MEPS HC-223:
2020 Person Round Plan Public Use File

August 2022

Due to the COVID-19 pandemic, changes were made to the 2020 MEPS data collection that analysts should keep in mind when doing trend analysis and pooling years of data. 1) The MEPS moved primarily to a phone rather than in-person survey. 2) Panels 23 and 24 were extended to nine rounds (four years) of data collection as opposed to the historical five rounds (two years). Because of the unforeseeable nature of the pandemic, data collection for 2020 included Round 5 interviews for Panel 23 that were fielded under the assumption that that interview would be the panel's last interview. Researchers using variables related to the first interview of the calendar year should read the documentation for their specific variables to understand the sources of the values for Panel 23.

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

Table of Contents

Section			Page
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 Description	C-1
		2.1 Complex File Structure with Examples	C-4
		2.2 Identifiers	C-6
		2.3 Reserved Codes	C-9
		2.4 Adding the Characteristics of Covered Persons	C-9
		2.5 Adding the Policyholder's Characteristics	C-9
		2.6 Choosing PRPL Records for Your Research Question	C-10
	3.0	Data File Contents	C-12
		3.1 ID Variables	C-12
		3.2 Person Variables	C-13
		3.3 Policyholder Variables	C-14
		3.4 Establishment Variables	C-15
		3.4.1 Employers, Exchanges/Marketplaces, and Other Establishments	C-16
		3.4.2 Types of Coverage through the Establishment	C-18
		3.4.3 Out-of-Pocket Premiums	C-19
		3.5 Plan Variables	C-21
		3.5.1 Household Reports of HMOs	C-21
		3.5.2 Change in Plan Name	C-22
		3.6 Links to Jobs Providing Insurance	C-22

Section			Page
	4.0	Linking to Other Files	C-24
		4.1 National Health Interview Survey	C-24
		4.2 Longitudinal Analysis	C-24
	5.0	Using MEPS Data for Trend Analysis	C-24
		References	C-27
D.		Variable-Source Crosswalk	D-1

A. Data Use Agreement

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

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

1. No one is to use the data in this data set in any way except for statistical reporting and analysis; and

2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) the Director Office of Management AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity; and

3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel Survey or the National Health Interview Survey. Furthermore, linkage of the Medical Expenditure Panel Survey and the National Health Interview Survey may not occur outside the AHRQ Data Center, NCHS Research Data Center (RDC) or the U.S. Census RDC network.

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

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

B. Background

1.0 Household Component

The Medical Expenditure Panel Survey (MEPS) provides nationally representative estimates of health care use, expenditures, sources of payment, and health insurance coverage for the U.S. civilian 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 and two additional rounds in 2020 covering a third year to compensate for the smaller number of completed interviews in Panel 25, 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. In 2006, the NHIS implemented a new sample design, which included Asian persons in addition to households with Black and Hispanic persons in the oversampling of minority populations. NHIS introduced a new sample design in 2016 that discontinued oversampling of these minority groups.

2.0 Medical Provider Component

Upon completion of the household CAPI interview and obtaining permission from the household survey respondents, a sample of medical providers are contacted by telephone to obtain information that household respondents cannot accurately provide. This part of the MEPS is called the Medical Provider Component (MPC) and information is collected on dates of 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 amounts of payment. The MPC is not designed to yield national estimates. It is primarily used as an imputation source to supplement/replace household reported expenditure information.

3.0 Survey Management and Data Collection

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

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of micro data files and tables via the MEPS website.

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

C. Technical and Programming Information

1.0 General Information

This public use data file contains data for each person with private health insurance reported in Rounds 6 and 7 of Panel 23, Rounds 3, 4, and 5 of Panel 24, and Rounds 1, 2, and 3 of Panel 25 (i.e., the rounds for the survey panels covering calendar year 2020) of the Medical Expenditure Panel Survey Household Component (MEPS HC). 2020 is the first data year to include three panels of data; Panel 23 was extended to include Rounds 6 and 7. Released as an ASCII file (with related R, SAS, SPSS, and Stata programming statements and data user information) and in a SAS data set, SAS transport format, Stata data set, and Excel file, this public use file (PUF) provides information collected on a nationally representative sample of the civilian noninstitutionalized population of the United States during the calendar year 2020. The HC-223 file (Person-Round-Plan Public Use File) contains records for persons insured through private establishments providing hospital/physician, Medigap, dental, vision, or prescription medication coverage and includes variables pertaining to HMOs. Records contained in this file are associated with persons in MEPS 2020 Full-Year Consolidated Public Use File (HC-224). Prior to 2020, the Person-Round-Plan Public Use File included persons who were in either the Full Year Consolidated or the Point-in-Time Round 1/Round 3 Population Characteristics Public Use File. Persons only in the point-in-time PUF are persons who did not respond in every round of the delivery year; these persons are excluded from the Full-Year Consolidated file. Because point-in-time data are no longer released to the public, coverage records for these persons are not releasable for delivery in the HC-223 final file. Note, however, that a policyholder record not present in the Full-Year Consolidated PUF will be retained if it is associated with a covered dependent with a positive weight in the Full-Year Consolidated PUF. The file contains 57 variables and has a logical record length of 233 with an additional 2-byte carriage return/line feed at the end of each record.

Users of the 2020 Person-Round-Plan (PRPL) file should take careful note of special considerations noted in the “Survey Sample Information” and the “Using MEPS Data for Trend Analysis” sections of the HC-244 documentation regarding modifications to sampling, collection, and trends in other federal samples due to COVID-19.

2.0 Data File Description

The Person-Round-Plan (PRPL) file for 2020 is a complex file of privately insured persons and their private health insurance plans and links to any jobs providing insurance. The PRPL file is designed to facilitate research on the sometimes complex and dynamic relationships between consumers and their private insurance. It is not a person-level file, and linking the PRPL file to a person-level file (such as the 2020 Full-Year Consolidated File (HC-224)) requires users to make analytic decisions based on understanding the complexity of the PRPL file.

Records contain the following types of information (Figure 1):

- Covered person
  - Flags for whether the person is the policyholder or a dependent
  - Whether enrolled at time of interview[1]
  - Months enrolled during the reference period of the interview
- Interview Round
- Policyholder
- Establishment providing insurance
  - Type of establishment (employer, union, insurance agent, etc.)
  - Types of coverage (hospital/physician, Medigap, dental, vision, prescription medication, Consolidated Omnibus Budget Reconciliation Act (COBRA), single or family)[2]
- Out-of-pocket premiums and employee contributions
- Links to the job providing insurance (for employment-based insurance only, HC-218, the JOBS Public Use File)[3]
- Plan (for hospital/physician and Medicare supplemental insurance coverage only)
  - Household reports of health maintenance organizations (HMOs)
  - High deductibles and Health Savings Accounts (HSAs) for hospital/physician insurance coverage only

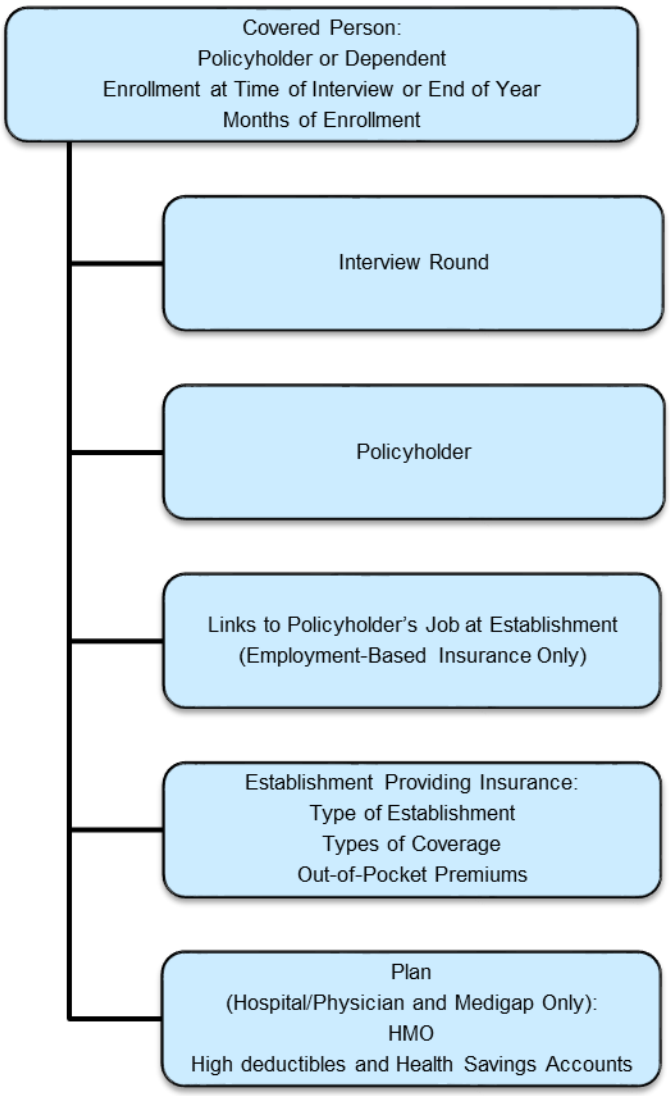
Notes pertaining to Figure 1 descriptions:

[1] Due to the extended length of the Round 6 reference period, the new variable EVALCOV5 indicating insurance status at the Round 5 interview date is available along with EVALCOVR. More information is provided in the "Person Variables" section.

[2] No effort has been made to validate variables representing type of coverage with external sources

[3] Users should take note of special Employment data processing information contained in the section below entitled "Links to Jobs Providing Insurance."

FIGURE 1
CONCEPTUAL OVERVIEW OF PRPL



For employer-sponsored health insurance coverage, on the records for dependents, variables link to the *policyholder's* job providing insurance, rather than the dependent's job.

"Establishment" refers to the organization through which the policyholder obtains private insurance. The establishment may be an employer, a union, an insurance agent, an insurance company, a professional association, or another type of organization. Many questions in the MEPS HC instrument were asked in reference to the establishment providing insurance to the policyholder. For example, the MEPS HC asked about the "types of health insurance" or covered

services, such as hospital/physician and dental coverage, the policyholder gets through the establishment.

For each establishment, a "plan" is the insurance company or HMO or self-insured company from which the policyholder receives hospital/physician or Medicare supplemental (Medigap) coverage. For some focused analyses, it may be important to recognize that information collected at the establishment level does not necessarily pertain to the plan level. For example, if a policyholder obtains from the establishment two separate plans, a hospital/physician plan and a dental plan, then the dental plan may not have the same HMO characteristics as the hospital/physician plan.

2.1 Complex File Structure with Examples

The PRPL file is designed to reflect the sometimes complex and dynamic relationships between people and their private insurance. It allows maximum flexibility for researchers, but it also requires that they make analytical decisions in their research.

The PRPL file is a person-round-policyholder-establishment-insurance plan level file with one record for each unique combination of establishment (source of private insurance), policyholder, interview round, 3-byte unique insurance plan identifier and covered person (policyholder or dependent). The 3-byte insurance plan number is helpful in differentiating between plans when more than one is reported through the same establishment-policyholder-interview round. Thus, the PRPL file contains at least one record for each person in each round with private health insurance, or 45,214 total records. (Panel 23 was extended to include Rounds 6 and 7, but the reference period for Round 5 ended on December 31, 2019, so the 2020 PRPL file does not include Round 5 records for Panel 23 respondents.) The PRPL file contains records for persons insured through establishments providing hospital/physician, Medigap, dental, vision, or prescription medication coverage. In most cases in this file, one person in the family has insurance from his or her employer, and this insurance covers everyone in the family. In this case, there is one record for each family member in each round, and each record flags the policyholder's current main job and links to one job record in the 2020 Jobs PUF (HC-218). However, other cases are more complex, and some hypothetical examples follow.

Multiple Establishments

- Juan and Maria are both employed parents, both have health insurance through their employers, and both parents choose family coverage. In this case, there are two PRPL records for each family member in each round.

- Fatema and Ali are both employed parents. Ali has single coverage from his employer. Fatema has family coverage from her employer covering the entire family. In this case, Fatema and the children each have one PRPL record for each round. Ali has two records for each round.

- Dexter has Medicare and Medicare supplemental insurance. In this case, Dexter has one PRPL record in each round for the Medicare supplemental insurance. There is no record for Medicare, because it is public insurance.

- Arlene is a child living with her mother. Both have Medicaid. Arlene's father, who does not live with them, has private insurance that covers Arlene. Arlene has one PRPL record in each round for the private insurance. There is no record for Medicaid, because it is public insurance.

No Private Insurance

- Jamie is uninsured. In this case, Jamie does not have any PRPL records.

- Cecilia has Medicaid instead of private coverage. In this case, Cecilia does not have any PRPL records.

Sources of Insurance: Employers and Other Establishments

- Rochelle is an employed parent with family coverage through her current main job. In this case, each family member's PRPL record flags Rochelle's current main job as the source of insurance, and each family member's PRPL record links to Rochelle's job record in PUF HC-218.

- Claire is employed, but she does not have insurance through her job. Instead, she buys a plan directly from an HMO. In this case, Claire has one PRPL record that does not flag her current main job and does not link to any job records in PUF HC-218.

- Felix has hospital/physician insurance through his employer, and he buys dental insurance through an insurance agent. In this case, Felix has two PRPL records, and only the employment-based insurance flags his current main job and links to a job record in PUF HC-218.

Policyholders Not in the Household

- Edith is a widow and has retiree insurance from her deceased husband's former job. In this case, Edith's PRPL record does not link to any employment information in the MEPS. There is a PRPL record for Edith's deceased husband, where he is flagged as the policyholder and flagged as deceased, but this record does not link to any records on any other MEPS PUFs.

- Matilda's parents are divorced. She lives with her father, but her insurance is through her mother's job. In this case, Matilda's PRPL record does not link to any employment information in the MEPS. There is also a PRPL record for Matilda's mother, where she is flagged as the policyholder and not residing in the Reporting Unit (RU), but this PRPL record does not link to any records on any other MEPS PUFs.

Changes in Insurance

- Isaac changes jobs between January 1st, 2020 and the date of his MEPS interview, and both jobs provide health insurance. In this case, Isaac has two PRPL records for the round, one for each source of coverage through a job. EVALCOVR shows whether one or both plans covered Isaac on the interview date.

- Julie reports that she worked at a job in Round 1 (Panel 25) but that she quit sometime before the interview. She began paying her previous employer to continue her health insurance while she looks for another job during Round 2. In this case, Julie's Round 1 PRPL record links to a job record in PUF HC-218. The Round 2 PRPL record has no job to link to in Round 2. Instead, the PRPL record is set to link with the Round 1 job. Even though she was actively employed at some point in Round 1, the PRPL record is not flagged as being through a current main job since she is not employed at the end of Round 1. The same is true for her Round 2 PRPL record.

Same Insurance Source Providing Multiple Plans

- Savannah reports working for a company that provides health insurance in which she enrolls. During the Round 2 interview, Savannah reports retiring from the job. Also in Round 2, her employer hires her back as a consultant. When reporting each job, Savannah indicates that she has insurance through both jobs. Therefore, in Round 2, Savannah has two unique PRPL records that are each linked to a separate job. Both 'plans' are through the same establishment-policyholder-round but each has its own unique insurance plan identifier.

2.2 Identifiers

Each record contains the following ID variables:

DUPERSID is the person identifier (either a dependent or a policyholder).

RN is the round of the interview in which the enrollment data were collected.

PHLDRIDX is the person identifier of the policyholder.

ESTBIDX is an ID number for the establishment - employer, union, insurance company, or other - i.e., the source of insurance coverage on the record.

InsurPrivIDEX is the concatenation of ESTBIDX and a 3-byte insurance number. It uniquely identifies establishment-insurance plans in a responding unit.

EPRSIDX is the concatenation of ESTBIDX, PHLDRIDX, RN, and 3-byte insurance number and it uniquely identifies the insurance plan that a policyholder obtains from an individual establishment.

EPCPIDX is the concatenation of EPRSIDX and DUPERSID, and it uniquely identifies each record.

JOBSIDX is the concatenation of the PHLDRIDX, a round identifier (RN), and a job number, and it uniquely identifies the policyholder's job at the establishment that provided insurance (for employment-based coverage).

For each person covered by a policyholder-establishment combination, the PHLDRIDX, ESTBIDX, EPRSIDX, and InsurPrivIDEX appear on each plan record for that coverage.

A person (DUPERSID) can be listed more than once on this file if (1) they are covered (as a policyholder or a dependent) by insurance policies from more than one establishment, (2) they are covered by multiple insurance policies from the same establishment, or (3) they are covered in more than one round. Within each round, establishment-policyholder pairs (EPRSIDXs) can be listed more than once if the health plan a policyholder obtains from a given establishment also covers their dependents.

As noted above, there is a PRPL record for each unique combination of establishment (source of insurance), round, insurance number, and covered person (policyholder or dependent). The following table presents a hypothetical example that illustrates the relationship between the ID variables on this file.

Row #	ESTBIDX	DUPERSID	PHLDRIDX	RN	Insurance #	**EPRSIDX*** PANEL #+ ESTBIDX+ PHLDRIDX+ RN+ Insurance #	**EPCPIDX*** PANEL#+ ESTBIDX+ PHLDRIDX+ RN+Insurance #+ DUPERSID	**JOBSIDX*** PANEL #+ PHLDRIDX+ RN+Job #
	A RU = 1							
1	101	104	104	1	101	231011041101	231011041101104	231041101
2	101	104	104	2	101	231011042101	231011042101104	231042101
3	101	104	104	3	101	231011043101	231011043101104	231043101
	B RU = 2							
4	201	206	206	1	202	232012061202	232012061202206	232061201
5	202	206	206	1	203	232022061203	232022061203206	-1
	C RU = 3							
6	301	301	301	3	302	233013013302	233013013302301	233013301
7	301	301	301	3	303	233013013303	233013013303301	233013301
8	301	302	301	3	303	233013013303	233013013303302	233013301
9	301	303	301	3	303	233013013303	233013013303303	233013301

Row #	ESTBIDX	DUPERSID	PHLDRIDX	RN	Insurance #	EPRSIDX* PANEL #+ ESTBIDX+ PHLDRIDX+ RN+ Insurance #	EPCPIDX* PANEL#+ ESTBIDX+ PHLDRIDX+ RN+Insurance #+ DUPERSID	JOBSIDX* PANEL #+ PHLDRIDX+ RN+Job #
	D RU = 4							
10	401	401	401	1	402	234014011402	234014011402401	234011401
11	401	402	401	1	402	234014011402	234014011402402	234011401
12	401	403	401	1	402	234014011402	234014011402403	234011401
	E RU = 5							
13	501	503	503	1	501	235015031501	235015031501503	235031502
14	501	503	503	2	501	235015032501	235015032501503	235031502

***NOTE: Panel number for the records are added as prefixes to each of these identifiers in the actual PRPL data.**

The first three rows of the table represent a situation where a person (DUPERSID=104) is listed three times in the PRPL file because she obtains insurance from the same establishment in all three rounds. Since the person is the policyholder, her DUPERSID is the same as the PHLDRIDX, which is repeated in the EPRSIDX, EPCPIDX, and JOBSIDX.

The fourth and fifth rows of the table represent a situation where a person (DUPERSID=206) is listed twice in the PRPL file because she obtains insurance from more than one establishment. In this example, the second establishment is not an employer or union, so JOBSIDX is -1 INAPPLICABLE.

The sixth, seventh, eighth, and ninth rows of the table represent a situation where a policyholder obtains coverage for themselves under one policy and obtains separate coverage for two dependents under another policy. Both policies are through the same ESTBIDX. The policyholder's PHLDRIDX appears in both EPRSIDX. These different plans are uniquely identified by the insurance number (EPRSIDX=233013013**302** and 233013013**303**). PRPL will create a policyholder record for situations where the policyholder is not covered under the plan.

The tenth, eleventh, and twelfth rows of the table represent a situation where a policyholder and two dependents obtain coverage through the policyholder's employer (a unique establishment-policyholder pair within each round, EPRSIDX=234014011402). The policyholder's PHLDRIDX appears in the EPRSIDX and the JOBSIDX for all three covered persons.

The last two rows of the table represent a situation where a person is retired and has retiree insurance through a job that ended prior to the current delivery year. In Round 1 of the first panel, the respondent reported the job from which the sample member retired, and MEPS does not ask about that job again.

However, in each round we ask about the health insurance. So in Round 2, the JOBSIDX contains round number 1, when the jobs data were last collected.

Finally, note that EPCPIDX uniquely identifies each record on the file.

In order to conduct person-level analyses, it is necessary to identify all policies that cover each individual either as a policyholder or as a dependent. Since each *person* in the PRPL file is uniquely identified by the variable DUPERSID, person-level analyses can be conducted by examining all PRPL records containing each DUPERSID.

2.3 Reserved Codes

The following reserved code values are used:

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

The value -15 (CANNOT BE COMPUTED) is assigned to MEPS constructed variables in cases where there is not enough information from the MEPS instrument to calculate the constructed variables. “Not enough information” is often the result of skip patterns in the data or from missing information resulting from MEPS responses of -7 (REFUSED) or -8 (DK). Note that reserved code -8 includes cases where the information from the question was “not ascertained” or where the respondent chose “don’t know”.

2.4 Adding the Characteristics of Covered Persons

The DUPERSID allows users to link on the age, sex, race, health status, or other person-level variables from the other HC files. However, this will result in multiple records per person, and estimates will not be nationally representative unless you use one PRPL record per person or summarize PRPL records to the person level (and use weights).

2.5 Adding the Policyholder’s Characteristics

The PHLDRIDX allows you to link characteristics of the policyholder onto the records of every person covered by the plan. For example, suppose you wanted to study persons whose private employment-based insurance is through an employee working full time at a current main job as of the first interview of 2020 (Panel 25 Round 1 or Panel 24 Round 3 or Panel 23 Round 6). You would select PRPL records matching HC-224 (FYFLG=1) where the insurance is through a current main job (CMJINS=1) and [(PANEL=25 and RN=1) or (PANEL=24 and RN=3) or

(PANEL=23 and RN=6)]. From HC-224, select the DUPERSID and HOUR31 variables and rename DUPERSID to PHLDRIDX. Merge HOUR31 onto the PRPL file by PHLDRIDX.

Some policyholders do not have records on HC-224. These include 1) deceased policyholders, 2) policyholders residing outside the RU, and 3) other policy holders not in HC-224 who covered one or more dependents with a positive weight in HC-224. All of the covered person records for these establishment-policyholder pairs are flagged with DECPHLDR, OUTPHLDR, or NOPUFLG equal to 1, respectively. FYFLG may be equal to 0 for these policyholders, depending on when the policyholder left the RU. Deceased policyholders complicate the estimation of nationally representative statistics on active policies. For these establishment-policyholder pairs, users must choose a covered person with a positive weight. However, when creating nationally representative estimates of policies and policyholders, establishment-policyholder pairs where the policyholder resides outside the RU should not be included in estimates. This is because MEPS policyholders include policies covering dependents outside the RU, and including RU members covered by a policyholder outside the RU will result in double counting policies that span RUs. Alternatively, a researcher could create nationally representative estimates of covered persons, regardless of whether the policyholder was in the RU, using all the covered persons in the MEPS.

2.6 Choosing PRPL Records for Your Research Question

In order to produce estimates from the data in this file, researchers must use the person-level or family-level weights released in the Consolidated PUF (HC-224). Researchers must consult the documentation for this PUF for guidance on creating nationally representative estimates for different time periods.

Note that if there are multiple records per person (DUPERSID) when you merge on weights, you will double count some people, and your estimates will not be nationally representative. There are two solutions: select only one record per person, or aggregate information across PRPL records.

Users will have additional considerations when linking the PRPL file to obtain job characteristics for third Panel 23 respondents. This is because a job may have ended on or before 2018 and therefore only exist in 2018 Jobs PUF (HC-203). As with prior years, some jobs for the second year panel, Panel 24, may have ended on or before 2019 and therefore only exist in 2019 Jobs PUF (HC-211).

How you develop your analytical file depends on your research question. The PRPL file is designed to help answer a wide variety of research questions. AHRQ cannot anticipate all these questions, so this section provides examples of how to use the PRPL file for four research questions.

How many people were covered by two or more private hospital/physician insurance plans at the end of 2020?

Select the Panel 25 Round 3, Panel 24 Round 5, and Panel 23 Round 7 records with PrivateCat>0 and MSUPINSX ne 1 and EVALCOVR=1. Count the number of records for each person (DUPERSID). Create one person-level record for each DUPERSID that has the number of plans (PRPL records). Merge the count variable onto PUF HC-224 and use weights, strata, and PSUs to create nationally representative estimates.

How many people reported private dental coverage from an employer at the end of 2020?

Select the Panel 25 Round 3, Panel 24 Round 5, and Panel 23 Round 7 records with DENTLINS=1 and PrivateCat in (1, 4) and EVALCOVR=1. Among these records, select one record for each person (DUPERSID). Merge each record onto PUF HC-224 and use weights, strata, and PSUs to create nationally representative estimates.

At the time of the first interview of 2020, how many private insurance policies for hospital/physician were not employment-based?

Select the Panel 25 Round 1 and Panel 24 Round 3 records with PrivateCat in (2, 3, 5, 6) and MSUPINSX ne 1 and EVALCOVR=1. Additionally, select the Panel 23 Round 6 records with PrivateCat in (2, 3, 5, 6) and MSUPINSX ne 1 and EVALCOV5=1. (Using EVALCOV5 for Panel 23 ensures the 'first interview' for this panel is of comparable length to the interviews for Panel 25 Round 1 and Panel 24 Round 3.) Next, select one record for each policyholder-establishment pair (EPRSIDX). To have a positive weight for the final count, we recommend choosing the covered person record of the policyholder (PHOLDER=1), unless the policyholder is deceased (DECPHLDR=1), in which case then the researcher should choose a different covered person's record. Merge each record onto PUF HC-224 and use weights, strata, and PSUs to create nationally representative estimates.

At the time of the first interview of 2020, how many people had insurance from jobs from which they retired?

Select the PRPL records for policyholders of employment-related insurance at the time of the first interview [(Panel 25 Round 1 or Panel 24 Round 3) and PHOLDER=1 and PrivateCat =1 and EVALCOVR=1]. From the 2020 Jobs PUF (HC-218), using JOBSIDX, select the records with jobs from which the person retired (SUBTYPE=6 or RETIRJOB=1 or YNOBUSN_M18=2 or WHY_LEFT_M18=3). Persons in Panel 24 may have reported retiring from a job in 2019, so, from the 2019 Jobs PUF (HC-211), using JOBSIDX select the records with PANEL=24 and (SUBTYPE=6 or RETIRJOB=1 or YNOBUSN=2 or WHY_LEFT=3). Combine the records from the two Jobs files, keeping only one record per JOBSIDX. Using the JOBSIDX, merge the selected JOBS records onto the selected PRPL records. Select the PRPL records with SUBTYPE=6 or RETIRJOB=1 or YNOBUSN_M18=2 or WHY_LEFT_M18=3 or EMPLSTAT=2. Select one record for each DUPERSID. Merge each record onto PUF HC-224 and use weights, strata, and PSUs to create nationally representative estimates of the number of people with one of these PRPL records. If users also wish to include Panel 23 Round 6 in this analysis, they should refer to JOBSFILE variable in the PRPL file to determine if links to the

2018 Jobs PUF (HC-203) or 2019 Jobs PUF (HC-211) are required to obtain job characteristics used in this example.

3.0 Data File Contents

3.1 ID Variables

In the MEPS Household Component, the definitions of Dwelling Units (DUs) and Group Quarters are generally consistent with the definitions employed for the National Health Interview Survey. The dwelling unit ID (DUID) is a seven-digit ID number consisting of a 2-digit panel number followed by a five-digit random number assigned (see below) after the case was sampled for MEPS. The three-digit person number (PID) uniquely identifies each person within the dwelling unit. The ten-character variable DUPERSID uniquely identifies each person represented on the file and is the combination of the variables DUID and PID. The MEPS HC-223 PRPL file can be linked to other public use files, such as MEPS HC-224, by using the DUPERSID.

For detailed information on dwelling units and families, please refer to the documentation for the 2020 Full-Year Consolidated PUF (HC-224).

Because all identifiers in this file begin with DUID, all identifiers begin with the panel number. In addition, the first byte of all CAPI-generated 3-byte identifiers represents the RU in numeric format (e.g. an A RU plan begins with '1' and a D RU plan begins with '4'). The last two bytes are generated from a counter *within* the RU.

PHLDRIDX is the person identifier (DUPERSID) of the policyholder of the private health insurance plan. Generally, the characteristics of the policyholder can be linked from person-level public use files by using the PHLDRIDX to match the DUPERSID on the person-level files. However, when the policyholder is deceased or resides outside the RU, then there are no person-level variables on public use files. Where the policyholder was alive and resided in the household at some point during the time periods covered by the interviews, the policyholder identifier in the EPRSIDX will change between rounds. InsurPrivIDEX can be used to track the insurance across rounds. The InsurPrivIDEX will identify all records, including the first reported through the plan so it can be used to track the change in policyholder. A more in-depth discussion of this can be found in Section 3.3.

ESTBIDX is a combination of DUID, "0", and a three-byte establishment number. The three-byte establishment number consists of the RU number of the RU where the establishment was reported, followed by a two-digit counter. It is assigned to place of employment and to sources of insurance. The combination of these elements ensures uniqueness within the RU.

EPRSIDX is a combination of ESTBIDX, PHLDRIDX, RN, and insurance number. In a few cases, more than one EPRSIDX may identify a policyholder-source of coverage pair. For example, when an RU splits, through divorce or because a child goes to college, each new RU separately reports insurance information, and hence MEPS cannot determine with certainty

whether members in both RUs have the same policy. Although both RUs may report coverage through the same policyholder, the RUs may have different EPRSIDXs and ESTBIDXs. (The RU number is embedded in the ESTBIDX and EPRSIDX.) For each RU (EPRSIDX), there is a PRPL record for the policyholder as a covered person, but for only one of the EPRSIDXs (the one in which the policyholder resides) is the policyholder coded as having coverage in the STAT## or EVALCOVR variables. The variable EVALCOV5 can be used if users wish to analyze Panel 23 insurance status in a time period comparable in length to Panel 25 Round 1 or Panel 24 Round 3.

InsurPrivIDEX uniquely identifies insurance policies from the same establishment and is made up of a combination of ESTBIDX and a 3-byte insurance number. The insurance number is a unique identifier within CAPI at the RU-level that identifies each insurance policy. InsurPrivIDEX is particularly helpful when tracing the history of a plan where a policyholder leaves a household (PID changes to 902) or returns to a household (PID changes from 902 to an RU member's PID). These records will not link based on ESTBIDX-PHLDRIDX. InsurPrivIDEX continues to be the most reliable way to identify all records across rounds.

JOBSIDX is a combination of the PHLDRIDX, a round identifier (RN), and a three-byte job number, and it uniquely identifies the policyholder's job at the establishment that provided insurance (for employment-based coverage) in the round. The job number is composed of the RU number of the RU where the job was reported, followed by a two-digit counter. The combination of these three preclude re-use of the two-digit counter. The round identifier embedded in JOBSIDX is the round in which the job was last reported, which is not necessarily the round in which the insurance was last reported (for example, when the job ended but the insurance continued). JOBSIDX can be used to link on characteristics of the policyholder's job providing insurance from the Jobs PUF in 2020 (HC-218) or in 2019 (HC-211) or, for Panel 23, 2018 (HC-203) as well.

Users should take special care when working with EPRSIDX and EPCPIDX in Excel. While these variables are formatted character values in the PRPL PUF, once outputted to Excel, Excel will interpret them as numeric since no character is part of the identifier. Excel uses the floating point standard, truncating after the 15th number. It rounds these identifiers (thus losing the complete identifier) and represents them in scientific notation. When importing into Excel, users should make sure to select a text data format within the import wizard for each identifier. Similarly, programmers should incorporate specific text formatting statements when importing and exporting in SAS, Stata, SPSS, and R. Lastly, if copying and pasting identifiers in Excel, a cell must be formatted as "Text" prior to pasting.

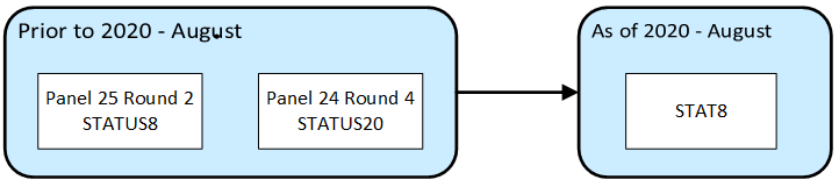
3.2 Person Variables

There are three person-level variables. Binary variables indicate whether the person is the policy holder (PHOLDER) or a dependent (DEPNDNT) on the coverage through the establishment. The variable FYFLG indicates whether the person has a record on HC-224.

There are 14 person-round-level variables. EVALCOVR is a binary variable indicating whether the person was covered by insurance from the establishment at the time of interview (Rounds 6

of Panel 23, Rounds 3 and 4 of Panel 24, and Rounds 1 and 2 of Panel 25) or on December 31 (Round 7 of Panel 23, Round 5 of Panel 24, and Round 3 of Panel 25). Round 6 spans from January 1, 2020 through the Round 6 interview date, which could be anytime between August 2020 and November 2020. If users wish to combine Round 6 coverage data with coverage records from Panel 25 Round 1 or Panel 24 Round 3, the variable EVALCOV5 indicates if coverage was available at the time of the Round 5 interview.

The STAT1-STAT12 variables indicate whether the respondent reported the person was covered by insurance from the establishment for at least one day during the month, January 2020 through December 2020. STAT1-STAT12 variables replace STATUS1-STATUS24. STAT1-STAT12 represent each month from January through December. They 'collapse' the panel-specific STATUS# variables into a single set of month-specific variables. For instance, instead of using STATUS8, STATUS20, and, with the introduction of Panel 23, STATUS32, to determine coverage across the file in August 2020, the single variable STAT8 may be used.



This change is designed to simplify data analysis. These variables continue to represent coverage reported for the interview reference period. For example, if a person from Panel 25 was first interviewed (Round 1) in February and reported she was covered in January and February, and then in the second interview (Round 2) in August she reported she was covered from March through August, then the PRPL record for the first round will have STAT1 and STAT2 set to 1 YES and the rest set to -1 INAPPLICABLE, and the PRPL record for the second round will have STAT3 through STAT8 set to 1 YES and the rest set to -1 INAPPLICABLE .

3.3 Policyholder Variables

The values of three variables describing the policyholder do not vary across the records of the persons covered by the plan, regardless of whether the covered person is the policyholder. The variable DECPHLDR indicates the policyholder is deceased. The variable OUTPHLDR indicates the policyholder resides outside the RU. In each case, there are no person-level records for the policyholder on any of the person-level PUF files, even though the PRPL file has a record for the policyholder as a covered person (that is, a record where PHOLDER=1). Instead, the person identifier portion of the coverage record identifier is set to either '901' indicating a deceased policyholder or '902' indicating a policyholder residing outside the RU.

Where a policyholder does not reside in the RU, and, in the next round, a member of the dwelling unit is selected as the policyholder at OE10, the person identifier of the policyholder will change from '902' to the person identifier of the selected dwelling unit member. As noted above, InsurPrivIDEX does not change across rounds for the same policy and thus may be used to link coverage records across rounds.

OUTPHLDR continues to identify only those policies where the person identifier of the policyholder is set to "902"; OUTPHLDR will be set to 2 NO in cases where a member of the dwelling unit is selected as the policyholder at OE10 in a subsequent round. The variable PHLDRCHNG has been added to the file to indicate if a policyholder changed in the current round from the previous round. It is set as follows:

PHLDRCHNG	Description
-1	No previous-round coverage exists or policyholder did not change from the previous round
1	Policyholder changed from 902 in previous round to non-902 in current round
2	Policyholder changed from non-902 in previous round to 902 in current round
3	Policyholder changed from non-902 in previous round to 901 in current round
4	Policyholder changed from non-902 in previous round to a different non-902 in current round
5	Dependents in different RUs review the same plan and report different persons at OE10 as policyholder

Policyholders that are "non-902" include both deceased policyholders and not deceased in-RU policyholders.

The variable NOPUFLG indicates there is another reason the policyholder does not have a record on a person-level PUF. The purpose of these flags is to explain any difficulty users may have linking policyholder information onto the PRPL file. These variables do not, however, measure mortality or policyholders' leaving the household, which should instead be obtained from the PSTATS variables on the person-level files. (For example, policyholders who die between Round 1 (Panel 25) or 3 (Panel 24) or Round 6 (Panel 23) and the end of 2020 will have records on HC-224, and FYFLG will be set to 1 YES.)

3.4 Establishment Variables

The values of establishment-level variables do not vary across the records of the persons insured through the policyholder-establishment pair.

3.4.1 Employers, Exchanges/Marketplaces, and Other Establishments

The type of establishment providing coverage (TYPEFLAG) is on the record. This variable is the source for types of establishments providing employer-based and non-employer-based, private coverage. In this file, TYPEFLAG includes the answers to HX200, HX300, and HP40. TYPEFLAG reflects the type of establishment when the establishment was first reported, but it is not necessarily updated. For example, users must link to the Jobs file to obtain information on employees who left their job since the interview in which the employer was first reported (see Section 3.6). For employment-based coverage through both an employer and a union (such as insurance through a labor-management committee), for most cases, only the coverage record through the employer establishment is retained. These cases are identifiable through the PROVDINS variable on the Jobs File. In some cases, the union and employer may provide different types of coverage. It is important for researchers to review coverage type variables - HOSPINSX, MSUPINSX, PMEDINS, DENTLINS, VISIONIN, and COBRA - to determine which establishment would be most helpful in answering analytic questions.

Note that when TYPEFLAG has a value of 10, "SPOUSE PREVIOUS EMPLOYER," and the spouse resides in the RU, and the respondent selects the spouse as the policyholder at HP90, then the policyholder's ID (PHLDRIDX) is the spouse's person ID (DUPERSID).

The MEPS HC asks about State Exchanges (also known as Marketplaces), so the PRPL File has three variables and one TYPEFLAG (HX200, HX300, and HP40) value related to State Exchanges. These variables and value pertain to sources of health insurance coverage created as part of the 2010 Affordable Care Act. The exchanges were launched in 2014 to simplify shopping for private health insurance coverage. Note that the terms "marketplace" and "exchange" are interchangeable.

The questions about State Exchanges are asked of respondents in every state. The name of the exchange in the respondent's state is used in the questions, but states are not identified on this file.

The variables are as follows:

Coverage through a State Exchange/Marketplace

STEXCH

In the CAPI instrument, HP50 and OE40 indicate whether insurance was obtained through an exchange/marketplace. HP50 provides this information for insurance obtained by a self-employed person with firm size = 1 reported through that job in the Employment section, as well as all other insurance not reported in the Employment section of CAPI. OE40 is asked in Round 3, Panel 24 Round 5, and Round 7 where coverage is reported as "DIRECTLY FROM AN INSURANCE AGENT," "DIRECTLY FROM INSURANCE COMPANY," or "DIRECTLY FROM AN HMO" at HX200, HX300, or HP40 and no State Exchange was reported through the establishment in the previous rounds. HP50 and OE40 are set to -1 INAPPLICABLE for insurance obtained through a current or former employer, union, school, or unknown source. Applicable values exist only for insurance obtained through other groups, insurance companies, insurance agents, HMOs, State Exchanges, or other private sources. These variables are

automatically set to 1 YES, indicating that the source of coverage was from a State Exchange when State Exchange was selected as the source of insurance at HX200, HX300, or HP40. Otherwise, these variables are the responses to HP50 or OE40.

After the interview, HP50 and OE40 were edited to 2 NO, indicating that the source of coverage is not from a State Exchange, when either of the following was true:

1. The respondent reported Medigap coverage (MSUPINSX = 1) was obtained through the State Exchange, or

2. For persons 65 years and older, the person had Medicare (reported elsewhere in the Health Insurance section of the survey).

Thus, these records are no longer included as State Exchange coverage.

Instead of delivering multiple State Exchange variables, the PRPL File creates one variable, STEXCH, to summarize whether coverage is through a State Exchange, using HP50 and OE40. STEXCH is set to -1 INAPPLICABLE for insurance obtained through a current or former employer, union, school, or unknown source. STEXCH is set to 1 YES, EXCHANGE COVERAGE if any of the three following conditions are met:

1. The respondent said a State Exchange is the source of insurance through a self-insured firm with firm size=1 (HP40).

2. The respondent said a State Exchange is the source of insurance not elsewhere reported (HX200/HX300).

3. The respondents said the insurance was through a group, association, insurance company, insurance agent, HMO, or other private source and the respondent said 1 YES, STATE EXCHANGE COVERAGE to "Is this coverage through STATE EXCHANGE NAME?" either in the interview the insurance was first reported (HP50) or in Round 3/Round 5/Round 7 for insurance reported in an earlier round as non-Exchange coverage (OE40).

All others are set to 2 NO, NOT EXCHANGE COVERAGE.

PLANMETL

There are five levels or "tiers" of coverage available through the Marketplace that identify how the policyholder and the insurance plan will split costs. To assist consumers in selecting a level of coverage, the tiers are named Catastrophic, Bronze, Silver, Gold, and Platinum, reflecting a graduated level of cost to the consumer for deductibles, copayments, coinsurance, and out-of-pocket maximums.

Persons are asked to identify the "metal" plan if 1) State Exchange coverage is indicated at HP40, HX200, HX300, HP50, or OE40; 2) the coverage provides hospitalization and physician benefits; and 3) the person is less than 65 years of age.

During editing, PLANMETL is set to -8 DON'T KNOW if hospitalization/physician benefit coverage is -7 REFUSED or -8 DON'T KNOW.

PREMSUBZ

PREMSUBZ (HX690 and OE200) indicates whether insurance was subsidized based on family income. PREMSUBZ is set to -1 INAPPLICABLE for insurance obtained through a current or former employer, union, school, high risk pool, or unknown source. Applicable values exist only for insurance obtained through other groups, insurance companies, insurance agents, HMOs, State Exchanges, or other private sources. For these sources of insurance, respondents are asked each round whether the insurance is subsidized.

3.4.2 Types of Coverage through the Establishment

The establishments in the PRPL file provide private health insurance covering hospital/physician, Medicare supplemental insurance, dental, vision, or prescription medication insurance. The variable PrivateCat identifies the type of source for hospital and physician or Medicare supplemental insurance.

HOSPINSX and MSUPINSX are edited establishment-policyholder flags for whether the policyholder has physician/hospital and Medigap coverage, respectively, through the establishment. However, even when PrivateCat indicates there is either hospital/physician or Medigap coverage, both HOSPINSX and MSUPINSX may have missing values. Note also that both HOSPINSX and MSUPINSX may be coded 1 YES on the same record. DENTLINS, VISIONIN, and PMEDINS flags indicate the establishment provides coverage for dental care, vision care, and prescription medications, respectively. Below are examples of how to use these variables to identify types of insurance:

Identifying Types of Insurance	**Variable and Values**
Hospital and physician or Medicare supplemental insurance	PrivateCat in (1, 2, 3, 4, 5, 6, 99)
Medicare supplemental insurance	MSUPINSX = 1
Hospital and physician insurance	PrivateCat in (1, 2, 3, 4, 5, 6, 99) & MSUPINSX ne 1
Dental insurance	DENTLINS = 1

The variable COBRA is a flag for whether the respondent reported that the coverage was obtained through the requirements of the COBRA of 1986. This act requires that certain employers allow some former employees to continue their employment-based coverage by paying the employer the premium (U.S. Department of Labor). This flag does not, however, indicate all the coverage through former employers, which can be determined using TYPEFLAG and links to former jobs in the Jobs file.

In the PRPL file, COBRA is set to 1 YES when COBRA coverage is indicated at HP140. COBRA is set to 2 NO when the insurance was not COBRA coverage. COBRA is set to -1 INAPPLICABLE when the coverage was not employment-based, and when the coverage was through a current job. COBRA is set to -15 CANNOT BE COMPUTED for retirement jobs first reported in the Employment section in Round 1 (EM380 or EM390), retirement jobs first reported in the Employment section for new RU members (EM380 or EM390), and insurance through unions reported in the Insurance section (HX200/HX300). In a few cases, self-employed persons with firm size = 1 reported buying coverage through a previous job, and these cases are coded as yes or no, while other insurance through self-employment with firm size = 1 is coded -1 INAPPLICABLE.

The variable COVTYPIN flags whether coverage was single or family, based on the number of persons covered in the RU, whether the establishment's insurance covers someone outside the household, and whether the policyholder is outside the household. For Panel 25 Rounds 1 and 2, Panel 24 Rounds 3 and 4, and Panel 23 Round 6, the number of covered persons was measured at the time of the interview (or end of the reference period). For Panel 25 Round 3, Panel 24 Round 5, and Panel 23 Round 7, the number is as of December 31st. When coverage ceased before the end of the reference period for every co-residing family member, COVTYPIN is set based on the number of persons ever covered during the round.

The variable COVROUT_M18 indicates whether out-of-RU persons were covered by the plan. COVROUT_M18 is asked if there are no dependents living in the RU covered under that policy. It is not asked if there is only one member of an RU and that person is covered by a policyholder who is deceased, no longer in the household, or not listed in the dwelling unit.

3.4.3 Out-of-Pocket Premiums

In the MEPS, questions on out-of-pocket premiums were asked of all policyholders with private insurance coverage for all establishments. The variable OOPPREM provides the monthly out-of-pocket premium paid by the policyholder for coverage through the establishment for Panel 25 as of Round 1 and Panel 24 as of Round 3.

Because Round 5 was administered for Panel 23 respondents as if it were their last MEPS interview, OOPPREM and related out-of-pocket premium variables are not available for Panel 23 respondents in the 2020 PRPL file. Out-of-pocket premium questions are now asked in select rounds of extended panels (Panel 23 Round 7), as well as Panel 24 Round 5 that now extends into the next calendar year, 2021. Premium amounts reported in Panel 23 Round 7 and Panel 24 Round 5 will be included in the 2021 PRPL file.

OOPELIG flags these covered-person-policyholder-establishment triples. OOPPREMX provides an edited version of OOPPREM and the variable OOPFLAG identifies which records were edited. OOPX12X is provided as a convenience to researchers and contains the edited monthly out-of-pocket premium amount multiplied by 12, representing the annual amount.

The edited variable OOPPREMX includes imputed values for records which contained missing values as well as for a limited number of records with values that were implausibly low or high. In most years, for policyholders in Round 3 of the second panel with missing out-of-pocket

premiums, if coverage is through a continuation job which was originally reported in Round 1 of the first year of the panel and type of coverage (COVTYPIN) is the same as in Round 1, then OOPPREMX is set equal to OOPPREMX from Round 1 times the growth rate in out-of-pocket premiums from the first delivery year to the next. The growth rate is assigned by type of coverage and is based on private sector out-of-pocket premiums reported in the MEPS Insurance Component in the current and prior year. Note that in 2019, however, the inflation method was not used. These records could enter other imputation processes.

Imputed values were typically assigned to these and other records by one of several imputation methods - hot-deck imputation or MEAN substitution, both of which consider the following person/plan characteristics: source of insurance (private employer, state and local government, federal government, Medigap, other non-group policy, State Exchanges), age of policyholder, educational attainment of policyholder, number of persons covered by the policy, if there is a high family deductible, size of employer, region and MSA, presence of supplemental benefits such as drug, dental and vision, whether the insurance was through a current or former job, and active or retired job. For employer-sponsored insurance where a link to a job is established in the PRPL file, a select group of the edited variables in HC-211 or HC-218 is used to define imputation classes for hotdeck imputation of out-of-pocket premiums.

Missing premium amounts on coverage purchased through a State Exchange continue to be hot-deck imputed in a separate process.

Both OOPPREM and OOPPREMX are coded as zero for group policyholders who reported paying none of their premium.

OOPPREM is created using the out-of-pocket amount reported and the frequency of payments (HX670, HX680, and HX680OS):

> HX670
>
> How much {{do/does}/did} {you/{POLICYHOLDER}} pay for the {INSURANCE SOURCE NAME}coverage?
>
> ENTER AMOUNT
>
> HX680/HX680OS
>
> {Is/Was} that per year, per month, per week, or what?
>
> UNIT OF COVERAGE:
>
> SPECIFY: UNIT OF COVERAGE

PREMLEVX shows whether the amount reflected in OOPPREM was the full premium or part of it. The PREMLEVL (HX660 or OE170) question is asked in all rounds for insurance obtained through other groups, insurance companies, insurance agents, HMOs, State Exchanges, or other private sources, but not insurance obtained through a current or former employer, union, school, or unknown source. Note that the premium amount is not collected in Rounds 2, 4, and 6.

For the entire set of 6 variables (OOPPREM, OOPPREMX, OOPX12X, OOPELIG, OOPFLAG, PREMLEVX), the same values are reported on the records of each dependent person covered through the policyholder's establishment, but the policyholder paid only once per establishment-policyholder.

3.5 Plan Variables

The values of plan-level variables do not vary across the records of the persons insured through the policyholder-establishment pair.

The variables ANNDEDCT (HX700/OE210) and HSAACCT (HX710/OE220) capture whether a private health plan has a high deductible and whether that plan is associated with a Health Savings Account (HSA) or a similar special fund/account. These questions are asked during the first report and during the review of insurance in Rounds 1 and 3 for all private plans except for individuals covered by disability, Workers' Compensation, accident insurance, or any combination thereof, and/or individuals covered by Medicare supplement/Medigap plans. These questions are now asked in select rounds of extended panels (Panel 23 Round 7), as well as Panel 24 Round 5 that now extends into the next calendar year, 2021. They are asked whether or not coverage status is known. Users should note ANNDEDCT is formatted as a numeric variable as of 2019.

3.5.1 Household Reports of HMOs

The variable UPRHMO identifies records for HMO coverage when the household respondent reported that the insurance was purchased through an HMO, reported the insurance company was an HMO, or described the plan as an HMO. In all cases the respondent answered a question using the term "HMO." The HMO question is asked of all private coverage records, even if the person also reports Medicaid or another government-sponsored plan with hospital/physician coverage in the round. UPRHMO is set to 1 YES if either of the following conditions are met:

1. If the respondent reported insurance purchased directly through an HMO (HX200, HX300, HP40)

2. If the respondent answered 1 YES to the following question (MC10):

   {Is/Was} {your/{POLICYHOLDER}'s } {NAME OF INSURER} an HMO {as of (END DATE)}? {When answering this question, do not consider {your/his/her} insurance through Medicare.}

   [With an HMO, you must generally receive care from HMO physicians. For other doctors, the expense is not covered unless you were referred by the HMO or there was a medical emergency.]

UPRHMO is set to 2 NO when the plan was not an HMO. UPRHMO is set to -1 INAPPLICABLE when the plan was not hospital/physician or Medicare supplemental coverage.

3.5.2 Change in Plan Name

The variable NAMECHNG indicates whether the name of the plan obtained through the establishment changed from the prior round. For Panel 25 Rounds 2 and 3 and Panel 24 Rounds 3, 4 and 5, and Panel 23 Rounds 6 and 7, NAMECHNG is set to 1 YES if four conditions were met: 1) someone in the RU had coverage through the establishment in the prior round; 2) either still had coverage at the time of the interview or the coverage status was unknown or refused; 3) a new dependent was added to the coverage and 4) the respondent answered 1 YES to the following question (OE110):

{Last time we recorded that {you/{POLICYHOLDER}} (were/was) covered by {PRELOAD.INSURANCE.INSURER}.}

{Since (START DATE), has there been/Between {START DATE} and {END DATE}, was there} any change in the plan name of the health insurance {you/{POLICYHOLDER}} {{have/has}/had}through {PRELOAD.INSURANCE.HISRCNAME}?

If the respondent answered no, then NAMECHNG is coded 2 NO. If no one in the RU had coverage through the establishment in the prior round, no one had coverage at the time of the interview, or it is a Round 1 record, then NAMECHNG is set to -1 INAPPLICABLE.

When the respondent answered 1 YES, then MEPS HC asked about types of benefits (OE130), which are updated on the PRPL file.

There are two important caveats to this variable. First, changes in plan name do not necessarily imply the plan itself changed. For example, the plan may have merely changed its name for marketing purposes. Second, the variable NAMECHNG pertains only to changes in plan names at the same establishment; a policyholder may switch plans if she or he switches the establishment (including employer) through which she or he obtains insurance. Switches in EPRSIDs and ESTBIDs between rounds indicate those other types of changes.

3.6 Links to Jobs Providing Insurance

For employment-based insurance, there are two variables linking the insurance to details about the jobs through which the insurance was obtained, CMJINS and JOBSIDX.

Most people with employment-based insurance have it through current main jobs. The variable CMJINS indicates whether the insurance is through a current main job. When insurance is through a previous job or through self-employment and there is only one employee, then CMJINS is set to 2 NO. When the insurance is not employment-based, CMJINS is set to -1 INAPPLICABLE. If CMJINS = 1 and the policyholder has a PUF record (FYFLG = 1), then many edited and imputed current main jobs variables are available on the 2020 Consolidated PUF (HC-224). If a policyholder is non-responding in 2020 but continues to cover a responding dependent in 2020, refer to the 2019 Consolidated PUF (HC-216). For a small portion of records, users may need to refer to the 2018 Consolidated PUF (HC-209).

For these and other types of jobs (for example, former jobs) with employer-sponsored health insurance, the Jobs files (2018 Jobs PUF (HC-203), 2019 Jobs PUF (HC-211), and 2020 Jobs PUF (HC-218)) contain edited variables describing the job. JOBSFILE indicates which Jobs file contains the most current information about the source of coverage. In most cases, information about the job is in HC-218, but for Panel 23 and Panel 24, if the job ended before 2020, information about the job is contained in HC-211. A small portion of Panel 23 records link to 2018 (HC-203) if the job ended before 2019. JOBSIDX allows users to link to the job that is the source of coverage in HC-203, HC-211, and HC-218.

Links between reported jobs and sources of coverage may be obtained directly from the respondent or inferred within the PRPL process. The variable JOBSINFR indicates if a link was directly reported or inferred. PRPL creates direct links using the link established in CAPI between a job and coverage reported in the Employment section. Prior to 2019 PRPL, linking in PRPL required a series of variables to define the link and the link was not always one-to-one. Because of this improved method of direct linking in CAPI and unique numbering of insurance plans (3-byte insurance number included in EPRSIDX), there are now fewer cases with inferred links and fewer cases of duplicative reports of coverage. As such, researchers may note differences in the distribution of some variables compared with previous years.

A link is inferred only when persons report employment-based health insurance at the end of the Insurance section (HX200 or HX300) or based on whether the insurance was through a current or former job (EMPLSTAT). Neither of these types of coverage has a direct link to a job reported within CAPI. An inferred link is established where a policyholder is employed at a job where insurance was not reported through the job in the Employment section of CAPI. Most inferred links are assigned where the employer and the insurance are to the same establishment.

For 2020, there are 96 Panel 23 Round 6 records that link to a Round 5 job instead of a Round 6 job because the Round 6 record was deleted during special 2020 processing. As collected in the interview, Panel 23 Round 5 respondents were asked to report employment information through 12/31/2019. However, Round 5 records for the Full Year 2020 Jobs PUF (HC-218) were changed to represent employment for the time period between 1/1/2020 and the Round 5 interview date. Round 6 records were changed to reflect the time period between the Round 5 interview date and the Round 6 interview date. (Refer to 2020 Jobs PUF (HC-218) Appendix 5 for a full description of the adjustment process and its impact on PUFs.) For this reason, some Round 5 records for former jobs actually originated in the Round 6 interview and the corresponding Round 6 job records were removed. For these records, no Round 6 job record exists for linking Round 6 insurance records in the Full Year 2020 PRPL PUF (HC-223). Therefore, PRPL established a link with the corresponding Round 5 former job.

The variable EMPLSTAT contains the answer to question HP120, which is asked only about the policyholders of employment-related insurance first mentioned at the end of the Insurance section of the interview (HX200/HX300), and HP120 is asked only in the interview round where the insurance was first reported. Thus, it is useful only for the cases where links to jobs could not be inferred. EMPLSTAT does not contain updated information about the policyholder's employment at each interview. However, EMPLSTAT is set on reviewed coverage to the value from the round where coverage was first reported.

4.0 Linking to Other Files

4.1 National Health Interview Survey

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

4.2 Longitudinal Analysis

Panel-specific longitudinal files are available for downloading in the data section of the MEPS website. As has been done routinely in past years, the longitudinal file for Panel 24 comprises MEPS survey data obtained in Rounds 1 through 5 of the panel and can be used to analyze changes over a two-year period. Unlike past years for MEPS, in 2020 Panel 23 had data collected for a third year. As such, two-year and three-year longitudinal files will be developed for Panel 23. These can be used to analyze changes over the corresponding two-year or three-year period. Variables in the file pertaining to survey administration, demographics, employment, health status, disability days, quality of care, patient satisfaction, health insurance, and medical care use and expenditures were obtained from the MEPS full-year Consolidated files from the years covered by each panel.

For more details or to download the data files, please see Longitudinal Weight Files on the AHRQ website.

5.0 Using MEPS Data for Trend Analysis

There are uncertainties associated with 2020 data quality for reasons discussed in the Survey Sample Information section of the Consolidated PUF document (HC-224). Evaluations described in that section suggest that care should be taken in the interpretation of estimates based on data collected in 2020 as well as in comparisons over time. Trend analyses are challenging since the advent of the COVID-19 pandemic resulted in uncertain data quality for MEPS as well as standard benchmark sources such as the CPS, ACS, and NHIS while the pandemic also had an impact on the health and access to health care of the U.S. population. For such reasons, the extent to which 2020 health care parameters may differ from those of prior years is difficult to assess.

In terms of other factors to be aware of, 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. The length of time being analyzed should be considered. In particular, large shifts in survey estimates over short periods of time (e.g. from one year to the next) that are statistically significant should be interpreted with caution unless they are attributable to known factors such as changes in public policy, economic conditions, or MEPS survey methodology.

With respect to methodological considerations, in 2013 MEPS introduced an effort focused on field procedure changes such as interviewer training to obtain more complete information about health care utilization from MEPS respondents with full implementation in 2014. This effort likely resulted in improved data quality and a reduction in underreporting starting in the second half of 2013 and throughout the 2014 full year files and have had some impact on analyses involving trends in utilization across years.

The aforementioned changes in the NHIS sample design in 2016 could also potentially affect trend analyses. The new NHIS sample design is based on more up-to-date information related to the distribution of housing units across the U.S. As a result, it can be expected to better cover the full U.S. civilian, noninstitutionalized population, the target population for MEPS, as well as many of its subpopulations. Better coverage of the target population helps to reduce potential bias in both NHIS and MEPS estimates.

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

Process changes, such as data editing and imputation, may also affect trend analyses. For example, users should refer to Section 2.5.11 in the 2020 Consolidated file (HC-224) and, for more detail, the documentation for the prescription drug file (HC-220A) when analyzing prescription drug spending over time.

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

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

Finally, statistical significance tests should be conducted to assess the likelihood that observed trends are not attributable to sampling variation. In addition, researchers should be aware of the impact of multiple comparisons on Type I error. Without making appropriate allowance for

multiple comparisons, undertaking numerous statistical significance tests of trends increases the likelihood of concluding that a change has taken place when one has not.

References

U.S. Department of Labor. Employee Benefits Security Administration. 2019. An Employee's Guide to Health Benefits under COBRA, the Consolidated Omnibus Budget Reconciliation Act of 1986. Washington, DC.

D. Variable-Source Crosswalk

FOR MEPS PUBLIC USE FILE HC-223: 2020 Person Round Plan

Health Insurance Variables

Variable	Description	Source
EPCPIDX	Unique record identifier (DUPERSID + EPRSIDX)	Constructed
DUPERSID	Person identifier of policyholder or dependent covered by policyholder (DUID+PID)	Constructed
PHLDRIDX	Person identifier of the policyholder	Constructed
ESTBIDX	Insurance source identifier	Constructed
EPRSIDX	Uniquely identifies insurance policy from an insurance source in a round (ESTBIDX+PHLDRIDX+RN+3-byte InsuranceID)	Constructed
InsurPrivIDEX	Uniquely identifies insurance policy from an insurance source (ESTBIDX+ 3-byte InsuranceID)	Constructed
PANEL	Panel number	Constructed
RN	Round number	Constructed
JOBSIDX	Link to Jobs file that uniquely identifies policyholder's job by round at the establishment that provided insurance	Constructed
JOBSINFR	JOBSIDX inferred rather than reported	Constructed
JOBSFILE	Public use Jobs file number indicating source of jobs information	Constructed
FYFLG	Person in full year public use file	Constructed
CMJINS	Current main job is the source of plan	PrivateCat, RJ20, EM50, EM80
EMPLSTAT	Policyholder's employment status at employer insurance coverage not reported in Employment section	HP120
PHOLDER	Policy holder record flag	HP70, 90
DEPNDNT	Dependent of policy holder record flag	PrivateCat, PHOLDER

Variable	Description	Source
PHLDRCHNG	Flag indicates policyholder identifier changed from prior round because policyholder left/returned to DU	OE10 Constructed
EVALCOVR	Covered at interview or on December 31st	HQ10_01,10_02
EVALCOV5	Covered at Round 5 interview	HQ10_01,10_02
STAT1-STAT12	Insurance active in month Where month = January-December	HQ10_01, 10_02, 10_03,10_04, 10_05
DECPHLDR	Deceased policyholder flag	Constructed
OUTPHLDR	Policyholder is not in RU	Constructed
NOPUFLG	Policyholder not in full year file	Constructed
COVROUT_M18	For policies without dependents covered in RU, identifies policy covers person not in RU	Constructed HP170, 180; OE100
TYPEFLAG	Type of insurance source	HP40; HX50_01, 60, 100_01, 120_01, 140, 170, 200, 225, 230, 270_01, 300_01, 320_01; EM50, 80, 100, 120, 180, 240, 340, 390, 720
STEXCH	Coverage obtained through State exchange	HP50; HX200, 300; OE40
PrivateCat	Category of private coverage (beginning in 2018)	Constructed HP40, 50, 130; HX200, 300, 620; OE40; EM710
HOSPINSX	Type of health insurance coverage received through plan: hospitalization & physician/HMO (edited)	HX620; OE130
MSUPINSX	Type of health insurance coverage received through plan: Medigap	HX620; OE130
DENTLINS	Type of health insurance coverage received through plan: dental	HX620; OE130
VISIONIN	Type of health insurance coverage received through plan: vision	HX620; OE130
PMEDINS	Type of health insurance coverage received through plan: prescription drug	HX620; OE130
COBRA	COBRA Coverage	HP40,120, 140; HX200, 300; OE70, 90; EM50, 80, 100, 270, 390; RJ20; PrivateCat

Variable	Description	Source
PLANMETL	Plan metal level	HX650, OE160
COVTYPIN	Coverage is single or family, based on number of persons within the RU ever covered in the round	HP160, 170, 180; OE100
OOPELIG	Flag indicates that the policyholder-insurance source has premium in the PRPL file	RN; TYPEFLAG; HP40, 160; HX200, 300
OOPPREM	Monthly out-of-pocket premium amount paid by policyholder	HX660, 670; OE170, 180
OOPPREMX	Monthly out-of-pocket premium amount paid by policyholder (edited/imputed)	Constructed
OOPX12X	Annual out-of-pocket premium amount paid by policyholder (edited/imputed)	Constructed
OOPFLAG	Flag indicates if premium was edited/imputed	Constructed
PREMLEVX	Portion of premium paid by family (edited)	HX660, 670; OE170, 180
PREMSUBZ	Cost of premium is subsidized based on family income	HX690; OE200
ANNDEDCT	Plan has high, low, or no deductible	HX700; OE210
HSAACCT	Plan is associated with Health Savings Account or similar special fund/account	HX710; OE220
UPRHMO	Coverage purchased through HMO, insurance company was an HMO, or insurance name indicates HMO	HP40; HX200, 300; MC10
NAMECHNG	Change in plan name from prior round	OE110