became a member of a MEPS reporting unit. MEPS data (e.g., utilization and income) were collected for the period of time a non-key person was part of the sampled unit to provide information for family-level analyses. However, non-key persons who leave a sample household unaccompanied by a key, in-scope member were not followed for subsequent interviews. Non-key individuals do not receive person-level sample weights and thus do not contribute to person-level national estimates. They may receive family-level weights if they are a member of a responding family.

The variable KEYNESS indicates a person's keyness status. This variable is not round-specific. Instead, it is set at the time the person enters MEPS, and the person's keyness status never changes. Once a person is determined to be key, that person will always be key.

It should be pointed out that a person may be key even though not part of the civilian, non-institutionalized portion of the U.S. population. For example, a person in the military may have been living with his or her civilian spouse and children in a household sampled for the NHIS. The person in the military would be considered a key person for MEPS. However, such a person would not be eligible to receive a person-level sample weight if he or she was never in-scope during 2009. He or she may receive a family weight if a member of a responding family.

Eligibility

The issue of a person's eligibility for MEPS is a data collection issue. Data are to be collected only for persons considered eligible for MEPS.

All key, in-scope persons of a sampled RU are eligible for data collection. The only non-key persons eligible for data collection are those who happen to be living in an RU with at least one key, in-scope person. Their eligibility continues only for the time that they are living with at least one such person. The only out-of-scope persons eligible for data collection are those persons serving full-time on active duty in the military who were living with key in-scope persons, and again only for the time they are living with such a person.

A person may be classified as eligible for an entire round or for some part of a round. For persons who are eligible for only part of a round, data are collected for that person only for the period of time for which that person was classified as eligible.

Person Disposition Status

The variable PSTAT13 indicates a person's response and eligibility status. The PSTAT13 variable indicates the reasons for either continuing data collection for a person or terminating data collection for each person in the MEPS. Using this variable, one could identify persons who moved during the reference period, died, were born, were institutionalized or were in the military. Note that some categories may be collapsed for confidentiality purposes.

The following codes specify the value labels for the PSTAT13 variable. Note that some values for PSTAT13 are round-specific, as indicated in the labels.

Value	Definition
-1	The person was not fielded during the round or the RU was non-response
0	Incorrectly listed in RU at NHIS – applies to MEPS Round 1 only
11	Person in original RU , not full-time active military duty
12	Person in original RU, full-time active military duty, out-of-scope for whole reference period
13	Full-time student living away from home, but associated with sampled RU
14	The person is full-time active military duty during round, is in-scope for part of the reference period and is in the RU at the end of the reference period
21	The person remains in a health care institution for the whole round – Round 3 only
22	The person leaves an institution (health care or non-health care) and rejoins the community – Round 3 only
23	The person leaves an institution (health care or non-health care), rejoins the community and then dies – Round 3 only
24	The person dies in a health care institution during the round (former RU member) – Round 3 only
31	Person from original RU, dies during reference period
32	Went to health care institution during reference period
33	Went to non-healthcare institution during reference period
34	Moved from original RU, outside U.S. (not as student)
35	Moved from original RU, to a military facility while on full-time active military duty
36	Went to institution (type unknown) during reference period
41	Moved from the original RU, to new RU within U.S. (new RUs include RUs originally classified as “Student RU” but which converted to “New RU”)
42	The person joins RU and is not full-time military during round
43	The person’s disposition as to why the person is not in the RU is unknown or the person moves and it is unknown whether the person moved inside or outside the U.S.

Value	Definition
44	The person leaves an RU and joins an existing RU and is not both in the military and coded as inscope during the round
51	Newborn in reference period
61	Died prior to reference period (not eligible)-Round 3/1 only
62	Institutionalized prior to reference period (not eligible)-Round 3/1 only
63	Moved outside U.S., prior to reference period (not eligible)-Round 3/1 only
64	Full-time military, living on a military facility, moved prior to reference period (not eligible)-Round 3/1 only
71	Student under 24 living away at school in grades 1-12 (Non-Key)
72	Person is dropped from the RU roster as ineligible: the person is a non-key student living away or the person is not related to reference person or the RU is the person's residence only during the school year
73	Not Key and not full-time military, moved without someone key and inscope (not eligible)
74	Moved as full-time military but not to a military facility and without someone key and inscope (not eligible this round)
81	Person moved from original RU, full-time student living away from home, did not respond

Geographic Variables

Two variables, REGION13 and MSA13, indicate the geographic location of the reporting unit. REGION13 indicates the Census region the RU resides in at the time of the Round 1/Round 3 interview. The Census regions are defined by the U.S. Department of Commerce, Bureau of the Census. MSA13 indicates whether or not the RU is in a Metropolitan Statistical Area (MSA) and reflects the most recent definitions of metropolitan statistical areas established by Office of Management and Budget (OMB), including the most recent updates. These updates are based on the application of the 2000 Standards for Defining Metropolitan Statistical Areas of OMB to Census Bureau population estimates for July 1, 2004 and July 1, 2005. For MEPS data releases prior to 2004 the MSA variables were coded in compliance with the definition of metropolitan statistical areas based on application of OMB standards to Census 1990 data.

The values and states for each region include the following:

Value	Label	States
1	Northeast	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont
2	Midwest	Indiana, Illinois, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin
3	South	Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia
4	West	Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming

2.5.2 Demographic Variables

These variables provide information about the demographic characteristics of each person. As noted below, some variables have edited and imputed values. Values of most demographic variables on this file are obtained during each round of the MEPS interview. These variables describe data for Panel 13 Round 3 and Panel 14 Round 1, as well as a number of characteristics that are not round specific.

Age

Date of birth and age for each RU member were asked or verified during each MEPS interview (AGE13X, DOBMM, DOBY). If date of birth was available, age was calculated based on the difference between date of birth and date of interview. Inconsistencies between the calculated age and the age reported during the CAPI interview were reviewed and resolved. For purposes of confidentiality, the variable AGE13X was top coded at 85 years, and DOBY bottom coded at 1924. When date of birth was not provided but age was (from either the MEPS or the NHIS data), the month and year of birth were assigned randomly from among the possible valid options. For any cases still not accounted for, age was imputed using (1) the mean age difference between MEPS participants with certain family relationships (where available) or (2) the mean age value for MEPS participants. For example, a mother's age is imputed as her child's age plus the mean age difference between MEPS mothers and their children, or a wife's age is imputed as the husband's age plus the mean age difference between MEPS wives and husbands.

Sex

The variable SEX contains data on the sex of each RU member (SEX), as determined during the NHIS interview; it was verified and, if necessary, corrected during each MEPS interview. The data for new RU members (persons who were not members of the RU at the time of the NHIS interview) were also obtained during each MEPS round. When sex of the RU member was not available from the NHIS interview and was not ascertained during one of the subsequent MEPS interviews, it was assigned in the following way. The person's first name was used to assign sex, if obvious. If the person's first name provided no indication of gender, then family relationships

were reviewed. If neither of these approaches made it possible to determine the individual's sex, sex was randomly assigned.

Race and Ethnicity Group

The race and the ethnic background questions were asked for each RU member during the MEPS interview. If the information was not obtained in Round 1, the questions were asked in subsequent rounds. It should be noted that race/ethnicity questions in the MEPS were revised starting in 2002. Race/ethnicity data from earlier years are not directly comparable. The following table shows the differences:

MEPS Race and Ethnicity Variables, by Years

FY PUF 1996 – 2001	FY PUF 2002 – current
RACEX	RACEX
1 American Indian	1 White – No other race reported
2 Aleut, Eskimo	2 Black – No other race reported
3 Asian or Pacific Islander	3 American Indian/Alaska Native – No other race reported
4 Black	4 Asian – No other race reported
5 White	5 Native Hawaiian/Pacific Islander – No other race reported
91 Other	6 Multiple race reported
HISPANX	HISPANX
1 Hispanic	1 Hispanic
2 Not Hispanic	2 Not Hispanic
RACETHNX	RACETHNX
1 Person is Hispanic	1 Person is Hispanic
2 Person is Black/Not Hispanic	2 Person is Black – No other race reported/Not Hispanic
3 Other/Not Hispanic	3 Person is Asian – No other race reported/Not Hispanic
	4 Other race/Not Hispanic
HISPCAT	HISPCAT

FY PUF 1996 – 2001	FY PUF 2002 – current
-9 Not Ascertained	-9 Not Ascertained
-7 Refused	-8 DK
1 Puerto Rican	-7 Refused
2 Cuban	1 Puerto Rican
3 Mexican/Mexican American/Mexicano/Chicano	2 Cuban/Cuban American
4 Other Latin American/Other Spanish	3 Dominican
5 Non-Hispanic	4 Mexican/Mexican American
	5 Central or South American
	6 Non-Hispanic
	91 Other Latin American
	92 Other Hispanic/Latino
	RACEAX
	1 Asian – No other race reported
	2 Asian – Other race(s) reported
	3 All other race assignments
	RACEBX
	1 Black – No other race reported
	2 Black – Other race(s) reported
	3 All other race assignments
	RACEWX
	1 White – No other race reported
	2 White – Other race(s) reported
	3 All other race assignments

Values for these variables were obtained based on the following priority order. If available, data collected were used to determine race and ethnicity. If race and/or ethnicity were not reported in the interview, then data obtained from the originally collected NHIS data were used. If still not ascertained, the race, and/or ethnicity were assigned based on relationship to other members of the DU using a priority ordering that gave precedence to blood relatives in the immediate family (this approach was used on 56 persons to set race and 32 persons to set ethnicity).

Starting in 2002, individuals were allowed to choose more than one race and, as a result, three new variables were constructed: RACEBX, RACEAX, and RACEWX. RACEBX identifies individuals as being: 1) Black – no other race reported, 2) Black – other race(s) reported, or 3) not black. RACEAX and RACEWX are constructed similarly but apply to Asians and Whites. All race and ethnicity variables reflect the imputations done for RACEX and HISPANX. RACETHNX summarizes both race and ethnicity information in a single variable.

Marital Status and Spouse ID

Current marital status was collected and/or updated during each round of the MEPS interview. This information was obtained in RE13 and RE97 and is reported as MARRY13X. Persons under the age of 16 were coded as 6 “Under 16 - Inapplicable”. In instances where there were discrepancies between the marital statuses of two individuals within a family, other person-level variables were reviewed to determine the edited marital status for each individual. For example, in Panel 13 Round 3, when one spouse was reported as married and the other spouse reported as widowed, the data were reviewed to determine if one partner should be coded as 8 “Widowed in Round”.

Four edits were performed to ensure some consistency across rounds for the Panel 13 Round 3 data. First, a person could not be coded as “Never Married” after previously being coded as any other marital status (e.g., “Widowed”). Second, a person could not be coded as “Under 16 - Inapplicable” after being previously coded as any other marital status. Third, a person could not be coded as “Married in Round” after being coded as “Married” in the round immediately preceding. Fourth, a person could not be coded as an “in Round” code (e.g., “Widowed in Round”) in two subsequent rounds.

The person identifier for each individual’s spouse is reported in SPOUID13. The variable is set to the PID (within each family) of the person identified as the spouse during the round. If no spouse was identified in the household, the variable was coded as 995 “no spouse in house”. Those with unknown marital status are coded as 996 “marital status unknown”. Persons under the age of 16 are coded as 997 “Less than 16 years old”.

The SPOUIN13 variable indicates whether a person’s spouse was present in the RU during the round. If the person had no spouse in the household, the value was coded as 2. For persons under the age of 16 the value was coded as 3. The SPOUID13 and SPOUIN13 variables were obtained from RE76A, where the respondent was asked to identify how each pair of persons in the household was related. Analysts should note that this information was collected in a question separate from the questions that asked about marital status. While editing was performed to ensure that SPOUID13 and SPOUIN13 are consistent within each round, there was no consistency check between these variables and marital status in a given round. Apparent

discrepancies between marital status and spouse information may be due to any of the following causes:

- Ambiguity as to when during a round a change in marital status occurred. This is a result of relationship information being asked for all persons living in the household at any time during the round, while marital status is asked as of the interview date (e.g., If one spouse died during the reference period, the surviving spouse's marital status would be "widowed in round", but SPOUIN13 and SPOUID13 for the same round would indicate that a spouse was present);
- Valid discrepancies in the case of persons who are married but not living with their spouse, or separating but still living together; or
- Discrepancies which cannot be explained by either of the previous reasons.

Student Status and Educational Attainment

The variable FTSTD13X indicates whether the person was a full-time student at the interview date. This variable has valid values for all persons between the ages of 17 - 23 inclusive.

The variables indicating completed years of education when first entered MEPS (EDUCYR) and highest degree when first entered MEPS (HIDEG) were obtained from questions RE103-105. For Panels 13 and 14 (panels from which data are based), questions RE103-105 were asked only when persons first entered MEPS, which was Round 1 for most people.

For the completed years of education variable (EDUCYR), children who are 5 years of age or older when they first entered MEPS and who never attended school were coded as 0; children under the age of 5 years were coded as -1 "Inapplicable" regardless of whether or not they attended school.

The highest degree (HIDEG) was obtained from two questions: high school diploma (RE104) and highest degree (RE105). Persons under 16 years of age when they first entered MEPS were coded as 8 "Under 16 -Inapplicable". In cases where the response to the highest degree question was "no degree" and highest grade was 13 through 17, the variable was coded as 3 "high school diploma". If highest grade completed for those with a "no degree" response was "refused" or "don't know", the variable was coded as 1 "no degree". The user should note that the EDUCYR and HIDEG variables are unedited variables and minimal data cleaning was performed on these variables. Therefore, discrepancies in data may remain for these two sets of variables. Decisions as to how to handle these discrepancies are left to the analyst.

Military Service and Service Era

Information on active duty military status was collected during each round of the MEPS interview. Persons currently on full-time active duty status are identified in the variable ACTDTY13. Those under 16 years of age were coded as 3 "under 16-inapplicable" and those over the age of 59 were coded as 4 "over 59-inapplicable".

HONRDC13 indicates whether a person has ever been honorably discharged from active duty in the United States Armed Forces. Those under 16 years of age are coded as 3 "16 or Younger -

Inapplicable”, and those who are currently serving on full-time active duty are coded as 4 “Now Active Duty”.

Relationship to the Reference Person within Reporting Units

For each reporting unit (RU), the person who owns or rents the dwelling unit is usually defined as the reference person. For student RUs, the student is defined as the reference person. (For additional information on reference persons, see the documentation on Reference Person Identifiers in the Survey Administration section.) The variable RFREL13X indicates the relationship of each individual to the reference person of the reporting unit (RU) in a given round. For the reference person, this variable has the value “self”; for all other persons in the RU, relationship to the reference person is indicated by codes representing “husband/spouse,” “wife/spouse,” “son,” “daughter,” “female partner,” “male partner,” etc. A code of 91, meaning “other related,” was used to indicate rarely observed relationship descriptions such as “mother of partner”. If the relationship of an individual to the reference person was not ascertained during the round-specific interview, relationships between other RU members were used, where possible, to assign a relationship to the reference person. If MEPS data were not sufficient to identify the relationship of an individual to the reference person, relationship variables from the NHIS data were used to assign a relationship. In the event that a meaningful value could not be determined or data were missing, the relationship variable was assigned a missing value code. Note that relationship values occurring fewer than six times in the file are recoded to “other related” for purposes of confidentiality.

2.5.3 Health Status and Priority Condition Variables

Health Status variables involved the construction of person-level variables based on information collected in the Condition Enumeration, Priority Condition Enumeration, and Health Status sections of the questionnaire. The majority of Health Status questions were initially asked at the family level to ascertain if anyone in the household had a particular problem or limitation. These were followed up with questions to determine which household member had each problem or limitation. Logical edits were performed in constructing the person-level variables to ensure that family-level and person-level values were consistent. Particular attention was given to cases where missing values were reported at the family level to ensure that appropriate information was carried to the person level. Inapplicable cases occurred when a question was never asked because of skip patterns in the survey (e.g., individuals who were 13 years of age or older were not asked some follow-up verification questions). Inapplicable cases are coded as -1. In addition, for all variables, deceased persons were coded as inapplicable and received a code of -1.

Perceived Health Status and Mental Health Status

Perceived health status (RTHLTH13) and perceived mental health status (MNHLTH13) were collected in the Priority Condition Enumeration section. These questions (PE00A and PE00B) asked the respondent to rate the physical and mental health of each person in the family according to the following categories: excellent, very good, good, fair, and poor. No editing was done to these variables.

IADL and ADL Help/Supervision

The Instrumental Activities of Daily Living (IADL) Help or Supervision variable (IADLHP13) was constructed from a series of three questions. The initial question (HE01) determined if anyone in the family received help or supervision with IADLs such as using the telephone, paying bills, taking medications, preparing light meals, doing laundry, or going shopping. If the response was “yes”, a follow-up question (HE02) was asked to determine which household member received this help or supervision. For persons under age 13, a final verification question (HE03) was asked to confirm that the IADL help or supervision was the result of an impairment or physical or mental health problem. If the response to the final verification question was “no”, IADLHP13 was coded as “no” for persons under the age of 13.

If no one in the family was identified as receiving help or supervision with IADLs, all members of the family were coded as receiving no IADL help or supervision. In cases where the response to the family-level question was “don’t know”, “refused”, or otherwise missing, all persons were coded according to the family-level response. In cases where the response to the family-level question (HE01) was “yes” but no specific individuals were identified in the follow-up question as having IADL difficulties, all persons were coded as “don’t know” (-8).

The Activities of Daily Living (ADL) Help or Supervision variable (ADLHLP13) was constructed in the same manner as IADLHP13, but using questions HE04-HE06. Coding conventions for missing data were the same as for IADLHP13.

Functional Limitations

A series of questions pertained to functional limitations, defined as difficulty in performing certain specific physical actions. WLKLIM13 was the filter question. It was derived from a question (HE09) that was asked at the family level: Does anyone in the family have difficulties walking, climbing stairs, grasping objects, reaching overhead, lifting, bending or stooping, or standing for long periods of time? If the answer was “no”, then all family members were coded as “no” (2) on WLKLIM13. If the answer was “yes”, then the specific persons who had any of these difficulties were identified and coded as “yes” (1) on WLKLIM13, and remaining family members were coded as “no”. If the response to the family-level question was “don’t know” (-8), “refused” (-7), “missing” (-9), or “inapplicable” (-1), then the corresponding missing value code was applied to each family member’s value for WLKLIM13. If the answer to HE09 was “yes”, but no specific individual was named as experiencing such difficulties, then each family member was assigned -8 for WLKLIM13. Deceased respondents were assigned a -1 code (“inapplicable”) for WLKLIM13.

If any family member was coded “yes” to WLKLIM13, a subsequent series of questions was administered. The series of questions for which WLKLIM13 served as a filter was as follows:

- LFTDIF13 – difficulty lifting 10 pounds
- STPDIF13 – difficulty walking up 10 steps
- WLKDIF13 – difficulty walking 3 blocks
- MILDIF13 – difficulty walking a mile
- STNDIF13 – difficulty standing 20 minutes

BENDIF13 – difficulty bending or stooping
RCHDIF13 – difficulty reaching over head
FNGRDF13 – difficulty using fingers to grasp

This series of questions was asked separately for each person who was coded “yes” to WLKLIM13. This series of questions was not asked for other individual family members for whom WLKLIM13 was “no”. In addition, this series was not asked about family members who were less than 13 years of age, regardless of their status on WLKLIM13. These questions were not asked about deceased family members. In such cases (i.e., WLKLIM13 = 2, or age < 13, or PSTAT13 = (23,24,31)), each question in the series was coded as “inapplicable” (-1). Finally, if responses to WLKLIM13 were “refused” (-7), “don’t know” (-8), “not ascertained” (-9), or otherwise inapplicable (-1), then each question in this series was coded as “inapplicable” (-1).

Analysts should note that, for WLKLIM13, there was no minimum age criterion that was used to determine a skip pattern, whereas, for the subsequent series of questions, persons less than 13 years old were skipped and coded as inapplicable. Therefore, it is possible for someone aged 12 or less to have a code of 1 (“yes”) on WLKLIM13, and also to have codes of inapplicable on the subsequent series of questions.

Use of Assistive Technology and Social/Recreational Limitations

The variables indicating use of assistive technology (AIDHLP13, from question HE07) and social/recreational limitations (SOCLIM13, from question HE22) were collected initially at the family level. If there was a “yes” response to the family-level question, a second question identified the specific individual(s) to whom the “yes” response pertained. Each individual identified as having the difficulty was coded “yes” on the appropriate variable; all remaining family members were coded “no”. If the family-level response was “don’t know”, “refused”, or otherwise missing, all persons were coded with the family-level response. In cases where the family-level response was “yes” but no specific individual was identified as having difficulty, all family members were coded as “don’t know” (-8).

Work, Housework, and School Limitations

The variable indicating any limitation in work, housework, or school (ACTLIM13) was constructed using questions HE19-HE20. Specifically, information was collected initially at the family level. If there was a “yes” response to the family-level question (HE19), a second question (HE20) identified the specific individual(s) to whom the “yes” response pertained. Each individual identified as having a limitation was coded “yes” on ACTLIM13; all remaining family members were coded “no”. If the family-level response was “don’t know”, “refused”, or otherwise missing, all persons were coded with the family-level response. In cases where the family-level response was “yes” but no specific individual was identified as having difficulty, all family members were coded as “don’t know” (-8). Persons less than five years old were coded as “inapplicable” (-1) on ACTLIM13.

If ACTLIM13 was “yes” and the person was 5 years of age or older, a follow-up question (HE20A) was asked to identify the specific limitation or limitations for each person. These included working at a job (WRKLIM13), doing housework (HSELIM13), or going to school

(SCHLIM13). Respondents could answer “yes” to each activity; one person could thus report limitation in multiple activities. WRKLIM13, HSELIM13, and SCHLIM13 have values of “yes” or “no” only if ACTLIM13 was “yes”; each variable was coded as “Inapplicable” (-1) if ACTLIM13 was “No” (2). When ACTLIM13 was “Refused” (-7), these variables were all coded as “Refused” (-7); when ACTLIM13 was “Don’t Know” (-8), these variables were all coded as “Don’t Know” (-8); and when ACTLIM13 was “Not Ascertained” (-9), these variables were all coded as “Not Ascertained” (-9). If a person was under 5 years old or was deceased, WRKLIM13, HSELIM13, and SCHLIM13 were each coded as “Inapplicable” (-1).

A second question (HE21) asked if the person was completely unable to work at a job, do housework, or go to school. Those respondents who were coded “no,” “Refused,” “Don’t Know,” or “Not Ascertained” on ACTLIM13, or were under 5 years of age, or were deceased, were coded as “inapplicable” (-1) on UNABLE13. UNABLE13 was asked once for whichever set of WRKLIM13, HSELIM13, and SCHLIM13 the respondent had limitations; if a respondent was limited in more than one of these three activities, UNABLE13 did not specify if the respondent was completely unable to perform all of them, or only some of them.

Cognitive Limitations

The variable COGLIM13 was collected at the family level as a three-part question (HE24-01 to HE24-03) indicating if any of the adults in the family (1) experience confusion or memory loss, (2) have problems making decisions, or (3) require supervision for their own safety. If a “yes” response was obtained to any item, the persons affected were identified in HE25 and COGLIM13 was coded as “yes”. Remaining family members not identified were coded as “no” for COGLIM13.

If responses to HE24-01 through HE24-03 were all “no”, or if two of three were “no” and the remaining was “don’t know”, “refused”, or otherwise missing, all family members were coded as “no”. If responses to the three questions were combinations of “don’t know”, “refused”, and missing, all persons were coded as “don’t know”. If the response to any of the three questions was “yes” but no individual was identified in HE25, all persons were coded as “don’t know”.

COGLIM13 reflects whether any of the three component questions is “yes”. Respondents with one, two, or three specific cognitive limitations cannot be distinguished. In addition, because the question asked specifically about “adult” family members, all persons less than 18 years of age are coded as “inapplicable” (-1) on this question.

2.5.4 Employment Variables

Employment questions were asked of all persons 16 years and older at the time of the interview. Employment variables consist of person-level indicators such as employment status and job-related variables such as hourly wage. All job-specific variables refer to a person’s current main job. The current main job, defined by the respondent, indicates the main source of employment.

Employment variables included on the Panel 13 Round 3/Panel 14 Round 1 2009 release are: EMPST13, HRWAG13X, HRWGRD13, HRWAY13, HOUR13, HELD13X, OFFER13X,

NUMEMP13 and SELFCM13. Most employment variables pertain to status as of the date of the interview.

Logical edits were performed on variables that indicate whether health insurance is held or offered at a current main job. In addition, some wage information was logically edited for consistency. Edits were performed under three circumstances:

- in cases where a respondent updated a wage, indicating as the reason for the change that the amount reported in a previous round was in error, and then provided the corrected amount for the previous round;
- in some cases where wages reported as less than \$1.00 per hour are updated in a subsequent round to greater than \$1.00, and the wage increased by a factor of 10 or 100 (for example, if a Round 4 wage is updated to \$20.00, the Round 3 wage of \$0.20 could logically be updated to \$20.00); in some of these cases, additional comments may have also indicated an error; and
- in some cases where wages changed substantially from round to round and a keying error was evident (for example, ‘the number of hours on which the salary is based’ is updated from ‘40’ to ‘4’; the ‘4’ could logically be updated to ‘40’).

In all cases that result in an edit, a complete review of wage and employment history is performed; in some cases, comparisons are made to employment at similar establishments within the MEPS as well as to data reported and summarized by the Bureau of Labor Statistics.

When missing, wage values were imputed for certain persons’ hourly wage; however, there was no editing performed on any values reported by the respondent (except as noted above). Hourly wages greater than or equal to \$67.31 were top-coded to –10. The number of employees variable was top-coded at 500.

Employment Status (EMPST13)

Employment status was asked for all persons aged 16 or older. Responses to the employment status question were: “currently employed” if the person had a job at the interview date, “has a job to return to” if the person did not work during the reference period but had a job to return to as of the interview date, “employed during the reference period” if the person had no job at the interview date but did work during 2009, and “not employed with no job to return to” if the person did not have a job at the interview date, did not work during the reference period, and did not have a job to return to. These responses are mutually exclusive. A current main job was defined for persons reporting that they were currently employed and who identified a current main job, and for persons who reported and identified a job to return to. Therefore, job-specific information, such as hourly wage, exists for persons not presently working at the interview date but who have a job to return to.

Hourly Wage (HRWAG13X, HRWGRD13, and HRWAY13)

Hourly wage was asked of all persons who reported a current main job that was not self-employment (SELFCM13). For reasons of confidentiality, the hourly wage variable (HRWAG13X) was top-coded. A value of –10 indicates that the hourly wage was greater than or

equal to \$67.31. The hourly wage on this file (HRWAG13X) should be considered along with its accompanying variables HRWGRD13 and HRWAY13.

HRWGRD13 is a flag that indicates the round in which the reported hourly wage was collected. This flag is always set to “1” for people who are a part of Panel 14 because the reported hourly wage is always from Round 1 as only Round 1 information is reported on this file. People who are a part of Panel 13 can have a current main job from a previous round and HRWGRD13 indicates the round in which the wage information was collected. For Round 3 current main jobs that continue as the current main job from Round 1, HRWGRD13 is “1”. For Round 3 current main jobs that continue as the current main job from Round 2 (but not Round 1), HRWGRD13 is “2”. For Round 3 current main jobs that are identified as current main for the first time in Round 3, HRWGRD13 is “3”.

For persons who did not indicate a wage amount but who did indicate a range into which the hourly wage falls, the reported hourly wage (HRWAG13X) is the median within that range. The medians were calculated using actual wages reported from the same round by persons of the same gender reporting hourly wages within each age range category. In some cases, particularly in the low wage range, gender was not used in the calculation of the median wage in order to provide a large enough base.

HRWAY13 indicates how the corresponding HRWAG13X was constructed. Hourly wage was derived, as applicable, from a large number of source variables. In the simplest case, hourly wage was reported directly by the respondent. For other persons, construction of the hourly wage was based upon their salary, the time period on which the salary was based, and the number of hours worked per time period. If the number of hours worked per time period was not available, a value of 40 hours per week was assumed, as identified in the HRWAY13 variable.

Health Insurance (HELD13X and OFFER13X)

There are two employment-related health insurance measures included in this release: health insurance held from a current main job (HELD13X) and health insurance offered from a current main job (OFFER13X). The held and offer variables were logically edited using health insurance information from the health insurance section not available for public release.

Persons under age 16 as well as persons aged 16 and older who do not hold a current main job, who are self-employed with no employees, or who are otherwise ineligible are coded as “inapplicable” for both the health insurance-related employment variables.

HELD13X is “yes” if the person reported having insurance coverage from the employer or union at the current main job and that coverage provides hospital/physician or Medigap benefits (as long as the person is not self-employed with no employees). HELD13X is also “yes” if the person’s current main job is with the armed forces.

HELD13X is “no” if the person either reported that insurance is not obtained through the current main job or reported insurance and then disavowed it. To disavow insurance is to initially report it but then to deny that it is provided later in the interview, or to confirm it but indicate that it does not include hospital/physician or Medigap benefits. As noted above, this does not apply to

self-employed persons with no employees (always “inapplicable”) and those with a current main job in the armed forces (always “yes”).

OFFER13X is always coded as “yes” if HELD13X is “yes”. In addition, except for certain self-employed persons with OFFER13X set to “inapplicable” (see above), OFFER13X is coded as “yes” if insurance was offered through the employer or union at the job. OFFER13X is “no” when HELD13X equals “no” and insurance was not offered by the employer or union at the job.

As indicated above, information collected in the health insurance section of the interview was considered in the construction of HELD13X and OFFER13X. For example, several persons indicated in the employment section of the interview that they held health insurance through a current main job and then denied this coverage later in the health insurance section. Such people were coded as “no” for HELD13X. Due to questionnaire skip patterns, the value for HELD13X was considered in constructing the OFFER13X variable. For example, if a person responded that health insurance was held from a current main job, they were skipped past the question relating to whether health insurance was offered at that job. If the person later disavowed this insurance in the health insurance section of the questionnaire, we would not be able to ascertain whether they were offered a policy. These individuals are coded as -9 for OFFER13X.

Hours (HOUR13)

HOUR13 is the number of hours worked per week at the current main job.

Number of Employees (NUMEMP13)

Due to confidentiality concerns, the variable indicating the number of employees at the establishment (NUMEMP13) has been top coded at 500 or more employees. NUMEMP13 indicates the number of employees at the location of the person’s current main job. For persons who reported a categorical size, we report a median estimated size from within the reported range.

2.5.5 Health Insurance Variables

Constructed and edited variables are provided that indicate any coverage during the MEPS Panel 14 Round 1 and Panel 13 Round 3 interviews for the sources of health insurance coverage collected during the MEPS interview. With the exception of private insurance (PRIV13), the insurance variables for the Panel 14 Round 1 observations have been edited. For both the Panel 14 Round 1 sample and the Panel 13 Round 3 sample, minimal editing was performed on the Medicare and Medicaid or State Children’s Health Insurance Program (SCHIP) variables to assign persons to coverage from these sources. Beginning October 1, 2001, persons 65 years and older can retain TRICARE coverage in addition to Medicare. Therefore, persons over age 65 will no longer have their reported TRICARE coverage (TRINW13X) overturned. TRICARE will act as a supplemental insurance for Medicare much as Medigap insurance does now. As mentioned above, private insurance coverage was unedited and unimputed for Panel 14 Round 1. For Panel 13 Round 3, most of the insurance variables have been logically edited to address issues that arose during Rounds 2 and 3 when reviewing insurance reported in earlier rounds. One edit corrects for possible respondent confusion with respect to a question about covered benefits asked of respondents who reported a change in their private health insurance plan name.

Additional edits were performed to address issues of missing data on the time period of coverage. Note that the Medicare and TRICARE variables indicate coverage at the time of the Panel 14 Round 1 or Panel 13 Round 3 interview dates. The private coverage and other public insurance variables indicate coverage at any time during Panel 14 Round 1 or Panel 13 Round 3.

Public sources include Medicare, TRICARE, Medicaid, SCHIP, and other public hospital/physician coverage. State-specific program participation in non-comprehensive coverage (STPRG13) was also identified but is not considered health insurance for the purposes of this survey.

Medicare

Medicare (MCARE13) coverage was edited (MCARE13X) for persons age 65 or over. Within this age group, individuals were assigned Medicare coverage if:

- They answered “yes” to a follow-up question on whether or not they received Social Security benefits; or
- They were covered by Medicaid/SCHIP, other public hospital/physician coverage, or Medigap coverage; or
- Their spouse was age 65 or older and covered by Medicare; or
- They reported TRICARE coverage.

Medicaid and Other Public Hospital/Physician Coverage

Questions about other public hospital/physician coverage were asked in an attempt to identify Medicaid or SCHIP recipients who may not have recognized their coverage as such. These questions were asked only if a respondent did not report Medicaid or SCHIP directly. Respondents reporting other public hospital/physician coverage were asked follow-up questions to determine if their coverage was through a specific Medicaid HMO or if it included some other managed care characteristics. Respondents who identified managed care from either path were asked if they paid anything for the coverage and/or if a government source paid for the coverage.

The Medicaid variables (MCAID13) have been edited (MCAID13X) to include persons who paid nothing for their other public hospital/physician insurance when such coverage was through a Medicaid HMO or reported to include some other managed care characteristics. The Medicaid variables also include those identified as covered by SCHIP.

To assist users in further editing sources of insurance, this file contains variables constructed from the other public hospital/physician series that measure whether:

- The respondent reported some type of managed care and paid something for the coverage, Other Public A Insurance (OTPUBA13); or
- The respondent did not report any managed care, Other Public B Insurance (OTPUBB13).

The variables OTPUBA13 and OTPUBB13 are provided only to assist in editing and should not be used to make separate insurance estimates for these types of insurance categories.

Any Public Insurance in Round 3/Round 1

The file also includes a summary measure that indicates whether or not a sample person has any public insurance during the early part of 2009 (PUB13X). Persons identified as covered by public insurance are those reporting coverage under TRICARE, Medicare, Medicaid or SCHIP, or other public hospital/physician programs. Persons covered only by state-specific programs that did not provide comprehensive coverage (STPRG13), for example, Maryland Kidney Disease Program, were not considered to have public coverage when constructing the variable PUB13X.

Private Insurance

Variables identifying private insurance in general (PRIV13) and specific private insurance sources such as employer/union group insurance (PRIEU13); non-group insurance (PRING13); and other group insurance (PRIOG13) were constructed. Private insurance sources identify coverage in effect at any time during the early part of 2009. Separate variables identify covered persons and policyholders (policyholder variables begin with the letter “H”, e.g., HPRIEU13). These variables indicate coverage or policyholder status within a source and do not distinguish between persons who are covered or policyholders on one or more policy within a given source. In some cases, the respondent was unable to characterize the source of insurance (PRIDK13). Covered persons (but not policyholders) are identified when the policyholder is living outside the RU (PRIOU13). An individual was considered to have private health coverage if, at a minimum, that coverage provided benefits for hospital and physician services (including Medigap coverage). Sources of insurance with missing information regarding the type of coverage were assumed to contain hospital/physician coverage. Persons without private hospital/physician insurance were not counted as privately insured.

Health insurance through a job or union (PRIEU13, PRIS13) was initially asked about in the Employment Section of the interview and later confirmed in the Health Insurance Section. Respondents also had an opportunity to report employer and union group insurance (PRIEU13) for the first time in the Health Insurance Section, but this insurance was not linked to a specific job.

All insurance reported to be through a job classified as self-employed with firm size of 1 (PRIS13) was initially reported in the Employment Section and verified in the Health Insurance Section. Unlike the other employment-related variable (PRIEU13), self-employed with firm size of 1 (PRIS13) insurance could not be reported in the Health Insurance Section for the first time. The variable PRIS13 has been constructed to allow users to determine if the insurance should be considered employment-related.

Private insurance that was not employment-related (PRING13, PRIOG13, PRIDK13, and PRIOU13) was reported in the Health Insurance Section only.

Any Insurance in Round 3 / Round 1

The file also includes a summary measure that indicates whether a sample person has any insurance during the early part of 2009 (INSRD13X). Persons identified as insured are those reporting coverage under TRICARE, Medicare, Medicaid, SCHIP, or other public

hospital/physician or private hospital/physician insurance (including Medigap plans). A person is considered uninsured if not covered by one of these insurance sources.

Persons covered only by state-specific programs that provide non-comprehensive coverage (STPRG13), for example, Maryland Kidney Disease Program, and those without hospital/physician benefits (for example, private insurance for dental or vision care only, accidents or specific diseases) were not considered to have public coverage when constructing the variable INSRD13X.

2.6 Linking to Other Files

2.6.1 National Health Interview Survey

Each MEPS panel can also be linked back to the previous year's National Health Interview Survey public use data files. For information on obtaining MEPS/NHIS link files please see www.meps.ahrq.gov/data_stats/more_info_download_data_files.jsp.

2.6.2 Longitudinal Analysis

For Panels 1 through 8, panel-specific files (called Longitudinal Weight Files) containing estimation variables to facilitate longitudinal analysis are available for downloading in the data section of the MEPS Web site. To create longitudinal files for these panels, it is necessary to link data from two subsequent annual files that contain data for the first and second years of the panel, respectively. Starting with Panel 9, it is not necessary to link files for longitudinal analysis because Longitudinal Data Files have been constructed and are available for downloading on the web.

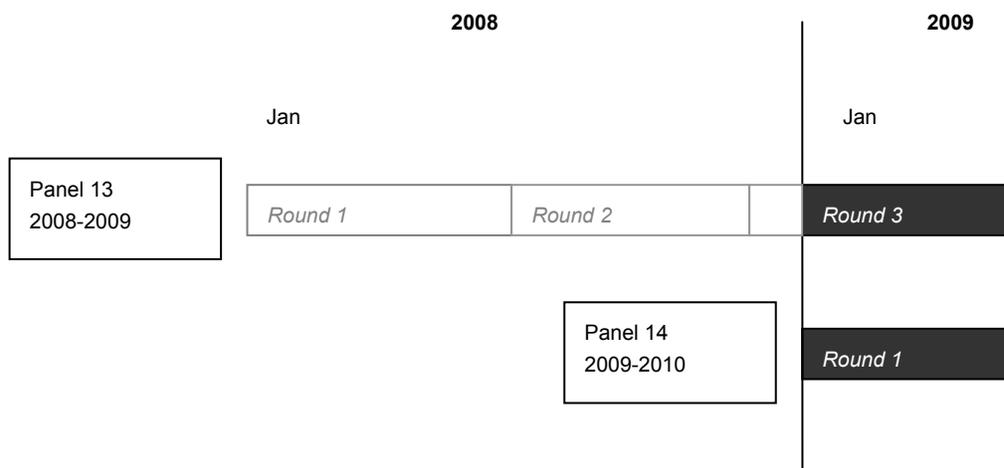
3.0 Survey Sample Information

3.1 Sample Design and Response Rates

The MEPS HC is designed to produce estimates at the national and regional levels over time for the civilian, non-institutionalized population of the United States and some subpopulations of interest. The MEPS HC uses an overlapping panel design in which data for two calendar years are obtained through five rounds of data collection.

A new sample (new Panel) of households for MEPS is selected each year from among household respondents to the previous year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics, Centers for Disease Control and Prevention. (The NHIS is an ongoing general health survey of the U.S. civilian non-institutionalized population.)

MEPS Panel 13 spans the two calendar years 2008 and 2009 while MEPS Panel 14 spans 2009 and 2010. This file consists of the subset of data from the thirteenth and fourteenth MEPS panels covering from January 1 through, roughly, the spring of calendar year 2009. More specifically, data from the 2009 portion of the third Round of data collection for the MEPS Panel 13 sample are pooled with data from the first Round of data collection for the MEPS Panel 14 sample (see illustration below).



Traditionally, the sample for the NHIS is redesigned and redrawn about every ten years. From 1995 to 2005 the NHIS used the same sample design, and thus the MEPS, which began in 1996, has been based on a single NHIS design through MEPS Panel 11 initiated in 2006. Since, in the NHIS, the same PSUs and second stage sampling units are used each year, the MEPS sample from its inception has likewise been clustered within these same sampling units. However, a new sample design for the NHIS was implemented in 2006. The fundamental structure of the new 2006 NHIS sample design is very similar to the previous 1995-2005 NHIS sample design. The sample PSUs and second stage sampling units for the new NHIS design were selected independent of the sample selection process under the previous design. Of course, there is some overlap between the area populations covered by the sampled PSUs selected under the two designs, mostly the larger ones selected with certainty. As households selected for MEPS participation are selected from among the previous year's NHIS respondents, the MEPS Panel 12, fielded in 2007, was the first MEPS Panel based on the new NHIS sample design. As a result, the 2007 PIT file consisted of two independent samples, serving to increase precision (due to decreased clustering since the sample was spread out among more PSUs and secondary sampling units) and increased degrees of freedom (due to many more variance strata available for variance estimation purposes).

As with the 2008 PIT file, both MEPS Panels (13 and 14) in the 2009 PIT file are based on a single NHIS sample design.

3.1.1 The MEPS Sampling Process and Response Rates: An Overview

Generally, a sample representing about three-eighths of the NHIS responding households is made available for use in MEPS. This was the case for Panels 13 and 14.

A subsample of the responding NHIS households is drawn for MEPS interviewing. 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 eligible for MEPS is slightly larger than the 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 to 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 dropped from the sample. For the NHIS, college students living in student housing are sampled independently of 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 the three informational components just discussed in Rows A, B, and C. 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). From these numbers unweighted response rates have been computed, allowing an assessment of the propensity to participate among those RUs actually sampled for MEPS. Those actually sampled for MEPS include a disproportionately high number of members of several minorities as well as those "predicted to be poor" (based on a statistical model) compared to the distribution of such groups in the general population. The "predicted to be poor" were oversampled for Panel 13 only.

Table 3-1. Unweighted response rates for Point-in-Time file (Panel 14 Round 1/Panel 13, Round 3)

	Panel 13	Panel 14	2009 Combined
A. Percentage of NHIS sample eligible for MEPS	87.4%	85.2%	
B. Number of households sampled from the NHIS	9,703	9,700	
C. Number of Households sampled from the NHIS and fielded for MEPS	9,688	9,672	
D. Round 1 – Number of RUs eligible for interviewing	10,325	10,227	
E. Round 1 – Number of RUs with completed interviews	8,017	7,650	
F. Round 2 – Number of RUs eligible for interviewing	8,252		
G. Round 2 – Number of RUs with completed interviews	7,809		
H. Round 3 – Number of RUs eligible for interviewing	7,982		
I. Round 3 – Number of RUs with completed interviews	7,684		
Overall response rates through the Spring of 2009 P13: $A \times (E/D) \times (G/F) \times (I/H)$ P14: $A \times (E/D)$ Combined: $.495 \times P13 + .505 \times P14$	61.8% (Panel 13 through Round 3)	63.7% (Panel 14) through Round 1)	62.8%

When an RU is visited for a round of data collection, changes in RU membership are identified. Such changes include RU members who have moved to another location in the U.S., thus creating a new RU to be interviewed for MEPS, as well as student RUs. Thus, the number of RUs known to be eligible for MEPS interviewing in a given round can only be 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 specifically 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 14, Round 1 the ratio of 7,650 (Row E) to 10,227 (Row D) multiplied by 100 is 74.8 and represents the unweighted response rate for the round conditioned on the set of RUs characterized as eligible for MEPS for Round 1, expressed as a percentage. Multiplying the percentage of the NHIS sample eligible for MEPS (row A) by the product of the ratios for a consecutive set of MEPS rounds beginning with round one produces the overall response rate through the last MEPS round specified. (It should be noted that the number of RUs with completed interviews is slightly higher than the number of RUs receiving family weights. RUs receiving family weights must satisfy additional criteria such as participation throughout the entire period of eligibility by all key, in scope RU members and the requirement that the RU reference person must be key.)

The overall response rate for the combined sample of Panel 14, Round 1 and Panel 13, Round 3 is obtained by taking the sum of the products of the relative sample sizes and the corresponding overall panel response rates. Panel 13, Round 3 represents about 49.5 percent of the combined sample size, while Panel 14, Round 1 represents approximately 50.5 percent. Thus, the combined response rate has been computed as .495 multiplied by the Panel 13 response rate through Round 3 plus .505 multiplied by the Panel 14 response rate through Round 1.

As a point of information, for Panel 13 an experiment has been conducted with respect to the effectiveness of various levels of incentives to increase response rates and/or the quality of the data collected. Analyses of the results of the experiment are currently being conducted.

3.1.2 Panel 13 Response

For MEPS Panel 13 Round 1 9,688 households were fielded in 2008 (row C of Table 3.1), a nationally representative subsample of the households responding to the 2007 National Health Interview Survey (NHIS).

Table 3-1 shows the number of RUs eligible for interviewing in each Round of Panel 13 as well as the number of RUs completing the MEPS interview. Computing the individual Round “conditional” Response Rates as described in section 3.1.1 and then taking the product of the resulting three “conditional” round response rates and the factor 87.4 (the percentage of the NHIS sampled households eligible for MEPS) yields an overall unweighted response rate of 61.8 percent for Panel 13 through Round 3.

3.1.3 Panel 14 Response

For MEPS Panel 14, 9,672 households were fielded in 2009 (again, found in Row C of Table 3.1), a nationally representative subsample of the households responding to the 2008 National Health Interview Survey (NHIS).

Table 3-1 shows the number of RUs eligible for interviewing (10,227) and the number completing the interview for Round 1 of Panel 14 (7,650). The overall unweighted response rate for Panel 14 through Round 1 of MEPS is thus computed as 85.2 percent times $(7,650/10,227)$, or 63.7 percent.

3.1.4 Combined Panel Response

A combined 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 13 response rate was weighted by a factor of .495 while that of Panel 14 by a factor of .505, reflecting approximately the distribution of the sample sizes between the two panels. The resulting unweighted response rate for the combined panels is $(.495 \times 61.8)$ plus $(.505 \times 63.7)$ or 62.8 percent (as shown in Table 3-1).

3.1.5 Oversampling

Oversampling was employed for selected subgroups of policy-level interest to help increase the precision of estimates associated with members of those subgroups. Before going into details, the concept of oversampling is 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 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, subgroups that are “oversampled” are represented at disproportionately high rates in the sample. Sample weights help ensure that population estimates account for this disproportionate contribution from oversampled subgroups, as the 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 a subgroup is done to improve the precision of survey estimates for that particular subgroup. 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.

For MEPS, some of the oversampling was achieved through its linkage to the NHIS. For the earlier sample design of the NHIS, 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 NHIS oversampling rate for black households was roughly 1.5 to 1. For the new NHIS sample design Asians are also oversampled. The estimated overall oversampling rates associated with each of the three minorities have not yet been reported.

The oversampling approaches and the sampling domains used for MEPS Panels 13 and 14 for subsampling among the NHIS respondents eligible for MEPS were slightly different.

From among the NHIS households eligible for the Panel 13, four domains (strata) were established in a hierarchical sequence but in essence there were only two sampling strata employed. The first stratum contained households with Asians and those “predicted to be poor”, a second stratum contained households with Hispanics not assigned to the first stratum, a third stratum contained households with black members in households not assigned to the first two strata, while the fourth stratum contained all remaining households. All households in the strata of “Asian/Predicted Poor”, Hispanics, and Blacks were sampled with certainty. The sampling rate for the “other” stratum was about 57 percent.

Similar to Panel 13, there were also four strata established for Panel 14, with the only difference that the first stratum only contained households with Asians. Those households “predicted to be poor” did not constitute a separate sampling domain for Panel 14. All households in the Asian stratum were sampled with certainty. The sampling rate was about 87 percent for the Hispanics stratum, about 90 percent for the Blacks stratum, and about 72 percent for the “other” stratum.

Within strata for both panels, responding NHIS households were selected for MEPS using a systematic sample selection procedure from among those eligible. With the subsampling, households that were oversampled for calendar year 2009 were the households containing Hispanics, Blacks, and Asians, based on their NHIS membership.

3.2 Sample Weights

The sample weights provided in this file can be used to produce cross-sectional estimates for the U.S. civilian, non-institutionalized population for the first half of 2009 and subgroups of this population based on the sample data. Two weights are provided: a person-level weight and a family-level weight.

3.2.1 Person-level Weight

The person-level weight variable (WGTS13) was constructed as a composite of separate panel specific weights. A positive person-level weight was assigned to all key members of the U.S. civilian, non-institutionalized population for whom MEPS data were collected, representing the corresponding U.S. population in early 2009. For the Panel 14 Round 1 participants, this weight reflects the original household probability of selection for the NHIS, a factor representing the proportion of the 16 NHIS panel-quarter combinations eligible for MEPS, the oversampling of the subgroups described earlier, ratio-adjustment to NHIS national population estimates at the household level, adjustment for non-participation in MEPS at the household or dwelling unit

level, and poststratification to U.S. civilian noninstitutionalized population estimates obtained from March 2009 CPS data at the family and person levels. For both panels 13 and 14 separately person-level poststratification reflected population distributions across census region, MSA status, race/ethnicity (Hispanic, black/non-Hispanic, Asian, other), sex, and age. Finally, a composite weight was assigned to each responding person and a final poststratification was undertaken across the variables of census region, MSA status, race/ethnicity, sex, and age.

Table 3-2 shows the number of persons with person weights for each of the two panels separately, as well as the combined total and the total population estimate represented by the weighted total for all persons with person-level weights. In terms of numbers of persons, there are 18,779 for Panel 13, Round 3 and 19,191 for Panel 14, Round 1. Thus, in total, there are 37,970 sample persons in the file with positive person-level weights (WGTSP13>0). The corresponding estimate for the civilian, noninstitutionalized population based on summing the weights found in the variable WGTSP13 for these 37,970 persons is 300,544,183.

Table 3-2. Persons with a person weight for the 2009 Point-in-Time file

	Panel 13	Panel 14	Combined	Population estimate (weighted total of combined sample)
Number	18,779	19,191	37,970	300,544,183

A Note on Population Estimates

Beginning with the 2001 Full Year data, MEPS transitioned to 2000 census-based population estimates for poststratification and raking. In addition, MEPS population estimates underwent some “discontinuities” due to adjustments made to the CPS estimates as of 2003 (CPS being the source of the control figures used for raking and poststratification in MEPS). Those who wish to learn about these changes in CPS population estimates may consult the report “Revisions to the Current Population Survey Effective in January 2003” from the January 2003 issue of the monthly Labor Review. This report was authored by Mary Bowler, Randy E. Ilg, Stephen Miller, Ed Robison, and Anne Polivka, all at the Bureau of Labor Statistics. Differences in the way racial categories are defined are also noted in this report.

3.2.2 Family-level Weight

3.2.2.1 Definition of MEPS Families

A family unit is defined in MEPS as two or more persons living together in the same household during the reference period (in this data set, from January 1, 2009 to the date of interview) who are related by blood, marriage, or adoption (including foster children). In addition, unrelated persons who identify themselves as a family (e.g., domestic partners) are also defined as a MEPS family unit. Persons who died during the Round 1 reference period and those who left the civilian, non-institutionalized population part way through the reference period due to institutionalization, emigration, or enrollment in the military were considered to be family members. Relatives identified as usual residents of the household but who were not present at the

time of the interview, such as college students living away from their parents' home during the school year, were considered as members of the family that identified them.

3.2.2.2 Assignment of Weights

If all key, in-scope members of a family responded to MEPS for their entire period of eligibility for Panel 13, Round 3 or for Panel 14, Round 1 and if the family had a key reference person, then that family received a family-level weight ($WGTRU13 > 0$). Reporting units consisting of an individual respondent who was both key and in-scope also received a family-level weight. These single person “family” units can be included or excluded from family-level analyses at the analyst’s discretion.

Family-level weights were poststratified to figures obtained from the March 2009 CPS. The family-level poststratification reflects population distributions across family type (reference person married, spouse present; male reference person, no spouse present; female reference person, no spouse present), size of family, age of reference person, location of family (census region and MSA status), and race/ethnicity of the family’s reference person.

Table 3-3 shows the number of families with family-level weights for each of the two panels separately, as well as the combined total and the total population estimate represented by the weighted total for all families with family-level weights. Included as families in these counts are individuals living in one person RUs. There are 7,346 such families for Panel 13, Round 3 and 7,488 for Panel 14, Round 1. Thus, in total, there are 14,834 sample families in the file with positive family-level weights ($WGTRU13 > 0$). The population estimate of the number of these “family” units (families plus single person “family” units) with family-level weights containing at least one member of the U.S. civilian, non-institutionalized population is 131,223,210 based on summing the family level weights across all 14,834 MEPS families where $WGTRU13$ is positive.

Table 3-3. Families with a family weight for the 2009 Point-in-Time file

	Panel 13	Panel 14	Combined	Population estimate (weighted total of combined sample)
Number	7,346	7,488	14,834	131,223,210

It should be noted that CPS and MEPS definitions of family units are slightly different. In particular, CPS does not include foster children in families or consider unmarried persons who live together as family units. Adjustments were made in the poststratification process to help compensate for some of these differences. Again, note that MEPS population estimates have undergone some “discontinuities” due to an adjustment in the 2003 CPS estimates.

3.2.2.3 Instructions to Create Family Estimates

To make estimates at the family level, it is necessary to prepare a family-level file containing one record per family. Each MEPS family unit is uniquely identified by the combination of the variables DUID and FAMID13. Only persons with positive, nonzero family weight values

(WGTRU13>0) are candidates for inclusion in family estimates. Following is a summary of steps that can be used for family-level estimation:

1. Concatenate the variables DUID and FAMID13 into a new variable (e.g., DUFAM13).
2. To create a family-level file, sort by DUFAM13 and then subset to one record per DUFAM13 value by retaining only the reference person record (RNDREF13=1) for each value of DUFAM13. If the analyst chooses to eliminate single person units from family analyses, it is also necessary to exclude records where FAMSIZ13=1. If aggregate measures for families are needed for analytic purposes (e.g., means or totals), then those measures need to be computed using person-level information within families and attached to the family record. For other types of variables, analysts frequently use characteristics of the reference person to represent family characteristics.
3. Apply the weight WGTRU13 to the analytic variable(s) of interest to obtain national family estimates.

3.2.3 Relationship between Person- and Family-Level Weights

Some persons with positive person-level weights do not have family-level weights because at least one member of their family was a non-participant in MEPS. Others with positive person weights did not receive a family weight because the family reference person was not key. In addition, some persons with positive family-level weights do not have person-level weights because they were either non-key or a member of the military or otherwise out-of-scope during the 2009 portion of the MEPS data collection round. Analysts should include only persons with positive person-level weights for analyses focused on the civilian, noninstitutionalized population or subgroups of this population. Analyses focused on members of families should include persons with positive family-level weights. Family level analyses can be undertaken as described in Section 3.2.2.3.

3.3 Variance Estimation

MEPS has 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. Replicate weights have not been developed for the MEPS data. Instead, the variables needed to calculate appropriate standard errors based on the Taylor-series linearization method are included on this point-in-time file as well as 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.

Initially, 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 rounds. However, beginning with the 2002 Point-in-Time PUF, the variance strata and PSUs were developed to be compatible with all future PUF until the NHIS design changed. As discussed, this change took place in 2006, effectively changing the MEPS design beginning with calendar year 2007, where Panel 12 was based on the new NHIS design while Panel 11 was based on the old one. Thus, in order 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 will have four digit values while those associated with the old design will have three digit values.

For the 2007 PIT data a temporary set of variance strata and PSUs were developed for use with data collected under the new NHIS sample design. The current set of variance strata and PSUs were re-established for the 2008 PIT data, and were carried over to the PIT 2009 data. Specifically, 165 variance estimation strata were created, each stratum with either two or three variance estimation PSUs. These have been numbered 1001-1165 for the 2009 PIT file.

Beginning with the 2002 PIT data base 203 variance strata were formed for use in developing variance estimates for all subsequent years and data bases under the old design. These were numbered 1-203. For data analyses where data pooling across calendar years is limited to 2002 and later, the numbering of the variance strata and variance PSUs now permits this with no further actions needed.

If pooled analyses involve data in calendar years earlier than 2002, a pooled linkage file has been created to permit assignment of variance strata and PSU values for any person sampled under the old NHIS sample design (the one used for the NHIS from 1995-2005, and thus associated with MEPS samples for MEPS Panels 1-11). This person-level file contains variance stratum and PSU variables for all respondents participating in MEPS, along with the standard MEPS person ID variables for linking to other MEPS files. This one file contains records for each person who is on any of the MEPS full-year consolidated files. It is found on PUF Number HC-036. (A Balanced Repeated Replicate or BRR version of this file is also available. See PUF Number HC-036BRR.)

3.4 Using MEPS Data for Trend Analysis

MEPS began in 1996 and the utility of the survey for analyzing health care trends expands with each additional year of data. However, it is important to consider a variety of factors when examining trends over time using MEPS. Statistical significance tests should be conducted to assess the likelihood that observed trends are attributable to sampling variation. The length of time being analyzed should also 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. Looking at changes over longer periods of time can provide a more complete picture of underlying trends. Analysts may wish to consider using techniques to smooth or stabilize analyses of trends using MEPS data such as comparing pooled time periods (e.g. 1996-97 versus 2004-05), working with moving averages, or using modeling techniques with several consecutive years of MEPS data to test the fit of specified patterns over time. Of course, researchers should be aware of the impact of multiple comparisons on Type I error. Performing numerous hypothesis tests to help identify existing trends increases the likelihood of inappropriately concluding there is a trend (because a test indicated statistical significance) when, in fact, there is not.

Finally, it should be noted that standard errors for differences over time should be computed reflecting the correlation between MEPS samples where it exists. MEPS panels from 2007 through 2009 share the same sample PSUs and secondary sampling units. As a result, the estimated standard error of the difference between two MEPS samples will generally be reduced due to this correlation, to the extent that there is a positive correlation between estimates over time. Failure to reflect this aspect of the MEPS sample design (i.e., treating MEPS estimates as having come from independent samples) can generally be expected to result in the estimated standard error of the difference overstating the actual standard error, thus reducing the power to detect existing differences over time. Variance estimation software packages designed for complex samples, such as SUDAAN, provide the capability to reflect the correlation between MEPS samples when estimating the standard error for an estimated difference over time.

D. Variable-Source Crosswalk

SURVEY ADMINISTRATION VARIABLES

VARIABLE	LABEL	SOURCE
DUID	Dwelling Unit ID	Assigned in Sampling
PID	Person Number	Assigned in Sampling or by CAPI
DUPERSID	Person ID (DUID + PID)	Assigned in Sampling
PANEL	Panel Number	Assigned by CAPI
FAMID13	Family Identifier (Student Merged In)	CAPI Derived
RULETR13	RU Letter	CAPI Derived
RUSIZE13	RU Size	CAPI Derived
RUCLAS13	RU Fielded As: Standard, New, Student	CAPI Derived
FAMSI13	RU Size Including Students	CAPI Derived
REGION13	Census Region	Assigned in Sampling
MSA13	MSA	Assigned in Sampling
RNDREF13	Reference Person	RE 42-45
RDRESP13	1st Respondent Indicator	RE 6, 8
PROXY13	Was Respondent A Proxy	RE 2
BEGRFD13	Reference Period Begin Date: Day	CAPI Derived
BEGRFM13	Reference Period Begin Date: Month	CAPI Derived
BEGRFY13	Reference Period Begin Date: Year	CAPI Derived
ENDRFD13	Reference Period End Date: Day	CAPI Derived
ENDRFM13	Reference Period End Date: Month	CAPI Derived
ENDRFY13	Reference Period End Date: Year	CAPI Derived
KEYNESS	Person Key Status	RE Section
INSCOP13	Inscope	RE Section
PSTAT13	Person Disposition Status	RE Section
RURSLT13	RU Result	Assigned by CAPI
RUENDD13	Date of Intv (Date Started: Day)	Assigned by CAPI
RUENDM13	Date of Intv (Date Started: Month)	Assigned by CAPI
RUENDY13	Date of Intv (Date Started: Year)	Assigned by CAPI

DEMOGRAPHIC VARIABLES

VARIABLE	LABEL	SOURCE
AGE13X	Age - (Edited/Imputed)	RE 12, 57-66
DOBMM	Date of Birth: Month	RE 12, 57-66
DOBY	Date of Birth: Year	RE 12, 57-66
SEX	Sex	RE 12, 57, 61
RACEBX	Black Among Races Rptd (Edited/Imputed)	RE 101A
RACEAX	Asian Among Races Rptd (Edited/Imputed)	RE 101A
RACEWX	White Among Races Rptd (Edited/Imputed)	RE 101A
RACEX	Race - (Edited/Imputed)	RE 101A
RACETHNX	Race/Ethnicity - (Edited/Imputed)	RE 98A-101A
HISPANX	Hispanic Ethnicity - (Edited/Imputed)	RE 98A-100A
HISPCAT	Specific Hispanic Ethnicity Group	RE 98A-100A
MARRY13X	Marital Status - (Edited/Imputed)	RE 13, 97
SPOUID13	Spouse ID	RE 76A
SPOUIN13	Marital Status with Spouse Present	RE 76A
EDUCYR	Years of Educ when First Entered MEPS	RE 103-105
HIDEG	Highest Degree when First Entered MEPS	RE 103-105
FTSTD13X	Student Status Ages 17-23 (Edit/Imputed)	RE 11A, 106-108
ACTDTY13	Military Full-Time Active Duty	RE14, 96
HONRDC13	Honorably Discharged from Active Duty	RE18A, RE96G
RFREL13X	Relation To Ref Pers (Edited/Imputed)	RE 76A

HEALTH STATUS VARIABLES

VARIABLE	LABEL	SOURCE
RTHLTH13	Perceived Health Status	PE 00A
MNHLTH13	Perceived Mental Health Status	PE 00B
IADLHP13	IADL Screener	HE 1, 2, 3
ADLHLP13	ADL Screener	HE 4, 5, 6
AIDHLP13	Uses Assistive Devices	HE 7,8
WLKLIM13	Limitation in Physical Functioning	HE 9,10
LFTDIF13	Difficulty Lifting 10 Pounds	HE 11
STPDIF13	Difficulty Walking Up 10 Steps	HE 12
WLKDIF13	Difficulty Walking 3 Blocks	HE 13
MILDIF13	Difficulty Walking a Mile	HE 14
STNDIF13	Difficulty Standing 20 Minutes	HE 15
BENDIF13	Difficulty Bending/Stooping	HE 16
RCHDIF13	Difficulty Reaching Over Head	HE 17
FNGRDF13	Difficulty Using Fingers to Grasp	HE 18
ACTLIM13	Limitation Work/Housework/School	HE 19,20
WRKLIM13	Work Limitation	HE 19,20
HSELIM13	Housework Limitation	HE 19,20
SCHLIM13	School Limitation	HE 19,20
UNABLE13	Completely Unable To Do Activity	HE 21
SOCLIM13	Social Limitation	HE 22,23
COGLIM13	Cognitive Limitation	HE 24,25

EMPLOYMENT VARIABLES

VARIABLE	LABEL	SOURCE
EMPST13	Employment Status	EM 1-3; RJ 1, 6
HRWAG13X	Hourly Wage at Current Main Job (Edited)	EW section; EM 104-105, 111
HRWGRD13	Hourly Wage Round Flag	Constructed
HRWAY13	Calculation Methods for Hourly Wage	EM 104-105, 111; EW section
HOUR13	Hours Worked Per Week at CMJ	EM 104-105, 111; EW 17; RJ 1
HELD13X	Health Insurance Held From CMJ (Ed)	EM, HX, RJ and HP sections
OFFER13X	Health Insurance Offered at CMJ (Ed)	EM, HX, RJ and HP sections
NUMEMP13	Number of Employees at Location of CMJ	EM 91-92, 124; RJ 8B
SELFCM13	Self-Employed at Current Main Job	EM 5, 11, 18, 27, 40, 53

HEALTH INSURANCE VARIABLES

VARIABLE	LABEL	SOURCE
TRINW13X	PID Cov By TRICARE/CHAMPVA at Int - Edited	HX 12, 13; PR 19 - 22; HQ section; RE 14, 96A
MCARE13	PID Cov By Medicare	HX 5 – 7
MCARE13X	PID Cov By Medicare - Edited	HX 5 - 7, 10 - 15; PRIV13 and (HX 48 or (OE 10, 24, 37)); PR 7-10, 19-26
MCAID13	PID Cov By Medicaid or SCHIP	HX 10, 11, 14, 15, 18, 19; HQ section; PR 7-10, 23-26, 39-42
MCAID13X	PID Cov by Medicaid or SCHIP - Edited	MCAID13, HX 10, 11, 14, 15, 18, 19; 41-43, 45; HQ section; PR 7-10, 11-14, 23-32, 39-42
OTPUBA13	PID Cov By/Pays Oth Gov Mcaid/SCHIP HMO	HX 14, 15, 41-45; HQ section; PR 23-30
OTPUBB13	PID Cov By Oth Public not Mcaid/SCHIP HMO	HX 14, 15, 41-43; HQ section; PR 23-30
STPRG13	PID Cov By State Specific Program	HX 16-19; HQ section; PR 35-38
PUB13X	PID Cov By Public Ins - Edited	TRINW13X, MCARE13X, MCAID13X, OTPUBA13, OTPUBB13
PRIEU13	PID Cov By Priv Empl/Union Plan	HX 2-4, 21-24 48; HP, OE, HQ, EM, and RJ sections

VARIABLE	LABEL	SOURCE
PRIDK13	PID Cov By Priv DK Plan	HX 21-24 48; HP, OE, and HQ sections
PRING13	PID Cov By NonGroup Plan	HX 21-24 48; HP, OE, and HQ sections
PRIOG13	PID Cov By Oth Group Plan	HX 21-24 48; HP, OE, and HQ sections
PRIS13	PID Cov By Self-Emp -1 Ins	HX 3, 4, 48; EM, RJ, OE, and HQ sections
PRIOUT13	PID Cov By Holder Outside RU	HX 21-24 48; HP, OE, and HQ sections
PRIV13	PID Cov By Private Ins	PRIEU13, PRIDK13, PRING13, PRIOG13, PRIS13, PRIOUT13
HPRIEU13	PID is Holder of Priv Empl/Union Plan	HX 2-4, 9, 11, 21-24 48; HP, OE, HQ, EM, and RJ sections
HPRIDK13	PID is Holder of Priv DK Plan	HX 11, 21-24 48; HP, OE, and HQ sections
HPRING13	PID is Holder of NonGroup Plan	HX 11, 21-24 48; HP, OE, and HQ sections
HPRIOG13	PID is Holder of Oth Group Plan	HX 11, 21-24 48; HP, OE, and HQ sections

VARIABLE	LABEL	SOURCE
HPRIS13	PID is Holder of Self-Emp -1 Ins	HX 3, 4, 9, 48; EM, RJ, OE, and HQ sections
HPRIV13	PID is Holder of Priv Ins Plan	HPRIEU13, HPRIDK13, HPRING13, HPRIOG13, HPRIS13
INSRD13X	PID is Insured - Edited	PUB13X, PRIV13

WEIGHTS VARIABLES

VARIABLE	LABEL	SOURCE
WGTSP13	Person Weight	Constructed
WGTRU13	Family Weight	Constructed
VARSTR	Variance Estimation Stratum	Constructed
VARPSU	Variance Estimation PSU	Constructed