MEPS HC-227
2021 Jobs File

March 2023

Table of Contents

Appendices		Page

A. Data Use Agreement

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

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

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

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

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

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

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

B.   Background

1.0   Household Component

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

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

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

2.0   Medical Provider Component

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

3.0   Survey Management and Data Collection

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

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

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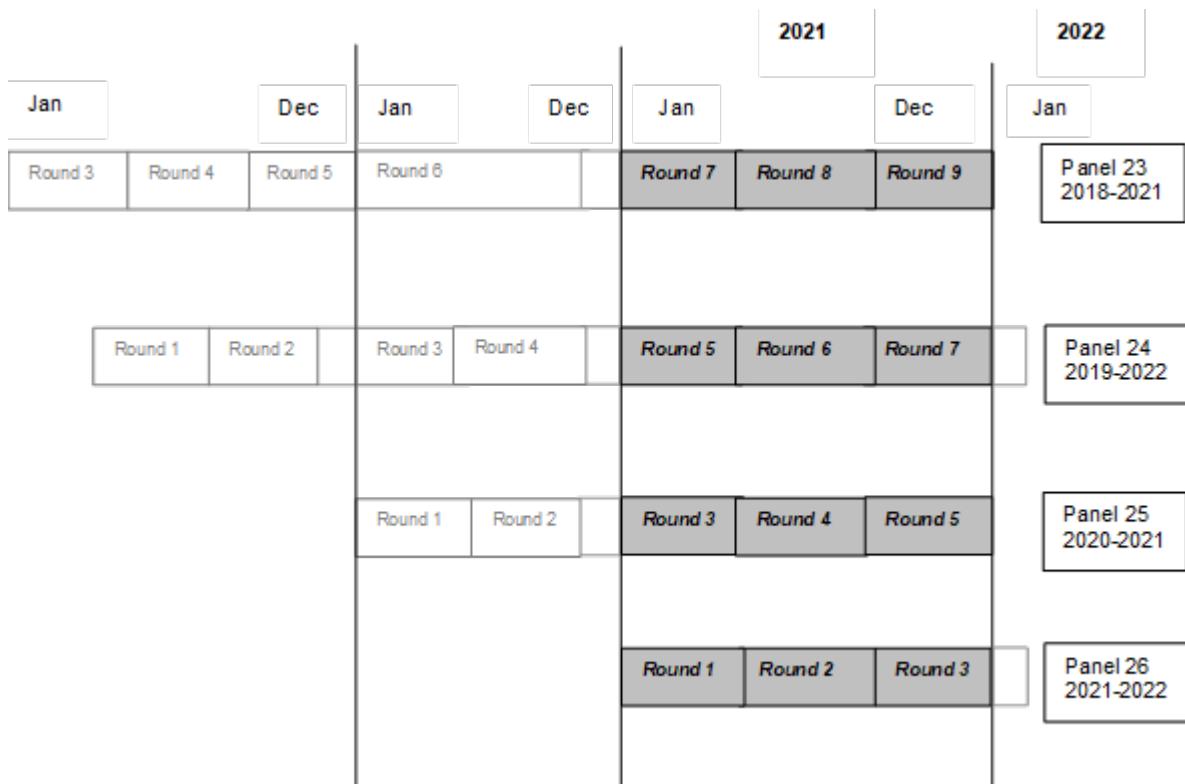
C. Technical and Programming Information

Section C of this document offers a brief overview of the data provided in MEPS public use release HC-227, as well as the content and structure of the codebook, reserved code values, and variable naming conventions. It is followed by Section D containing the Variable-Source Crosswalk, Appendix 1 containing sample SAS program code, and Appendix 2 containing sample Stata program code. A copy of the survey instrument used to collect the information on this file is available on the [MEPS website](#).

1.0 General Information

This file is being released as a research file and has undergone the standard quality control procedures usually performed on MEPS data files. The file includes a total of 48,353 records, with each record representing a unique job for a person by round. This file presents information about jobs starting on or before 12/31/2021 only. The 2022 Jobs file release will present information on Panel 24 jobs and Panel 26 jobs starting in 2022.

In the Employment section, MEPS collects complete job-related information in the round in which a job is first reported. While the details collected vary by job type (see Section 2.0), the data reported for a job in its first survey round may include earnings by type (gross salary, tips, etc.), start and stop dates, hours and weeks worked, establishment size, industry and occupation codes, presence of retirement and other benefits, self-employment versus other status, temporary or seasonal situations, and health insurance availability. Minimal data updates are collected for later rounds in which the job continues.

This Full-Year Jobs file contains job records from four MEPS panels. The 2021 Jobs file provided in this release, MEPS HC-227, contains job-level information collected in Rounds 7 through 9 for Panel 23, Rounds 5 through 7 for Panel 24, Rounds 3 through 5 for Panel 25, and Rounds 1 through 3 for Panel 26 of the Medical Expenditure Panel Survey.

				2021			2022	
Jan		Dec	Jan	Dec	Jan	Dec	Jan	
Round 3	Round 4	Round 5	Round 6		Round 7	Round 8	Round 9	Panel 23 2018-2021
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7	Panel 24 2019-2022
		Round 1	Round 2		Round 3	Round 4	Round 5	Panel 25 2020-2021
					Round 1	Round 2	Round 3	Panel 26 2021-2022

Due to the impact of the COVID-19 pandemic on MEPS collection methods and lower response rates in the 2020 and 2021 calendar years, AHRQ extended fielding in 2021 for Panel 23 and Panel 24 persons to include two additional rounds. Panel 23 Round 8 and Round 9 and Panel 24 Round 6 and Round 7 collected information about 2021. As a result, the 2021 MEPS includes four panels of data: Panel 26 Rounds 1, 2, and 3; Panel 25 Rounds 3, 4, and 5; Panel 24 Rounds 5, 6, and 7; and Panel 23 Rounds 7, 8, and 9. For 2021 data collection, Panel 23 Round 9 and Panel 25 Round 5 are treated as terminal rounds, referring back to the period between the Round 8 interview and December 31, 2021 for Panel 23, and to the period between the Round 4 interview and December 31, 2021 for Panel 25. Alternatively, Round 5 from Panel 24 and Round 7 from Panels 23 and 24 were collected as cross-year rounds, covering the entire period between the current interview date and the prior interview date (regardless of calendar year and no truncation at December 31). These cross-year measures are collected/coded the same as Round 3 cross-year variables from all panels.

In order to obtain complete information for a job, users must note the round in which the job is first reported. This is because MEPS collects complete Jobs information in that round only, as noted above.

For the first year panel, jobs from Panel 26 Rounds 1, 2, and 3 are included in the 2021 Jobs file. Complete information for any Panel 26 job is available for jobs that started before 1/1/2022, whether that job was first reported in Round 1, 2, or 3. This is the case for any first year panel (the panel that began its first year of interviewing in the given year) in a Full-Year Jobs file.

For the second year panel (the panel that continued with its second year of interviewing in the given year), jobs from Panel 25 Rounds 3, 4, and 5 are included in the 2021 Jobs file. If the

Round 3, 4, or 5 job continued from Round 1 or Round 2, users must look back to the Jobs file from the previous year (2020) to obtain complete information for the job.

For the third year panel (the panel that continued with its third year of interviewing in the given year), jobs from Panel 24 Rounds 5, 6, and 7 are included in the 2021 Jobs file. If the Round 5, 6, or 7 job continued from Round 3 or Round 4, users must look back to the Jobs file from the previous year (2020) to obtain complete information for the job. If the Round 5, 6, or 7 job continued from Round 1 or Round 2, users must also look back to the Jobs file from two years prior (2019) to obtain complete information for the job.

For the fourth year panel (the panel that continued with its fourth year of interviewing in the given year), jobs from Panel 23 Rounds 7, 8, and 9 are included in the 2021 Jobs file. If the Round 7, 8, or 9 job continued from Round 5 or Round 6, users must look back to the Jobs file from the previous year (2020) to obtain complete information for the job. If the Round 7, 8, or 9 job continued from Round 3 or Round 4, users must also look back to the Jobs file from two years prior (2019) to obtain complete information for the job. If the Round 7, 8, or 9 job continued from Round 1 or Round 2, users must also look back to the Jobs file from three years prior (2018) to obtain complete information for the job.

Appendix 1 includes sample SAS code and Appendix 2 contains sample Stata code to assist users in obtaining this information. Users should note that, because of differences in sample composition between the current year and the previous year files (i.e., a person was included in the previous year's delivery but not the current year or vice versa), or because more accurate information was received in subsequent round comments following the delivery of the Jobs records in the previous year, there occasionally may not be a corresponding job in the previous year file.

2.0   Data File Information

2.1   File Contents

Each record in the 2021 Jobs file represents one job reported by a person in a round. All persons age 16 and older in the MEPS are asked to report on jobs held. Depending on an individual's job history, these reported jobs may be held:

- at the interview date,

- in the round but prior to the interview date, or

- prior to the round.

Only those persons reporting a job in a round will have a record in the 2021 Jobs file.

Record Identifiers

The unique record identifier is the variable JOBSIDX, which is composed of a person identifier (DUID + PID), a round identifier (RN), and a job number (JOBNUM). The similarly named variable JOBIDX (without "S") has the same structure as JOBSIDX but without the round identifier (RN). JOBIDX allows users to easily select all rounds of the same job for the same person. A panel indicator (PANEL) is included on the file to distinguish (a) Round 7 jobs held by Panel 23 persons from Round 7 jobs held by those in Panel 24, (b) Round 5 jobs held by Panel 24 persons from Round 5 jobs held by those in Panel 25, and (c) Round 3 jobs held by Panel 25 from Round 3 jobs held by those in Panel 26. The DUID identifier in this data release includes a 2-digit code to identify the panel and, as a result, JOBSIDX includes a panel identifier via DUID. The variable OrigRnd indicates the round in which a job was first created. Therefore, it may or may not contain the same value as RN.

In 2021, an establishment identifier, ESTBIDX, is added to the file. ESTBIDX is comprised of DUID + "an establishment number" and can help data users to (a) determine potential duplication of job records (i.e., a person reports multiple jobs to the same establishment in the same round with many or all of the same characteristics), and (b) better understand job changes, since job holders may leave an establishment and return to the same establishment in any round.

Each identifier variable (JOBSIDX, JOBIDX, ESTBIDX, DUID, DUPERSID) begins with the 2-digit panel number. This allows analysts to easily identify records delivered in a previous year Jobs file (when panel is used in conjunction with other variables, such as RN and OrigRnd). In addition, CAPI assigns a unique job number that *may not be used in subsequent rounds* on different jobs. This 3-byte number, JOBNUM, is unique to the *reporting unit* (RU) and is set to a value that corresponds with the RU in which a person's job was first reported (e.g. A RU is '1', B RU is '2', C RU is '3', etc).

Initial Reporting Round

Most persons held only one job at the first interview date - their "Current Main Job." For persons who held more than one job at the round's interview date (a current job), respondents were asked to identify the main job. This job was classified as the "Current Main Job" and any other simultaneously held job was classified as a "Current Miscellaneous Job." The MEPS also obtained some information on any former jobs (Former Main Job or Former Miscellaneous Job) held in the reference period but not at the interview date. For those persons neither working at the interview date nor earlier in the reference period, limited information on the last job the person held was collected. Additionally, for those persons age 55 or older who were identified as having retired from a job, the MEPS obtained some job-level information (Retirement Job).

The variable SUBTYPE indicates the type of job record - current main (1), current miscellaneous (2), former main (3), former miscellaneous (4), last job outside reference period (5), or retirement job (6). When a job is initially reported, MEPS asks for detailed information about any "Current Main Job" and basic information about other job types. Refer to the questionnaire to see which information was asked for each job type. The following variable list identifies when a variable could be set based on the job SUBTYPE. Self-employed and wage-earner status at a

job also defines when a variable may be set. (Note: wage-earner is used to describe workers who are not self-employed.) The last column indicates if the variable is populated in the round in which the job is first reported (collection only), when the job is reviewed (review only), or both (collection and review).

Variables Set for Each SUBTYPE

Variable	Self-Employed Jobs	Wage Earner Jobs	Current Main	Current Miscellaneous	Former Main	Former Miscellaneous	Last Job Outside Reference Period	Retirement	When Populated
JOBTYPE	x	x	x	x	x	x	x	x	collection only
JSTRTM	x	x	x	x	x	x			collection only
JSTRTY	x	x	x	x	x	x			collection only
JSTOPM	x	x			x	x	x	x	collection and review
JSTOPY	x	x			x	x	x	x	collection and review
RETIRJOB	x	x						x	collection and review
SUBTYPE	x	x	x	x	x	x	x	x	collection and review
JOBHASHI	x	x		x	x	x	x	x	collection only
NUMEMPS		x	x		x				collection only
ESTMATE1_M19		x	x		x				collection only
MORELOC		x	x		x				collection only
BUSINC	x		x		x				collection only
PROPRIET	x		x		x				collection only
TYPEEMPL		x	x		x	x if not self-employed & retired	x	x	collection only
YLEFT_M18		x			x		x		collection only
YNOBUSN_M18	x				x		x		collection only

Variable	Self-Employed Jobs	Wage Earner Jobs	Current Main	Current Miscellaneous	Former Main	Former Miscellaneous	Last Job Outside Reference Period	Retirement	When Populated
HRSPRWK	x	x	x	x	x				collection only
HRS35WK	x	x	x		x				collection only
SICKPAY		x	x		x				collection only
PAYDRVST		x	x		x				collection only
PAYVACTN		x	x		x				collection only
RETIRPLN		x	x		x				collection only
SESNLJOB	x	x	x	x	x				collection only
TEMPJOB	x	x	x	x	x				collection only
WKLYAMT	x	x		x					collection only
EMPLINS	x	x	x						collection only
OFFRDINS	x	x	x	x	x	x	x	x	collection only
DIFFPLNS	x	x	x	x	x	x	x	x	collection only
ANYINS	x	x	x	x	x	x	x	x	collection only
INUNION	x	x	x	x	x	x	x	x	collection only
PROVDINS	x	x	x	x	x	x	x	x	collection only
EmplUnionProv	x	x	x	x	x	x	x	x	collection only
HHMEMBER_M18	x		x	x	x	x	x	x	collection only
TOTLEMP_M18	x		x	x	x	x	x	x	collection and review
TotNumEmp	x		x	x	x	x	x	x	collection only
RvwTotNumEmp	x		x	x					review only
SALARIED		x	x		x				collection and review

Variable	Self-Employed Jobs	Wage Earner Jobs	Current Main	Current Miscellaneous	Former Main	Former Miscellaneous	Last Job Outside Reference Period	Retirement	When Populated
HOWPAID		x	x		x				collection and review
DAYWAGE		x	x		x				collection and review
HRSPRDY		x	x		x				collection and review
MAKEAMT		x	x		x				collection and review
PERUNIT_M18		x	x		x				collection and review
MORE10		x	x		x				collection and review
MORE15		x	x		x				collection and review
MOREMINM		x	x		x				collection and review
GROSSPAY		x	x		x				collection and review
GROSSPER		x	x		x				collection and review
SALRYWKS		x	x		x				collection and review
HRSALBAS		x	x		x				collection and review
EARNTIPS		x	x		x				collection and review
EARNBONS		x	x		x				collection and review
EARNCOMM		x	x		x				collection and review
TIPSUNIT_M18		x	x		x				collection and review
TIPSAMT		x	x		x				collection and review
BONSUNIT		x	x		x				collection and review
BONSAMT		x	x		x				collection and review
COMMUNIT		x	x		x				collection and review
COMMAMT		x	x		x				collection and review

Variable	Self-Employed Jobs	Wage Earner Jobs	Current Main	Current Miscellaneous	Former Main	Former Miscellaneous	Last Job Outside Reference Period	Retirement	When Populated
HRLYWAGE		x	x		x				collection and review
STILLAT	x	x	x						review only
MAIN_JOB	x	x	x						review only
DIFFWAGE		x	x						review only
StillWorkFTPT	x	x	x						review only
WhyChngPTToFT	x	x	x						review only
WhyChngFTToPT	x	x	x						review only
STILLWRK	x	x		x					review only
OFFTAKEI	x	x	x	x					review only
NOWTAKEI	x	x	x	x					review only
NOWTAKEI_M22	x	x	x	x					review only
ESTBTHRU	x	x	x	x					review only
INSESTB	x	x	x	x					review only
WHY_LEFT_M18	x	x			x	x			review only

For last jobs outside of reference period and retirement jobs that ended more than two years prior to the beginning of the reference period, certain questions (HHMEMBER_M18 and TOTLEMP_M18) are not asked. The precise calculation of the two-year cut-off date is not possible for some persons due to allowed negative values on stop year, stop month, and reference period start month. Therefore, HHMEMBER_M18 and TOTLEMP_M18 may be collected for some jobs that ended more than two years prior to the reference period.

Users will note the inclusion of two new variables - EmplUnionProv and NOWTAKEI_M22. They are described in more detail in subsequent sections.

Skip Patterns

Due to many skip patterns, it is recommended that users of the 2021 Jobs file become familiar with the Employment section in the MEPS questionnaire. To aid users, a crosswalk between variables and MEPS questionnaire numbers is provided in this release. The following examples of variables involved in skip patterns are presented to be illustrative; these examples do not represent the full range of variables affected by questionnaire skip patterns.

In one example of a skip pattern, the MEPS does not obtain job-related benefits such as vacation, sick leave, and pension information for self-employed jobs, so those variables are coded as "Inapplicable" (-1) for those types of jobs. Nor does the MEPS attempt to obtain wage, salary, and information regarding whether the job was in the private sector, federal or local government (TYPEEMPL) for the self-employed. So again, due to the skip pattern, TYPEEMPL is coded as "Inapplicable" (-1) for self-employed jobs.

Conversely, the questions relating to business organization type (BUSINC, PROPRIET) are asked only of the self-employed, so the skip pattern results in those variables being coded as "Inapplicable" (-1) for jobs performed by wage earners.

Job Updates and "Inapplicable" (-1) Values

The MEPS used dependent interviewing in Rounds 7, 8, and 9 for Panel 23, Rounds 5, 6, and 7 for Panel 24, Rounds 3, 4, and 5 for Panel 25, and in Rounds 1, 2, and 3 for Panel 26 (see Section RJ in the Employment section of the questionnaire). In these rounds, the MEPS asked about current main and current miscellaneous jobs held at the previous round interview date to determine whether the jobholder continued to work at these jobs. For other job types (former, last, or retirement) reported in the previous round, MEPS does not ask any follow-up questions. These jobs, by definition, are no longer held by the person and therefore are not included on the file except in the round they are first reported.

With dependent interviewing, if a person still held a Current Main Job from the previous round, the MEPS asked whether the job was still the main job. For most jobholders, it was reported that they still worked at the same job and it was still their main job. If, in a subsequent interview, a job was no longer held, it was designated as a former job for that follow-up round. It is also possible, although unusual, for a job to change from main to miscellaneous (or vice versa) in a round subsequent to the initial report.

If job status remained the same for a continuing job (either main or miscellaneous), the MEPS asked only a subset of the employment questions as a review. Because the MEPS asked only this subset of questions if job status for a person did not change in later rounds, many job-level variables on the subsequent round's job records are coded as "Inapplicable" (-1); the complete information for a continued job is located on the record for the job in the first round in which it was reported. Thus, it is important to determine whether a job continues from the previous round when working with the job records. In rounds where this applies, the variables STILLAT (for jobs that were current main in the previous round) and STILLWRK (for jobs that were current miscellaneous in the previous round) indicate whether a person still holds the job at the subsequent round interview date. The variable SUBTYPE on the subsequent round record

indicates whether the job is main or miscellaneous in that subsequent round. Note that if a Panel 25 job included in this 2021 file is continued from a job first reported in Round 1 or 2, or a Panel 24 job included in this 2021 file is continued from a job first reported in Round 3 or 4, or a Panel 23 job included in this 2021 file is continued from a job first reported in Round 5 or 6 (in the 2020 file), much of the information will be contained in the 2020 jobs file (HC-218). Likewise, if a Panel 24 job included in this 2021 file is continued from a job first reported in Round 1 or 2, or a Panel 23 job included in this 2021 file is continued from a job first reported in Round 3 or 4 (in the 2019 file), much of the information will be contained in the 2019 jobs file (HC-211). Finally, if a Panel 23 job included in this 2021 file is continued from a job first reported in Round 1 or 2 (in the 2018 file), much of the information will be contained in the 2018 Jobs file (HC-203). Users should access these prior year files to obtain the desired job characteristics. Appendix 1 provides a sample SAS program showing how to do this, and Appendix 2 provides a sample Stata program showing how to do this. Both sample programs take into account the addition of the fourth panel.

Any new job reported in a round following the initial interview is collected the same way as in the first interview round.

Variables that relate only to the review of a job reported in a previous round (DIFFWAGE, ESTBTHRU, INSESTB, MAIN_JOB, NOWTAKEI, NOWTAKEI_M22 (in applicable rounds), OFFTAKEI, STILLAT, StillWorkFTPT, STILLWRK, RvwTotNumEmp, WHY_LEFT_M18, WhyChngPTToFT, WhyChngFTToPT) were not asked in Round 1, and these variables are coded as "Inapplicable" (-1) on a Jobs record for the round in which the job is initially reported.

Another type of job update pertains to situations where a reviewed current miscellaneous job becomes the current main job in the round. The flag variable TYPECHGD indicates if a job changed from a current miscellaneous job to a current main job. For these types of jobs, questions asked when the job was first reported as a current miscellaneous job are not re-asked, with three exceptions.

1. Responses to either EM540 or EM620 (typical hours worked per week) are used to populate the variable HRSPRWK. When originally reported, the current miscellaneous job was asked EM620 (but not asked EM540). As a current main job, it will now be asked EM540 instead of EM620. Consequently, there may be different values on HRSPRWK between rounds.

2. Responses to either EM560 or EM630 (whether job is temporary) are used to populate the variable TEMPJOB. When originally reported, the current miscellaneous job was asked EM630 (but not asked EM560). As a current main job, it will now be asked EM560 instead of EM630. Consequently, there may be different values on TEMPJOB between rounds.

3. Responses to either EM570 or EM640 (whether job is seasonal) are used to populate the variable SESNLJOB. When originally reported, the current miscellaneous job was asked EM640 (but not asked EM570). As a current main job, it will now be asked EM570 instead of EM640. Consequently, there may be different values on SESNLJOB between rounds.

Exceptions to the "Inapplicable" (-1) Rule

Unlike the situation explained above (applicable for most variables on the file), for certain variables a value other than "Inapplicable" (-1) does not necessarily mean that a job is newly reported. For a small subset of variables, previous round variables are carried forward to the next round, even if there have been no updates to the variables since they were originally reported. There are two distinct situations in which this special treatment is used, due to internal processing needs.

The first type of exception occurs when questions related to the affected variables are skipped over as "Inapplicable" (-1) during the interview in rounds subsequent to the one in which the job was initially reported, but have their originally reported response carried forward from round to round. This group includes the following 15 variables: EMPLINS, HRSPRWK, HRS35WK, JOBTYPE, JSTRTY, JSTRTM, MORELOC, NUMEMPS, OFFRDINS, PROVDINS, EmplUnionProv (in applicable rounds), TYPEEMPL, JOBHASHI, HRSALBAS, and RETIRJOB. Note that HRSALBAS and RETIRJOB may also be updated in subsequent rounds.

The second type of exception occurs for certain questions that are asked during the review of a job in rounds following the round in which the job was initially reported. If there is no change based on the review, the value for the affected variable is copied forward from the previous round. If there is a change, the variable is updated to reflect the new information. These six variables are: JSTOPY, NOWTAKEI, NOWTAKEI_M22 (in applicable rounds), OFFTAKEI, SUBTYPE, and TOTLEMP_M18.

Variables related to earnings (such as HRLYWAGE, GROSSPAY, SALARIED) are treated similarly to the six variables just discussed. In the review section, the MEPS attempted to obtain information regarding changes in wages for the same job from round to round. If there were no wage changes (indicated by the DIFFWAGE variable), then the most recent round's information was carried forward. If changes were recorded, then the relevant variables were updated. For every new job reported for a person, the MEPS attempted to obtain current wage information.

Top-Coding, Bottom-Coding, Editing, and Confidentiality

Outlier Wage Editing on Current Main Jobs

In 2021, wage information on current main job records is logically edited for consistency using established rules and guidance from AHRQ. Outliers are checked for persons who report a wage change and the new reported wage (a) is substantially different from prior wage (change >=100%), (b) is no different than prior wage, (c) is low in value ($0 < wage < $1) or, (d) has a value higher than prior year's top code value. There are numerous sources for these types of errors, including keystroke or respondent error. In 2021, approximately 100 wages were reviewed per panel, resulting in approximately 50 wage edits (overall).

Users should keep in mind that such edits were not performed in 2020. To help users identify cases that would have been reviewed (but not necessarily edited) in this process, the 2020 Population Characteristics Public Use file data includes wage outlier flag variables,

OUTFLAGrr. These round-specific wage outlier flag variables OUTFLAG31, OUTFLAG42, and OUTFLAG53 indicate that a person's updated wage at the current main job would have been programmatically selected for review during the 2020 wage outlier editing process (but not necessarily edited). More information on these variables may be found in MEPS HC-224: 2020 Full Year Consolidated Data File documentation. These variables were not constructed for the 2021 Population Characteristics file since outlier reviews were performed.

Wage Top-Coding

For reasons of confidentiality, earnings variables on the 2021 Jobs file were top-coded. The earnings variables include HRLYWAGE, BONSAMT, COMMAMT, TIPSAMT, DAYWAGE, WKLYAMT, GROSSPAY, and MAKEAMT. A value of "TOP CODED" (-10) for one of these variables on a record indicates that the variable had a positive value and that the hourly rate for that earnings variable for the record was greater than or equal to $105.77. The process by which the top-code value for the Jobs file is derived incorporates the wage top-code process used in the 2021 Population Characteristics Public Use file top-coding process. The purpose of top-coding is to ensure confidentiality for each person across files.

In addition to using wages from the first report of a current main job, updated wages from that job reported in any subsequent round are also included in deriving the wage top-code value. On the 2021 Population Characteristics Public Use file, any person who has a wage for any job in any round that is greater than or equal to the top-code value will have all wages for all jobs top-coded, regardless of round. Any person whose wages are top-coded on the 2021 Population Characteristics Public Use file also has *all* wages on *all* jobs top-coded in the 2021 Jobs file.

Moreover, because other jobs where wages are reported are included in the 2021 Jobs file but not summarized in the 2021 Population Characteristics Public Use file (i.e., newly reported former main jobs and current/former miscellaneous jobs), and these wages may exceed the current year top-code value, wages for these jobs and all jobs belonging to the same jobholder are top-coded on the 2021 Jobs file. In turn, the wages of these persons are top-coded in the 2021 Population Characteristics Public Use file as well.

Note that there are also some jobs where respondents indicate that a supplemental wage, such as a commission, tip, or bonus, is greater than or equal to the wage top-code value but, at that same job, base wage such as the annual salary is not. For these cases, only the tips, commissions, or bonus amounts were top-coded on the job where they are greater than or equal to the wage top-code value (note, these supplemental wages reside on the 2021 Jobs file but not on the 2021 Population Characteristics file). All other wage amounts for all jobs for these persons were left as reported. (This applies to wages and jobs on both the 2021 Population Characteristics Public Use and 2021 Jobs files.)

As of the 2020 Jobs file, wages are also top-coded to -10 on the Jobs file for two situations where wages were formerly reset to "Cannot Be Computed" (-15) in prior year Jobs files (2019 and earlier). These situations are:

1.   If wages at a current main job were imputed on the Full Year Population Characteristics file to a value less than the top code value but calculated on the Full Year Jobs file greater than or equal to the top code value, or

2.   If wages at a current main job that changes to a current miscellaneous job are greater than or equal to the top code value. Note that wages earned through a miscellaneous job are not reported on the Full Year Population Characteristics file.

Additional Wage Information

To improve the quality of wage reports, CAPI prompts the respondent to confirm wages reported in the Employment Wage section if a wage amount falls outside a specified wage range. Ranges vary depending on the unit of pay as follows:

Unit of Pay	Wage Range
Per year	$5,000.00 - $200,000.00
Per month	$375.00 - $20,000.00
Per 2-week period	$150.00 - $10,000.00
Per week	$75.00 - $5,000.00
Per day	$10.00 - $750.00
Per hour	$1.00 - $125.00

To calculate the hourly rate for earnings types not reported on an hourly basis, the number of hours per week worked and in some cases the number of weeks worked were used in conjunction with the various amounts. These hours and weeks are included on the file along with the reported earnings amounts, but not the calculated hourly rates. (Earnings variables were not reconciled with income data collected elsewhere in the MEPS.)

Establishment Size Information

The establishment size variable for the self-employed is TOTLEMP_M18. In addition, two variables are available containing the individual responses collected at RJ110 and EM740 (number of employees at a self-employed job). They are RvwTotNumEmp (establishment size at continuing self-employed job) and TotNumEmp (establishment size at newly reported self-employed job), respectively.

The establishment size for wage-earners can be found in NUMEMPS (establishment size at non-self-employed job); this value is collected at EM430 (number of employees). Respondents who did not know the actual establishment size (NUMEMPS) are asked in question EM440 to choose approximate establishment size from a number of size ranges. These responses are used to create

the variable ESTMATE1_M19. The categorical values available to respondents in EM440 are as follows.

Value	Category
-1	Inapplicable
-7	Refused
-8	Don't Know
2	2-9
3	10-25
4	26-49
5	50-100
6	101-500
7	501-1,000
8	1001-5,000
9	5001+

The value "Cannot Be Computed" (-15) is not an allowed value for ESTMATE1_M19.

For confidentiality reasons, NUMEMPS, TOTLEMP_M18, RvwTotNumEmp and TotNumEmp were top coded to "-10 # OF EMP >= 18,000" for establishment sizes greater than or equal to 18,000 employees.

Job Start/Stop Year

In addition to top coding wages and establishment size, the start year of job (JSTRTY) and the stop year of job (JSTOPY) are bottom-coded. This is done because a person's age may be calculated using the job start or stop year and that age may indicate that the jobholder is older than 85 years, the age top-code value. This value is calculated by taking the delivery year in which the job is first reported (e.g. 2021), subtracting the age top-code value (i.e. 85 years of age), then adding back 15 (i.e. the age of a person in the year before entering the work force as defined in MEPS). For the 2021 Jobs file, the bottom code value for the job start and stop year on jobs first reported in Panel 26 Round 1, Round 2, or Round 3; Panel 25 Round 4 or Round 5; Panel 24 Round 6 or Round 7; or Panel 23 Round 8 or Round 9 is 1951. Jobs that were first reported in Panel 25 Round 1, Round 2, or Round 3; Panel 24 Round 4 or Round 5; or Panel 23 Round 6 or Round 7 were delivered in the 2020 Jobs file and have a bottom code value of 1950. Jobs that were first reported in Panel 24 Round 1, Round 2, or Round 3, or Panel 23 Round 4 or Round 5 were delivered in the 2019 Jobs file and have a bottom code value of 1949. Lastly, jobs that were first reported in Panel 23 Round 1, Round 2 or Round 3 were delivered in the 2018 Jobs file and have a bottom code value of 1948.

Temporary and Seasonal Jobs

Two variables on the file pertain to the temporary and seasonal nature of a person's main or miscellaneous job. The variable TEMPJOB indicates whether a main or miscellaneous job is temporary (i.e., is a current main job for a limited amount of time or until the completion of a project). The variable SESNLJOB indicates either that a main or miscellaneous job is available only during certain times of the year or that the individual is working throughout the entire year at that job. Teachers and other school personnel who work only during the school year are considered to work year round. These questions are asked of newly reported jobs only. These variables are set to "-1 INAPPLICABLE" for all subsequent rounds. These questions are not asked of newly reported former miscellaneous jobs, last jobs outside of reference period, and retirement jobs.

Reason No Longer at Place of Employment

In cases where a former job is newly reported, questions are asked regarding why the person is no longer at that place of work. For wage earners, this information is found in YLEFT_M18. For self-employed persons, this information is collected in YNOBUSN_M18.

It is important to note that the retirement job classification in the variable SUBTYPE is independent of any retirement response in the following variables:

- YNOBUSN_M18, which relates to the question why a person no longer has a self-employed business;

- WHY_LEFT_M18, which relates to the question why a person left a job in the current round.

Health Insurance Data

Questions about employment-related health insurance are asked both when any type of job is newly reported and when any continuing job is reviewed. For main jobs, either newly reported or changing from miscellaneous, the variable that indicates whether insurance is held through that establishment is EMPLINS. For all non-main jobs, the variable JOBHASHI indicates whether insurance is held through that establishment.

For a newly reported job, depending on whether employment-related insurance is held or not, there may be follow-up information gathered which is contained in the following variables:

- OFFRDINS, which notes whether health insurance is offered through the job in cases where the jobholder reports that they do not hold health insurance through the job;

- DIFFPLNS, which notes whether a choice of health insurance plans is available for cases where the jobholder reports that health insurance is either offered or held through the job;

- ANYINS, which notes whether health insurance coverage through the job is available to any other employees at the establishment in cases where the jobholder does not hold health insurance through the job and is not offered health insurance coverage through the job.

If a job holder holds insurance at the employer ("Yes" (1) at EM660, EMPLINS or JOBHASHI) and that person belongs to a union ("Yes" (1) at EM700, INUNION), and the job is first reported in the round, respondents are asked to indicate if the health insurance is from the employer/business or the union at EM710. Either or both establishments may be the source of insurance. Through Panel 23 Round 8, Panel 24 Round 6, Panel 25 Round 4, and Panel 26 Round 2, both establishments could be selected at EM710 (PROVDINS), and two sets of private insurance coverage were created in the Health Insurance (HX) section of MEPS.

1 EMPLOYER

2 UNION

3 BOTH EMPLOYER AND UNION

Beginning in Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3, response options at EM710 have changed. Respondents are now required to identify the *primary* source of health insurance - either the employer/business or the union - if the respondent indicates both provide insurance at EM710 (PROVDINS renamed EmplUnionProv for these cases).

1 EMPLOYER

2 UNION

3 BOTH EMPLOYER AND UNION (EMPLOYER IS PRIMARY)

4 BOTH EMPLOYER AND UNION (UNION IS PRIMARY)

Only the primary source of insurance coverage will be created in the HX section. The result is that persons who report insurance via both union and employer sources will no longer have the secondary source insurance coverage recorded in HX.

The variable set at EM710 is renamed from PROVDINS to EmplUnionProv beginning in Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9 to reflect this change. EmplUnionProv is "Inapplicable" (-1) in prior rounds. Note that PROVDINS was constructed for all rounds of 2021 using responses collected in EmplUnionProv for jobs newly reported in Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9. Where EmplUnionProv = "Both Employer and Union (Employer is Primary)" (3) or "Both Employer and Union (Union is Primary)" (4), PROVDINS = "Both" (3).

Users should be mindful that 2021 is a transition year, since jobs in the 2021 file report primary/secondary union/employer insurance differently depending on which panel/round the job was first reported. Users combing multiple years of MEPS should also be mindful that prior

years in MEPS contained separate insurance records of both primary and secondary private insurance.

For a continuing job, when no health insurance was held through the job in the round in which the job was first reported but health insurance was offered through the job, the question RJ70 OFFTAKEI is asked in later rounds to determine whether the employee now holds the health insurance that is offered through the job. (Note: if health insurance through this job was reported as being held via RJ70 in the prior round, RJ70 is not asked in the current round.)

Similarly, the insurance status question RJ80 (responses stored on NOWTAKEI through Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8, and on NOWTAKEI_M22 beginning in Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9) is asked to determine whether health insurance is now held through the job in the following cases:

- insurance through the job ended in a prior round or

- insurance coverage was never reported through the job and the person was not offered insurance through the job in the round a job was first reported or

- the respondent disavows coverage through the job in the Health Insurance section that was previously indicated in the Employment section of the interview or

- the respondent reports new employer-sponsored health insurance in the prior round but coverage was not active at the interview date (see below).

Beginning in 2021 Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9, RJ80 is now asked if the respondent reports new employer-sponsored health insurance in the prior round but that coverage was not active at the interview date, that is, a response of "No" (2) in the Health Insurance section of MEPS at HQ01 "Was {PERSON} covered the whole time from {START DATE} until {END DATE}" and at HQ02 "Is {PERSON} covered now?" Prior to this change, persons for whom health insurance was not active at the interview date in the prior round skipped RJ80.

To reflect the new CAPI flow, the variable set at RJ80 is renamed from NOWTAKEI to NOWTAKEI_M22 in the 2021 Jobs file. The 2021 Jobs file contains both variables. NOWTAKEI reflects responses at RJ80 through Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8, and is "Inapplicable" (-1) in later rounds. NOWTAKEI_M22 reflects responses at RJ80 beginning in Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9 and is "Inapplicable" (-1) in prior rounds.

MEPS then includes several clarifying questions regarding health insurance availability at an employer. Where the person does not report, does not know, or refuses to indicate the insurance coverage status through the job at RJ70 or reports no insurance coverage through the job at RJ80, the respondent is asked if the person was offered insurance through the job at RJ90 (ESTBTHRU).

Lastly, when a respondent indicates that the jobholder of a reviewed job neither holds insurance through the job nor was offered health insurance at the job, the respondent is asked if *any other*

employees were offered health insurance through the job at RJ100 (INSESTB). The Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9 CAPI change at RJ80 means that more persons could be asked if the person was offered insurance at RJ90 (ESTBTHRU) or if other employees were offered insurance at the employer establishment at RJ100 (INSESTB), discussed below.

In some cases, respondents will indicate in the Health Insurance section that health insurance reported in the Employment section was either wholly or partially reported in error. This is referred to as insurance being "disavowed." If newly reported health insurance through the job is disavowed in the Health Insurance section, follow-up questions (HX21, HX22, HX23) regarding whether health insurance is offered at the job, whether more than one plan is available, and whether health insurance is offered to any employees are asked in the Health Insurance section. This information is used in an editing process whereby responses in the Health Insurance section are transferred into the Employment or Review of Jobs sections. As a result, the disavowal process may result in a change to values originally collected in the Employment or Review of Jobs section (wherever the health insurance was initially reported). The complete list of variables potentially impacted includes: EMPLINS, JOBHASHI, OFFRDINS, DIFFPLNS, ANYINS, PROVDINS, and EmplUnionProv, collected in the Employment section, and NOWTAKEI, NOWTAKEI_M22, OFFTAKEI, ESTBTHRU, and INSESTB, collected in the Review of Jobs section. In some cases, a disavowal may result only in a change to the value of PROVDINS/EmplUnionProv.

Through Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8, health insurance through an employer can be disavowed in MEPS based on a respondent's answer to one of four questions (HX14, HX15, HX20, HP70). Beginning in Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9, disavowal is possible at one of two questions (HX20 and HP70).

To help users understand the source of the disavowal, the variable HIDISAVW indicates which of the following questions resulted in the disavowal. HIDISAVW will include only one source among these options. Please note, however, that through Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8, it is possible for a respondent to disavow one source of coverage at HX15 and then later disavow the second source of coverage at HP70. In these cases, HIDISAVW will be set to HP70.

1. HX14 - This question is asked through Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8 if both employer and union coverage are reported at EM710 (PROVDINS) to determine if there is 1 ONE PLAN, 2 TWO PLANS, or if 3 INSURANCE WAS REPORTED IN ERROR. HIDISAVW = HX14 indicates that HX14 = 3 and that there is neither insurance coverage through the employer nor insurance coverage through the union and that updates were made to the insurance variables collected in the Employment section (EMPLINS, JOBHASHI, OFFRDINS, DIFFPLNS, ANYINS, NOWTAKEI, OFFTAKEI, ESTBTHRU, INSESTB, PROVDINS) during the disavowal clean-up process. HX14 is omitted beginning in Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9 due to the CAPI change that requires respondents to select the primary source of health insurance at EM710 (EmplUnionProv).

2. HX15 - This question is asked through Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8 if, at HX14, the respondent indicates 1 ONE PLAN (HX14 = 1). At HX15, the respondent must select either insurance coverage through the employer or insurance coverage through the union. Depending on which of these are chosen (employer or union) the other source of coverage was disavowed. For example, if HX14 = 1 and HX15 = employer, the insurance coverage through the union will be disavowed. The originally reported value of PROVDINS = 3, both employer and union, will be edited to PROVDINS = 1, employer only. Conversely, if HX15 = union, the insurance coverage through the employer will be disavowed, and the originally reported value of PROVDINS = 3 will be edited to PROVDINS = 2, union only. HX15 is omitted beginning in Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9 due to the CAPI change that requires respondents to select the primary source of health insurance at EM710 (EmplUnionProv).

3. HX20 - This question is asked if either insurance coverage through the employer only or insurance coverage through the union only are reported at EM660 ((EMPLINS or JOBHASHI=1) and INUNION<>1) or, through Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8, EM710 (PROVDINS = 1 EMPLOYER ONLY or 2 UNION ONLY), or, beginning in Panel 26 Round 3, Panel 25 Round 5, Panel 24 Round 7, and Panel 23 Round 9, EM710 (EmplUnionProv = 1 EMPLOYER, 2 UNION, 3 BOTH EMPLOYER AND UNION (EMPLOYER IS PRIMARY), or 4 BOTH EMPLOYER AND UNION (UNION IS PRIMARY). If the respondent volunteers that the job-related insurance coverage reported at HX20 was in error, the insurance coverage reported in the Employment or Review of Jobs section is removed during the disavowal clean-up process.

4. HP70 - This question is asked of private health insurance coverage through a job that was reported in the Employment section. The respondent is asked to verify that the jobholder is the policyholder of the job related insurance coverage. If the response is NO, REFUSED, DON'T KNOW, the job-related insurance coverage is removed during the disavowal clean-up process.

Industry and Occupation Coding

Industry and occupation codes were assigned by professional coders at the Census Bureau based on verbatim descriptions provided by respondents during the survey interview. The codes are determined at a detailed 4-digit level and then collapsed into broader groups on the file to ensure the confidentiality of the records. INDCODEX contains industry information and OCCCODEX contains occupation information. Appendices 3 and 4 contain crosswalks between the detailed and collapsed codes for industry and occupation.

With the 2010 file, the Census Bureau began using 2007 Industry and 2010 Occupation codes, which were developed for the Bureau's Current Population Survey and American Community Survey. These updated coding schemes incorporate minor changes from the 2003 industry and occupation codes used for the 2002-2009 files; therefore, INDCODEX and OCCCODEX for

2010 and later files will be comparable to those variables on the 2002-2009 files. (Industry and occupation variables for pre-2002 files are not comparable to those for later files.)

2.2 Other 2021 File Considerations

Round-to-Round Changes to Job Rosters

As seen starting in 2020, COVID-19 continues to greatly impact response rates, increasing the likelihood that job characteristics of MEPS job holders vary more than typical MEPS round-to-round changes. Jobholder sex, race, educational attainment, industry, occupation, establishment size, and job place flexibility all impact employment stability more so since the advent of COVID-19.

In 2020, non-responding households in Panel 25 produced a higher drop-off in jobs reviewed in Round 2 than in prior first year panels. For instance, between Round 1 and Round 2, there was a 54% drop-off in job records in Panel 25, 14% more than the drop-off seen in Panel 24 between Round 1 and 2, and 15% more than the drop-off seen in Panel 23 between Round 1 and Round 2. Therefore, in 2021, record counts of jobs available for review is substantially lower than typical second year panel. Also in 2021, the drop-off rate continues at a high rate in Panel 25. Between Round 3 and Round 4, there was a 16% drop-off in job records reported for Panel 25. This is comparable to the drop-off in jobs in Panel 24 between Round 3 and Round 4, and 13% higher than the drop-off seen in Panel 23 between Round 3 and Round 4.

The drop-off rate seen in Panel 26 between Round 1 and Round 2 is comparable to that seen in Panel 25 at approximately 52%. Users should also note that the proportion of new jobs reported in Panel 26 Round 1 as compared to all new jobs reported in Panel 26 Rounds 1-3 is comparable to other panels. In a typical year, over 80% of all jobs in a panel are reported in the first round of the survey. This is true for Panels 23, 24, 25, and 26. For these reasons, it will be important for users to compare jobholder and job characteristics in their analysis. Given the low response rates in Panel 26 Round 1 and continuing drop-off of reviewed jobs in Panel 25, COVID-19 will have a sustained impact on Panel 25 and Panel 26 employment data more so than in any other panels.

Questions where Respondents Note COVID-19 Impacts

Users of the 2021 Jobs file may find it helpful to know where in Employment sections field interviewers documented COVID-related impacts from respondent comments. Although a smaller subset of respondent comments than those noted in 2020, this list demonstrates the continuing impact of COVID-19 on employment data:

- temporary increase or reduction in hours worked (EM540 HRSPRWK, RJ40 StillWorkFTPT, RJ50/RJ55 WhyChngFTToPT/PTToFT)

- layoffs (RJ110 WHY_LEFT_M18, EM520 YLEFT_M18, RJ40 StillWorkFTPT, RJ50/RJ55 WhyChngFTToPT/PTToFT, RJ120_01 JSTOPM, RJ120_03 JSTOPY)

- wage changes (both wage reduction and bonuses or other 'special' pay obtained for work during the COVID pandemic)

- impact on health insurance (EM660 EMPLINS)

2.3   Person-Level Estimates

This 2021 Jobs file does not include any weights necessary to extrapolate this data to the U.S. population. To make person-level estimates, link to any of the 2021 MEPS files and use the person-level weight for the appropriate panel. The link should be made through the variable DUPERSID. Note that not all persons in the MEPS have positive weights and job records; only those persons who have either a positive person-level or family-level weight in the 2021 Population Characteristics Public Use file are included in the 2021 Jobs file.

2.4   Codebook Structure

For each variable on the 2021 Jobs file, an unweighted frequency is provided in the accompanying codebook file.

2.5   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
-10	TOP CODED	Variable was top-coded for confidentiality, as described above
-15	CANNOT BE COMPUTED	Value cannot be derived from data

The value -15 (CANNOT BE COMPUTED) assigned to MEPS constructed variables in cases where there is not enough information from the MEPS instrument to calculate the constructed variable. "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".

2.6   Codebook Format

This codebook describes an ASCII dataset (with related SAS, SPSS, R, and Stata programming statements and data user information), although the data are also provided in a SAS data set, SAS transport file, Stata data set, and Excel file. The file contains 86 variables and has a logical record length of 283 with an additional 2-byte carriage return/line feed at the end of each record. The following codebook items are provided for each variable:

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

2.7   Variable Source and Naming Conventions

As variable collection, universe, or categories are altered, the variable name will be appended with "_Myy" to indicate in which year the alterations took place. Details about these alterations can be found throughout this document.

In general, variable names reflect the content of the variable. Due to system changes, variable names are no longer restricted to 8 characters. Variables contained on this file were derived from the questionnaire itself or from the CAPI. The source of each variable is identified in Section D. Variable-Source Crosswalk. Sources for each variable are indicated in one of two ways:

1. Variables derived from CAPI or assigned in sampling are so indicated as "CAPI Derived" or "Assigned in Sampling," respectively;

2. Variables that come from one or more specific questions have those questionnaire sections and/or question numbers listed in the "Source" column.

3.0   Discussion of Pandemic Effects on Quality of 2021 MEPS Data

3.1   Summary

The challenges associated with MEPS data collection in 2020 after the onset of the COVID-19 pandemic continued into 2021. The major modifications to the standard MEPS study design remained in effect, permitting data to be collected safely but with accompanying concerns related to the quality of the data obtained. These data quality issues are discussed below. The suggestion

made in the documentation for the FY 2020 MEPS Consolidated PUF data (as well as for most federal major in-person surveys conducted in 2021 and 2020) still holds. Researchers are counseled to take care in the interpretation of estimates based on data collected from these two calendar years. This includes the comparison of such estimates to those of other years and corresponding trend analyses.

3.2   Overview

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

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

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

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

Households fielded for Round 1 of MEPS in each year have been selected as a subsample from among the NHIS responding households from the prior year. The MEPS first year panel in 2021 was Panel 26. The households fielded for MEPS in Round 1 of Panel 26 were thus selected from NHIS responding households in 2020. It is important to note here that the NHIS households eligible for use in MEPS are restricted to the first three quarters of the NHIS as the fourth quarter households cannot be made available in time for MEPS data collection early in the next calendar year.

The onset of the pandemic in 2020 at a national level took place in mid-March of that year, when the NHIS data collection for the first quarter of 2020 was virtually completed and that of the second quarter was about to begin. The NHIS had to make a rapid transition from in-person to telephone interviewing in order to attempt to gather NHIS data for the second quarter of 2020. While NCHS was able to make the transition, assessments made by NCHS at the time indicated a much lower response rate than is typically experienced during Quarter 2 and the quality of Quarter 2 data was of particular concern. NCHS thus modified the 2020 NHIS sample design for Quarters 3 and 4. A randomly selected subsample of the sampled housing units originally

selected for fielding in Quarters 3 and 4 of 2020 was removed from the sample to be fielded. This reduced sample for Quarters 3 and 4 was then enhanced by randomly selecting responding households from the 2019 NHIS for interviewing in 2020 as well. In consideration of the data quality issues and sample design modifications associated with the 2020 NHIS, the MEPS sample design for FY 2021 was modified, as will be discussed shortly.

With respect to the CPS, the quality of CPS data is always of particular importance to MEPS as March CPS-ASEC estimates serve as the basis of control totals for the raking component of the MEPS weighting process. These control totals incorporate the following demographic variables: age, sex, race/ethnicity, region, MSA status, educational attainment, and poverty status. The CPS estimates of educational attainment and poverty status used in the development of the FY 2021 MEPS PUFs were of particular concern. Evaluations of these estimates undertaken by the Census Bureau have shown that they suffered from bias due to survey nonresponse with CPS income estimates being on the high side and the estimate of those under poverty being on the low side. The impact of these CPS estimates on the quality of MEPS estimates has been carefully considered. The approach used for the MEPS FY 2021 Population Characteristics PUF sample weights, where educational attainment is employed in the raking process but poverty status is not, is discussed in Section 3.5.

A set of references discussing the fielding of these surveys during the pandemic and since, including possible bias concerns, can be found in the References section of this document.

3.4   Modifications to the MEPS HC 2021 Sample Design

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

AHRQ contacted NCHS, reviewing the various issues and asking if it would be possible that responding households in NHIS Panels 2 and 4 from Quarter 1 of 2020 be made available for MEPS sample selection. Virtually all of these households were interviewed in-person prior to the major onset of the pandemic, so the Quarter 1 response rates for all four NHIS panels were consistent with prior years and the data quality issues associated with the pandemic could be avoided. NCHS was fully supportive of this approach and made NHIS Panels 2 and 4 for Quarter 1 available for use by MEPS. Thus, for MEPS Panel 26, the NHIS responding households subsampled from MEPS were selected from among all NHIS responding households in Quarter 1 as well as those responding in Quarter 3 that were not originally sampled for the 2019 NHIS.

As an adjunct to this modification, it was decided to take advantage of the additional PSUs (sampled localities) available from NHIS Panels 2 and 4 and appearing in the MEPS sample for the first time. State level estimation is of interest to MEPS, and the added PSUs would serve to

increase the precision for state level estimates. State estimates that would be expected to benefit the most from these added PSUs were the "middle-sized" states. The largest states already had large sample sizes while precision for the smallest states would remain low. As a result, the MEPS sample focused on oversampling the "middle-sized" states rather than Hispanics, Blacks, and Asians, as has usually been the practice.

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

3.5   Data Quality Issues for MEPS for FY 2021

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

With this in mind, there were a number of uncertainties for FY 2021 warranting examination. Would Panel 23 data quality increase substantially once the issue of an extensive recall period was eliminated? Would the switching from in-person to telephone interview in Round 1 continue to impact Panel 25 estimates? Since Panel 26 was the first year MEPS panel in 2021, would Panel 26 estimates tend to be different than those of the other three panels?

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

The analyses undertaken thus far suggest no major differences between the four panels for the distribution of health insurance status. Even though slight differences were observed with Panel 25 (e.g., the distribution associated with the age range 18-64 showed a higher percentage of all public insurance compared to the other three panels while those at least 65 years of age showed a lower percentage of some private insurance compared to the other three panels), no statistically significant differences were detected.

In summary, the FY 2021 Population Characteristic PUF weights used to select jobholders for the 2021 Jobs Public Use release file can be expected to produce useful estimates for initial analyses of MEPS 2021 data. Further analyses of MEPS estimates will be conducted as part of the production of the FY 2021 Consolidated PUF to be released later in 2023. This will help identify any additional data quality issues as well as possible improvements that could be implemented.

3.6   Discussion and Guidance

The various actions taken in the development of the person-level weights for the MEPS FY 2021 Population Characteristics PUF, used to select jobholders for the 2021 Jobs Public Use release file, were designed to limit the potential for bias in the data due to changes in data collection and response bias. However, evaluations of MEPS data quality in 2021, consistent with those of other Federal surveys fielded in 2021, suggest that users of the MEPS FY 2021 Jobs PUF should exercise caution when interpreting estimates and assessing analyses based on these data as well as in comparing 2021 estimates to those of prior years.

4.0   Longitudinal Analysis

Panel-specific longitudinal files are available for downloading in the data section of the MEPS website. For all four panels (Panel 23, Panel 24, Panel 25, and Panel 26), the longitudinal file comprises MEPS survey data obtained in Rounds 1 through 5 of the panel and can be used to analyze changes over a two-year period. In addition, for Panel 23 a longitudinal file that comprises MEPS survey data obtained in Rounds 1 through 7 of the panel can be used to analyze changes over a three-year period. For Panel 24 a file representing a three-year period will also be established, and for Panel 23 a file representing a four-year period will be established. Variables in the file pertaining to survey administration, demographics, employment, health status, disability days, quality of care, health insurance, and medical care use and expenditures were obtained from the MEPS full-year Consolidated files from the years covered by that panel. For more details or to download the data files, please see Longitudinal Weight Files.

4.1   Using MEPS Data for Trend Analysis

First, of course, we note that there are uncertainties associated with 2020 and 2021 data quality for reasons discussed in Section 3.5. Preliminary evaluations of a set of MEPS estimates of particular importance suggest that they are of reasonable quality. Nevertheless, analysts are advised to exercise caution in interpreting these estimates, particularly in terms of trend analyses since access to health care was substantially affected by the COVID-19 pandemic as were related factors such as health insurance and employment status for many people.

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

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

trends in utilization across years. The changes in the NHIS sample design in 2016 and 2018 could also potentially affect trend analyses. The new NHIS sample design is based on more up-to-date information related to the distribution of housing units across the U.S. As a result, it can be expected to better cover the full U.S. civilian, noninstitutionalized population, the target population for MEPS, as well as many of its subpopulations. Better coverage of the target population helps to reduce the potential for bias in both NHIS and MEPS estimates.

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

Process changes, such as data editing and imputation, may also affect trend analyses. For example, users should refer to Section 2.5.11 in the 2021 Consolidated Public Use file (HC-232) and, for more detail, the documentation for the prescription drug file (HC-228A) when analyzing prescription drug spending over time.

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

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

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

References

Bramlett, M.D., Dahlhamer, J.M., & Bose, J. (2021, September). Weighting Procedures and Bias Assessment for the 2020 National Health Interview Survey. Centers for Disease Control and Prevention.

Current Population Survey: 2021 Annual Social and Economic (ASEC) Supplement. (2021). U.S. Census Bureau.

Dahlhamer, J.M., Bramlett, M.D., Maitland, A., & Blumberg, S.J. (2021). Preliminary evaluation of nonresponse bias due to the COVID-19 pandemic on National Health Interview Survey estimates, April-June 2020. National Center for Health Statistics.

Lau, D.T., Sosa, P., Dasgupta, N., & He, H. (2021). Impact of the COVID-19 Pandemic on Public Health Surveillance and Survey Data Collections in the United States. *American Journal of Public Health, 111* (12), pp. 2118-2121.

Rothbaum, J. & Bee, A. (2021, May 3). Coronavirus Infects Surveys, Too: Survey Nonresponse Bias and the Coronavirus Pandemic. U.S. Census Bureau.

Rothbaum, J. & Bee, A. (2022, September 13). How Has the Pandemic Continued to Affect Survey Response? Using Administrative Data to Evaluate Nonresponse in the 2022 Current Population Survey Annual Social and Economic Supplement. U.S. Census Bureau.

Zuvekas, S.H. & Kashihara, D. (2021). The Impacts of the Covid-19 Pandemic on the Medical Expenditure Panel Survey. *American Journal of Public Health, 111 (12)*, pp. 2157-2166.

D. Variable-Source Crosswalk

FOR MEPS PUBLIC USE RELEASE HC-227

SURVEY ADMINISTRATION VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
JOBSIDX	Job-round identifier	CAPI Derived/Encrypted
JOBIDX	Person's unique job identifier	CAPI Derived/Encrypted
JOBNUM	Unique DU-job identifier	CAPI Derived
ESTBIDX	Establishment identifier	CAPI Derived/Encrypted
DUPERSID	Person ID (DUID + PID)	Assigned in Sampling
DUID	Panel # + encrypted DU identifier	Assigned in Sampling
PID	Person Number	Assigned in Sampling
RN	Round	CAPI Derived
OrigRnd	Round job first reported	CAPI Derived
PANEL	Panel to which Jobholder Belongs	Assigned in Sampling

EMPLOYMENT VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
JSTRTM	Job start date - month	EM60_02, EM90_02, EM110_02, EM130_02, EM190_02, EM250_02
JSTRTY	Job start date - year	EM60_01, EM90_01, EM110_01, EM130_01, EM190_01, EM250_01
JSTOPM	Job stop date - month	EM140_02, EM200_02, EM260_02, EM310_02, EM400_02, RJ120_02

VARIABLE	DESCRIPTION	SOURCE
JSTOPY	Job stop date - year	EM140_01, EM200_01, EM260_01,EM310_01, EM400_01, RJ120_01
RETIRJOB	Person retired from this job	EM50, EM80, EM100, EM270, EM380
SUBTYPE	Job sub-type	EM50, EM80, EM100, EM120, EM180, EM270, EM340, EM380, EM390, EM410, RJ10/RJ60
STILLAT	Still works at main job establishment	RJ10
TYPECHGD	Job sub-type changed between rounds	Constructed
MAIN_JOB	Still main job or business	RJ20
DIFFWAGE	Any change in wage amount	RJ30
StillWorkFTPT	Still works full or part time	RJ40
WhyChngPTToFT	Why change part to full time	RJ50
WhyChngFTToPT	Why change full to part time	RJ55
STILLWRK	Still works at misc job establishment	RJ60
OFFTAKEI	Offered insurance and now take	RJ70
NOWTAKEI	Now offered and take insurance	RJ80 [through Fall 2021 (Panel 23 Round 8, Panel 24 Round 6, Panel 25 Round 4, Panel 26 Round 2)]
NOWTAKEI_M22	Now has health insurance through employer	RJ80 [as of Spring 2022 (Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, Panel 26 Round 3)]
ESTBTHRU	Offered insurance, did not take (review)	RJ90
INSESTB	Insurance offered to any employees (review)	RJ100
HIDISAVW	Capi q where health insur thru emp/union disavowed	Constructed from HX responses
RvwTotNumEmp	Establishment size at continuing self-employed job	RJ110

VARIABLE	DESCRIPTION	SOURCE
WHY_LEFT_M18	Reason why no longer at job now	RJ130
JOBTYPE	Self-employed or works for someone else	EM420
NUMEMPS	Establishment size at not self-employed job	EM430
ESTMATE1_M19	Categorical approximate establishment size	EM440
MORELOC	Employer has more than one location	EM450
BUSINC	Business incorporated	EM460
PROPRIET	Proprietorship or partnership	EM470
TYPEEMPL	Employee type	EM480
YLEFT_M18	Reason why no longer at job	EM520
YNOBUSN_M18	Reason why no longer has business	EM530
HRSPRWK	Number of hours worked per week	EM540, EM620
HRS35WK	Works at least 35 hours per week	EM550
TEMPJOB	Job at employer is temporary	EM560, EM630
SESNLJOB	Job is available certain time of year	EM570, EM640
SICKPAY	Has paid sick leave thru job	EM580
PAYDRVST	Has paid sick leave for doc visit thru job	EM590
PAYVACTN	Has paid vacation leave thru job	EM600
RETIRPLN	Has pension/retirement plan thru job	EM610
WKLYAMT	Usual weekly gross income at misc job	EM650
EMPLINS	Has health insurance thru current main job	EM660
JOBHASHI	Has health insurance thru job	EM660
OFFRDINS	Offered insurance but chose not to take	EM670
DIFFPLNS	Choice of different health insurance plans	EM680
ANYINS	Health insurance offered to any employees	EM690
INUNION	Belongs to labor union	EM700

VARIABLE	DESCRIPTION	SOURCE
PROVDINS	Is health ins provided by employer, union, or both	EM710 [through Fall 2021 (Panel 23 Round 8, Panel 24 Round 6, Panel 25 Round 4, Panel 26 Round 2)]
EmplUnionProv	Employer or union is primary health insurer	EM710 [as of Spring 2022 (Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, Panel 26 Round 3)]
HHMEMBER_M18	Any other hh member wrk at this business	EM730
TOTLEMP_M18	Current establishment size at self-employed job	Constructed from EM740 and RJ110
TotNumEmp	Establishment size at new self-employed job	EM740
SALARIED	Person salaried, paid by hour, some other way	EW10
HOWPAID	How is person paid	EW20
DAYWAGE	Person's daily wage rate	EW30
HRSPRDY	Number of hours person worked in one day	EW40
MAKEAMT	How much money does person make	EW50
PERUNIT_M18	Period for which person is paid	EW60
HRLYWAGE	How much person makes per hour	EW70, EW140, EW190
MORE10	Person makes more or less than $10/hour	EW80, EW150, EW200
MORE15	Person makes more or less than $15/hour	EW90, EW160, EW210
MOREMINM	Person makes more or less than min. wage	EW100, EW170, EW220
GROSSPAY	Person's salary before taxes (gross)	EW110
GROSSPER	Period in which gross salary was earned	EW120
SALRYWKS	Number of weeks per year salary is based	EW130
HRSALBAS	Hours per week salary based on	EW180
EARNTIPS	Person earns tips	EW230A
EARNBONS	Person earns bonuses	EW230B
EARNCOMM	Person earns commission	EW230C

VARIABLE	DESCRIPTION	SOURCE
TIPSAMT	How much are person's tips	EW240
TIPSUNIT_M18	Period which tip earnings are based on	EW250
BONSAMT	How much are person's bonuses	EW260
BONSUNIT	Period which bonuses are based on	EW270
COMMAMT	How much are person's commissions	EW280
COMMUNIT	Period which commissions are based on	EW290
INDCODEX	Condensed industry code	Constructed from EM490
OCCCODEX	Condensed occupation code	Constructed from EM500, EM510

Appendix 1
Sample SAS Program

```
8              *** APP21.sas ***;
9
10                OPTIONS LS=132 PS=79;
11
12              ********************************************************************************
13              ***    Program Name:  SAMPLE.SAS                                           ***
14              ***                                                                        ***
15              ***    Description:  This job provides an example of how to get job info  ***
16              ***                  from Round 1 or Round 2 in the FY2018 JOBS file       ***
17              ***                  or Round 3 or Round 4 in the FY2019 JOBS file         ***
18              ***                  or Round 5 or Round 6 in the FY2020 JOBS file when a  ***
19              ***                  continuation current main job in the FY2021 JOBS file ***
20              ***                  is first reported in the FY2018, FY2019 or FY2020     ***
21              ***                  JOBS File.                                            ***
22              ***                                                                        ***
23              ***                  This example creates a dataset of continuation JOBS   ***
24              ***                  records with a SICKPAYX variable copied from the      ***
25              ***                  Round 1, 2, 3, 4, 5 or 6 newly reported job.          ***
26              ***                                                                        ***
27              ********************************************************************************;
28
29               libname  jobs18   "c:\mydata\jobs18";
30               libname  jobs19   "c:\mydata\jobs19";
31               libname  jobs20   "c:\mydata\jobs20";
32               libname  jobs21   "c:\mydata\jobs21";
33
34              ***  a.                                                                    ***
35              ***  Select continuing Panel 23 Round 7 or Panel 24 Round 5 or Panel 25 Round 3 ***
36              ***  Current Main Jobs (SUBTYPE=1, STILLAT=1) from the FY 2021 JOBS file   ***
37              ***  and print selected variables from the first 20 observations          ***;
38
39               data j21r753;
40                  set jobs21.jobs21;
41                     if    ((panel=23 and rn=7 and origrnd<7)
42                            or (panel=24 and rn=5 and origrnd<5)
43                            or (panel=25 and rn=3 and origrnd<3))
44                     and    subtype=1
45                     and    stillat=1
46                     and    sickpay=-1
47                     ;
48              run;

NOTE: There were 48353 observations read from the data set JOBS21.JOBS21.
NOTE: The data set WORK.J21R753 has 7290 observations and 86 variables.
NOTE: Compressing data set WORK.J21R753 decreased size by 5.00 percent.
      Compressed is 38 pages; un-compressed would require 40 pages.
NOTE: DATA statement used (Total process time):
      real time           1.21 seconds
      cpu time            0.14 seconds


49
50              proc print data=j21r753 (obs=20);
51                  title1 'Print Sample of Continuation Current Main Jobs';
52                  title2 'Panel 23 Round 7 or Panel 24 Round 5 or Panel 25 Round 3 Records';
53                     var jobidx panel rn origrnd subtype stillat sickpay;
54              run;

NOTE: There were 20 observations read from the data set WORK.J21R753.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


55
56
57              ***  b.                                                                    ***
58              ***  Select newly reported Panel 23 or Panel 24 or Panel 25 Current Main Jobs ***
59              ***  records from the FY 2020 JOBS file and print selected variables       ***
60              ***  from the first 20 observations.                                       ***;
61
62               data j20;
63                  set jobs20.jobs20;
64                     if    ((panel=23 and rn in (5,6))
65                            or (panel=24 and rn in (3,4))
66                            or (panel=25 and rn in (1,2)))
67                     and    subtype=1
68                     and    stillat=-1
69                     ;
70              run;
```

```
NOTE: There were 47776 observations read from the data set JOBS20.JOBS20.
NOTE: The data set WORK.J20 has 6456 observations and 84 variables.
NOTE: Compressing data set WORK.J20 decreased size by 2.94 percent.
      Compressed is 33 pages; un-compressed would require 34 pages.
NOTE: DATA statement used (Total process time):
      real time           2.46 seconds
      cpu time            0.03 seconds


71
72                    proc print data= j20 (obs=20);
73                       title1 'Print Sample of Newly Reported Current Main Jobs';
74                       title2 'Panel 23 Round 5 or 6 or Panel 24 Round 3 or 4 or Panel 25 Round 1 or 2 Records';
75                    var jobidx panel rn origrnd subtype stillat sickpay;
76                    run;

NOTE: There were 20 observations read from the data set WORK.J20.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


77
78                    proc freq data= j20 ;
79                       tables sickpay/list missing;
80                       title1 'Sickpay Value of FY2020 Newly Reported Current Main Jobs';
81                       title2 'Panel 23 Round 5 or 6 or Panel 24 Round 3 or 4 or Panel 25 Round 1 or 2 Records';
82                    run;

NOTE: There were 6456 observations read from the data set WORK.J20.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


83
84                    title2;
85
86
87                    ***  c.                                                  ***
88                    ***  Select newly reported Panel 23 or Panel 24 Current Main Jobs     ***
89                    ***  records from the FY 2019 JOBS file and print selected variables  ***
90                    ***  from the first 20 observations.                     ***;
91
92                    data j19;
93                       set jobs19.jobs19;
94                          if   ((panel=23 and rn in (3,4))
95                               or (panel=24 and rn in (1,2)))
96                          and    subtype=1
97                          and    stillat=-1
98                          ;
99                    run;

NOTE: There were 50334 observations read from the data set JOBS19.JOBS19.
NOTE: The data set WORK.J19 has 8924 observations and 84 variables.
NOTE: Compressing data set WORK.J19 decreased size by 6.38 percent.
      Compressed is 44 pages; un-compressed would require 47 pages.
NOTE: DATA statement used (Total process time):
      real time           2.16 seconds
      cpu time            0.06 seconds


100
101                    proc print data= j19 (obs=20);
102                       title1 'Print Sample of Newly Reported Current Main Jobs';
103                       title2 'Panel 23 Round 3 or 4 or Panel 24 Round 1 or 2 Records';
104                    var jobidx panel rn origrnd subtype stillat sickpay;
105                    run;

NOTE: There were 20 observations read from the data set WORK.J19.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


106
107                    proc freq data= j19 ;
108                       tables sickpay/list missing;
109                       title1 'Sickpay Value of FY2019 Newly Reported Current Main Jobs';
110                       title2 'Panel 23 Round 3 or 4 or Panel 24 Round 1 or 2 Records';
111                    run;

NOTE: There were 8924 observations read from the data set WORK.J19.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.02 seconds
      cpu time            0.03 seconds
```

```
112
113               title2;
114
115
116               ***  d.                                                    ***
117               ***  Select newly reported Panel 23 Current Main Jobs records from    ***
118               ***  the FY 2018 JOBS file and print selected variables from the      ***
119               ***  first 20 observations.                                 ***;
120
121               data j18;
122                   set jobs18.jobs18;
123                   if       subtype=1
124                   and      stillat=-1
125                   and      panel=23
126                   and      rn in (1,2);
127               run;

NOTE: There were 53323 observations read from the data set JOBS18.JOBS18.
NOTE: The data set WORK.J18 has 7774 observations and 85 variables.
NOTE: Compressing data set WORK.J18 decreased size by 7.14 percent.
      Compressed is 39 pages; un-compressed would require 42 pages.
NOTE: DATA statement used (Total process time):
      real time           1.32 seconds
      cpu time            0.12 seconds


128
129
130               proc print data=j18 (obs=20);
131                   title1 'Print Sample of Newly Reported Current Main Jobs';
132                   title2 'Panel 23 Round 1 or 2 Records';
133                   var jobidx panel rn subtype stillat sickpay;
134               run;

NOTE: There were 20 observations read from the data set WORK.J18.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


135
136               proc freq data=j18;
137                   tables sickpay/list missing;
138                   title1 'Sickpay Value of FY2018 Newly Reported Current Main Jobs';
139                   title2 'Panel 23 Round 1 or 2 Records';
140               run;

NOTE: There were 7774 observations read from the data set WORK.J18.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


141
142
143               ***  e.                                                    ***
144               ***  Sort and merge datasets into J21R753F                 ***
145               ***  Prepare FY2018, FY2019, FY2020 and FY2021 data for merge    ***;
146
147               proc sort data=j21r753;
148                   by jobidx;
149               run;

NOTE: There were 7290 observations read from the data set WORK.J21R753.
NOTE: SAS sort was used.
NOTE: The data set WORK.J21R753 has 7290 observations and 86 variables.
NOTE: Compressing data set WORK.J21R753 decreased size by 5.00 percent.
      Compressed is 38 pages; un-compressed would require 40 pages.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


150
151               proc sort data=j20;
152                   by jobidx;
153               run;

NOTE: There were 6456 observations read from the data set WORK.J20.
NOTE: SAS sort was used.
NOTE: The data set WORK.J20 has 6456 observations and 84 variables.
NOTE: Compressing data set WORK.J20 decreased size by 2.94 percent.
      Compressed is 33 pages; un-compressed would require 34 pages.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


154
```

```
155                 proc sort data=j19;
156                     by jobidx;
157                 run;

NOTE: There were 8924 observations read from the data set WORK.J19.
NOTE: SAS sort was used.
NOTE: The data set WORK.J19 has 8924 observations and 84 variables.
NOTE: Compressing data set WORK.J19 decreased size by 6.38 percent.
      Compressed is 44 pages; un-compressed would require 47 pages.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.02 seconds
      cpu time            0.03 seconds


158
159                 proc sort data=j18;
160                     by jobidx;
161                 run;

NOTE: There were 7774 observations read from the data set WORK.J18.
NOTE: SAS sort was used.
NOTE: The data set WORK.J18 has 7774 observations and 85 variables.
NOTE: Compressing data set WORK.J18 decreased size by 7.14 percent.
      Compressed is 39 pages; un-compressed would require 42 pages.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


162
163
164             ***  f.                                                    ***
165             ***  Create a dataset (J21R753F) that includes all variables for the    ***
166             ***  continuation Panel 23 Round 7 or Panel 24 Round 5 or Panel 25 Round 3    ***
167             ***  Current Main Jobs and create the new variable SICKPAYX by         ***
168             ***  copying SICKPAY from the corresponding Round 1, Round 2, Round 3,    ***
169             ***  Round 4, Round 5 or Round 6 newly reported job record. Users may    ***
170             ***  prefer to drop "yy" variables at this point            ***;
171
172               data j21r753f;
173                 merge j21r753 (in=a)
174                             j20   (in=b  keep = jobidx sickpay rename=(sickpay=SICKPAY20))
175                             j19   (in=c  keep = jobidx sickpay rename=(sickpay=SICKPAY19))
176                             j18   (in=d  keep = jobidx sickpay rename=(sickpay=SICKPAY18));
177                 by jobidx;
178
179             if a and b and SICKPAY20 ^= .
180                 then SICKPAYX = SICKPAY20;
181
182             else if a and c and SICKPAY19 ^= .
183                 then SICKPAYX = SICKPAY19;
184
185             else if a and d and SICKPAY18 ^= .
186                 then SICKPAYX = SICKPAY18;
187
188             if a and (b or c or d);
189                 run;

NOTE: There were 7290 observations read from the data set WORK.J21R753.
NOTE: There were 6456 observations read from the data set WORK.J20.
NOTE: There were 8924 observations read from the data set WORK.J19.
NOTE: There were 7774 observations read from the data set WORK.J18.
NOTE: The data set WORK.J21R753F has 7289 observations and 90 variables.
NOTE: Compressing data set WORK.J21R753F decreased size by 4.76 percent.
      Compressed is 40 pages; un-compressed would require 42 pages.
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.04 seconds


190
191                 proc freq data=j21r753f;
192                     tables panel*rn*sickpay*sickpayx/list missing;
193                     title1  'Diagnostic Post-Merge - Sickpay * Sickpayx';
194                     title2  'Panel 23 Round 7 or Panel 24 Round 5 or Panel 25 Round 3 Continuation Current Main Jobs ';
195                 run;

NOTE: There were 7289 observations read from the data set WORK.J21R753F.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds
```

Print Sample of Continuation Current Main Jobs
Panel 23 Round 7 or Panel 24 Round 5 or Panel 25 Round 3 Records

Obs	JOBIDX	PANEL	RN	ORIGRND	SUBTYPE	STILLAT	SICKPAY
1	2320022103106	23	7	3	1	1	-1
2	2320024102102	23	7	1	1	1	-1
3	2320027102103	23	7	1	1	1	-1
4	2320038101101	23	7	1	1	1	-1
5	2320041101102	23	7	4	1	1	-1
6	2320043102102	23	7	1	1	1	-1
7	2320045101104	23	7	4	1	1	-1
8	2320045102103	23	7	2	1	1	-1
9	2320050101101	23	7	1	1	1	-1
10	2320051101101	23	7	1	1	1	-1
11	2320057101101	23	7	1	1	1	-1
12	2320057102103	23	7	2	1	1	-1
13	2320063102102	23	7	1	1	1	-1
14	2320065102202	23	7	5	1	1	-1
15	2320074102103	23	7	3	1	1	-1
16	2320075101101	23	7	1	1	1	-1
17	2320078101101	23	7	1	1	1	-1
18	2320081101101	23	7	1	1	1	-1
19	2320081103102	23	7	1	1	1	-1
20	2320091101104	23	7	4	1	1	-1

Print Sample of Newly Reported Current Main Jobs
Panel 23 Round 5 or 6 or Panel 24 Round 3 or 4 or Panel 25 Round 1 or 2 Records

Obs	JOBIDX	PANEL	RN	ORIGRND	SUBTYPE	STILLAT	SICKPAY
1	2320006102204	23	5	5	1	-1	1
2	2320006102203	23	6	6	1	-1	2
3	2320019103301	23	6	6	1	-1	-8
4	2320065102202	23	5	5	1	-1	1
5	2320069101104	23	6	3	1	-1	2
6	2320112104103	23	6	6	1	-1	2
7	2320123102103	23	5	5	1	-1	2
8	2320125103104	23	5	5	1	-1	2
9	2320134102105	23	5	5	1	-1	1
10	2320149101103	23	5	5	1	-1	-1
11	2320158101106	23	5	5	1	-1	-1
12	2320184102105	23	6	6	1	-1	-1
13	2320206101103	23	5	5	1	-1	1
14	2320221104407	23	5	5	1	-1	1
15	2320222101068	23	6	6	1	-1	1
16	2320229103201	23	5	5	1	-1	-1
17	2320271101102	23	5	5	1	-1	2
18	2320271102103	23	5	5	1	-1	2
19	2320279102105	23	6	6	1	-1	1
20	2320280104301	23	5	5	1	-1	1

Sickpay Value of FY2020 Newly Reported Current Main Jobs
Panel 23 Round 5 or 6 or Panel 24 Round 3 or 4 or Panel 25 Round 1 or 2 Records

HAS PAID SICK LEAVE THRU JOB

SICKPAY	Frequency	Percent	Cumulative Frequency	Cumulative Percent
-8	230	3.56	230	3.56
-7	9	0.14	239	3.70
-1	791	12.25	1030	15.95
1	3210	49.72	4240	65.68
2	2216	34.32	6456	100.00

Print Sample of Newly Reported Current Main Jobs
Panel 23 Round 3 or 4 or Panel 24 Round 1 or 2 Records

Obs	JOBIDX	PANEL	RN	ORIGRND	SUBTYPE	STILLAT	SICKPAY
1	2320002101202	23	3	1	1	-1	2
2	2320002102102	23	4	4	1	-1	2
3	2320006102201	23	4	4	1	-1	2
4	2320019101103	23	3	3	1	-1	2
5	2320019103207	23	3	3	1	-1	2
6	2320022103106	23	3	3	1	-1	2
7	2320028101105	23	3	3	1	-1	1
8	2320028102108	23	4	4	1	-1	1
9	2320028103107	23	3	3	1	-1	2
10	2320034101104	23	3	3	1	-1	1
11	2320034102105	23	3	3	1	-1	2
12	2320034102108	23	4	4	1	-1	2
13	2320034107107	23	3	3	1	-1	1
14	2320036102103	23	4	4	1	-1	2
15	2320041101102	23	4	4	1	-1	2
16	2320045101104	23	4	4	1	-1	1
17	2320069101103	23	3	3	1	-1	-1
18	2320074102103	23	3	3	1	-1	2
19	2320091101104	23	4	4	1	-1	1
20	2320102101102	23	3	3	1	-1	1

Sickpay Value of FY2019 Newly Reported Current Main Jobs
Panel 23 Round 3 or 4 or Panel 24 Round 1 or 2 Records

HAS PAID SICK LEAVE THRU JOB

SICKPAY	Frequency	Percent	Cumulative Frequency	Cumulative Percent
-8	276	3.09	276	3.09
-7	14	0.16	290	3.25
-1	1004	11.25	1294	14.50
1	4566	51.17	5860	65.67
2	3064	34.33	8924	100.00

Print Sample of Newly Reported Current Main Jobs
Panel 23 Round 1 or 2 Records

Obs	JOBIDX	PANEL	RN	SUBTYPE	STILLAT	SICKPAY
1	2320002101201	23	1	1	-1	2
2	2320002101203	23	2	1	-1	2
3	2320002102101	23	1	1	-1	1
4	2320003102102	23	1	1	-1	-8
5	2320008102101	23	1	1	-1	1
6	2320019101101	23	1	1	-1	1
7	2320019102203	23	1	1	-1	2
8	2320019102205	23	2	1	-1	-1
9	2320019103201	23	1	1	-1	2
10	2320019104204	23	1	1	-1	1
11	2320022103103	23	1	1	-1	2
12	2320022104104	23	1	1	-1	2
13	2320022104105	23	2	1	-1	1
14	2320024102102	23	1	1	-1	1
15	2320027102103	23	1	1	-1	1
16	2320028102102	23	1	1	-1	2
17	2320032101101	23	1	1	-1	1
18	2320032102102	23	1	1	-1	1
19	2320034101101	23	1	1	-1	1
20	2320034102102	23	1	1	-1	2

Sickpay Value of FY2018 Newly Reported Current Main Jobs
Panel 23 Round 1 or 2 Records

HAS PAID SICK LEAVE THRU JOB

SICKPAY	Frequency	Percent	Cumulative Frequency	Cumulative Percent
-8	246	3.16	246	3.16
-7	18	0.23	264	3.40
-1	921	11.85	1185	15.24
1	4066	52.30	5251	67.55
2	2523	32.45	7774	100.00

Diagnostic Post-Merge - Sickpay * Sickpayx
Panel 23 Round 7 or Panel 24 Round 5 or Panel 25 Round 3
Continuation Current Main Jobs

PANEL	RN	SICKPAY	SICKPAYX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
23	7	-1	-8	73	1.00	73	1.00
23	7	-1	-7	9	0.12	82	1.12
23	7	-1	-1	344	4.72	426	5.84
23	7	-1	1	1426	19.56	1852	25.41
23	7	-1	2	698	9.58	2550	34.98
24	5	-1	-8	53	0.73	2603	35.71
24	5	-1	-7	3	0.04	2606	35.75
24	5	-1	-1	350	4.80	2956	40.55
24	5	-1	1	1447	19.85	4403	60.41
24	5	-1	2	610	8.37	5013	68.77
25	3	-1	-8	43	0.59	5056	69.36
25	3	-1	-1	328	4.50	5384	73.86
25	3	-1	1	1402	19.23	6786	93.10
25	3	-1	2	503	6.90	7289	100.00

Appendix 2
Sample Stata Program

Convert SAS Datasets to .dat Files

```
libname jobs18 "c:\mydata\jobs18";
libname jobs19 "c:\mydata\jobs19";
libname jobs20 "c:\mydata\jobs20";
libname jobs21 "c:\mydata\jobs21";

proc export data=jobs18.jobs18 outfile= jobs18.dta;
run;

proc export data=jobs19.jobs19 outfile= jobs19.dta;
run;

proc export data=jobs20.jobs20 outfile= jobs20.dta;
run;

proc export data=jobs21.jobs21 outfile= jobs21.dta;
run;
```

Sample Stata Program

```
*#delimit ;

set linesize 100

log using "c:\mydata\APPdofile.log", replace

--------------------------------------------------------------------------------
*a. Select continuing Panel 23 Round 5 or Panel 24 Round 3
*   Current Main JOBS (SUBTYPE=1, STILLAT=1) from the FY 2020 JOBS file
*   and print selected variables from the first 20 observations
--------------------------------------------------------------------------------

use "c:\mydata\jobs20.dta", clear

format PANEL SUBTYPE STILLAT SICKPAY %3.0f

keep if (PANEL==23 & RN==5 & ORIGRND < 5 & SUBTYPE==1 & STILLAT==1 & SICKPAY==-1) | (PANEL==24 &
RN==3 & ORIGRND < 3 & SUBTYPE==1 & STILLAT==1 & SICKPAY==-1)

----------------------------------------------------------------------------
*Print Sample of Continuation P23 R5 and P24 R3 Current Main Job Records
----------------------------------------------------------------------------

asdoc list JOBIDX PANEL RN ORIGRND SUBTYPE STILLAT SICKPAY if _n<=20, font(arial) fs(8)
separator(0) noobs, save(stata_output.doc) title(Print Sample of Continuation P23 R5 or P24 R3
Records)

sort JOBIDX

save "c:\mydata\j20.dta", replace

--------------------------------------------------------------------------------
*b. Select newly reported Panel 23 or Panel 24 Current Main JOBS
*   records from the FY 2019 JOBS file and print selected variables
*   from the first 20 observations.
--------------------------------------------------------------------------------

use "c:\mydata\jobs19.dta", clear

format PANEL SUBTYPE STILLAT SICKPAY %3.0f

keep if ((PANEL==23 & (RN==3 | RN==4) ) | (PANEL==24 & (RN==1 | RN==2))) & SUBTYPE==1 &
STILLAT==-1

----------------------------------------------------------------------
*Print Sample of Newly Reported P23 R3 or 4 and P24 R1 or 2 Records
----------------------------------------------------------------------

asdoc list JOBIDX PANEL RN ORIGRND SUBTYPE STILLAT SICKPAY if _n<=20, font(arial) fs(8)
separator(0) noobs, save(stata_output.doc) title(Print Sample of Newly Reported P23 R3 or 4 and
P24 R1 or 2 Records)

sort JOBIDX

rename SICKPAY SICKPAY19

keep JOBIDX SICKPAY19

save "c:\mydata\j19.dta", replace

----------------------------------------------------------------------
*Sickpay Value of FY2019 P23 R3 or 4 and P24 R1 or 2 Newly Reported CMJs
----------------------------------------------------------------------

asdoc tabulate SICKPAY19, font(arial) fs(8), save(stata_output.doc) title(Sickpay Value of FY2019
P23 R3 or 4 and P24 R 1 or 2 Newly Reported CMJs)

--------------------------------------------------------------------------------
```

```
*c. Select newly reported Panel 23 Current Main JOBS records from
*   the FY 2018 JOBS file and print selected variables from the
*   first 20 observations.
--------------------------------------------------------------------------------

use "c:\mydata\jobs18.dta", clear

format PANEL SUBTYPE STILLAT SICKPAY %3.0f

keep if PANEL==23 & (RN==1 | RN==2) & SUBTYPE==1 & STILLAT==-1

-------------------------------------------------------------------
*Print Sample of Newly Reported P23 R1 or 2 Records
-------------------------------------------------------------------

asdoc list JOBIDX PANEL RN SUBTYPE STILLAT SICKPAY if _n<=20, font(arial) fs(8) separator(0)
noobs, save(stata_output.doc) title(Print Sample of Newly Reported P23 R1 or 2 Records)

-------------------------------------------------------------------
*Sickpay Value of FY2018 P23 R1 or 2 Newly Reported CMJs
-------------------------------------------------------------------

sort JOBIDX

rename SICKPAY SICKPAY18

keep JOBIDX SICKPAY18

save "c:\mydata\j18.dta", replace

asdoc tabulate SICKPAY18, font(arial) fs(8), save(stata_output.doc) title(Sickpay Value of FY2018
P23 R1 or 2 Newly Reported CMJs)

--------------------------------------------------------------------------------
*d. Create a dataset (J20R53F) that includes all variables
*   for the continuation Panel 23 Round 5 or Panel 24 Round 3
*   Current Main JOBS and create the new variable SICKPAYX by
*   copying SICKPAY from the corresponding Round 1, Round 2,
*   Round 3, or Round 4 newly reported job record.
--------------------------------------------------------------------------------

use "c:\mydata\j20.dta", clear

merge 1:m JOBIDX using "c:\mydata\j19.dta", generate(matchvar19)

gen SICKPAYX = .
keep if matchvar19 == 1 | matchvar19 == 3
replace SICKPAYX =  SICKPAY19 if SICKPAY19 != .

merge 1:m JOBIDX using "c:\mydata\j18.dta", generate(matchvar18)

keep if matchvar18 == 3 | matchvar19 == 3
replace SICKPAYX = SICKPAY18 if SICKPAY19 == .

save "c:\mydata\j20r53f.dta", replace

------------------------------------------------------
* Diagnostic Post-Merge - Sickpay * Sickpayx
* Continuation P23 R5 and P24 R3 Current Main Jobs Only
------------------------------------------------------

asdoc tabulate SICKPAY SICKPAYX, save(stata_output.doc) font(arial) fs(8) title(Diagnostic Post-
Merge - Sickpay * Sickpayx)

log close
--------------------------------------------------------------------------------------------
```

```
---------------------------------------------------------------------------------------------------
      name:  <unnamed>
       log:  c:\mydata\APPdofile.log
  log type:  text
.
. ********************************************************************************************
. * a. Select continuing Panel 23, Round 5 or Panel 24, Round 3 Current Main Jobs
. *(SUBTYPE=1, STILLAT=1) from the FY 2020 JOBS file and print selected variables from first 20 obs
. ********************************************************************************************
.
. use "c:\mydata\jobs21.dta", clear


.
. format PANEL ORIGRND SUBTYPE STILLAT SICKPAY %3.0f


.
. keep if (PANEL==23 & RN==7 & ORIGRND < 7 & SUBTYPE==1 & STILLAT==1 & SICKPAY==-1) | (PANEL==24 & R
> N==5 & ORIGRND < 5 & SUBTYPE==1 & STILLAT==1 & SICKPAY==-1) | (PANEL==25 & RN==3 & ORIGRND < 3 & S
> UBTYPE==1 & STILLAT==1 & SICKPAY==-1)
(41,063 observations deleted)


.
. *******************************************************************
. *Print Sample of Continuation P23 R7 and P24 R5 and P25 R3 Records
. *******************************************************************
.
. asdoc list JOBIDX PANEL RN ORIGRND SUBTYPE STILLAT SICKPAY if _n<=20, font(arial) fs(8) separator(
> 0) noobs, save(stata_output.doc) title(Print Sample of Continuation P23 R7 or P24 R5 or P25 R3 Rec
> ords)
(File stata_output.doc already exists, option append was assumed)


.
. sort JOBIDX


.
. save "c:\mydata\j21.dta", replace
file c:\mydata\j21.dta saved


.
.
. ******************************************************************************************
. * b. Select newly reported Panel 23 or Panel 24 or Panel 25 Current Main Jobs
. * records from the FY 2020 JOBS file and print selected variables from first 20 obs
. ******************************************************************************************
.
. use "c:\mydata\jobs20.dta", clear


.
. format PANEL ORIGRND SUBTYPE STILLAT SICKPAY %3.0f


.
. keep if ((PANEL==23 & (RN==5 | RN==6) ) | (PANEL==24 & (RN==3 | RN==4) ) | (PANEL==25 & (RN==1 | R
> N==2))) & SUBTYPE==1 & STILLAT==-1
(41,320 observations deleted)


.
. *************************************************************************************
. *Print Sample of Newly Reported P23 R5 or 6 and P24 R3 or 4 and P25 R1 or 2 Records
. *************************************************************************************
.
. asdoc list JOBIDX PANEL RN ORIGRND SUBTYPE STILLAT SICKPAY if _n<=20, font(arial) fs(8) separator(
> 0) noobs, save(stata_output.doc) title(Print Sample of Newly Reported P23 R5 or 6 and P24 R3 or 4
> and P25 R1 or 2 Records)
(File stata_output.doc already exists, option append was assumed)


.
. sort JOBIDX


.
. rename SICKPAY SICKPAY20


.
. keep JOBIDX SICKPAY20


.
. save "c:\mydata\j20.dta", replace
file c:\mydata\j20.dta saved

.
```

```
. ****************************************************************************************
. *Sickpay Value of FY2020 P23 R5 or 6 and P24 R3 or 4 and P25 R1 or 2 Newly Reported CMJs
. ****************************************************************************************
.
. asdoc tabulate SICKPAY20, font(arial) fs(8), save(stata_output.doc) title(Sickpay Value of FY2020
> P23 R5 or 6 and P24 R3 or 4 and P25 R 1 or 2 Newly Reported CMJs)
(File stata_output.doc already exists, option append was assumed)

.
.
. ******************************************************************************
. * c. Select newly reported Panel 23 or Panel 24 Current Main Jobs records from
. * the FY 2019 JOBS file and print selected variables from first 20 obs
. ******************************************************************************
.
. use "c:\mydata\jobs19.dta", clear

.
. format PANEL ORIGRND SUBTYPE STILLAT SICKPAY %3.0f

.
. keep if ((PANEL==23 & (RN==3 | RN==4) ) | (PANEL==24 & (RN==1 | RN==2))) & SUBTYPE==1 & STILLAT==-
> 1
(41,410 observations deleted)

.
. ****************************************************************************
. *Print Sample of Newly Reported P23 R3 or 4 and P24 R1 or 2 Records
. ****************************************************************************
.
. asdoc list JOBIDX PANEL RN ORIGRND SUBTYPE STILLAT SICKPAY if _n<=20, font(arial) fs(8) separator(
> 0) noobs, save(stata_output.doc) title(Print Sample of Newly Reported P23 R3 or 4 and P24 R1 or 2
> Records)
(File stata_output.doc already exists, option append was assumed)

.
. sort JOBIDX

.
. rename SICKPAY SICKPAY19

.
. keep JOBIDX SICKPAY19

.
. save "c:\mydata\j19.dta", replace
file c:\mydata\j19.dta saved

.
. ****************************************************************************
. *Sickpay Value of FY2019 P23 R3 or 4 and P24 R1 or 2 Newly Reported CMJs
. ****************************************************************************
.
. asdoc tabulate SICKPAY19, font(arial) fs(8), save(stata_output.doc) title(Sickpay Value of FY2019
> P23 R3 or 4 and P24 R 1 or 2 Newly Reported CMJs)
(File stata_output.doc already exists, option append was assumed)

.
.
. ****************************************************************************
. * d. Select newly reported Panel 23 Current Main Jobs records from
. * the FY 2018 JOBS file and print selected variables from first 20 obs
. ****************************************************************************
.
. use "c:\mydata\jobs18.dta", clear

.
. format PANEL SUBTYPE STILLAT SICKPAY %3.0f

.
. keep if PANEL==23 & (RN==1 | RN==2) & SUBTYPE==1 & STILLAT==-1
(45,549 observations deleted)

.
. **************************************************
. *Print Sample of Newly Reported P23 R1 or 2 Records
. **************************************************
.
. asdoc list JOBIDX PANEL RN SUBTYPE STILLAT SICKPAY if _n<=20, font(arial) fs(8) separator(0) noobs
> , save(stata_output.doc) title(Print Sample of Newly Reported P23 R1 or 2 Records)
(File stata_output.doc already exists, option append was assumed)
```

```
.
. *********************************************************
. *Sickpay Value of FY2018 P23 R1 or 2 Newly Reported CMJs
. *********************************************************
.
. sort JOBIDX

.
. rename SICKPAY SICKPAY18

.
. keep JOBIDX SICKPAY18

.
. save "c:\mydata\j18.dta", replace
(note: file c:\mydata\j18.dta not found)
file c:\mydata\j18.dta saved

.
. asdoc tabulate SICKPAY18, font(arial) fs(8), save(stata_output.doc) title(Sickpay Value of FY2018
> P23 R1 or 2 Newly Reported CMJs)
(File stata_output.doc already exists, option append was assumed)

.
. ***************************************************************************
. * e. Create a dataset (J21R753F) that includes all variables for the
. * continuation Panel 23 Round 7 or Panel 24 Round 5 or Panel 25 Round 3
. * Current Main Jobs and create the new variable SICKPAYX by copying
. * SICKPAY from the corresponding Round 1, Round 2, Round 3, Round 4,
. * Round 5 or Round 6 newly reported job record.
. ***************************************************************************
.
. use "c:\mydata\j21.dta", clear

.
. merge 1:m JOBIDX using "c:\mydata\j20.dta", generate(matchvar20)

    Result                         # of obs.
    -----------------------------------------
    not matched                        7,346
        from master                    4,090  (matchvar20==1)
        from using                     3,256  (matchvar20==2)

    matched                            3,200  (matchvar20==3)
    -----------------------------------------

.
. gen SICKPAYX = .
(10,546 missing values generated)

. keep if matchvar20 == 1 | matchvar20 == 3
(3,256 observations deleted)

. replace SICKPAYX =  SICKPAY20 if SICKPAY20 != .
(3,200 real changes made)

.
.
. merge 1:m JOBIDX using "c:\mydata\j19.dta", generate(matchvar19)

    Result                         # of obs.
    -----------------------------------------
    not matched                       11,538
        from master                    4,952  (matchvar19==1)
        from using                     6,586  (matchvar19==2)

    matched                            2,338  (matchvar19==3)
    -----------------------------------------

.
. keep if matchvar19 == 3 | matchvar20 == 1 | matchvar20 == 3
(6,586 observations deleted)

. replace SICKPAYX =  SICKPAY19 if SICKPAY19 != . & SICKPAY20 == .
(2,334 real changes made)

.
.
. merge 1:m JOBIDX using "c:\mydata\j18.dta", generate(matchvar18)
```

```
     Result                          # of obs.
     -----------------------------------------
     not matched                        11,550
         from master                     5,533  (matchvar18==1)
         from using                      6,017  (matchvar18==2)

     matched                             1,757  (matchvar18==3)
     -----------------------------------------


.
. keep if matchvar18 == 3 | matchvar19 == 3 | matchvar20 == 3
(6,018 observations deleted)

. replace SICKPAYX = SICKPAY18 if SICKPAY18 != . & SICKPAY19 == . & SICKPAY20 == .
(1,755 real changes made)


.
.
. save "c:\mydata\j21r753f.dta", replace
(note: file c:\mydata\j21r753f.dta not found)
file c:\mydata\j21r753f.dta saved


.
.
. *******************************************************************
. * Diagnostic Post-Merge - Sickpay * Sickpayx
. * Continuation P23 R7 and P24 R5 and P25 R3 Current Main Jobs Only
. *******************************************************************
.
. asdoc tabulate SICKPAY SICKPAYX, save(stata_output.doc) font(arial) fs(8) title(Diagnostic Post-Me
> rge - Sickpay * Sickpayx)
(File stata_output.doc already exists, option append was assumed)


.
. log close
       name:  <unnamed>
        log:  c:\mydata\APPdofile.log
  log type:  text

------------------------------------------------------------------------------------------
```

Appendix 3
MEPS Industry Codes Condensing Rules

MEPS Industry Codes Condensing Rules
FY2010 and Subsequent Files

Condensed Industry Code	Census Industry Code Range	Description
1	0170 - 0290	Natural Resources
2	0370 - 0490	Mining
3	0770	Construction
4	1070 - 3990	Manufacturing
5	4070 - 5790	Wholesale and Retail Trade
6	0570 - 0690, 6070 - 6390	Transportation and Utilities
7	6470 - 6780	Information
8	6870 - 7190	Financial Activities
9	7270 - 7790	Professional and Business Services
10	7860 - 8470	Education, Health, and Social Services
11	8560 - 8690	Leisure and Hospitality
12	8770 - 9290	Other Services
13	9370 - 9590	Public Administration
14	9890	Military
15	9990	Unclassifiable Industry

MEPS uses the 4-digit Census occupation and industry coding systems developed for the Current Population Survey and the American Community Survey.

For industry coding, MEPS uses the 2007 4-digit Census industry codes. Descriptions of the 4-digit Census industry codes can be found at the U.S. Census Bureau website.

See Census IO Index for more information on the Census coding systems used by MEPS.

Appendix 4
MEPS Occupation Codes Condensing Rules

MEPS Occupation Codes Condensing Rules
FY2010 and Subsequent Files

Condensed Occupation Code	Census Occupation Code Range	Description
1	0010 - 0950	Management, Business, and Financial Operations Occupations
2	1005 - 3540	Professional and Related Occupations
3	3600 - 4650	Service Occupations
4	4700 - 4965	Sales and Related Occupations
5	5000 - 5940	Office and Administrative Support Occupations
6	6005 - 6130	Farming, Fishing, and Forestry Occupations
7	6200 - 7630	Construction, Extraction, and Maintenance Occupations
8	7700 - 9750	Production, Transportation, and Material Moving Occupations
9	9840	Military Specific Occupations
10	9920	Not in Labor Force
11	9990	Unclassifiable Occupation

MEPS uses the 4-digit Census occupation and industry coding systems developed for the Current Population Survey and the American Community Survey.

For occupation coding, MEPS uses the 2010 4-digit Census occupation codes. Descriptions of the 4-digit Census occupation codes can be found at the U.S. Bureau of Labor Statistics website.

See the Census IO Index for more information on the Census coding systems used by MEPS.